

## Research article

# Age-related changes in structural connectivity are improved using subject-specific thresholding



Corinna M. Bauer<sup>a,\*</sup>, Lauren E. Zajac<sup>b</sup>, Bang-Bon Koo<sup>b</sup>, Ronald J. Killiany<sup>b</sup>, Lotfi B. Merabet<sup>a</sup>

<sup>a</sup> Laboratory for Visual Neuroplasticity, Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, MA, USA

<sup>b</sup> Center for Biomedical Imaging, Department of Anatomy and Neurobiology, Boston University School of Medicine, Boston, MA, USA

## HIGHLIGHTS

- Diffusion tractography using HARDI is significantly affected by quantitative anisotropy (QA) threshold values.
- Tractography of specific fasciculi can be inadvertently influenced by QA threshold.
- Network-based analyses are also susceptible to the effects of QA threshold.
- Subject-specific QA thresholds are recommended to provide meaningful results.

## ARTICLE INFO

## Article history:

Received 27 February 2017

Received in revised form 12 June 2017

Accepted 16 June 2017

Available online 21 June 2017

## Keywords:

High angular resolution diffusion imaging (HARDI)

Tractography

Quantitative anisotropy threshold

Arcuate fasciculus

Network analysis

Normal aging

## ABSTRACT

**Background:** Deterministic diffusion tractography obtained from high angular resolution diffusion imaging (HARDI) requires user-defined quantitative anisotropy (QA) thresholds. Most studies employ a common threshold across all subjects even though there is a strong degree of individual variation within groups. We sought to explore whether it would be beneficial to use individual thresholds in order to accommodate individual variance. To do this, we conducted two independent experiments.

**Method:** First, tractography of the arcuate fasciculus and network connectivity measures were examined in a sample of 14 healthy participants. Second, we assessed the effects of QA threshold on group differences in network connectivity measures between healthy young ( $n = 19$ ) and old ( $n = 14$ ) individuals.

**Results:** The results of both experiments were significantly influenced by QA threshold. Common thresholds set too high failed to produce sufficient reconstructions in most subjects, thus decreasing the likelihood of detecting meaningful group differences. On the other hand, common thresholds set too low resulted in spurious reconstructions, providing deleterious results.

**Comparison with existing methods:** Subject specific thresholds acquired using our QA threshold selection method (QATS) appeared to provide the most meaningful networks while ensuring that data from all subjects contributed to the analyses.

**Conclusions:** Together, these results support the use of a subject-specific threshold to ensure that data from all subjects are included in the analyses being conducted.

© 2017 Elsevier B.V. All rights reserved.

## 1. Introduction

Diffusion MRI can be used to assess the movement of water molecules within the brain, making it possible to infer structural properties of underlying neural tissue. Deterministic fiber tracking methods provide the ability to reconstruct models of white matter pathways obtained from diffusion MRI data. High angular resolution diffusion imaging (HARDI) and diffusion spectrum imaging (DSI) are becoming popular because of their ability to model pathways through the complicated geometry imposed by crossing, kissing, or other complex fiber orientations found in the brain (Wedeen et al., 2008). HARDI samples the diffusion MR signal at

**Abbreviations:** HARDI, high angular resolution diffusion imaging; ODF, orientation distribution function; QA, quantitative anisotropy; QATS, quantitative anisotropy threshold selection; ROI, region(s) of interest.

\* Corresponding author at: Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, MA, 02114, USA.

E-mail addresses: [Corinna.Bauer@meei.harvard.edu](mailto:Corinna.Bauer@meei.harvard.edu) (C.M. Bauer), [zajac@bu.edu](mailto:zajac@bu.edu) (L.E. Zajac), [bbkoo@bu.edu](mailto:bbkoo@bu.edu) (B.-B. Koo), [killiany@bu.edu](mailto:killiany@bu.edu) (R.J. Killiany), [Lotfi.Merabet@meei.harvard.edu](mailto:Lotfi.Merabet@meei.harvard.edu) (L.B. Merabet).

<http://dx.doi.org/10.1016/j.jneumeth.2017.06.010>

0165-0270/© 2017 Elsevier B.V. All rights reserved.

a set of points on the surface of a sphere at a high b-value and is sensitive to the microstructural differences in the underlying tissue (Nagy et al., 2013). HARDI, as a single B-shell sequence, offers image acquisition times that are amenable for scanning clinical populations (Rodrigues et al., 2013).

In HARDI, the orientation distribution function (ODF) characterizes the relative likelihood of diffusion along any given direction within a voxel and is based on the spin density function (Yeh et al., 2010). For each reconstructed fiber within a voxel, a quantitative anisotropy (QA) value is calculated. QA is defined for each peak of the spin distribution function, making this technique rather insensitive to partial voluming. Because of this, QA provides a useful index for filtering fiber populations and for defining track termination in deterministic tractography (Yeh et al., 2013).

A critical step in the reconstruction of white matter projections with deterministic tractography is choosing the appropriate threshold parameters. These include anisotropy threshold (e.g. QA, FA), termination angle, step size, smoothing, and length constraints. The anisotropy threshold plays an important role in determining when tracts end. When the anisotropy value within a voxel falls below the threshold, the tracking algorithms terminate. Thus, anisotropy threshold ultimately determines which voxels are included in a tract and the choice of this value can significantly alter fiber trajectory (Mukherjee et al., 2008). One of the biggest challenges faced is in determining the optimal QA threshold. For example, setting the value too high can deflate the number of connections found, producing false negative results (Kunimatsu et al., 2004), while setting the QA threshold too low can result in spurious connections and cause false-positive results (Seehaus et al., 2012). Despite the importance of this parameter, finding a way to objectively determine the optimal QA value has received little attention in the literature. Furthermore, because the scaling of QA is different from FA (Yeh et al., 2010), previously determined FA thresholds (Domin et al., 2014; Kunimatsu et al., 2004; Parizel et al., 2007; Seehaus et al., 2012; Taoka et al., 2009) cannot be directly applied in studies using HARDI forms of imaging.

The impact that anisotropy threshold parameters can have on the resultant tractography models has not been studied within the context of ODF-based approaches, such as HARDI. This is concerning given the growing number of studies that are using ODF-based techniques to study a wide array of disorders such as neurological conditions (Abhinav et al., 2014; Dennis et al., 2011; Muhlert et al., 2013) or normal aging and development (Dennis et al., 2013; Lee et al., 2015). Moreover, the impact that these thresholds have on the fiber tract models generated has not been examined in the context of network-based connectivity (Rubinov and Sporns, 2010), which is also becoming increasingly popular. In the realm of network-based connectivity, there is evidence that early processing steps can significantly influence the outcome of network-based models. For example, studies have investigated the influence of different parcellation schemes with respect to a variety of network properties (Cammoun et al., 2012; Cheng et al., 2012b) and weighting schemes (Cheng et al., 2012a). Further, the specific tractography algorithm used has also been shown to impact network characteristics (Bastiani et al., 2012). However, it is unknown how QA threshold influences these measures.

Most studies being conducted today employ a common anisotropy threshold across all subjects even though there is a strong degree of individual variation in anisotropy values. In doing this, there is the potential for individual subjects to offer little to the analysis because their fiber reconstruction models have few to no connections. In this study, we sought to explore whether it would be beneficial to accommodate individual variation through the use of individual thresholds. To do this, we conducted two separate experiments. First, the effect of varying QA threshold in the reconstruction and anatomical properties of the arcuate fasciculus was

examined. The arcuate fasciculus has been characterized extensively *in-vivo* using diffusion based imaging tractography (Catani and Thiebaut de Schotten, 2008) and *ex-vivo* using the Klinger dissection technique (Fernández-Miranda et al., 2015). Thus, there exists an anatomical/true “gold standard” that can be compared to the fiber reconstruction models that we generate. Second, the role of QA threshold on the network-based properties of clustering coefficient, node strength, nodal efficiency, network global efficiency, network density, and number of network hubs (Rubinov and Sporns, 2010) was investigated in two phases. The first phase involved a network-based exploration of the overall effects of QA threshold; independent of any clinical application. The second phase involved a network-based study on the effect of QA threshold on the ability to detect meaningful differences between healthy young and healthy aged participants.

## 2. Materials and methods

### 2.1. Subjects

For the first experiment, data were collected from fourteen healthy subjects (mean = 26.46 years, range = 15–44, 6 females). The second experiment consisted of a completely separate cohort of 19 young (mean = 24.9 years, range = 19–32, 8 females) and 14 aged individuals without subjective cognitive complaints (mean = 72.4 years, range = 65–84, 7 females). All subjects were free of any neurological or psychiatric disorders and had normal structural MRI scans. Written informed consent was obtained from all subjects (and/or their guardians where appropriate) prior to participation. The experimental procedures were approved by the Institutional Review Board at the Massachusetts Eye and Ear Infirmary, Boston, MA, USA or the Institutional Review Board at Boston University School of Medicine, Boston, MA as appropriate.

### 2.2. MRI acquisition parameters

All MRI scanning was performed using a 3T Philips Achieva system (Best, the Netherlands) with an eight-channel phased array head coil. In the first experiment, two T1-weighted structural images were acquired using a FFE pulse sequence (TE 3.1 ms, TR 6.8 ms, flip angle 9°, 1 × 1 × 1.2 mm voxel size) and one high angular diffusion imaging (HARDI) scan using a single shot EPI sequence (TE 73 ms, TR 17844 ms, 64 directions,  $b_{\max}$  3000 s/mm<sup>2</sup>, a single b<sub>0</sub>, 2 × 2 × 2 mm voxel size, 80 mT/m, slew rate 100 mT/m/ms, diffusion time 36.3 ms, diffusion gradient duration 19.1 ms), and field map (TE 2.3 and 4.6 ms, TR 20 ms, flip angle 10°, 3 × 3 × 3 mm voxel size). For the second experiment, data were obtained as part of the Healthy Outreach Program for the Elderly (HOPE) study using a 32 channel phased array head coil. The T1 w parameters were identical though only one scan was acquired and the single shot EPI HARDI parameters were as follows: TE 100 ms, TR 8789 ms, 64 directions,  $b_{\max}$  3000 s/mm<sup>2</sup>, a single b<sub>0</sub>, 2 × 2 × 2 mm voxel size, 40 mT/m, slew rate 200 mT/m/ms, diffusion time 50.3 ms, diffusion gradient duration 33.4 ms. Dynamic stabilization was used during all HARDI scans to adjust for B0 drift (Benner et al., 2006). Because of the differences in acquisition parameters, datasets for each experiment were assessed completely separately and the results of each analysis were independent.

### 2.3. Processing of T1-weighted data

T1-weighted data was processed using FreeSurfer 5.3.0 (<https://surfer.nmr.mgh.harvard.edu/>), which has been described in detail elsewhere (Dale et al., 1999; Fischl et al., 1999, 2002, 2004). Briefly, the T1-weighted images were skull stripping, intensity normalization, and Talairach registration. Grey matter was parcellated into 68

cortical and subcortical regions of interest (ROIs) from the Desikan atlas (Desikan et al., 2006). White matter was segmented into 68 regions that were derived from the cortical parcellations using a Voronoi diagram approach (Salat et al., 2009). A distance constraint of 2 mm into the white matter from the grey/white matter boundary was applied to avoid substantial overlap between white matter segmentations.

#### 2.4. Preprocessing of HARDI data

The HARDI data were skull-stripped using the brain extraction tool (BET) and corrected for eddy currents using Eddy-correct and from FSL 5.0.8 (FMRIB Software Library, <http://fsl.fmrib.ox.ac.uk/fsl>). Motion was assessed using MCFLIRT. No subject's motion exceeded 2 mm, which is less than the voxel size. For the first HARDI experiment, the average absolute motion was 1.13 mm and the average relative motion was 0.66 mm. For the second HARDI aging experiment, the average absolute motion was 1.32 mm (aging = 1.44 mm, young = 1.23 mm,  $t = -2.06$ ,  $p = 0.048$ ) and the average relative motion was 0.805 mm (aging = 0.89 mm, young = 0.75 mm,  $t = -2.27$ ,  $p = 0.03$ ). Some motion correction occurs as part of the eddy-current correction process. For both studies, the ODF in each voxel was reconstructed in DSI-Studio (<http://dsi-studio.labsolver.org>) using generalized q-sampling imaging (GQI) (Yeh et al., 2010) with a diffusion sampling length ratio of 1.25, ODF sharpening via decomposition (Yeh and Tseng, 2013), a decomposition fraction of 0.04, and a maximum fiber population of 8. Three fibers per voxel were resolved with an 8-fold ODF tessellation.

#### 2.5. Parcellation of HARDI data

Boundary based registration (BBR) in FreeSurfer (Greve and Fischl, 2009) was used to align the HARDI data to the corresponding T1-weighted structural image. Each white matter, cortical grey matter, and subcortical grey matter ROI from the T1-weighted parcellation scheme was reverse transformed into subject-specific HARDI space. This resulted in 136 regions (68 grey and 68 white matter with 34 in each hemisphere) that served as seed (white matter) and end target (grey matter) points for tractography. This choice ensured that reconstructions ended in a grey matter region of interest (Zajac et al., 2017). This also allowed us to examine the effects of QA threshold on both grey and white matter regions.

#### 2.6. Tractography and connectivity matrix generation

An in-house MATLAB script was used to generate connectivity matrices between each of the 68 grey and 68 white matter regions. Tractography parameters were as follows: termination angle = 45°, step size = 0.5 mm, min/max length = 5/300 mm, 100,000 seeds, random fiber orientation, and Gaussian radial interpolation. Tracking was performed using the Runge-Kutta method (Basser et al., 2000). Six different QA threshold values were tested in the initial experiment. The first, upper-end threshold was QA = 0.1, followed by two-midrange thresholds of QA = 0.07263 and QA = 0.062. The fourth threshold used was QA = 0.025 as a low-end constant. The fifth threshold was individually-derived using  $0.6 \cdot (\text{Otsu's threshold})$  (Otsu, 1979) (as utilized in DSI-Studio). For the purposes of this study, this is referred to as the DSI-Studio method. The final threshold was determined using an in-house tool, the Quantitative Anisotropy Threshold Selection Tool (QATS), to identify a QA threshold specific to each subject based on the percentage of suprathreshold voxels remaining within each ROI. The number of suprathreshold voxels within each ROI (as a percentage of the total number of voxels) was determined for a variety of QA thresholds ranging from 0.005 to 0.3 in increments of 0.05. For each subject,

the QA threshold value that was chosen was the one which kept on average 70% of the available voxels within each grey matter ROI as well as approximately 90% or more of the voxels within each white matter ROI for each subject. Thus, the threshold was based on ODF coverage (i.e. the percentage of suprathreshold voxels within each ROI whose ODF is included in the search algorithm) of both the white and grey matter ROIs. Based on the results of the initial experiment, the second experiment only examined high- (0.1), low- (0.025), and mid-range (0.062) QA thresholds, along with subject-specific QATS thresholds (using the same guidelines as above) and the average QATS (0.0455). Finally, all analyses of young and old subjects were also run at the average QATS threshold of 0.0455.

#### 2.7. Experiment 1

##### 2.7.1. Parcellation and MNI transformation of the arcuate fasciculus

The arcuate fasciculus was defined using start and end seeds of the pars opercularis, pars triangularis, middle temporal, superior temporal, and transverse temporal gyri (Catani and Thiebaut de Schotten, 2008; Fernández-Miranda et al., 2015; Kamali et al., 2014; Makris et al., 2005). For simplicity, only the left arcuate fasciculus was examined as it has been shown that this pathway has significant lateralization favoring the left hemisphere (Fernández-Miranda et al., 2015).

Using FLIRT linear registration, for every subject, the left arcuate fasciculus at each QA threshold was transformed into MNI space (152, 2 mm voxels) using the corresponding T1w data as a high-resolution reference image. Once in MNI space, the reconstructions were added together as a weighted measure of tract density across subjects. Total fiber counts per voxel of the arcuate fasciculus were determined and sorted into eleven 500-fiber bins, ranging from 0 to 5500 fibers per voxel, providing the number of voxels containing  $n$  fibers across all subjects. To examine the spatial overlap of the left arcuate fasciculus between subjects as a function of QA threshold, reconstructions were binarized and summed. Each voxel in these volumes was then sorted and placed into one of fourteen bins, representing the total number of subjects. This was repeated for each QA threshold.

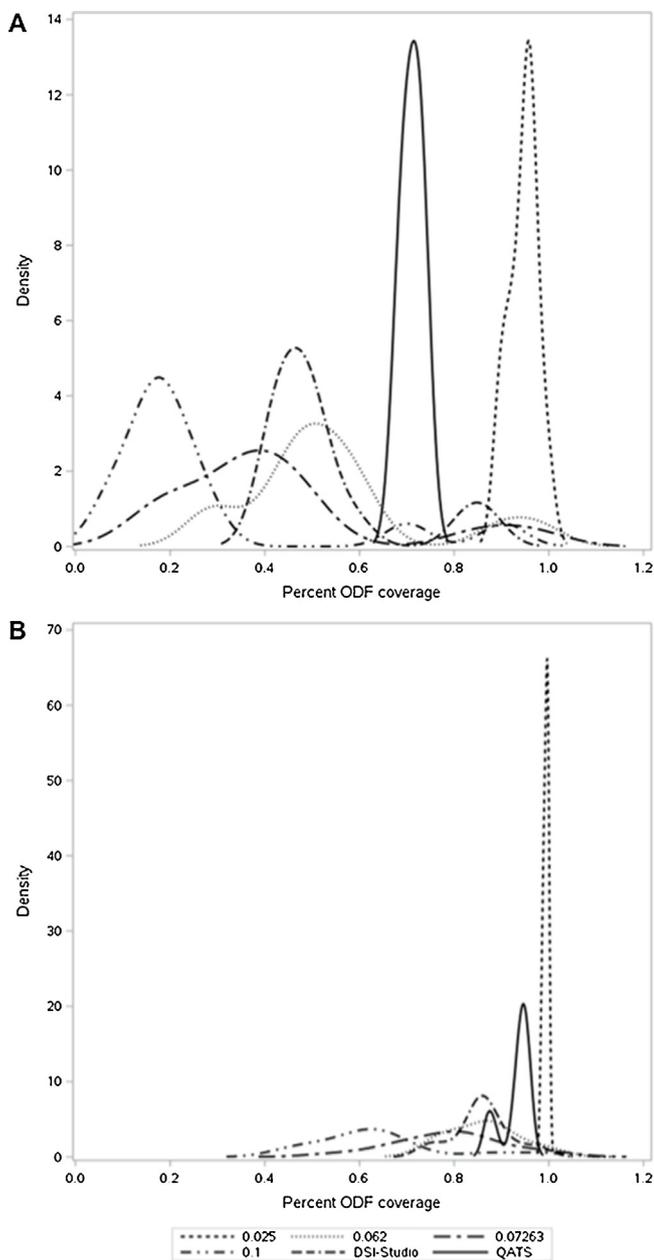
##### 2.7.2. Statistical analysis

The fraction of voxels within each ROI covered by the ODF was compared as a function of QA threshold using one-way repeated measures ANOVA. The variance of ODF coverage across the ROIs of each subject was compared using a series of F-tests, while the range of ODF coverage was compared qualitatively.

#### 2.8. Experiment 2

##### 2.8.1. Phase 1: effect of QA threshold on whole brain connectivity metrics within a single group

The effect of QA threshold on network reconstructions of the whole cortex was examined using network-based measures. The same in-house MATLAB script referenced in Section 2.4 was used to generate matrices for each of the QA thresholds tested (0.1, 0.07263, 0.062, 0.025, DSI-Studio, and QATS). These were then made symmetrical about the diagonal axis according to the maximum value. Clustering coefficient, node strength, network density, and network efficiency were calculated using Brain Connectivity Toolbox (Version 2015/01/25) (Rubinov and Sporns, 2010). Matrices were not normalized and, where possible, weighted undirected functions were used. Connectivity matrices were assessed for QA-related differences in clustering coefficient, node strength, network density, and network efficiency using one-way repeated measures ANOVA. Significance was determined based on a  $p$ -value of 0.05.



**Fig. 1.** A) Distribution of the mean grey matter ODF voxel coverage as a function of QA threshold. B) Distribution of mean white matter ODF voxel coverage as a function of QA threshold. Notice that for both GM and WM ROIs, the coverage variability is much larger for thresholds of 0.1, 0.07263, 0.062, and DSI-Studio compared to either QATS or 0.025. Also, the peak coverage occurs at a larger percentage for QATS and 0.025 compared to other thresholds.

### 2.8.2. Phase 2: age-related changes in network connectivity measures

To demonstrate how the choice of QA threshold can influence between group results, we applied five QA thresholds to a cohort of young and aged individuals. As previously mentioned, the QA thresholds examined in this experiment were 0.1, 0.025, 0.062, QATS, and average QATS (0.0455). After preprocessing and whole brain tractography (as described above), measures of connectivity were calculated using Brain Connectivity Toolbox (Version 2015/01/25). Significant differences in node strength, clustering coefficient, nodal efficiency, network density, and network global efficiency were assessed using a repeated-measures ANOVA followed by Tukey's post-hoc analysis in SAS (University Edition). The number of hubs based on nodal efficiency were determined

for young and old groups at each QA threshold, whereby a hub was defined where local efficiency  $\geq$  mean nodal efficiency + the standard deviation (i.e. global efficiency + S.D.) (Lo et al., 2010). In addition, a series of *t*-tests were performed at each QA threshold to determine how threshold influenced age-related changes in connectivity. To assess for overall differences in connectivity, two sample *t*-tests were performed at each node. These *p*-values were then corrected for multiple comparisons using false discovery rate (FDR) (Benjamini and Hochberg, 1995). Only the FDR corrected results are presented.

## 3. Results

### 3.1. Experiment 1

#### 3.1.1. Effects of QA threshold on ODF coverage within ROIs

QA threshold influenced the overall percentage of voxels included in the tractography search algorithm (Table 1). There was an overall significant effect of QA threshold on mean grey matter ROI ODF coverage ( $F(5,65) = 61.64$ ,  $p < 0.0001$ ), with no overall change in the within-subject variance of grey matter ODF coverage (e.g. the consistency of ODF coverage across all grey matter ROIs in a single subject,  $F(5,65) = 0.9$ ,  $p = 0.49$ ). Individual *F*-tests for between-subjects variance of grey matter coverage between pairs of QA thresholds revealed that the QATS and 0.025 provided significantly lower variance than other thresholds. However, between-subjects variance with QATS and 0.025 were not significantly different (Supplemental Table 1). Qualitatively, QATS showed the lowest standard deviation and range for grey matter ODF coverage across subjects (Table 1), as well as low variability between subjects as indicated by the narrow full-width-at-half-maximum in Fig. 1a.

There was also an overall significant effect of QA threshold on mean white matter ODF coverage ( $F(5,65) = 47.92$ ,  $p < 0.0001$ ). An overall significant effect of QA threshold on within-subjects variance of white matter ODF coverage was observed (e.g. the consistency of ODF coverage across all white matter ROIs in a single subject,  $F(5,65) = 47.01$ ,  $p < 0.0001$ ). Both QATS and 0.025 thresholds provided significantly decreased amounts of variance compared to all other thresholds, but they were not significantly different from one another (Table 1). Individual *F*-tests for between-subjects variance of white matter ODF coverage between pairs of QA thresholds revealed that a threshold of 0.025 provided significantly less variance than other thresholds, followed by QATS (Supplemental Table 1). Qualitatively, the range and standard deviation of white matter ODF coverage across subjects were lowest using a QA threshold of 0.025, followed by QATS (Table 1). The variability in white matter ODF coverage between subjects across the QA thresholds tested is demonstrated in Fig. 1b.

#### 3.1.2. Arcuate fasciculus reconstruction

Qualitatively, the arcuate fasciculus reconstructions were visually compared to anatomical dissections that have been previously published (Catani and Thiebaut de Schotten, 2008; Fernández-Miranda et al., 2015). For three subjects, no fibers from the arcuate fasciculus could be reconstructed at a QA threshold of 0.1 (e.g. subjects 2 and 4 in Fig. 2), and the presence of spurious connections was observed in a number of subjects at a threshold of 0.025 (e.g. subjects 1 and 3 in Fig. 2). QATS consistently generated reconstructions in all subjects that resembled those in the literature.

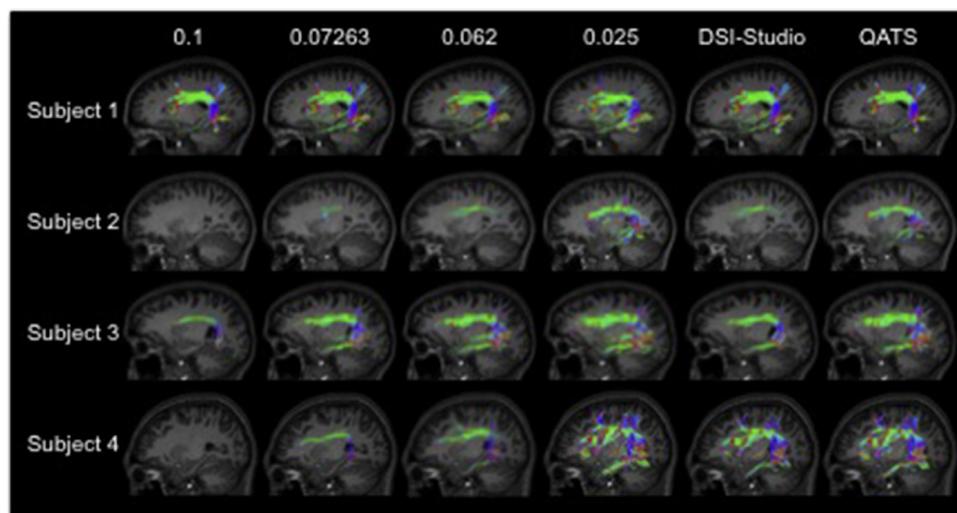
#### 3.1.3. Conjunction data for arcuate fasciculus

The maximum fiber density per voxel was greatest for QATS and 0.025, at  $>5000$  fibers. This was approximately double that observed for DSI-Studio, 0.062, and 0.07263 (Table 2). The total number of voxels present in the reconstructed group arcuate ranged between

**Table 1**

ODF coverage data for each QA threshold tested. Where applicable, a one-way repeated measures ANOVA was utilized with Tukey's posthoc analysis. For Tukey's results, cells with the same letter are not significantly different. \* Significance for between subjects variance was calculated using a series of F-tests and Tukey's style grouping applied afterwards. Full F-test results are found in Supplemental material

	0.025	0.062	0.07263	0.1	QATS	DSI-Studio	F-value	Pr > F
mean GM coverage	0.95 (a)	0.54 (c)	0.42 (c)	0.26 (d)	0.71 (b)	0.53 (c)	61.64	<0.0001
sd GM coverage	0.030	0.19	0.23	0.24	0.021	0.14		
between subjects variance GM coverage*	0.00089 (b)	0.038 (a)	0.053 (a)	0.057 (a)	0.00043 (b)	0.021 (a)		
average within subject variance	0.0045 (a)	0.025 (a)	0.023 (a)	0.048 (a)	0.034 (a)	0.026 (a)	0.9	0.4885
range GM coverage	0.10	0.69	0.79	0.83	0.063	0.46		
min GM coverage	0.89	0.28	0.18	0.075	0.68	0.41		
max GM coverage	1.00	0.98	0.96	0.90	0.74	0.87		
mean WM coverage	0.99 (a)	0.86 (b)	0.79 (c)	0.64 (d)	0.93 (ab)	0.86 (bc)	47.92	<0.0001
sd WM coverage	0.0050	0.073	0.11	0.14	0.031	0.062		
between subjects variance WM coverage*	0.000025 (e)	0.0054 (cb)	0.013 (ba)	0.021 (a)	0.00093 (d)	0.0039 (c)		
average within subjects variance WM coverage	0.00026 (c)	0.012 (b)	0.017 (b)	0.027 (a)	0.0045 (c)	0.012 (b)	47.01	<0.0001
range WM coverage	0.016	0.24	0.44	0.53	0.089	0.24		
min WM coverage	0.98	0.76	0.55	0.46	0.87	0.74		
max WM coverage	1.00	1.00	1.00	0.99	0.96	0.98		



**Fig. 2.** Tractography of the arcuate fasciculus in four subjects at differing QA thresholds: 0.1, 0.07263, 0.062, 0.025, and individually derived thresholds using DSI-studio QATS. Starting seeds were pars opercularis and pars triangularis, while end seeds were superior, middle, and inferior temporal gyri, supramarginal gyrus, and the transverse temporal regions, as per anatomical dissection studies. Fibers which were part of the uncinate fasciculus were manually removed. Notice that the QATS threshold reconstructs the arcuate the most accurately in each subject, whereas in each of the conditions whereby the same QA threshold was used across all subjects, some subjects show robust connections, perhaps verging on inaccurate (i.e. subject 1 at 0.07263, 0.062, and 0.025) while other subjects show very minimal or no reconstruction (i.e. subject 4 at 0.07263 and 0.062), and others still show anatomically accurate reconstructions (i.e. subjects 2 and 3 at 0.07263 or 0.062). This clearly demonstrates the need to use a subject-specific QA threshold for tract tracing studies.

10914–32634, depending on the threshold chosen. Of these, 0.025 and QATS were the highest, while 0.1 and 0.07263 were the lowest. The degree of overlap across subjects was visualized in a series of binary maps, where yellow indicates a greater overlap (Fig. 3). It is evident that there was more overlap between subjects using thresholds of 0.025 and QATS. A greater number of voxels were represented across all 14 subjects using QA thresholds of 0.025 or QATS compared to the other tested thresholds (Table 3).

### 3.2. Experiment 2 – phase 1

#### 3.2.1. Network connectivity measures within a single group

Network-based measures were examined for effects of QA threshold. Network efficiency showed a statistically significant effect of QA threshold  $F(5,65) = 51.47$ ,  $p < 0.0001$ . Tukey's post-hoc analysis revealed significantly larger values of network efficiency for QATS compared to all other thresholds, except for 0.025, with which it was not significantly different. Network density also showed a statistically significant effect of QA threshold

$F(5,65) = 84.97$ ,  $p < 0.0001$ . QATS resulted in significantly larger network density values over all other thresholds, with the exception of 0.025 (Table 4).

A significant main effect of QA threshold on clustering coefficient was observed for the majority of ROIs (Supplemental Table 2). In general, the lower the QA threshold, the higher the clustering coefficient. Tukey's post-hoc analysis revealed that there was not a significant difference between QATS and 0.025 for the majority of regions, with the exception of six ROIs in which QATS provided a significantly greater clustering coefficient.

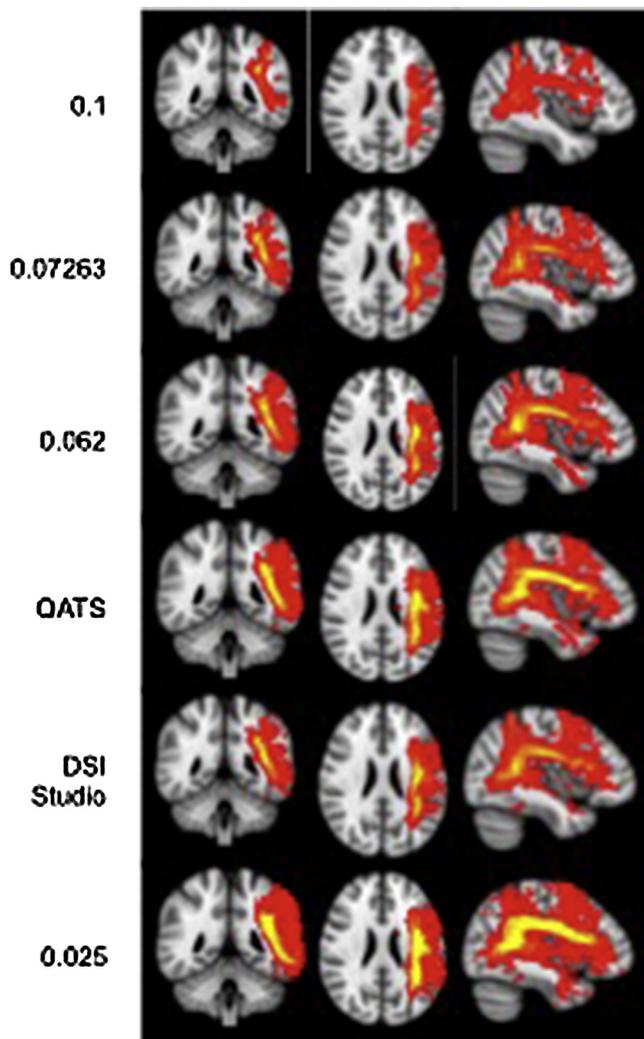
A significant main effect of QA threshold on node strength was observed for all ROIs ( $p < 0.0001$ ) (Supplemental Table 3). Tukey's post-hoc analysis revealed that node strength using a QA threshold of 0.025 and QATS were not significantly different from one another for approximately half of the ROIs within each hemisphere, while for the other half a threshold of 0.025 provided a significantly higher node strength than QATS. Node strength was the lowest for a threshold of 0.1, while it was typically not significantly different using a threshold of 0.062, 0.07263, or that provided by DSI-Studio.

**Table 2**  
Number of voxels with streamlines present within the arcuate fasciculus reconstructions at each of the six tested QA thresholds. The measures of fiber density were obtained from the weighted group sum of arcuate reconstructions for each threshold independently. The total fiber counts per voxel were calculated and sorted into eleven 500-fiber bins, ranging from 0 to 5500 fibers per voxel. The number of voxels containing  $n$  fibers across all subjects is shown.

	Number of fibers per voxel											total voxels	
	Threshold	0–500	501–1000	1001–1500	1501–2000	2001–2500	2501–3000	3001–3500	3501–4000	4001–4500	4501–5000		5001–5500
Number of voxels	0.025	29983	1425	664	275	124	82	47	28	5	1	0	32634
containing $n$ fibers	0.062	21663	829	236	66	15	3	0	0	0	0	0	22812
	0.07263	18404	639	152	16	2	0	0	0	0	0	0	19213
	0.1	10650	201	53	10	0	0	0	0	0	0	0	10914
	QATS	25255	1284	516	148	71	25	19	4	1	1	1	27325
	DSI-Studio	20305	829	263	41	4	0	0	0	0	0	0	21442

**Table 3**  
The number of voxels at which  $n$  subjects (1 through 14) have streamlines present in reconstructions of the left arcuate fasciculus. To quantify the spatial overlap, individual arcuate reconstructions were binarized and summed across all subjects. This was done independently for each QA threshold tested. Thus, voxels represented in all 14 subjects would have a value of 14, etc.

	Number of Subjects Represented within $n$ voxels														total voxels	
	Threshold	1	2	3	4	5	6	7	8	9	10	11	12	13		14
Number of voxels	0.025	9512	5799	4474	3335	2509	1717	1282	976	732	687	591	511	350	159	32634
within the group	0.062	10383	4358	2411	1553	1033	717	574	461	419	358	291	182	68	4	22812
arcuate reconstructions	0.07263	9262	3575	1925	1293	915	605	457	406	380	268	106	21	0	0	19213
	0.1	6203	2061	972	654	404	330	215	73	2	0	0	0	0	0	10914
	QATS	10668	5017	3315	2252	1496	1107	800	605	547	459	450	339	193	77	27325
	DSI-Studio	9181	4026	2316	1655	1050	738	576	518	408	412	318	178	61	5	21442



**Fig. 3.** Conjunction maps in diffusion space of the reconstructed left arcuate fasciculus of each subject at each tested QA threshold shown in radiological convention. Yellow indicates regions with the most overlap between subjects. The optimal threshold will have the highest amount of overlap and show the least amount of red, which indicates presence of tracts in fewer subjects. Using a higher threshold, such as 0.1 or 0.07263 shows low amounts of overlap. A low threshold (i.e. 0.025) shows high amount of overlap as well as high amount of individual subject variability as shown by the increased red voxels. Both DSI-Studio and QATS are subject-specific thresholds, however it is clear that QATS provides greater overlap in the arcuate between subjects. The mid-range threshold of 0.062 performs well, although QATS provides greater overlap between subjects. Min = 0, Max = 1000, slice information  $x=65$ ,  $Y=40$ ,  $Z=48$ . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

When taken together with results from these two independent experiments demonstrate that QA threshold significantly impacts not only the reconstructions of individual fasciculi, but also whole brain network connectivity measures.

### 3.3. Experiment 2 – phase 2

#### 3.3.1. Age-related changes in network connectivity measures

The second experiment determined the effect of QA threshold on identifying age-related changes in network measures of structural connectivity. Significant age-related group differences were observed for each network characteristic (node strength, clustering coefficient, network efficiency, and network density), however these differences varied substantially based on QA threshold.

When examining connectivity matrices, significant age-related increases and decreases in number of fibers reconstructed (i.e. con-

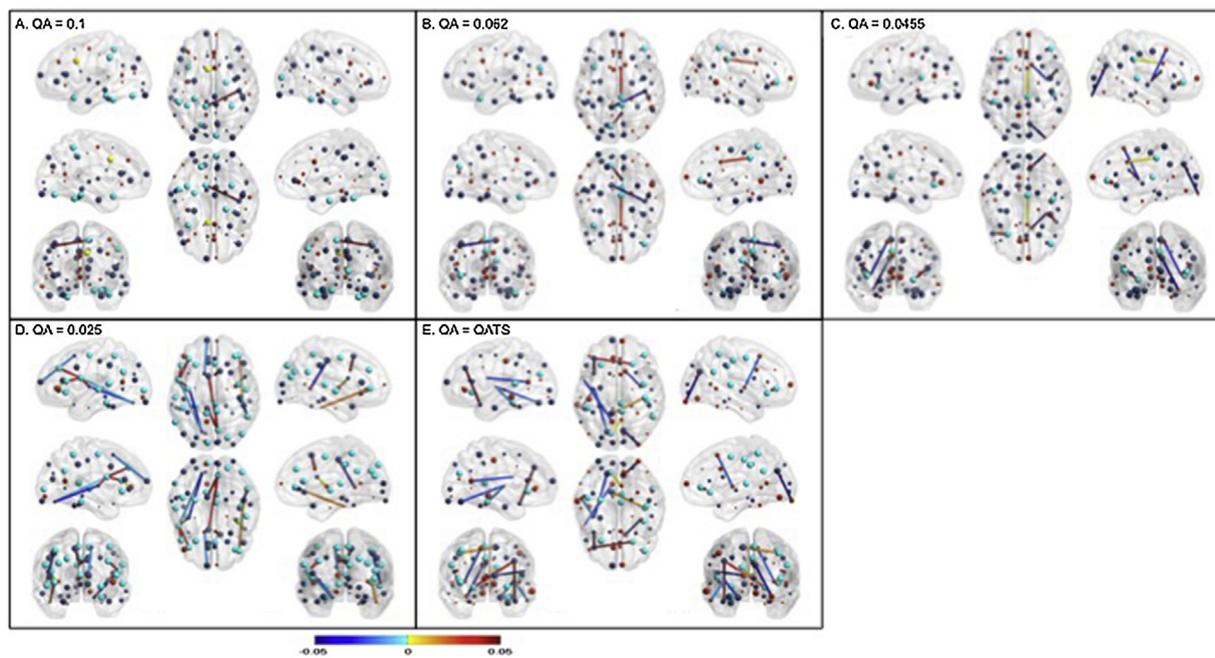
nection strength) were revealed (Figs. 4–5). The extent, direction, and location of these significant differences varied depending on which QA threshold was chosen. For example, at a QA threshold of 0.1, one connection showed a significant increase in fiber number in old versus young subjects. This same inter-hemispheric connection was significantly decreased between old and young groups using QA thresholds of 0.063 and QATS, but was not significant using QA 0.025. As the QA threshold decreases one would expect the majority of connections from higher thresholds to be replicated at lower thresholds. To this end, 100% of the significantly different connections observed using 0.1 were replicated at the lower thresholds of 0.062 and QATS, while no overlap was observed with 0.0455 or 0.025; 67% of the significantly different connections observed using 0.062 were replicated with QATS, 33% with 0.0455, and 0% with 0.025; 40% of the significantly different connections observed using 0.045 were replicated with QATS and 20% with 0.025; and lastly, only 11% of the significantly different connections observed using QATS were replicated using 0.025. In total, 100% of the significantly different connections observed using 0.1 and 0.062 were shared with at least one other threshold. Similarly, 80% of those observed using 0.0455, 56% of those observed using QATS, and 15% of the significant differences observed at 0.025 were shared with at least one other threshold (Figs. 4–5). Note that the lack of overlap at a threshold of 0.025 was not due to an increase in total connections, but rather an absence of connections observed at higher thresholds. Despite more connections showing differences at the lowest threshold tested, the low degree of overlap with other thresholds indicates that there may be a mixture of biologically based changes in connectivity and artificial changes in connectivity based solely on the effect of QA threshold on tracking algorithm performance.

Age-related differences in node strength also varied depending on which threshold was selected (Supplemental Table 4, Fig. 4). Using QATS, 12 nodes showed a significant decrease in node strength in old compared to young subjects, while a threshold of 0.1 revealed a significant decrease in nine nodes, a threshold of 0.062 or 0.0455 revealed significant decreases in five nodes, and a threshold of 0.025 revealed a significant decrease in 23 nodes. A significant increase in node strength was observed in the old group in the left caudal anterior cingulate using a threshold of 0.1; otherwise, no ROI demonstrated increased node strength between old and young groups.

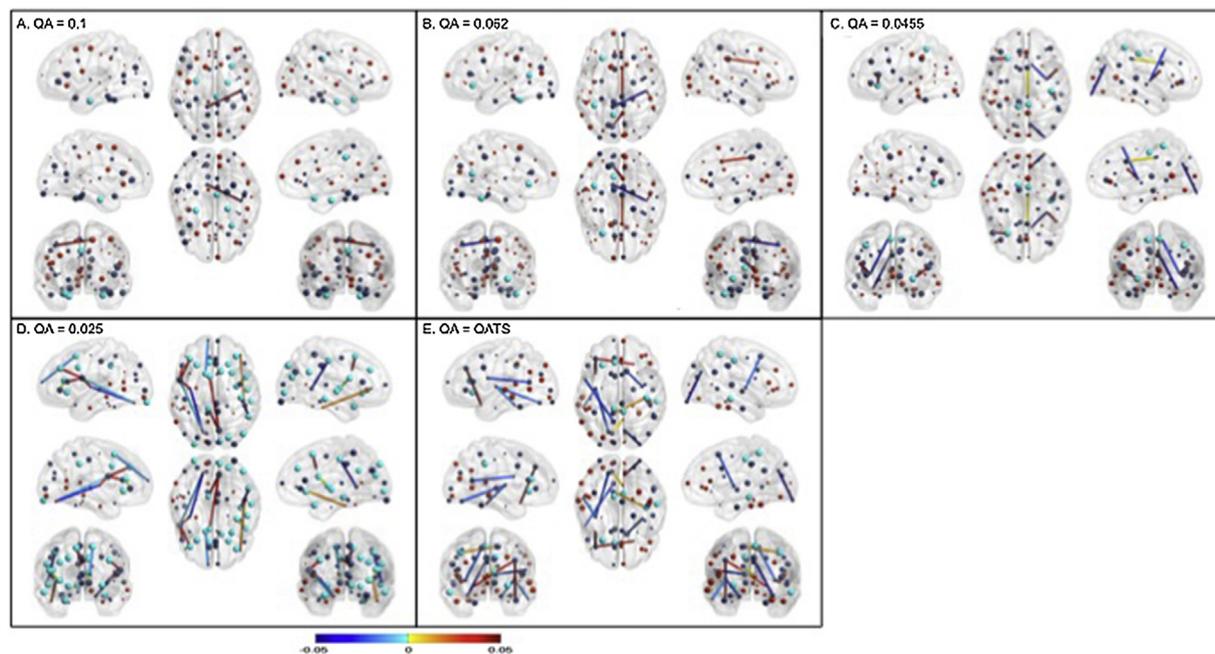
When examining age differences in clustering coefficient, the results again varied based on which QA threshold was selected (Supplemental Table 5, Fig. 5). Using QATS, six nodes showed a significant age-related decrease in clustering coefficient. This was observed in four nodes using a threshold of 0.1, three nodes using 0.062, four nodes using 0.0455, and 20 nodes using 0.025. No nodes revealed significant increases in clustering coefficient in old compared to young groups.

Both network density and global network efficiency also showed significant age-related differences which varied with QA threshold. Network efficiency was reduced in the old group compared to young subjects, but this reached significance only for QATS and QA = 0.025 (Fig. 6, Table 5). Network density was also reduced in old compared to young subject groups, but this did not reach statistical significance at any of the thresholds tested ( $p > 0.05$ ). Across thresholds, similar to the results of the first experiment, network density was the greatest for 0.025, followed by QATS, with 0.1 demonstrating the lowest density (Supplemental Fig. 1, Table 5).

Age-related differences in nodal efficiency were also influenced by choice of QA threshold (Supplemental Table 6). Specifically, using QATS, 22 nodes showed significant age-related decreases in nodal efficiency. Significant decreases were also observed in seven nodes using a threshold of 0.1, two nodes using a threshold of 0.062, eight nodes using a threshold of 0.0455, and 52 nodes using a threshold of 0.025. The number of hubs based on nodal effi-



**Fig. 4.** Ball and stick representations of network differences between old and young subject groups. Differences in node strength are represented by the spheres, whereby size and color reflect the p-value. Dark blue represents non-significant decreases in node strength, cyan represents significant decreases in node strength, red represents non-significant increases in node strength, while yellow represents significant increases in node strength in old compared to young subjects. The larger the sphere, the more significant the difference is. Differences in number of fibers between ROIs is represented by the “sticks”. These are FDR corrected for multiple comparisons. Decreases in connectivity are represented by cool colors (e.g. blue/cyan), whereas increases in connectivity in the aged group are represented by warm colors (e.g. red/yellow). This scheme is the same for each of the four conditions. Changes in node strength and connectivity using a QA threshold of A) 0.1 for all subjects, B) QA threshold of 0.062 for all subjects, C) QA threshold of 0.0455 for all subjects, D) QA threshold of 0.025 for all subjects, and E) subject-specific QATS. Note that as QA threshold decreases, more significant changes (both increases and decreases) with aging are observed. Furthermore, the connections observed at 0.1 and 0.062 are not represented using QA 0.025, suggesting that these differences at the lowest threshold tested may not be truly representative of age-related changes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

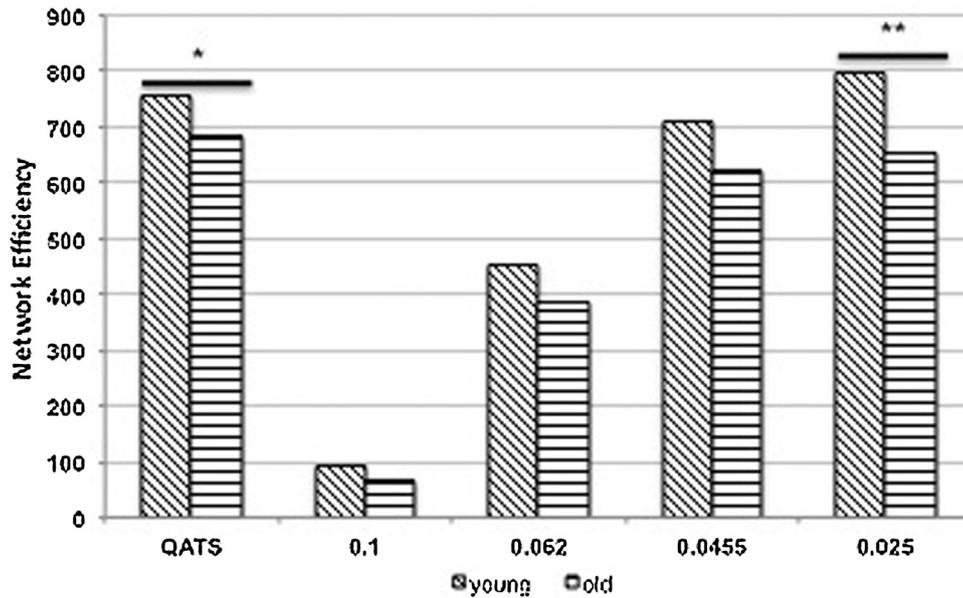


**Fig. 5.** Ball and stick representations of network differences between old and young subject groups. Differences in clustering coefficient are represented by the spheres, whereby size and color reflect the p-value. Dark blue represents non-significant decreases in clustering coefficient, cyan represents significant decreases in clustering coefficient, red represents non-significant increases in clustering coefficient, while yellow represents significant increases in clustering coefficient in old compared to young subjects. The larger the sphere, the more significant the difference is. Differences in number of fibers between ROIs is represented by the “sticks”. These are FDR corrected for multiple comparisons. Specifically, decreases in connectivity are represented by cool colors (e.g. blue/cyan), whereas increases in connectivity in the aged group are represented by warm colors (e.g. red/yellow). This scheme is the same for each of the four conditions. Changes in clustering coefficient and connectivity using a QA threshold of A) 0.1 for all subjects, B) QA threshold of 0.062 for all subjects, C) QA threshold of 0.0455 for all subjects, D) QA threshold of 0.025 for all subjects, and E) subject-specific QATS. Note that as QA threshold decreases, more significant changes (both increases and decreases) with aging are observed. Furthermore, the connections observed at 0.1 and 0.062 are not represented using QA 0.025, suggesting that these differences at the lowest threshold tested may not be truly representative of age-related changes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Table 4**

Mean network density and efficiency for connectivity matrices at each of the six QA thresholds tested. A one-way repeated measures ANOVA was utilized with Tukey's posthoc analysis. For Tukey's results, cells with the same letter are not significantly different.

Threshold	0.025	0.062	0.07263	0.1	QATS	DSI-Studio	F-value	Pr > F
Density	0.77 (a)	0.43 (c)	0.35 (c)	0.20 (d)	0.55 (b)	0.42 (c)	84.97	<0.0001
Efficiency	1696.71 (a)	1013.10 (b)	746.60 (b)	346.11 (c)	1459.49 (a)	1004.20 (b)	51.47	<0.0001



**Fig. 6.** A significant reduction in network efficiency was observed in the aged group compared to young controls only using QA thresholds of 0.025 or QATS. \*  $p < 0.05$ , \*\*  $p < 0.0001$ .

ciency was also determined for young and old groups for each QA threshold (Fig. 7). The number of hubs identified for each group and threshold were as follows: 22 young QATS, 21 old QATS; 6 young, 3 old 0.1; 4 young, 0 old 0.062; 10 young, 5 old 0.0455; 18 young, 14 old 0.025.

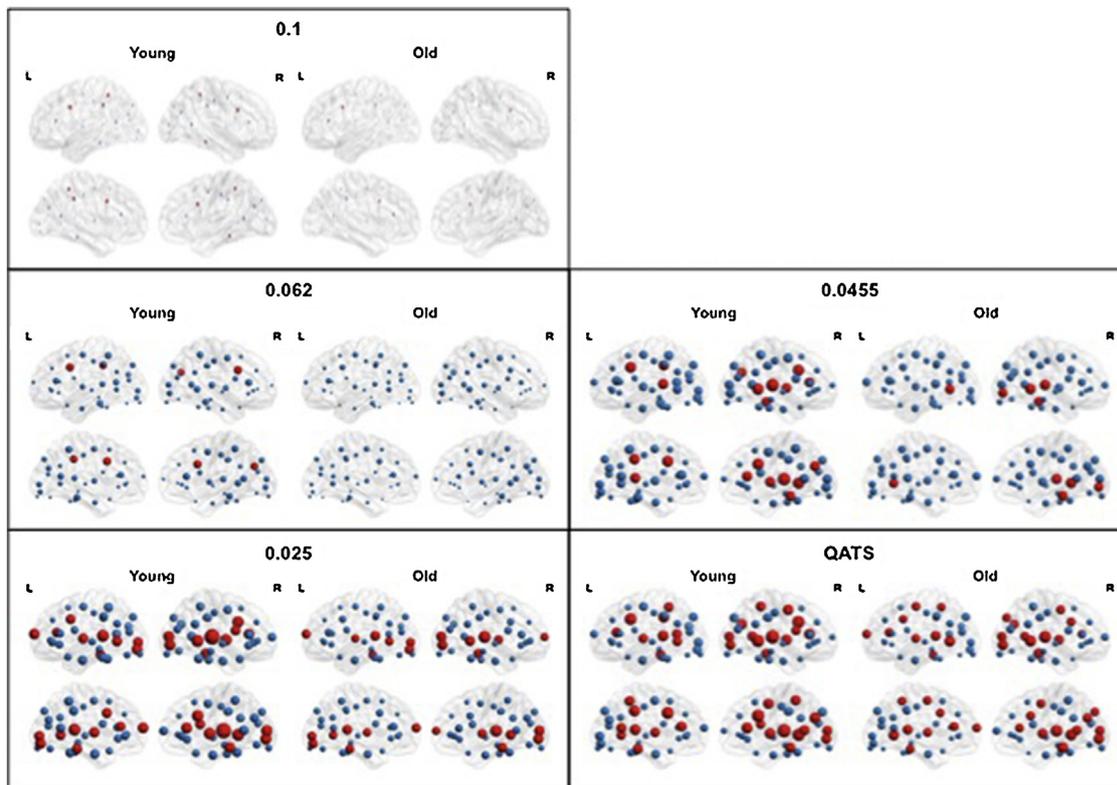
#### 4. Discussion

The choice of tracking thresholds is critical to obtaining meaningful fiber reconstructions from diffusion data (Mukherjee et al., 2008) and, as the present results indicate, the choice of threshold can alter the differences observed when comparing two subject groups. Previous studies examining these parameters have only been performed using FA measures derived from DTI data (Domin et al., 2014; Kunimatsu et al., 2004; Parizel et al., 2007; Seehaus et al., 2012; Taoka et al., 2009), but never using QA-based HARDI approaches. To our knowledge, this is the first study to systematically examine the effects of anisotropy threshold values on network-based measures.

In the first experiment reported herein, HARDI data were used to empirically test the effects of varying ODF-derived QA thresholds on the anatomical reconstruction of the arcuate fasciculus, as well as whole-brain network-based measures in two independent samples. The QA threshold determines which ODFs to include in the search for possible fiber tracts (Yeh et al., 2013). Using an individually derived QA threshold obtained with the QATS method provided significantly different results than other common or subject-specific thresholds tested. Specifically, the QATS method provided greater grey and white matter coverage compared to DSI-Studio and all the common QA values, with the exception of a threshold value of 0.025. However, qualitative assessment of recon-

structed arcuate fasciculi using a threshold of 0.025 indicated that there was an increase in spurious and anatomically infeasible fibers (e.g. subjects 1, 2, and 3 in Fig. 3), rendering this threshold value sub-optimal. This also highlights the importance of assessing the accuracy of how well reconstructions match published anatomical dissections. The inverse was also true, whereby setting the threshold too high diminished the number and quality of reconstructed fibers, resulting in a paucity of reconstructed fibers in many subjects (e.g. subjects 2 and 4 in Fig. 2). This is in agreement with previous FA tractography studies, which determined that a poor choice in anisotropy threshold can be detrimental to the tractography outcomes (Kunimatsu et al., 2004; Seehaus et al., 2012).

As a secondary analysis, QA threshold was examined in the context of network-based measures. The choice of QA threshold was shown to have a significant effect on measures of network analysis in both experiments, including network density and efficiency, and node-specific measures of clustering coefficient, nodal efficiency, and node strength. Hubs were also strikingly different in both young and old groups depending on the QA threshold chosen. The highest number of nodes was identified using QATS, with the fewest using the mid-range threshold value of 0.062. In the second experiment, the effects of QA threshold choice on age-related differences in structural connectivity network measures were determined. A high- (0.1), low- (0.025), and two mid-range (0.062 and 0.0455) common QA thresholds, as well as a subject-specific (QATS) threshold were utilized. The results indicate that when the threshold is too high, as in the case of 0.1 and 0.062, there is the potential for some subjects to not be represented in the sample because too few fibers are reconstructed between ROIs, diminishing the likelihood of observing group differences. When the threshold is too low, as was in the case for QA = 0.025, although more group dif-



**Fig. 7.** Nodal hubs for young and old subjects group by QA threshold. Nodes surpassing the threshold for hub identification are shown in red, while non-hub nodes are shown in blue. Nodes size indicates their nodal efficiency values. The threshold for nodal hubs was mean nodal efficiency + S.D. (e.g. global efficiency + S.D.) for each group and threshold. Note that nodal efficiency increases at lower thresholds, with the young subjects having overall higher nodal efficiency compared to old subjects. Furthermore, the number of hubs varies as a function of QA threshold and age. Young subjects also have more hubs than older subjects. The most hubs were observed for both young and old groups using QATS, while the fewest hubs were observed using 0.062, one of the mid-range threshold values. Similar to what was observed for other nodal measures, 0.025 shows less overlap with higher threshold, whereas every hub identified at higher threshold is present using QATS. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ferences were observed, many of these were not meaningful and likely represented mixed non-biological effects. The final threshold tested was the average QATS value. This was in between the two lowest common thresholds, but still represented a deviation from subject-specific values, both in terms of node and connection changes. Specifically, 0.0455 (the average QATS value) did not reach the same extent of group differences observed with QATS. The further downside is that this value could not be fully determined until all subjects were scanned and analyzed.

The results of this study suggest that there is a benefit to setting a QA threshold on an individual basis in order to ensure that data from each subject is included in the analysis. If a more traditional common threshold is used, the percentage of voxels included in the search algorithm in the seed and end target ROIs will differ greatly from subject to subject, making it impossible to parse out differences related to disease (or the condition of interest) from that of individual variation of ODF coverage at the chosen threshold. Individualized QA thresholds, such as those obtained with QATS, ensure relatively consistent levels of white and grey matter coverage across subjects and potentially minimize the effects of the scanning environment across sessions that may contribute to the diffusion signal (i.e. field inhomogeneity, coil sensitivity, etc.) (Abhinav et al., 2014). As a result, QATS likely provides between group results that are more meaningful (i.e. the results observed will likely be because of true group differences, not because of an artifact related to thresholding).

Although we are confident in the conclusions drawn from this study, a few limitations exist. Due to feasibility constraints, a

restricted number of common thresholds were examined. However, the current results provide support for a subject-specific threshold using QATS, as it provides more similar ODF coverage across ROIs compared to a common threshold. When the tractography algorithm is limited to similar amounts of voxels/ODFs across subjects, one can be confident that tractography outcomes are derived from the variable of interest, rather than a spurious outcome related to the specific post-processing design.

In conclusion, this study highlights the importance of choosing the optimal QA termination thresholds in ODF based tractography using HARDI data. We provide support for an individualized QA threshold, using the QATS approach. Using an individualized threshold reduced the inter-subject tractography variability for both grey and white matter seeds, which is a direct result of the number of voxels included within each seed region after applying a QA threshold, while maintaining the similarities or differences relevant to the research question of interest. This could help isolate more accurate differences between clinical and control populations and limit the ability to artificially skew the results.

## Acknowledgements

The authors wish to thank the subjects for participating in this study. We would also like to thank Yansong Zhao, Ph.D., a clinical scientist working for Philips Healthcare, for setting up the HARDI sequence on our MRI scanner and for his willingness to share his insights on diffusion imaging. This work was supported by an

**Table 5**  
Differences in network density and efficiency for aging compared to control subjects at each of the tested QA thresholds. Mean (sd) and p-values are reported. (S) represented satterthwaite test rather than student's *t*-test due to unequal variance.

Influence of Group	GLOBAL				QATS				0.1				0.062				0.0455				0.025				
	F(4155)	p	Young	Old	Young	Old	Young	Old	Young	Old	Young	Old	Young	Old	Young	Old	Young	Old	Young	Old	Young	Old	p	p	
Density	3.55	0.00613	0.35819163	0.33710648	0.4227	0.4128	0.0005	0.647	0.1074	0.0348	0.0348	0.0348	0.0348	0.0348	0.0348	0.0348	0.0348	0.0348	0.0348	0.0348	0.0348	0.0348	0.0348	0.0348	0.0348
Efficiency	12.18	0.00007	525.446049	445.934221	755	682.6	(89.67)	647	94.9754	85.69	65.5437	64.888	65.5437	64.888	65.5437	64.888	65.5437	64.888	65.5437	64.888	65.5437	64.888	65.5437	64.888	65.5437

NIH/NEI R01EY019924 (L.B.M). There are no conflicts of interest for any of the authors on this manuscript.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jneumeth.2017.06.010>.

## References

- Abhinav, K., Yeh, F.-C., El-Dokla, A., Ferrando, L.M., Chang, Y.-F., Lacomis, D., Friedlander, R.M., Fernandez-Miranda, J.C., 2014. Use of diffusion spectrum imaging in preliminary longitudinal evaluation of amyotrophic lateral sclerosis: development of an imaging biomarker. *Front. Hum. Neurosci.* 8, 270, <http://dx.doi.org/10.3389/fnhum.2014.00270>.
- Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A., 2000. *In vivo* fiber tractography using DT-MRI data. *Magn. Reson. Med.* 44, 625–632.
- Bastiani, M., Shah, N.J., Goebel, R., Roebroeck, A., 2012. Human cortical connectome reconstruction from diffusion weighted MRI: the effect of tractography algorithm. *Neuroimage* 62, 1732–1749, <http://dx.doi.org/10.1016/j.neuroimage.2012.06.002>.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* 57, 289–300.
- Cammoun, L., Gigandet, X., Meskaldji, D., Thiran, J.P., Sporns, O., Do, K.Q., Maeder, P., Meuli, R., Hagmann, P., 2012. Mapping the human connectome at multiple scales with diffusion spectrum MRI. *J. Neurosci. Methods* 203, 386–397, <http://dx.doi.org/10.1016/j.jneumeth.2011.09.031>.
- Catani, M., Thiebaut de Schotten, M., 2008. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex J. Devoted Study Nerv. Syst. Behav.* 44, 1105–1132, <http://dx.doi.org/10.1016/j.cortex.2008.05.004>.
- Cheng, H., Wang, Y., Sheng, J., Kronenberger, W.G., Mathews, V.P., Hummer, T.A., Saykin, A.J., 2012a. Characteristics and variability of structural networks derived from diffusion tensor imaging. *NeuroImage* 61, 1153–1164, <http://dx.doi.org/10.1016/j.neuroimage.2012.03.036>.
- Cheng, H., Wang, Y., Sheng, J., Sporns, O., Kronenberger, W.G., Mathews, V.P., Hummer, T.A., Saykin, A.J., 2012b. Optimization of seed density in DTI tractography for structural networks. *J. Neurosci. Methods* 203, 264–272, <http://dx.doi.org/10.1016/j.jneumeth.2011.09.021>.
- Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis I: segmentation and surface reconstruction. *NeuroImage* 9, 179–194.
- Dennis, E.L., Jahanshad, N., Rudie, J.D., Brown, J.A., Johnson, K., McMahon, K.L., de Zubicaray, G.I., Montgomery, G., Martin, N.G., Wright, M.J., Bookheimer, S.Y., Dapretto, M., Toga, A.W., Thompson, P.M., 2011. Altered structural brain connectivity in healthy carriers of the autism risk gene, CNTNAP2. *Brain Connect.* 1, 447–459, <http://dx.doi.org/10.1089/brain.2011.0064>.
- Dennis, E.L., Jahanshad, N., McMahon, K.L., de Zubicaray, G.I., Martin, N.G., Hickie, I.B., Toga, A.W., Wright, M.J., Thompson, P.M., 2013. Development of brain structural connectivity between ages 12 and 30: a 4-Tesla diffusion imaging study in 439 adolescents and adults. *NeuroImage* 64, 671–684, <http://dx.doi.org/10.1016/j.neuroimage.2012.09.004>.
- Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* 31, 968–980, <http://dx.doi.org/10.1016/j.neuroimage.2006.01.021>.
- Domin, M., Langner, S., Hosten, N., Lotze, M., 2014. Comparison of parameter threshold combinations for diffusion tensor tractography in chronic stroke patients and healthy subjects. *PLoS One* 9, e98211, <http://dx.doi.org/10.1371/journal.pone.0098211>.
- Fernández-Miranda, J.C., Wang, Y., Pathak, S., Stefaneau, L., Verstynen, T., Yeh, F.-C., 2015. Asymmetry, connectivity, and segmentation of the arcuate fascicle in the human brain. *Brain Struct. Funct.* 220, 1665–1680, <http://dx.doi.org/10.1007/s00429-014-0751-7>.
- Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical Surface-Based Analysis – II: Inflation, Flattening, and a Surface-based Coordinate System [WWW Document], URL <http://discovery.ucl.ac.uk/145122/> (accessed 9.21.11).
- Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation. *Neuron* 33, 341–355, [http://dx.doi.org/10.1016/S0896-6273\(02\)00569-X](http://dx.doi.org/10.1016/S0896-6273(02)00569-X).
- Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. *Cereb. Cortex* 14, 11–22, <http://dx.doi.org/10.1093/cercor/bhg087>.
- Greve, D.N., Fischl, B., 2009. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage* 48, 63–72, <http://dx.doi.org/10.1016/j.neuroimage.2009.06.060>.
- Kamali, A., Sair, H.I., Radmanesh, A., Hasan, K.M., 2014. Decoding the superior parietal lobule connections of the superior longitudinal fasciculus/arcuate fasciculus in the human brain. *Neuroscience* 277C, 577–583, <http://dx.doi.org/10.1016/j.neuroscience.2014.07.035>.

- Kunimatsu, A., Aoki, S., Masutani, Y., Abe, O., Hayashi, N., Mori, H., Masumoto, T., Ohtomo, K., 2004. The optimal trackability threshold of fractional anisotropy for diffusion tensor tractography of the corticospinal tract. *Magn. Reson. Med. Sci. MRMS Off. J. Jpn. Soc. Magn. Reson. Med.* 3, 11–17.
- Lee, A., Ratnarajah, N., Tuan, T.A., Chen, S.-H.A., Qiu, A., 2015. Adaptation of brain functional and structural networks in aging. *PLoS One* 10, e0123462, <http://dx.doi.org/10.1371/journal.pone.0123462>.
- Lo, C.-Y., Wang, P.-N., Chou, K.-H., Wang, J., He, Y., Lin, C.-P., 2010. Diffusion tensor tractography reveals abnormal topological organization in structural cortical networks in Alzheimer's disease. *J. Neurosci. Off. J. Soc. Neurosci.* 30, 16876–16885, <http://dx.doi.org/10.1523/JNEUROSCI.4136-10.2010>.
- Makris, N., Kennedy, D.N., McInerney, S., Sorensen, A.G., Wang, R., Caviness, V.S., Pandya, D.N., 2005. Segmentation of subcomponents within the superior longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. *Cereb. Cortex N. Y. N* 1991 15, 854–869, <http://dx.doi.org/10.1093/cercor/bhh186>.
- Muhlert, N., Sethi, V., Schneider, T., Daga, P., Cipolotti, L., Haroon, H.A., Parker, G.J.M., Ourselin, S., Wheeler-Kingshott, C.A.M., Miller, D.H., Ron, M.A., Chard, D.T., 2013. Diffusion MRI-based cortical complexity alterations associated with executive function in multiple sclerosis. *J. Magn. Reson. Imaging* 38, 54–63, <http://dx.doi.org/10.1002/jmri.23970>.
- Mukherjee, P., Chung, S.W., Berman, J.L., Hess, C.P., Henry, R.G., 2008. Diffusion tensor MR imaging and fiber tractography: technical considerations. *Am. J. Neuroradiol.* 29, 843–852, <http://dx.doi.org/10.3174/ajnr.A1052>.
- Nagy, Z., Alexander, D.C., Thomas, D.L., Weiskopf, N., Sereno, M.I., 2013. Using high angular resolution diffusion imaging data to discriminate cortical regions. *PLoS One* 8, e63842, <http://dx.doi.org/10.1371/journal.pone.0063842>.
- Otsu, N., 1979. A Threshold Selection Method From Gray-Level Histograms, 9, pp. 62–66.
- Parizel, P.M., Van Rompaey, V., Van Loock, R., Van Hecke, W., Van Goethem, J.W., Leemans, A., Sijbers, J., 2007. Influence of user-defined parameters on diffusion tensor tractography of the corticospinal tract. *Neuroradiol. J.* 20, 139–147.
- Rodrigues, P., Prats-Galino, A., Gallardo-Pujol, D., Villoslada, P., Falcon, C., Prčková, V., 2013. Evaluating structural connectomics in relation to different Q-space sampling techniques. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013, Lecture Notes in Computer Science*. Springer, Berlin Heidelberg, pp. 671–678.
- Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. *NeuroImage* 52, 1059–1069, <http://dx.doi.org/10.1016/j.neuroimage.2009.10.003>.
- Salat, D.H., Greve, D.N., Pacheco, J.L., Quinn, B.T., Helmer, K.G., Buckner, R.L., Fischl, B., 2009. Regional white matter volume differences in nondemented aging and Alzheimer's disease. *NeuroImage* 44, 1247–1258, <http://dx.doi.org/10.1016/j.neuroimage.2008.10.030>.
- Seehaus, A.K., Roebroek, A., Chiry, O., Kim, D.-S., Ronen, I., Bratzke, H., Goebel, R., Galuske, R.A.W., 2012. Histological validation of DW-MRI tractography in human postmortem tissue. *Cereb. Cortex* bhs036, <http://dx.doi.org/10.1093/cercor/bhs036>.
- Taoka, T., Morikawa, M., Akashi, T., Miyasaka, T., Nakagawa, H., Kiuchi, K., Kishimoto, T., Kichikawa, K., 2009. Fractional anisotropy–threshold dependence in tract-based diffusion tensor analysis: evaluation of the uncinate fasciculus in Alzheimer disease. *Am. J. Neuroradiol.* 30, 1700–1703, <http://dx.doi.org/10.3174/ajnr.A1698>.
- Wedeen, V.J., Wang, R.P., Schmahmann, J.D., Benner, T., Tseng, W.Y.I., Dai, G., Pandya, D.N., Hagmann, P., D'Arceuil, H., de Crespigny, A.J., 2008. Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *NeuroImage* 41, 1267–1277, <http://dx.doi.org/10.1016/j.neuroimage.2008.03.036>.
- Yeh, F.-C., Tseng, W.-Y.I., 2013. Sparse solution of fiber orientation distribution function by diffusion decomposition. *PLoS One* 8, e75747, <http://dx.doi.org/10.1371/journal.pone.0075747>.
- Yeh, F.-C., Wedeen, V.J., Tseng, W.-Y.I., 2010. Generalized q-sampling imaging. *IEEE Trans. Med. Imaging* 29, 1626–1635, <http://dx.doi.org/10.1109/TMI.2010.2045126>.
- Yeh, F.-C., Verstynen, T.D., Wang, Y., Fernández-Miranda, J.C., Tseng, W.-Y.I., 2013. Deterministic diffusion fiber tracking improved by quantitative anisotropy. *PLoS One* 8, e80713, <http://dx.doi.org/10.1371/journal.pone.0080713>.
- Zajac, L., Koo, B.-B., Bauer, C.M., Killiany, R., 2017. Behalf of the Alzheimer's disease neuroimaging initiative, null, seed location impacts whole-brain structural network comparisons between healthy elderly and individuals with Alzheimer's disease. *Brain Sci.* 7, <http://dx.doi.org/10.3390/brainsci7040037>.