A time-course study of actively stained mouse brains: DTI parameter and connectomic stability over one year

Jaclyn Xiao[1+], Kathryn J. Hornburg PhD[2+], Gary Cofer[2], James J. Cook[2], Yi Qi MD[2], and G Allan Johnson PhD[1,2*]

[1] Department of Biomedical Engineering, Duke University, Durham, North Carolina, USA
[2] Department of Radiology, Duke University Medical Center, Durham, North Carolina, USA

[+] These two authors contributed equally and should be seen as co-first author

[*] Corresponding Author:
G Allan Johnson
Department of Radiology
DUMC Box 3302, Durham, NC 27710
gjohnson@duke.edu

WORD COUNT:
7798 total

ABSTRACT

While the application of diffusion tensor imaging (DTI), tractography, and connectomics to fixed ex-vivo tissue is a common practice today, there have been limited studies examining the effects of fixation on brain microstructure over extended periods. This time-course study reports the changes of regional brain volumes and diffusion scalar parameters, such as fractional anisotropy across twelve representative brain regions as measures of brain structural stability. The scalar DTI parameters and regional volumes were highly variable over the first two weeks after fixation. The same parameters were stable over a two to eight-week window after fixation which means confounds from tissue stability over that scanning window are minimal. Quantitative connectomes were analyzed over the same time period with extension out to one year. While there is some change in the scalar metrics at one year after fixation, these changes are sufficiently small, particularly in white matter to support reproducible connectomes over a period ranging from two weeks to one year post fixation. These findings delineate a stable scanning period during which brain volumes, diffusion scalar metrics and connectomes are remarkably stable.

KEYWORDS
diffusion tensor imaging, magnetic resonance histology, time course, tissue fixation

INTRODUCTION

The use of diffusion tensor imaging (DTI) has grown over the past few years, spawning the development of technical advancements[1] and numerous applications in connectomics[2-4]. In multiple human populations studies, eg. the human connectome project, the ENGIMA consortium and the ADNI initiative, DTI has proven particularly useful in understanding human cognition, teasing out genetic differences and understanding pathology [5-8]. The application of DTI has expanded into animal studies, where the scalar parameters have proven useful in understanding of structural changes of the brain during maturation changes that are genetically driven and changes due to disease or injury [9 10].

The development of disease models and the genetic control that mouse studies offer provide a controlled approach to the understanding of disease mechanisms not possible in a clinical setting. The postmortem mouse allows for a level of imaging precision otherwise unachievable in human or live animal studies, since one can employ specialized staining methods and long scan times[11-13]. However, the study of the postmortem tissue brings its own unique challenges. Diffusion in unfixed postmortem tissues change over time due to autolysis[14]. Fixation, which affects water diffusivity in tissue and promotes cross-linking of proteins[15], can be used to preserve and stabilize tissue architecture. Consequently, the errors associated with a postmortem study, like tissue decomposition, can be minimized.

Active staining introduces fixative and contrast agent through direct cardiac infusion of the live animal minimizing tissue destruction from autolysis[12,16,17]. Active staining

reduces the spin lattice relaxation time (T1) through the use of a gadolinium contrast agent, leading to improved signal to noise, spatial resolution, and image contrast. However, until the fixative diffuses uniformly throughout the tissue, the brain remains susceptible to autolysis[15,18]. Crosslinking induced by fixation takes place on a time scale of minutes to years during which the relaxation and diffusion properties can undergo profound changes[18,19]. Furthermore, the effects of fixation on spin-spin relaxation (T2) over time post fixation[19-22] can lead to additional uncertainties in DTI parameters in fixed tissue.

Previous studies have examined diffusion parameters in live versus ex vivo fixed tissue. Most have concluded that while the absolute values may change, relative values of MR parameters in fixed tissue (eg. ratio fractional anisotropy of gray matter to white matter) are similar to those in vivo[23-26]. However, few studies have looked at the variation of DTI parameters at extended time post fixation. Most reports suggest scalar DTI parameters (fractional anisotropy, FA; axial diffusivity, AD; and radial diffusivity, RD) remain unchanged over time post fixation[27-30]. The impact of post mortem interval (PMI), i.e. the time between death and fixation during which the tissue undergoes autolysis has been studied [14]. But scan interval (SI), the time post fixation until scanning, has been seen as something of minimal impact[31].   Other studies have investigated the effects of SI on T1 and T2. Most of these studies found that the T2 decreases sharply within the immediate period following fixation – a period found to last a week - before reaching a plateau up to 1-year postmortem[20-22,30,32]. It could thus be hypothesized that the DTI parameters, which are affected by T1 and T2, should stabilize following an initial period after fixation, once the brain has fully fixed. As emphasized through these dichotomous hypotheses, the previous literature observations with regards to DTI stability over time have been mixed.

While studies have separately addressed the changes in tissue variables such as volume or diffusion metrics over time post fixation, no study has yet to cohesively analyzed all these metrics across a large and consistent set of time points to allow group comparative analysis. Furthermore, no study has investigated the effects of active staining perfusion fixation on parameters that contribute to connectome tracking nor analyzed statistical differences in connectomes across time.

Our current work flow for quantitative connectomics involves active staining of a batch (n=6-24) of animals at a time. Our scan protocol acquires 51 three-dimensional (3D) volumes requiring a total scan time per specimen of ~12 hrs. The practical consequence is that there exists a spread in the interval between perfusion and scanning. The goal of this study was to understand the sources of variability in actively stained mouse brain over time to establish a scan interval during which changes from fixation would have little effect on group comparisons of scalar DTI measures, volumes and connectomes.

EXPERIMENTAL

All animal experimentation was performed in accordance with the National Institutes of Health Guidelines for Animal Care and Use of Laboratory Animals and the protocols of the Duke University Animal Care and Use Committee.

Specimen Preparation
Eight male C57BL/6J adult (90 day) mice were acquired (The Jackson Laboratory, Bar Harbor, ME, Cat# 000664) for MR imaging. The mice were prepared using an active staining technique [17] in which the mice were transcardially perfused through the left heart ventricle with a solution of 10% ProHance (Bracco Diagnostics, Princeton NJ), and 10% buffered formalin. Following fixation, the heads were removed and placed in 10% buffered formalin for 24 hours to allow fixation to continue. At the end of 24 hours, the brains, still in the cranial vault were removed from the formalin, rinsed and immersed in a 0.5% ProHance/phosphate buffered saline (PBS) solution to restore the T2.

Data Acquisition
All data were acquired on a 9.4-Tesla Agilent Direct Drive MRI system utilizing a Stesjkal-Tanner spin echo sequence[33]. Acquisition time was reduced using a compressed sensing algorithm with a compression factor of 8 [34]. A 420 x 256 x 256 image matrix was acquired over an 18.9 x 11.52 x 11.52 mm^3 field of view with a native isotropic spatial resolution of 45 μm^3 using MRI scan parameters (repetition time (TR) = 100 ms; echo time (TE) = 12.7 ms, b-value = 4000 s/mm^2). Forty-six diffusion-weighted 3D volumes were acquired with the gradient direction of each of volume uniformly distributed on the unit sphere. Five base line (b0) volumes were acquired. A post processing pipeline registered the five b0 volumes together to generate the average baseline. The diffusion weighted volumes were registered to this baseline to correct for eddy currents[13].

MR images were gathered at seven time points post-fixation (2-3 days, 1 week, 2 weeks, 3 weeks, 4 weeks, 8 weeks, and nominally 1 year). The scan intervals for each specimen are shown in Figure 1. Two specimens were scanned for the 2-3 day, 2-week, 3-week, 4-week, and 8-week time period. Four specimens were scanned at all time points. Two additional specimens were scanned only at the 1-year time point.
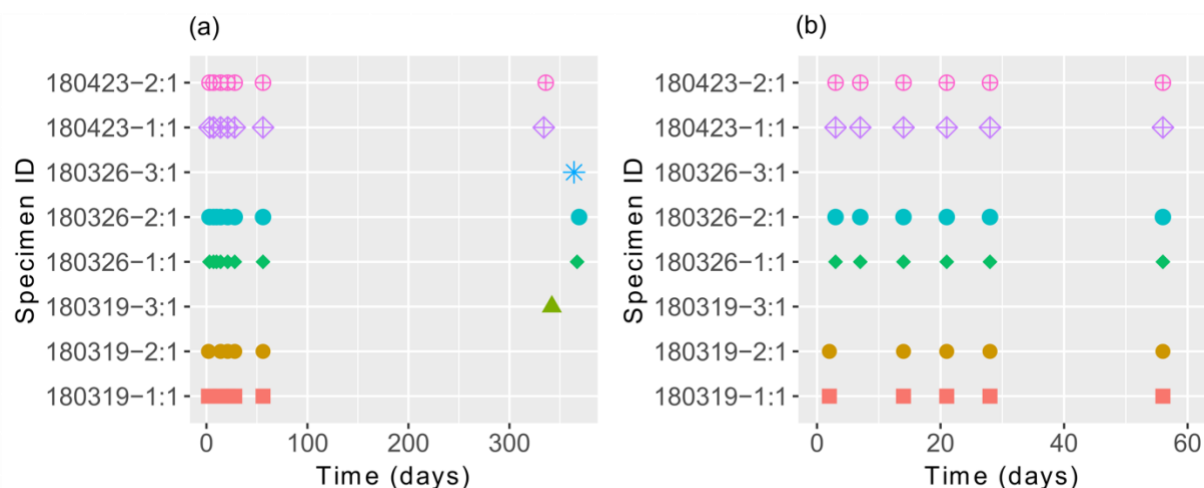
Figure 1: Scan time after perfusion for the 8 specimens used in this study: (a) all the time points (2-3 days to 1 year post fixation), and (b) detail of the time region from 2-3 days to 8 weeks post fixation.

Atlas Mapping to MR Images

The SAMBA pipeline [35] was used to register the Waxhom Space (WHS) atlas[36] with labels onto each specimen. The pipeline uses two steps to create the label mapping: an affine step using the diffusion weighted image (DWI) and diffeomorphic registration utilizing FA image. The WHS atlas includes 166 (symmetric) regions on each hemisphere in an adult C57BL/6J mouse which are mapped onto all the diffusion data in this study[36]. This provides ready extraction of the mean and standard deviation of the volume and diffusion scalar metrics from all 332 regions of interest in all the scans. It also provides the nodes used in defining the connectome for each specimen.

Selection of Characteristic Regions

Twelve regions of interest (ROI) were followed across the entire time span; nine white matter and three gray matter regions. These regions are shown in Figure 2. The white matter regions: the anterior commissure (AC), corpus callosum (CC), cingulum (CG), fimbria (FI), internal capsule (IC), raphe nucleus (RN), cerebral peduncle (CP), middle cerebellar peduncle (MCP), and inferior cerebellar peduncle (ICP), were selected because of their relatively large volumes, to minimize partial volume errors, while still being representative for the whole brain. Some of these regions represent challenges with mapping: proximity to the ventricles (FI) and smaller volume (cerebellar peduncles). The grey matter regions: somatosensory cortex (S1), cingulate cortex (C), and primary visual cortex (V), were selected because these represent bulk cortical regions which also represent a challenge with automated mapping[37,38]. The contrast between cortical regions (eg. motor and sensory) is not particularly high in the MR images. In addition, layers within some cortical sections are not well delineated. So, we have combined some of these regions to simplify the segmentation.
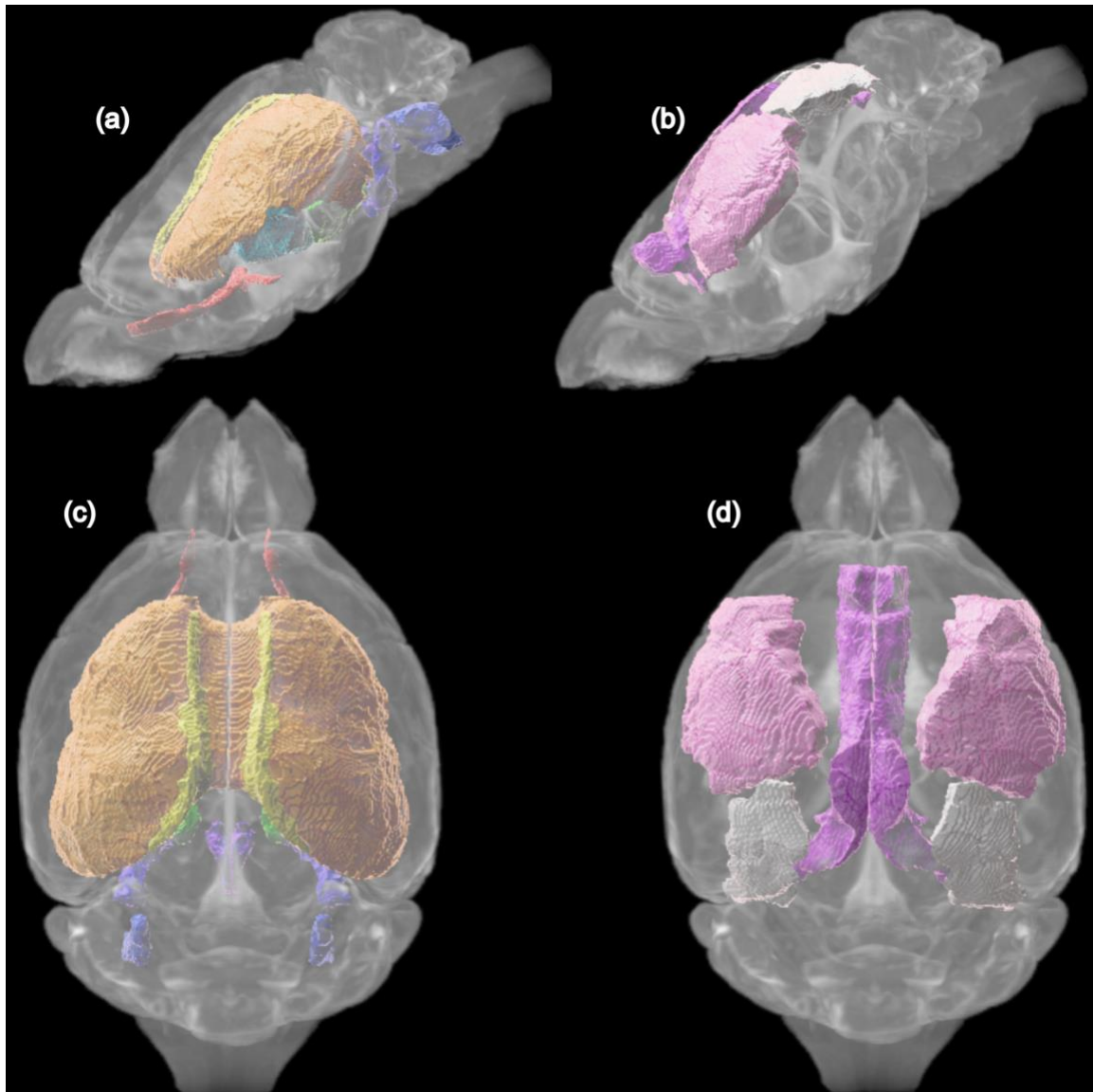
Figure 2: The regions analyzed in this study visualized on an example C57BL/6J specimen with (a and c) showing the white matter regions – AC (red), CC (orange), CG (yellow), FI (teal), IC (blue), RN (purple), CP (green), MCP (violet), and ICP (dark blue) – viewed from superior and sagittal axis respectively and (b and d) showing the grey matter regions – S1 (pink), C (dark pink), V (white) – viewed from superior and sagittal axis respectively.

Characterization of DTI Scalars and Volume Measurements
In order to minimize partial volume effects at the periphery, ImageJ was used to erode each ROI at every time point prior to statistical analysis of the scalar DTI parameter. Figure 3 demonstrates the impact of the erosion on the histogram of FA over three ROI, exemplifying the narrowing and normalization of voxel histograms, and a gradual increase of peak FA with increasing erosion. We chose to use erosion level 1 to balance the removal of edge effects while maintaining an adequate sample of the ROI.
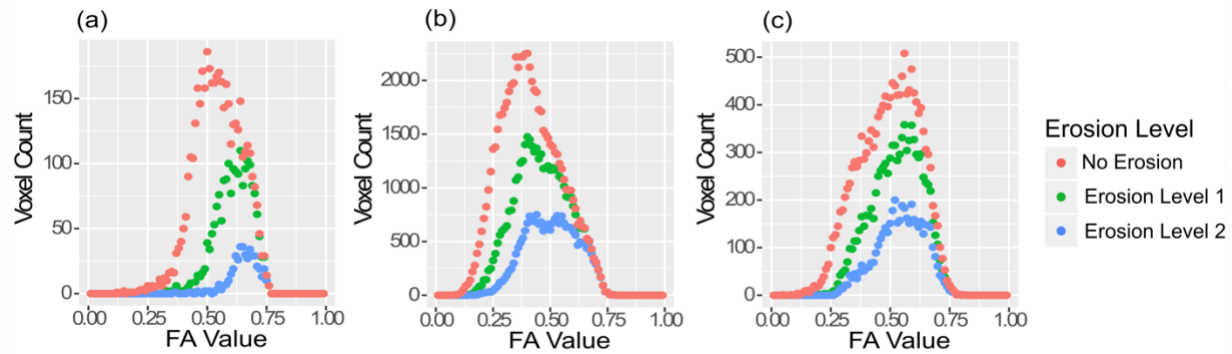
Figure 3: Effects of erosion on the histogram of FA for the (a) anterior commissure, (b) corpus callosum, and (c) internal capsule at the 3 weeks after fixation.

The peak values of each histogram were extracted at each time point as a main study metric. The means and standard deviations of these peaks were the basis of statistical comparison. Data were compared using an ANOVA test to determine if there was any significant variation in time with the given diffusion scalar or volume for the given region- - with posthoc pairwise comparison via Tukey's test to see the most notable time point variations.

The volume variation of each region was calculated in a similar way as the DTI metrics (using an ANOVA test then posthoc pairwise testing). No erosion was applied to the region sizes and the mean percent volume differences, compared to three week time point after fixation, is presented as the key study metric.

Tractography
DSI studio [39] version 20181115 was used to generate tractography for the connectomes. We used the GQI (Generalized Q-Sampling Imaging) model which generates a spin distribution function using a model free reconstruction[40]. Tractography was generated using a tracking plan which terminated after generating 2 million fiber tracts with lengths between 0.5 – 200 mm utilizing all fiber directions at a turning angle of 45 degrees with smoothing and step size of 0.01. The GQI algorithm generates a quantitative anisotropy (QA) analogous to FA. The QA can be normalized by its maximum value of QA, generating the normalized QA (NQA) on a range [0 1]. The NQA threshold for tracking is found for each specimen from rank ordering of the NQA histogram and finding the NQA value corresponding to 10% of the total number of points in the histogram. The mean NQA threshold value for all the specimens over all the times point was 0.105 ± 0.011 (mean ± standard deviation). The connectome values were calculated by counting the number of tracts passing through each region. All connections were included, i.e. there was no thresholding.

Similar methods were used in a recent study of connectome differences between four different strains of mice[9]. In that analysis, approximately five million tracts were generated for each specimen in each strain. In that study, there was little difference in connectomes generated with two million tracts and five million tracts. Thus, we standardized on two million tracts in this new study normalized to an idealized

C57BL/6J brain volume (435 mm^3) to account for potential changes in brain size across specimens. The mean scale factor is 2.60 +/- 0.07 and the average brain volume of this data set is 453 +/- 12.0 mm^3.

The degree of connection - a measurement of the number of connections from one brain region to another - was extracted for all the regions in the atlas [41]. As noted above, several cortical areas were combined leaving a total number of possible connections at 296. The degree measurement was normalized to this value so that degree of every region ranged between [0 1]. The variation in degree for the 12 example regions is plotted as a function of time after perfusion in the same manner as was used for the scalar DTI metrics.

Omni-MANOVA
Connectome differences were compared using a graph theory process called OMNI-MANOVA[42]. The connectome is a matrix that expresses the strength of connectivity between each node (i.e. ROI) of the brain. A profile of a specific node, eg. CC, can be constructed based on all the tracts passing through that node. Constructing these profiles allows all the nodes to be expressed as a smaller vector in a common space reducing the complex connectome (332x332 for each specimen) to a more manageable 332x ~3-5 for each specimen based on selecting the principle eigenvalues.

The reduction of terms allows use of multi-variate statistics like MANOVA to determine if the ideally-null tractography fail to reject the null hypothesis between sets of time points. The MANOVA test is a multi-variate form of the ANOVA (a one-way analysis of variance) applied to the node profile described above. We first check that the connectomes are equivalent across the time points and then, by looking at the regions that are significantly different, investigating if that region is known to be problematic, i.e could produce varying tractography due to its structure and features. The p-values corresponding to each region of interest are corrected for multiple comparisons using a Benjamini-Hochberg correction[43]. As we can compare visually the tractography differences, we can utilize less conservative correction methods and then check which nodes are truly different [44].This is unlike other statistics applications where the concern for false positive results outweighs benefits of additional contributions obtained by other multiple comparison corrections and where there is no additional check [45], such as afforded by our visualization of tractography.

A classical multidimensional scaling was formed in 2-D based on the distances of the Omni-MANOVA embedding tensor via a frobenius norm[46,47] as a check of how each specimen and group relate to the entire set of study animals. The classical multidimensional scaling is similar to principle component analysis[47]. Rather than focusing on illustrating maximum variation in the set (principle component analysis), the multidimensional scaling illustrates the similarity of the scans. This embedding generates a simplified view, removing region-based analysis of the embedding utilized in the Omni-Manova, so we can make assumptions of the general relationship of our set and compare the spread of groups scanned at different times within the same embedding space.

Differential Tractography

Once we determined the periods for comparison using the DTI scalar metrics and graph theory based connectomic analysis, we investigated tract differences between the early period (<2 weeks) and the 1 year time point and the stable period using differential tractography in DSI Studio[48]. The method was developed for comparing scans of the same patient at different time points by mapping data sets into a common space and tracking the difference in anisotropy along all the tracks. By looking at the difference in anisotropy, one can determine the stability of the tracking across those two periods. Differential tractography was performed on 4 different specimens in which the 2 week and 8 week time points were compared and the 8 week and 1 year time points were compared.

The initial tracts for differential tractography were generated using DSI Studio version 20200627 to create a series of 50,000 seeds with the same parameters as the initial tracking. We used QA instead NQA, because the percent change criteria in differential tractography is based on QA. The QA thresholding was determined at 10% of the actual QA values. For comparisons between time points, we looked at the changes in QA with four different thresholds, ±20% and ±50%, which had been used previously in a clinical study of demyelination and axonal loss[49].

RESULTS

A representative dataset for this work is available on VoxPort (https://civmvoxport.vm.duke.edu/voxbase/studyhome.php?studyid=746). To access the database, a user will need to register an account. Additional data used in this study can be obtained via contact with corresponding author (gjohnson@duke.edu).

Fractional Anisotropy

The mean peak FA and standard deviation across the specimens at each time point are shown in Figure 4. The mean peak FA of the gray matter regions, Figure 4j-4l with range 0.04-0.24 (-), is substantially lower than the white matter regions, Figure 4a-4i with range 0.23-0.83 (-). The mean peak FA of all ROIs broadly follows an exponential decay, with the most dramatic change from 2-3 days to 2 weeks, decreasing to a plateau from 2 to 8 weeks. For most regions, the FA is relatively unchanged up to a year after fixation.

An ANOVA test revealed all ROIs changed significantly ($p < 0.05$) over the entire time period (2-3 days to 1 year). Posthoc Tukey testing (with $p < 0.05$) determined the intervals of lowest FA variability: the stable time period. All twelve regions showed no significant change in the peak FA from 2 to 8 weeks after fixation.

Ten regions (S1, RN, V, IC, MCP, FI, CC, CG, C, AC) underwent significant changes in peak FA from 2-3 days to 2 weeks after fixation, and nine regions (S1, RN, V, IC, MCP, FI, CC, C, AC) showed a statistically significant difference between at least two points in

the stable time period (2 weeks to 1 year) and the 2-3 day period. In Figure 4, significant differences in at least two points in the stable time period are marked by red asterisks for each region. The FA peak of only one region (C) changed significantly from 1 week to 2 weeks and also differed from at least two points in the stable time period. In contrast, the peak values significantly decreased in nine regions (S1, RN, V, IC, MCP, FI, CC, C, AC) from 2-3 days to 1 week after fixation.

The FA in five of the regions (S1, RN, MCP, CG, C) differed between at least two points in the stable period and the one-year time point. But only two of the regions (RN and C) were statistically different from 8 weeks to 1 year. Further investigation revealed that for the five regions that were statistically different at the one year time point, the FA differed by a maximum of 20% relative to the stable period (S1: -18%, RN: -14%, MCP: -5%, CG: -17%, C: -20%), with an overall average change of -8% across all regions.
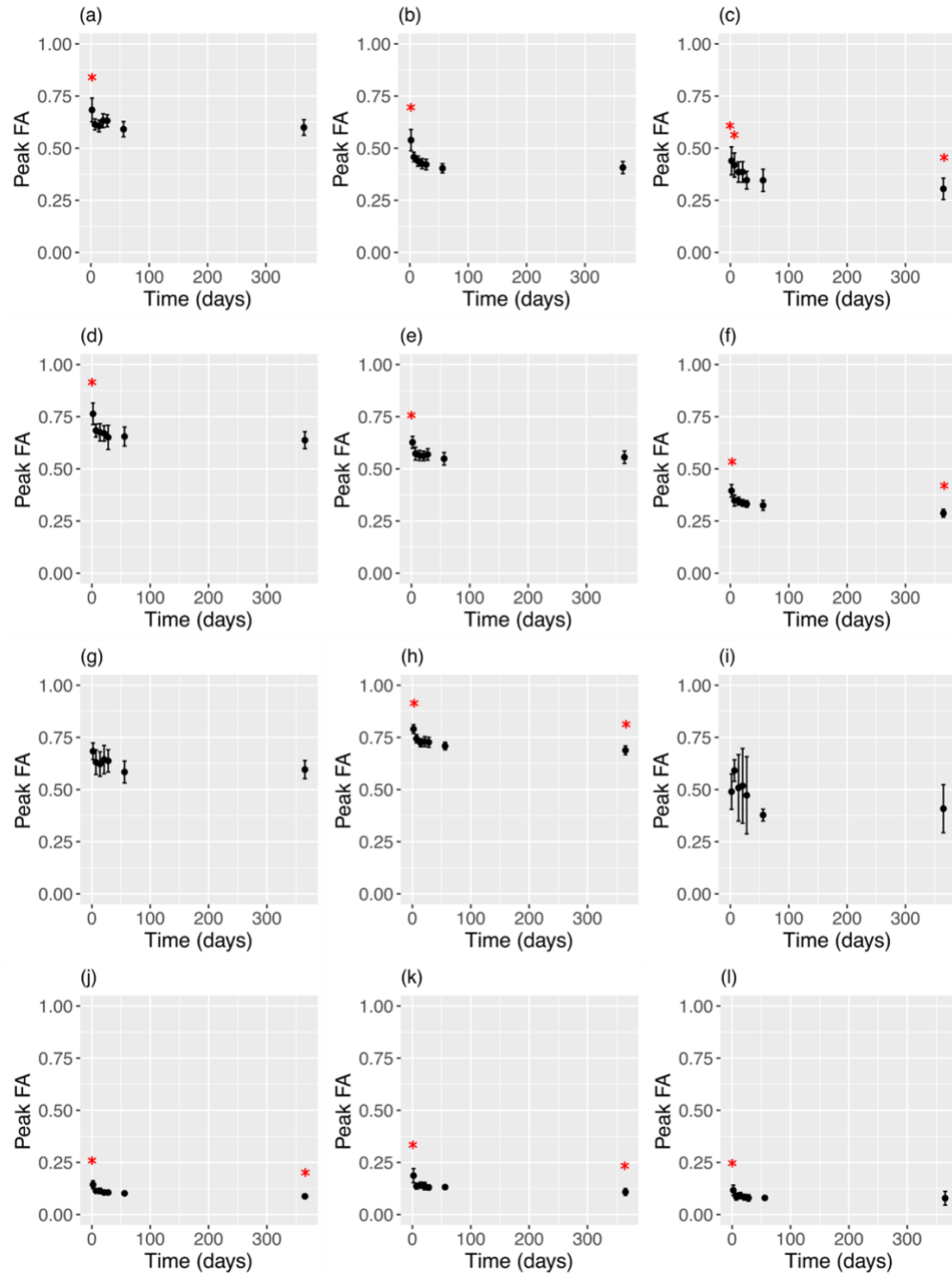
Figure 4: Mean and standard deviation of peak FA across all time points for white matter regions: (a) AC, (b) CC, (c) CG, (d) FI, (e) IC, (f) RN, (g) CP, (h) MCP, and (i) ICP; and gray matter regions: (j) S1, (k) C, and (l) V. Red asterisks denote the time points for a region that differs significantly (p<0.05) in peak FA from at least 2 other time points in the stable period (2 to 8 week post fixation).

Volume
We investigated the volume changes across the 12 selected regions by normalizing to 3 weeks after fixation, which is around the middle of the stable period in the FA. Figure 5 shows the average and standard deviation of the volume change across all time points.

There were no statistically discernable changes in volume for any of structures during the stable period with a small increase in the volumes at 1 year after fixation.

An ANOVA test determined that eight of the twelve regions had statistically significant change in volume across time ($p < 0.05$; S1, RN, V, IC, MCP, ICP, C, CP). The four regions without significant volume differences over time (FI, CC, CG, AC) were white matter regions. Post-hoc testing was performed on the eight regions that changed significantly ($p < 0.05$) in volume over time in order to discern the time period in which the volume is stable. Similar to the results for the diffusion metrics, there were no significant changes in volume between 2 and 8 weeks after perfusion. Outside of this stable period, we see with the pairwise comparison that volumes at the one year time point varied significantly from the stable period in five of the eight regions (S1, RN, V, IC, CP). While the number of regions with volumes that were different at one year is large, the percent difference of volume is small, with a mean change of 5% (S1: 5%, RN: 5%, V: 6%, IC: 5%, CP: 4%). Unlike the FA, only the volumes of two of the eight regions (MCP, ICP) differed significantly at the 2-3 day time point compared to at least 2 other time points in the stable period. No volume of any region at 1 week after fixation differed significantly from at least two time points in the stable period nor significantly differed from the volume at the 2 week time point. In Figure 4, the significant changes from at least two points in the stable time period are marked by red asterisks for each region.
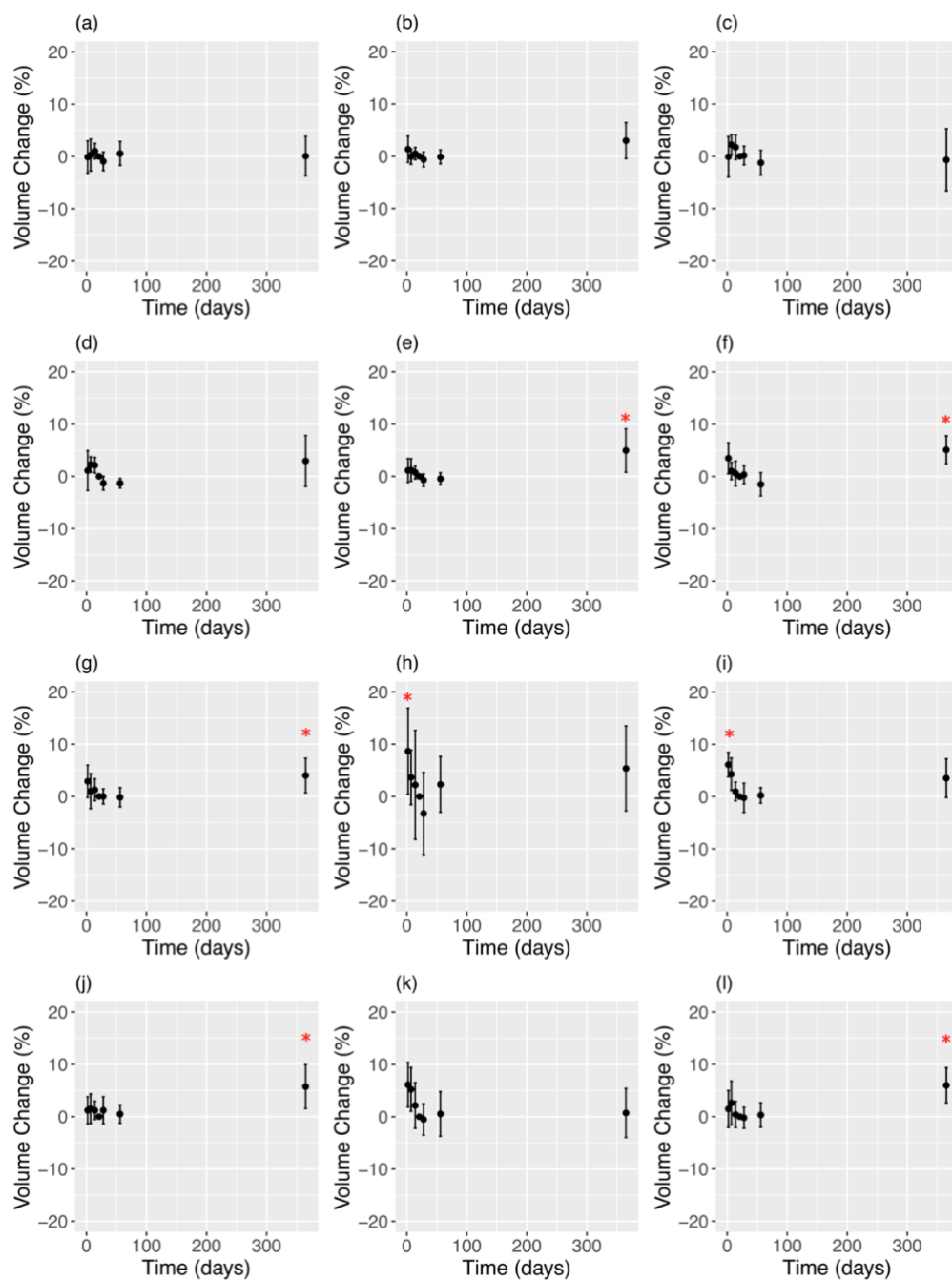
Figure 5: Mean and standard deviations of percent change in volume compared to the 3-week time point across all scans for white matter regions: (a) AC, (b) CC, (c) CG, (d) FI, (e) IC, (f) RN, (g) CP, (h) MCP, (i) ICP; and gray matter regions: (j) S1, (k) C, (l) V. Red asterisks denote the time points at a region that significantly (p<0.05) differs in volume (mm³) from at least 2 other time points in the stable period (2 to 8 weeks post fixation).

Supplemental Scalar Metric Analysis

Four additional scalar metrics (mean diffusivity (MD), radial diffusivity (RD), axial diffusivity (AD), and normalized degree) were analyzed for this work in the same way as FA and volume. The resulting figures and analysis are included within the supplement. The results for those additional metrics follow that of the FA and volume.

Omni-MANOVA

The Omni-MANOVA analysis in Figure 6 provides an expanded analysis of the stability of the connectome by showing the reduced vector comparisons at multiple time points. In Figure 6a, in which all the time points are compared, there is a marked separation of the 2-3 day and 1 week data (red and green) from the rest of the groups (blue, cyan, magenta, yellow and black). This indicates the 2, 3, 4, 8 week, and 1 year measurements are more similar to one another than the early time points. The rank ordered p-value results of the MANOVA test, identified 160 of the 332 regions that reject the null hypothesis that the connectome is unchanged, i.e. the vectors describing connections of 160 nodes may have changed. In Figure 6b, the 2-3 day and 1 week time points have been removed. The 1 year time point does not overlap the stable period as closely as in Figure 6a but the difference is less than one standard deviation of the constituent groups and the reconstruction covers less area showing that the data are more similar than all the time groups together. The MANOVA test indicates that profiles (vectors) of 4 out of the 332 regions of interest reject the null hypothesis, Table 1. All are cortical and none are bilateral. Boundaries of these regions are poorly defined in our atlas so the use of FA as an automated mapping criterion results in higher variability in mapping. Suspecting this variation is due to mainly the boundaries of the included data within the embedding, 2 weeks and 1 year, we continue by trying two different embeddings each removing one of these boundary time points. In Figure 6c, the 2-week data has been removed, the plot shows 3 weeks to 1 year. There is little change from Figure 6b and no determined significant changes in connectivity. This helps us conclude that the stable period is 2-8 weeks. Figure 6d plots the data for this 2-8 week stable period. Note the spread (standard deviation) in this plot is small and not that different than Figure 6c. In addition, there were no profiles which had changed significantly. This suggests that the 1 year data is not drastically different from that gathered during this stable period.
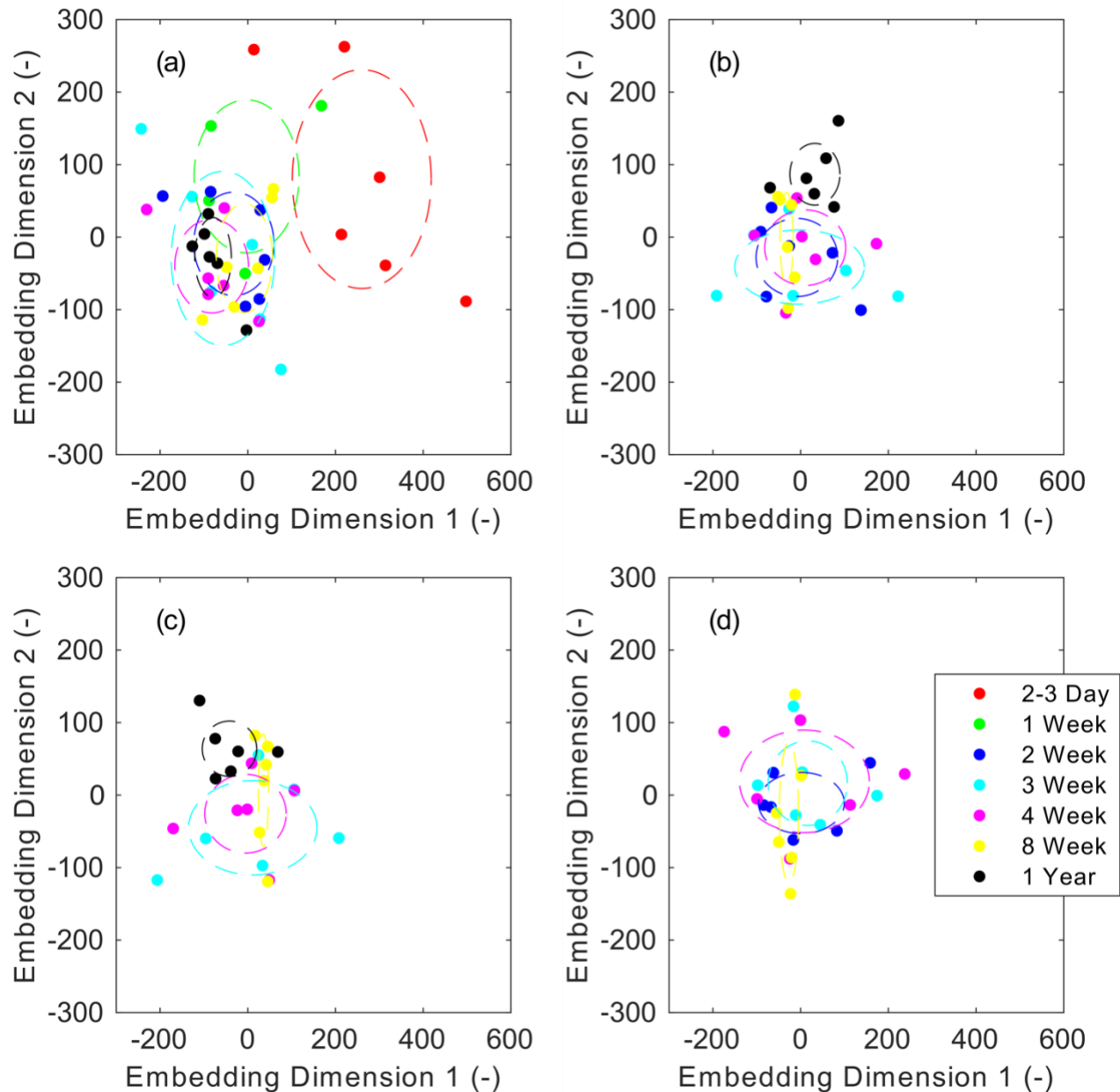
Figure 6: Reconstructions of the adjacency spectral embedding created using Omni-MANOVA framework resulting from (a) all time points, (b) 2, 3, 4, 8 week, and 1 year time points, (c) 3, 4, 8 week, and 1 year time points, and (d) 2, 3, 4, and 8 week time points. The dotted lines are centered on the mean of individual points in the specific time periods with the spread corresponding to the standard deviation in embedding dimension 1 and 2. Individual points in the plots represent each specimen in this study.

Table 1: The regions associated with 2, 3, 4, 8 week and 1 year embedding that failed to reject the null hypothesis according to BH correction, sorted by raw p-value.

Region Abbreviation	Region Name	Hemisphere
V1B	primary visual cortex, Binocular Area	Right
DLEnt	dorsolateral entorhinal cortex	Right
A24aPrime	cingulate cortex, area 24a prime	Right
S1FL	primary somatosensory cortex, forelimb region	Left

Differential Tractography

Differential tractography was employed as a final test of tissue stability. We compared four specimens at 2, 8 week, and 1 year time points. The percent change in QA at 2 weeks and 1 year utilized the 8 week scans as the base point. Using 50,000 seeds with whole brain tracking, the average track number and length for the groups were nominally the same, 21482 +/- 240 and 5.03 +/- 0.111 mm respectively, at the three time points, indicating that there is not a drastic change in the diffusivity on which the tractography is based. The average numbers associated with each time point are shown in Supplemental Table 1.

When QA *change* is used as the thresholding condition, the mean tract length is approximately a fifth of the tract length in the QA images. The mean tract number and mean tract length at ±20% and ±50% thresholds are shown in Supplemental Table 2. The differential tracts are small, scattered, and likely are related to edge effects. Differences between the 2 and 8 week scans are negligible. The differences between the 8 week and 1 year scan are most notable at -20% suggesting some systematic though small reduction in FA as the tissue rests in PBS for a year. One can relate the differential tracts to the global length and density of tracts comparison to all tracts passing through the left portion of the CC for the same specimen (180423-2:1) at the three time points from the 50,000 seeded whole brain track data. We picked 180423-2:1 because in the whole brain tracking, the response of number of tracts and track length was most close to the average result. Example images of this tractography at all three time points are shown in Figure 7b-7d. Both Figure 7b and Figure 7d used Q-Space Diffeomorphic Reconstruction[50] (QSDR) to place the 2 week and 1 year data into the same orientation alignment as the 8 week data (Figure 7c), and thus needed to use the 8 week label for segmentation of CC tracts. We see more grouping of key tracts, such as the arching structure of the CC and projections into the olfactory bulb, than illustrated by the difference tractography. As a comparison, the average number of tracks shown in Figure 7b-7d is 2520 +/- 60 which is approximately the same number of average tracks found for all specimen at -20% threshold comparison of 1 year and 8 weeks. Interestingly, while the number of tracts was the same, the average tract length of Figure 7b-7d is 11.5 +/- 0.796 mm which is an order of magnitude longer than the multi-specimen averaged -20% threshold comparison of 1 year and 8 weeks. Thus, this

confirms that our remaining differences in tractography shown in the differential tractography are negligible on the whole.
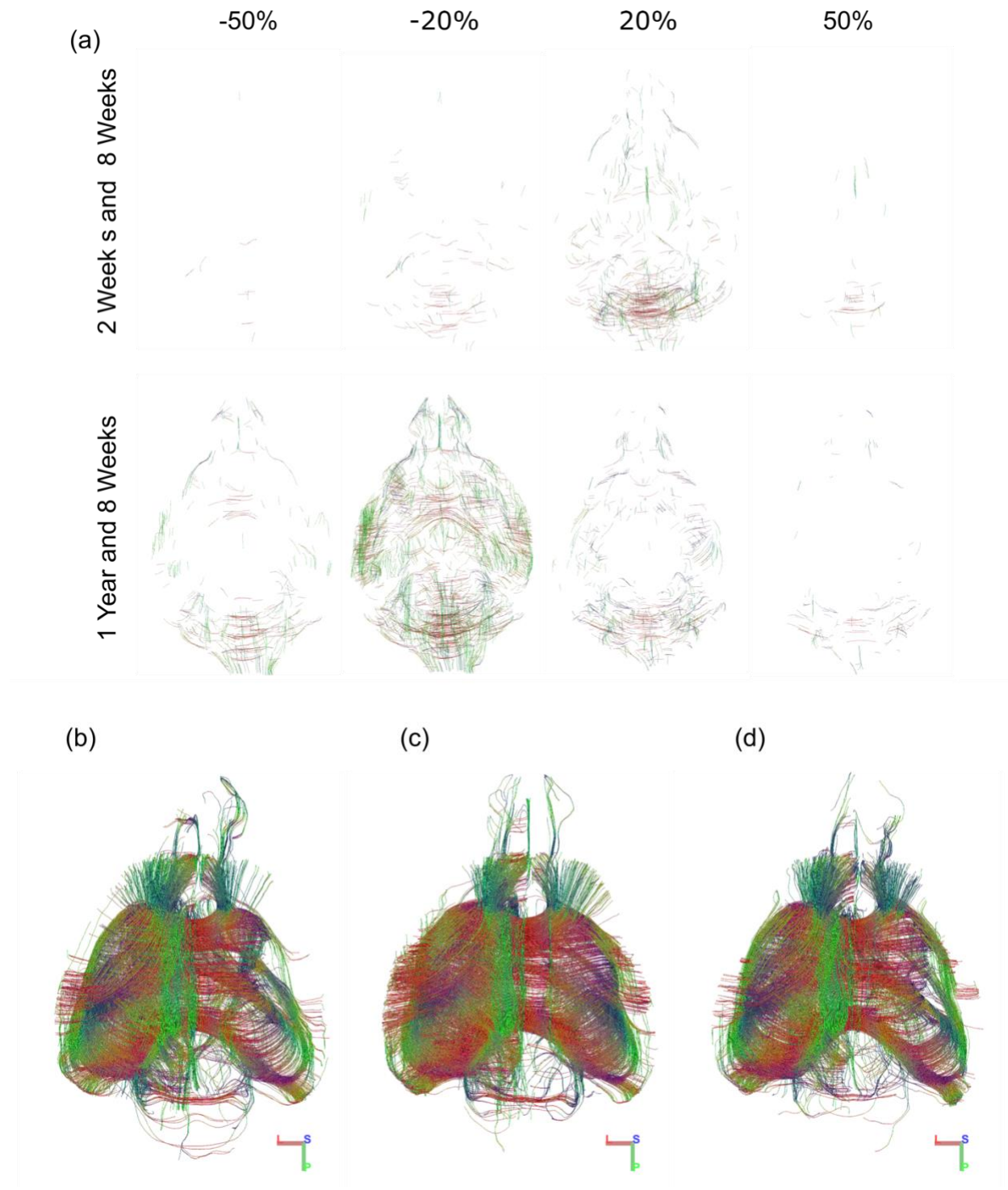

Figure 7: A comparison of (a) differential tractography measuring the change of QA at 2 weeks and 1 year using 8 weeks as a basis and tracking parsed to only those that pass

through the CC at (b) 2 weeks (c) 8 week, and (d) 1 year post fixation. All images are produced with specimen 180423-2:1 which had the largest difference in track length and number at ±20% and ±50% in the differential QA data.

DISCUSSION

Through our analysis of fractional anisotropy, axial diffusivity, radial diffusivity, and mean diffusivity, we have sought to understand the changes in diffusion tensor scalars due to fixation over time. Similar to the literature results for T1 and T2 changes post-fixation [21] [22,51] [20], we found that some of the DTI scalar values varied significantly during the first two weeks following perfusion fixation, but eventually stabilized to a nominally constant value. Changes in the diffusion metrics were statistically insignificant between 2 and 8 weeks providing a stable period. Specimens scanned during this window can be compared without confound from fixation. Conversely the scalar DTI metrics in multiple gray and white matter regions changed significantly over the first two weeks of fixation suggesting our standard operating procedure should include a stabilization period of at least 2 weeks prior to scanning.

There are statistically detectable differences in the scalar metrics between the stable period and 1 year. But the DTI parameters at one-year post fixation have changed only marginally from the measurements during the stable period. The largest change was in the AD at 13% with an average change of 7% in AD across all the structures measured. This suggests that specimens may be reexamined for a longer period without serious confound to the study.

Volume of the 12 structures was statistically unchanged during the 2 to 8 week stable time period. Volume change out to one year post fixation was ~ 3% over all the regions. De Guzman et al have performed a nicely detailed study of volume changes using protocols similar to ours (i.e. perfusion fixation, 1-5 day formalin immersion, long term storage in saline or water)[52]. The period of observation (to 150 days) is comparable to ours. They found a rapid change in volume (3.5%/day) when tissue is stored in formalin and ~3%/month when tissue is stored in PBS. Our data is consistent with theirs though the time points of comparison are not identical. De Guzman showed volume changes at 5 months of white matter ranging between ±4% and grey matter between -4% and +5%. For our data, the maximum volume change in any volume at 1 year is 6% and an average change of 3% over all regions. While small, these changes are still larger than the whole brain volume change, which increases to a maximum of 2% volume change from 2 weeks to one year. But it bears repeating that during the stable period (2-8 weeks) there were no statistically discernable changes in volume in the 12 structures surveyed. In a recent publication, we measured the coefficient of variation between left and right hemispheric structures in four different strains of mice as a measure of technical error [9]. That coefficient of variation was <5% for structures with volumes > 1 mm^3. This variation is probably an indication of the reproducibility of the registration pipeline.

The Omni-MANOVA results further are consistent with changes in the scalar metrics during the first two weeks and stable time period. The points associated with the early (2/3 day and 1 week) time period stand out as particularly different than the other groups. Figure 6a maps all the data into a common space. In this mapping, there are 160 of 322 regions with connectivity that differs. In Figure 6b, the 2/3 day and 1 week data have been removed and there are only 4 regions (all cortical) that differ with p values that are significant. The connectomes are remarkably stable from 2 weeks out to one year. Similar to the conclusion derived from the parametric analysis of the scalar DTI metrics, the one year connectome data is similar enough to the stable period that the one year time point can be considered in practice invariant from the stable time period.

The differential tractography results also confirms the stability of the tissue over the 2-8 week period with minor changes between 8 weeks and 1 year. There is a suggestion of increased anisotropy (+20%) particularly in the cerebellum in comparison to the 2 and 8 week data but the mean track length and number (0.91 mm, 570 tracks) in the difference image vs mean track length and number (5 mm, 21516 racks) in the base images suggest this is probably an edge effect. The mean track length in the 8 week and 1 year difference image is also small (0.97mm @ +20%) suggesting an edge effect. But there is an increased in number and mean length at -20% from 0.82 mm with 111 tracks @ -20% to 1.18 mm with 2490 tracks when comparing 8 weeks and 1 year. This is consistent with the bulk scalar metrics in which the AD at one year was statistically different from the stable period in eight of the twelve regions (Figure S3).

CONCLUSION

Numerous researchers have studied the effects of autolysis, fixation (post mortem interval after death, length and type of fixative and storage (in fixative, water, saline) on the diffusion properties of tissue [19,24,25,27,30,53-55]. Our *very specific protocol for quantitative connectomics in the rodent brain* has been constructed with the results of these studies in mind. For example, it is clear that fixation is essential to limit autolysis. Perfusion fixation is considerably more effective than immersion. Extended fixation (>24 hrs) causes continuing reduction in T2 (and therefore reduced signal to noise). Storage in buffered saline is desirable to minimize tissue swelling/changing. And active staining [16,56], i.e. perfusion with a contrast agent is essential for microscopic imaging. Finally, our application differs considerably from much of the previous literature. We are scanning the mouse brain. Perfusion fixation is easily performed which is generally not possible in the human studies. We are doing so with spatial resolution (45 um^3 i.e 91 pl). The spatial resolution of DTI in clinical MRI is generally >1 mm^3, i.e. a difference of nearly 11,000 X in voxel volume.

Quantitative connectomics is now being applied in both clinical and basic science settings. It provides remarkable insight into brain structure and health. But is has limitations. Acquisition protocols differ widely in spatial resolution, angular sampling, gradient weighting, and processing algorithms. Maier-Hein and a consortium of many of the leaders in the clinical use of connectomes raised significant concerns over false

positive connections even with optimized clinical protocols and pipelines [57]. And as noted above, comparison of clinical protocols with methods for the mouse brain is fraught. Retroviral tracers continue to be the gold standard for quantitative differentiation of afferent and efferent connections in the mouse brain [58]. Yet diffusion derived connectomes at microscopic resolution provide intriguing applications that cannot be done with tracers [9,13,59]. But like any new methodology, the protocols must be optimized and the limits understood. We have provided a carefully constructed comparison of our protocol with retroviral tracers [60]. We have optimized that protocol for routine use [34]. And in this work, we have begun to define how to use it with confidence.

CONFLICT OF INTEREST

The authors declare no conflict of interest in the publishing of this work.

ACKNOWLEDGEMENTS

ABBREVIATIONS AND UNITS

Selected Regions of Interest (ROI):
AC - anterior commissure
CC - corpus callosum
CG - cingulum
CP - cerebral peduncle
FI - fimbria
IC - internal capsule
ICP - inferior cerebral peduncle
MCP - middle cerebral peduncle
RN - raphe nucleus
C - cingulate cortex
S1 - primary somatosensory cortex
V - primary visual cortex

DTI (Diffusion Tensor Imaging) Measure Abbreviations:
FA - fractional anisotropy [unitless]
AD - axial diffusivity [mm^3/s]
RD - radial diffusivity [mm^3/s]
MD - mean diffusivity [mm^3/s]

REFERENCES

[Dataset] Jaclyn Xiao, Kathryn J. Hornburg, Gary Cofer, James J. Cook, Yi Qi, and G Allan Johnson; 2020; "Stained Time-course Study Representative Data"; VoxPort; https://civmvoxport.vm.duke.edu/voxbase/studyhome.php?studyid=746

1. Pecheva D, Kelly C, Kimpton J, et al. Recent advances in diffusion neuroimaging: applications in the developing preterm brain. *F1000Res.* 2018;7.
2. Fornito A, Zalesky A, Breakspear M. The connectomics of brain disorders. *Nat Rev Neurosci.* 2015;16(3):159-172.
3. Fornito A, Bullmore ET. Connectomics: a new paradigm for understanding brain disease. *Eur Neuropsychopharmacol.* 2015;25(5):733-748.
4. Sporns O. The human connectome: a complex network. *Ann N Y Acad Sci.* 2011;1224:109-125.
5. Thompson PM, Andreassen OA, Arias-Vasquez A, et al. ENIGMA and the individual: Predicting factors that affect the brain in 35 countries worldwide. *Neuroimage.* 2017;145(Pt B):389-408.
6. Van Essen DC, Smith SM, Barch DM, et al. The WU-Minn Human Connectome Project: an overview. *Neuroimage.* 2013;80:62-79.
7. Chauhan G, Adams HHH, Bis JC, et al. Association of Alzheimer's disease GWAS loci with MRI markers of brain aging. *Neurobiol Aging.* 2015;36(4):1765 e1767-1765 e1716.
8. Nir TM, Jahanshad N, Villalon-Reina JE, et al. Effectiveness of regional DTI measures in distinguishing Alzheimer's disease, MCI, and normal aging. *Neuroimage Clin.* 2013;3:180-195.
9. Wang NA, RJ; Ashbrook, DG; Gopalakrishnan, V; Park, Y; Priebe, CE; Laoprasert, R; Vogelstein, JT; Williams, RW; Johnson, GA. Variability and Heritability of Mouse Brain Structure: Microscopic MRI Atlases and Connectomes for Diverse Strains. *Neuroimage.* 2020;222.
10. Boretius S, Escher A, Dallenga T, et al. Assessment of lesion pathology in a new animal model of MS by multiparametric MRI and DTI. *Neuroimage.* 2012;59(3):2678-2688.
11. Johnson GA, Benveniste H, Black RD, Hedlund LW, Maronpot RR, Smith BR. Histology by magnetic resonance microscopy. *Magnetic Resonance Quarterly.* 1993;9(1):1-30.
12. Johnson GA, Hedlund LW, Inventors; Duke University, assignee. Three-dimensional morphology by magnetic resonance imaging. US patent 6,023,1622000.
13. Johnson GA, Wang N, Anderson RJ, et al. Whole mouse brain connectomics. *Journal of Comparative Neurology.* 2018;12:26.
14. D'Arceuil H, de Crespigny A. The effects of brain tissue decomposition on diffusion tensor imaging and tractography. *Neuroimage.* 2007;36(1):64-68.
15. Fishbein KW, Gluzband YA, Kaku M, et al. Effects of formalin fixation and collagen cross-linking on T2 and magnetization transfer in bovine nasal cartilage. *Magn Reson Med.* 2007;57(6):1000-1011.
16. Johnson GA, Cofer GP, Fubara B, Gewalt SL, Hedlund LW, Maronpot RR. Magnetic resonance histology for morphologic phenotyping. *Journal of Magnetic Resonance Imaging.* 2002;16(4):423-429.
17. Petiet A, Johnson GA. Active staining of mouse embryos for magnetic resonance microscopy. *Methods Mol Biol.* 2010;611:141-149.

18. Puchtler H, Meloan SN. On the chemistry of formaldehyde fixation and its effects on immunohistochemical reactions. *Histochemistry.* 1985;82(3):201-204.

19. Shepherd TM, Flint JJ, Thelwall PE, et al. Postmortem interval alters the water relaxation and diffusion properties of rat nervous tissue--implications for MRI studies of human autopsy samples. *Neuroimage.* 2009;44(3):820-826.

20. Dawe RJ, Bennett DA, Schneider JA, Vasireddi SK, Arfanakis K. Postmortem MRI of human brain hemispheres: T2 relaxation times during formaldehyde fixation. *Magn Reson Med.* 2009;61(4):810-818.

21. Thickman DI, Kundel HL, Wolf G. Nuclear magnetic resonance characteristics of fresh and fixed tissue: the effect of elapsed time. *Radiology.* 1983;148(1):183-185.

22. Yong-Hing CJ, Obenaus A, Stryker R, Tong K, Sarty GE. Magnetic resonance imaging and mathematical modeling of progressive formalin fixation of the human brain. *Magn Reson Med.* 2005;54(2):324-332.

23. Haga Y, Hata J, Uematsu A, et al. MR Imaging Properties of ex vivo Common Marmoset Brain after Formaldehyde Fixation. *Magn Reson Med Sci.* 2019;18(4):253-259.

24. Sun SW, Neil JJ, Song SK. Relative indices of water diffusion anisotropy are equivalent in live and formalin-fixed mouse brains. *Magn Reson Med.* 2003;50(4):743-748.

25. Sun SW, Neil JJ, Liang HF, et al. Formalin fixation alters water diffusion coefficient magnitude but not anisotropy in infarcted brain. *Magn Reson Med.* 2005;53(6):1447-1451.

26. Guilfoyle DN, Helpern JA, Lim KO. Diffusion tensor imaging in fixed brain tissue at 7.0 T. *NMR Biomed.* 2003;16(2):77-81.

27. Dyrby TB, Baare WF, Alexander DC, Jelsing J, Garde E, Sogaard LV. An ex vivo imaging pipeline for producing high-quality and high-resolution diffusion-weighted imaging datasets. *Hum Brain Mapp.* 2011;32(4):544-563.

28. Miller KL, Stagg CJ, Douaud G, et al. Diffusion imaging of whole, post-mortem human brains on a clinical MRI scanner. *Neuroimage.* 2011;57(1):167-181.

29. Pattany PM, Puckett WR, Klose KJ, et al. High-resolution diffusion-weighted MR of fresh and fixed cat spinal cords: evaluation of diffusion coefficients and anisotropy. *AJNR Am J Neuroradiol.* 1997;18(6):1049-1056.

30. Shatil AS, Uddin MN, Matsuda KM, Figley CR. Quantitative Ex Vivo MRI Changes due to Progressive Formalin Fixation in Whole Human Brain Specimens: Longitudinal Characterization of Diffusion, Relaxometry, and Myelin Water Fraction Measurements at 3T. *Front Med (Lausanne).* 2018;5:31.

31. Roebroeck A, Miller KL, Aggarwal M. Ex vivo diffusion MRI of the human brain: Technical challenges and recent advances. *NMR Biomed.* 2019;32(4):e3941.

32. Blamire AM, Rowe JG, Styles P, McDonald B. Optimising imaging parameters for post mortem MR imaging of the human brain. *Acta Radiol.* 1999;40(6):593-597.

33. Stejskal EO, Tanner JE. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *Journal of Chemical Physics.* 1965;42(1):288.

34. Wang N, Anderson RJ, Badea A, et al. Whole mouse brain structural connectomics using magnetic resonance histology. *Brain Structure and Function.* 2018;223(9):4323-4335.

35. Anderson RJ, Cook JJ, Delpratt N, et al. Small Animal Multivariate Brain Analysis (SAMBA) – a High Throughput Pipeline with a Validation Framework. *Neuroinformatics.* 2019;17(3):451-472.

36. Johnson GA, Badea A, Brandenburg J, et al. Waxholm space: an image-based reference for coordinating mouse brain research. *Neuroimage.* 2010;53(2):365-372.

37. Van Essen DC, Glasser MF. Parcellating Cerebral Cortex: How Invasive Animal Studies Inform Noninvasive Mapmaking in Humans. *Neuron.* 2018;99(4):640-663.

38. Van Essen DC, Donahue CJ, Coalson TS, Kennedy H, Hayashi T, Glasser MF. Cerebral cortical folding, parcellation, and connectivity in humans, nonhuman primates, and mice. *Proceedings of the National Academy of Sciences.* 2019;116(52):26173.

39. Yeh F-C. A Tractography Software Tool for Diffusion MRI Analysis. DSI Studio Web site. http://dsi-studio.labsolver.org/ Accessed October 21, 2020.

40. Yeh F-C, Wedeen VJ, Tseng WI. Generalized Q-Sampling Imaging. *IEEE Transactions on Medical Imaging.* 2010;29(9):1626-1635.

41. Wallis W. *A beginners guide to graph theory.* Boston: Birkhauser, Inc; 2007.

42. Levin K, Athreya A, Tang M, Lyzinski V, Priebe CE. A Central Limit Theorem for an Omnibus Embedding of Multiple Random Dot Product Graphs. Paper presented at: 2017 IEEE International Conference on Data Mining Workshops (ICDMW); 18-21 Nov. 2017, 2017.

43. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological).* 1995;57(1):289-300.

44. Pike N. Using false discovery rates for multiple comparisons in ecology and evolution. *Methods in Ecology and Evolution.* 2011;2(3):278-282.

45. McDonald JH. Multiple comparisons. In: *Handbook of Biological Statistics.* 3 ed. Baltimore, Maryland: Sparky House Publishing; 2014:254-260.

46. Levin K, Athreya A, Tang M, Lyzinski V, Park Y, Priebe CE. A central limit theorem for an omnibus embedding of multiple random graphs and implications for multiscale network inference. 2019:arXiv:1705.09355. https://ui.adsabs.harvard.edu/abs/2017arXiv170509355L. Published June 26 2019. Accessed October 21, 2020.

47. Wang J. *Geometric Structure of High-Dimensional Data and Dimensionality Reduction.* Berlin Heidelberg: Springer-Verlag; 2012.

48. Yeh F-C. Differential tractography for longitudinal data. DSI Studio. Documentation and Tutorials Web site. http://dsi-studio.labsolver.org/Manual/differential-tractography/ Accessed August 20, 2020.

49. Yeh F-C, Zaydan IM, Suski VR, et al. Differential tractography as a track-based biomarker for neuronal injury. *Neuroimage.* 2019;202:116131.

50. Yeh F-C, Tseng WY. NTU-90: a high angular resolution brain atlas constructed by q-space diffeomorphic reconstruction. *Neuroimage.* 2011;58(1):91-99.

51. Tovi ME, A Measurement of T1 and T2 over time in formalin fixed human whole brain specimens. *Acta Radiol.* 1992;33:400.

52. de Guzman AE, Wong MD, Gleave JA, Nieman BJ. Variations in post-perfusion immersionfixation and storage alter MRI measurements of mouse brain morphometry. *Neuroimage.* 2016;142:687-695.

53. MacFall JR, Maki JH, Johnson GA, Hedlund LW, Cofer GP. Pre-and post-mortem diffusion coefficients in rat neural and muscle tissues. *Magnetic Resonance in Medicine.* 1991;20(1):89-99.

54. Thelwall PS, TM; Stanisz, GJ; Blackband, SJ. Effects of temperature and aldehyde fixation on tissue water diffusion properties, studied in an erythrocyte ghost tissue model. *Magn Reson Med.* 2006;56(2):282-289.

55. D'Arceuil H dCA. The effects of brain tissue decomposition on diffusion tensor imaging and tractography. *Neuroimage.* 2007;36(1):64-68.

56. Johnson GA, Cofer GP, Gewalt SL, Hedlund LW. Morphologic phenotyping with magnetic resonance microscopy: the visible mouse. *Radiology.* 2002;222(3):789-793.

57. Maier-Hein KH, Neher PF, Houde J-C, et al. The challenge of mapping the human connectome based on diffusion tractography. *Nature Communications.* 2017;8(1):1349.

58. Oh SW, Harris JA, Ng L, et al. A mesoscale connectome of the mouse brain. *Nature.* 2014;508:207.

59. Badea A, Wu W, Shuff J, et al. Identifying Vulnerable Brain Networks in Mouse Models of Genetic Risk Factors for Late Onset Alzheimer's Disease. *Frontiers in Neuroinformatics.* 2019;13(72).

60. Calabrese E, Badea A, Cofer G, Qi Y, Johnson GA. A diffusion MRI tractography connectome of the mouse brain and comparison with neuronal tracer data. *Cereb Cortex.* 2015;25(11):4628-4637.