METHODOLOGICAL IMPROVEMENTS IN VOXEL BASED ANALYSIS OF DIFFUSION TENSOR IMAGES: APPLICATIONS TO STUDY THE IMPACT OF APOE ON WHITE MATTER INTEGRITY

Shawn M. Newlander, MS[1], **Alan Chu, BS**[1], **Usha S. Sinha, PhD**[1], **Po H Lu, PsyD**[2], and **George Bartzokis, MD**[3,4]

[1]Department of Physics, San Diego State University, San Diego, CA 92182

[2]Mary S. Easton Center for Alzheimer's Disease Research, University of California, Los Angeles, CA 90095

[3]Department of Psychiatry, The David Geffen School of Medicine at UCLA, Los Angeles, CA 90095

[4]Brain Research Institute, University of California, Los Angeles, CA 90095

Abstract

Purpose—To identify regional differences in apparent diffusion coefficient (ADC) and fractional anisotropy (FA) using customized preprocessing prior to voxel-based analysis (VBA) in 14 normal subjects with the specific genes that decrease (APOE ε2) and that increase (APOE ε4) the risk of Alzheimer's Disease.

Materials and Methods—Diffusion Tensor images (DTI) acquired at 1.5 T were denoised with a total variation tensor regularization algorithm prior to affine and nonlinear registration to generate a common reference frame for the image volumes of all subjects. Anisotropic and isotropic smoothing with varying kernel sizes was applied to the aligned data prior to VBA to determine regional differences between cohorts segregated by allele status.

Results—VBA on the denoised tensor data identified regions of reduced FA in APOE ε4 compared to the APOE ε2 healthy older carriers. The most consistent results were obtained using the denoised tensor and anisotropic smoothing prior to statistical testing. In contrast, isotropic smoothing identified regional differences for small filter sizes alone, emphasizing that this method introduces bias in FA values for higher kernel sizes.

Conclusion—Voxel-based DTI analysis can be performed on low SNR images to detect subtle regional differences in cohorts using the proposed pre-processing techniques.

Keywords

Diffusion Tensor Imaging; Voxel Based Analysis; Tensor Denoising; Anisotropic smoothing; APOE

Introduction

The advent of volumetric, high resolution MRI coupled with high throughput imaging has resulted in an unprecedented amount of data. The development of automated analysis of

Corresponding Author Usha Sinha, Ph.D., 5500 Campanile Drive, Department of Physics, San Diego State University, San Diego, CA 92182-1233, usinha@sciences.sdsu.edu, (619) 594-1791 (phone), (619) 594-5485 (fax).

differences in morphology and more recently, of diffusion indices, between population cohorts is critical to the accurate analysis of subtle differences. A widely used automated method is termed 'voxel based morphometry' (VBM) that detects differences in gray matter in population cohorts (1). An extension of this method is 'voxel based analysis' (VBA) in which image volumes of MRI derived indices (T1, T2, ADC, FA) are analyzed to detect differences in these indices at the voxel level (2, 3, 4). Many studies have focused on detecting voxel based differences in the DTI indices in population cohorts segregated by age, genetic influences, and disease (5, 6, 7). Recent studies have also extended this to identifying alterations in brain fiber pathways in disease states using diffusion tensor based tractography (8). Most of these studies to date use protocols that are optimized to yield a high signal to noise and are thus acquired with a high number of averages and/or at higher fields (3 Tesla). However, an overwhelming number of clinical studies are acquired at 1.5T under constraints of clinically relevant scan times. These routinely acquired clinical scans are difficult to process with automated techniques due to the poorer image quality. The ability to analyze these imaging studies using VBA will enable automated techniques to be available in the clinical setting.

Though a number of studies have explored VBA to quantify white matter integrity in normal aging and in disease conditions, many technical questions relating to optimal denoising and smoothing are still unresolved. A general rule of thumb that arose from gray scale morphometry and functional MR studies is that the full width half maximum of the smoothing kernel should be of the order of 2 to 3 times the voxel size. Extending this to DTI analysis requires Gaussian smoothing kernels of 6 to 12 mm or so. However, such large smoothing kernels and the high contrast between gray and white matter on FA images can lead to false positives. More recently, two approaches based on anisotropic diffusion edge preserving filters have been suggested as alternates to Gaussian smoothing (9, 10). However, while anisotropic smoothing is effective in reducing the noise in the data as well as in decreasing the number of independent voxels in the analysis, it does not address the effect of residual misregistration. Anisotropic smoothing is designed to preserve edges while smoothing homogeneous regions. The Gaussian filter, on the other hand, smooths across edges and homogeneous regions. Thus, if there are errors in structure (edges) alignment, the Gaussian smoothing will be able to blur these edges and reduce the error in misregistration. The anisotropic filter preserves edges and will not be able to mitigate the effects of misregistration as smoothing does not occur across these edges. However, it should be noted that though the anisotropic filter is better than the Gaussian filter in preserving edges, some amount of edge blurring occurs even with the anisotropic filter. Thus the anisotropic filter may also address small misregistration effects.

The APOE ε4 allele has been identified as a risk factor for developing Alzheimer's Dementia (AD). Neuro-imaging studies focused on gray matter changes in healthy elderly subjects with and without the ε4 allele report hippocampal atrophy in one study (11), and in another, lower gray matter density in the right medial temporal and bilateral fronto-temporal regions associated with the ε4 allele (12). A recent tensor based morphometry longitudinal study revealed that the APOE ε4 genotype is associated with greater temporal and hippocampal atrophy rates in healthy elderly adults well before the onset of cognitive deficits (13). In addition to investigations of gray matter changes, a few studies have also focused on white matter changes. In the sequence of events leading to cognitive decline in degenerative disorders such as AD, it has been postulated that white matter changes precede changes in gray matter and may thus provide early markers of AD (14, 15). DTI studies have reported APOE ε4 dependent differences in the posterior corpus callosum, occipito-fronto fasciculus, left hippocampus, and in the parahippocampal gyrus (16). A recent study combined gray matter (GM) voxel based morphometry and Tract Based Spatial Statistics (TBSS) of white matter FA and revealed APOE dependent differences consisting of atrophy

in the hippocampus and amygdala and reduced FA in the left parahippocampal gyrus WM (17).

This paper explores the feasibility of VBA on whole brain DTI studies acquired in 3 minutes at 1.5 Tesla to investigate differences in white matter integrity in a non-symptomatic healthy elderly cohort segregated by the ε4 allele. The focus is on (i) evaluating a novel tensor anisotropic denoising algorithm to reduce noise in the diffusion tensors and (ii) comparing Gaussian and anisotropic image smoothing with different kernels applied prior to the statistical analysis. The hypothesis is that the initial anisotropic tensor smoothing will reduce noise and voxel independence sufficiently, so that a small kernel for isotropic smoothing can be used prior to statistical analysis to correct for registration mismatch alone. The goal is to develop robust pre-processing methods to support VBA of DTI studies acquired in clinically relevant scan times.

Methodology

Subject Details

Fourteen subjects imaged as part of a clinical DTI study exploring white matter integrity of healthy subjects who are carriers of APOE ε4 compared to non-carriers were used in the current analysis. All subjects received written and oral information about the study and signed written informed consents approved by the local institutional review board (University of California at Los Angeles IRB) prior to study participation. The APOE profile of the subjects were as follows: ε4ε4: 1, ε3ε4: 6, ε2ε3: 7 who were grouped as ε4 carriers (ε4ε4 and ε3ε4) or ε4 non-carriers (ε2ε3). The mean age for all the study subjects was 72.1 years with a mean age of 76.2 ± 6.7 and 69.1 ± 5.5 years for the APOE ε2 and APOE ε4 subjects respectively; this age difference was not significant (p=0.10). Further, we confirmed with VBA on age segregated (and allele status matched) cohorts that there were no age related differences in FA and ADC. All subjects denied problems with memory, were living independently, and had no evidence of functional deficits at the time of scanning and their mean MMSE score was 27.2 (SD=2.1). A single subject refused cognitive testing however he had no change in function or complaints on a follow-up visit 3 years later.

Image Acquisition

Imaging studies were performed on a Siemens 1.5 Tesla Avanto scanner in a 12 channel head coil; parallel imaging was not used. High resolution T1 weighted anatomical images were acquired using the magnetization-prepared rapid gradient echo (MPRAGE) sequence with the following parameters: $1 \times 1 \times 1$ mm^3 voxels, TE/TR/TI:: 2.49 ms / 2000 ms/ 900 ms, field of view 256 (RO)×192 (PE), flip angle = 8° and processed for voxel-based morphometry. The DTI sequence was a single shot EPI sequence with dual 180° pulses for eddy current reduction. Twelve non-collinear gradient directions with a b-factor of 800 s/ mm^2 were used to map the direction dependent diffusion. Imaging parameters were TE/TR: 104 ms/ 5900 ms with 2 signal averages. Images were not averaged in the scanner but were stored separately resulting in 26 volumes, two of them were b=0 s/mm^2 volumes and the rest, diffusion weighted. All image volumes were registered to the baseline image to correct for eddy current and motion related artifacts. The field of view, slice thickness, and acquisition matrix were 240×240 mm^2, 3 mm, and 128×128 respectively; yielding a voxel resolution of 1.875 mm × 1.875 mm × 3.00 mm (clinical imaging with limitation on scan time will not support acquisition of diffusion weighted images with isotropic resolution).

Preprocessing and DTI analysis

Corrections for the eddy current induced distortions were made using FSL software (18). The images were then processed using the FSL Brain Extraction Tool (BET) in order to

remove the non-brain tissue (19). The BET extracted, eddy current/motion corrected image volumes were then fit on a voxel basis to the diffusion tensor equation (20).

Tensor Denoising

A total variation (TV) regularization method for image noise reduction was incorporated to enhance the SNR ratio in the diffusion tensor images which inherently have low SNR. In the tensor regularization, the total variation of a scalar quantity is extended to the 3 × 3 tensor, D, and the algorithm ensures the positive definiteness of D (21). The tensor, D, is treated implicitly as the product:

$$D = LL^T$$

with L being the lower triangular matrix. This guarantees the symmetry, positive definiteness, and orthogonality of the eigenvectors which is required by the diffusion tensor model (21). The regularization consists of solving a minimization problem in terms of l_{kl}. For each l_{kl}, the following function is minimized:

$$\min \left\{ \sqrt{\sum_{ij} TV[d_{ij}(l_{kl})]^2} + \frac{\lambda}{2} \sum_{ij} \left\| d_{ij} - \hat{d}_{ij} \right\|_2^2 \right\}$$

where l_{kl} are the lower triangular matrix; the subscript d_{ij} denotes the tensor component; $\hat{d}_{ij}$ refers to the noisy tensor component; and λ is the fidelity term. The concept of the total variation algorithm for denoising is to smooth the image gradient (noise is emphasized in the gradient images) to suppress noise (first term in TV equation). However, the true edges in the image also have a large contribution to the gradient image. In order to preserve the edges, it is important to check that the smoothed image does not deviate too much from the noisy image. The second term in the TV equation refers to the 'fidelity' term which ensures that the image edges are not smoothed. The relative contribution of the two terms is adjusted by the constant, λ, which is determined by the relative strengths of the edge gradients and the noise terms (small values of λ result in more smoothing). The numerical implementation of the algorithm uses a steepest descent gradient algorithm with fixed time step.

Apparent Diffusion Coefficient (ADC) and Fractional Anisotropy (FA) images were calculated from the original and the denoised tensors. The tensor denoising was evaluated for the extent of smoothing and introduction of bias into mean ADC and FA values. For quantitative assessment of the denoising algorithm performance, an average deviation angle (ADA) index was used. The ADA is a measure of the lead eigenvector (angular) similarity in the immediate neighborhood of a pixel and is calculated from (22):

$$ADA = \frac{\Delta\alpha_{i-1} + \Delta\alpha_{i+1} + \Delta\alpha_{j-1} + \Delta\alpha_{j+1} + \Delta\alpha_{k-1} + \Delta\alpha_{k+1}}{6}$$

where,

$$\Delta\alpha_{i-1} = \cos^{-1}(|V_{ijk}, V_{i-1jk}|)$$

Here, V_{ijk} is the lead eigenvector direction at location i,j,k. For a noise free homogeneous region, all the eigenvectors will have the same direction and the angle between these

eigenvectors will be close to zero. The ADA is calculated in 3D using the average of the six nearest neighbors of a given voxel. In conformance with Chen et al (22), only the nearest neighbors were used in the ADA computation in order to ensure that regions with different fiber orientations were not included. A global ADA for all brain voxels was determined as well as in smaller ROI's for each subject before and after denoising. A FA threshold of (FA<0.25) was used to remove gray matter voxels in the determination of the global ADA.

Generation of a common reference frame and atlas creation

In order to perform VBA, a common reference frame of reference is required. This was generated by creating an inter-subject DTI atlas by non-linear alignment of the ADC and FA images of all 14 subjects to a common reference space. The subjects' FA and ADC images were first interpolated to isotropic voxels (1.875 mm × 1.875 mm × 1.875 mm) using a cubic b-spline interpolation. A reference ADC and FA image was randomly selected from the 14 subjects and each subject was aligned to the reference by first performing an affine transformation using the FSL Linear Registration Image Tool (FLIRT) (18). To create an atlas of the cohorts, each affine matched image volume was then aligned to the reference subject using a nonlinear registration algorithm (23). A dual channel nonlinear multi-scale registration algorithm based on optical flow is implemented in an iterative hierarchical manner and calculates the displacement from ADC and FA image volumes (23):

$$\nu_{n+1} = G_\sigma \otimes \left(\nu_n + \frac{1}{2} \sum_{i=1}^{2} \left[\frac{C(T_i - S_i) \|\nabla S_i\| \|\nabla T_i\|}{(\|\nabla T_i\|^2 + \|\nabla S_i\|^2)(\|\nabla S_i\|^2 + \|\nabla T_i\|^2 + 2(T_i - S_i)^2)} \nabla S_i \right] \right)$$

where ν_{n+1} is the deformation vector field at iteration n+1, $G\sigma$ is the Gaussian filter with variance σ, $\otimes$ denotes the convolution, i represents the two channels (ADC map and FA map), C is the scaling factor and T and S are the target and transformed images respectively. A Gaussian filter of 5 × 5 was used along with a scaling factor of 16. The hierarchical implementation is accomplished by subsampling and starting the registration with an image reduced by a factor of 8; at each level the algorithm performs preset number of iterations with more at the lower resolution and less at the higher resolution. The equation above includes the average of deformation fields calculated from two channels (i=1,2 in the summation corresponds to the ADC and FA channels). The morphology in the atlas is preserved by computing the forward and backward deformation vectors. The average deformation field is applied to the average intensity image to create an average shape atlas which encodes the average intensity. After the first iteration, the reference subject is replaced with the average shape atlas which after a few iterations converges to the centroid of the population. The convergence criterion was based on the magnitude of the deformation field vector summed over all the voxels in the brain and all subjects to provide a global measure which is a minimum when the atlas converges to the centroid. The iterations were stopped when the index did not change by 5% from the previous trial; this was determined from empirical observations that the index fluctuated ±5% when convergence was reached.

The ADC and FA atlases were created using both the original ADC and FA volumes and the denoised ADC and FA volumes from each subject and evaluated both qualitatively and quantitatively. Qualitative evaluation included a visual inspection of (i) anatomical contours traced on the atlas overlaid on individual images and (ii) alignment of smaller fiber tracts.

Voxel Based Analysis

VBA was performed on the DTI indices calculated from the original images and denoised images using SPM8 (Wellcome Department of Cognitive Neurology, London, UK) software. Prior to VBA, the FA and ADC images were smoothed to ensure normal

distribution of the data, to decrease the number of independent comparisons, as well as to account for limitations in the registration accuracy. The effects of isotropic and anisotropic smoothing at different filter sizes were evaluated; the different combinations are shown in the flowchart (Figure 1). For isotropic smoothing, a three-dimensional Gaussian filter provided with SPM8 was used. An edge preserving anisotropic diffusion filter implemented in house was used for anisotropic smoothing (24). Smoothing was performed with a full-width half maximum (FWHM) of 0, 3, 6, 9 and 12 mm for both the Gaussian and anisotropic diffusion filter. VBA was performed on the aligned ADC and FA volumes using a two-sample (APOEε4 and ε2) one-tailed t-test. In order to account for multiple comparisons, family-wise error was used with a p-value < 0.05 and a cluster extent threshold (k) of 30 voxels. The two groups (APOE ε4 andε2) were not age matched, so an additional analysis of age related effects on FA and ADC was performed on cohorts separated by age (threshold of 74 years to form two equal cohorts with equal distribution of ε2 and ε4 in each group).

Voxel Based Morphometry

Morphometric analysis was conducted on the high resolution structural MR images using SPM8 (Wellcome Department of Cognitive Neurology, London, UK). Standard VBM steps were used: segmentation was performed using the FSL FAST (FMRIB's Automated Segmentation Tool), followed by affine registration to a reference image from the population cohort and realignment, and finally smoothing with a 8 mm FWHM Gaussian kernel. Additionally, all images were visually checked after segmentation and registration for inaccuracies. Gray matter maps of APOEε4 carrier and non-carrier subjects were compared to identify areas of ε4-related gray matter difference using a 2-sample t-test with age as confounding variable. SPM implements the General Linear Model (GLM) which accommodates confounding variables (25). It is a well-accepted practice in comparing imaging data of different cohorts to integrate the confounding variable as a covariate in the SPM statistical model (26); in the current paper, age was included as 1×14 vector.

Results

SNR was measured in two anatomical locations in the baseline (b0) and diffusion weighted images (b=800 s/mm^2). The SNR (b0/b=800 s/mm^2) of ventricles was 20.6/4.0 and that of the corpus callosum was 12.6/7.7.

Denoising

The optimum value of the fidelity constant of the TV algorithm for the current data was 15 which maximized SNR while preserving edges. The computational time to denoise a single volume of tensors and to calculate the ADC and FA maps was 120 minutes on a Macintosh with a Quad Core processor and 2 Giga-bytes of RAM. Figure 2 shows an example of the results of the denoising algorithm; ADC and FA images are shown at the mid-brain level for one of the subjects. The variance in FA maps is very sensitive to the noise in the baseline and diffusion weighted image volumes. This can be seen in the zoomed FA inset of the original images and the reduction in variance with denoising can be appreciated (Figure 2, inset).

ROI (4 pixels) measurements of ADC and FA values were taken in the corpus callosum (genu and callosum) in the original and denoised images to evaluate the denoising algorithm. Paired t-test of the mean ADC and FA before and after denoising showed no statistically significant differences in the data (ADC: p= 0.375 (genu), p=0.537 (splenium); FA: p=0.359 (genu), p=0.373 (splenium)). This confirms that the denoising does not alter the values of the diffusion indices. A one-tailed F-test showed that the variance of the ADC

and FA measurements from corresponding noisy and denoised ROI data was significantly different for ADC and FA of the genu and splenium at p=0.05. This confirms that the tensor denoising reduces variance in homogeneous regions.

Table 1 lists the ADA values in select ROIs from an axial slice (ROIs were chosen in the splenium and in the right and left posterior limb of the internal capsule) as well as in the whole brain. The ADA of the denoised data is reduced compared to that from the original data. As the ADA is a measure of the orientational coherence of the lead eigenvector, a lower value implies less variance in the eigenvector orientation. Further, a paired t-test between original and denoised image volumes yielded a p-value < 0.001 for global and regional ADA values which confirms that the reduction in ADA on denoising is significant.

Atlas

The atlas converged to the centroid of the population in five iterations and the computational time was approximately 40 minutes running on a Macintosh with a quad core processor with 2 Giga-bytes of RAM. Separate atlases were created using the original and the denoised datasets. Figure 3 shows one slice from the ADC and FA atlases created from the original data (top row), and corresponding slices from the denoised data are shown in the bottom row. In both sets of atlases, the accuracy of the non-linear registration can be qualitatively appreciated by the sharp edges of even the smaller white matter tracts.

The standard deviation of FA in the atlas (variance in values across subjects at each voxel) will ultimately determine the ability of VBA to localize differences between cohorts. Figure 4 shows the standard deviation of FA in atlas voxels (original: row 1, denoised: row 3) which are color coded red if they exceed a standard deviation of 10% of the mean. Clearly, both original and denoised data contain several voxels that exceed 10%; however the original data has substantially more voxels exceeding the 10% threshold. The effect of smoothing with a 3 mm × 3 mm × 3 mm Gaussian smoothing filter is clearly evident in the reduction of the number of such voxels in both datasets (rows 2 and 4). This underlines the need for a smoothing filter prior to VBA. However, the denoised data with smoothing shows the least variance (compare row 4 with other rows, Figure 4).

Voxel Based Analysis

Voxel based analysis was conducted for various combinations of isotropic and anisotropic smoothing with different kernel sizes using the original and denoised data. Figure 5 is a plot of the FA from a ROI in the genu for one of the subjects to illustrate the effect of anisotropic vs. isotropic smoothing (FWHM: 0 to 12 mm) on FA values. The FWHM of the anisotropic filter was calculated from κ in the anisotropic diffusion equation (exp- $|\nabla|/\kappa$). The standard deviation, σ is related to κ as: $\sigma = \kappa/\sqrt{2}$ and the FWHM of the filter is given by: 2.354 σ. The graph demonstrates that the Gaussian smoothing drastically reduces the FA values even for moderate values of the FWHM. Figure 6 shows example images of one subject smoothed with isotropic and anisotropic smoothing, both with a FWHM of 6 mm. The images clearly demonstrate that the anisotropic smoothing is able to preserve edges within the image while the Gaussian smoothing blurs the edges. Preliminary analysis was conducted to explore differences in ADC and FA in cohorts segregated by the allele APOE ε4. The only analysis that yielded regions with significant differences was FA of APOE ε2 subjects greater than that of APOE ε4 subjects. Figure 7 shows the region with significant difference in FA consistently in the genu of the corpus callosum (CC) and in the brain stem across the range of kernel sizes for the anisotropic diffusion filter. Table 2 summarizes the results of the VBA analysis using the two types of smoothing at different FWHM. Only results for the denoised data are shown here as the original (noisy) data did not show any regions of difference between the cohorts for the filters listed in Table 2. Comparing the different

smoothing filters, the denoised data showed voxels of significant differences when processed with smaller Gaussian smoothing kernels or with anisotropic smoothing kernels to the maximum kernel size (Table 2). It should be noted that Table 2 lists the cluster size for the voxels at the highest significance within the cluster identified with the criteria of FWE p<0.05 and cluster size>30.

Voxel Based Morphometry

ε4-carriers had a trend for lower gray matter regional volumes in the right hippocampus and in the posterior cingulate gyrus (p < 0.01, uncorrected, Figure 8). The uncorrected test only allowed inference on trends; a larger sample size will be required to reach significance when correction for multiple comparisons is included. No regions of significant differences were found for the test of ε4 non-carriers with lower gray matter regional volumes than ε4-carriers.

Discussion

Voxel based analysis of diffusion tensor derived indices has been increasingly employed to automatically localize regions with differences between two population cohorts (7). The VBA method is a powerful analysis tool but has to be used with an awareness of its potential pitfalls (20). The success of the method lies in the accuracy of the registration used in aligning the diffusion tensor images to a common frame of reference. Zhang et al performed a comparison of three normalization procedures and found that the high dimensional normalization of FA images was only marginally less accurate than the computationally intensive high dimensional tensor normalization (27). In the current paper, a high dimensional nonlinear registration algorithm that uses FA and ADC image volumes is employed. The inclusion of the ADC channel as opposed to the more routinely used FA channel alone increased the accuracy of the registration especially in the ventricle and CSF/ gray matter interface regions. The choice of the ADC and FA channels (as opposed to the eigenvalues) is based on the high contrast between brain/ CSF in ADC and the high contrast between white/gray matter and white matter/CSF in the FA images. A combination of these two channels thus provided the best discrimination between white, gray and CSF to drive the intensity based registration. The accuracy of the registration employed in the current work was tested earlier and showed registration accuracies to within a voxel in the splenium and body but not in the genu (23). The decreased accuracy at the level of the genu arose from residual distortions near the frontal lobes due to susceptibility artifacts at 3T (23). Given that the current data was acquired at 1.5 T which has smaller artifacts from susceptibility differences, the registration accuracy for the current dataset is closer to 2 mm (voxel resolution of 1.875 mm) globally. This is confirmed visually by the sharpness of the smaller white matter fibers in the atlas (Figure 3).

The next step in VBA that determines the detection sensitivity of regional differences is the smoothing prior to the statistical analysis. Smoothing is performed for four reasons: to increase the SNR by averaging neighborhood pixels, to compensate for residual misalignment, to reduce the number of multiple comparisons, and to make the data more normally distributed which ensures the assumptions of the Gaussian general linear model underlying the statistical inferences (1). Most DTI studies have employed a range of FWHM isotropic smoothing from 6 mm to 12 mm (28). However, as shown in Figure 6, isotropic smoothing with kernel sizes of 6 mm and greater results in a large reduction in FA values as data is averaged across WM (high FA) with structures of low FA (GM and CSF). This has serious implications in the ability to detect regional differences. Jones et al performed VBA comparing FA of schizophrenic subjects to healthy controls varying the size of the smoothing kernel (28). Their findings showed that the detection sensitivity and the extent of the areas of abnormality were a function of filter size. However, it should be noted that

Jones et al used a low dimensional registration algorithm that may have introduced complex effects from the presence of misregistration (28). Further, Jones et al found the cluster size to increase with kernels up to 16 mm FWHM and this may be a false positive effect that is a consequence of isotropic smoothing with large kernel sizes. As shown in Figure 5, even a kernel size of 6 mm reduces the FA by more than 50 %.

There have been a few attempts to address the problems related to smoothing. In one study by Lee et al., white matter was segmented prior to the spatial normalization and this reduces partial volume effects that are likely to occur during resampling of the data after registration (29). After spatial normalization, the smoothed DTI data (FA and ADC) are then compensated by dividing them by the smoothed WM mask. This compensation allowed for values in the smoothed data to be closer to the original value and reduces the potential confounds of morphological differences. However, an important assumption is that white matter segmentation errors are negligible and do not contribute to errors in the VBA analysis.

Two other approaches address smoothing by replacing the isotropic Gaussian smoothing by anisotropic smoothing. Moraschi et al. tested isotropic smoothing vs. anisotropic smoothing comparing AD subjects with healthy controls (9). The anisotropic smoothing used in the latter paper performed smoothing individually for each diffusion weighted image. However, the method does not guarantee a positive definite tensor, as does the TV algorithm proposed in the current paper. Comparing anisotropic to Gaussian filter, they show that the anisotropically denoised data showed cohort differences comparable to their standard reference. However, their statistical analysis is limited to uncorrected p<0.001 and cluster size of 19 voxels. Even these clusters (using the anisotropic smoothing filter) may not have been present if a correction for multiple comparisons was included. The second article by Hecke et al. presents a comprehensive comparison of Gaussian vs. non-anisotropic smoothing using simulated data (10). Anisotropic and isotropic smoothing was applied to the FA maps just prior to the statistical testing. The atlas used in the study was based on a high dimensional full tensor reorientation which yields the most accurate alignment. They also simulated misalignments up to 6 mm and examined both types of smoothing. The anisotropic smoothing out-performed the isotropic smoothing even in the presence of the simulated misalignments. But it should be noted that anisotropic smoothing does not satisfy the requirement of smoothing to reduce registration inaccuracies. This resulted in a decreased sensitivity and specificity of the anisotropic smoothing for the larger simulated registration inaccuracies.

The present work is an attempt to combine both types of smoothing: anisotropic and isotropic. The anisotropic tensor smoothing increases SNR, reduces the number of multiple comparisons, and makes the data more normally distributed. The residual misalignment from the registration is addressed by a small isotropic filter prior to the statistical analysis. The effectiveness of this proposed combination is seen in Table 2 where tensor denoised data smoothed by a $3\times3\times3$ mm^3 isotropic filter shows the regional differences at the genu. However, the ability to detect regional differences is lost for Gaussian smoothing filters greater than 3 mm. This may arise from the drastic reduction in FA when using larger size filters.

Comparing the effectiveness of the TV and anisotropic diffusion filters, it should be noted that the TV reduced the noise in the tensor whereas the anisotropic filter works only on the scalar FA values. The tensor denoising reduces the noise in all tensor components and is thus more effective in reducing noise than the scalar anisotropic denoising filter. Further, the denoising TV algorithm is more effective in preserving edges than the anisotropic diffusion filter. This can be seen in Figure 5 where the FA values are reduced due to blurring across

edges for the anisotropic filter; however the results of ROI measurements in the corpus callosum showed that the tensor denoising does not introduce any significant bias in the FA values.

Surprisingly, the incorporation of anisotropic smoothing filter prior to statistical testing was also able to identify regional differences in FA between the two cohorts. It was anticipated that since an anisotropic filter at this stage does not address the residual misalignments, it will be less effective in identifying regional differences compared to a small kernel isotropic filter. One reason for the current finding may arise from the fact that the registration used here is accurate to a voxel, so that smoothing to remove registration inaccuracies is not required. It remains to be seen if there will be a different response if a low dimensional registration is used. An important result of this analysis is that regional differences persist even when the anisotropic diffusion filter kernel size is increased. This is clearly related to the observation that the anisotropic filter does not significantly bias the FA values even for large filters (Figure 5).

Another important finding from this analysis is that no regional differences were found in the noisy (original) data for any of the smoothing filters (Gaussian or anisotropic) applied prior to statistical testing. For isotropic filtering, this can be explained by the fact that the large size filters required to increase SNR (for the 'noisy' data) cause a significant reduction in FA. However, the negative finding for the anisotropic filter is harder to explain as the larger filters should have been able to increase SNR to a level where differences are detectable. A possible explanation is that the tensor anisotropic denoising applied prior to registration is more effective than the scalar anisotropic filter applied to the FA images. These clearly indicate the efficacy of the tensor denoising for VBA analysis of clinical image data that have a low SNR.

Few studies have been reported comparing the WM integrity of non-demented healthy subjects who are carriers/ noncarriers of APOE ε4 allele. Persson et al. report altered brain WM integrity in healthy carriers of the APOE ε4 allele (16) and showed a decline in FA in the posterior of the corpus callosum and medial temporal lobe of ε4 carriers compared to noncarriers. More recently, Honea et al. investigated differences in brain structure associated with APOE ε4 using VBM and TBSS to study morphometry and FA changes respectively (17). They showed that ε4 carriers had lower gray matter regional volumes in the left lingual/parahippocampal gyrus (p-value < 0.001, uncorrected) and decreased FA in the left parahippocampal gyrus white matter (p<0.001 uncorrected) (17).

The current VBA study consistently identified a decline in FA in the genu of APOE ε4 carriers compared to noncarriers (Figure 7). This contrasts with the results of the two studies outlined above. Methodological differences in analysis may be one reason for the conflicting reports. However, it should be noted that Honea et al. report the results of TBSS with uncorrected p values which is known to give rise to false positives (17). The small sample size in the current study may have been statistically underpowered to detect smaller changes outside the CC. However, the genu was identified in an earlier study on a larger group of ε4 carriers and non-carriers using T2 relaxometry maps (15). In the latter relaxometry study, late myelinating white matter regions such as the genu were found to be affected the most; this corresponds well with the current findings of reductions in FA of the genu in ε4 carriers compared to non-carriers.

A recent report of a longitudinal study of gray matter atrophy rates of healthy older individuals using tensor based morphometry revealed that the ε4 group demonstrated significantly greater annual atrophy rates in the temporal lobes ($p = 0.048$) and hippocampus ($p = 0.016$); greater volume loss was observed in the right hippocampus than the left (13).

The finding of greater right hippocampal atrophy in the current cross-sectional study is in accordance with the longitudinal study. Given the significant right hippocampal atrophy, and the possibility that such gray matter loss may be due at least in part to myelin loss (13), it is surprising that the adjacent parahippocampal region in ε4 carriers was not identified in the current study as having a compromised white matter integrity (e.g., lower FA values). However, other studies with larger number of subjects have identified the parahippocampal gyrus white matter as a region with lower FA in ε4 carriers compared to non carriers. The other region of significant atrophy difference in the current study was the posterior cingulate gyrus. While this region has not been identified in other morphometric studies comparing ε4 carriers to non-carriers, studies on AD and MCI patients have identified the posterior cingulum as a region that shows loss of white matter integrity (30, 31). The change in the integrity of the posterior cingulum WM found in other studies may be associated with the GM atrophy of the posterior cingulate gyrus found in the current study. Such apparent inconsistencies between volumetric and DTI data could be due to differences in gray/white contrasts and DTI parameters respectively as well as the differing analytic methods employed.

In conclusion, the total variation regularization denoising algorithm allows accurate voxel-based analysis of images acquired in clinically relevant scan times (2.5 minutes) as obtained in the current study. The most robust VBA results were obtained for tensor denoised data smoothed with an anisotropic filter. The finding of decreased FA in the genu for APOE ε4 subjects is in agreement with the findings from relaxation rate (1/T2) studies on ε4 carriers/ non carriers; the relaxometry study also found that in healthy and cognitively intact older subjects late myelinating structures were first to be selectively compromised (15).

Acknowledgments

References

1. Ashburner J, Friston KJ. Voxel-based morphometry--the methods. Neuroimage. 2000; 11:805–821. [PubMed: 10860804]

2. Pell GS, Briellmann RS, Waites AB, Abbott DF, Jackson GD. Voxel-based relaxometry: a new approach for analysis of T2 relaxometry changes in epilepsy. Neuroimage. 2004; 21:707–713. [PubMed: 14980573]

3. Morgen K, Sammer G, Weber L, et al. Structural Brain Abnormalities in Patients with Parkinson Disease: A Comparative Voxel-Based Analysis Using T1-Weighted MR Imaging and Magnetization Transfer Imaging. AJNR Am J Neuroradiol. 2011; 11:2080–2086. [PubMed: 22081675]

4. Li F, Huang X, Yang Y, et al. Microstructural brain abnormalities in patients with obsessive-compulsive disorder: diffusion-tensor MR imaging study at 3. 0 T. Radiology. 2011; 260:216–223. [PubMed: 21474704]

5. Draganski B, Ashburner J, Hutton C, et al. Regional specificity of MRI contrast parameter changes in normal ageing revealed by voxel-based quantification (VBQ). Neuroimage. 2011; 55:1423–1434. [PubMed: 21277375]

6. Lee AD, Leporé N, Brun C, et al. Tensor-based analysis of genetic influences on brain integrity using DTI in 100 twins. Med Image Comput Comput Assist Interv. 2009; 12(Pt 1):967–974. [PubMed: 20426082]

7. Abe O, Takao H, Gonoi W, et al. Voxel-based analysis of the diffusion tensor. Neuroradiology. 2010; 52:699–710. [PubMed: 20467866]

8. Singh M, Jeong J, Hwang D, Sungkarat W, Gruen P. Novel diffusion tensor imaging methodology to detect and quantify injured regions and affected brain pathways in traumatic brain injury. Magn Reson Imaging. 2010; 28:22–40. [PubMed: 19608369]

9. Moraschi M, Hagberg G, Di Paola M, Spalletta G, Maraviglia B, Giove F. Smoothing that does not blur: Effects of the approach for evaluating diffusion tensor imaging data in the clinic. J Magn Reson Imag. 2010; 31:690–697.

10. Hecke W, Leemans A, De Becker S, Jeurissen B, Parizel P, Sijbers J. Comparing isotropic and anisotropic smoothing for voxel-based DTI analyses: A simulation study. Hum Brain Mapp. 2010; 31:98–114. [PubMed: 19593775]

11. Lemaître H, Crivello F, Dufouil C, et al. No epsilon4 gene dose effect on hippocampal atrophy in a large MRI database of healthy elderly subjects. Neuroimage. 2005; 24:1205–1213. [PubMed: 15670698]

12. Wishart HA, Saykin AJ, McAllister TW, et al. Regional brain atrophy in cognitively intact adults with a single APOE epsilon4 allele. Neurology. 2006; 67:1221–1224. [PubMed: 17030756]

13. Lu PH, Thompson PM, Leow A, et al. Apolipoprotein E genotype is associated with temporal and hippocampal atrophy rates in healthy elderly adults: a tensor-based morphometry study. J Alzheimers Dis. 2011; 23:433–442. [PubMed: 21098974]

14. Bartzokis G. Alzheimer's Disease as homeostatic responses to age-related myelin breakdown. Neurobiol Aging. 2011; 32:1341–1371. [PubMed: 19775776]

15. Bartzokis G, Lu PH, Geschwind D, Edwards N, Mintz J, Cummings J. Apolipoprotein E genotype and age-related myelin breakdown in healthy individuals. Arch Gen Psychiat. 2006; 63:63–72. [PubMed: 16389198]

16. Persson J, Lind J, Larsson A, et al. Altered brain white matter integrity in healthy carriers of the APOE epsilon4 allele: a risk for AD? Neurology. 2006; 66:1029–1033. [PubMed: 16606914]

17. Honea R, Vidoni E, Harsha A, Burns JM. Impact of APOE on the healthy aging brain: A voxel-based MRI and DTI study. J Alzheimers Dis. 2009; 18:553–564. [PubMed: 19584447]

18. Jenkinson M, Bannister PR, Brady JM, Smith SM. Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage. 2002; 17:825–841. [PubMed: 12377157]

19. Smith S. Fast robust automated brain extraction. Hum Brain Mapp. 2002; 17:143–155. [PubMed: 12391568]

20. Jones D, Cercignani M. Twenty-five pitfalls in the analysis of diffusion MRI data. NMR Biomed. 2010; 23:803–820. [PubMed: 20886566]

21. Christiansen O, Lee T, Lie J, Sinha U, Chan T. Total variation regularization of matrix-valued images. Int J Biomed Imag. 2007; 27432:1–11.

22. Chen B, Hsu E. Noise removal in magnetic resonance diffusion tensor imaging. Magn Reson Med. 2005; 54:393–401. [PubMed: 16032670]

23. Ardekani S, Sinha U. Statistical representation of mean diffusivity and fractional anisotropy brain maps of normal subjects. J Magn Reson Imag. 2006; 24:1243–1251.

24. Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion. IEEE Trans Patt Anal Mach Intell. 1990; 12:629–639.

25. Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, et al. Statistical Parametric Maps in Functional Imaging: A General Linear Approach. Human Brain Mapping. 1995; 2:189–210.

26. Dukart J, Schroeter ML, Mueller K. Alzheimer's Disease Neuroimaging Initiative. Age correction in dementia--matching to a healthy brain. PLoS One. 2011; 6:e22193. [PubMed: 21829449]

27. Zhang H, Avants BB, Yushkevich PA, et al. High-dimensional spatial normalization of diffusion tensor images improves the detection of white matter differences: an example study using amyotrophic lateral sclerosis. IEEE Trans Med Imaging. 2007; 26:1585–97. [PubMed: 18041273]

28. Jones D, Symms M, Cercignani M, Howard R. The effect of filter size on VBM analyses of DT-MRI data. NeuroImage. 2005; 26:546–554. [PubMed: 15907311]

29. Lee JE, Chung MK, Lazar M, et al. A study of diffusion tensor imaging by tissue-specific smoothing-compensated voxel-based analysis. NeuroImage. 2009; 44:870–883. [PubMed: 18976713]

30. Stricker NH, Schweinsburg BC, Delano-Wood L, et al. Decreased white matter integrity in late-myelinating fiber pathways in Alzheimer's disease supports retrogenesis. Neuroimage. 2009; 45:10–16. [PubMed: 19100839]

31. Medina D, deToledo-Morrell L, Urresta F, et al. White matter changes in mild cognitive impairment and AD: A diffusion tensor imaging study. Neurobio Aging. 2006; 27:663–672.
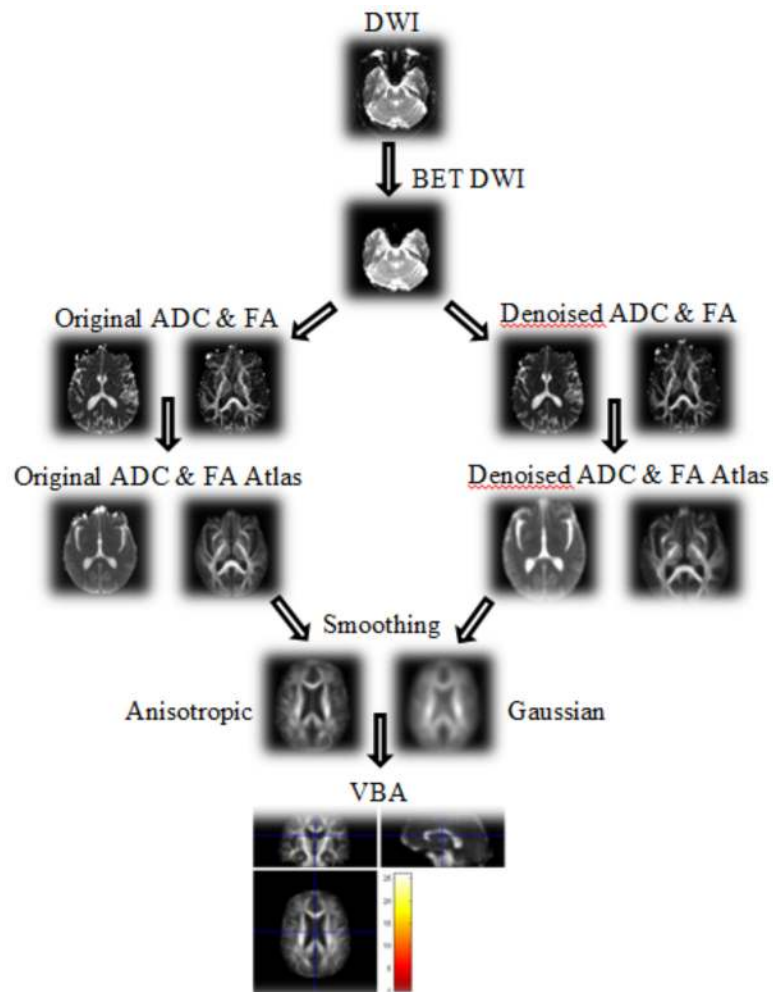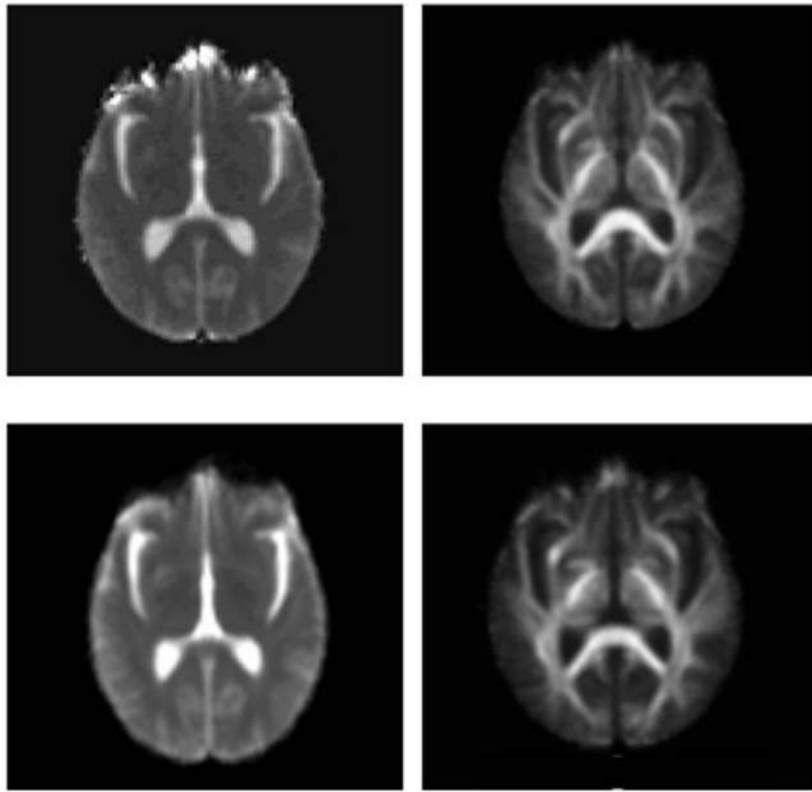
Figure 1.
Flowchart of the steps involved in the preprocessing of the DTI data and in the voxel based analysis (VBA). Atlases were created from the original ADC, FA and the denoised ADC and FA image volumes using nonlinear registration. Both datasets were analyzed with VBA and the effects of smoothing with anisotropic and Gaussian filters of varying kernel size were tested on both datasets.
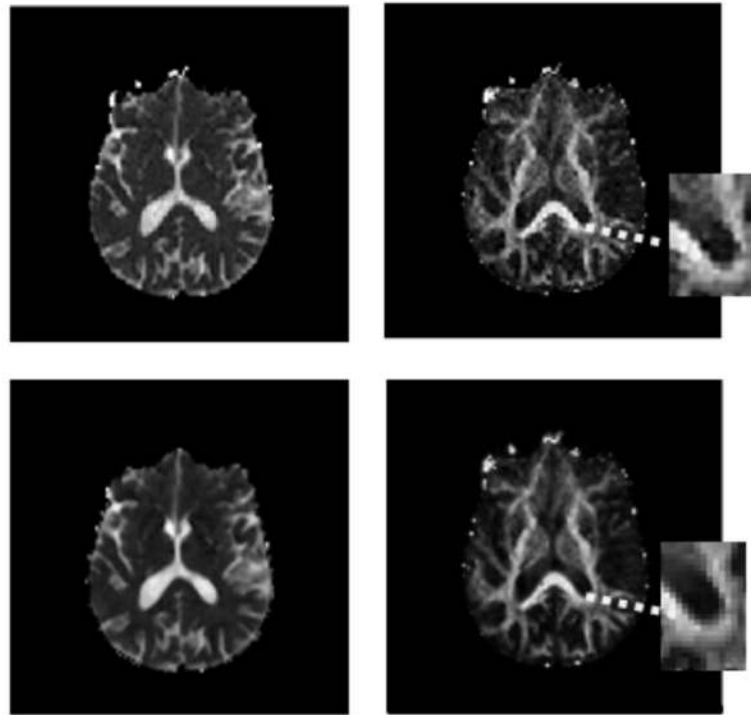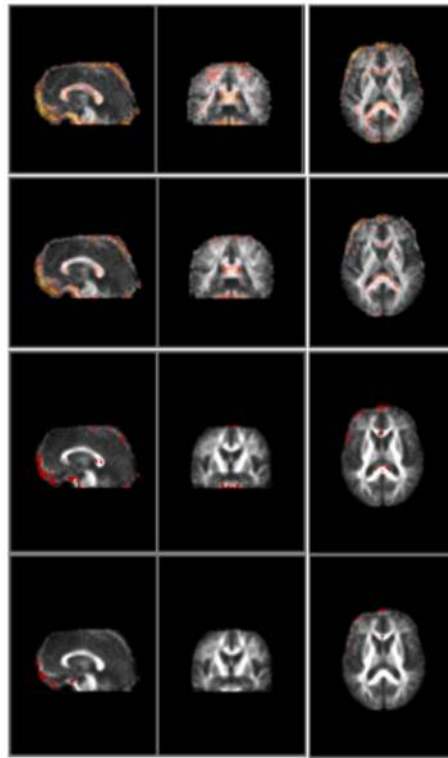
Figure 2.
A slice at the mid-volume level of ADC (left column) and FA (right column) is shown here for one subject. The top row shows the ADC and FA images obtained from the original data and the bottom row shows corresponding images obtained from the denoised tensor. The zoomed inset allows one to appreciate the reduction in variance in the FA images in homogeneous regions.
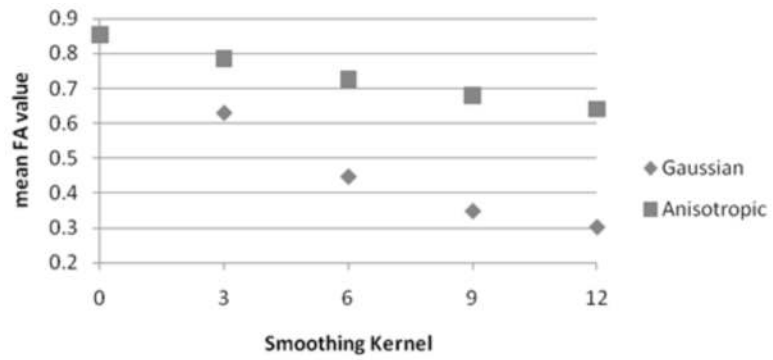
Figure 3.
Select slice from the ADC (left) and FA (right) atlases created from the original data set (top row). The bottom row shows corresponding slice from the denoised tensor data. The atlas is an average of the 14 subjects registered to a common frame of reference and this is evident in the high SNR. The fiber tracts in the FA image are well defined confirming the accuracy of the underlying non-linear registration.
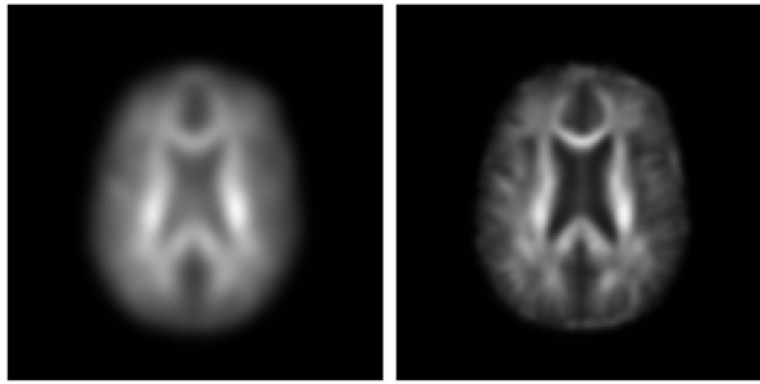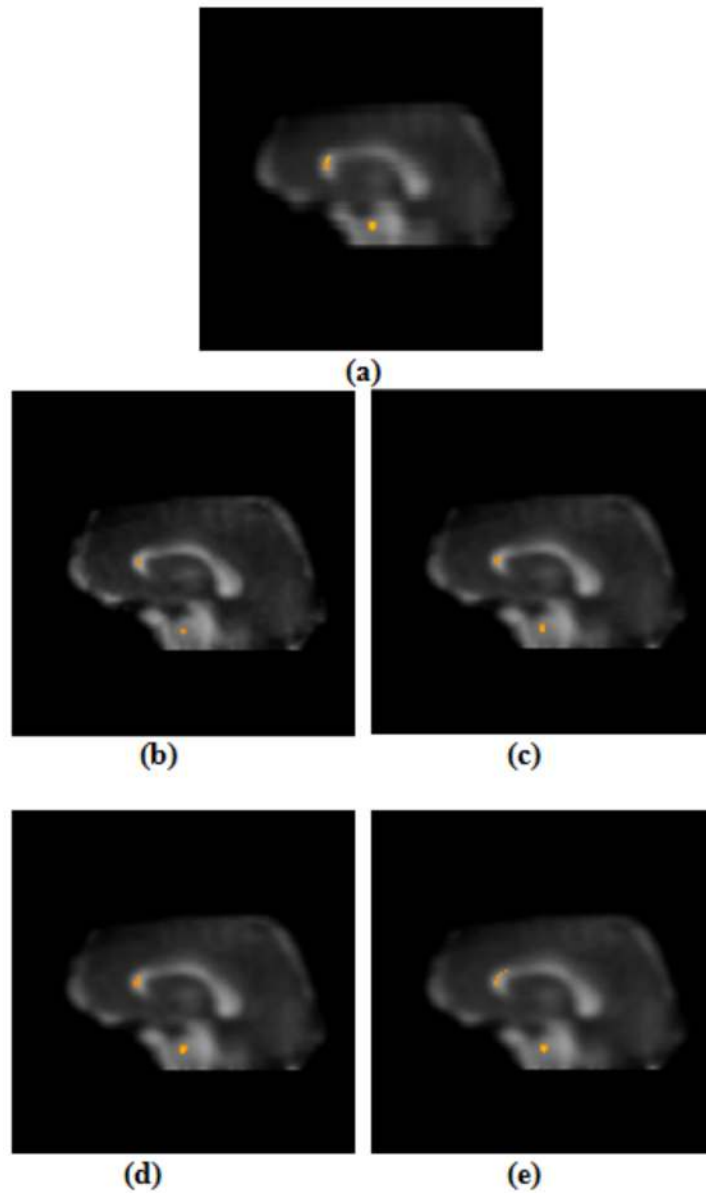
Figure 4.
Standard Deviation maps (across subjects) for the FA atlas created from the original data
(row 1), denoised data (row 3). Regions in red are those that exceed 10%. The effect of a
3×3×3 mm Gaussian filter on the standard deviation maps on corresponding slices and
orientations are shown in rows 2 and 4 (smoothed noisy and smoothed denoised data
respectively).

Figure 5.
Plot of the FA in a ROI (placed in the genu) as a function of the FWHM of the smoothing filter for the Gaussian and anisotropic filter. The strong blurring and partial voluming effect with even a moderate Gaussian smoothing can be appreciated.

Figure 6.
One slice from a subject image volume smoothed with a Gaussian filter (left) and with an anisotropic filter (right) illustrates the severe blurring introduced by Gaussian filters. Both filters used a FWHM of 6×6×6 mm.

Figure 7.
Clusters identified as significantly different between the two cohorts (APOE ε2 FA> APOE ε4 FA) is shown color coded and overlaid on the mean FA image for: (a) 3×3×3 mm isotropic Gaussian smoothing, (b), 3×3×3 mm anisotropic smoothing (c), 6×6×6 mm anisotropic smoothing, (d)9×9×9 mm anisotropic smoothing, and (e) 12×12×12 mm anisotropic smoothing. Image (a) using the Gaussian smoothing was processed using the internal filter in SPM 8 while the others (b through e) were smoothed using the custom anisotropic filter.
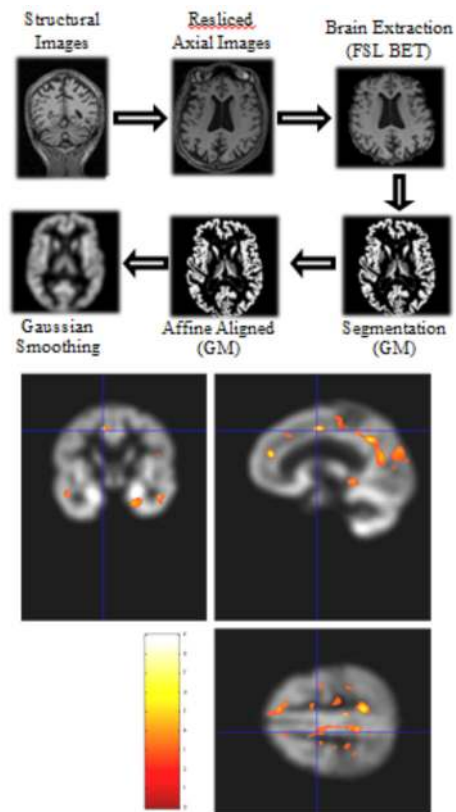
Figure 8.
Flowchart shows the processing steps in detecting gray matter differences between cohorts. Voxel based morphometry analysis of gray matter differences between ε4 carriers and non carriers shows differences in the right hippocampus and the posterior cingulate gyrus regions (ε4<ε2; p<0.01, uncorrected).

Table 1

Average deviation angle (Ada) in DEGREES for the whole brain and select rois.

	GLOBAL ADA	ROI 1ADA	ROI 2ADA	ROI 3ADA
Noisy Tensor	23.53±8.29	26.16±5.70	24.47±6.21	20.94±6.61
Denoised Tensor	18.05±7.79	18.60±6.85	20.16±6.90	13.14±7.12

Table 2

VBA Results for Isotropic and Anisotropic Smoothing Effects on the Denoised Images

Smoothing Filter Size		VBA	Cluster Size	T value	p value
None		positive	10–15	19.55	0.000
Gaussian	3	positive	18–20	15.19	0.002
Gaussian	6	negative	0	-	-
Gaussian	9	negative	0	-	-
Gaussian	12	negative	0	-	-
Anisotropic	3	positive	15–20	15.46	0.001
Anisotropic	6	positive	10–15	14.32	0.003
Anisotropic	9	positive	8–9	13.68	0.006
Anisotropic	12	positive	7–8	13.43	0.007