NeuroImage: Clinical 33 (2022) 102937

Contents lists available at ScienceDirect

NeuroImage: Clinical

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

General additive models address statistical issues in diffusion MRI: An example with clinically anxious adolescents

Nathan M. Muncy [a,*], Adam Kimbler [a], Ariana M. Hedges-Muncy [b], Dana L. McMakin [a], Aaron T. Mattfeld [a]

[a] Center for Children and Families, Florida International University, Miami, Florida, USA
[b] Brigham Young University, Provo, UT, USA

ARTICLE INFO

Keywords:
Adolescence
Uncinate
Anxiety
DWI
MRI
GAM

ABSTRACT

Statistical models employed to test for group differences in quantized diffusion-weighted MRI white matter tracts often fail to account for the large number of data points per tract in addition to the distribution, type, and interdependence of the data. To address these issues, we propose the use of Generalized Additive Models (GAMs) and supply code and examples to aid in their implementation. Specifically, using diffusion data from 73 peri-adolescent clinically anxious and no-psychiatric-diagnosis control participants, we tested for group tract differences and show that a GAM allows for the identification of differences within a tract while accounting for the nature of the data as well as covariates and group factors. Further, we then used these tract differences to investigate their association with performance on a memory test. When comparing our high versus low anxiety groups, we observed a positive association between the left uncinate fasciculus and memory overgeneralization for negatively valenced stimuli. This same association was not evident in the right uncinate or anterior forceps. These findings illustrate that GAMs are well-suited for modeling diffusion data while accounting for various aspects of the data, and suggest that the adoption of GAMs will be a powerful investigatory tool for diffusion-weighted analyses.

1. Introduction

Diffusion-weighted imaging (DWI) is a magnetic resonance imaging technique capitalizing on constrained water diffusion to approximate anatomical features such as axonal bundles across voxels. Modeling axonal bundles is desirable as modern neuroscience conceptualizes the functioning of the central nervous system as a dynamic graph where individual functional nodes are connected within a network (Feldman et al., 2010). Distinct regions form the functional nodes which are connected via axonal projections, constituting the network edges (Sotiropoulos and Zalesky, 2019). Accordingly, a systems-level description of the functioning central nervous system necessitates accurate characterization of both micro- and macro-anatomic pathways. Clinically, mental health, injury, and/or disease often are associated with white matter disruption (e.g. traumatic brain injury, multiple sclerosis, anxiety), and classification of these differences may serve as strong etiological biomarkers (Harrison et al., 2011; De Santis et al., 2019; Mesaros et al., 2012; Shenton et al., 2012; Hutchinson et al., 2018; Raizman et al., 2020; Delouche et al., 2016; Adluru et al., 2017; Jamieson et al., 2021). Additionally, developmental research will benefit from careful modeling given the role of myelination in both development and DWI metrics (Dumontheil, 2016; Østby et al., 2009). Utilizing DWI data to sensitively investigate group differences remains a non-trivial task fraught with issues of multiple comparisons, data distribution and type, and interdependence. Accordingly, it is our goal to articulate a number of extant issues in modeling DWI data, propose a statistical approach that will address these issues, and demonstrate the utility of such an approach. To this end, we analyzed data collected from 73 periadolescent participants of both clinically anxious and no-psychiatric-diagnosis control populations. We demonstrate that our statistical proposal accounts for many of the troublesome aspects of DWI data and that such analyses are capable of identifying group differences that correlate with behavior outcomes.

Recent advancements in the quantification and analysis of DWI data has resulted in the development of the Automated Fiber Quantification (AFQ) software (Kruper et al., 2021; Yeatman et al., 2012; Yeatman

* Corresponding author.
E-mail address: nmuncy@fiu.edu (N.M. Muncy).

https://doi.org/10.1016/j.nicl.2022.102937
Received 13 September 2021; Received in revised form 10 December 2021; Accepted 3 January 2022
Available online 5 January 2022

et al., 2018). AFQ utilizes a conjunction of techniques with the goal of robustly producing sensitive axonal pathway models, and results in tractographic profiles from which multiple diffusion metrics may be derived. Critically, AFQ resamples each tract into *N* equidistant nodes, typically ≈ 100, allowing for the calculation of diffusion metrics (scalars, such as FA or MD) at each node along the tract. This resampling facilitates a more sensitive comparison between groups on the tract diffusion metrics as it becomes possible to interrogate whether group differences exist *within* a tract rather than simply comparing averaged tract metrics. Generating 100 scalar values per tract, however, can be problematic for statistical analyses.

Researchers using AFQ are often motivated to investigate group differences within a tract, conducting an analysis on each tract node (termed a "point-wise" analysis) which results in a non-trivial number of comparisons for which one must properly correct. Of the 36 papers we identified that used AFQ for DWI analyses (using a PubMed search query "automated fiber quantification" OR "AFQ" AND "diffusion" [Table 1]), 28 employed either a permutation-based multiple comparisons correction method (Nichols and Holmes, 2002), a false discover rate correction, or a Bonferroni adjustment. Unfortunately, such correction methods may trade sensitivity for proper family-wise error rates, potentially inflating the probability of Type-II errors. A second issue with AFQ-derived diffusion statistics we note is the distribution of values for any given tract may not meet the requisite normality assumptions of Student's *t* or Analysis of Variance testing, even if multiple comparisons are properly controlled. While certain studies noted and accounted for this distribution issue (Table 1), these studies are the minority, and failing to meet basic assumptions of statistical models in the broader use of AFQ remains an issue. The non-normality complication is compounded with yet one more problem not unique to AFQ but relevant for all studies using diffusion data: diffusion metrics like FA values are not continuous but proportional, being bounded between 0 (perfect isotropy) and 1 (perfect anisotropy). Failing to account for the type of data may yet constitute another violation of the basic assumptions for the statistical models employed. Finally, and most seriously, there is a violation of independence due to the spatial correlation in the nodes which drastically inflates the resulting *p*-values. Together, then, we note three points and potential issues which must be considered during test selection for AFQ-derived DWI data: multiple comparisons, non-Gaussian distributions and proportional data type, and spatial dependencies.

A Generalized Additive Model (GAM; Wood, 2017) is an ideal method for modeling AFQ DWI values and addressing the aforementioned issues. From the family of regression models, a GAM models the relationship between independent and dependent variables utilizing a set of smoothing functions. These smoothing functions allow for a proper model fit to "wiggly" data, or data with an X-Y relationship that cannot be readily described with a polynomial (e.g. linear, quadratic, etc.) Additionally, and similar to generalized linear models, a range of link functions facilitate appropriate modeling of non-Gaussian and bounded data. The resulting GAM generates a smoothed spline fit to the data. For use with AFQ, separate splines may be produced for different groups when modeling a tract and by comparing group splines, it will be possible to identify the tract nodes which differ between groups,

Table 1
Articles from the PubMed search terms "automated fiber quantification" OR "AFQ" AND "diffusion". Test Column: statistic employed to test between groups, among others; node = statistic tested each node, mean = FA values were averaged across nodes, tract = node was treated as a within-subject variable, M-W = Mann–Whitney, K-W = Kruskal–Wallis. MCC Method Column: multiple comparison correction (MCC) method employed; Perm = permutation based, α/N = reduced alpha, FWE = family-wise error, FDR = false discovery rate, FPC = false positive correction, Bonf = Bonferroni. "*" indicates that the same analyses were reported twice.

1st Author, Year	Sample Size	Test	MCC Method	DOI
Angelopoulou et al. (2019)	57	node t-test	Perm	10.3389/fnins.2019.01424
Banfi et al. (2019)	69	node ANOVA	$\alpha/3$	10.1002/hbm.24410
Cai et al. (2019)	55	node t-test	Perm	10.1007/s11682-019–00160-1
Carbine et al. (2020)	87	Spline		10.1007/s11682-019–00036-4
Chen et al. (2020)	70	node Anova node GLM	FWE	10.1111/cns.13283
Chen et al. (2020)	81	node GLM	FDR	10.3389/fnins.2020.570123
Clocksin et al. (2021)	43	tract LMM		10.1016/j.ymgme.2020.12.001
Deng et al. (2018)	104	node t-tests	Perm, FPC	10.1016/j.pnpbp.2017.09.006
Dou et al. (2020)	120	node ANOVA	FDR	10.1016/j.cortex.2020.03.032
Goodrich-Hunsaker et al. (2018)	153	node ANCOVA	FDR	10.1002/jnr.24142
Hall et al. (2016)	20	tract ANCOVA	Bonf	10.1016/j.nicl.2016.01.013
Huang et al. (2020)	244	node GLM	Perm	10.3389/fnagi.2020.598242
Huang et al. (2019)	6	*Case studies*		10.3171/2019.5.PEDS19117
Jiang et al. (2019)	72	node M-W U	Bonf	10.3174/ajnr.A5914
Jossinger et al. (2021)	42	mean Wilcoxon node Wilcoxon	Perm	10.1007/s00429-020–02210-7
Kreilkamp et al. (2019)	64	mean K-W ANOVA along-the-tract	Bonf, FDR	10.1016/j.nicl.2019.102024
Li et al. (2020)	37	mean t-test node t-test	FDR	10.1038/s41598-020–73305-8
Li et al. (2017)	107	node t-test		10.3760/cma.j.issn.0376–2491.2017.13.003
Li et al. (2020)	42	node t-test		10.3760/cma.j.issn.0376–2491.2020.03.003
Lin et al. (2020)	14	node t-test	Perm	10.1007/s11682-018–0010-2
Pascual-Diaz et al. (2020)	120	node t-test	Bonf	10.1016/j.neuroimage.2020.117260
Sacchet et al. (2014)*	32	mean t-test node t-test	FDR	10.1186/2045–5380-4–8
Sacchet et al. (2014)*	32	mean t-test node t-test	FDR	10.1109/ISBI.2014.6867940
Sarica et al. (2017)	24	node t-test	Perm	10.1002/hbm.23412
Sarica et al. (2019)	44	node t-test	Perm	10.1159/000503970
Sommer et al. (2017)	16	mean NRMSE	Perm	10.1002/brb3.588
Unterrainer et al. (2019)	153	node t-test	Perm	10.3389/fpsyt.2019.00667
Vakhtin et al. (2020)	246	mean ANCOVA		10.1089/neu.2019.6487
Van Der Auwera et al. (2021)	87	node t-test	$\alpha/5$, FDR	10.1016/j.neuroimage.2021.118087
Xue et al. (2019)	13	mean Pearson		10.1109/EMBC.2019.8857590
Yeatman et al. (2012)	74	*Original Paper node t-test*	Perm	10.1371/journal.pone.0049790
Yeatman et al. (2018)		*Methods*		10.1038/s41467-018–03297-7
Zeineh et al. (2015)	29	node t-test	Perm	10.1148/radiol.14141079
Zhang (2018)	25	mean t-test node t-test	Bonf	10.3389/fneur.2018.00089
Zhang et al. (2019)	158	node t-test	Perm	10.1016/j.nicl.2019.101723
Zhou et al. (2018)	54	node t-test	FDR	10.1016/j.brainres.2018.07.003

mitigating the point-wise multiple comparisons problem. Additionally, covariates may be included in the GAM thereby allowing one to model a tract for multiple groups while controlling for various factors and parameters. It is our aim, then, to demonstrate the use of a GAM to model AFQ-derived FA values, and to facilitate future implementations, we supply and describe the requisite code.

2. Methods

2.1. Participants

Periadolescent participants, ages 10–13 years, were recruited from both community sources and pediatric anxiety clinics as part of an ongoing R01 study. Data from the subset of these participants who contributed a diffusion-weighted MR image was used in the current work, totaling 73 participants (44 female, age $= 11.2 \pm 1.1$ years). Institutional Review Board approval was obtained at the outset of the study, and prior to beginning experimental procedures participants completed informed consent and assent. The clinical population had an inclusion criteria of an anxiety disorder diagnosis, and participants were evaluated for any MR contraindications and exclusionary major medical issues or diagnosed psychiatric disorders (e.g., past or current depressive disorder, bipolar disorder, post-traumatic stress disorder, conduct disorder, oppositional defiant disorder, psychotic disorders, obsessive compulsive disorder). All participants were right-handed and had normal or corrected-to-normal vision. Each participant was assessed for anxiety severity using the Pediatric Anxiety Rating Scale (PARS-6, described below), and pubertal development was assessed via the Pubertal Development Scale (PDS, Petersen et al., 1988). Scoring for the PDS followed Shirtcliff et al. (Shirtcliff et al., 2009), which is an approach that aims to approximate Tanner staging (Tanner, 1962). Following completion of the study protocol, participants were remunerated for their time.

2.2. Pediatric anxiety rating scale

The Pediatric Anxiety Rating Scale (PARS; Storch, 2012) is a semi-structured interview in which clinicians assess anxiety severity over the past week by probing 50 anxiety symptoms with both parent and child (interviewed separately), and then rating seven global severity items. The PARS-6 (range 0–30) is an established computation that eliminates symptom number from the total severity score given lack of direct contribution to severity, as severity can be driven by a single symptom dimension and frequency of all symptoms is already captured in the score (Caporino et al., 2013). To achieve a full distribution of symptom severity in the primary study, randomization (Sleep, Wake conditions, see below) was stratified by 3 levels of anxiety severity as assessed by PARS-6: 0–3 (Low), 4–12 (Medium), and 13–30 (High). Severity ranges were determined by a review of literature defining cut-points with high specificity and sensitivity for a clinical diagnosis (Ginsburg et al., 2011) and for likely remission in clinical trials (Caporino et al., 2013; Johnco et al., 2015), as well as a review of severity distributions in 3 archival datasets that included clinic and non-clinic samples in this age-range. Of the 73 participants whose data were used in this experiment, 34 participants were classified as Low, 20 as Medium, and 19 as High.

2.3. Memory experiment

Participants took part in an emotional version of the mnemonic similarity task (eMST; Stark et al., 2019; Leal et al., 2014), which consisted of Study and Test sessions: the Study session involved an incidental encoding task during which participants viewed pictures of everyday scenes for two seconds and were instructed to endorse each scene as either emotionally negative, neutral, or positive. These stimuli were separated by a jittered inter-stimulus interval (2–6 s) during which a white central fixation was presented on a black background. Each scene was presented once resulting in a total of 175 images (58 negative, 57 neutral, 60 positive). Participants then returned one week later between the hours of 11:00 and 15:00 for a surprise memory test. This test consisted of presenting participants a random order of stimuli that were either identical to one encountered in the Study session (Targets), similar to but different from a Study session stimulus (Lures), or entirely novel (Foils). Participants were instructed to endorse each stimulus as either 'Old' or 'New', where 'Old' indicated they remembered the exact stimulus from the encoding session. Stimulus duration and inter-stimulus intervals were identical to the encoding session, and a total of 251 stimuli were presented: 48 Targets (16 negative, 15 neutral, 17 positive), 91 Lures (30 negative, 30 neutral, 31 positive), and 112 Foils (33 negative, 40 neutral, 39 positive). A complete description of the paradigm and an analysis of R01 pilot data are reported elsewhere (McMakin et al., 2021).

2.4. MRI

Imaging was conducted on a 3 Tesla Siemens MAGNETOM Prism at the Florida International University Center for Imaging Science utilizing a 32-channel coil. Each participant contributed T1 and diffusion weighted images. T1-weighted structural scans were acquired using a magnetization-prepared rapid acquisition with gradient echo (MP-RAGE) sequence with the following parameters: TE = 29 ms, flip angle = 8°, TR = 2500 ms, FoV = 256 × 256, slices = 176 interleaved, voxel size = 1 mm^3. Diffusion weighted scans were acquired with the following parameters: TE = 89 ms, flip angle = 90°, TR = 4200 ms, voxel size = 1.7 mm^3, 103 slices, 7 reference volumes (b-value = 0 s/mm^2), 4 shells (b-values = 500, 1000, 2000, and 3000 s/mm^2), 96 directions, multi-band acceleration factor = 3, bandwidth = 1700 Hz/Px; all shells were utilized in DWI pre-processing. A diffusion field map was acquired utilizing the same protocol, save that TR = 12600 ms and reversed acquisition direction.

2.5. DWI pre-processing

Pre-processing of DWI data was conducted using FreeSurfer (version 7.1; Fischl, 2012) alongside FSL's FDT suite (version 6.0.3; Behrens et al., 2003, 2007). First, a field map acquired in the P≫A direction was combined with the extracted b_0 images from the diffusion acquisition (acquired A≫P). Next, FSL's TOPUP (Ashburner, 2003; Smith et al., 2004) then utilized this combined b_0 to calculate the susceptibility distortion of the image at each b_0 image. Finally, FSL's Eddy (Andersson and Sotiropoulos, 2016) used the resulting distortion map in conjunction with slice-timing information to produce a motion-corrected diffusion image for subsequent analyses.

2.6. Automated fiber quantification

AFQ utilizes pre-processed DWI data to generate node-specific diffusion metrics for a predetermined set of white matter tracts. Generally, there are six main steps in the process, and using the python version of AFQ (pyAFQ version 0.7.1; Kruper et al., 2021) allows for controlling the software via a single configuration toml file. Our implementation of AFQ capitalized on the command-line interface of pyAFQ and largely utilized the default options for the various parameters.

To briefly describe the approach implemented in AFQ (see Yeatman et al., 2012 for a full description), we first generated a whole-brain fiber tractography map utilizing a probabilistic approach; in tracking an individual bundle, the probabilistic approach accounts for local uncertainty and incorporates variability of the diffusion metrics into the tract calculation, resulting in maps that more accurately describe the geometric properties of fibers and their intermixing. Second, this fiber map was then parcellated for individual tracts. To accomplish this task AFQ

employs the Wakana et al. (Wakana et al., 2007) method of classifying fibers according to whether they pass through a number of *a priori* waypoint regions of interest (ROIs). These waypoint ROIs are defined in an MNI atlas space, and are moved into participant space via a symmetric, non-linear diffeomorphic transformation. Third, fiber tracts were refined by incorporating the probability maps of Hua et al. (Hua et al., 2008). Any tract fibers which take an idiosyncratic pathway between the two waypoint ROIs will have traversed a lower probability space for the tract, decreasing the likelihood that the fiber is a member of the tract. These low-probability fibers were removed from the tract bundle. Fourth, the tract was then cleaned of fibers that significantly differ from the average of the tract bundle. To determine which fibers constitute outliers, a Mahalanobis distance metric for the core of the bundle is computed, and any fiber that had a length >4 standard deviations or a pathway that deviated >5 standard deviations from the average path was removed. Next, the bundle may be clipped such that the terminal projections of the bundle are removed, and only the portion between the waypoints are retained. This is done because the inter-waypoint section typically has a more consistent trajectory resulting in a larger averaged FA value whereas the terminal fibers differ greatly in their orientation and so their corresponding average FA values approach zero. We did not clip the fibers at the waypoints but instead modeled the entire tract. Finally, the bundle tracts were up-sampled by a factor of 100 in order to generate 100 equidistant nodes for which diffusion metrics were derived. These metrics are penalized averages such that voxels with a smaller Mahalanobis distance will have a larger weight but the converse holds true for voxels with a larger distance.

By default pyAFQ generates a set of 18 tracts and, given the aims of the parent R01 (McMakin et al., 2021), three pathways were selected for modeling and testing: the left uncinate fasciculus, the right uncinate fasciculus, and the anterior forceps. The left and right uncinate were selected for their connection with sleep, anxiety, development, and potential link with memory (Jamieson et al., 2021; Granger et al., 2021), while the anterior forceps were selected as a control region due to its similarity in curvature with the uncinate fasciculi (Fig. 1).

2.7. Statistical analyses

All statistical models and analyses were conducted in R version 4.0.3 (R Core Team, 2020), and the main packages used in this work include *fitdistrplus* version 1.1.6 (Delignette-Muller and Dutang, 2015), *mgcv* version 1.8.38 (Wood, 2011), and *itsadug* version 2.4 (van Rij et al., 2020). Analyses reported in this main text focus on modeling FA values via a GAM, but we note that GAMs are equally appropriate to model the other DWI scalars and supply a model of mean diffusivity (MD) values in Supplemental Materials Section 5.1. First, in order to utilize a GAM to model FA values for white matter tracts of interest, we first organized the AFQ output into a data frame which contained the FA value of each tract node for each subject, factors for group and sex, and continuous covariate values; see Supplemental Materials, Table S1 for an example data frame. Next, we assessed the distribution of the FA values in order to select the proper family and link function (R Code 1).

R Code 1: Determining distribution of FA values. df_tract = data frame for a specific tract, dti_fa = column name corresponding to FA values.

```
library("fitdistrplus")
descdist(df_tract$dti_fa, discrete=F)
```

For each tract, we found a number of distributions could be used to model the data. Accordingly, we constructed a separate GAM to fit each potential distribution using the appropriate family (e.g. beta or gamma). These GAMs modeled FA values by fitting a spline for each group while controlling for sex, where each subject was modeled as a random effect. The logit link function was used to account for the bounded (ratio) FA values. Finally, the residual estimates of maximum likelihood (REML) method was used to assess model fit (R Code 2).

R Code 2: GAM of tract FA values using a gamma distribution. For the beta model, the "betar" argument was used instead of "Gamma". In this model, the "bam" function was used to generate the GAM in order to reduce computation time on a large data set, but "gam" could be employed for smaller data sets. dti_fa = column name of FA values, Group & Sex = factors, nodeID = AFQ node, *k* = basis dimension (knots), subjectID = subject identifier, df_tract = data frame for tract.

```
library("mgcv")
fit_gamma <- bam(dti_fa ~ Group +
    Sex +
    s(nodeID, by=Group, k=40) +
    s(subjectID, bs="re"),
    data = df_tract,
    family = Gamma(link = "logit"),
    method = "REML")
```

The resulting model fit was then assessed via the command *gam.check(fit_gamma, rep = 500)* to verify the basis dimension (*k*) employed was appropriate, and the *k* parameter was iteratively updated until the factor *k*-index > 1. Next, the two GAM models using different families were compared directly (R Code 3), and the model with the best fit was used in subsequent analyses.

R Code 3: Determining which GAM produces the best fit. fit_gamma = GAM fit using a gamma distribution, fit_beta = GAM fit using a beta distribution.

```
library("itsadug")
compareML(fit_gamma, fit_beta)
```

Additionally, we opted to include a continuous covariate for pubertal development (PDS) as we were modeling data from an adolescent population (R Code 4). Adding a smooth term for the covariate resulted in a new GAM, and then the process of determining the *k* parameter was repeated. This covariate GAM, which modeled the tract FA values for each group while controlling for sex and PDS, was then compared

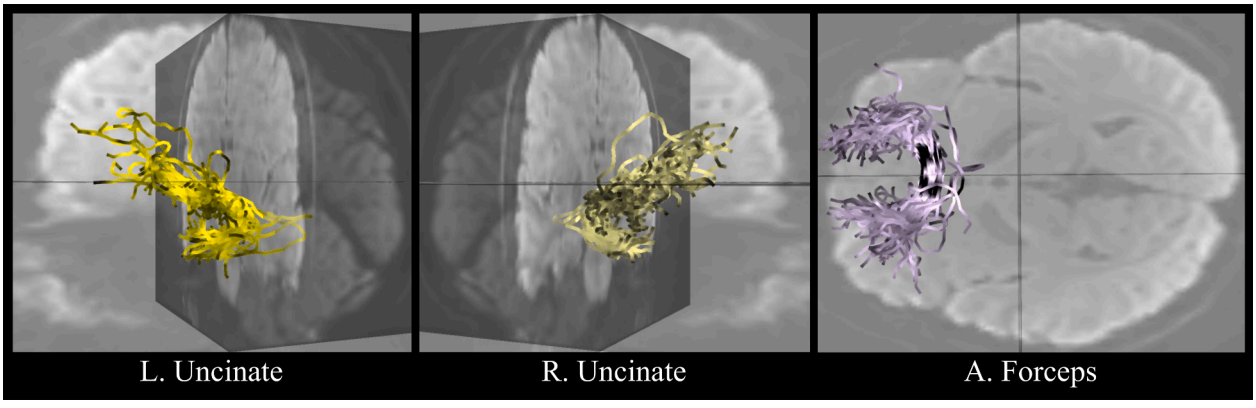

Fig. 1. Representative AFQ tract bundles overlaid on an extracted b_0 image.

against the non-covariate model (fit_gamma) in order to determine whether adding the covariate improved model fit.

R Code 4: GAM of FA data utilizing PDS as a covariate.

```
fit_cov_pds <- bam(dti_fa ~ Group +
    Sex +
    s(nodeID, by=Group, k=40) +
    s(PDS, by=Sex) +
    s(subjectID, bs="re"),
    data = df_tract,
    family = Gamma(link = "logit"),
    method = "REML")
```

Finally, the spline fit estimates and standard error at each node were compared between groups (R Code 5) which identified the tract nodes that differed in their FA values between groups. The averaged FA value of these nodes was then extracted for regression analyses with memory measures. Finally, given that multiple statistical tests were conducted in this manuscript, the interpretation of significance utilized a Bonferroni corrected α-level, except where otherwise stated. Code, data, and documents used in this project are available at https://github.com/nmuncy/emu_AFQ.

R Code 5: Testing whether two splines differ at each node. fit_cov_pds = GAM object; nodeID = AFQ node; 0 & 1 = factor values of group.

```
library("itsadug")
plot_diff(fit_cov_pds,
    view="nodeID",
    comp=list(Group=c("0", "1")),
    rm.ranef=T)
```

3. Results

3.1. GAM of the Left Uncinate

Diffusion-weighted MRI data of 73 periadolescent participants were modeled with pyAFQ to generate a set of white matter tracts for the left uncinate, right uncinate, and anterior forceps pathways. One benefit of the AFQ approach is that it allows to test for group differences along a tract. To this end, participants were grouped into Low, Medium, and High tertiles according to their PARS-6 score. These groups did not differ in age ($F_{(2,70)} = .47,\ p = .62,\ \eta_p^2 = .013$) or PDS ($F_{(2,70)} = .08,\ p = .92,\ \eta_p^2 = .002$). For the left uncinate, assessing the distribution of FA values revealed a beta or gamma function fit the data (Supplemental Materials, Fig. S1; distribution mean $= .45 \pm .07$, skewness $= -.84$, kurtosis $= 3.55$). Accordingly, we conducted two GAMs, one utilizing the "gamma" family (GAM_γ) and one using "beta" (GAM_β; R Code 2); a basis dimension (k) of 40 in both GAMs resulted in k-indices > 1 for each group factor. Next, we tested GAM_γ against GAM_β to determine which model best fit the data (Table 2, L. Uncinate). AIC and model fit difference testing (via *compareML*) indicated the Gamma family had better fit (GAM_γ: $-REML = -8802.6,\ R^2_{adj} = .82$; GAM_β: $-REML = 10561,\ R^2_{adj} = .82$; AIC difference $= -436.1$); χ^2 testing was not conducted as the two models had equal degrees of freedom.

The GAM_γ model for the left uncinate indicated both the Medium and High PARS-6 groups had significantly higher intercepts than Low PARS-6. Also Males did not differ in their intercept from Females (Table 3, L. Uncinate GAM_γ Parametric Coefficients). For the smooth terms, the interaction of tract node with each group was significant, indicating non-flatness of the spline and that an interactive, and not additive, structure is necessary for modeling the data. This was not the case for the subject term (Table 3, Approximate Significance of Smooth Terms [top]).

GAM_γ modeled the left uncinate tract with separate splines for each group while controlling for sex, but as periadolescence is a sensitive period for development, we computed another GAM which included PDS scores as a covariate ($GAM_{\gamma c}$). That is, as a sex × puberty onset age interaction exists such that females typically enter puberty at an early age than males, and this sex × pubertal onset age affects myelination (see Discussion), we attempted to control for developmental variance in tract FA values by incorporating a measure of puberty given that no age difference was detected between males and females (Welch Two Sample $t_{(62.3)} = -0.05,\ p = 0.95$) and that females had higher developmental scores as measured by the PDS (Welch Two Sample $t_{(69.1)} = 2.61,\ p = 0.011$, 95% CI [0.13, 1.01]). In this regard we reasoned that PDS would better control for developmental related variance than age, thereby affording us greater sensitivity to investigate whether tract

Table 2
Model fit of each GAM and deviance explained (Dev Ex). GAM_β = GAM using a beta distribution, GAM_γ = GAM using gamma distribution, $GAM_{\gamma c}$ = GAM with covariate using a gamma distribution, GAM_G = GAM with a Gaussian family, and GAM_{Gc} = GAM with covariate using a Gaussian distribution. REML = residual estimation of maximum likelihood.

L. Uncinate	**-REML**	R^2_{adj}	**Dev Ex**
GAM_β	10561	0.82	0.83
GAM_γ	−8802.6	0.82	0.85
$GAM_{\gamma c}$	−9336	0.85	0.88
R. Uncinate	**-REML**	R^2_{adj}	**Dev Ex**
GAM_β	10578	0.89	0.9
GAM_γ	−10450	0.89	0.92
$GAM_{\gamma c}$	−10912	0.91	0.93
A. Forceps	**-REML**	R^2_{adj}	**Dev Ex**
GAM_β	10659	0.95	0.96
GAM_γ	−10644	0.95	0.96
GAM_G	−14824	0.95	0.96
GAM_{Gc}	−14954	0.96	0.96

Table 3
Model statistics for the GAM_γ (Top) and GAM with a PDS covariate ($GAM_{\gamma c}$, bottom) of left uncinate FA values. Est = model estimate, SE = standard error, edf = estimated degrees of freedom, Ref.df = reference degrees of freedom, Sig = significance, n.s. = not significant.

L. Uncinate GAM_γ Parametric Coefficients

	Est	**SE**	**t-stat**	**p-value**	**Sig**
(Intercept)	−0.239	0.003	−79.66	<.001	***
Med	0.028	0.004	7.64	<.001	***
High	0.023	0.004	6.06	<.001	***
Male	0.006	0.003	1.92	0.055	n.s.

Approximate Significance of Smooth Terms

	edf	**Ref.df**	**F-stat**	**p-value**	**Sig**
s(nodeID):Low	27.91	32.84	757.40	<.001	***
s(nodeID):Med	26.27	31.27	457.50	<.001	***
s(nodeID):High	24.64	29.61	505.40	<.001	***
s(subjectID)	0.00	1	0	0.57	n.s.

L. Uncinate $GAM_{\gamma c}$ Parametric Coefficients

	Est	**SE**	**t-stat**	**p-value**	**Sig**
(Intercept)	−1.996	0.155	−12.91	<.001	***
Med	0.042	0.004	11.14	<.001	***
High	0.052	0.004	13.11	<.001	***
Male	0.002	0.025	0.08	0.93	n.s.

Approximate Significance of Smooth Terms

	edf	**Ref.df**	**F-stat**	**p-value**	**Sig**
s(nodeID):Low	28.70	33.56	867.37	<.001	***
s(nodeID):Med	27.16	32.13	521.00	<.001	***
s(nodeID):High	25.53	30.52	574.01	<.001	***
s(PDS):Female	7.89	8.00	105.47	<.001	***
s(PDS):Male	6.81	6.98	90.16	<.001	***
s(subjectID)	0.99	1	126.25	<.001	***

differences were associated with measures of anxiety (R Code 4).

This produced a set of splines that modeled tract FA values while also controlling for group, sex, and a measure of development, potentially allowing for more precise between-group testing (Fig. 2, top). The resulting covariate model $GAM_{\gamma c}$ was then compared against GAM_{γ}, and these analysis indicated that incorporating the developmental covariate increased model fit despite the added complexity ($GAM_{\gamma c}$ $-REML = -9336$; AIC difference $= 1222$; $\chi^2_{(4)} = 533.4$, $p < .001$). Accordingly, while GAM_{γ} demonstrated an $R^2_{adj} = 0.82$, adding a covariate resulted in an $R^2_{adj} = 0.85$ (Table 2, L. Uncinate). Parametric coefficients and smooth terms for $GAM_{\gamma c}$ were largely identical to GAM_{γ} (Table 3, bottom), while PDS interacted with sex.

By incorporating a developmental covariate into the GAM, we were able to produce a spline best fitting the tract FA data for each group while controlling for sex and PDS (Fig. 2, top). Next, we aimed to identify aspects of the tract that differed significantly between groups.

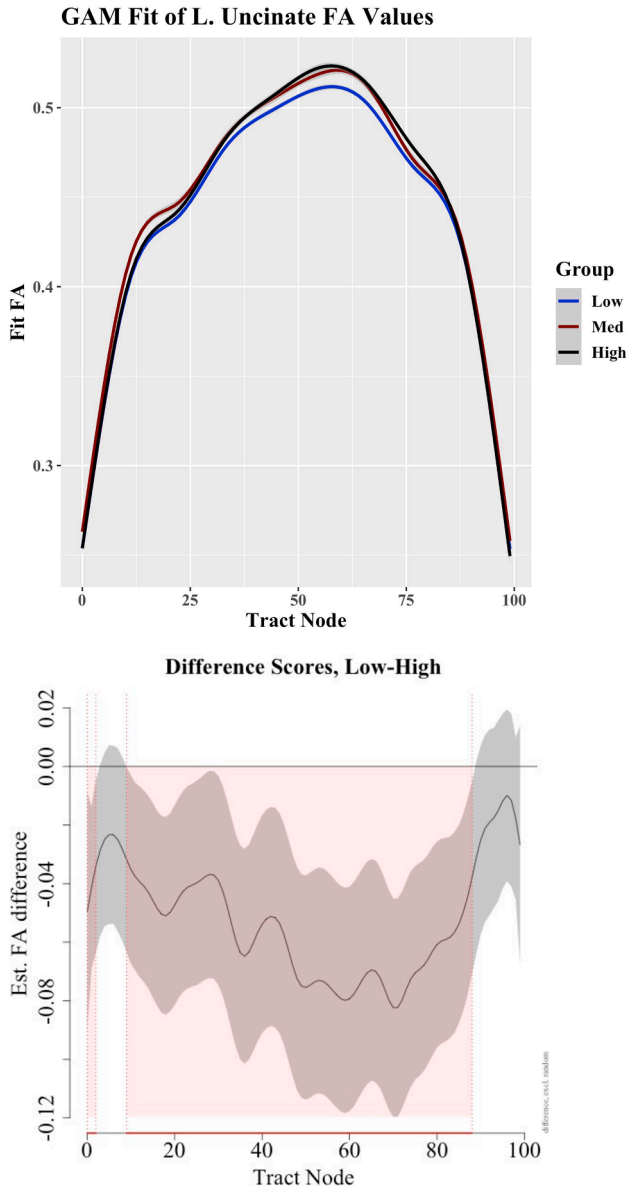

Fig. 2. $GAM_{\gamma c}$ modeling of left uncinate FA values. Top, separate splines are produced for each PARS-6 tertile group. Bottom, estimated FA differences of nodes which differ between Low and High group splines. Red shading indicates regions of significant difference, comparison is Low minus High PARS-6 group (Low–High).

The R tool *plot_diff* (R Code 4) compares only two groups, and so we elected to compare the Low PARS-6 group to the High (although note that *plot_diff* also generates a data frame of model fit estimates and standard error which could be used to investigate interactions with >2 groups). This Low–High spline comparison identified all nodes for which the two group splines differed significantly (Fig. 2, bottom). Specifically, 83 nodes were significantly greater in FA fit estimates for the High compared to the Low group while controlling for sex and PDS. Node number 71 demonstrated the greatest FA fit estimate in the High relative to the Low PARS-6 group.

3.2. GAM of the right uncinate and anterior forceps

We used the same methods as detailed above to model both the right uncinate and anterior forceps. For the right uncinate, the distribution of FA values could be described with either a gamma or beta function just as the left uncinate (distribution mean $= .47 \pm .08$, skewness $= -.92$, kurtosis $= 3.77$). Fit statistics indicated that GAM_{γ} better fit the data than GAM_{β} (Table 2, R. Uncinate), and that adding a covariate ($GAM_{\gamma c}$) also improved model fit ($\chi^2_{(4)} = 461.8$, $p < .001$). The distribution for the anterior forceps had less skewness, however, resulting in a distribution that could be described with either a Gaussian, beta, or gamma function (distribution mean $= .57 \pm .14$, skewness $= .03$, kurtosis $= 2.34$). We used each distribution in separate GAM models, and we found the Gaussian function best fit the distribution of the data (Table 2, A. Forceps). Finally, as with the bi-hemisphere uncinate fasciculi, adding a covariate increased model fit (GAM_{Gc}: $\chi^2_{(4)} = 129.4$, $p < .001$).

For parametric coefficients, each factor in the right uncinate $GAM_{\gamma c}$ significantly differed from the intercept while sex (Male) did not differ from the intercept for the anterior forceps GAM_{Gc} model (Table 4), similar to the left uncinate $GAM_{\gamma c}$ (See Table 3). Additionally, each smooth term differed significantly from flatness, and the interactions of

Table 4
Statistics for the right uncinate fasciculus $GAM_{\gamma c}$ (Top) and anterior forceps GAM_{Gc} (Bottom). Est = model estimate, SE = standard error, edf = estimated degrees of freedom, Ref.df = reference degrees of freedom, Sig = significance, n.s. = not significant.

R. Uncinate $GAM_{\gamma c}$ Parametric Coefficients					
	Est	**SE**	**t-stat**	**p-value**	**Sig**
(Intercept)	−1.24	0.13	−9.56	<.001	***
Med	0.06	0.00	18.44	<.001	***
High	0.08	0.00	22.14	<.001	***
Male	0.05	0.01	3.33	<.001	***
Approximate Signficance of Smooth Terms					
	edf	**Ref.df**	**F-stat**	**p-value**	**Sig**
s(nodeID):Low	30.96	35.43	1464.12	<.001	***
s(nodeID):Med	28.48	33.35	999.09	<.001	***
s(nodeID):High	27.49	32.43	1030.13	<.001	***
s(PDS):Female	7.93	8.00	99.84	<.001	***
s(PDS):Male	6.59	6.90	47.17	<.001	***
s(subjectID)	0.99	1	68.24	<.001	***
A. Forceps GAM_{Gc} Parametric Coefficients					
	Est	**SE**	**t-stat**	**p-value**	**Sig**
(Intercept)	−2.97	0.17	−16.97	<.001	***
Med	−0.03	0.004	−6.23	<.001	***
High	0.02	0.004	4.85	<.001	***
Male	0.04	0.03	1.51	.13	n.s.
Approximate Signficance of Smooth Terms					
	edf	**Ref.df**	**F-stat**	**p-value**	**Sig**
s(nodeID):Low	35.34	38.13	1446.61	<.001	***
s(nodeID):Med	32.84	36.76	872.65	<.001	***
s(nodeID):High	33.13	36.95	792.39	<.001	***
s(PDS):Female	7.52	7.94	21.40	<.001	***
s(PDS):Male	6.82	6.98	14.42	<.001	***
s(subjectID)	0.99	1	358.45	<.001	***

group × tract node as well as PDS × sex were significant (Table 4). The resulting splines modeled the right uncinate and anterior forceps for each group while controlling for sex and PDS (Fig. 3, top). Finally, a large number of nodes ($n = 78$) in the right uncinate showed significantly greater FA fit estimates for the High compared to the Low group while controlling for sex and PDS, whereas the anterior forceps had demonstrably fewer nodes which exhibited a similar significant difference ($n = 39$, Fig. 3, bottom).

R Code 6: General Linear Model to test for group differences in tract FA values at each node, controlling for sex and PDS. dti_fa = FA value of node, group = PARS-6 tertile group, pds = PDS, <family> = a gamma family was used for the uncinate tracts and a Gaussian family for the anterior forceps, df_node = data frame for tract node, "glm" from stats version 4.0.3.

```
fit_glm <- glm(dti_fa ~ group + sex + sex * pds,
    family = <family>(link = "logit"),
    data = df_node)
```

3.3. Point-wise analyses via a general linear model

To illustrate differences between the proposed GAM model and the more traditional method of conducting 'point-wise' or corrected multiple comparison analyses, we modeled the data at each node utilizing a General Linear Model (GLM). A GLM is capable of accounting for the distribution and (bounded) nature of the scalar data while investigating group differences and controlling for covariates and factors. To this end, a GLM was utilized to investigate PARS-6 group differences in tract node FA values while controlling for sex and PDS (R Code 6).

Next, to better relate the results of the GLM 'point-wise' analyses to the outcomes of comparing GAM splines in the sections above, we then extracted the High versus Low PARS-6 group statistic from each GLM and applied an FDR correction. The 'point-wise' method detected nodes that exhibited significantly greater average FA in the High compared to the Low group in only the right uncinate but not the left uncinate or anterior forceps (Fig. 4).

3.4. Correlating memory performance with FA differences

While the main aim of this work was to address limitations in DWI analyses, we additionally desired to demonstrate the utility of our proposed approach within the context of a larger experiment where DWI metrics were not the final aim. Accordingly, we next sought to determine whether the group tract differences detected above were associated with memory performance. The eMST paradigm elicits correct and incorrect

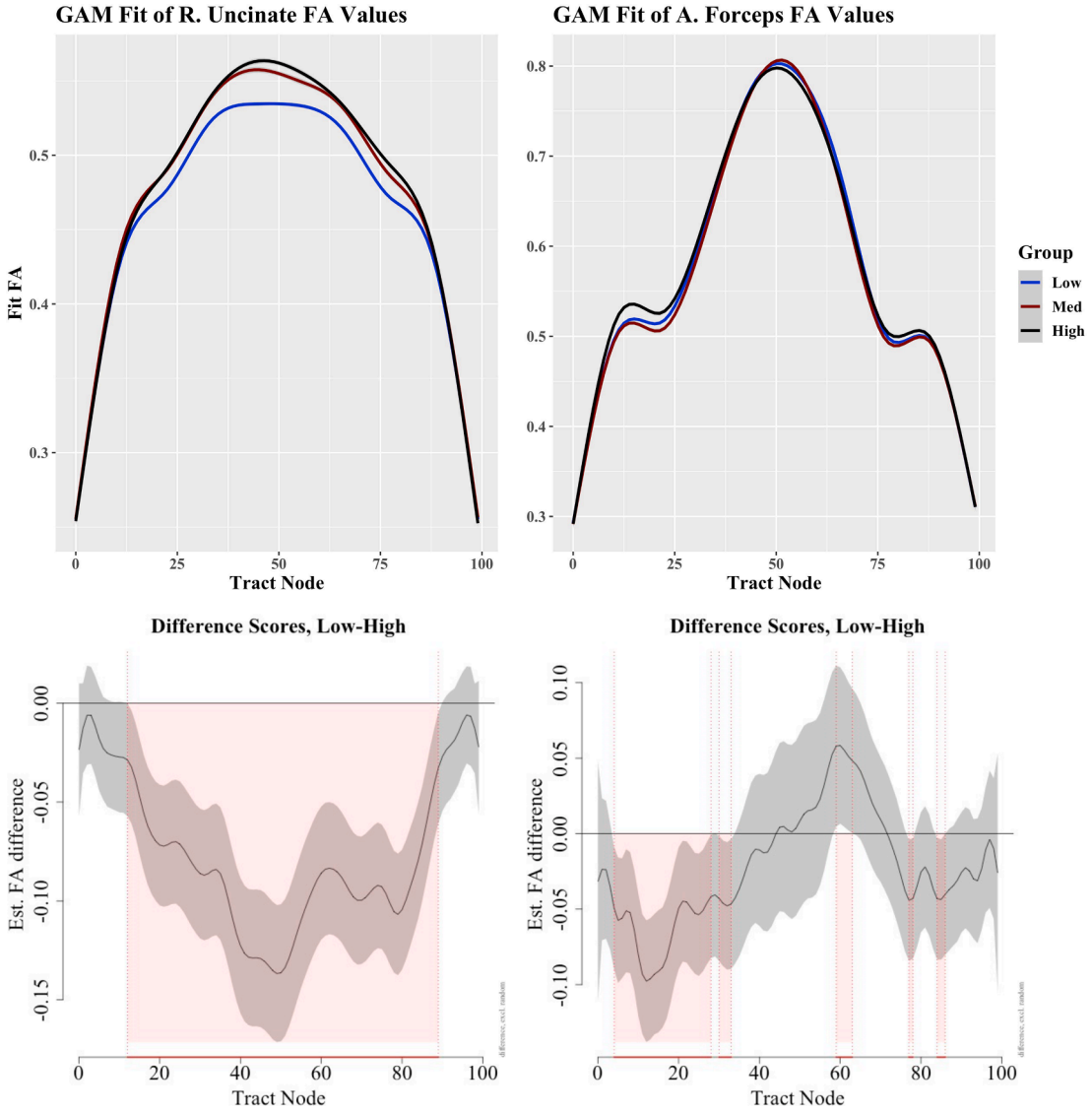

Fig. 3. Top, splines modeling the right uncinate (left) and anterior forceps (right) FA values for each PARS-6 tertile group. Bottom, associated estimated FA differences between Low and High PARS-6 groups, red shading indicates regions of significant differences.

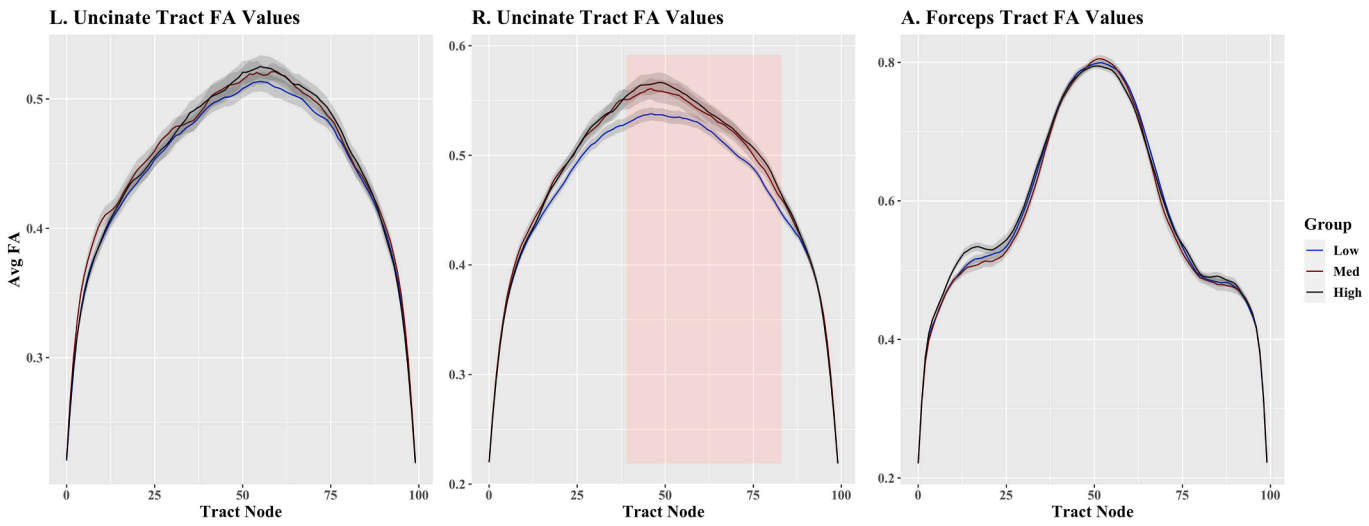

Fig. 4. Average tract FA values for each node plotted by PARS-6 tertile group with standard error margins. Red shading = nodes which significantly differ between Low and High groups following an FDR correction.

Target, Lure, and Foil test responses for negative, neutral, and positive stimuli, and analyses of such a task as well as our preliminary data are reported elsewhere (Leal et al., 2014; McMakin et al., 2021). An investigation into group × memory performances, as measured by lure generalization and detection indices (LGI and LDI, respectively, where LGI = $p(Old|Lure) - p(Old|Foil)$ and LDI = $p(New|Lure) - p(New|Target)$), was conducted via MANOVA testing utilizing two within-subject factors (valence, memory) and one between-subject factor (group). While this analysis revealed a significant interaction of group × memory ($F_{(2,70)} = 5.7$, $p < .01$, $\eta^2_G = .034$) as well as valence × memory ($F_{(1,70)} = 28.12$, $p < .001$, $\eta^2_G = .077$), a valence × memory × group interaction was not detected ($F_{(2,70)} = 1.38$, $p = .25$, $\eta^2_G = .008$). Consequently, we conducted an exploratory analysis wherein we removed the main effects of valence and memory in order to test only the interaction of valence-memory × group. That is, we combined the valence and memory factors in order to have one within-subject factor (valence-memory) and one between-subject factor (group) for MANOVA testing; this adjustment, while not recommended, adds a degree of freedom and gives more power to detect an interaction as the model is not testing main effects. This analysis revealed group differences on negative LGI performance ($F_{(2,70)} = 3.85$, $p = .025$, $\eta^2_p = .1$), where the High PARS-6 group demonstrated significantly higher LGI scores than the Low group (*Tukey's HSD* $p = 0.02$, 95% *CI* $[-.23, -.01]$, Fig. 5); all other statistics were not significant. For the sake of transparency, we explicitly note that the omnibus *F*-statistic of this exploratory model does not survive a Bonferroni correction ($F_{(8,136)} = 2.14$, $p = .035$, $\eta^2_p = .11$). The nature of this exploratory analysis tempers any conclusions or interpretations, but we remark that a subtle group difference, if true, may be "washed out" in a large MANOVA, and that any behavioral differences are impressive given that memory test data were acquired from periadolescent participants seven days after initial encoding. Additionally, such group differences on LGI are consistent with previous work, given the elevated likelihood of negative overgeneralization in anxious populations (Dymond et al., 2015; Greenberg et al., 2013), a developmental predisposition towards generalization (Keresztes et al., 2017; Lavenex and Banta Lavenex, 2013; Leal et al., 2014), and our prior work elucidating potential neural mechanisms of negative overgeneralization in a similar sample of anxious youth (McMakin et al., 2021). Nevertheless, critical evaluation is warranted and we reiterate that these analyses are supplied in order to demonstrate how the outcome of modeling tract profiles with GAMs may be utilized to investigate other questions.

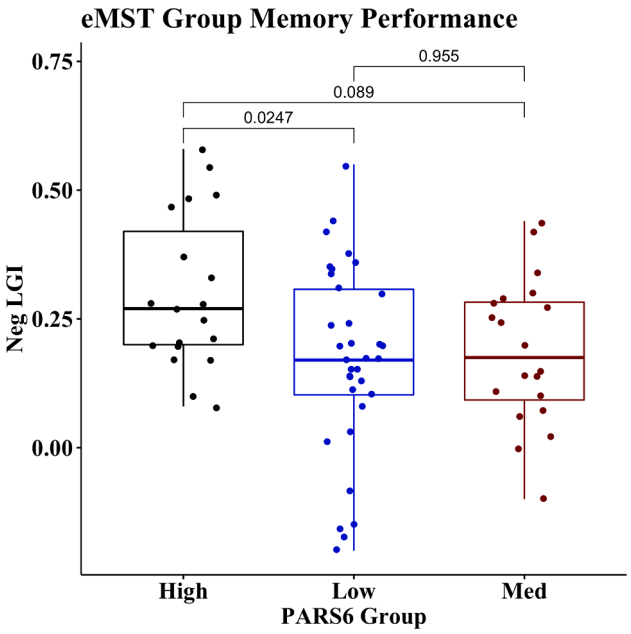

Fig. 5. Box plot of PARS-6 group (Low, Medium [Med], High) of Lure Generalization Indices (LGI) on negative stimuli at one week. Tukey's HSD *p*-values are reported.

To determine whether spline differences were associated with negative LGI performance, the average FA value for the nodes which differed between Low and High PARS-6 groups was calculated for each tract. A multiple linear regression then tested each tract for whether the averaged FA value was associated with negative LGI. Of the three tracts tested, only values from the left uncinate were significantly positively related with negative LGI ($R^2_{adj} = .18$; Table 5). Further, significance testing of the slope differences, given the average FA × group interaction, was significant ($F_{(3,49)} = 4.75$, $p < .01$, $\eta^2_p = .08$; Fig. 6).

4. Discussion

Automated fiber quantification (AFQ) was used to generate bilateral uncinate and anterior forceps tracts for 73 periadolescent participants. The fractional anisotropic (FA) values from nodes along the tracts were

Table 5
Multiple linear regression coefficients for predicting Low and High PARS-6 group negative LGI performance from tract FA differences. FAavg = averaged FA of nodes which differed between splines. Est = model estimate, SE = standard error.

	L. Uncinate FA differences				
	Est	**SE**	**t-stat**	**p-value**	**Sig**
(Intercept)	−0.9509	0.4389	−2.167	0.035	*
FAavg	2.4262	0.9458	2.565	0.013	*
Group	0.8339	0.3763	2.216	0.031	*
FAavg:Group	−1.6583	0.8030	−2.065	0.044	*
	R. Uncinate FA differences				
	Est	**SE**	**t-stat**	**p-value**	**Sig**
(Intercept)	−0.7405	0.6390	−1.159	0.252	n.s.
FAavg	1.8335	1.2817	1.431	0.159	n.s.
Group	0.8167	0.5227	1.562	0.125	n.s.
FAavg:Group	−1.4926	1.0256	−1.455	0.152	n.s.
	A. Forceps FA differences				
	Est	**SE**	**t-stat**	**p-value**	**Sig**
(Intercept)	0.1601	0.7216	0.222	0.825	n.s.
FAavg	0.0230	1.3257	0.017	0.986	n.s.
Group	−0.0387	0.9907	−0.039	0.969	n.s.
FAavg:Group	0.1841	1.7983	0.102	0.919	n.s.

then modeled using a generalized additive model (GAM) that (a) accounted for the distribution and non-continuous nature of FA values, (b) accounted for the interdependence of the data points, and (c) avoided the multiple comparisons issue by comparing differences between spline fits. Using these methods, we identified tract nodes which were significantly greater in the High compared to Low anxiety groups while accounting for sex and pubertal status (PDS). We then tested whether the tract differences were associated with performance on a memory task, and demonstrated that only FA values in the left uncinate were significantly positively related to negative memory generalization.

In this paper, we elucidated three limitations in using traditional ANOVAs to analyze FA values and introduced generalized additive models as a solution to said limitations. Here, we extended the smoothing spline method presented in Carbine et al. (Carbine et al., 2020) to a GAM, which allows for multiple smoothing functions, random effects, and linear covariates. Accordingly, utilizing a GAM to model diffusion-weighted imaging data is a potential resolution to the various issues noted earlier.

Beginning with the multiple comparison issue, a univariate smoothing spline modeling the FA values along the 100 equidistant nodes significantly reduces the number of multiple comparisons needed to compare results between groups as noted in Carbine et al. (Carbine et al., 2020). Using the nodes as a covariate to predict the FA values from the AFQ method, the smooth spline helps researchers avoid the pointwise analysis issues while still allowing for the comparison of penalized regressions between groups. Without including smoothing functions, the model may under-estimate the true relationship between nodes and FA values, increasing the chance of a Type-II error. Finally, using a penalized regression with linear covariates addresses the multiple comparison problem as researchers can include an indicator variable for groups, allowing for the comparison of group-specific splines which model the data while accounting for other factors and covariates.

Another issue often occurring with FA values is the non-normal, nonlinear, and proportional nature of the data. With a sum of smoothing functions, we are more able to account for the non-linearities in the data. As a generalization of an additive model, a GAM uses link functions to account for non-Gaussian distributions of data. Proper distribution modeling improves the sensitivity and specificity of the analysis, where using statistical tests with normality assumptions to model non-normal data will lead to confidence intervals that are either too narrow (inflated Type-I error) or too wide (inflated Type-II error). Next, the utility of link functions in the GAM can address the proportional nature of the FA data by transforming the data into an unbounded, continuous scale. Additionally, we note that GAMs are appropriate for use with scalars aside from FA (see Supplemental Materials, Section 5.1).

Third, and most importantly, a GAM is able to take the spatial dependencies inherent in DWI data into account. In linear regression (including ANOVAs), violating the independence assumption can seriously inflate both Type-I and Type-II error rates. In fact, Forstmeier et al. (Forstmeier et al., 2017) argue the non-independence of data points is one cause of the current replication crisis. By ignoring the spatial dependence when testing AFQ-derived node FA values, the resulting *p*-values may be artificially inflated. A GAM accounts for this issue by modeling nodes as a covariate or random effect when predicting FA values. Further, a GAM in R allows for the inclusion of a correlation matrix if there are more serious spatial dependencies in the data. In addition, including random effects in a GAM can help alleviate another often violated linear regression assumption of homoscedasticity. While minor heteroscedasticity is not much of a concern, exceedingly large heteroscedasticity (often seen in DWI data) will increase the Type-I error. Including random effects in a GAM can help mediate the heterogeneous variance by accounting for subject-level and trial-level variance. In short, GAMs have a strong potential and utility for modeling diffusion-weighted imaging data; the flexibility in specifying the type of smoothing function as well as the ability to include several smoothing methods in a single model can encompass interactions of factors while controlling for covariates.

Finally, we utilized methods more commonly employed with AFQ-

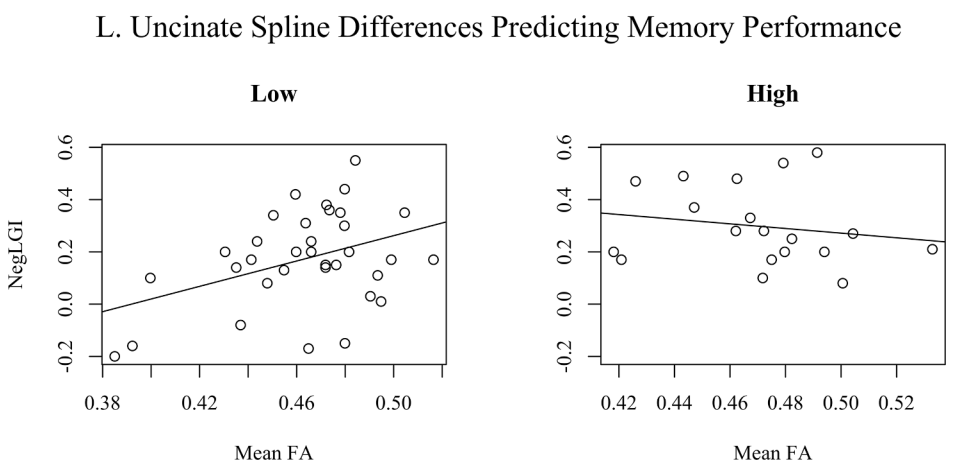

Fig. 6. Multiple linear regression of the average (mean) FA values for nodes which differed between group splines and participant lure generalization index to negative stimuli (NegLGI).

derived DWI data, namely by conducting a series of 'point-wise' analyses along the tract and then applying a multiple comparison correction (Section 3.3). In these analyses we elected to utilize a General Linear Model (GLM), rather than the more common ANOVA-styled tests (Table 1), because in doing so we were still able to account for the non-normal and proportional nature of the data while also including factors and covariates. Fewer nodes were identified as differing between the Low and High PARS-6 groups, implying a difference in sensitivities between the GLM and GAM approaches. As a GAM can also additionally account for multiple comparisons and spatial dependencies in the data, we propose that a GAM is the superior approach but also encourage the use of GLM-styled analyses if a certain research question cannot utilize the GAM approach, perhaps due to limitations of the technique.

As with all statistical techniques, there are limitations in modeling DWI data with GAMs. First, while GAMs are well-suited to model the distribution, bounded, and interdependent nature of individual DWI scalars (such as FA) derived from a single tract, they are not currently capable of modeling multiple scalars across multiple tracts in a single, multivariate model, as would be ideal when investigating traumatic brain injuries. In these instances, multiple models would still be required for each scalar × tract, and appropriate corrections applied. Here, we modeled FA values to demonstrate the utility of GAMs as FAs are a common metric, but incorporating all scalars into a statistical model while accounting for their interdependence remains an open issue and is beyond the scope of the current proposal. A second limitation is the propensity of a GAM to overfit the data, particularly when using a generalized cross-validation method. Fortunately, we can place a heavier penalty on the effective degrees of freedom to counteract a potential overfit; this correction can be used in the GAM function in the *mgcv* package in R (option $gamma = x$). Third, the GAM approach that we detail here involves the comparison between group splines, as the R package *plot_diff* is only written to compare the parameter estimates and standard deviation of one spline against another. If the research question necessitates the comparison of multiple splines, then a multiple comparison correction would still need to occur. This is still an improvement, however, as such a correction would be smaller than the traditional 'point-wise' method. Additionally, as it is possible to construct a data frame of model estimates and standard errors, a multivariate comparison could be constructed. Finally, a GAM may not be ideal when the interaction of multiple factors is central to the aims of the study. We recommend both "Mixed effects models and extensions in ecology with R" by Alain Zuur (Zuur et al., 2009) and "Generalized additive models: An introduction with R" by Simon Wood (Wood, 2017) for more examples, code, and information on analyzing data with a GAM. In sum, we note that a "perfect" statistical model does not exist, but that each approach has strengths and weaknesses. Here, we detail how a GAM is likely a stronger method for modeling diffusion scalars derived from a tract than the majority of extant approaches, but are also well aware that such an approach does not resolve every statistical issue with DWI data. We propose that utilizing these approaches will better fit the data, thereby increasing sensitivity in analyses, but acknowledge difficulties remain.

The example analyses employed to demonstrate modeling DWI data with GAMs had a number of results. First, we note that incorporating a covariate for pubertal development (PDS) increased model fit in all tracts, and correspondingly, the GAMs explained a large portion of variance in modeling tract FA values (Table 2). Such strong model fits demonstrate the flexibility of the GAM method and its utility for DWI analyses. Further, these results show the appropriateness of accounting for pubertal stage in modeling periadolescent data: as myelination processes are demonstrably correlated with development, and pubertal onset is related to sex, including a covariate helped model fit by accounting for this sex × developmental interaction that would impact tract FA values (Dumontheil, 2016; Østby et al., 2009). We also note that we elected to not include age as a covariate in the model; the age range in this sample is narrow, age is collinear with puberty, and age may have lower explanatory power given the interaction of sex, age, and pubertal onset. Second, and correspondingly, we did not detect a main effect of sex (Male) in the parametric coefficients of either the left uncinate (Table 3) or the anterior forceps (Table 4). In these analyses, instead, the pubertal development term (PDS) interacted significantly with both sexes, suggesting that these structures are largely similar between the two sexes and any difference is a function of development. The Male parametric coefficient of the right uncinate, however, differed significantly from the intercept (Female) and potentially implicates a sexual dimorphism in this tract.

Third, group × tract node interactions were detected between Low and High anxiety (PARS-6) groups, particularly in the left uncinate. In all instances, the Low anxiety group had lower FA values at the differential nodes than the High group (Fig. 2). This result was surprising. Previous work has consistently detected a negative relationship between the left uncinate FA values and measures of anxiety, and often interpret such a relationship as suggestive of insufficient executive down-regulation of limbic systems (e.g. Tromp et al., 2012; Liao et al., 2014; Hanson et al., 2015; Ho et al., 2017; Hein et al., 2018; Jamieson et al., 2021). A number of differences set our work apart from previous studies, however, and while a full methodological investigation of the impact of differing processing pipelines, software, and statistical models on the uncinate-anxiety outcomes is beyond the scope of this paper, we postulate that the unique characteristics of this study and dataset explain our discrepant findings: (A) Rather than recruit convenience samples of an adolescent population and subsequently administer anxiety assessments, we recruited clinically anxious periadolescent individuals from pediatric anxiety clinics in addition to the no-psychiatric-diagnosis control participants. During recruitment, we specifically excluded current or past diagnosed mood disorders given the parent grant interest in periadolescence as a sensitive developmental window when processes relevant to the progression from anxiety to depression may be taking shape. Our exclusion of mood disorders was to ensure that we would be able to characterize the developmental progression, and accordingly our sampling was uniquely focused on anxiety. In our review of the literature, these recruitment efforts yielded a unique data set which rather specifically targets clinical anxiety at a narrow window in development. Previous work investigating anxiety and the uncinate fasciculus have largely utilized either young adult clinical populations (Phan et al., 2009; Tromp et al., 2012; Modi et al., 2013; Hanson et al., 2015), community samples (Ho et al., 2017), or individuals in a different pediatric age range (Liao et al., 2014; Hein et al., 2018). (B) We opted to use probabilistic tractography when modeling the DWI data. Although a deterministic approach would still be well-modeled with a GAM, and was the major tractographic method employed in the reviewed literature, recent work suggests that a probabilistic approach may be superior when tracing a specific tract (Petersen et al., 2017; Sarwar et al., 2019; Sotiropoulos and Zalesky, 2019). Follow-up work is planned to more fully elucidate the role of tractography in the study of the uncinate fasciculus, development, anxiety, and memory. (C) The AFQ method produces weighted FA values for each node along the tract. Weighted FA values account for the Mahalanobis distance of nearby voxels and this results in node FA values which account for local tract properties. Further, when modeling tracts from FA values to test for group differences, we accounted for the interaction of sex and pubertal development stage. As far as we are aware, this is a unique approach in the study of the uncinate fasciculus-anxiety relationship, and even similar work done by Ho et al., (Ho et al., 2017), who also utilized AFQ to study early life stress in a periadolescent group, averaged FA values for regions of the uncinate fasciculus and did not include sex or developmental covariates. (D) We utilized a DWI protocol that capitalized on multi-band acceleration factors, multiple shells, and a large number of directions. We also note that, somewhat uniquely, all participants were scanned within an early-afternoon window (between 11:00 and 15:00). As both scanning protocol and time-of-day are known to affect diffusion metrics (Tanner, 1962; Jiang et al., 2014; Bernardi et al., 2016; Celik, 2016; Jones, 2010;

Barrio-Arranz et al., 2015), we specifically established our protocol to best fit our population (Pines et al., 2020) while controlling for diurnal effects.

In sum, we interpret greater FA values in High relative to Low anxiety groups as indicative of aberrant tract development that is associated with clinical anxiety in a periadolescent population. While greater myelination in High anxiety could explain the elevated FA values, a reduction in intermixing factors such as merging, kissing, branching, and/or crossing along the insular portion of the tract would also result in increased FA values, but the investigation of these possibilities are beyond the scope of this work. If greater uncinate FA values in High relative to Low anxiety groups are indeed the result of "increased structural integrity", rather than a reflection of decreased local intermixing, then one possible interpretation would be that clinical anxiety in this population is associated with greater availability of emotionally valenced information resulting from more efficient amygdaloid-frontal connectivity via uncinate projections (Heide et al., 2013; Eden et al., 2015; Baur et al., 2012). To find convergence with the extant literature, then, it is possible that myelination could subsequently be affected in later pubertal development, resulting in the typically observed negative correlation of uncinate FA values and anxiety, possibly due to the adverse effects of hypercortisolaemia on myelination (Piasecka et al., 2020; Chen et al., 2020; Wong et al., 2013; Garg and Mittal, 2020; Jamieson et al., 2021), but this is conjecture.

Fourth, and finally, when correlating tract differences with negative memory overgeneralization (LGI), only the left uncinate group differences were related in any way with memory generalization despite the fact that group FA differences were detected in each tract. We note that these test statistics did not survive a Bonferroni correction despite decent effect sizes, and so while we will interpret them it is possible that we are merely capitalizing on noise. The detection of the left uncinate association with memory generalization relates to recent work by Granger et al., (Granger et al., 2021), who demonstrated in an identical eMST paradigm that only differences in uncinate diffusion metrics were associated with medial temporal processes during correct lure discrimination. Where Granger et al., (Granger et al., 2021) did not detect a relationship between uncinate integrity and lure discrimination, however, we demonstrate a group × left uncinate FA interaction during negative lure generalization. This interaction is interesting as (a) the High anxiety group had significantly larger FA values (Fig. 2), (b) High anxiety participants were significantly more likely to overgeneralize to negatively valenced stimuli (Fig. 5), and (c) a positive correlation between negative memory generalization and uncinate FA was detected only in the Low anxiety group (Fig. 6). Taken together, it appears that greater left uncinate FA values are deleterious to negative Lure test performance and this corresponds with our postulation above where prefrontal regions receive more emotionally valenced information in clinically anxious periadolescent populations. In this interpretation, the information carried by the uncinate would largely reflect generalization processes, possibly driven by amygdaloid responses to negatively valenced information. As a positive association of negative generalization and uncinate FA values are detected in Low but not High anxiety groups, it is possible that data from clinically anxious children demonstrate a ceiling effect in the FA × negative generalization interaction, given their high uncinate FA values, increased propensity to generalize, and flat FA × negative LGI correlation.

In conclusion, we demonstrated the utility of modeling AFQ-derived diffusion metrics with a generalized additive model. We found such models to be robust in their application, fitting the aspects of DWI data well while addressing substantial, extant issues in DWI analyses. We also found GAMs to be externally valid, detecting tract differences which were predictive of independent behavior measures.

CRediT authorship contribution statement

Nathan M. Muncy: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing - original draft, Visualization. **Adam Kimbler:** Software, Formal analysis, Data curation, Writing - review & editing. **Ariana M. Hedges-Muncy:** Conceptualization, Methodology, Writing - review & editing. **Dana L. McMakin:** Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Aaron T. Mattfeld:** Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgments

We would like to thank the Center for Imaging Sciences at Florida International University and the EMU project staff for their aid in collecting these data. We also thank the National Institute of Mental Health (NIMH) for funding which supported this project (R01 MH116005). Finally, we thank Paola Parrales for her help with editing the manuscript and Stephanie N. Hedges M.S. for the inspiration.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.nicl.2022.102937.

References

Adluru, N., Luo, Z., Van Hulle, C.A., Schoen, A.J., Davidson, R.J., Alexander, A.L., Goldsmith, H.H., 2017. Anxiety-related experience-dependent white matter structural differences in adolescence: A monozygotic twin difference approach. Sci. Rep. 7 (1), 8749. https://doi.org/10.1038/s41598-017-08107-6 Bandiera abtest: a Cc license type: cc by Cg type: Nature Research Journals Primary atype: Research Subject term: Anxiety;Statistics Subject term id: anxiety;statistics.

Ashburner, J., 2003. How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. NeuroImage 20 (2), 870–888. https://doi.org/10.1016/S1053-8119(03)00336-7.

Andersson, J.L.R., Sotiropoulos, S.N., 2016. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. NeuroImage 125, 1063–1078. https://doi.org/10.1016/j.neuroimage.2015.10.019.

Angelopoulou, G., Meier, E.L., Kasselimis, D., Pan, Y., Tsolakopoulos, D., Velonakis, G., Karavasilis, E., Kelekis, N.L., Goutsos, D., Potagas, C., Kiran, S., 2019. In vestigating Gray and White Matter Structural Substrates of Sex Differences in the Narrative Abilities of Healthy Adults. Front. Neurosci. 13, 1424. https://doi.org/10.3389/fnins.2019.01424.

Banfi, C., Koschutnig, K., Moll, K., Schulte-Körne, G., Fink, A., Landerl, K., 2019. White matter alterations and tract lateralization in children with dyslexia and isolated spelling deficits. Hum. Brain Mapp. 40 (3), 765–776. https://doi.org/10.1002/hbm.24410.

Barrio-Arranz, G., de Luis-García, R., Tristáan-Vega, A., Martín-Fernaández, M., Aja Fernáandez, S., 2015. Impact of MR Acquisition Parameters on DTI Scalar Indexes: A Tractography Based Approach. PLOS ONE 10 (10), e0137905. https://doi.org/10.1371/journal.pone.0137905.

Baur, V., Hänggi, J., Jäncke, L., 2012. Volumetric associations between uncinate fasciculus, amygdala, and trait anxiety. BMC Neuroscience, 13 (1), 4. https://doi.org/10.635 1186/1471-2202-13-4.

Behrens, T.E., Berg, H.J., Jbabdi, S., Rushworth, M.F.S., Woolrich, M.W., 2007. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? NeuroImage 34 (1), 144–155. https://doi.org/10.1016/j.neuroimage.2006.09.018.

Behrens, T.E., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., Matthews, P.M., Brady, J.M., Smith, S.M., 2003. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. Magnet. Resonance Med. 50 (5), 1077–1088.

Bernardi, G., Cecchetti, L., Siclari, F., Buchmann, A., Yu, X., Handjaras, G., Bellesi, M., Ricciardi, E., Kecskemeti, S.R., Riedner, B.A., Alexander, A.L., Benca, R.M., Ghilardi, M.F., Pietrini, P., Cirelli, C., Tononi, G., 2016. Sleep reverts changes in human gray and white matter caused by wake-dependent training. NeuroImage 129 (649), 367–377. https://doi.org/10.1016/j.neuroimage.2016.01.020.

Cai, W., Zhao, M., Liu, J., Liu, B., Yu, D., Yuan, K., 2019. Right arcuate fasciculus and superior longitudinal fasciculus abnormalities in primary insomnia. Brain Imaging Behav. 13 (6), 1746–1755. https://doi.org/10.1007/s11682-019-00160-1.

Caporino, N.E., Brodman, D.M., Kendall, P.C., Albano, A.M., Sherrill, J., Piacentini, J., Sakolsky, D., Birmaher, B., Compton, S.N., Ginsburg, G., 2013. Defining treatment response and remission in child anxiety: Signal detection analysis using the pediatric anxiety rating scale. J. Am. Acad. Child Adoles-cent Psychiatry 52 (1), 57–67.

Carbine, K.A., Duraccio, K.M., Hedges-Muncy, A., Barnett, K.A., Kirwan, C.B., Jensen, C. D., 2020. White matter integrity disparities between normal-weight and overweight/obese adolescents: An automated fiber quantification tractography study. Brain Imaging Behavior 14 (1), 308–319. https://doi.org/10.1007/s11682-019-00036-4.

Celik, A., 2016. Effect of imaging parameters on the accuracy of apparent diffusion coefficient and optimization strategies. Diagnost. Intervent. Radiol. 22 (1), 101. https://doi.org/10.5152/dir.2015.14440.

Chen, H.-F., Huang, L.-L., Li, H.-Y., Qian, Y., Yang, D., Qing, Z., Luo, C.-M., Li, M.-C., Zhang, B., Xu, Y., 2020. Microstructural disruption of the right inferior fronto occipital and inferior longitudinal fasciculus contributes to WMH-related cognitive impairment. CNS Neurosci. Therapeutics, 26 (5), 576–588. https://doi.org/10.6701111/cns.13283.

Chen, H., Sheng, X., Qin, R., Luo, C., Li, M., Liu, R., Zhang, B., Xu, Y., Zhao, H., Bai, F., 2020. Aberrant White Matter Microstructure as a Potential Diagnostic Marker in Alzheimer's Disease by Automated Fiber Quantification. Front. Neurosci. 14, 570123 https://doi.org/10.3389/fnins.2020.570123.

Chen, Y., Zhang, J., Tan, H., Li, J., Yu, Y., 2020. Detrimental effects of hypercortisolism on brain structure and related risk factors. Sci. Rep. 10 (1), 12708. https://doi.org/10.1038/s41598-020-68166-0.

Clocksin, H.E., Hawks, Z.W., White, D.A., Christ, S.E., 2021. Inter- and intra-tract analysis of white matter abnormalities in individuals with early-treated phenylke tonuria (PKU). Molecular Genetics and Metabolism, 132 (1), 11–18. https://doi.org/10.1016/j.ymgme.2020.12.001.

De Santis, S., Bastiani, M., Droby, A., Kolber, P., Zipp, F., Pracht, E., Stoecker, T., Groppa, S., Roebroeck, A., 2019. Characterizing Microstructural Tissue Properties in Multiple Sclerosis with Diffusion MRI at 7 T and 3 T: The Impact of the Experimental Design. Neuroscience 403, 17–26. https://doi.org/10.1016/j.neuroscience.2018.03.048.

Delignette-Muller, M.L., Dutang, C., 2015. Fitdistrplus: An R package for fitting distributions. J. Stat. Software, 64 (4), 1–34. https://www.jstatsoft.org/ article/view/v064i04.

Delouche, A., Attyé, A., Heck, O., Grand, S., Kastler, A., Lamalle, L., Renard, F., Krainik, A., 2016. Diffusion MRI: Pitfalls, literature review and future directions of research in mild traumatic brain injury. Eur. J. Radiol. 85 (1), 25–30. https://doi.org/10.1016/j.ejrad.2015.11.004.

Deng, F., Wang, Y., Huang, H., Niu, M., Zhong, S., Zhao, L., Qi, Z., Wu, X., Sun, Y., Niu, C., He, Y., Huang, L., Huang, R., 2018. Abnormal segments of right uncinate fasciculus and left anterior thalamic radiation in major and bipolar depression. Progr. Neuro-Psychopharmacol. Biol. Psychiatry 81, 340–349. https://doi.org/10.1016/j.pnpbp.2017.09.006.

Dou, X., Yao, H., Feng, F., Wang, P., Zhou, B., Jin, D., Yang, Z., Li, J., Zhao, C., Wang, L., An, N., Liu, B., Zhang, X., Liu, Y., 2020. Characterizing white matter connectivity in Alzheimer's disease and mild cognitive impairment: An automated fiber quantification analysis with two independent datasets. Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 129, 390–405. https://doi.org/10.1016/j.cortex.2020.03.032.

Dumontheil, I., 2016. Adolescent brain development. Curr. Opin. Behav. Sci. 10, 39–44. https://doi.org/10.1016/j.cobeha.2016.04.012.

Dymond, S., Dunsmoor, J.E., Vervliet, B., Roche, B., Hermans, D., 2015. Fear Generalization in Humans: Systematic Review and Implications for Anxiety Disorder Research. Behav. Ther. 46 (5), 561–582. https://doi.org/10.1016/j.beth.2014.10.001.

Eden, A.S., Schreiber, J., Anwander, A., Keuper, K., Laeger, I., Zwanzger, P., Zwitser lood, P., Kugel, H., Dobel, C., 2015. Emotion Regulation and Trait Anxiety Are Predicted by the Microstructure of Fibers between Amygdala and Prefrontal Cortex. J. Neurosci., 713 35 (15), 6020–6027. doi: 10.1523/JNEUROSCI. 3659-14, 2015.

Feldman, H.M., Yeatman, J.D., Lee, E.S., Barde, L.H., Gaman-Bean, S., 2010. Diffusion tensor imaging: A review for pediatric researchers and clinicians. J. Devel. Behav. Pediatrics: JDBP 31 (4), 346.

Fischl, B., 2012. FreeSurfer. NeuroImage 62 (2), 774–781.

Forstmeier, W., Wagenmakers, E.-J., Parker, T.H., 2017. Detecting and avoiding likely false-positive findings-a practical guide. Biol. Rev. 92 (4), 1941–1968.

Garg, M.K., Mittal, M., 2020. Structural and Functional Consequences of Hypercorti solism on Brain: Are the Brain and Psycho-neuro-cognitiveManifestations Reversible? Indian J. Endocrinol. Metab. 24 (6), 507–508. https://doi.org/10.4103/ijem.IJEM 542 20.

Ginsburg, G.S., Keeton, C.P., Drazdowski, T.K., Riddle, M.A., 2011. The Utility of Clinicians Ratings of Anxiety Using the Pediatric Anxiety Rating Scale (PARS). Child Youth Care Forum 40 (2), 93–105. https://doi.org/10.1007/s10566-010-9125-3.

Goodrich-Hunsaker, N.J., Abildskov, T.J., Black, G., Bigler, E.D., Cohen, D.M., Mihalov, L.K., Bangert, B.A., Taylor, H.G., Yeates, K.O., 2018. Age- and sex-related effects in children with mild traumatic brain injury on diffusion magnetic resonance imaging properties: A comparison of voxelwise and tractography methods. J. Neurosci. Res. 96 (4), 626–641. https://doi.org/10.1002/jnr.24142.

Granger, S.J., Leal, S.L., Larson, M.S., Janecek, J.T., McMillan, L., Stern, H., Yassa, M.A., 2021. Integrity of the uncinate fasciculus is associated with emotional pattern separation-related fMRI signals in the hippocampal dentate and CA3. Neurobiol. Learn. Mem. 177, 107359 https://doi.org/10.1016/j.nlm.2020.107359.

Greenberg, T., Carlson, J.M., Cha, J., Hajcak, G., Mujica-Parodi, L.R., 2013. Ventro medial prefrontal cortex reactivity is altered in generalized anxiety disorder during fear generalization. Depression Anxiety 30 (3), 242–250.

Hall, S.S., Dougherty, 740 R.F., Reiss, A.L., 2016. Profiles of aberrant white matter microstructure in fragile X syndrome. NeuroImage. Clinical, 11, 133–138. https://doi. org/10.1016/j.nicl.2016.01.013.

Hanson, J.L., Knodt, A.R., Brigidi, B.D., Hariri, A.R., 2015. Lower structural integrity of the uncinate fasciculus is associated with a history of child maltreatment and future psychological vulnerability to stress. Devel. Psychopathol. 27, 1611–1619.

Harrison, D.M., Caffo, B.S., Shiee, N., Farrell, J. a. D., Bazin, P.-L., Farrell, S.K., Ratchford, J.N., Calabresi, P.A., Reich, D.S., 2011. Longitudinal changes in diffusion tensor-based quantitative MRI in multiple sclerosis. Neurology, 76 (2), 179-186. https://doi.org/10.1212/WNL.0b013e318206ca61.

Heide, R.J.V.D., Skipper, L.M., Klobusicky, E., Olson, I.R., 2013. Dissecting the uncinate fasciculus: Disorders, controversies and a hypothesis. Brain 136 (6), 1692. https://doi.org/10.1093/brain/awt094.

Hein, T.C., Mattson, W.I., Dotterer, H.L., Mitchell, C., Lopez-Duran, N., Thomason, M.E., Peltier, S.J., Welsh, R.C., Hyde, L.W., Monk, C.S., 2018. Amygdala habituation and uncinate fasciculus connectivity in adolescence: A multi-modal approach. NeuroImage 183, 617–626.

Ho, T.C., King, L.S., Leong, J.K., Colich, N.L., Humphreys, K.L., Ordaz, S.J., Gotlib, I.H., 2017. Effects of sensitivity to life stress on uncinate fasciculus segments in early adolescence. Social Cogn. Affective Neurosci. 12 (9), 1460–1469.

Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D.S., Calabresi, P.A., Pekar, J.J., van Zijl, P.C., Mori, S., 2008. Tract probability maps in stereotaxic spaces: Anal yses of white matter anatomy and tract-specific quantification. Neuroimage 39 (1), 336–347.

Huang, L., Chen, X., Sun, W., Chen, H., Ye, Q., Yang, D., Li, M., Luo, C., Ma, J., Shao, P., Xu, H., Zhang, B., Zhu, X., Xu, Y., 2020. Early Segmental White Matter Fascicle Microstructural Damage Predicts the Corresponding Cognitive Domain Impairment in Cerebral Small Vessel Disease Patients by Automated Fiber Quantification. Front. Aging Neurosci. 12, 598242 https://doi.org/10.3389/fnagi.2020.598242.

Huang, Y., Yecies, D., Bruckert, L., Parker, J.J., Ho, A.L., Kim, L.H., Fornoff, L., Winter mark, M., Porter, B., Yeom, K.W., Halpern, C.H., Grant, G.A., 2019. Stereo tactic laser ablation for completion corpus callosotomy. J. Neurosurgery. Pe-diatrics, 1-9. https://doi.org/10.3171/2019.5.PEDS19117.

Hutchinson, E.B., Schwerin, S.C., Avram, A.V., Juliano, S.L., Pierpaoli, C., 2018. Diffusion MRI and the detection of alterations following traumatic brain injury. J. Neurosci. Res. 96 (4), 612–625. https://doi.org/10.1002/jnr.24065.

Jamieson, D., Shan, Z., Lagopoulos, J., Hermens, D.F., 2021. The role of adolescent sleep quality in the development of anxiety disorders: A neurobiologically-informed model. Sleep Med. Rev. 101450.

Jiang, C., Zhang, L., Zou, C., Long, X., Liu, X., Zheng, H., Liao, W., Diao, Y., 2014. Diurnal Microstructural Variations in Healthy Adult Brain Revealed by Diffusion Tensor Imaging. PLOS ONE 9 (1), e84822. https://doi.org/10.1371/journal.pone.0084822.

Jiang, H., Li, X., Jin, C., Wang, M., Liu, C., Chan, K.C., Yang, J., 2019. Early Diagnosis of Spastic Cerebral Palsy in Infants with Periventricular White Matter Injury Using Diffusion Tensor Imaging. AJNR 40 (1), 162–168. https://doi.org/10.3174/ajnr.A5914.

Johnco, C.J., Salloum, A., Lewin, A.B., Storch, E.A., 2015. Refining clinical judgment of treatment response and symptom remission identification in childhood anxiety using a signal detection analysis on the Pediatric Anxiety Rating Scale. J. Child Adolescent Psychopharmacol. 25 (9), 674–683.

Jones, D.K., 2010. Precision and Accuracy in Diffusion Tensor Magnetic Resonance Imaging. Top. Magn. Reson. Imaging 21 (2), 87–99. https://doi.org/10.1097/RMR.0b013e31821e56ac.

Jossinger, S., Kronfeld-Duenias, V., Zislis, A., Amir, O., Ben-Shachar, M., 2021. Speech rate association with cerebellar white-matter diffusivity in adults with persistent developmental stuttering. Brain Struct. Funct, 226 (3), 801–816. https://doi. org/10.1007/s00429-020-02210-7.

Keresztes, A., Bender, A.R., Bodammer, N.C., Lindenberger, U., Shing, Y.L., Werkle-Bergner, M., 2017. Hippocampal maturity promotes memory distinctiveness in childhood and adolescence. Proc. Natl. Acad. Sci., 114 (34), 802 9212–9217. doi: 10.1073/pnas.1710654114.

Kreilkamp, B.A.K., Lisanti, L., Glenn, G.R., Wieshmann, U.C., Das, K., Marson, A.G., Keller, S.S., 2019. Comparison of manual and automated fiber quantification tractography in patients with temporal lobe epilepsy. NeuroImage. Clinical 24, 102024. https://doi.org/10.1016/j.nicl.2019.102024.

Kruper, J., Yeatman, J.D., Richie-Halford, A., Bloom, D., Grotheer, M., Caffarra, S., Kiar, G., Karipidis, I.I., Roy, E., Chandio, B.Q., Garyfallidis, E., Rokem, A., 2021. Evaluating the reliability of human brain white matter tractometry. Aperture (in press).

Lavenex, P., Banta Lavenex, P., 2013. Building hippocampal circuits to learn and remember: Insights into the development of human memory. Behav. Brain Res. 254, 8–21. https://doi.org/10.1016/j.bbr.2013.02.007.

Leal, S.L., Tighe, S.K., Yassa, M.A., 2014. Asymmetric effects of emotion on mnemonic interference. Neurobiol. Learn. Mem. 111, 41–48.

Li, J., Wei, X.H., Liu, Y.K., Chen, L.S., Zhu, Z.Q., Hou, S.Y., Fang, X.K., Wang, Z.Q., 2020. Evidence of motor injury due to damaged corticospinal tract following acute hemorrhage in the basal ganglia region. Sci. Rep. 10 (1), 16346. https://doi.org/10.1038/s41598-020-73305-8.

Li, W.P., Wang, F.F., Zhang, X., Li, M., Lu, J.M., Wu, S.C., Zhang, B., 2017. Study on age and white matter neuronal integrity in healthy volunteers based on automating fiber-tract quantification. Zhonghua Yi Xue Za Zhi 97 (13), 976–981. https://doi.org/10.3760/cma.j.issn.0376-2491.2017.13.003.

Li, W.P., Zhao, H., Zhang, X., Liang, X., Liu, Y., Zhang, W., Zhang, B., 2020. Study on the white matter neuronal integrity in amnestic mild cognitive impairment based on automating fiber-tract quantification. Zhonghua Yi Xue Za Zhi 100 (3), 172–177. https://doi.org/10.3760/cma.j.issn.0376-2491.2020.03.003.

Liao, M., Yang, F., Zhang, Y., He, Z., Su, L., Li, L., 2014. White matter abnormalities in adolescents with generalized anxiety disorder: A diffusion tensor imaging study. BMC Psychiatry 14 (1), 1–6.

Lin, Q., Bu, X., Wang, M., Liang, Y., Chen, H., Wang, W., Yi, Y., Lin, H., Zhou, J., Lu, L., Hu, X., Yang, C., Huang, X., 2020. Aberrant white matter properties of the callosal tracts implicated in girls with attention-deficit/hyperactivity disorder. Brain Imaging Behavior, 14 (3), 728-735. https://doi.org/10.1007/s11682-018-0010-2.

McMakin, D.L., Kimbler, A., Tustison, N.J., Pettit, J.W., Mattfeld, A.T., 2021. Negative Overgeneralization is Associated with Anxiety and Mechanisms of Pattern

Completion in Peripubertal Youth. Soc. Cognit. Affective Neurosci. https://doi.org/10.1093/scan/nsab089.
Mesaros, S., Rocca, M.A., Kacar, K., Kostic, J., Copetti, M., Stosic-Opincal, T., Preziosa, P., Sala, S., Riccitelli, G., Horsfield, M.A., Drulovic, J., Comi, G., Filippi, M., 2012. Diffusion tensor MRI tractography and cognitive impairment in multiple sclerosis. Neurology 78 (13), 969–975. https://doi.org/10.1212/WNL.0b013e31824d5859.
Modi, S., Trivedi, R., Singh, K., Kumar, P., Rathore, R.K.S., Tripathi, R.P., Khushu, S., 2013. Individual differences in trait anxiety are associated with white matter tract 841 integrity in fornix and uncinate fasciculus: Preliminary evidence from a DTI based tractography study. Behavioural Brain Research, 238, 188–192. https://doi.org/10. 1016/j.bbr.2012.10.007.
Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: A primer with examples. Hum. Brain Mapp. 15 (1), 1–25. https://doi.org/10.1002/hbm.1058.
Østby, Y., Tamnes, C.K., Fjell, A.M., Westlye, L.T., Due-Tønnessen, P., Walhovd, K.B., 2009. Heterogeneity in subcortical brain development: A structural magnetic resonance imaging study of brain maturation from 8 to 30 years. J. Neurosci. 850, 11772–11782, 29 (38).
Pascual-Diaz, S., Varriano, F., Pineda, J., Prats-Galino, A., 2020. Structural character ization of the Extended Frontal Aslant Tract trajectory: A ML-validated laterality study in 3T and 7T. NeuroImage 222, 117260. https://doi.org/10.1016/j.neuroimage. 54 2020.117260.
Petersen, A., Crockett, L., Richards, M., Boxer, A., 1988. A self-report measure of puber tal status: Reliability, validity, and initial norms. J. Youth Adolesc. 17 (2), 117–133. https://doi.org/10.1007/BF01537962.
Petersen, M., Lund, T.E., Sunde, N., Frandsen, J., Rosendal, F., Juul, N., Østergaard, K., 2017. Probabilistic versus deterministic tractography for delineation of the cortico subthalamic hyperdirect pathway in patients with Parkinson disease selected for deepbrain stimulation. J. Neurosurgery, 126 (5), 1657–1668. https://doi.org/10.3171/2016.4.JNS1624.
Phan, K.L., Orlichenko, A., Boyd, E., Angstadt, M., Coccaro, E.F., Liberzon, I., Arfanakis, K., 2009. Preliminary evidence of white matter abnormality in the uncinate fasciculus in generalized social anxiety disorder. Biol. Psychiatry 66 (7), 691–694.
Piasecka, M., Papakokkinou, E., Valassi, E., Santos, A., Webb, S.M., de Vries, F., Pereira, A.M., Ragnarsson, O., 2020. Psychiatric and neurocognitive consequences of endogenous hypercortisolism. J. Intern. Med. 288 (2), 168–182. https://doi.org/10.1111/joim.13056.
Pines, A.R., Cieslak, M., Larsen, B., Baum, G.L., Cook, P.A., Adebimpe, A., Dávila, D.G., Elliott, M.A., Jirsaraie, R., Murtha, K., Oathes, D.J., Piiwaa, K., Rosen, A.F.G., Rush, S., Shinohara, R.T., Bassett, D.S., Roalf, D.R., Satterthwaite, T.D., 2020. Leveraging multi-shell diffusion for studies of brain development in youth and young adulthood. Developmental Cognitive Neuroscience, 43, 100788. https://doi.org/10. 1016/j.dcn.2020.100788.
R Core Team, 2020. R: A language and environment for statistical computing. Manual. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/.
Raizman, R., Tavor, I., Biegon, A., Harnof, S., Hoffmann, C., Tsarfaty, G., Fruchter, E., Tatsa-Laur, L., Weiser, M., Livny, A., 2020. Traumatic Brain Injury Severity in a Network Perspective: A Diffusion MRI Based Connectome Study. Sci. Rep. 10 (1), 9121. https://doi.org/10.1038/s41598-020-65948-4.
Sacchet, M.D., Prasad, G., Foland-Ross, L.C., Joshi, S.H., Hamilton, J.P., Thompson, P.M., Gotlib, I.H., 2014. Characterizing White Matter Connectivity in Major Depressive Disorder: Automated Fiber Quantification and Maximum Density Paths. Proceedings. IEEE International Symposium on Biomedical Imaging 11, 592–595. https://doi.org/10.1109/ISBI.2014.6867940.
Sacchet, M.D., Prasad, G., Foland-Ross, L.C., Joshi, S.H., Hamilton, J.P., Thompson, P.M., Gotlib, I.H., 2014b. Structural abnormality of the corticospinal tract in major depressive disorder. Biol. Mood Anxiety Disorders, 4, 8. https://doi. org/10.1186/2045-5380-4-8.
Sarica, A., Cerasa, A., Valentino, P., Yeatman, J., Trotta, M., Barone, S., Granata, A., Nisticoó, R., Perrotta, P., Pucci, F., Quattrone, A., 2017. The corticospinal tract profile in amyotrophic lateral sclerosis. Hum. Brain Mapp. 38 (2), 727–739. https://doi.org/10.1002/hbm.23412.
Sarica, A., Valentino, P., Nisticó, R., Barone, S., Pucci, F., Quattrone, A., Cerasa, A., Quattrone, A., 2019. Assessment of the Corticospinal Tract Profile in Pure Lower Motor Neuron Disease: A Diffusion Tensor Imaging Study. Neuro-Degenerative Diseases 19 (3–4), 128–138. https://doi.org/10.1159/000503970.
Sarwar, T., Ramamohanarao, K., Zalesky, A., 2019. Mapping connectomes with diffusion MRI: Deterministic or probabilistic tractography? Magn. Reson. Med. 81 (2), 1368–1384. https://doi.org/10.1002/mrm.27471.
Shenton, M.E., Hamoda, H.M., Schneiderman, J.S., Bouix, S., Pasternak, O., Rathi, Y., Vu, M.-A., Purohit, M.P., Helmer, K., Koerte, I., Lin, A.P., Westin, C.-F., Kikinis, R., Kubicki, M., Stern, R.A., Zafonte, R., 2012. A review of magnetic resonance imaging and diffusion tensor imaging findings in mild traumatic brain injury. Brain Imaging Behav. 6 (2), 137–192. https://doi.org/10.1007/s11682-012-9156-5.
Shirtcliff, E.A., Dahl, R.E., Pollak, S.D., 2009. Pubertal development: Correspondence between hormonal and physical development. Child Devel. 80 (2), 327–337.
Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23, S208–S219. https://doi.org/10.1016/j.neuroimage.2004.07.051.
Sommer, S., Kozerke, S., Seifritz, E., Staempi, P., 2017. Fiber up-sampling and quality assessment of tractograms - towards quantitative brain connectivity. Brain Behav. 7 (1), e00588 https://doi.org/10.1002/brb3.588.
Sotiropoulos, S.N., Zalesky, A., 2019. Building connectomes using diffusion MRI: Why, how and but. NMR Biomed. 32 (4), e3752 https://doi.org/10.1002/nbm.3752.
Stark, S.M., Kirwan, C.B., Stark, C.E.L., 2019. Mnemonic similarity task: A tool for assessing hippocampal integrity. Trends Cognit. Sci. 23 (11), 938–951. https://doi.org/10.1016/j.tics.2019.08.003.
Storch, E.A., Wood, 923 J.J., Ehrenreich-May, J., Jones, A.M., Park, J.M., Lewin, A.B., Murphy, T.K., 2012. Convergent and Discriminant Validity and Reliability of the Pediatric Anxiety Rating Scale in Youth with Autism Spectrum Disorders. J. Autism Developmental Disorders, 42 (11), 2374-2382. https://doi.org/10.1007/ s10803-012-1489-9.
Tanner, J.M., 1962. Growth at adolescence.
Thomas, C., Sadeghi, N., Nayak, A., Treer, A., Sarlls, J., Baker, C.I., Pierpaoli, C., 2018. Impact of time-of-day on diffusivity measures of brain tissue derived from diffusion tensor imaging. NeuroImage, 173, 25-34. https://doi.org/10.1016/j.neuroimage.2018.02.026.
Tromp, D.P., Grupe, D.W., Oathes, D.J., McFarlin, D.R., Hernandez, P.J., Kral, T.R., Lee, J.E., Adams, M., Alexander, A.L., Nitschke, J.B., 2012. Reduced structural connectivity of a major frontolimbic pathway in generalized anxiety disorder. Arch. General Psychiatry 69 (9), 925–934.
Unterrainer, H.F., Hiebler-Ragger, M., Koschutnig, K., Fuchshuber, J., Ragger, K., Perch told, C.M., Papousek, I., Weiss, E.M., Fink, A., 2019. Brain Structure Alterations in Poly-Drug Use: Reduced Cortical Thickness and White Matter Impairments in Regions Associated With Affective, Cognitive, and Motor Functions. Front. Psychiatry, 10, 667. https://doi.org/10.3389/fpsyt.2019.00667.
Vakhtin, A.A., Zhang, Y., Wintermark, M., Massaband, P., Robinson, M.T., Ashford, J.W., Furst, A.J., 2020. White Matter Asymmetry: A Reection of Pathology in Traumatic Brain Injury. J. Neurotrauma 37 (2), 373–381. https://doi.org/10.1089/neu.2019.6487.
Van Der Auwera, S., Vandermosten, M., Wouters, J., Ghesquiéere, P., Vanderauwera, J., 2021. A three-time point longitudinal investigation of the arcuate fasciculus throughout reading acquisition in children developing dyslexia. NeuroImage 118087. https://doi.org/10.1016/j.neuroimage.2021.118087.
van Rij, J., Wieling, M., Baayen, R.H., & van Rijn, H. (2020). Itsadug: Interpreting time series and autocorrelated data using GAMMs.
Wakana, S., Caprihan, A., Panzenboeck, M.M., Fallon, J.H., Perry, M., Gollub, R.L., Hua, K., Zhang, J., Jiang, H., Dubey, P., 2007. Reproducibility of quantitative tractography methods applied to cerebral white matter. Neuroimage 36 (3), 630–644.
Wong, E.K., Krishnadas, R., Cavanagh, J., 2013. The interface between neurology and psychiatry: The case of multiple sclerosis. Adv. Psychiatric Treatment 19 (5), 370–377.
Wood, S.N., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J. R. Stat. Soc.: Series B (Statistical Methodology) 73 (1), 3–36.
Wood, S.N., 2017. Generalized additive models: An introduction with R. CRC Press.
Xue, K., Wang, D., Wang, T., Li, Y., 2019. Posterior Corpus Callosal Integrity based on Automated Fiber Quantification Predicts Age-related Decline of Cognitive Performance. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society Annual International Conference 2019, 446–449. https://doi.org/10.1109/EMBC.2019.8857590.
Yeatman, J.D., Dougherty, R.F., Myall, N.J., Wandell, B.A., Feldman, H.M., 2012. Tract Profiles of White Matter Properties: Automating Fiber-Tract Quantification. PLOS ONE 7 (11), e49790. https://doi.org/10.1371/journal.pone.0049790.
Yeatman, J.D., Richie-Halford, A., Smith, J.K., Keshavan, A., Rokem, A., 2018. A browser-based tool for visualization and analysis of diffusion MRI data. Nature Commun. 9 (1), 940. https://doi.org/10.1038/s41467-018-03297-7.
Zeineh, M.M., Kang, J., Atlas, S.W., Raman, M.M., Reiss, A.L., Norris, J.L., Valencia, I., Montoya, J.G., 2015. Right arcuate fasciculus abnormality in chronic fatigue syndrome. Radiology 274 (2), 517–526. https://doi.org/10.1148/radiol.14141079.
Zhang, J., Wei, 976 X., Xie, S., Zhou, Z., Shang, D., Ji, R., Yu, Y., He, F., Du, Y., Ye, X., Luo, B., 2018. Multifunctional Roles of the Ventral Stream in Language Models: Advanced Segmental Quantification in Post-Stroke Aphasic Patients. Front. Neurol., 9, 89. https://doi.org/10.3389/fneur.2018.00089.
Zhang, X., Sun, Y., Li, W., Liu, B., Wu, W., Zhao, H., Liu, R., Zhang, Y., Yin, Z., Yu, T., Qing, Z., Zhu, B., Xu, Y., Nedelska, Z., Hort, J., Zhang, B., Initiative, Alzheimer's Disease Neuroimaging, 2019. Characterization of white matter changes along fibers by automated fiber quantification in the early stages of Alzheimer's disease. NeuroImage. Clinical 22, 101723. https://doi.org/10.1016/j.nicl.2019.101723.
Zhou, S., Jin, L., He, J., Zeng, Q., Wu, Y., Cao, Z., Feng, Y., 2018. Distributed performance of white matter properties in chess players: A DWI study using automated fiber quantification. Brain Res. 1700, 9–18. https://doi.org/10.1016/j.brainres.2018.07.003.
Zuur, A., Ieno, E.N., Walker, N., Saveliev, A.A., Smith, G.M., 2009. Mixed effects models and extensions in ecology with R. Springer Science & Business Media.