

White Matter Differences between Healthy Young ApoE4 Carriers and Non-Carriers Identified with Tractography and Support Vector Machines

Laurence O'Dwyer^{1*}, Franck Lambertson², Silke Matura¹, Monika Scheibe¹, Julia Miller¹, Dan Rujescu^{3,4}, David Prvulovic¹, Harald Hampel¹

1 Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, Goethe University, Frankfurt, Germany, **2** Centre for Imaging-Neurosciences and Applications to Pathologies, UMR 6232, CNRS, CEA, University of Caen and Paris Descartes, Caen, France, **3** Department of Psychiatry, Ludwig-Maximilians-University, Munich, Germany, **4** Department of Psychiatry, University of Halle-Wittenberg, Halle, Germany

Abstract

The apolipoprotein E4 (ApoE4) is an established risk factor for Alzheimer's disease (AD). Previous work has shown that this allele is associated with functional (fMRI) changes as well structural grey matter (GM) changes in healthy young, middle-aged and older subjects. Here, we assess the diffusion characteristics and the white matter (WM) tracts of healthy young (20–38 years) ApoE4 carriers and non-carriers. No significant differences in diffusion indices were found between young carriers (ApoE4+) and non-carriers (ApoE4–). There were also no significant differences between the groups in terms of normalised GM or WM volume. A feature selection algorithm (ReliefF) was used to select the most salient voxels from the diffusion data for subsequent classification with support vector machines (SVMs). SVMs were capable of classifying ApoE4 carrier and non-carrier groups with an extremely high level of accuracy. The top 500 voxels selected by ReliefF were then used as seeds for tractography which identified a WM network that included regions of the parietal lobe, the cingulum bundle and the dorsolateral frontal lobe. There was a non-significant decrease in volume of this WM network in the ApoE4 carrier group. Our results indicate that there are subtle WM differences between healthy young ApoE4 carriers and non-carriers and that the WM network identified may be particularly vulnerable to further degeneration in ApoE4 carriers as they enter middle and old age.

Citation: O'Dwyer L, Lambertson F, Matura S, Scheibe M, Miller J, et al. (2012) White Matter Differences between Healthy Young ApoE4 Carriers and Non-Carriers Identified with Tractography and Support Vector Machines. PLoS ONE 7(4): e36024. doi:10.1371/journal.pone.0036024

Editor: Francisco José Esteban, University of Jaén, Spain

Received: December 22, 2011; **Accepted:** March 26, 2012; **Published:** April 25, 2012

Copyright: © 2012 O'Dwyer et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by institutional research funds of the School of Medicine, Goethe University, Frankfurt, Germany, to HH. This work was also supported by the Neurodegeneration & Alzheimer's disease research grant of the LOEWE (Landesoffensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz) program "Neuronal Coordination Research Focus Frankfurt" (NeFF) to HH and DP. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No additional external funding was received for this study.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: larodwyer@gmail.com

Introduction

Apolipoprotein E (ApoE) regulates the metabolism of lipids by coordinating their transport and redistribution from one cell type to another via ApoE receptors and proteins associated with lipid transfer and lipolysis [1]. ApoE is thought to play a key role in neuronal development and brain plasticity [2]. There are three allelic variants of the ApoE gene in humans (E2, E3, E4) [3]. Although the frequency of the ApoE4 allele is low in humans, primate studies suggest that it is the ancestral allele and it has been hypothesized that the mutation leading to ApoE3 may have been selected for during the evolution of the unique role of grandmothering in early humans which led to the spread of the E3 allele [3].

The E4 allele has been shown to confer a higher risk of developing both early onset and late onset AD [4,5]. Brain structure and function have been found to be altered in ApoE4 carriers, both in AD patients [6–9] and in healthy subjects [10–14]. Studies have found greater rates of temporal lobe atrophy in AD patients with greater load of E4 allele [6,15–17] and reduced medial temporal lobe volumes in healthy ApoE4 carriers across

the age spectrum [14,18–21]. However, a number of studies have also failed to replicate these findings [22–24]. A mixed picture is also found in terms of functional studies, with both increased [6,9,25] and decreased [26,27] task-related BOLD signals reported in carrier groups relative to non-carriers.

As the E4 allele has been related to several deleterious biological effects, the question arises as to why it has persisted. Recent work has suggested that the ApoE4 genotype might exert an antagonistic pleiotropic effect [28] whereby carriers have contrasting effects across the lifespan, with significant negative effects occurring in middle to old age, but with specific advantages noted in healthy young carriers. This pleiotropic effect of E4 may account for some of the discrepancies noted in the studies above. E4 has also been associated with higher IQ [29], a higher education level [30], reduced cardiovascular response to experimentally induced stress [31], a protective effect against spontaneous abortion during embryogenesis [32], and against perinatal death [33]. Additionally, hippocampal LTP was enhanced at a young age in knock-in mice expressing human APOE4 [34].

While the majority of studies to date have focussed on GM changes, few studies have investigated the WM characteristics of

ApoE4 carriers and non-carriers. As the ApoE gene is involved in the transport of lipids which in turn are used for the construction of myelin sheaths [35,36], the role of WM development in ApoE4 carriers and non-carriers will clearly be of great importance. In Alzheimer disease (AD) WM pathology has been shown to correlate with disease severity [37–39]. WM changes may also be a key indicator of early pathology [40] and WM damage in AD has been highlighted both in postmortem studies [41] as well as in *in vivo* MRI studies [42–44]. Degradation of WM microstructure can occur secondary to grey matter (GM) pathology via an accumulation of aggregated hyperphosphorylated tau protein and the deposition of A β [45]. There is also evidence for direct WM damage occurring as a result of oligodendrocyte death and reactive gliosis [46].

To date two previous DTI studies in healthy older people found subtle but significant decreases in FA in the parahippocampal gyrus [47] and in the splenium of the corpus callosum [48] in ApoE4 carriers. A recent whole brain tract based spatial statistics (TBSS) study found widespread significant differences between healthy young carriers and non-carriers [49]. Therefore, the few studies that exist indicate the possibility that ApoE4 carriers exhibit WM damage in early adulthood.

The aim of the current study was to investigate the effect of APOE genotype on WM structure in healthy young people. Whole brain TBSS was applied to detect significant differences that may be present between the groups. Diffusion data derived from TBSS was also run through a feature selection algorithm (ReliefF) [50] which identified the most salient voxels for group classification that were subsequently employed in the training of a support vector machine (SVM) classifier. The SVM was then used to classify the two groups. The voxels selected by the ReliefF algorithm were also used as seeds for probabilistic tractography.

We hypothesized that tractography initiated from seed voxels selected by ReliefF would identify a WM network that may be structurally different between carrier and non-carrier groups. We also hypothesized that the SVM would be able to classify the two groups using diffusion information from voxels selected by the ReliefF algorithm.

Methods

Ethics Statement

The study was approved by the Ethics Committee of Goethe University and was in accordance with the Declaration of Helsinki. All participants provided informed written consent.

Participants

44 cognitively intact persons between 20 and 38 years of age, all without any history of neurological or psychiatric disease were assessed in the current study. These 44 subjects were drawn from a larger cohort of 96 subjects. All of the 44 selected subjects were right-handed, as assessed with the Edinburgh Handedness Inventory [51] and provided written informed consent. Ethics approval was obtained from the local ethics committee of JWG University Frankfurt. All subjects underwent neuropsychological assessment. Verbal learning and memory was assessed using the German Version of the California Verbal Learning Test (CVLT) [52,53], visual memory was tested with the Brief Visual Memory Test - R (BVMT R) [54]. Additionally, measures of working memory and attention were obtained using the Letter Number Sequencing (LNS) [55], Spatial Span of the Wechsler Memory Scale 3 (WMS SS) [56] and Trail Making Test A (TMT). The verbal IQ was tested with a German verbal intelligence test (Mehrfachwahl-Wortschatz-Test B; MWTB), in which subjects

had to indicate real words within lists of pseudo-words [57]. Depressive Symptoms were measured with the German Version of the Beck Depression Inventory (BDI 2) [58,59].

All participants from the larger cohort ($n = 96$) underwent APOE genotyping using PCR and sequencing. For the current analysis, 21 subjects who were heterozygote for ApoE4 ($\epsilon 3/\epsilon 4$) and one subject who was homozygote for ApoE4 were included into the $\epsilon 4+$ group. 22 subjects, matched for age, gender and education who were $\epsilon 4$ negative ($\epsilon 3/\epsilon 3$) were included into the $\epsilon 4$ group. Group characteristics are summarized in Table 1.

ApoE4 Genotyping

APOE genotyping of the two determining variants rs7412 and rs429358 was analyzed using pre-designed TaqMan SNP Genotyping assays (Applied Biosystems, Foster City, CA). Briefly for each SNP 20 μ l reaction mix contained 15 ng genomic DNA, unlabeled PCR primers, MGB labeled probes (VIC, 6FAM), 10 μ l of 2 \times TaqMan universal PCR Master Mix (Applied Biosystems, Foster City, CA). PCR was performed on an ABI 7000 instrument (Applied Biosystems, Foster City, CA) with the following cycling program: 95°C for 15 s, 40 cycles of 95°C for 15 s and 60°C for 60 s. The ABI 7000 genotyping software was used for allelic discrimination.

Imaging Methods

All MR images were acquired using a Trio 3-T scanner (Siemens, Erlangen, Germany) with a standard head coil for radiofrequency transmission and signal reception. Participants were outfitted with protective earplugs to reduce scanner noise and a hand-held response device. For T1 weighted structural brain imaging, an optimized 3D modified driven equilibrium Fourier transform (3D MDEFT) sequence was used with the following

Table 1. Demographic and cognitive characteristics of the sample groups.

Variable	APOE4 non-carriers		APOE4 carriers		T-value	P-value
	Mean	SD	Mean	SD		
	n = 22		n = 22			
Age (years)	26.7	4.0	26.9	5.3	5.28	0.92
Gender (m/f)	13/9		13/9			0.76
Education (years)	16.8	4.5	17.0	4.3	−0.15	0.88
MWTB	29.7	3.6	30.3	4.3	−0.46	0.65
MWTB IQ	106.7	23.6	114.7	15.3	−1.28	0.21
TMT (sec)	22.0	5.8	19.3	3.9	1.81	0.08
WMS SS	19.1	2.0	19.4	2.8	−0.31	0.76
LNS	18.7	3.2	17.8	2.7	1.07	0.29
BVMT R	32.7	3.5	32.0	3.6	0.62	0.54
BDI 2	3.2	3.6	2.4	2.9	0.83	0.41
CVLT	66.8	8.0	64.5	9.1	0.90	0.37

Values are mean \pm standard deviation. Significance was set at $p < 0.05$; thus no significant differences were found between the groups. Values denote mean and standard deviation or number of subjects. P-values refer to t-tests (parametric tests) and chi-square tests (for categorical data). Abbreviations: MWTB: Mehrfachwahl-Wortschatz-Test B, a German Verbal intelligence test; TMT: trail making test; WMS SS: Spatial Span of the Wechsler Memory Scale; LNS: Letter Number Sequencing; BVMT R: Brief Visual Memory Test R; BDI 2: Beck Depression Inventory 2; CVLT: California Verbal Learning Test. doi:10.1371/journal.pone.0036024.t001

parameters: acquisition matrix = 256×256 , repetition time (TR) = 7.92 ms, echo time (TE) = 2.48 ms, field of view = 256 mm, 176 slices, 1.0 mm slice thickness.

DTI was acquired using an echo planar imaging (EPI) sequence with the following pulse sequence: TR = 8760 ms; TE = 100 ms; acquisition voxel size = $2 \times 2 \times 2$ mm³; 60 axial adjacent slices; slice thickness = 2 mm (no gap); FOV = 192 mm \times 192 mm \times 120 mm; acquisition matrix = 96×96 ; 10 images without diffusion weighting and 60 diffusion-encoding gradients applied in 60 noncollinear directions; b-value = 1000 s/mm²; both the b0 and the 60 diffusion weighted images were averaged three times, bandwidth = 1302 Hz/pixel; total acquisition time = 10 min 31 sec. This sequence was acquired using generalised auto-calibrating parallel acquisitions (GRAPPA; Griswold et al., 2002) with an acceleration factor of 2.

For each subject a total of three consecutive DTI scans were acquired. An average of these three scans was then created for use in subsequent DTI processing.

A T2-weighted fluid attenuation inversion recovery (FLAIR) sequence was also acquired to ensure that vascular pathology was not significant. For all 44 subjects selected from the larger cohort, no hyperintense white matter lesions were seen in the FLAIR scans.

High Resolution T1W Structural Image Processing

Images were skull stripped with the Brain Extraction Tool (BET) from the FSL library [60]. Brain tissue volume, normalized for subject head size, was estimated with SIENAX [61,62], which is part of the FSL library. SIENAX starts by extracting brain and skull images from the single whole-head input data. The brain image is then affine-registered to MNI152 space [63,64] (using the skull image to determine the registration scaling); this is primarily in order to obtain the volumetric scaling factor, to be used as a normalization for head size. Next, tissue-type segmentation with partial volume estimation is carried out [65] in order to calculate total volume of brain tissue including separate estimates of volumes of WM and GM.

DTI Processing

DTI analysis was performed using TBSS [66]. Images were skull stripped with the Brain Extraction Tool (BET) from the FSL library [60]. Raw DTI images were first corrected for motion and eddy current effects. The diffusion tensor was then calculated with the DTIFIT program for whole brain volumes and the resulting FA maps, together with the DA (λ_1) and DR $((\lambda_2 + \lambda_3)/2)$ and MD $((\lambda_1 + \lambda_2 + \lambda_3)/3)$ maps, were used in subsequent TBSS analysis.

TBSS performs a non-linear registration that aligns each FA image to every other one and calculates the amount of warping needed for the images to be aligned. The most representative image is determined as the one needing the least warping for all other images to align to it. The FSL library also provides a 1 mm isotropic FA target image (FMRIB58_FA) in standard space, which is sometimes used instead of the most representative image from the study cohort. This can be problematic as the target image is based on a young healthy brain. Using the method of “all subject to all subject” registration is more computationally intensive, but highly desirable when dealing with populations other than young healthy controls. After this registration step, warped versions of each subject’s FA image were generated which were then averaged and a white matter “skeleton” was then created suppressing all non-maximum FA values in each voxel’s local-perpendicular direction and subsequently comparing all remaining non-zero voxels with their nearest neighbours, thus searching for the centre of fibre bundles. The skeleton was then

thresholded at an FA value of 0.2 which limits the effects of poor alignment across subjects and ensures that GM and CSF voxels are excluded from the skeleton. The resulting skeleton contained WM tracts common to all subjects. A “distance map” is then created which is used to project each FA image onto the mean FA skeleton that is common to all subjects [66]. The same non-linear transformations derived for the FA maps were applied to the DA, DR and MD maps.

Following TBSS processing, a global region of interest was created using the white matter skeleton that is common to all subjects. Mean values of FA, DA, DR and MD were extracted from each subject using this global ROI.

For statistical analysis, the images were analyzed using the “randomise” tool from FSL using a standard GLM design which controlled for the effect of age and gender. Randomise computes permutation tests on the assumption that the null hypothesis implies complete exchangeability of the observations [67]. Using this setup voxelwise differences between groups were then assessed, setting the number of permutations at 5000 permutations. Significance was tested at $p < 0.05$ levels, corrected for multiple comparisons using the “2D” parameter settings with threshold-free cluster enhancement (TFCE), a method which avoids using an arbitrary threshold for the initial cluster-formation [68].

Support Vector Machines Analysis

Classification of individual subjects was undertaken using the freely available WEKA software package (<http://www.cs.waikato.ac.nz/ml/weka>, Version 3.6.4) [69,70]. Following TBSS analysis, the skeletonised FA, DA, DR and MD data was analysed in Matlab (program written by FL and available on request), which extracted the diffusion values from the WM skeleton. There were 132,208 voxels in the WM skeleton and diffusion values for all indices of diffusion were extracted from each voxel in the WM skeleton. Classification between carrier and non-carrier groups was undertaken using each index of diffusion separately in order to determine the most efficient index for classification.

The first step of the WEKA analysis was to reduce the number of voxels to those that are most relevant for classification. This step eliminates non-discriminative voxels which would reduce classification accuracy. The feature selection algorithm “ReliefF” was used to extract the most important voxels [50]. For each classification group and also for each index of diffusion, eight reduced datasets were created as follows:

1. 250 voxel dataset
2. 500 voxel dataset
3. 750 voxel dataset
4. 1000 voxel dataset
5. 2000 voxel dataset
6. 3000 voxel dataset
7. 4000 voxel dataset
8. 5000 voxel dataset

Therefore in total, 32 reduced datasets were created; i.e. 8 reduced datasets for each index of diffusion (FA, MD, DR and DA). The choice of identifying these 8 datasets is based on previous studies that have shown that this is an optimum range of data reduction for successful classification [71–74].

The aim of the ReliefF algorithm is to estimate the quality of voxels according to how well the value of a voxel distinguishes between instances that are near to each other. The central idea of ReliefF is to estimate the quality of voxels according to how well their values distinguish between instances that are near to each

other [50,75]. Given a randomly selected instance Ins_m from class L , ReliefF searches for K of its nearest neighbours from the same class called nearest hits H , and also K of its nearest neighbours from each of the different classes, called nearest misses M . It then updates the quality estimate Wi for voxel i based on the values for Ins_m , H and M . If instance Ins_m and those in H have different values for voxel i , then the quality estimation Wi is decreased. Alternatively, if instance Ins_m and those in M have different values on the voxel i , then Wi is increased [50,75].

After reducing the data into the three datasets of differing sizes, classification was then performed using the SVM algorithm “sequential minimal optimization” (SMO) [76] with a radial basis function (RBF) kernel [77]. SVMs are algorithms that learn how to assign labels to objects [78]. They use linear models to implement nonlinear class boundaries by transforming the input into a new higher dimensional space. In this way, a straight line in the new space can be curved or non-linear when transformed back to the original lower-dimensional space. Following transformation, a linear model called the maximum margin hyperplane is created. To visualise this, imagine a dataset with two-classes that are linearly separable. The maximum margin hyperplane is the one that gives the greatest separation between the classes. The hyperplane describes a straight line in a high-dimensional space, and therefore a separating hyperplane is a line that separates the classes. The instances that are closest to the maximum margin hyperplane are called support vectors. A unique set of support vectors defines the maximum margin hyperplane for the learning problem. Once the support vectors for the two classes are established, a maximum margin hyperplane can be constructed.

The projection of the data from low dimensional space to higher dimensional space is achieved with a kernel function. The optimal kernel function is usually found by trial and error [78]. In the current study a radial basis function (RBF) kernel was used to nonlinearly map samples into a higher dimensional space. RBF kernels use two parameters: C and $GAMMA$. $GAMMA$ represents the width of the radial basis function, and C represents the error/trade-off parameter that adjusts the importance of the separation error in the creation of the separation surface [79]. C was fixed to 1 and $GAMMA$ was fixed to 0.01. During each trial, tenfold cross validation was used for classification. This involves dividing the data into 10 parts. Nine parts were used for training and one part for testing. This was repeated 10 times, resulting in the learning algorithm being implemented 100 times on datasets that are all nine-tenths the size of the original [69,70]. This is a standard procedure in machine learning which reduces the variation related to data selection and allows results to be averaged to yield robust calculations of the performance of the SVM. During each of the ten repetitions within a given trial, the same values of C and $GAMMA$ are retained for each repetition.

For the analysis of results, measures of sensitivity, specificity and an F-measure were used. An F-measure is used to give a single measure that can characterise overall performance [70]. An F-measure is the harmonic mean of Precision and Recall.

A workflow, outlining the creation of the WM skeleton, the extraction of diffusion values from each point on the WM skeleton, the reduction of the data using ReliefF and the application of the SVM is shown Fig. 1.

Tractography

Seeds were created from the significant clusters created from the ReliefF algorithm, and these seeds were used for probabilistic tractography [80]. In order to perform the probabilistic tractography in standard space for each subject, we fed both warpfields generated with FNIRT when registering individual subjects to

standard space, and their corresponding inversed warpfields into the tractographic algorithm. A multi-fibre model was fit to diffusion data at each voxel which allows tracing of fibres through regions of fibre crossing and complexity [81]. At each voxel, a probability distribution function (pdf) is estimated on each fibre direction. Tractography then proceeds by drawing multiple (in this case 5,000) streamline samples through these pdfs from each seed voxel to create an estimate of the distribution of connections from each seed voxel. When these streamlines reach a voxel in which more than one direction is estimated they follow the direction that is closest to parallel with the direction at which the streamline arrives. Generated pathways are volumes in which values at each voxel represent the number of samples passing through that voxel and, therefore, the probability of connection to the seed voxel. To remove spurious connections, pathways in individual subjects were thresholded to include only voxels which had at least 250 samples passing through them (out of 5,000) generated from each seed voxel. These pathways were then binarised and overlaid to create population probability maps in which voxel values represent the number of subjects in whom a pathway is present. As is noted in the Results and Discussion sections, tractography was performed using the seed points from the FA dataset because this diffusion index produced the greatest classification accuracy.

Results

Demographic and Cognitive Characteristics

There were no significant differences between the groups in terms of any of the demographic or psychological measures taken (Table 1).

Normalised Grey Matter and White Matter Volumes

Normalised Grey Matter and White Matter Volumes were not significantly different between groups ($p > 0.05$) (Fig. 2).

Differences in Multiple Indices of Diffusion between ApoE4 Carriers and Non-Carriers

After correcting for multiple comparisons no significant voxelwise differences were found between the two groups using randomise permutation testing of the TBSS images. Global boxplots for the four indices of diffusion can be seen in Figure 3. The diffusion values in these boxplots are taken from the entire white matter skeleton.

SVM Classification of ApoE4 Carrier and Non-Carrier

The ReliefF algorithm was used to reduce the FA, DA, DR and MD datasets to datasets that contain the most salient voxels for group classification. Figure 4(a) presents a graphical depiction of the reduction of the FA dataset by ReliefF to the 500 most important voxels as judged by the weighting procedure described in the methods. The volume of clusters is noted on the y-axis. Figure 4(b) depicts the anatomical location of some of the voxels selected by ReliefF following the reduction of the FA dataset. The largest cluster of voxels is also indicated in the figure. The same process was followed when reducing the FA data to 250, 750, 1000, 2000, 3000, 4000 and 5000 voxel datasets. The entire procedure was repeated for the DR, MD and DA datasets.

The reduced datasets were then used by the SVM for training and subsequent classification of ApoE4 carriers and non-carriers. An SVM analysis was performed separately using FA, DA, DR and MD datasets. The results from these four different SVM analyses are shown in Figure 5. The FA data produced the best overall performance, with sensitivity and specificity ranging from 93 to 100%. For the DA, DR and MD indices of diffusion,

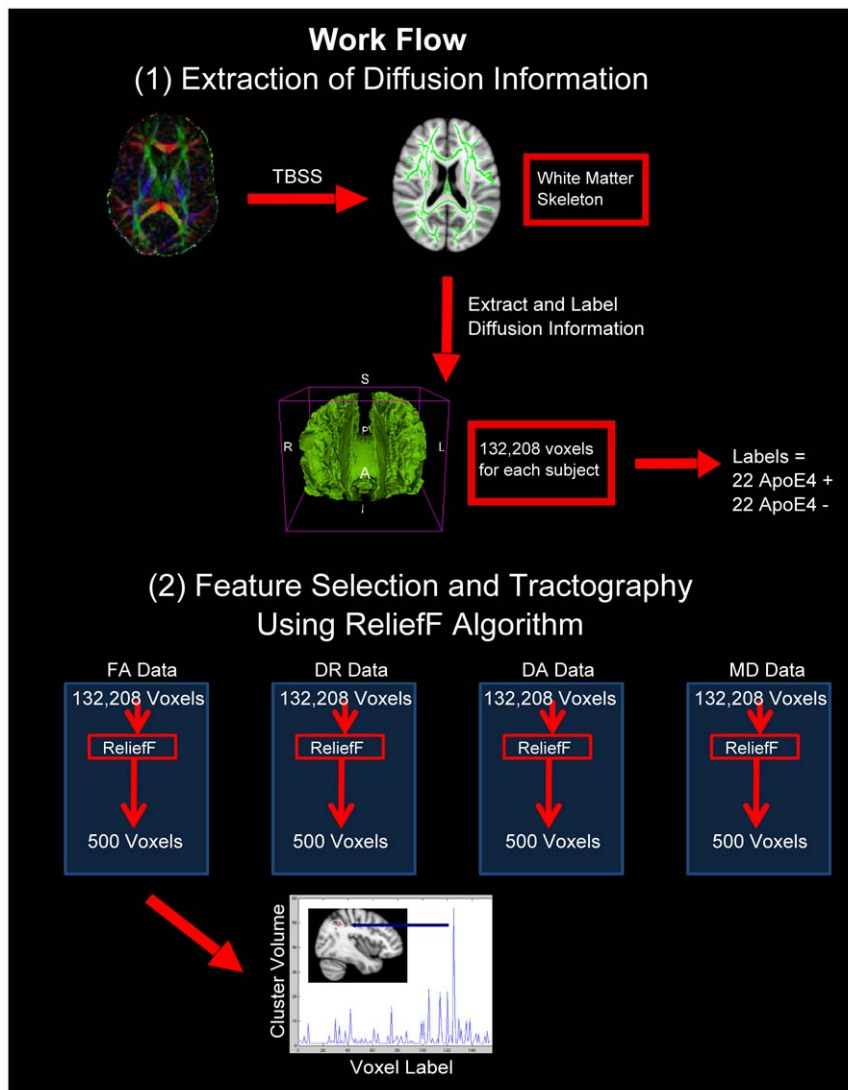


Figure 1. Workflow for the processing of white matter diffusion information. (1) Following TBSS processing a white matter skeleton is created which is common to all subjects. Diffusion values are then extracted from every voxel in the white matter skeleton using a custom made matlab script. Data is then labelled as either ApoE4 carrier or non-carrier. (2) Following preparation of the data into a WEKA compatible format the data is reduced using a feature selection algorithm, ReliefF. This algorithm selects the most salient voxels for group classification. These voxels are then used for training a support vector machine (SVM) classifier and subsequently the classification is performed with this classifier. The lowermost figure shows a readout of top 500 voxels selected from the full FA dataset. The y-axis represents the volume of particular clusters and the location of the largest cluster is shown circled in white in the anatomical images.
doi:10.1371/journal.pone.0036024.g001

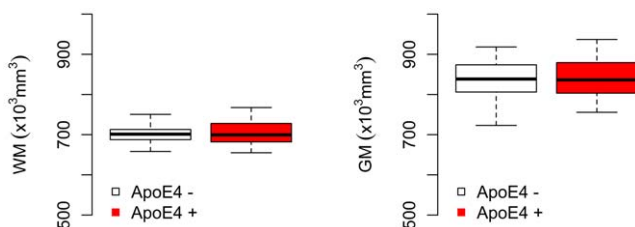


Figure 2. Normalised WM (left) and GM (right) volumes for healthy young ApoE4 carrier and non-carrier groups. No significant differences were found between the groups for either WM or GM volumes.
doi:10.1371/journal.pone.0036024.g002

classification performance was slightly lower, with sensitivity and specificity in the range of 86–100% (Fig. 5).

Tractography

As the voxels selected by ReliefF from the FA dataset produced the best overall performance in terms of group classification, the top 500 voxels from the FA data were used as seed points for tractography. Thus, tracts were initiated from all 500 of these voxels using probabilistic tractography in each ApoE4 carrier and non-carrier subject. These tracts (binarised at a threshold of 250 out of 5000 streamlines) are overlaid to create population probability heatmaps in which each voxel value represents the number of subjects in whom a pathway is present (Fig. 6). These population probability heatmaps highlight the WM network that is initiated from the 500 voxel seed created with the ReliefF

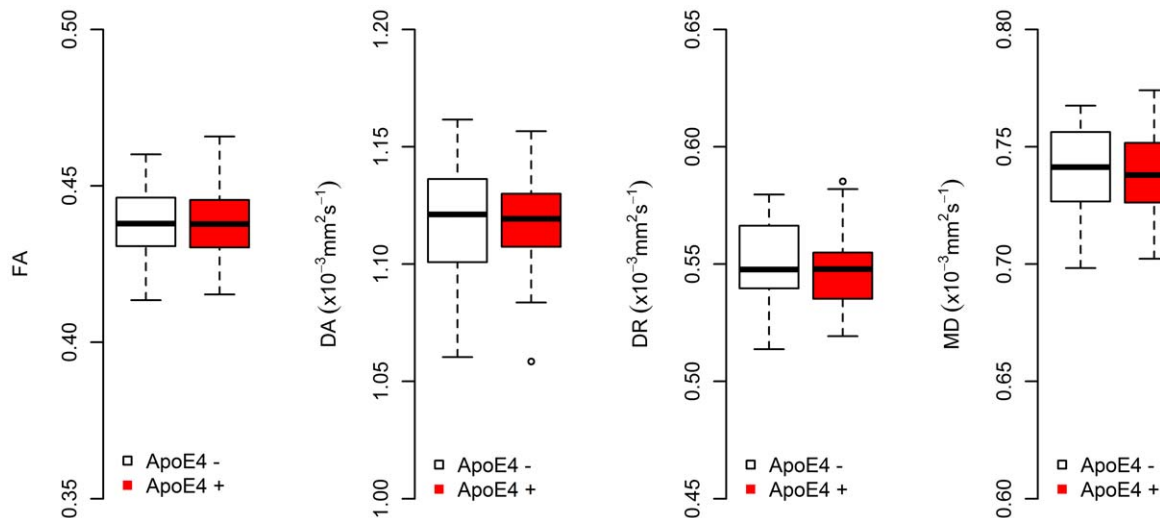


Figure 3. Boxplots showing the distribution of diffusion tensor MRI indices for the global WM skeleton in ApoE4 carriers (red) and non-carriers (white). The boxplots represent the interquartile ranges, which contain 50% of individual subjects' values. The whiskers are lines that extend from the box to the highest and lowest values. No significant differences were found between the groups.
doi:10.1371/journal.pone.0036024.g003

algorithm. As these tracts were initiated from seed points that were chosen by ReliefF as being the most useful for locating differences between the groups, these heatmaps give an indication of tracts that may be susceptible to micro and macrostructural damage in ApoE4 carriers. The WM network implicated in both carrier and non-carrier groups extends into much of the parietal lobe, including the posterior cingulum and precuneus. The forceps major, forceps minor, superior longitudinal fasciculus, as well as the uncinate fasciculus and large portions of the prefrontal cortex are also implicated. Finally, an analysis of the total volume of tracts

created from this 500 voxel seed is shown in Figure 7. ApoE4 carriers exhibited a non-significant ($p = 0.45$, $t = 0.76$) decrease in tract volume relative to non-carriers (Fig. 7). Mean tract volume was 190311.3 (SD 13151.6) mm^3 for ApoE4 carriers and 193441.9 (SD 14105.5) mm^3 for ApoE4 non-carriers.

Discussion

The current results show that it is possible to classify healthy young ApoE4 carriers and non-carriers using an automated

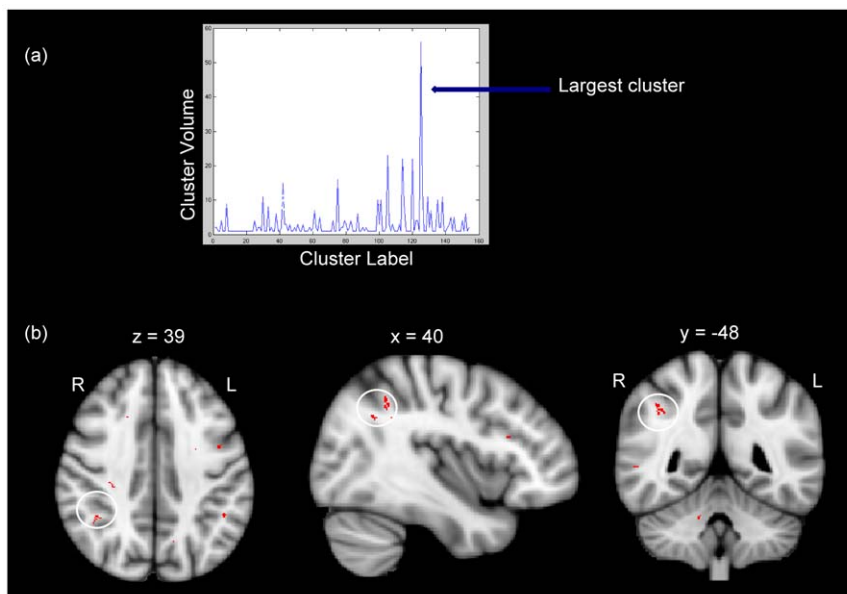


Figure 4. Voxels selected by the ReliefF algorithm as the most salient for group classification. (a) A graph of the most significant clusters of voxels selected by the ReliefF algorithm. This graph shows the cluster label on the x-axis and the volume of the contiguous clusters on the y-axis. (b) A selection of the top 500 voxels identified by the ReliefF algorithm. The white circles highlight the location of the largest cluster as identified by the previous graph. Note that this figure highlights the reduction of the entire FA dataset to the top 500 voxels. This process is followed for the top 250, 500, 750, 1000, 2000, 3000, 4000 and 5000 voxels of the FA dataset. This entire procedure is also repeated for DA, MD and DR datasets. The coordinates shown for each slice are MNI coordinates in millimetres. The labels L and R indicate left and right respectively. The middle panel ($x = 40$) indicates a sagittal slice in the right hemisphere.
doi:10.1371/journal.pone.0036024.g004

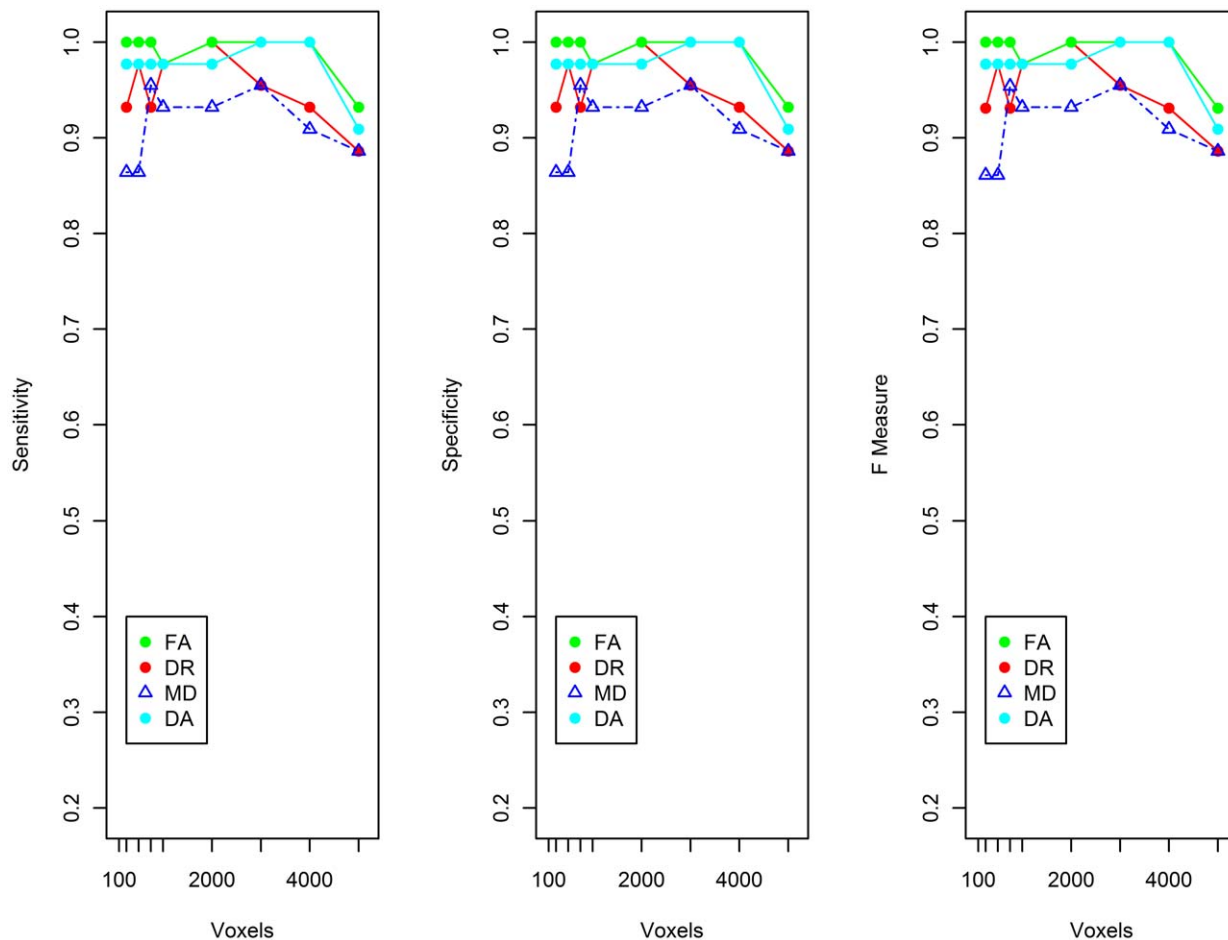


Figure 5. Sensitivity, Specificity and F-measure for SVM run on ApoE4+ and ApoE4- subjects. The F-measure is the harmonic mean of precision and sensitivity and can be used as a single measure of performance of the classification. doi:10.1371/journal.pone.0036024.g005

procedure based on structural WM differences between the two groups. Previous work has distinguished between healthy older subjects and subjects with mild cognitive impairment (MCI) using a combination of DTI and SVMs [74] but to the best of our knowledge no machine learning paradigm has attempted to classify healthy young ApoE4 carriers and non-carriers. Using diffusion data in combination with an SVM, our machine learning approach was able to classify ApoE4 carrier and non-carrier groups with an extremely high degree of accuracy.

Of note is the fact that the FA index produced the best classification performance. This is in line with a previous SVM classification study which also found that FA gave the greatest accuracy for the classification of healthy control and MCI subjects [74]. The reduced FA dataset was studied in further detail to assess the locations of the voxels selected by the ReliefF algorithm. This feature selection algorithm is able to effectively provide quality estimates for each voxel in the white matter skeleton. It has previously been used successfully in a number of different domains including genetics [82] and proteomics [83], and has recently begun to be used in MRI [74]. In the current study we see that it filters out irrelevant voxels and selects the most useful voxels which can be used by an SVM for training and classification. The largest cluster identified by ReliefF from the FA dataset was found in the right anterior parietal region. Interestingly, this region has been

shown to demonstrate increased BOLD response in older ApoE4 carriers relative to non-carriers during picture learning [13].

It has been suggested, on the basis of functional changes in ApoE4 carriers, that the E4 allele modulates neuronal activity decades before any expression of disease [8,84]. A significant interaction between age and E4 status has also been noted in the frontal lobe and other areas, such that aging was associated with decreased activity in E4 carriers and increased activity in non-carriers [85]. Therefore the ApoE genotype determines age-related changes in brain function that may reflect the increased vulnerability to cognitive decline in later life. Our finding of a subtle reduction in tract volumes in carriers may also be an indicator of an increased vulnerability of ApoE4 carriers in the tracts identified by tractography.

Also of note is the fact that 63% of the voxels (315 out of 500 voxels) selected by ReliefF were located in the right hemisphere. Laterality effects have been noted in a number of ApoE4 studies and evidence suggests that frontal-executive processes may be responsible for mediating a right hemisphere compensation in older age [86] [10] [87]. Bondi and colleagues found that the anterior cingulate was directly implicated in a right hemisphere compensatory mechanism response in healthy older ApoE4 carriers [87]. Laterality effects have also been noted in older ApoE4 carriers during an object naming task, where carriers were found to have greater right perisylvian activation [88]. Overall,

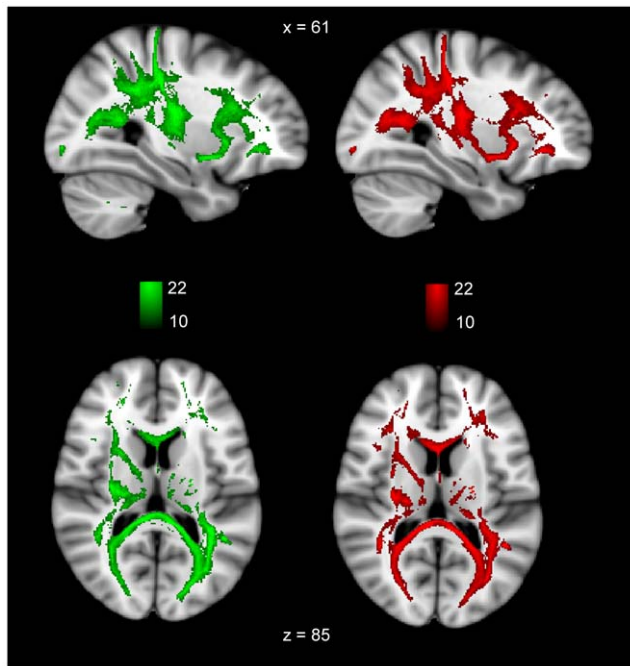


Figure 6. Tractography using seed points selected by the ReliefF algorithm for ApoE4 carriers and non-carriers, after reducing the full FA dataframe. Here we have heatmaps created from tracts that were thresholded at 250 out of 5,000 streamlines. At one end of the scale, the bright green (ApoE4 non-carrier) or bright red (ApoE4 carrier) indicates that tracts were present in all 22 subjects in each group, whereas at the other end, the dark green or dark blue indicates that tracts were present in 10 out of the total number of 22 subjects in each group.
doi:10.1371/journal.pone.0036024.g006

evidence suggests that the right hemisphere may be activated to a heightened degree in older ApoE4 carriers in order to compensate for early declines in episodic memory. Although the underlying reasons for this laterality are not known it is of interest that a model of aging, termed HAROLD (hemisphere asymmetry reductions in older adults) [89] posits that there is an age-related loss of hemispheric asymmetry of task-related activation in the prefrontal cortex (PFC). While healthy young subjects show greater activation in left PFC compared with the right PFC during the execution of a variety of cognitive tasks, many studies indicate that this laterality is reduced in older subjects [89–93]. This loss of laterality is frequently related to increased right PFC activation in older adults which has been suggested to serve as a possible compensatory mechanism in order to maintain cognitive performance in older age. The large number of functional activation studies showing increased right hemisphere functional activation in older age [89–93] may serve as indirect evidence for chronically stressed right hemisphere structures being predilection sites for structural alterations in neurodegenerative diseases such as AD. Subjects at genetic risk of AD (e.g. ApoE4 carriers) may display an increased vulnerability of these same predilection sites to structural damage even at younger ages. A number of recent MCI and AD dementia classification studies have also noted a right-more-than-left pattern WM damage [94–96]. It is tempting to consider that ReliefF's identification of a large number of clusters in the right hemisphere, may represent key WM nodes in ApoE4 carriers which may be particularly sensitive to WM damage in later life when the negative effects of the ApoE4 gene become more pronounced.

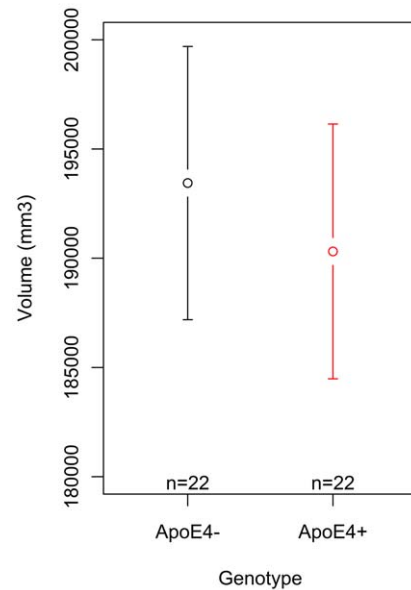


Figure 7. Volume of tracts produced using a seed of the top 500 voxels created by the ReliefF algorithm from the full FA dataset. The tracts were thresholded to include only WM tracts that contained 250 or more streamlines, out of a total of 5000 streamlines. ApoE4 carriers exhibited a non-significant ($p = 0.45$, $t = 0.76$) decrease in tract volume relative to non-carriers (Fig. 7). Mean tract volume was 190311.3 (SD 13151.6) mm³ for ApoE4 carriers and 193441.9 (SD 14105.5) mm³ for ApoE4 non-carriers.
doi:10.1371/journal.pone.0036024.g007

The WM tracts initiated from the 500 voxels selected by ReliefF include significant portions of the parietal lobe, the cingulum, the dorsolateral prefrontal cortex, forceps major, forceps minor and the uncinate fasciculus. It is of interest that the posterior parietal lobe including the posterior cingulum/precuneus is part of the network identified, as this is also an important hub which sustains information transfer between the parahippocampal gyrus and the prefrontal cortex [94,97]. A number of studies have proposed that damage to the posterior cingulum/precuneus leads to dysfunction of a network that is responsible for sustaining memory function [42,94,97–99]. The precuneus is connected to the temporal lobe via the retrosplenial cortex, and the hippocampus may not be able to communicate with the neocortex if it is sufficiently damaged [94,97,98]. PET studies have also highlighted reduced metabolism in the posterior cingulum and the prefrontal cortex in both MCI and AD subjects [100,101].

Analysis of the total volume of tracts initiated from the 500 voxel seed indicates a non-significant decrease in volume in the ApoE4 carrier group. Overall, these results suggest that there are subtle structural WM differences between ApoE4 carriers and non-carriers. When considered together with evidence from previous studies, it is possible that while the structural differences are subtle at a young age, the WM network that has been highlighted by tractography may be more susceptible to further structural damage in ApoE4 carriers as they progress into middle and old age.

It is of particular interest that although no significant differences were found between the groups for any index of diffusion when inspecting TBSS results that are corrected for multiple comparisons, the voxels selected by ReliefF could still be employed by an SVM for successful classification. The reason for this lies in methodological differences between the two approaches. For TBSS, each voxel is assigned a t-value that is corrected for

multiple comparisons, taking into account the fact that there are a total of 132,208 voxels in the entire white matter skeleton. The level of correction for multiple comparisons is therefore extremely high due to the large number of total voxels to be considered. However, the SVM approach is underpinned by a different theoretical framework. First, the most salient voxels are identified using the Relief algorithm. After this data reduction step the majority of voxels in the WM skeleton are discarded as irrelevant for subsequent SVM classification. Therefore, the SVM only considers the voxels identified by Relief in its calculation of a single parameter per subject, and does not run into the problem of multiple comparison biases [74].

It should be stressed that possession of the ApoE4 allele does not automatically imply the possibility of cognitive or structural deficits within the carrier group. A growing body of evidence suggests that the ApoE4 genotype may confer some beneficial effects at a young age, while being detrimental in older age [29–34]. Thus, there may be evolutionary reasons for the retention of the ApoE4 genotype because of its possible beneficial effects at a young age, which is offset by an increased risk of cognitive deficits in older age. Results from the current study however suggest that the ApoE4 gene may have no beneficial effects on WM structure in healthy young people.

On the basis of the current results, subtle microstructural WM differences do appear to be present in healthy young carrier subjects, which can be utilised by SVMs for group classification. The approach outlined in this study may offer a way to study a particular WM network longitudinally, based on the hypothesis that the network identified will be preferentially susceptible to structural damage in ApoE4 carriers in later life.

Only one previous study has examined structural WM differences between healthy young carrier and non-carrier groups [49]. The current results which found no significant corrected TBSS differences between the groups are at odds with this previous study which found significant FA decreases in young carriers relative to non-carriers. In fact, the FA reductions in young carriers relative to non-carriers found by Heise and colleagues appeared to be as widespread as the FA reductions noted previously in Alzheimer's patients relative to healthy older controls [42,43,98]. The age, sample size and overall demographics of the current cohort are comparable to those of Heise and colleagues; therefore, it might be possible that larger sample sizes need to be employed in future studies in order to determine whether or not widespread FA reductions in healthy young ApoE4 carriers relative to non-carriers can be replicated in the future.

To the best of our knowledge this is the first attempt to classify healthy young ApoE4 carrier and non-carrier groups with machine learning techniques. The use of 10 times 10-fold cross-validation [70] ensures the robustness and generality of the results by running the learning algorithm 100 times during each experiment. This method also reduces the effect of random variation when different folds are selected [69]. The use of a joint TBSS/SVM analysis allows information to be used from the entire brain which is a significant advantage over ROI based approaches for classification. The current methodology also obviates the need for the labour intensive selection and creation of ROIs.

Some limitations of the study should be noted. While the TBSS approach strives to avoid the problems of voxel based morphology

relating to partial voluming, small WM tracts may still be contaminated with GM if the tract width is smaller than the original voxel size [66]. However, applying an FA threshold of 0.2 or less to the WM skeleton in the TBSS preprocessing steps is thought to remove the potential occurrence of GM. It should also be noted that probabilistic tractography is influenced not only by the true underlying probability of an anatomical connection, but also by factors such as the distance from a seed voxel and the geometry of the pathways being tracked with more tortuous pathways or pathways involving crossing fibers being more difficult to track [81]. Probabilistic tractography is also insensitive to the polarity of axonal projections. However, the probabilistic method employed in the current study uses multi-fibre tractography which offers significant advantages over deterministic methods when tracking difficult white matter tracts such as those that include crossing fibres [81]. It is important to note that due to the difficulties of obtaining sufficient numbers of ApoE4 carriers, it has not been possible to validate the current SVM method in an independent sample. While 10 times 10-fold cross validation was used to ensure performance generalisation, validation in an independent sample will be essential to determine how robust the current approach is when applied to a fully independent dataset. Recent work adopting the current pipeline has been successfully used to classify DTI images from older subjects into MCI and healthy older categories [72]. While these findings from an older cohort support the validity of the SVM approach outlined here, caution should nevertheless be exercised before the method is validated in an independent sample. This is currently being pursued as part of the European DTI Study in Dementia (EDSD) initiative.

Overall, the current study demonstrates the efficacy of using DTI in conjunction with SVMs to classify ApoE4 carrier and non-carrier groups. The feature selection algorithm, ReliefF, allowed us to develop tractography from key seed regions that were identified as having significantly different diffusion properties between carrier and non-carrier groups. The current results suggest that there are very subtle structural differences between healthy ApoE4 carrier and non-carrier subjects, with carriers of the ApoE4 allele showing slight reductions in the volume of white matter tracts initiated from the seed regions identified by ReliefF. The white matter network developed from these seed regions, which incorporates the posterior and anterior cingulum, the precuneus, and the prefrontal cortex, may represent a WM network that is more susceptible to structural damage over the life course of ApoE4 carriers. Longitudinal studies would be needed to assess the diffusion properties and development of these tracts over the life span of ApoE4 carriers and non-carriers.

Acknowledgments

We thank Céline Bourdon and Felix Geser for comments, discussion and assistance relating to the manuscript.

Author Contributions

Conceived and designed the experiments: LOD FL SM DP HH. Performed the experiments: LOD SM JM MS. Analyzed the data: LOD FL. Contributed reagents/materials/analysis tools: LOD DP HH DR. Wrote the paper: LOD.

References

1. Verghese PB, Castellano JM, Holtzman DM (2011) Apolipoprotein E in Alzheimer's disease and other neurological disorders. *The Lancet Neurology* 10: 241–252. doi:10.1016/S1474-4422(10)70325-2.
2. Finch CE, Sapolsky RM (1999) The evolution of Alzheimer disease, the reproductive schedule, and apoE isoforms. *Neurobiol Aging* 20: 407–428.

3. Zannis VI, Breslow JL, Utermann G, Mahley RW, Weisgraber KH, et al. (1982) Proposed nomenclature of apoE isoproteins, apoE genotypes, and phenotypes. *J Lipid Res* 23: 911–914.
4. Okuizumi K, Onodera O, Tanaka H, Kobayashi H, Tsuji S, et al. (1994) ApoE-epsilon 4 and early-onset Alzheimer's. *Nat Genet* 7: 10–11. doi:10.1038/ng0594-10b.
5. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, et al. (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261: 921–923.
6. Lehtovirta M, Laakso MP, Soininen H, Helisalmi S, Mannermaa A, et al. (1995) Volumes of hippocampus, amygdala and frontal lobe in Alzheimer patients with different apolipoprotein E genotypes. *Neuroscience* 67: 65–72.
7. Cosentino S, Scarmeas N, Helzner E, Glymour MM, Brandt J, et al. (2008) APOE epsilon 4 allele predicts faster cognitive decline in mild Alzheimer disease. *Neurology* 70: 1842–1849. doi:10.1212/01.wnl.0000304038.37421.cc.
8. Filippini N, MacIntosh BJ, Hough MG, Goodwin GM, Frisoni GB, et al. (2009) Distinct patterns of brain activity in young carriers of the APOE-epsilon 4 allele. *Proc Natl Acad Sci U S A* 106: 7209–7214. doi:10.1073/pnas.0811879106.
9. Filippini N, Rao A, Wetten S, Gibson RA, Borrie M, et al. (2009) Anatomically-distinct genetic associations of APOE epsilon4 allele load with regional cortical atrophy in Alzheimer's disease. *Neuroimage* 44: 724–728. doi:10.1016/j.neuroimage.2008.10.003.
10. Bookheimer SY, Strojwas MH, Cohen MS, Saunders AM, Pericak-Vance MA, et al. (2000) Patterns of brain activation in people at risk for Alzheimer's disease. *N Engl J Med* 343: 450–456. doi:10.1056/NEJM200008173430701.
11. Small GW, Ercoli LM, Silverman DH, Huang SC, Komo S, et al. (2000) Cerebral metabolic and cognitive decline in persons at genetic risk for Alzheimer's disease. *Proc Natl Acad Sci U S A* 97: 6037–6042. doi:10.1073/pnas.090106797.
12. Reiman EM, Chen K, Alexander GE, Caselli RJ, Bandy D, et al. (2004) Functional brain abnormalities in young adults at genetic risk for late-onset Alzheimer's dementia. *Proc Natl Acad Sci U S A* 101: 284–289. doi:10.1073/pnas.2635903100.
13. Bondi MW, Houston WS, Eyler LT, Brown GG (2005) fMRI evidence of compensatory mechanisms in older adults at genetic risk for Alzheimer disease. *Neurology* 64: 501–508. doi:10.1212/01.WNL.0000150885.00929.7E.
14. Wishart HA, Saykin AJ, McAllister TW, Rabin LA, McDonald BC, et al. (2006) Regional brain atrophy in cognitively intact adults with a single APOE epsilon4 allele. *Neurology* 67: 1221–1224. doi:10.1212/01.wnl.0000238079.00472.3a.
15. Frisoni GB, Laakso MP, Beltramello A, Geroldi C, Bianchetti A, et al. (1999) Hippocampal and entorhinal cortex atrophy in frontotemporal dementia and Alzheimer's disease. *Neurology* 52: 91–100.
16. Du A-T, Schuff N, Chao LL, Kornak J, Jagust WJ, et al. (2006) Age effects on atrophy rates of entorhinal cortex and hippocampus. *Neurobiol Aging* 27: 733–740. doi:10.1016/j.neurobiolaging.2005.03.021.
17. Filippini N, Rao A, Wetten S, Gibson RA, Borrie M, et al. (2009) Anatomically-distinct genetic associations of APOE epsilon4 allele load with regional cortical atrophy in Alzheimer's disease. *Neuroimage* 44: 724–728. doi:10.1016/j.neuroimage.2008.10.003.
18. Shaw P, Lerch JP, Pruessner JC, Taylor KN, Rose AB, et al. (2007) Cortical morphology in children and adolescents with different apolipoprotein E gene polymorphisms: an observational study. *The Lancet Neurology* 6: 494–500. doi:10.1016/S1474-4422(07)70106-0.
19. Barboriak DP, Doraiswamy PM, Krishnan KR, Vidyarthi S, Sylvester J, et al. (2000) Hippocampal sulcal cavities on MRI: relationship to age and apolipoprotein E genotype. *Neurology* 54: 2150–2153.
20. den Heijer T, Oudkerk M, Launer LJ, van Duijn CM, Hofman A, et al. (2002) Hippocampal, amygdalar, and global brain atrophy in different apolipoprotein E genotypes. *Neurology* 59: 746–748.
21. Burggren AC, Zeineh MM, Ekstrom AD, Braskie MN, Thompson PM, et al. (2008) Reduced cortical thickness in hippocampal subregions among cognitively normal apolipoprotein E epsilon 4 carriers. *Neuroimage* 41: 1177–1183. doi:10.1016/j.neuroimage.2008.03.039.
22. Schmidt H, Schmidt R, Fazekas F, Semmler J, Kapeller P, et al. (1996) Apolipoprotein E epsilon 4 allele in the normal elderly: neuropsychologic and brain MRI correlates. *Clin Genet* 50: 293–299.
23. Reiman EM, Uecker A, Caselli RJ, Lewis S, Bandy D, et al. (1998) Hippocampal volumes in cognitively normal persons at genetic risk for Alzheimer's disease. *Ann Neurol* 44: 288–291. doi:10.1002/ana.410440226.
24. Cherbuin N, Anstey KJ, Sachdev PS, Maller JJ, Meslin C, et al. (2008) Total and regional gray matter volume is not related to APOE*E4 status in a community sample of middle-aged individuals. *J Gerontol A Biol Sci Med Sci* 63: 501–504.
25. Wishart HA, Saykin AJ, Rabin LA, Santulli RB, Flashman LA, et al. (2006) Increased brain activation during working memory in cognitively intact adults with the APOE epsilon4 allele. *Am J Psychiatry* 163: 1603–1610. doi:10.1176/appi.ajp.163.9.1603.
26. Mondadori CRA, de Quervain DJ-F, Buchmann A, Mustovic H, Wollmer MA, et al. (2007) Better memory and neural efficiency in young apolipoprotein E epsilon4 carriers. *Cereb Cortex* 17: 1934–1947. doi:10.1093/cercor/bhl103.
27. Lind J, Persson J, Ingvar M, Larsson A, Cruts M, et al. (2006) Reduced functional brain activity response in cognitively intact apolipoprotein E epsilon4 carriers. *Brain* 129: 1240–1248. doi:10.1093/brain/awl054.
28. Han SD, Bondi MW (2008) Revision of the apolipoprotein E compensatory mechanism recruitment hypothesis. *Alzheimers Dement* 4: 251–254. doi:10.1016/j.jalz.2008.02.006.
29. Yu YW, Lin CH, Chen SP, Hong CJ, Tsai SJ (2000) Intelligence and event-related potentials for young female human volunteer apolipoprotein E epsilon4 and non-epsilon4 carriers. *Neurosci Lett* 294: 179–181.
30. Hubacek JA, Pitha J, Skodová Z, Adámková V, Lánská V, et al. (2001) A possible role of apolipoprotein E polymorphism in predisposition to higher education. *Neuropsychobiology* 43: 200–203.
31. Ravaja N, Räikkönen K, Lyytinen H, Lehtimäki T, Keltikangas-Järvinen L (1997) Apolipoprotein E phenotypes and cardiovascular responses to experimentally induced mental stress in adolescent boys. *J Behav Med* 20: 571–587.
32. Zetterberg H, Palmér M, Ricksten A, Poirier J, Palmqvist L, et al. (2002) Influence of the apolipoprotein E epsilon4 allele on human embryonic development. *Neurosci Lett* 324: 189–192.
33. Becher J-C, Keeling JW, McIntosh N, Wyatt B, Bell J (2006) The distribution of apolipoprotein E alleles in Scottish perinatal deaths. *J Med Genet* 43: 414–418. doi:10.1136/jmg.2005.033936.
34. Kitamura HW, Hamanaka H, Watanabe M, Wada K, Yamazaki C, et al. (2004) Age-dependent enhancement of hippocampal long-term potentiation in knock-in mice expressing human apolipoprotein E4 instead of mouse apolipoprotein E. *Neurosci Lett* 369: 173–178. doi:10.1016/j.neulet.2004.07.084.
35. Mahley RW (1988) Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science* 240: 622–630.
36. Han X (2007) Potential mechanisms contributing to sulfate depletion at the earliest clinically recognizable stage of Alzheimer's disease: a tale of shotgun lipidomics. *J Neurochem* 103 Suppl 1: 171–179. doi:10.1111/j.1471-4159.2007.04708.x.
37. Rose SE, Chen F, Chalk JB, Zelaya FO, Strugnell WE, et al. (2000) Loss of connectivity in Alzheimer's disease: an evaluation of white matter tract integrity with colour coded MR diffusion tensor imaging. *J Neurol Neurosurg Psychiatr* 69: 528–530.
38. Duan J-H, Wang H-Q, Xu J, Lin X, Chen S-Q, et al. (2006) White matter damage of patients with Alzheimer's disease correlated with the decreased cognitive function. *Surg Radiol Anat* 28: 150–156. doi:10.1007/s00276-006-0111-2.
39. Ferrarini L, Palm WM, Olofsen H, van der Landen R, Jan Blauw G, et al. (2008) MMSE scores correlate with local ventricular enlargement in the spectrum from cognitively normal to Alzheimer disease. *Neuroimage* 39: 1832–1838. doi:10.1016/j.neuroimage.2007.11.003.
40. Bartzokis G (2004) Age-related myelin breakdown: a developmental model of cognitive decline and Alzheimer's disease. *Neurobiol Aging* 25: 5–18; author reply 49–62.
41. Brun A, Englund E (1986) A white matter disorder in dementia of the Alzheimer type: a pathoanatomical study. *Ann Neurol* 19: 253–262. doi:10.1002/ana.410190306.
42. Acosta-Cabrero J, Williams GB, Pengas G, Nestor PJ (2010) Absolute diffusivities define the landscape of white matter degeneration in Alzheimer's disease. *Brain* 133: 529–539. doi:10.1093/brain/awp257.
43. Bosch B, Arenaza-Urquijo EM, Rami L, Sala-Llonch R, Junqué C, et al. (2010) Multiple DTI index analysis in normal aging, amnesic MCI and AD. Relationship with neuropsychological performance. *Neurobiol Aging* Available: <http://www.ncbi.nlm.nih.gov/pubmed/20371138>. Accessed 9 Nov 2010.
44. O'Dwyer L, Lambertson F, Bokde ALW, Ewers M, Fahuy YO, et al. (2011) Multiple Indices of Diffusion Identifies White Matter Damage in Mild Cognitive Impairment and Alzheimer's Disease. *PLoS ONE* 6: e21745. doi:10.1371/journal.pone.0021745.
45. Englund E (1998) Neuropathology of white matter changes in Alzheimer's disease and vascular dementia. *Dement Geriatr Cogn Disord* 9 Suppl 1: 6–12.
46. Englund E, Brun A (1990) White matter changes in dementia of Alzheimer's type: the difference in vulnerability between cell compartments. *Histopathology* 16: 433–439.
47. Nierenberg J, Pomara N, Hoptman MJ, Sidtis JJ, Ardekani BA, et al. (2005) Abnormal white matter integrity in healthy apolipoprotein E epsilon4 carriers. *Neuroreport* 16: 1369–1372.
48. Persson J, Lind J, Larsson A, Ingvar M, Cruts M, et al. (2006) Altered brain white matter integrity in healthy carriers of the APOE epsilon4 allele: a risk for AD? *Neurology* 66: 1029–1033. doi:10.1212/01.wnl.0000204180.25361.48.
49. Heise V, Filippini N, Ebmeier KP, Mackay CE (2010) The APOE [epsilon]4 allele modulates brain white matter integrity in healthy adults. *Mol Psychiatry* Available: <http://dx.doi.org/10.1038/mp.2010.90>. Accessed 22 Mar 2011.
50. Robnik-Šikonja M, Kononenko I (2003) Theoretical and Empirical Analysis of Relief and RRelief. *Mach Learn* 53: 23–69. doi:10.1023/A:1025667309714.
51. Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9: 97–113.
52. Delis DC, Kramer JH, Kaplan E, Ober BA (1987) California Verbal Learning Test: Adult version. San Antonio, Texas: Psychological Corporation.

53. Niemann H, Sturm W, Thone-Otto AIT, Willmes K (2008) California Verbal Learning Test (CVLT). Deutsche adaptation. Frankfurt: Pearson Clinical & Talent Assessment.
54. Benedict R (1997) BriefVisuospatial Memory Test-Revised professional manual. Odessa, FL: Psychological Assessment Resources, Inc.
55. Gold JM, Carpenter C, Randolph C, Goldberg TE, Weinberger DR (1997) Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia. *Arch Gen Psychiatry* 54: 159–165.
56. Wechsler D (1997) WMS-III administration and scoring manual. San Antonio: The Psychological Corporation.
57. Lehl S (1995) Mehrfachwahl-Wortschatz-Test (MWT-B). Erlangen: Straube.
58. Beck AT, Steer RA, Brown GK (1996) Beck Depression Inventory. Second ed. San Antonio, Texas: The Psychological Corporation.
59. Hautzinger M, Keller F, Kuhner C (2006) "BDI II" Beck Depressions-Inventar. Second ed. Frankfurt: Harcourt Test Services.
60. Smith SM (2002) Fast robust automated brain extraction. *Hum Brain Mapp* 17: 143–155. doi:10.1002/hbm.10062.
61. Smith SM, De Stefano N, Jenkinson M, Matthews PM (2001) Normalized accurate measurement of longitudinal brain change. *J Comput Assist Tomogr* 25: 466–475.
62. Smith SM, Zhang Y, Jenkinson M, Chen J, Matthews PM, et al. (2002) Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage* 17: 479–489.
63. Jenkinson M, Smith S (2001) A global optimisation method for robust affine registration of brain images. *Med Image Anal* 5: 143–156.
64. Jenkinson M, Bannister P, Brady M, Smith S (2002) Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17: 825–841.
65. Zhang Y, Brady M, Smith S (2001) Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* 20: 45–57. doi:10.1109/42.906424.
66. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, et al. (2006) Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31: 1487–1505. doi:10.1016/j.neuroimage.2006.02.024.
67. Nichols TE, Holmes AP (2002) Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp* 15: 1–25.
68. Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage* 44: 83–98. doi:10.1016/j.neuroimage.2008.03.061.
69. Frank E, Hall M, Trigg L, Holmes G, Witten IH (2004) Data mining in bioinformatics using Weka. *Bioinformatics* 20: 2479–2481. doi:10.1093/bioinformatics/bth261.
70. Witten I, Frank E, Hall M (2011) Data Mining (Third Edition). Boston: Morgan Kaufmann Available: <http://www.sciencedirect.com/science/article/pii/B9780123748560000183>. Accessed 17 Jun 2011.
71. Graña M, Termonen M, Savio A, Gonzalez-Pinto A, Echeveste J, et al. (2011) Computer aided diagnosis system for alzheimer disease using brain diffusion tensor imaging features selected by Pearson's correlation. *Neurosci Lett* 502: 225–229. doi:10.1016/j.neulet.2011.07.049.
72. O'Dwyer L, Lamberton F, Bokde ALW, Ewers M, Faluy YO, et al. (2012) Using Support Vector Machines with Multiple Indices of Diffusion for Automated Classification of Mild Cognitive Impairment. *PLoS ONE* 7: e32441. doi:10.1371/journal.pone.0032441.
73. Cui Y, Wen W, Lipnicki DM, Beg MF, Jin JS, et al. (2011) Automated detection of amnesic mild cognitive impairment in community-dwelling elderly adults: A combined spatial atrophy and white matter alteration approach. *NeuroImage* Available: <http://www.ncbi.nlm.nih.gov/pubmed/21864688>. Accessed 21 Nov 2011.
74. Haller S, Nguyen D, Rodriguez C, Emch J, Gold G, et al. (2010) Individual prediction of cognitive decline in mild cognitive impairment using support vector machine-based analysis of diffusion tensor imaging data. *J Alzheimers Dis* 22: 315–327. doi:10.3233/JAD-2010-100840.
75. Kononenko I, Šimec E, Robnik-Šikonja M (1997) Overcoming the myopia of inductive learning algorithms with RELIEFF. *Applied Intelligence* 7: 39–55.
76. Platt J (1999) Sequential minimal optimization: A fast algorithm for training support vector machines. *Advances in Kernel Methods-Support Vector Learning* 208: 98–112.
77. Scholkopf B, Sung KK, Burges CJ, Girosi F, Niyogi P, et al. (1997) Comparing support vector machines with Gaussian kernels to radial basis function classifiers. *Signal Processing, IEEE Transactions on* 45: 2758–2765.
78. Noble WS (2006) What is a support vector machine? *Nat Biotechnol* 24: 1565–1567. doi:10.1038/nbt1206-1565.
79. Bishop C (1991) Improving the Generalization Properties of Radial Basis Function Neural Networks. *Neural Computation* 3: 579–588. doi:10.1162/neco.1991.3.4.579.
80. Behrens TEJ, Woolrich MW, Jenkinson M, Johansen-Berg H, Nunes RG, et al. (2003) Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn Reson Med* 50: 1077–1088. doi:10.1002/mrm.10609.
81. Behrens TEJ, Berg HJ, Jbabdi S, Rushworth MFS, Woolrich MW (2007) Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* 34: 144–155. doi:10.1016/j.neuroimage.2006.09.018.
82. Zhang Y, Ding C, Li T (2008) Gene selection algorithm by combining reliefF and mRMR. *BMC Genomics* 9 Suppl 2: S27. doi:10.1186/1471-2164-9-S2-S27.
83. Supek F, Peharec P, Krsnik-Rasol M, Smuc T (2008) Enhanced analytical power of SDS-PAGE using machine learning algorithms. *Proteomics* 8: 28–31. doi:10.1002/pmic.200700555.
84. Dennis NA, Browndyke JN, Stokes J, Need A, Burke JR, et al. (2010) Temporal lobe functional activity and connectivity in young adult APOE varepsilon4 carriers. *Alzheimers Dement* 6: 303–311. doi:10.1016/j.jalz.2009.07.003.
85. Filippini N, Ebmeier KP, Macintosh BJ, Trachtenberg AJ, Frisoni GB, et al. (2010) Differential effects of the apoe genotype on brain function across the lifespan. *Neuroimage* Available: <http://www.ncbi.nlm.nih.gov/pubmed/20705142>. Accessed 17 Aug 2010.
86. Han SD, Houston WS, Jak AJ, Eyler LT, Nagel BJ, et al. (2007) Verbal paired-associate learning by APOE genotype in non-demented older adults: fMRI evidence of a right hemispheric compensatory response. *Neurobiol Aging* 28: 238–247. doi:10.1016/j.neurobiolaging.2005.12.013.
87. Bondi MW, Wierenga CE, Stricker JL, Houston WS, Eyler LT, et al. (2006) Functional connectivity comparisons of learning by APOE genotype among nondemented older adults. *Alzheimer's and Dementia* 2: S308–S309. doi:10.1016/j.jalz.2006.05.1085.
88. Wierenga CE, Stricker NH, McCauley A, Simmons A, Jak AJ, et al. (2010) Increased functional brain response during word retrieval in cognitively intact older adults at genetic risk for Alzheimer's disease. *Neuroimage* 51: 1222–1233. doi:10.1016/j.neuroimage.2010.03.021.
89. Cabeza R (2002) Hemispheric asymmetry reduction in older adults: the HAROLD model. *Psychol Aging* 17: 85–100.
90. Davis SW, Kragel JE, Madden DJ, Cabeza R (2012) The architecture of cross-hemispheric communication in the aging brain: linking behavior to functional and structural connectivity. *Cereb Cortex* 22: 232–242. doi:10.1093/cercor/bhr123.
91. Dennis NA, Kim H, Cabeza R (2007) Effects of aging on true and false memory formation: an fMRI study. *Neuropsychologia* 45: 3157–3166. doi:10.1016/j.neuropsychologia.2007.07.003.
92. Morcom AM, Good CD, Frackowiak RSJ, Rugg MD (2003) Age effects on the neural correlates of successful memory encoding. *Brain* 126: 213–229.
93. Reuter-Lorenz PA, Jonides J, Smith EE, Hartley A, Miller A, et al. (2000) Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. *J Cogn Neurosci* 12: 174–187.
94. Vincent JL, Snyder AZ, Fox MD, Shannon BJ, Andrews JR, et al. (2006) Coherent Spontaneous Activity Identifies a Hippocampal-Parietal Memory Network. *J Neurophysiol* 96: 3517–3531. doi:10.1152/jn.00048.2006.
95. Fan Y, Rao H, Hurt H, Giannetta J, Korczykowski M, et al. (2007) Multivariate examination of brain abnormality using both structural and functional MRI. *Neuroimage* 36: 1189–1199. doi:10.1016/j.neuroimage.2007.04.009.
96. Wee C-Y, Yap P-T, Li W, Denny K, Browndyke JN, et al. (2011) Enriched white matter connectivity networks for accurate identification of MCI patients. *Neuroimage* 54: 1812–1822. doi:10.1016/j.neuroimage.2010.10.026.
97. Buckner RL, Sepulcre J, Talukdar T, Krienken FM, Liu H, et al. (2009) Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *J Neurosci* 29: 1860–1873. doi:10.1523/JNEUROSCI.5062-08.2009.
98. Salat DH, Tuch DS, van der Kouwe AJW, Greve DN, Pappu V, et al. (2010) White matter pathology isolates the hippocampal formation in Alzheimer's disease. *Neurobiol Aging* 31: 244–256. doi:10.1016/j.neurobiolaging.2008.03.013.
99. Seeley WW, Crawford RK, Zhou J, Miller BL, Greicius MD (2009) Neurodegenerative diseases target large-scale human brain networks. *Neuron* 62: 42–52. doi:10.1016/j.neuron.2009.03.024.
100. Fouquet M, Desgranges B, Landeau B, Duchesnay E, Mézenge F, et al. (2009) Longitudinal brain metabolic changes from amnesic mild cognitive impairment to Alzheimer's disease. *Brain* 132: 2058–2067. doi:10.1093/brain/awp132.
101. Nestor PJ, Fryer TD, Smielewski P, Hodges JR (2003) Limbic hypometabolism in Alzheimer's disease and mild cognitive impairment. *Ann Neurol* 54: 343–351. doi:10.1002/ana.10669.