Tract-specific statistics based on diffusion-weighted probabilistic tractography

Andrew T. Reid[1*], **Julia A. Camilleri**[2,3], **Felix Hoffstaedter**[2,3], **Simon B. Eickhoff**[2,3]

***Corresponding author:**
andrew.reid@nottingham.ac.uk

[1]School of Psychology, University of Nottingham, UK; [2]Institute for Neuroscience and Medicine (INM-7), Jülich Research Center, Jülich, Germany; [3]Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University, Düsseldorf, Germany

Abstract Diffusion-weighted neuroimaging approaches provide rich evidence with which to estimate the structural integrity of white matter in vivo. Tract-based spatial statistics (TBSS), for instance, allows us to assess the spatial distribution of statistical tests performed on diffusion properties such as fractional anisotropy, which are proposed to reflect the integrity of white matter tracts. However, such methods do not provide a direct assessment of white matter integrity for connections between two specific regions of the brain. Here, we present a novel method for deriving tract-specific diffusion statistics, based upon arbitrarily-defined regions of interest. Our approach makes use of an empirically-derived population distribution based on probabilistic tractography, using the Nathan Kline Institute (NKI) Enhanced Rockland sample. We use a heuristic method to determine the most likely geometry of a path between two regions, if one exists, and express this as a spatial distribution. We then estimate the average orientation of streamlines traversing this path, at discrete distances along its trajectory, and compute the fraction of diffusion directed along this orientation for each participant. This allows us to obtain participant-wise metrics along the specific tract (tract-specific anisotropy; TSA), which can then be used to perform statistical analysis on any comparable population. Based on this method, we report both negative and positive associations between age and TSA for two networks derived from published meta-analytic studies (the “default mode” and “what-where” networks), along with more moderate sex differences and age-by-sex interactions in both networks.

Introduction

Diffusion-weighted imaging (DWI) is a promising non-invasive in vivo technique for evaluating the integrity of myelinated axonal projections in the brain. DWI is based on the attenuation of the T2-weighted MRI signal in the presence of a field gradient, which indicates the degree to which diffusion is unrestricted in brain tissue in the direction of that gradient. Methods that reconstruct DWI maps across multiple gradient orientations can be used to model the diffusion of water molecules in discrete compartments (or voxels) of brain tissue, and the **anisotropy** of this diffusion can be used to estimate the orientation(s) along which diffusion is biased in that voxel. Because myelin is highly lipid-based, and forms a strong hydrophobic barrier, the degree of directed anisotropy in a white matter voxel is presumed to indicate the degree to which it is comprised of coherently oriented myelinated axons (***Assaf and Pasternak, 2007***). Moreover, variance in this anisotropy can be used to estimate the relative integrity of myelinated fibres that run through a voxel – with the assumption being that decreased anisotropy indicates a decrease in myelination or myelinated axons.

The simplest model of diffusion used for DWI analysis is the **diffusion tensor**, which assumes a Gaussian distribution with one principal and two secondary axes, corresponding to the first three eigenvectors of observed diffusion across gradient orientations (***Basser and Jones, 2002***; ***Chung et al., 2010***). From this model, we can obtain a summary measure of anisotropy (fractional anisotropy; FA), based on the relative magnitudes of the eigenvalues associated with each axis of the tensor. FA ranges from 0 to 1, where 0 indicates perfect isotropy (as would be expected of a uniform substance such as water or cerebrospinal fluid), and 1 indicates diffusion exclusively in the principal direction. The diffusion tensor can also be used to perform **deterministic tractography**, in which streamlines are generated by starting at a pre-specified set of "seed" voxels, and propagated through a series of neighbouring voxels by reorienting at each according to its principal orientation of diffusion. Tractography can also be set in a probabilistic framework, by generating many streamlines and sampling at each voxel from a posterior probability distribution of orientations based on the tensor model (***Behrens et al., 2003***).

While the basic diffusion tensor is an adequate model for voxels through which fibres are essentially oriented in a single direction (e.g., fibres traversing the corpus callosum), it fails to model more complex situations, such as the case where two or more fibres are crossing, diverging, or converging. This presents a strong bias in favour of finding certain pathways over others. One way of addressing this issue is to explicitly model multiple fibre directions, for example by using Bayesian model estimation to determine whether multiple fibre orientations are present, and if so how strongly they contribute to the observed diffusion signal (***Behrens et al., 2007***). Such an approach greatly improves the ability of probabilistic tractography to discover tracts which traverse areas of uncertainty. As an example from Behrens et al. (2007): while seeding in the internal capsule led only to a prominent primary motor projection in the single-fibre approach, numerous other cortical targets were reached when crossing fibres were explicitly modelled.

It is often desirable to relate voxel-wise DWI-based metrics such as FA to other phenotypical observations, such as behavioural or cognitive measures, or clinical status. Voxel-wise analyses can be highly confounded by the individual geometry of white matter tracts, and one way to address this issue is tract-based spatial statistics (TBSS), in which FA measures are projected onto a population-based FA "skeleton" with a high probability of being white matter in all participants (***Smith et al., 2006***, ***2007***). The presence of crossing fibres, however, also has implications for the interpretation of FA (***Jbabdi et al., 2010***). As an example, for two identical fibres oriented along the anteroposterior (AP) axis, FA would be inversely proportional to the number of fibres crossing each along the (perpendicular) mediolateral (ML) axis. Interpreting FA in terms of the underlying microstructure of white matter in a voxel is thus inherently ambiguous. This ambiguity can be improved if crossing fibres are explicitly modelled, for example using the Bayesian approach described above. Such a crossing fibre model has been proposed as an extension to the TBSS approach (***Jbabdi et al., 2010***).

Although TBSS provides a means of assessing the spatial distribution of statistical effects on white matter integrity, it is often difficult to apply this distribution to specific connectivity in the brain. Suppose, for instance, that we are interested in whether the white matter comprising the physical connection between brain regions $\boldsymbol{R}_a$ and $\boldsymbol{R}_b$ is altered in condition $\boldsymbol{C}$. Using TBSS, we observe that the $\boldsymbol{C}+$ group has decreased FA in an area of white matter that *could* be intermediary to $\boldsymbol{R}_a$ and $\boldsymbol{R}_b$. However, as most major white matter tracts host a mixture of numerous projection, association, and commissural fibres, we can only really speculate about the possibility of a compromised $\boldsymbol{R}_a\boldsymbol{R}_b$ tract. To improve interpretability, we require an explicit approximation of the geometry of tract $\boldsymbol{R}_a\boldsymbol{R}_b$, and an estimate of the diffusion specifically oriented along this tract.

In this study, we introduce a novel methodology to address both of these issues. We first perform probabilistic tractography on a representative sample of participants (N=130, aged 18-80), using high angular resolution DWI data from the Nathan Klein Institute Enhanced Rockland sample (***Nooner et al., 2012***), and two sets of regions-of-interest (ROIs) obtained from published meta-analytic neuroimaging studies. For each pair of ROIs $\boldsymbol{R}_a$ and $\boldsymbol{R}_b$, we then compute the probability of a tract passing through a given voxel, and use a heuristic approach to determine whether a tract likely exists between $\boldsymbol{R}_a$ and $\boldsymbol{R}_b$, and what its most probable trajectory is. Finally, we use an approach similar to Behrens et al. (2007) to determine for each participant, and each voxel, the degree of diffusion in the direction of the tract at that voxel. This yields

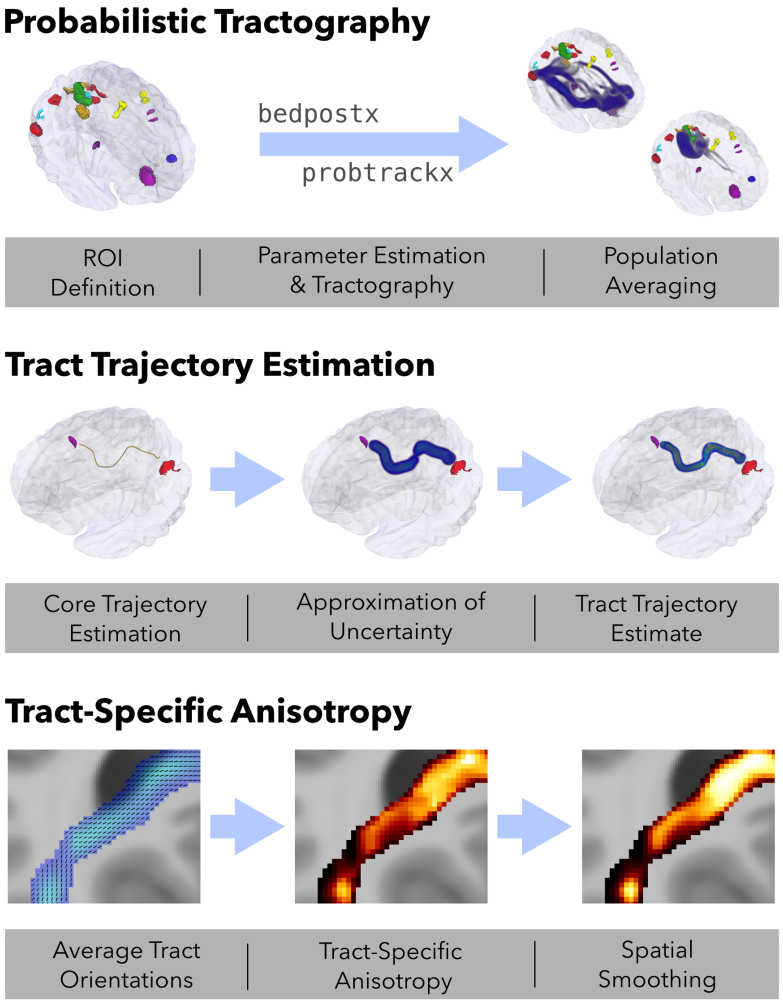

Figure 1. Schematic of the procedure followed in this study. *Probabilistic tractography*: ROIs were obtained from previous meta-analytic studies and used as seed/target regions for diffusion tensor modelling with `bedpostx` and probabilistic tract tracing performed with `probtrackx`, across all participants. The resulting probability distributions were averaged across directions for each ROI pair ($\mathbf{P}_{ab}$). *Tract trajectory estimation*: For each ROI pair, a "core" trajectory was estimated from these bidirectional averages, and represented as a 3-dimensional polyline. An uncertainty field ($\mathbf{\Phi}_{ab}$) was then generated from this polyline using an anisotropic Gaussian kernel. Finally, a "core" tract estimate ($\mathbf{P}_{ab-\text{tract}}$) was generated as the element-wise product of $\mathbf{P}_{ab}$ and $\mathbf{\Phi}_{ab}$. *Tract-specific anisotropy*: Average tract orientations were computed for each voxel in a given tract, and these orientations were then regressed against the diffusion evidence for each individual participant. This produced a tract-specific anisotropy (TSA) distribution for each participant, that can be regressed against variables of interest.

a tract-specific anistropy (TSA) metric that can be regressed against variables of interest. Here, we report the tract-wise and 3D distributions of TSA regressed against age, sex, and their interaction.

Results

Tract determination

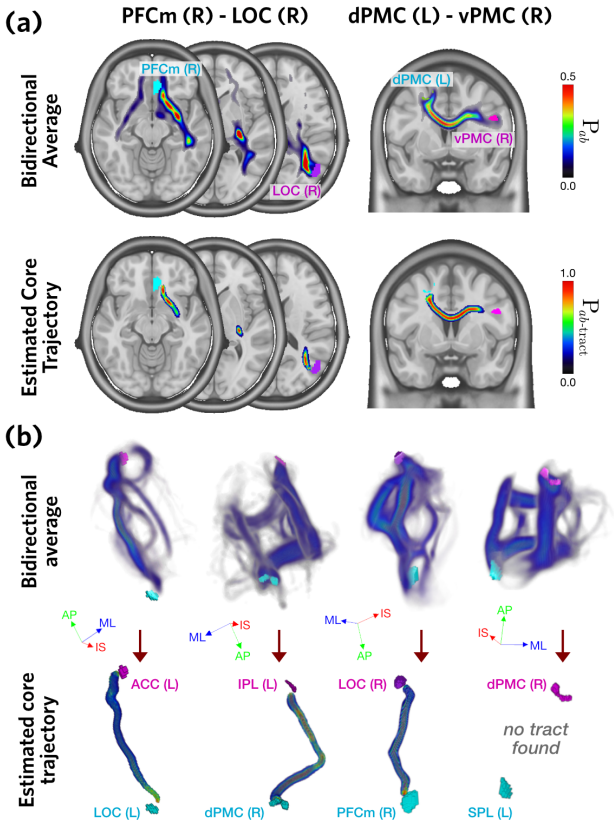

Figure 2. (a) Core tract trajectory estimation steps, shown for two exemplar tracts. The top row shows the streamline probability determined across all participants, averaged in both directions, $\mathbf{P}_{ab}$. The bottom row shows the estimated core trajectory for each ROI pair, $\mathbf{P}_{ab-\text{tract}}$. **(b)** Three-dimensional distributions of $\mathbf{P}_{ab}$ (top) and $\mathbf{P}_{ab-\text{tract}}$ (bottom), for four exemplar tracts. Axis labels: ML, medial-lateral; AP, anterior-posterior; IS, inferior-superior.

We used a heuristic approach (see ***Figure 1***) to determine the most probable trajectory of a white matter projection between two regions of interest (ROIs). For a given pair of ROIs $\boldsymbol{R}_a$ and $\boldsymbol{R}_b$, we performed probabilistic tractography twice, seeding in one of these regions and terminating in the other. The resulting directed probability distributions were averaged to obtain a bidirectional probability average, which was then thresholded in order to identify a set of contiguous voxels connecting the two regions. If this step did not result in such a pathway, a connection between $\boldsymbol{R}_a$ and $\boldsymbol{R}_b$ was rejected; otherwise, we identified the "core" of the pathway and used this to obtain a final bidirectional tract trajectory estimate. Tract determination was carried out on two sets of ROIs, derived from previously published meta-analytic studies: the default mode network (DMN), and the what-where network (WWN).

Figure 2a illustrates the tract trajectory estimation process for example tracts PFCm(R) - LOC(R) and dPMC(L)-vPMC(R). The horizontal and coronal slice renderings show the initial minimal bidirectional averages (top row) and final tract trajectory estimates (bottom row). Figure 2b shows three-dimensional renderings of these steps, including the failure to estimate a core trajectory for SPL(L)-dPMC(R). For the vast majority of tracts, this method was able to isolate a single, core trajectory from a variety of alternatives.

For the DMN, 33 of 36 ROI pairs (92%) produced core tract estimates (***Figure 3***). Two of the failed tracts involved the left PFCm, with the more posterior LOC(R) and PCm(L) regions. Tracts connecting PCm and PCG to PFC and ACC traversed the more superior cingulum bundle, while those connecting LOC to PFC and ACC traversed the more inferior fronto-occipital fasciculus. Contralateral DMN connections traversed the splenium of the corpus callosum (CC; Figure 3d) for posterior ROIs, and the genu of the CC for anterior ROIs and anteriorly projecting LOC connections.

For the WWN, 43 of 45 tracts were estimated (96%), with only 2 tracts failing because no core tract could

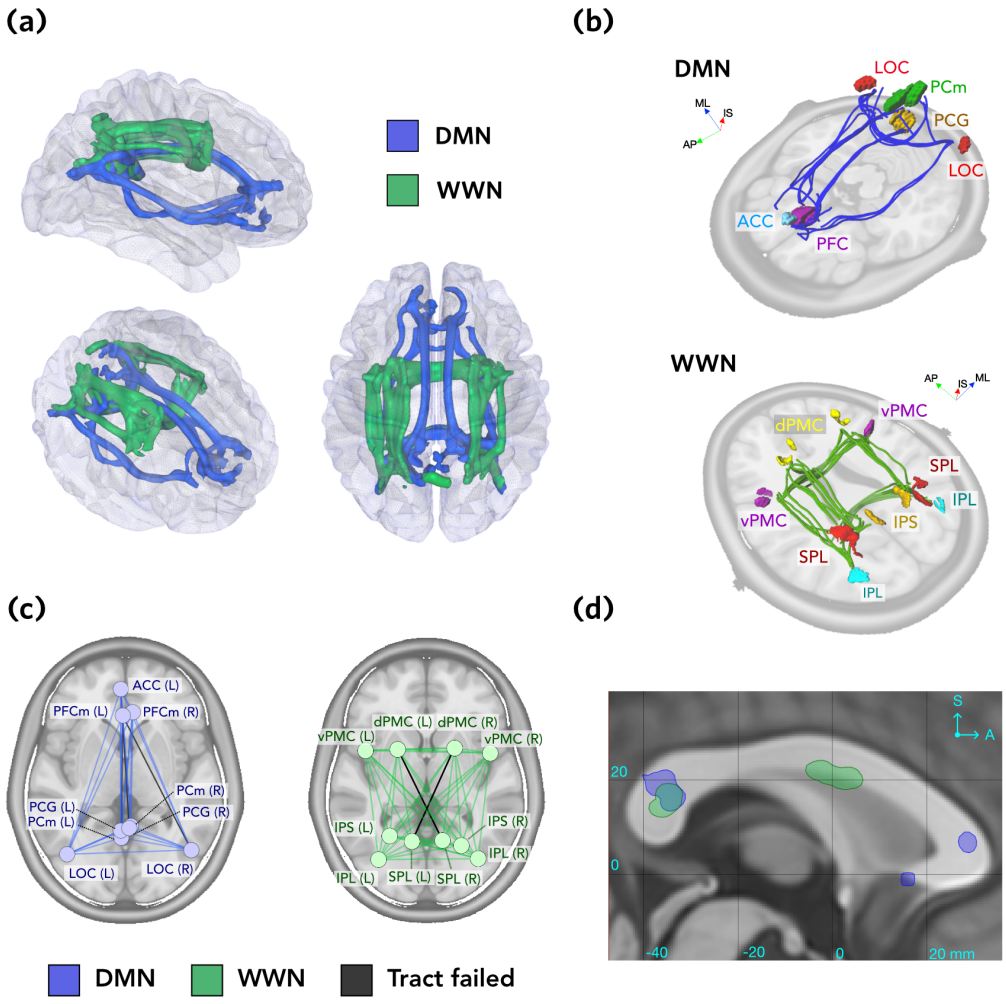

Figure 3. Tract determination. **(a)** All accepted tracts shown for the DMN (blue) and WWN (green) networks, rendered as isosurfaces thresholded at 0.5 (top, right, and oblique perspectives). **(b)** Core polylines representing the trajectories of accepted tracts in the DMN (top) and WWN (bottom) networks. Images are rendered with oblique perspectives next to planar sections of the ICBM non-linear T1 template image, for reference. See Materials and Methods for names of the ROIs. **(c)** Graph representations of the DMN (blue) and WWN (green) networks, with failed edges shown in grey. Graph vertices represent the center points of each ROI, projected onto the transverse plane. **(d)** Sagittal section at midline, showing where contralateral tracts (thresholded at 0.5) for each network traverse the corpus callosum. Coordinates are ICBM152. Axis labels: ML, medial-lateral; AP, anterior-posterior; IS, inferior-superior.

be identified (***Figure 3***). The failed tracts were the two contralateral connections between SPL and dPMC (Figure 3c). Successfully estimated tracts consisted of anteroposterior projections traversing the superior longitudinal fasciculus, and contralateral connections traversing the splenium of the CC and the inferior part of the body of the CC (Figure 3d).

Tract-specific anisotropy

Given a tract trajectory estimate, as described above, we next wanted to determine how strongly diffusion profiles of individual participants were oriented along that tract. For a given voxel, we determined the average streamline orientation derived from the previous probabilistic tractography step, and then computed, for each participant, a "tract-specific anisotropy" (TSA) estimate, indicating the degree to which this average orientation loaded onto the DWI intensities along each gradient direction (TSA values are the β parameters estimated from ***Equation 3***).

Kernel density estimates for TSA, for both networks, are shown in ***Figure 4***a. These show the distribution of scores for tracts thresholded at $\mathbf{P}_{ab-\text{tract}} > 0.5$. While most distributions were roughly Gaussian, with a heavy positive tail, this varied across tracts in terms of both kurtosis and skewness, and some shorter tracts (e.g., PCG(L)-PCm(R)) had very little variance. For a number of tracts (e.g., SPL(R)-IPS(L)), a bimodal distribution was evident. The overall distribution for each network is shown in the insets; WWN showed a broader, flattened distribution relative to DMN.

Figure 4b shows the two-dimensional spatial distributions of TSA scores for two exemplar participants, along with the average TSA across all 130 participants. These distributions demonstrate variability between participants, but in general the highest TSA scores were observed in the center of estimated tract trajectories - reflecting the strongest loading of diffusion profiles - which tapered off towards the edges.

Tract-specific Age and Sex effects

Having obtained TSA scores for each individual participant, we were next interested in whether these scores were associated with age, sex, and their interaction. We performed voxel-wise regressions of the form $TSA = \beta_0 + \beta_1 \cdot \text{Age} + \beta_2 \cdot \text{Sex} + \beta_3 \cdot \text{Age} \times \text{Sex} + \varepsilon$. T-statistics for each coefficient were obtained, and summarised at each distance along the tract. For each tract, we then used one-dimensional random field theory to identify significant clusters of t-statistics (p<0.05). To control for family-wise error, we limited the false discovery rate (FDR) over all tracts to 0.05.

For the DMN, there were significant associations of TSA with age and sex, as well as the interaction between these two factors (***Figure 5***, top row). Age effects were fairly diffuse. Negative effects of age were found in 15/32 (47%) of DMN tracts, including ipsi- and contralateral connections between PFC and LOC (traversing the fronto-occipital fasciculus), and PCm(L)-PFCm(R). Modest positive effects of age were also found for 4/32 (12%) of tracts. These were smaller, and comprised exclusively of ipsilateral tracts (three left and one right). Both positive (female > male) and negative sex differences were also found. ***Figure 6*** shows scatter and violin plots for selected tracts. Notably, tract PCm(L)-PCG(L) showed effects for age, sex, and their interaction; and tract ACC(L)-LOC(L) and PFCm(R)-LOC(R) showed both positive and negative age effects.

The WWN also showed significant associations for age, sex, and their interaction (***Figure 5***, bottom row). Surprisingly, there were strong posirtive and negative age effects for this network, with 16/32 (50%) having showiung negsative effects and 14/32 tracts (44%) showing positive effects. Compared to DMN, the age effects for WWN were more focal. Negative age effects were found proximal to dPMC(L), and in collosal fibres traversing the body of the corpus callosum (see ***Figure 3***d), involving contralateral tracts between dPMC and vPMC. There were also negative associations found in contralateral tracts between IPS, dPMC, and vPMC. The largest positive age effects occurred along anteroposterior-oriented tracts in the longitudinal fasciculus, particularly those involving left and right dPMC. All sex differences were negative (male > female) for this network, on were found on contralateral tracts. ***Figure 6*** shows scatter and violin plots for selected tracts. Notably, the homotopic dPMC(L)-dPMC(R) tract showed both a sex difference and an Age×Sex interaction.

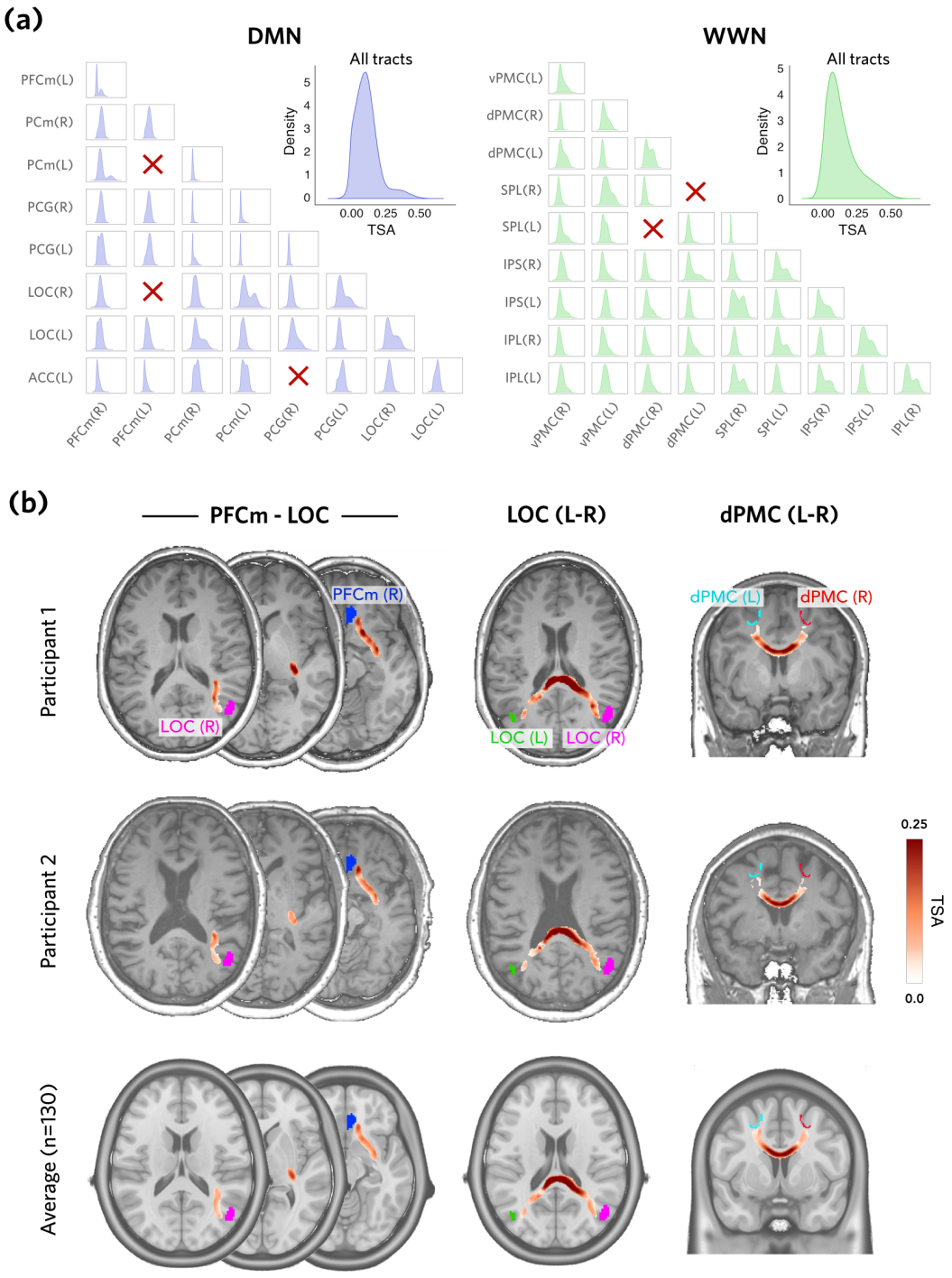

Figure 4. Tract-specific anisotropy (TSA). **(a)** Distributions of TSA across voxels and participants, for each estimated tract (ROI pair), shown as kernel density estimates for DMN (blue) and WWN (green). Insets show the distribution of TSA values over all tracts; x-axis scales are the same for all plots. Red crosses indicate that no tract was generated. **(b)** Spatial distributions of TSA values for two exemplar participants and averaged over all participants, and three examplar tracts, shown in horizontal and coronal section. Anatomical images are the ICBM152 nonlinear template.

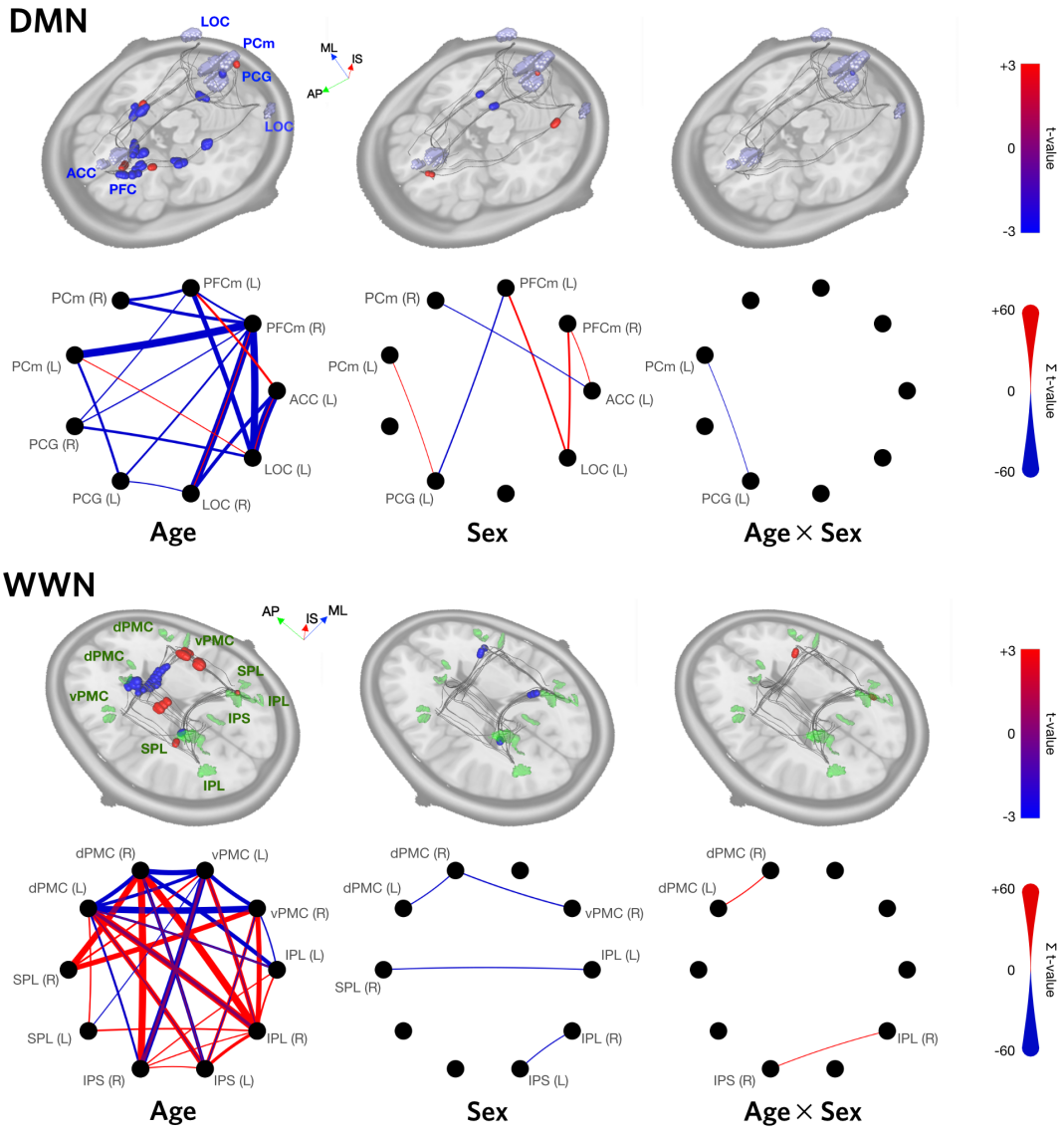

Figure 5. Regression results for both networks, for linear models of the form $TSA = Age + Sex + Age \times Sex + \epsilon$. 3D renderings show the maximal t-values (thresholded using cluster-wise inference, with FDR<0.05) at each distance along the core trajectories of each tract in the network. Circular graph representations show the sum of significant positive (red) and negative (blue) t-values for each tract. The thickness of an edge is proportional to its sum, and the absence of an edge indicates that no significant clusters were found for that tract. Axis labels: ML, medial-lateral; AP, anterior-posterior; IS, inferior-superior

Figure 5–Figure supplement 1. GLM t-value distance traces for all ROI pairs in the DMN.

Figure 5–Figure supplement 2. GLM t-value distance traces for all ROI pairs in the WWN.

Scatter and violin plots for all significant effects, and line plots and tables showing raw and cluster-thresholded t-values for all tracts, are provided in Supplementary Material.

Discussion

Estimation of core tract trajectories

In this study we introduce a novel method for obtaining tract-specific spatial estimates of the "core" trajectories of projections between two arbitrarily defined ROIs, on the basis of probabilistic DWI tractography performed across a representative cohort of healthy individuals. This approach yielded plausible trajectory estimates for 33 of 36 ROI pairs (92%) in the DMN network, and 43 of 45 ROI pairs (96%) in the WWN network. In order to achieve sufficient coverage of the space of possible trajectories, we derived our probability estimates from a large number of streamlines (50,000 × ROI) over 130 participants. Our results demonstrate the feasibility of using DWI-based probabilistic tractography for delineating individual tracts between arbitrary pairs of distal brain regions, as well as obtaining participant- and tract-specific anisotropy values from these tracts.

ROI pairs failed to generate tract trajectory estimates when the thresholding applied to the bidirectional population average distribution broke the contiguous path between them. Notably, this is not evidence that the tract does not exist, but rather that there is ambiguity about the location of its trajectory, based on the diffusion evidence. It is important to note that the threshold applied here was determined somewhat arbitrarily, i.e., by observing what value reduced the number of alternative pathways to a single, core one. In cases where this failed, there were typically two or more alternatives that could not be disambiguated (see ***Figure 2***b for an illustration of failed tract SPL(L)-dPMC(R)). Further investigation into these ROI pairs, e.g., by comparing them to known connectivity evidence from tract tracing studies (in non-human primates) or different modalities (in humans), may be useful for determining whether a tract indeed exists between them.

Previous studies have also evaluated white matter tract geometry using DWI. In a now-classic study, ***Catani et al.*** (***2002***) used deterministic tractography to identify and "dissect" individual tracts based on seed ROIs. This was extended by ***Jones et al.*** (***2005***) in order to map diffusion metrics onto the trajectories of specific tracts. More recently, ***Colby et al.*** (***2012***) used B-spline resampling to map FA to specific tracts in individual participants, allowing effects to be mapped to points along its trajectory. The present approach extends the work of these previous studies, in that it facilitates the investigation of white matter pathways connecting specific ROI pairs, rather than coarse-scale fasciculi, in terms of tract geometry and anisotropy estimation. Another important advantage is its use of a probabilistic tractography framework to describe the "core" trajectories of specific tracts across a population of interest.

Tract-specific anisotropy

It is common to use streamline counts from probabilistic tractography as an estimate of connection strength between two ROIs (e.g., ***Daianu et al., 2013***; ***Lohse et al., 2014***; ***Reid et al., 2016b***). This approach, however, suffers from a number of seemingly intractable biases, including: the nature of the diffusion profile through which a given tract traverses (anisotropy bias); the length of the tract (distance bias); and the position of seed and target ROIs relative to gryi or sulci (***Reveley et al., 2015***; ***Reid, 2016***). It is thus uncertain how to interpret a relative streamline count, or change in this quantity in association with some covariate of interest, with respect to biophysical properties such as white matter integrity or connectivity strength. The present approach is arguably less susceptible to these biases, because it only models the geometry of streamlines that *do* connect two ROIs. However, future research efforts should be directed at validating (1) the interpretability of TSA with respect to axonal white matter integrity, and (2) the spatial specificity of trajectory estimates. This could be done, for example, through the use of phantoms (***Zhu et al., 2011***; ***Caspers and Axer, 2017***) or histological approaches (***Mollink et al., 2017***; ***Schaeffer et al., 2018***).

The present method might be considered an extension of the well-known TBSS approach (***Smith et al., 2007***), in which statistics are projected onto a pre-established population-based white matter "skeleton",

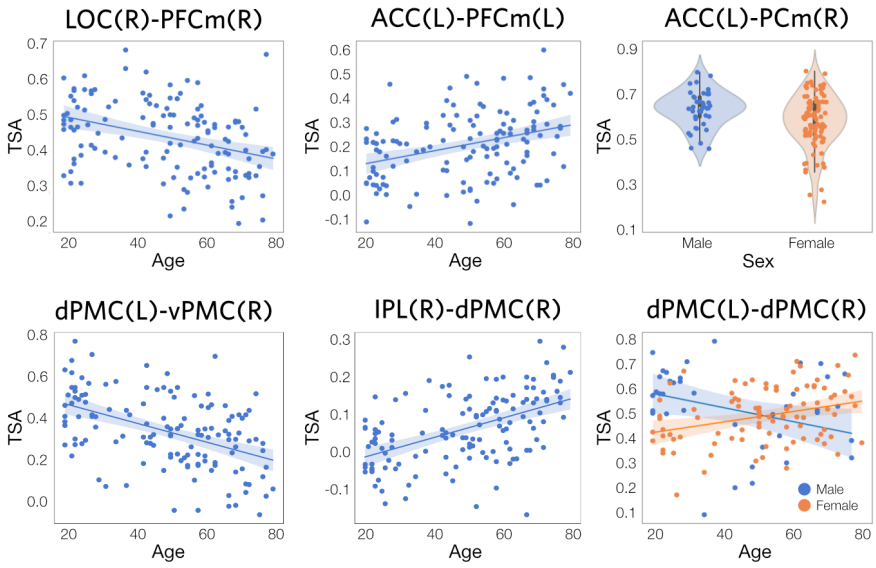

Figure 6. Plots of selected TSA regression results. *Top row*: DMN tracts with (from left to right) negative and positive age effects, and a sex difference (male > female). *Bottom row*: WWN tracts with negative and positive age effects, and an Age×Sex interaction. Data points are the mean TSA values within significant clusters (as shown in ***Figure 5***). Shaded areas on scatterplots represent 95% confidence intervals.

Figure 6–Figure supplement 1. Scatter and violin plots for all significant effects in the DMN.

Figure 6–Figure supplement 2. Scatter and violin plots for all significant effects in the WWN (Part A).

Figure 6–Figure supplement 3. Scatter plots for all significant effects in the WWN (Part B).

which serves a similar function to population-based anatomical grey matter templates, such as the linear and nonlinear ICBM-152 templates (***Fonov et al., 2009***; ***Mazziotta et al., 1995***). There are two major advantages of the present method over TBSS: (1) it allows a population-based tract estimate to be *derived specifically for the white matter tract connecting any two arbitrarily-defined ROIs*, if that tract is likely to exist; and (2) it *allows participant- and tract-specific anisotropy to be estimated*, based on the orientation of streamlines defining that tract in each voxel along its trajectory. Conversely, one major limitation with the TSA approach is that it is substantially more computationally intensive, and scales quadratically with the number of ROIs in a network of interest.

Age and sex are associated with tract-specific anisotropy

Positive and negative age/TSA associations were found for both networks. There is an abundance of evidence for age-related decreases in DMN connectivity, including a decrease in fMRI-based resting-state functional covariance (***Hafkemeijer et al., 2012***; ***Marstaller et al., 2015***; ***Sambataro et al., 2010***), and age-related decreases in FA in white matter tracts proximal to DMN regions (***Marstaller et al., 2015***). Marstaller and colleagues found a reduction in the extent to which posterior cingulate and precuneus activity covaried with wider brain networks, along with reduced FA in numerous tracts, including the fronto-occipital fasciculus, where the current negative age association is most prominent. Sambataro et al. reported decreased covariance between PCC and PFC, which predicted working memory performance. Decreases in DMN functional covariance also appear to be accelerated in Alzheimer's disease (***Jones et al., 2011***; also reviewed in ***Reid and Evans, 2013***).

For the WWN, age-related decreases in TSA occurred mainly in the body, but not the splenium, of the corpus callosum. This pattern is in agreement with several DWI-based studies of age-related connectivity changes. ***Burzynska et al.*** (***2010***) used TBSS to show an age-related reduction in FA (and increase in radial and mean diffusivity) in the genu and body of the corpus callosum, but not the splenium. Using a DTI

approach, ***Bennett and Madden*** (***2014***) found decreased FA in older versus younger participants for both the genu and splenium, but with a more pronounced effect in the former. The same authors report an anterior-to-posterior gradient in age-related FA changes, with these being more pronounced in frontal white matter, consistent with the pattern found in the current study (***Bennett et al., 2009***).

Positive age associations were a more surprising finding, as numerous articles report negative age/FA associations (e.g., ***Kodiweera et al., 2016***) and postmortem evidence of white matter loss and decrease in the proportion of small myelinated fibres with age (***Tang et al., 1997***; ***Aboitiz et al., 1997***). Both positive and negative age/FA associations have been reported in at least one previous brain-wide TBSS study (***Kochunov et al., 2007***), however. For DMN in particular, we found that most positive associations occurred in regions proximal to ROIs, where the potential confound of crossing fibres is likely more pronounced (***Jeurissen et al., 2012***). The positive relationships in the WWN were especially prominent, and suggest a (paradoxical) increase in white matter integrity in these tracts. The majority of positive relationships in WWN were found in the middle of the superior longitudinal fasciculus (SLF). One TBSS study focusing on the SLF in a healthy cohort found no effect of age on FA (***Madhavan et al., 2014***), while another whole-brain study found SLF among tracts with negative age effects (***Marstaller et al., 2015***). Increased FA in Alzheimer's disease patients has also been reported in SLF (***Douaud et al., 2011***). The authors of this study suggest that a relative sparing of crossing motor fibres may account for this effect, but this is consistent with our observed increase in TSA; on the contrary, our findings might be explained by a relative *decrease* in WM integrity in these crossing fibres.

It is possible that the increased specificity of the current approach permits a more fine-grained spatial and angular dissection of effects than does TBSS with FA, which uses a more coarse-grain white matter skeleton, and is not orientation-specific. If so, then the positive effects observed here may reflect a real age-related increase in white matter integrity for specific tracts. This possibility is supported by reports of fairly widespread increases in fMRI-based functional covariance with aging (***Tomasi and Volkow, 2011***), which have been proposed to reflect compensatory changes in response to degeneration or dysfunction of other brain regions. Given the conflicting evidence, however, these effects should be interpreted with caution. It will be important in future TSA studies to increase the number of ROIs, or query specific crossing tracts, in order to obtain a more complete picture of age-related effects across white matter.

More modest sex differences, and age-by-sex interactions, were also observed for both networks. Previous DWI-based studies have also found sex differences in FA, typically with males showing higher FA values (***Takao et al., 2013***; ***Rathee et al., 2016***; ***Menzler et al., 2011***). For DMN, we found negative effects (males > females) for the PCG(L) with PFCm(L), which accords with findings from ***Menzler et al.*** (***2011***), who report prominent differences in the cingulum bundle. We also found positive effects (females > males) that were mostly left lateralized and included LOC(L), which have not commonly been reported in the literature. Although sex differences in LOC function have been hypothesised on the basis of differential object perception (***Vanston and Strother, 2016***) – a function that has been associated with this region – there does not currently appear to be much direct evidence for this.

It is also notable that, for numerous tracts, we found both positive and negative age effects in different parts of the same tract (see ***Figure 5***). It is not immediately clear how to interpret such a result. If our inference is that a TSA value represents the number of intact axons projecting between two grey matter ROIs, this is contradicted by the observation of both increases and decreases in this region - damage to an axon anywhere along its length will trigger Wallerian degeneration in both directions. On the other hand, this finding is consistent with the idea that TSA reflects the degree of (de)myelination, which may be increased on average in one part of the tract and decreased in another. Importantly, there is evidence that demyelinated axons, while they may be functionally impaired, are not necessarily at higher risk of degeneration (***Smith et al., 2013***). A third possibility is that TSA is influenced by the degree of crossing fibres in a particular region along its length; changes to diffusion in directions other than the average orientation of interest will influence the regression fit used to estimate this metric.

Limitations and future directions

The NKI Rockland dataset was chosen due to its large size, age range, and the use of a single MRI scanner and protocol. To ensure the cohort was as representative of the general population as possible, and to enable the analysis of age over the lifespan, we chose to use close to the full age range (18-80), and to exclude participants with clinical diagnoses. As with most population templates, however, the choice of cohort is an important consideration when interpreting a derived result. The human brain is known to show systematic anatomical grey matter changes across the lifespan (***Sowell et al., 2003***; ***Giorgio et al., 2010***), and this will almost certainly bias normalization in a way that may account for a portion of the TSA effects reported here. It will be important in future studies to assess the influence of this bias, use cohorts that are more targeted to a particular phenomenon under investigation, and compare the predictions of TSA to in vivo or post mortem analyses of white matter (e.g., as in ***Reveley et al., 2015***).

The ability to estimate spatial trajectories of white matter projections between specific pairs of ROIs could lead to a number of important applications. For instance, this information could be used to predict the functional outcomes of age-related white matter lesions (WML), which have a prevalence of ~18% in people aged 60-69 and ~40% in people over 80 (***Vernooij et al., 2008***). Conceivably, clinicians would identify WML locations using T2-FLAIR imaging, estimate which tracts intersect these lesions, and predict functional deficits on the basis of the brain regions connected by these tracts. This approach could augment previous studies investigating connectivity (***Tuladhar et al., 2014***), cortical thickness (***Reid et al., 2010***), and gait changes related to WML (***de Laat et al., 2012***).

Materials and Methods

Participants

Participant data were obtained from the publicly available Nathan Klein Institute (NKI) Enhanced Rockland sample (***Nooner et al., 2012***), through the 1000 Functional Connectomes Project (www.nitrc.org/projects/fcon_1000/). We included participants from the first four data releases, while excluding any participant with an existing clinical diagnosis at the time of scanning. In total, 130 participants (86 female, age range 18 to 80) were analyzed.

Neuroimaging data and metadata

All imaging data in the Rockland sample were acquired from the same scanner (Siemens Magnetom TrioTim, 3.0T). T1-weighted images were obtained using a MPRAGE sequence (TR = 1900 ms; TE = 2.52 ms; voxel size = 1 mm isotropic). DWI was collected with a high spatial and angular resolution (TR = 2400 ms; TE = 85 ms; voxel size = 2 mm isotropic; b = 1500 s/mm2; 137 gradient directions). Age and sex data were also obtained, with the Sex factor encoded as 1=Male, 2=Female.

Preprocessing

DWI images for all participants were preprocessed using FMRIB software (***Smith et al., 2004***; ***Woolrich et al., 2009***); specifically, FSL version 6.0.0. Participant datasets were processed in parallel using a Linux-based SLURM computing cluster, located at the University of Nottingham. Raw diffusion data were first corrected for eddy current artifacts using the `eddy` command. All $\boldsymbol{B}_0$ (zero-gradient) images were then averaged and used to extract a brain mask using the `bet` command. Next, for each voxel in the brain mask, a diffusion tensor model was fit to the data using `dtifit`, and subsequently passed to the `bedpostx` command, which infers the existence of crossing fibres and estimates the contribution of each crossing fibre to the diffusion-weighted signal (***Behrens et al., 2007***). Finally, to express the voxel-wise diffusion models in standard space, linear and nonlinear transforms between native diffusion and MNI-152 space were estimated, by applying the `flirt` (12 degrees of freedom) and `fnirt` commands, respectively, using individual FA images as input and a mean template FA image as the reference. Inverse transforms were also estimated, using the `invwarp` command, for use in later steps.

Regions of interest

Seed regions of interest (ROIs) were obtained from two previously-defined networks (see ***Figure 3***b). For each network, all possible ROI pairs were considered as potential tracts.

The *default-mode network* (DMN) was derived from ***Schilbach et al.*** (***2012***), who performed an Activation Likelihood Estimation (ALE; ***Eickhoff et al., 2009***, ***2012***) meta-analysis using 533 experiments from the BrainMap database (***Fox and Lancaster, 2002***), querying for regions that were consistently deactivated across tasks. The ROIs were obtained by downloading the DMN network from the ANIMA database (https://anima.inm7.de/studies/Schilbach_SocialNetworks_2012; ***Reid et al., 2016a***). The network was comprised of 9 ROIs: anterior cingulate cortex (ACC; left only), lateral occipital cortex (LOC; bilateral), posterior cingulate gyrus (PCG; bilateral), precuneus (PCm; bilateral), and medial prefrontal cortex (PFCm; bilateral).

The *what/where* network (WWN) was derived from ***Rottschy et al.*** (***2012***), with data obtained from the ANIMA database (https://anima.inm7.de/studies/Rottschy_WorkingMemory_2012). The authors performed a conjunction on ALE analyses for tasks testing memory for object identity ("what"; 42 experiments) and object location ("where"; 13 experiments). The WWN was comprised of 10 ROIs, with 5 brain regions represented bilaterally. These ROIs were: dorsal and ventral premotor cortex (dPMC, vPMC), superior parietal lobule (SPL), inferior parietal lobule (IPL), and infraparietal sulcus (IPS).

Probabilistic tractography

Probabilistic tractography was performed in MNI-152 space, using the `probtrackx` command (***Behrens et al., 2007***). For both networks, we generated 50,000 streamlines per seed voxel, with the other seed regions as target masks. We applied a step length of 0.5 mm, a curvature threshold of 0.2, a minimal path distance of 5 mm, and a fibre threshold of 0.01. A distance correction was also applied. For each voxel `probtrackx` recorded the number of streamlines which encountered that voxel, producing a separate count for streamlines terminating in each target ROI. Additionally, the voxel-wise mean orientation of streamlines between each seed/target pair was computed.

Tract determination

Tracts were determined separately for each ROI pair. Voxel-wise streamline counts for each direction ($\boldsymbol{R}_a$ to $\boldsymbol{R}_b$ or $\boldsymbol{R}_b$ to $\boldsymbol{R}_a$) were first normalized by dividing by the total number of streamlines reaching target $\boldsymbol{R}_b$ from seed $\boldsymbol{R}_a$, and the minimum across both directions was obtained. These minimum images were subsequently averaged over participants, smoothed, and normalized to the range [0,1], yielding a probability $\boldsymbol{P}_{ab}(i)$ of voxel i being included in tract $\boldsymbol{T}_{ab}$. It is noteworthy that in some cases, the matrix $\mathbf{P}_{ab}$ of such values is constrained to a single, geometrically confined tract, while in others, it reflects many possible routes between two ROIs (see examples in ***Figure 2***b). Our objective was to identify the single most probable tract, or reject the tract altogether if such could not be determined. This was done in a heuristic manner. Firstly, we thresholded $\mathbf{P}_{ab}$ by setting all values where $\mathbf{P}_{ab} < \alpha$ to zero. A threshold of $\alpha = 0.07$ was applied by experimentation, in order to both disconnect most low-probability alternative pathways, and ensure that (for cases where the existence of a tract is known or likely) at least one pathway remained intact. Additionally, because $\mathbf{P}_{ab}$ proximal to ROIs tended to be lower that for the main tract trajectory, ROIs were dilated by 3 voxels to ensure they remained connected after thresholding.

We next discretized $\mathbf{P}_{ab}$ into distance assignments $\mathbf{D}$, using a flood-fill approach, assigning all voxel neighbours of seed region $\boldsymbol{R}_a$ a value of $d = 1$, all subsequent neighbours $d = 2$, and so on, until all voxels in $\boldsymbol{T}_{ab}$ were assigned a distance. We then identified the voxel of maximal $\boldsymbol{P}_{ab}$ for each distance d, constructing a polyline $\mathbf{L}_{ab}$ with center points of these voxels as its vertices. Inclusion of a vertex in $\mathbf{L}_{ab}$ was conditional on two constraints: (1) segment length $|\mathbf{x}_{i-1,i}| < 4$ mm; and (2) vertex angle $\theta_i < \pi/3$. Where a constraint was violated, the voxel with the next highest $\boldsymbol{P}_{ab}$ was tested, and so on until a voxel was found satisfying the constraints. In cases where no appropriate voxels were found, the maximal voxel was added and the violation was recorded as a flag for visual inspection. $\mathbf{L}_{ab}$ was extended in this fashion until: (1) the target region $\boldsymbol{R}_b$ was encountered, and the tract was accepted for further analysis; or (2) the maximal distance was encountered, but not the target region $\boldsymbol{R}_b$, and the tract was rejected for further analysis (i.e., the thresholding broke all routes, and thus a "true" route between $\boldsymbol{R}_a$ and $\boldsymbol{R}_b$ could not be determined).

With the assumption that an accepted polyline represents the geometric center of a given tract T_{ab}, we then modelled the probability $\mathbf{P}_{ab-\text{tract}}$ of a voxel being in that tract, as the product of an uncertainty field $\mathbf{\Phi}_{ab}$ oriented around $\mathbf{L}_{ab}$, and the original probability field $\mathbf{P}_{ab}$ (see ***Figure 1***).

$\mathbf{\Phi}_{ab}$ was constructed as follows. For every vertex $v \in \mathbf{L}_{ab}$, an orientation vector ω_v was computed as the sum of the segment vectors $\mathbf{x}_{v-1,v}$ and $\mathbf{x}_{v,v+1}$ (at the endpoints of $\mathbf{L}_{ab}$, only one segment was used). From ω_v, an anisotropic Gaussian kernel $\phi(i, \omega_v, \mu, \sigma_a, \sigma_r)$ was applied to the subset of voxels within a radius of 8 mm of v. Parameters defining the shape of the Gaussian function were fixed at $\mu = 0$ mm, axial $\sigma_a = 10$ mm, and radial $\sigma_r = 4$ mm; where axial and radial axes were parallel and perpendicular to ω_v, respectively. For a given voxel i in T_{ab}, the maximal value of ϕ across all vertices v was assigned:

$$\mathbf{\Phi}_{ab}(i) = \max_v \Big(\phi(i, \omega_v, \mu, \sigma_a, \sigma_r) \Big) \tag{1}$$

$\mathbf{P}_{ab-\text{tract}}$ was determined as:

$$\mathbf{P}_{ab-\text{tract}} = f(\mathbf{\Phi}_{ab} \odot \mathbf{P}_{ab}, d) \tag{2}$$

The function $f(g, d)$ normalizes g, at each discrete distance d, to values between 0 and 1.

Tract-specific anisotropy estimation

Having defined (or rejected) a tract T_{ab}, we were next interested in extracting meaningful diffusion metrics from it, which can allow us to perform statistical inference on specific tracts. At each voxel, we thus wanted to estimate how strongly its diffusion weighed onto the orientation of the tract at that voxel. To do this, we obtained the average orientation (across all participants) of streamlines going through voxels in T_{ab} using `probtrackx` (see **Probabilistic tractography**). These average orientation images, along with the $\mathbf{P}_{ab-\text{tract}}$ images, were warped from standard to individual participants' diffusion space, using the inverse transforms computed in the preprocessing step (see **Preprocessing**).

Next, for each participant, and for each voxel j, the fraction of the diffusion-weighted signal along the average tract orientation was estimated by fitting the following linear regression using the `statsmodels` Python library (https://www.statsmodels.org):

$$\mathbf{s}_j / s_j^0 = \beta_j \cdot e^{-b\delta(\mathbf{R}^\intercal \bar{\mathbf{v}}_j)^2} + c_j \tag{3}$$

where $\bar{\mathbf{v}}_j$ is the average streamline orientation, $\mathbf{R}$ is the $M \times 3$ matrix of gradient orientation vectors (where M is the number of orientations, here $M = 137$), s_j^0 is the non-diffusion-weighted signal, $\mathbf{s}_j$ is the observed signal at each gradient orientation, b is the gradient strength (b-value), and δ is the diffusivity.

Notably, this formulation is equivalent to that presented in ***Behrens et al.*** (***2003***, 2007), but applying only to the average orientation $\bar{\mathbf{v}}_j$. β_j is the regression coefficient (with c_j being an intercept term), and is analogous to the f value from the crossing fibres model; i.e., the fraction of signal contributed by $\bar{\mathbf{v}}_j$. We refer to these coefficients as *tract-specific anisotropy* (TSA).

Statistical analyses

The TSA values obtained in the preceding step were warped back into standard space and smoothed with a Gaussian kernel with a full-width at half-maximum (FWHM) of 1.5 mm. Subsequently, for each voxel in a tract, the following linear regression model was tested, for the first-order effects of Age and Sex, and their interaction:

$$\text{TSA} = \beta_0 + \beta_1 \cdot \text{Age} + \beta_2 \cdot \text{Sex} + \beta_3 \cdot \text{Age} \times \text{Sex} + \epsilon \tag{4}$$

For each contrast, we next wanted to summarize the resulting t-statistics at each discrete distance along the trajectory of the tract. To do this, following an approach similar to TBSS (***Smith et al., 2006***), we computed weighted t-statistics for each voxel j as $t'(j) = t(j) \cdot \mathbf{P}_{ab-\text{tract}}(j)^\lambda$, where λ determines the rate of decay (here, we set $\lambda = 1.0$). At each discrete distance d along the tract, we assigned (unweighted) t_d as the t-statistic corresponding to the maximal t'_d.

Because anatomical properties along a tract can be assumed to form a continuous random field (i.e., neighbouring vertices have a spatial dependence), we analyzed the resulting distance-wise summary statistics as a one-dimensional random field, using the `rft1d` Python library (http://www.spm1d.org/rft1d/; ***Pataky, 2016***). Firstly, the spatial smoothness of model residuals was estimated for each tract as a FWHM value, and averaged across tracts. For each tract, this mean FWHM was used to compute t^*, the critical t-value at $\alpha = 0.05$ for a Gaussian field, using an inverse survival function. Secondly, cluster-wise inference was performed to identify significant clusters along the tract, with a minimum cluster size of 3. Thirdly, p-values for all clusters, across all tracts, were corrected for family-wise error using a false discovery rate (FDR) threshold of 0.05. FDR was performed using the `statsmodels` Python library, with a two-stage non-negative FDR method (`fdr_tsbky`). Finally, to estimate the spatial extent of effects for a given tract, significant t-values were summed over that tract, for positive and negative t-values separately.

Code and data sharing

All code for this procedure was written in Python 3 (Anaconda build), making system calls of native FSL packages, as indicated. The following third-party Python libraries were also used: `nibabel`, `nilearn`, and `statsmodels`. Source code for all procedures used in this study is freely available at https://github.com/neurocoglab/dwi-tracts. This includes: scripts to run FSL preprocessing and DWI analyses, using standard SGE- or SLURM-style parallel environments; a module to estimate core trajectories and TSA values; a module to fit general linear models to TSA values; and a module to plot the results of these. The output of these modules (polylines, surface meshes, graphs) can be visualised using ModelGUI software, an open source Java library available at https://github.com/neurocoglab/mgui-core. The NKI-Rockland data is freely available at www.nitrc.org/projects/fcon_1000/. Participant- and sample-wise data will be made freely available on the University of Nottingham Research Data Management Repository at https://doi.org/10.17639/nott.7102.

Acknowledgements

We thank Saad Jbabdi for extremely helpful feedback, in particular regarding the TSA approach, as well as Marije ter Wal and Christopher Madan for their invaluable advice on various parts of the manuscript. 2D and 3D renderings for this publication were generated using ModelGUI, a free and open source software package (http://www.modelgui.org/). Parallel processing at the University of Nottingham was provided through its Augusta HPC service and the Beacon of Excellence for Precision Imaging, which provide access to high performance computing resources for the University's neuroimaging research community.

This study was supported by the the National Institute of Mental Health (R01-MH074457), the Helmholtz Portfolio Theme "Supercomputing and Modeling for the Human Brain" and the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 945539 (HBP SGA3).

References

Aboitiz F, Rodriguez E, Olivares R, Zaidel E (1997). Age-related changes in fibre composition of the human corpus callosum: sex differences. *Neuroreport*. 7(11):1761–1764. https://oce.ovid.com/article/00001756-199607290-00013/HTML.

Assaf Y, Pasternak O (2007). Diffusion Tensor Imaging (DTI)-based White Matter Mapping in Brain Research: A Review. *Journal of Molecular Neuroscience*. 34(1):51–61. doi: 10.1007/s12031-007-0029-0.

Basser PJ, Jones DK (2002). Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review. *NMR in Biomedicine*. 15(7-8):456–467. doi: 10.1002/nbm.783.

Behrens TEJ, Berg HJ, Jbabdi S, Rushworth MFS, Woolrich MW (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?. *NeuroImage*. 34(1):144–155. doi: 10.1016/j.neuroimage.2006.09.018.

Behrens TEJ, Woolrich MW, Jenkinson M, Johansen-Berg H, Nunes RG, Clare S, Matthews PM, Brady JM, Smith SM (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*. 50(5):1077–1088. doi: 10.1002/mrm.10609.

Bennett IJ, Madden DJ (2014). Disconnected aging: Cerebral white matter integrity and age-related differences in cognition. *Neuroscience*. 276:187–205. doi: 10.1016/j.neuroscience.2013.11.026.

Bennett IJ, Madden DJ, Vaidya CJ, Howard DV, Howard JH (2009). Age-related differences in multiple measures of white matter integrity: A diffusion tensor imaging study of healthy aging. *Human Brain Mapping*. p. NA–NA. doi: 10.1002/hbm.20872.

Burzynska AZ, Preuschhof C, Bäckman L, Nyberg L, Li SC, Lindenberger U, Heekeren HR (2010). Age-related differences in white matter microstructure: Region-specific patterns of diffusivity. *NeuroImage*. 49(3):2104–2112. doi: 10.1016/j.neuroimage.2009.09.041.

Caspers S, Axer M (2017). Decoding the microstructural correlate of diffusion MRI. *NMR in Biomedicine*. 32(4):e3779. doi: 10.1002/nbm.3779.

Catani M, Howard RJ, Pajevic S, Jones DK (2002). Virtual in Vivo Interactive Dissection of White Matter Fasciculi in the Human Brain. *NeuroImage*. 17(1):77–94. doi: 10.1006/nimg.2002.1136.

Chung HW, Chou MC, Chen CY (2010). Principles and Limitations of Computational Algorithms in Clinical Diffusion Tensor MR Tractography. *American Journal of Neuroradiology*. 32(1):3–13. doi: 10.3174/ajnr.a2041.

Colby JB, Soderberg L, Lebel C, Dinov ID, Thompson PM, Sowell ER (2012). Along-tract statistics allow for enhanced tractography analysis. *NeuroImage*. 59(4):3227–3242. doi: 10.1016/j.neuroimage.2011.11.004.

Daianu M, Jahanshad N, Nir TM, Toga AW, Jack CR, Weiner MW, Paul M Thompson ftADNI (2013). Breakdown of Brain Connectivity Between Normal Aging and Alzheimer's Disease: A Structural k-Core Network Analysis. *Brain Connectivity*. 3(4):407–422. doi: 10.1089/brain.2012.0137.

Douaud G, Jbabdi S, Behrens TEJ, Menke RA, Gass A, Monsch AU, Rao A, Whitcher B, Kindlmann G, Matthews PM, Smith S (2011). DTI measures in crossing-fibre areas: Increased diffusion anisotropy reveals early white matter alteration in MCI and mild Alzheimer's disease. *NeuroImage*. 55(3):880–890. doi: 10.1016/j.neuroimage.2010.12.008.

Eickhoff SB, Bzdok D, Laird AR, Kurth F, Fox PT (2012). Activation likelihood estimation meta-analysis revisited. *NeuroImage*. 59(3):2349–2361. doi: 10.1016/j.neuroimage.2011.09.017.

Eickhoff SB, Laird AR, Grefkes C, Wang LE, Zilles K, Fox PT (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human Brain Mapping*. 30(9):2907–2926. doi: 10.1002/hbm.20718.

Fonov V, Evans A, McKinstry R, Almli C, Collins D (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*. 47:S102. doi: 10.1016/s1053-8119(09)70884-5.

Fox PT, Lancaster JL (2002). Mapping context and content: the BrainMap model. *Nature Reviews Neuroscience*. 3(4):319–321. doi: 10.1038/nrn789.

Giorgio A, Santelli L, Tomassini V, Bosnell R, Smith S, Stefano ND, Johansen-Berg H (2010). Age-related changes in grey and white matter structure throughout adulthood. *NeuroImage*. 51(3):943–951. doi: 10.1016/j.neuroimage.2010.03.004.

Hafkemeijer A, van der Grond J, Rombouts SARB (2012). Imaging the default mode network in aging and dementia. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*. 1822(3):431–441. doi: 10.1016/j.bbadis.2011.07.008.

Jbabdi S, Behrens TEJ, Smith SM (2010). Crossing fibres in tract-based spatial statistics. *NeuroImage*. 49(1):249–256. doi: 10.1016/j.neuroimage.2009.08.039.

Jeurissen B, Leemans A, Tournier JD, Jones DK, Sijbers J (2012). Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Human Brain Mapping*. 34(11):2747–2766. doi: 10.1002/hbm.22099.

Jones DT, Machulda MM, Vemuri P, McDade EM, Zeng G, Senjem ML, Gunter JL, Przybelski SA, Avula RT, Knopman DS, Boeve BF, Petersen RC, Jack CR (2011). Age-related changes in the default mode network are more advanced in Alzheimer disease. *Neurology*. 77(16):1524–1531. doi: 10.1212/wnl.0b013e318233b33d.

Jones DK, Travis AR, Eden G, Pierpaoli C, Basser PJ (2005). PASTA: Pointwise assessment of streamline tractography attributes. *Magn Reson Med*. 53(6):1462–1467. doi: 10.1002/mrm.20484.

Kochunov P, Thompson PM, Lancaster JL, Bartzokis G, Smith S, Coyle T, Royall DR, Laird A, Fox PT (2007). Relationship between white matter fractional anisotropy and other indices of cerebral health in normal aging: Tract-based spatial statistics study of aging. *NeuroImage*. 35(2):478–487. doi: 10.1016/j.neuroimage.2006.12.021.

Kodiweera C, Alexander AL, Harezlak J, McAllister TW, Wu YC (2016). Age effects and sex differences in human brain white matter of young to middle-aged adults: A DTI, NODDI, and q-space study. *NeuroImage*. 128:180–192. doi: 10.1016/j.neuroimage.2015.12.033.

de Laat KF, Reid AT, Grim DC, Evans AC, Kötter R, van Norden AGW, de Leeuw FE (2012). Cortical thickness is associated with gait disturbances in cerebral small vessel disease. *NeuroImage*. 59(2):1478–1484. doi: 10.1016/j.neuroimage.2011.08.005.

Lohse C, Bassett DS, Lim KO, Carlson JM (2014). Resolving Anatomical and Functional Structure in Human Brain Organization: Identifying Mesoscale Organization in Weighted Network Representations. *PLoS Computational Biology*. 10(10):e1003712. doi: 10.1371/journal.pcbi.1003712.

Madhavan KM, McQueeny T, Howe SR, Shear P, Szaflarski J (2014). Superior longitudinal fasciculus and language functioning in healthy aging. *Brain Research*. 1562:11–22. doi: 10.1016/j.brainres.2014.03.012.

Marstaller L, Williams M, Rich A, Savage G, Burianová H (2015). Aging and large-scale functional networks: White matter integrity, gray matter volume, and functional connectivity in the resting state. *Neuroscience*. 290:369–378. doi: 10.1016/j.neuroscience.2015.01.049.

Mazziotta JC, Toga AW, Evans A, Fox P, Lancaster J (1995). A Probabilistic Atlas of the Human Brain: Theory and Rationale for Its Development. *NeuroImage*. 2(2):89–101. doi: 10.1006/nimg.1995.1012.

Menzler K, Belke M, Wehrmann E, Krakow K, Lengler U, Jansen A, Hamer HM, Oertel WH, Rosenow F, Knake S (2011). Men and women are different: Diffusion tensor imaging reveals sexual dimorphism in the microstructure of the thalamus, corpus callosum and cingulum. *NeuroImage*. 54(4):2557–2562. doi: 10.1016/j.neuroimage.2010.11.029.

Mollink J, Kleinnijenhuis M, van Cappellen van Walsum AM, Sotiropoulos SN, Cottaar M, Mirfin C, Heinrich MP, Jenkinson M, Pallebage-Gamarallage M, Ansorge O, Jbabdi S, Miller KL (2017). Evaluating fibre orientation dispersion in white matter: Comparison of diffusion MRI, histology and polarized light imaging. *NeuroImage*. 157:561–574. doi: 10.1016/j.neuroimage.2017.06.001.

Nooner KB, Mennes M, Li Q, Hinz CM, Kaplan MS, Rachlin AB, Cheung B, Yan C, Calhoun V, Courtney W, King M, Kelly AMC, Martino AD, Petkova E, Biswal B, Hoptman MJ, Javitt DC, Milham MP (2012). The NKI-Rockland sample: a model for accelerating the pace of discovery science in psychiatry. *Frontiers in Neuroscience*. 6:152. doi: 10.3389/fnins.2012.00152.

Pataky TC (2016). rft1d : Smooth One-Dimensional Random Field Upcrossing Probabilities in Python. *Journal of Statistical Software*. 71(7). doi: 10.18637/jss.v071.i07.

Rathee R, Rallabandi VPS, Roy PK (2016). Age-related Differences in White Matter Integrity in Healthy Human Brain: Evidence from Structural Mri and Diffusion Tensor Imaging. *Magnetic Resonance Insights*. 9:MRI.S39666. doi: 10.4137/mri.s39666.

Reid AT (2016). The quest for the human connectome: a progress report [blog post]. http://andrew.modelgui.org/blog/8.

Reid AT, Bzdok D, Genon S, Langner R, Müller VI, Eickhoff CR, Hoffstaedter F, Cieslik EC, Fox PT, Laird AR, Amunts K, Caspers S, Eickhoff SB (2016). ANIMA: A data-sharing initiative for neuroimaging meta-analyses. *NeuroImage*. 124:1245–1253. doi: 10.1016/j.neuroimage.2015.07.060.

Reid AT, Evans AC (2013). Structural networks in Alzheimer's disease. *European Neuropsychopharmacology*. 23(1):63–77. doi: 10.1016/j.euroneuro.2012.11.010.

Reid AT, Lewis J, Bezgin G, Khundrakpam B, Eickhoff SB, McIntosh AR, Bellec P, Evans AC (2016). A cross-modal, cross-species comparison of connectivity measures in the primate brain. *NeuroImage*. 125:311–331. doi: 10.1016/j.neuroimage.2015.10.057.

Reid AT, van Norden AGW, de Laat KF, van Oudheusden LJB, Zwiers MP, Evans AC, de Leeuw FE, Kötter R (2010). Patterns of cortical degeneration in an elderly cohort with cerebral small vessel disease. *Human brain mapping*. 31(12):1983–92. doi: 10.1002/hbm.20994.

Reveley C, Seth AK, Pierpaoli C, Silva AC, Yu D, Saunders RC, Leopold DA, Ye FQ (2015). Superficial white matter fiber systems impede detection of long-range cortical connections in diffusion MR tractography. *Proceedings of the National Academy of Sciences*. 112(21):E2820–E2828. doi: 10.1073/pnas.1418198112.

Rottschy C, Langner R, Dogan I, Reetz K, Laird AR, Schulz JB, Fox PT, Eickhoff SB (2012). Modelling neural correlates of working memory: A coordinate-based meta-analysis. *NeuroImage*. 60(1):830–846. doi: 10.1016/j.neuroimage.2011.11.050.

Sambataro F, Murty VP, Callicott JH, Tan HY, Das S, Weinberger DR, Mattay VS (2010). Age-related alterations in default mode network: Impact on working memory performance. *Neurobiology of Aging*. 31(5):839–852. doi: 10.1016/j.neurobiolaging.2008.05.022.

Schaeffer DJ, Johnston KD, Gilbert KM, Gati JS, Menon RS, Everling S (2018). In vivo manganese tract tracing of frontal eye fields in rhesus macaques with ultra-high field MRI: Comparison with DWI tractography. *NeuroImage*. 181:211–218. doi: 10.1016/j.neuroimage.2018.06.072.

Schilbach L, Bzdok D, Timmermans B, Fox PT, Laird AR, Vogeley K, Eickhoff SB (2012). Introspective Minds: Using ALE Meta-Analyses to Study Commonalities in the Neural Correlates of Emotional Processing, Social & Unconstrained Cognition. *PLoS ONE*. 7(2):e30920. doi: 10.1371/journal.pone.0030920.

Smith CM, Cooksey E, Duncan ID (2013). Myelin Loss Does Not Lead to Axonal Degeneration in a Long-Lived Model of Chronic Demyelination. *Journal of Neuroscience*. 33(6):2718–2727. doi: 10.1523/jneurosci.4627-12.2013.

Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM, Behrens TEJ (2006). Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*. 31(4):1487–1505. doi: 10.1016/j.neuroimage.2006.02.024.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, Luca MD, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, Stefano ND, Brady JM, Matthews PM (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*. 23:S208–S219. doi: 10.1016/j.neuroimage.2004.07.051.

Smith SM, Johansen-Berg H, Jenkinson M, Rueckert D, Nichols TE, Miller KL, Robson MD, Jones DK, Klein JC, Bartsch AJ, Behrens TEJ (2007). Acquisition and voxelwise analysis of multi-subject diffusion data with Tract-Based Spatial Statistics. *Nature Protocols*. 2(3):499–503. doi: 10.1038/nprot.2007.45.

Sowell ER, Peterson BS, Thompson PM, Welcome SE, Henkenius AL, Toga AW (2003). Mapping cortical change across the human life span. *Nature Neuroscience*. 6(3):309–315. doi: 10.1038/nn1008.

Takao H, Hayashi N, Ohtomo K (2013). Sex dimorphism in the white matter: Fractional anisotropy and brain size. *Journal of Magnetic Resonance Imaging*. 39(4):917–923. doi: 10.1002/jmri.24225.

Tang Y, Nyengaard JR, Pakkenberg B, Gundersen HJG (1997). Age-Induced White Matter Changes in the Human Brain: A Stereological Investigation. *Neurobiology of Aging*. 18(6):609–615. doi: 10.1016/s0197-4580(97)00155-3.

Tomasi D, Volkow ND (2011). Aging and functional brain networks. *Molecular Psychiatry*. 17(5):549–558. doi: 10.1038/mp.2011.81.

Tuladhar AM, Reid AT, Shumskaya E, de Laat KF, van Norden AGW, van Dijk EJ, Norris DG, de Leeuw FE (2014). Relationship between white matter hyperintensities, cortical thickness, and cognition. *Stroke*. 46(2):425–432. doi: 10.1161/STROKEAHA.114.007146.

Vanston JE, Strother L (2016). Sex differences in the human visual system. *Journal of Neuroscience Research*. 95(1-2):617–625. doi: 10.1002/jnr.23895.

Vernooij MW, de Groot M, van der Lugt A, Ikram MA, Krestin GP, Hofman A, Niessen WJ, Breteler MMB (2008). White matter atrophy and lesion formation explain the loss of structural integrity of white matter in aging. *NeuroImage*. 43(3):470–477. doi: 10.1016/j.neuroimage.2008.07.052.

Woolrich MW, Jbabdi S, Patenaude B, Chappell M, Makni S, Behrens T, Beckmann C, Jenkinson M, Smith SM (2009). Bayesian analysis of neuroimaging data in FSL. *NeuroImage*. 45(1):S173–S186. doi: 10.1016/j.neuroimage.2008.10.055.

Zhu T, Hu R, Qiu X, Taylor M, Tso Y, Yiannoutsos C, Navia B, Mori S, Ekholm S, Schifitto G, Zhong J (2011). Quantification of accuracy and precision of multi-center DTI measurements: A diffusion phantom and human brain study. *NeuroImage*. 56(3):1398–1411. doi: 10.1016/j.neuroimage.2011.02.010.

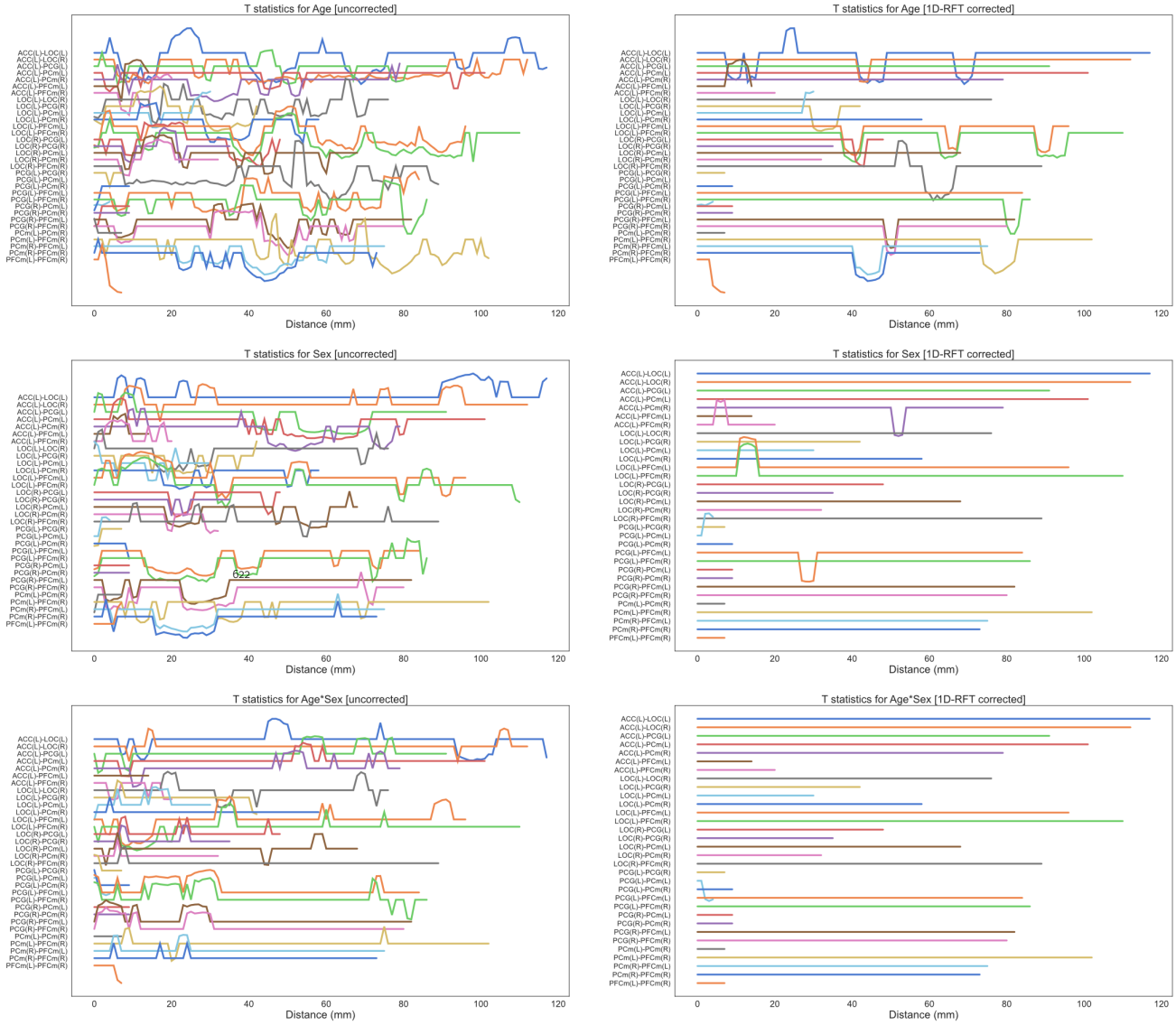

Figure 5–Figure supplement 1. Distance traces show the t-values for Age, Sex, and $Age \times Sex$, for all ROI pairs in the DMN. Plots on the left show uncorrected t-values, and plots on the right show t-values derived from one-dimensional random field theory (1D-RFT) and false-discovery rate (FDR) < 0.05. Line separation is $t = 1$.

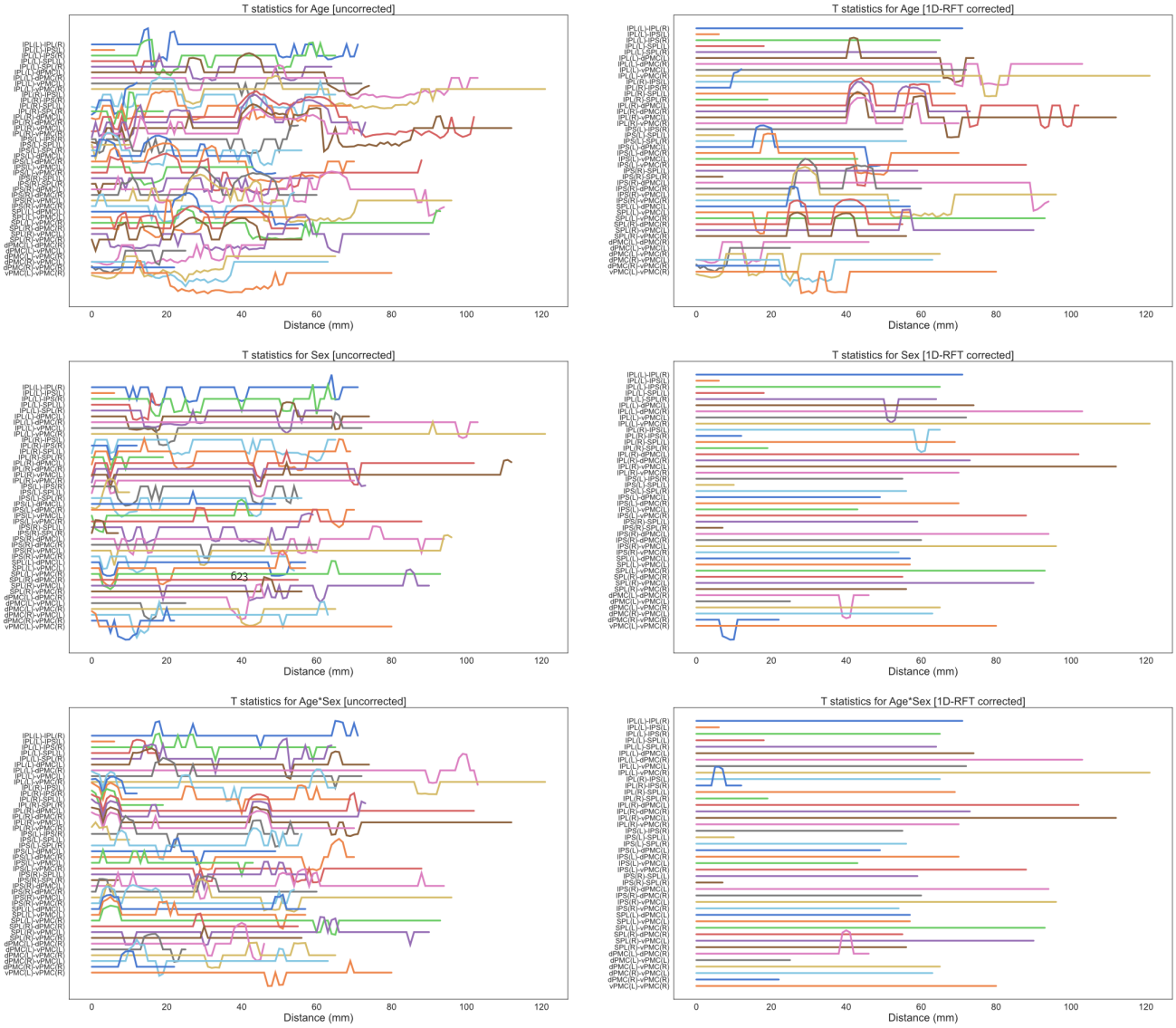

Figure 5–Figure supplement 2. Distance traces show the t-values for Age, Sex, and $Age \times Sex$, for all ROI pairs in the WWN. Plots on the left show uncorrected t-values, and plots on the right show t-values derived from one-dimensional random field theory (1D-RFT) and false-discovery rate (FDR) < 0.05. Line separation is $t = 1$.

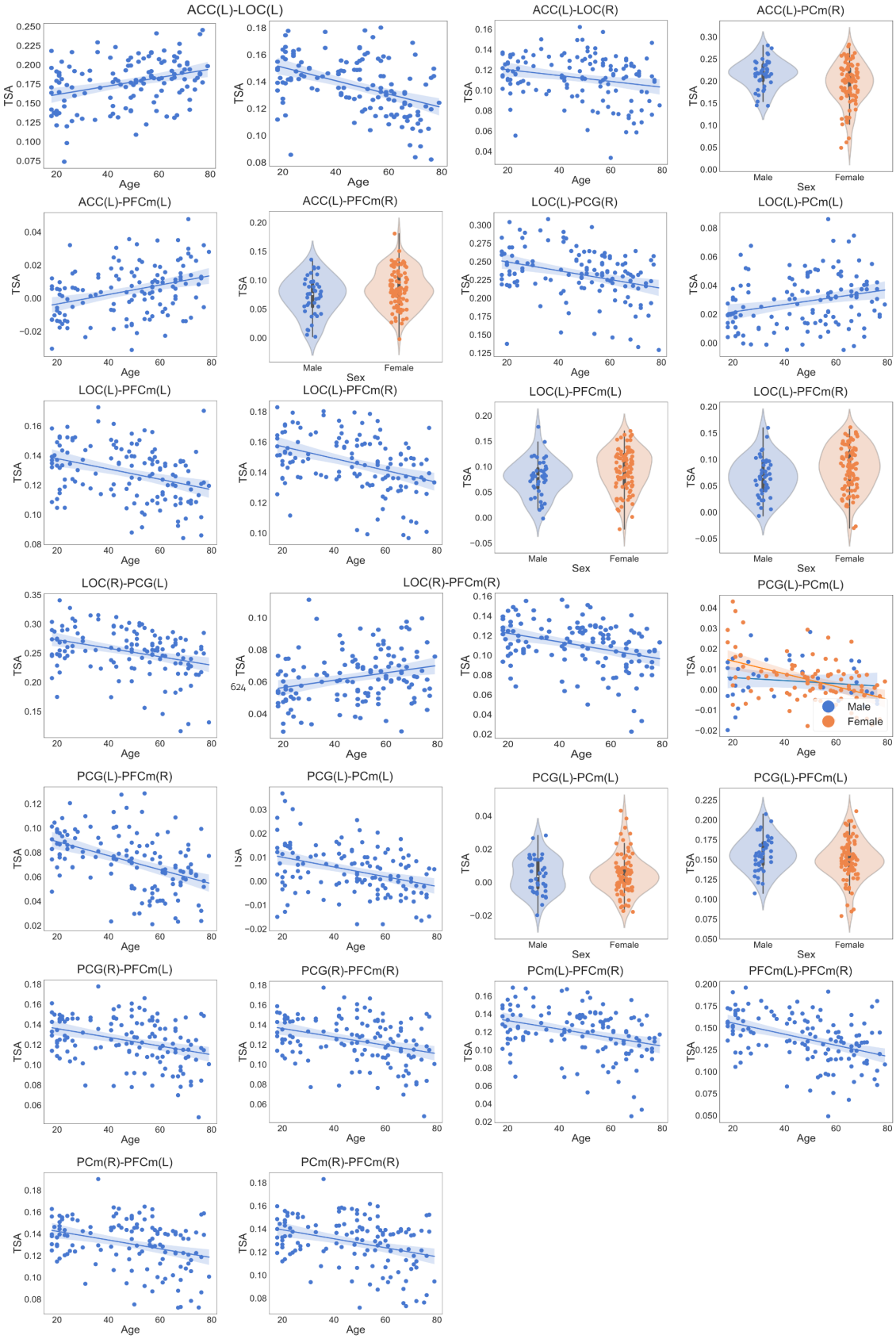

Figure 6–Figure supplement 1. Scatter and violin plots showing significant effects for all tracts in the DMN.

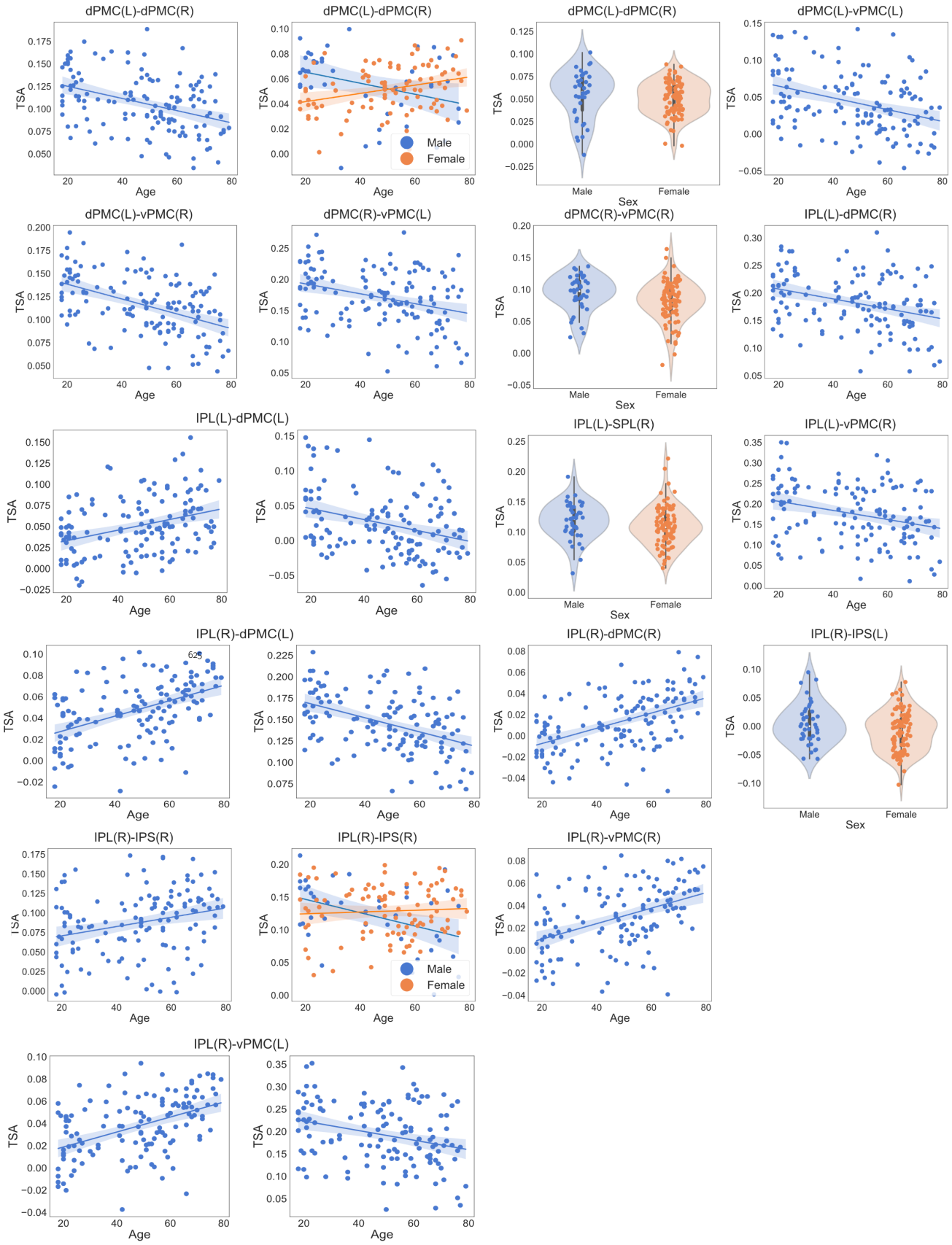

Figure 6–Figure supplement 2. Scatter and violin plots showing significant effects for all tracts in the WWN (Part A).

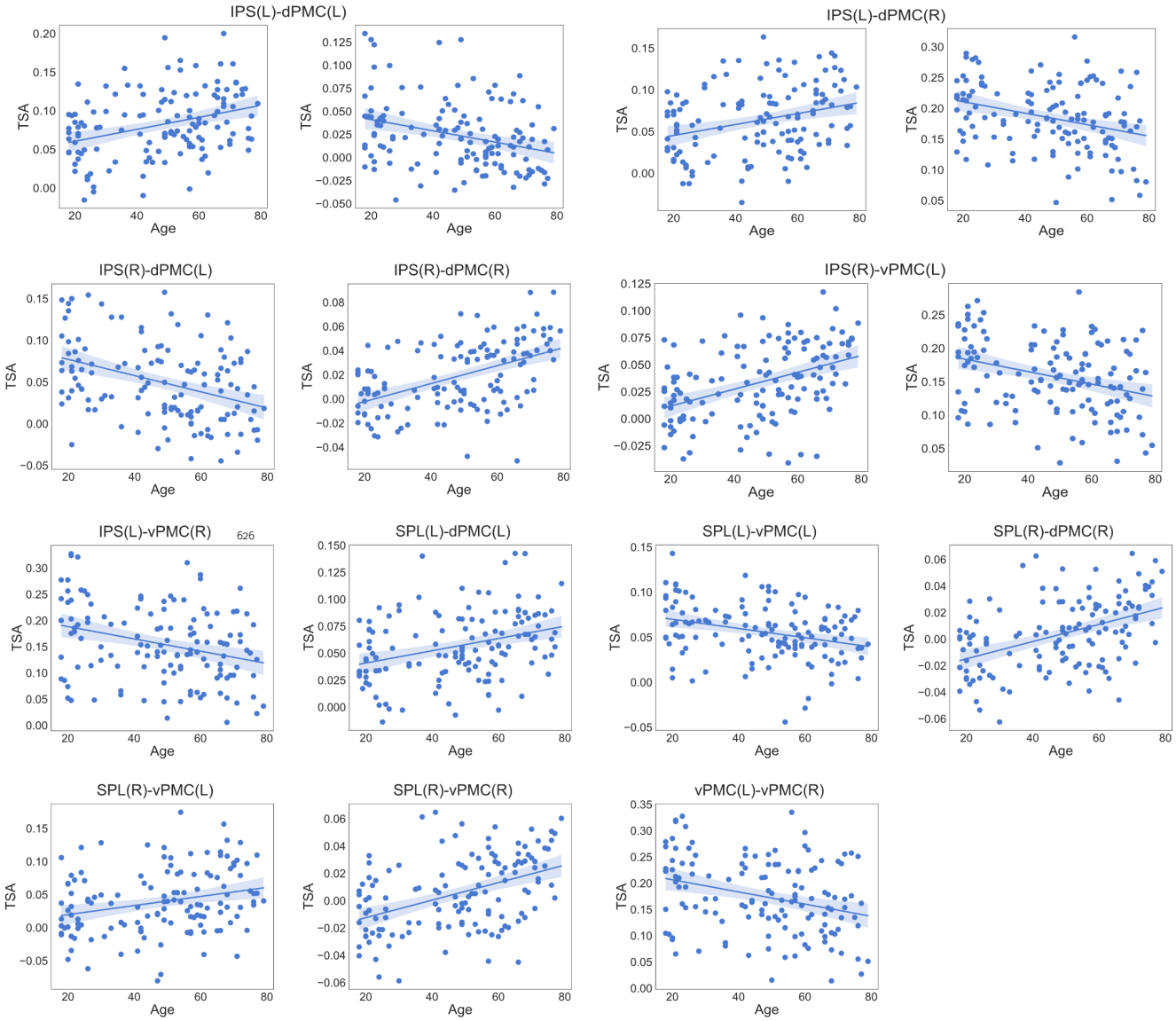

Figure 6–Figure supplement 3. Scatter plots showing significant effects for all tracts in the WWN (Part B).