

Diffusion MRI abnormalities detection with orientation distribution functions: A multiple sclerosis longitudinal study

Olivier Commowick, Adil Maarouf, Jean-Christophe Ferré, Jean-Philippe Ranjeva, Gilles Edan, Christian Barillot

▶ **To cite this version:**

Olivier Commowick, Adil Maarouf, Jean-Christophe Ferré, Jean-Philippe Ranjeva, Gilles Edan, et al.. Diffusion MRI abnormalities detection with orientation distribution functions: A multiple sclerosis longitudinal study. Medical Image Analysis, 2015, 22 (1), pp.114-123. 10.1016/j.media.2015.02.005 . inserm-01134107

HAL Id: inserm-01134107

https://inserm.hal.science/inserm-01134107

Submitted on 22 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Diffusion MRI Abnormalities Detection with Orientation Distribution Functions: a Multiple Sclerosis Longitudinal Study

Olivier Commowick[a], Adil Maarouf[b], Jean-Christophe Ferré[c], Jean-Philippe Ranjeva[d], Gilles Edan[c], Christian Barillot[a]

[a] *VISAGES: INSERM U746 - CNRS UMR6074 - INRIA - University of Rennes I, France*
[b] *Neurology Department, University Hospital of Reims, France*
[c] *University Hospital of Rennes, 2 Rue Henri le Guilloux 35000 Rennes, France*
[d] *CRMBM-CNRS-Aix-Marseille University, 27 Bd. Jean Moulin 1305 Marseille, France*

Abstract

We propose a new algorithm for the voxelwise analysis of orientation distribution functions between one image and a group of reference images. It relies on a generic framework for the comparison of diffusion probabilities on the sphere, sampled from the underlying models. We demonstrate that this method, combined to dimensionality reduction through a principal component analysis, allows for more robust detection of lesions on simulated data when compared to classical tensor-based analysis. We then demonstrate the efficiency of this pipeline on the longitudinal comparison of multiple sclerosis patients at an early stage of the disease: right after their first clinically isolated syndrome (CIS) and three months later. We demonstrate the predictive value of ODF-based scores for the early detection of lesions that will appear or heal.

Keywords: Diffusion MRI, patient to controls comparison, Orientation distribution functions

1. Introduction

Diffusion MRI has been widely used in recent studies for the detection of abnormalities between brains of patients suffering from different diseases and those of healthy controls. To do so, several methodologies have been proposed that may be split into two categories. First, several groups (Whitcher et al. (2007); Lepore et al. (2008)) have presented studies comparing a population

Email address: Olivier.Commowick@inria.fr (Olivier Commowick)

of patient brains with respect to a population of healthy controls. Such approaches allow to see, at the population level, how a disease affects the brain white matter structure and are therefore of great interest to understand globally a disease. The second category focuses more on the study of a specific patient with respect to a population of controls (Commowick et al. (2008)), allowing to detect for that specific person where the brain may be affected (i.e. presence of lesions or other diffuse pathology-related change).

All these methods have been widely used to perform studies on multiple sclerosis (MS) with the goal of better understanding the disease and possibly find early biomarkers of the pathology evolution that would allow an adaptation of the treatment to the patient. Among those studies, Filippi et al. (2001) studied, in different regions of interest, differences of diffusion derived scalar coefficients (fractional anisotropy - FA, apparent diffusion coefficient - ADC) between controls and MS patients. Mani et al. (2013) on the other hand performed studies on the shape of specific white matter fiber tracts, corpus callosum, and found evidence of shape changes related to multiple sclerosis. Longitudinal studies of MS in diffusion tensor imaging have also been performed, trying to find correlation between clinical scores such as EDSS (Expanded Disability Status Scale) and DTI derived scalar values. Giorgio et al. (2010) or Liu et al. (2012) performed tract-based spatial statistics and found some correlation in several white matter regions, such as the splenium of the corpus callosum, between decreasing FA and EDSS in relapsing remitting patients. Rocca et al. (2000) conducted an interesting study on the predictive value of diffusion MRI in a relatively small number of patients with scans every week, showing an increase in mean diffusivity in normal appearing white matter areas subsequently transforming into lesions. Werring et al. (2000) studied the evolution of ADC in relationship with the apparition of lesions with contrast in Gadolinium-enhanced MR images, concluding that structural damage in lesions can cause damage in connected areas of normal appearing white matter. Finally, Grigis et al. (2012) presented a longitudinal study of change detection in diffusion MRI, based on DWI signal bootstrap, demonstrating the ability to detect lesions or diffuse damage as a change of diffusion properties. Only few studies have looked at diffusion imaging in multiple sclerosis from the beginning of the disease (i.e. from the first clinically isolated syndrome - CIS). Such studies could bring important insights into the early development of the disease, leading to a better patient management in daily practice.

Almost all previously mentioned MS studies have considered the comparison of scalar values derived from DTI, such as ADC or FA, and are thereby losing part of the information contained in the tensor. Whitcher et al. (2007); Lepore et al. (2008) showed in studies on other pathologies that considering

the whole tensor can be more sensitive to detect differences. Moreover, it has been demonstrated that the tensor model is not sufficient to capture regions of crossing fibers in the brain, as it allows to code only for one principal direction of diffusion when there may be several. Recently, several papers proposed higher order diffusion models able to capture several diffusion directions (Assaf and Basser (2005); Descoteaux et al. (2007); Scherrer and Warfield (2012); Stamm et al. (2012); Zhang et al. (2012)). Utilizing the full information coming from those diffusion models could potentially improve further the detection power of abnormalities in patients. Following this idea, recent works have studied frameworks based on multiple fascicle models. Jbabdi et al. (2010) presented an extension of the popular tract-based statistics framework to better handle fiber crossing regions, based on a ball-and-stick model. Taquet et al. (2014) presented a framework to perform group comparison between patients and controls relying on the multi-tensor model. Both approaches compare information along tracts of interest at the group level to detect any abnormalities related to a disease. However, as group comparison frameworks, these papers are not directly applicable to a single patient to help predict disease evolution. Moreover, they consider tract-based statistics while in prospective studies we may be also interested in detecting differences at the voxel level over the whole brain.

We propose a new method to detect abnormalities by comparing the orientation distribution functions (ODF) at the voxel level between a patient and a population of controls, in a similar fashion to Commowick et al. (2008). It relies on a statistical comparison, combined to a principal component analysis (PCA) for robustness, of the diffusion probability values sampled on the sphere from virtually any model that has a probability density function defined on the sphere. We applied this method first to simulated data with lesions to evaluate the detection power of ODF-based scores with respect to DTI-based scores. Then, we applied our pipeline to the longitudinal detection of differences between a database of CIS patients with scans at two time points, to evaluate the ability of ODF-based scores to highlight early differences that may lead to the appearance or disappearance of lesions at the second time point.

We present in Section 2 our new methodology for the comparison of higher order models on the sphere, the databases on which our method was applied and our overall pipeline for longitudinal comparison. Then, we present in Section 3 our experiments, first based on simulated data with lesions where the ground truth of differences is known. This first section demonstrates the evaluation of our new framework compared to existing algorithms. Then, we focus on the results of the longitudinal study of MS patients. Finally, we discuss our results and conclude in Section 4.

2. Material and Methods

2.1. Patient to Group Diffusion MRI Comparison

The central part of our framework, illustrated later on with experimental data in Section 2.3, relies on the voxelwise comparison of a diffusion model to a group of diffusion models, e.g. the comparison of the DTI value of a patient vs the DTI values of a set of controls. In the following, we assume that all images (patient and controls) are aligned into a common space where each voxel of each image describes the exact same spatial position. Based on this assumption, we present a change detection method that applies both to diffusion tensor images and higher order models such as orientation distribution functions (ODF). We detail in the following the general comparison framework and its application to tensor and ODF images.

2.1.1. General Abnormalities Detection Framework

In Commowick et al. (2008), we introduced a methodology for the comparison of diffusion tensors between a patient and a group of controls. Here, we generalize and extend it to virtually any model. We wish to compare, at a given voxel, the patient model M_P and models from the control subjects database $M_j,\ j = 1 \dots N$. From each of these diffusion models, let us assume that it can be represented as a K-dimensional vector of features V (V_P or V_j). We could then apply any comparison framework to those vectors. However, artifactual detections may appear due to noise in the images or registration errors. This is particularly true for complex diffusion models whose estimation from diffusion MRI may be more sensitive to these errors. The comparison is therefore preceded by a Principal Component Analysis (PCA) as follows:

- Perform a PCA from the control database to retain the h most meaningful eigenvectors
- Project V_P and the V_j onto a reduced h-dimensional space utilizing the eigenvector basis
- Compute a comparison score and p-value

To perform the PCA analysis, the covariance matrix Σ_C is derived from the controls database. The eigenvalues and eigenvectors of Σ_C are then computed and we retain only the h largest eigenvalues and eigenvectors. These vectors are in turn utilized to project the input data (patient and controls) on an h-dimensional space ($h \ll K$). After projection, we therefore seek to compare a vector $V_{P,h}$ and a set of vectors $V_{j,h}$ from the controls

database. The controls population $\mathbf{V} = \{V_{j,h},\ j = 1 \dots N\}$ is assumed to follow a multivariate Normal distribution $N(\bar{V}, \Sigma_V)$, where $\bar{V}$ and Σ_V denote respectively the average and covariance matrix of the population $\mathbf{V}$. Under this assumption, we compute the difference statistic between $V_{P,h}$ and $N(\bar{V}, \Sigma_V)$ as a Mahalanobis distance:

$$d^2(V_{P,h}) = \left(V_{P,h} - \bar{V}\right)^T \Sigma_V^{-1} \left(V_{P,h} - \bar{V}\right) \tag{1}$$

d^2 varies between 0 and infinity, getting larger as the patient vector gets less likely to belong to the multivariate Normal distribution of the controls. A p-value can then be computed from this distance by recalling that the statistic $T = \frac{N(N-h)}{h(N^2-1)} d^2$ follows a Fisher distribution with parameters h and $N - h$: $T \sim F(h, N - h)$. The difference test p-value is therefore written as:

$$p(V_{P,h}) = 1 - F_{h,N-h}\left(d^2(V_{P,h})\right) \tag{2}$$

where $F_{h,N-h}$ is the cumulative distribution function of a Fisher distribution with parameters h and $N - h$.

2.1.2. DTI Abnormalities Detection

In the particular case of DTI, we wish to compare tensor values. Tensors lie in the space of 3×3 positive definite matrices and some care should be taken so that computations account for this specific space. We utilize the log-Euclidean framework proposed by Arsigny et al. (2006b) to perform the comparison. Arsigny et al. have indeed defined a Lie group for tensors (relying on their matrix principal logarithm), and Euclidean operations can therefore be computed on their logarithms. For DTI, we therefore chose to compare the vector representations of the logarithms of tensor values $\mathrm{Vec}\left[\log(D)\right]$ - where D is a tensor, log is the matrix logarithm and Vec the vector operator as proposed by Arsigny et al. (2006b). PCA and vector comparisons are then computed on these values, and in this case h may vary in between 0 and 6.

2.1.3. ODF Abnormalities Detection

Higher order models may allow us to capture more differences as they are better able to capture regions where fibers are crossing. In the following experiments, we will study Orientation Distribution Functions (ODF) represented on a modified Spherical Harmonics (SH) basis as proposed by Descoteaux et al. (2007):

$$Y_j(\theta,\phi) = \begin{cases} \frac{\sqrt{2}}{2}\left((-1)^m Y_l^{-m}(\theta,\phi) + Y_l^m(\theta,\phi)\right), & if\ m < 0 \\ Y_l^0, & if\ m = 0 \\ i\frac{\sqrt{2}}{2}\left((-1)^m Y_l^{-m}(\theta,\phi) - Y_l^m(\theta,\phi)\right), & if\ m > 0 \end{cases} \tag{3}$$

where the couple (θ, ϕ) denotes a point on the sphere, $l = 0, 2, 4, \ldots, L$ is the order of the SH basis, $m = -l, \ldots, 0, \ldots, l$ is an index, $j := j(l, m) = (l^2 + l + 1) + m$. Y_j is the j-th element of the SH basis, and Y_l^m is the complex SH value. An ODF M expressed in this basis is therefore a vector of $L^2 + L + 1$ parameters c_j: $M(\theta, \phi) = \sum_j c_j Y_j(\theta, \phi)$.

From each of the diffusion models, we sample the probability distribution function (pdf) of diffusion in K directions regularly placed over the unit sphere thereby constructing for each model a K-dimensional vector V (V_P or V_j). The comparison of the patient with respect to the control database is then performed using Eqs (1) and (2), with the projected data from PCA. Interestingly, following such an approach allows us to be independent of the diffusion model used: if a probability density function is available for the model, then the presented statistical test may be used.

2.2. Diffusion MRI Databases

2.2.1. Simulated Data

To quantitatively evaluate the detection performance of our "one patient vs control group" comparison framework, we have created two simulated databases where lesions are known. The first database consists of crossing fibers with and without lesions (illustrated in Fig. 1). A noise-free DWI image of two fibers crossing at 90 degrees was first constructed using a multi-tensor model, the relative weights of each fiber being equal in the crossing region. The simulated diffusion weighted image consists of one B0 and 81 directions with a b-value of 1000 $s.mm^{-2}$ (image size of 128×128). Then, a simulated lesion was added inside the crossing region, so that each compartment ADC (respectively FA) was increased (respectively decreased). Then, Rician noise was added repeatedly to these two images so that a database of 50 simulated controls and 100 simulated cases was created. Different noise levels were tested to study its influence on detection results: low level (3% of the image peak value), mid level (5% of the image peak value) and high level of noise (7% of the image peak value).

The second dataset (illustrated in Fig. 2) was created from a real control subject acquisition, for whom a diffusion sequence with 30 directions (image size: 128×128×60, b-value: 1000 $s.mm^{-2}$, 2×2×2 mm^3 voxel resolution) was acquired. We first added a lesion in the control DWI volume, located in a known region of crossing fibers: fibers going through the corpus callosum and corticospinal tracts. The diffusion properties of the lesion were then computed by comparing values of DWI signals (for each gradient direction) in known patient lesions and corresponding DWI signals in controls at the same positions (after registration). For each DWI image, we therefore computed a ratio indicating the relative change between a lesion signal and a normal

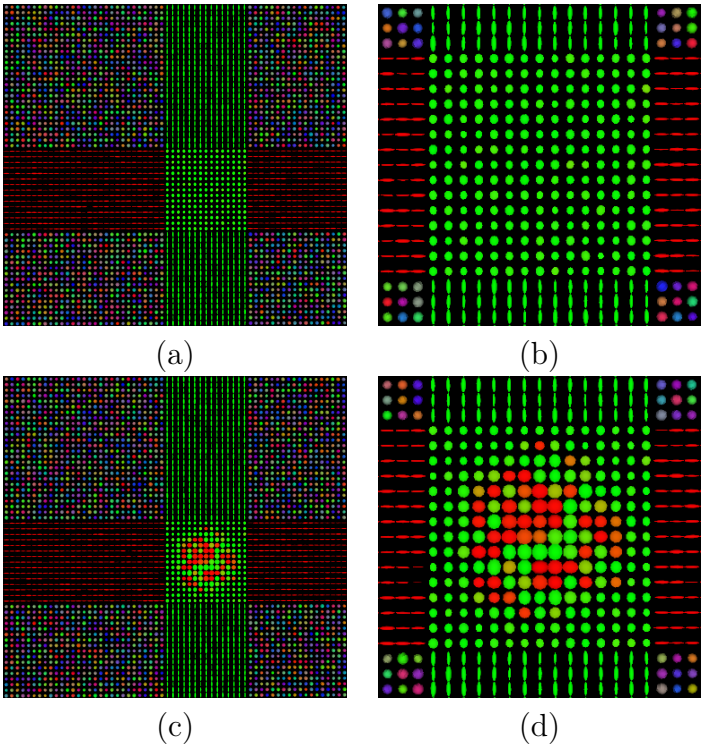


Figure 1: **Illustration of Simulated Data from the First Database**. Tensor maps derived from simulated diffusion weighted images. Legend: Tensor map without lesions (a,b) and with a simulated lesion in the crossing region (c,d). (b,d) are close-ups on (a,c).

signal. This intensity ratio was then applied to the selected control subject in the lesion region delineated. Rician noise was also added repeatedly to the two images so that a database of 50 simulated cases and 50 simulated controls were obtained. As for the first dataset, three different levels of noise were also tested (3%, 5% and 7% of the image peak value). While the first database aims at proving the interest of higher order models for regions of crossing fibers when HARDI sequences are acquired, the second database will be used to evaluate the proposed approach in a realistic frame where lower angular resolution images are acquired.

2.2.2. Multiple Sclerosis CIS Patients

We have used a database of 15 patients acquired from their first Clinically Isolated Syndrome (CIS), suggestive of multiple sclerosis. Among those 15 patients, 53.3 % were women and the average age was 29.8 ± 8.5 years old. MRI data was acquired using the same protocol at two time points: right after the first CIS event, hereafter called M_0, and 3 months later, hereafter called M_3. At each of these dates, T1w, T2w, PDw and FLAIR images were

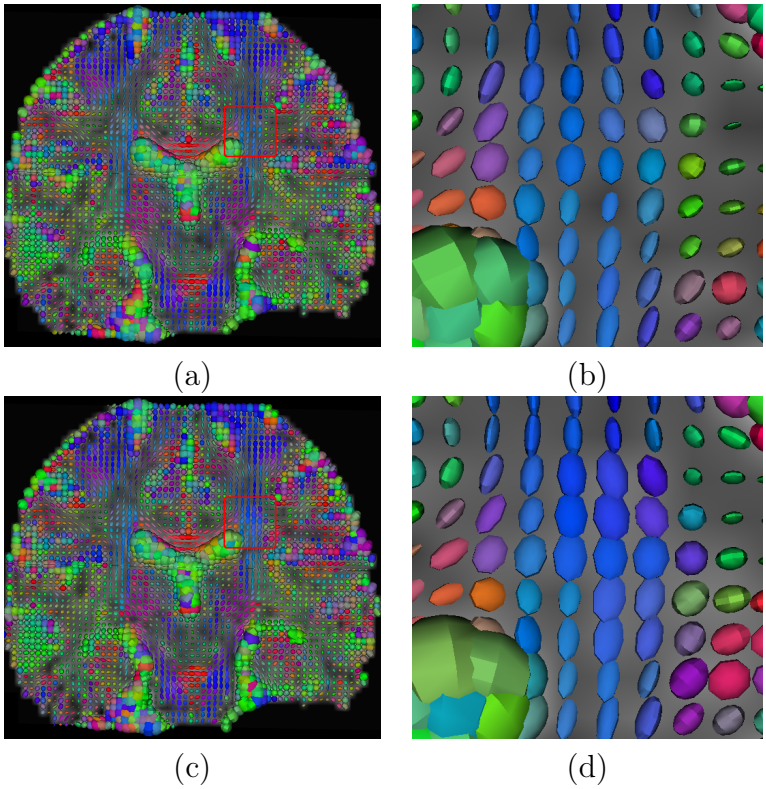


Figure 2: **Illustration of Simulated Data from the Second Database**. Tensor maps derived from simulated diffusion weighted images overlaid on top of average diffusion coefficient images. Legend: Tensor map without lesions (a,b) and with a simulated lesion in the crossing region (c,d). (b,d) are close-ups on the lesion region (a,c).

acquired on a Siemens Verio 3T scanner, as well as diffusion weighted MRI on a single shell. The T1w sequence was a 3D-MPRAGE with a 1 mm^3 isotropic voxel resolution, on a 256×256×160 matrix. T2w and PDw were acquired using a dual Turbo Spin Echo sequence with a 1×1×3 mm^3 voxel resolution, on a 192×256×44 matrix. The diffusion weighted MRI acquisition was a 30 directions acquisition on a single shell of b-value 1000 $s.mm^{-2}$, with a 2×2×2 mm^3 voxel resolution, on a 128×128×60 matrix.

A trained radiologist then delineated all lesions at all time points for those patients, based on T2w and FLAIR images which are conventionally used in clinical routine for lesion delineation.

2.2.3. Healthy Controls

In addition to the patients, a database of 46 matched control subjects was acquired. These subjects were selected so that their age and gender matched those from the patients (60.9 % of the controls were women and the average

age was 29.3 ± 6.9 years old). The same protocol was performed on the same scanner for those subjects but at only one time point since no evolution was expected. T1w, T2w, PDw and FLAIR images were acquired on a Siemens Verio 3T scanner. The T1w sequence was a 3D-MPRAGE with a 1 mm^3 isotropic voxel resolution, on a $256 \times 256 \times 160$ matrix. T2w and PDw were acquired using a dual Turbo Spin Echo sequence with a $1 \times 1 \times 3$ mm^3 voxel resolution, on a $192 \times 256 \times 44$ matrix. A diffusion weighted MRI acquisition was added with 30 directions on a single shell of b-value 1000 $s.mm^{-2}$, with a $2 \times 2 \times 2$ mm^3 voxel resolution, on a $128 \times 128 \times 60$ matrix.

2.3. Multiple Sclerosis Processing Pipeline

The processing pipeline we developed is divided into two parts. First, the input images were pre-processed as illustrated in Fig. 3. All gradient volumes of each DWI image was first registered onto the corresponding B0 image to correct for motion during the acquisition. To do so, a global rigid transform was computed using block-matching registration as proposed by Ourselin et al. (2000). Then, EPI distortion correction was applied by non-linearly registering each gradient image to the B0, utilizing the FFD algorithm from Rueckert et al. (1999) with few control points (5 in each direction). Then, tensors were estimated as well as ODF using Descoteaux et al. (2007) estimation method. To conclude this pre-processing pipeline, T1w images and diffusion images were registered together up to a global rigid transformation to get an anatomical reference for our results.

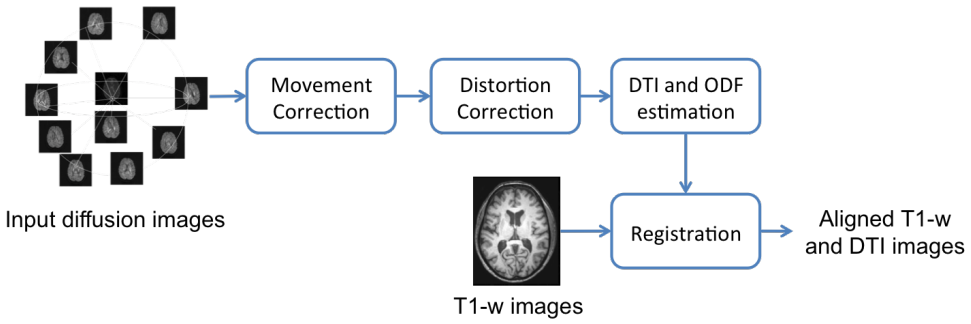


Figure 3: **Preprocessing Pipeline Illustration**.

Once this preprocessing was performed for all input images (controls and patients at M_0 and M_3), we then utilized the pipeline illustrated in Fig. 4 to compare, voxel by voxel, each patient to the database of controls. First, a DTI atlas was computed from the controls images, utilizing an updated method derived from Guimond et al. (2000) to handle transformations in the log-Euclidean framework on diffeomorphisms proposed by Arsigny et al.

(2006a). Non linear registration for atlas construction was performed based on full tensors utilizing a block-matching strategy introduced by Suarez et al. (2012). This registration proceeds by iteratively computing matches between blocks in the tensor images to estimate dense field corrections. We used the following parameters for the registration: three pyramid levels, block size $5 \times 5 \times 5$ voxels, block neighborhood of 2 voxels in each direction. The obtained transformations were then applied on T1w images, and on ODF images, using the local rigid rotations extracted from the transformation fields (Blanco et al. (1997); Ritchie and Kemp (1999); Geng et al. (2009)).

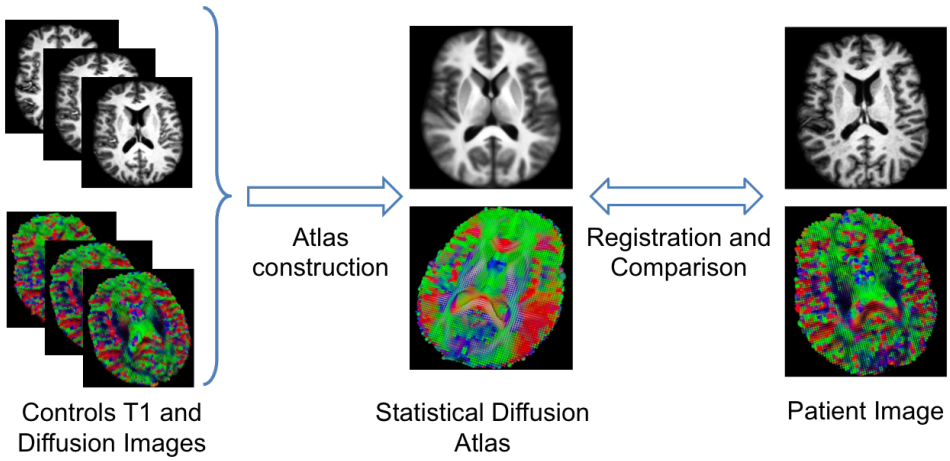


Figure 4: **Processing Pipeline Illustration**.

The first time point (M_0) of each patient was then registered onto the average atlas image utilizing the same registration methods with the same parameters. Then, for subsequent time points, only a supplementary global rigid registration to M_0 is necessary to bring the DTI and ODF image onto the atlas, as we can assume that the global brain shape of the patient did not change within three months. Once M_0 and M_3 time points are aligned onto the atlas, we applied the framework proposed above to compare the ODFs extracted from the diffusion weighted images and report the results in the next section.

3. Experiments and Results

3.1. Diffusion Models Comparison on Simulated Data

Based on the simulated datasets presented in Section 2.2, we have estimated tensors and ODFs in a spherical harmonics basis of order 4. We have then performed an abnormality detection for each of the simulated cases independently, each time comparing the detection results with the known ground

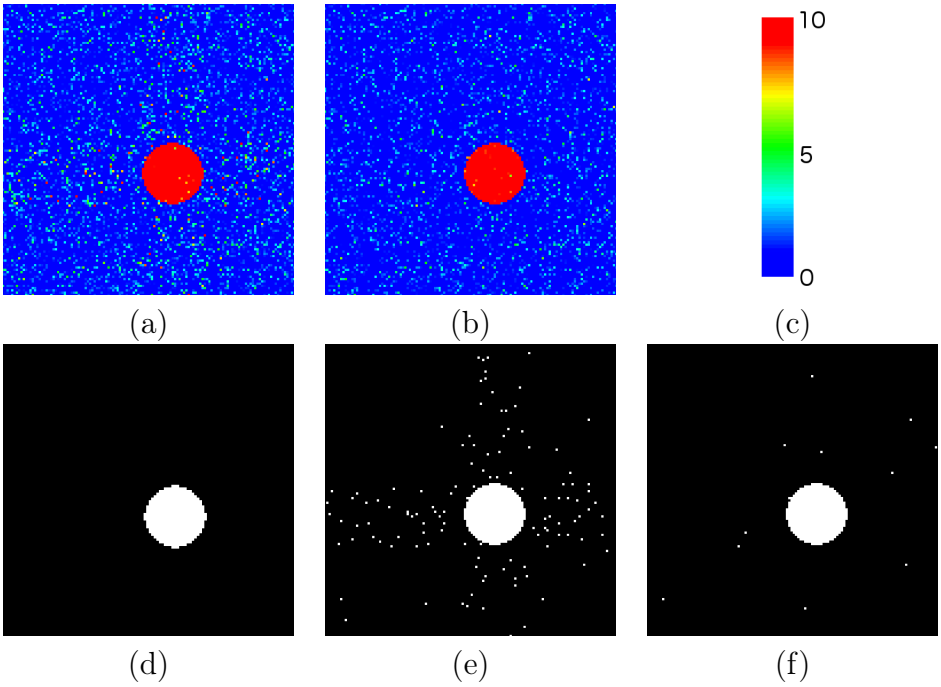


Figure 5: **Lesion Difference Evaluation on Simulated Fiber Crossing Data**. Z-score maps (top row) and thresholded p-value maps (bottom row) illustrating the differences of detection between tensor and ODF-based abnormality detection. Legend: difference score maps for tensor-based (a) and ODF-based (b) methods, (c): color range of values, (d): abnormality ground truth, detection maps ($p < 0.05$ after FDR correction) for tensor-based (e) and ODF-based (f) methods.

truth. Such an approach was applied on the two databases and results are reported in the following sub-sections.

3.1.1. Results on Simulated Crossing Fibers with Lesions

Fig. 5 presents a representative example of the FDR-corrected detection results at $p < 0.05$ for the mid-level of noise, with our approach (when keeping four eigenvectors from PCA analysis) based either on tensors or on ODFs. In addition, we tested the validity of the proposed statistical tests on this dataset: we calculated the the ratio of false positive detections in the background before FDR correction. The numbers are close to 5% (respectively 5.39% and 5.27%), confirming the validity of the proposed tests. From Fig. 5, it appears that both methods are well capable to detect the true lesion. However, tensor-based detection seems more prone to false positives than ODF-based detection.

In addition to visual inspection and individual tests, we have then computed Dice scores between the detections obtained by both algorithms and

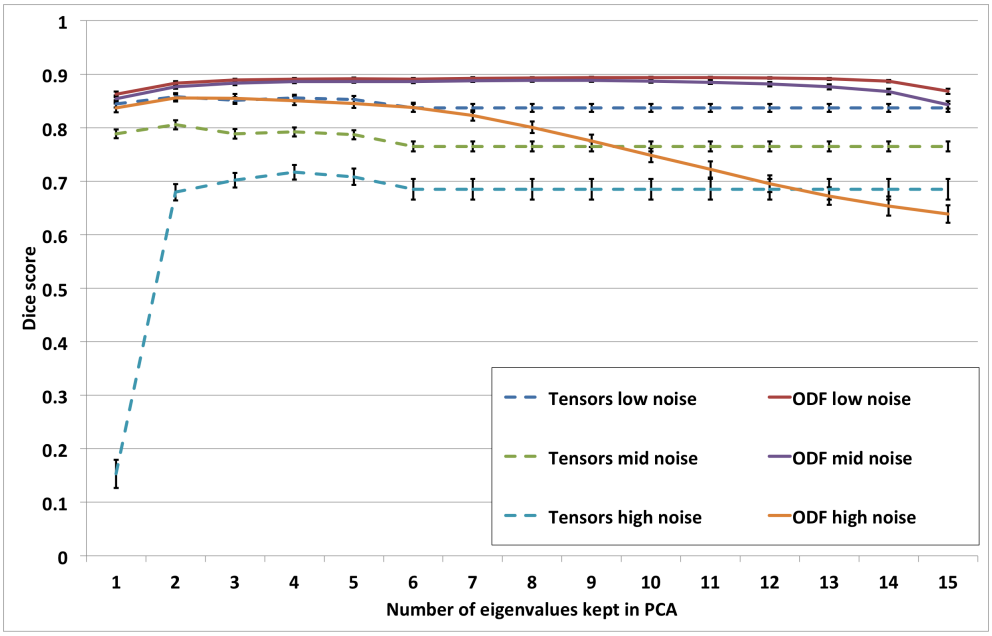


Figure 6: **Lesion Detection Dice Scores on Simulated Fiber Crossing Data**. Dice overlap scores for lesion detection on simulated crossing fibers with known lesion ground truth, for tensor and ODF-based detection

, on three different levels of noise. Results for tensor-based detection above six eigenvalues are set to the results with six eigenvalues for comparison purposes.

the known ground truth of lesion. We report in Fig. 6 the mean and standard deviation of these scores over the 100 simulated cases and for the three different levels of noise. As the ODF-based method is dependent on the number of eigenvectors kept from PCA, we are also reporting the results depending on this number for ODF-based detection.

This figure illustrates a clear difference between the ODF and tensor-based method, with an advantage to ODF for all noise levels. This demonstrates that, due to the inability of DTI to estimate water diffusion in crossing fibers regions, DTI-based comparison may result in degraded detection scores. On the contrary, ODF, being able to model multiple fiber directions in a voxel, is more able to capture differences between a patient and controls. ODF-based detection achieves very good results for low and mid-levels of noise, with only a small decrease for 14 and 15 eigenvalues, which comes from the fact that adding more eigenvectors is only bringing more noise at that stage. For a high level of noise, all detection results get lower scores but ODF detection performs still better than tensor-based comparison and gets high scores (above 0.8), although it is more impacted by noise when reaching

a high number of kept eigenvectors, compared to other levels of noise.

3.1.2. Results on Simulated Images of Patients

Fig. 7 presents FDR-corrected detection results at $p < 0.05$ obtained on the second database of simulated data with a mid-level of noise (on real images) with our proposed approach based on either tensors or ODFs (when keeping three or six eigenvectors from PCA analysis). Even with lower angular resolution, both methods are well capable to detect the true lesion. Tensor-based detection is more prone to false positives than ODF-based detection. In addition, it is clear that the number of eigenvalues kept is important: keeping only three eigenvalues leads to a clear under-estimation of the abnormality both for tensor and ODF-based detection.

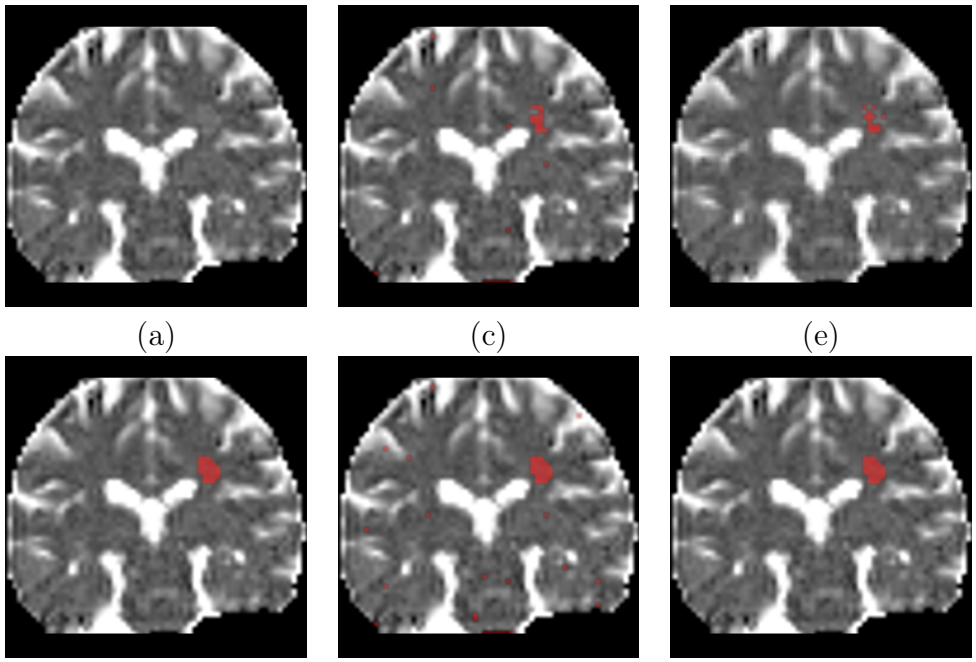


Figure 7: **Abnormality Detection on Simulated Patient Data**. Abnormalities detected at a $p < 0.05$ level after FDR correction with the evaluated methods. Patient ADC image (a), patient ADC image overlayed with (b): true simulated lesion, tensor-based detections keeping the 3 (c) or 6 (d) eigenvectors in the PCA step, ODF-based detections keeping the 3 (e) or 6 (f) eigenvectors.

We have then computed Dice scores between the detections obtained for each algorithm and the known ground truth of lesion. We report in Fig. 8 the mean and standard deviation of these scores over the 50 simulated cases for each noise level. This figure illustrates again the advantage of

ODF-based detection, which outperforms tensor-based detection in almost all cases. The ODF-based detection best result is reached for 6 eigenvalues (Dice score respectively of 0.98, 0.97 and 0.86 for low, mid and high levels of noise, compared to tensor-based scores of 0.82, 0.73 and 0.56). Interestingly, this number of eigenvalues where the maximum is reached differs from previous experiments, indicating that either the change to a real image or to a lower angular resolution is affecting the optimal number of eigenvalues kept after PCA. Overall, although scores are lower with respect to the previous experiment, this figure confirms that ODF performs better than tensor-based comparison even for low angular resolution images.

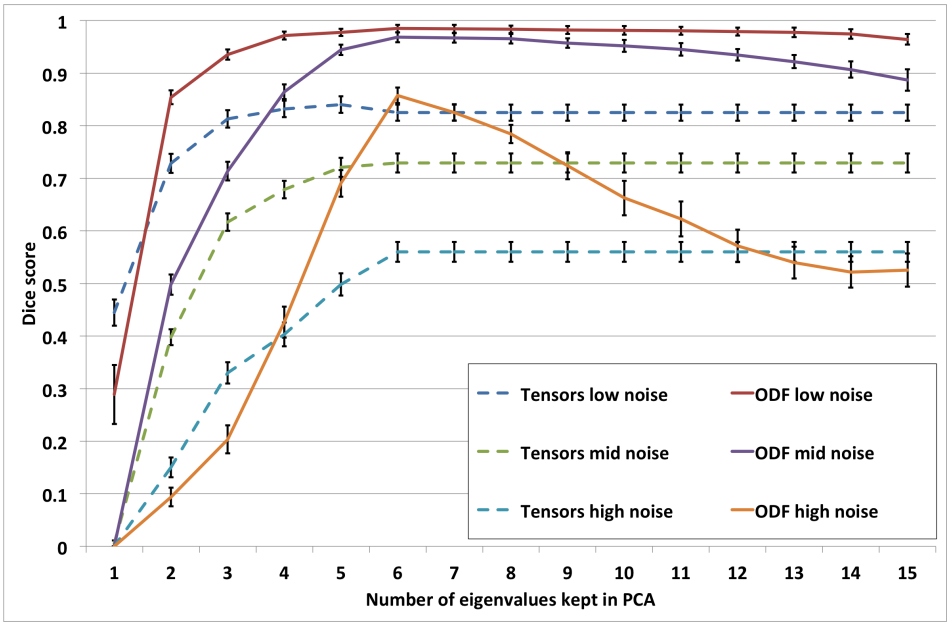


Figure 8: **Lesion Detection Dice Scores on Simulated Patient Data**. Dice overlap scores for lesion detection on a simulated patient with known ground truth, on three different levels of noise. Results for tensor-based detection above six eigenvalues are set to the results with six eigenvalues for comparison purposes.

3.2. Multiple Sclerosis Patients Analysis

We have then applied our processing pipeline to our database of multiple sclerosis patients, running it on each patient and each time point to obtain difference scores as expressed in Eq. 1. In addition, we have derived p-values from these scores and detect abnormalities as those voxels having a p-value $p < 0.05$ after FDR correction. We present in Figs. 9 and 10 results obtained for two representative patients out of the 15 patients studied. Appearing lesions at M_3 are highlighted in these figures with arrows.

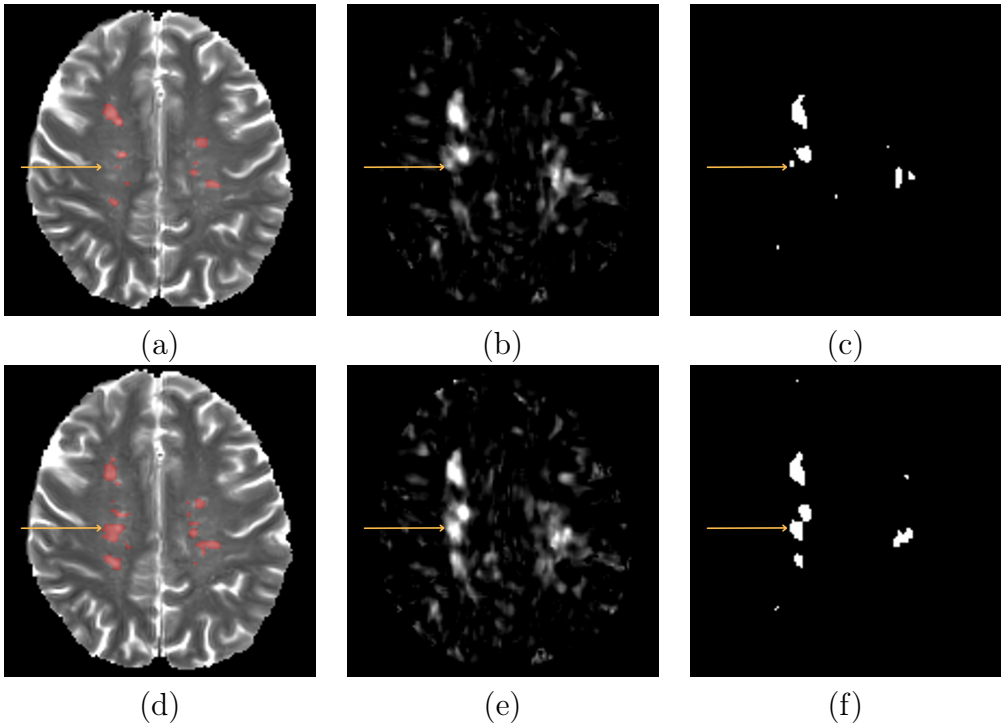


Figure 9: **ODF-Based Difference Scores on one MS Patient with Appearing Lesions**. Illustration of the ODF scores at the two time points for one patient with appearing lesions. First row: at M_0, second row: at M_3. (a,d) are axial views of T2-w images overlaid with the manual lesion delineation from the expert. (b,e) correspond to the same axial views of the ODF scores as computed from our pipeline. (c,f) correspond to the abnormalities with respect to normal anatomy detected (at $p < 0.05$ after correction). The ODF scores at M_0 show a higher score when a lesion will appear at M_3 (see arrows).

Both figures clearly show detections of abnormalities inside the regions of lesions in diffusion weighted MRI. In addition, it may be seen on Fig. 9 that some regions of the brain that appear as normal on both time points on T2w images are showing abnormalities in diffusion weighted imaging (right part of Fig. 9.c and 9.f). Comparing the results at time points M_0 and M_3 on regions of appearing lesions at M_3 (arrows on images b and e on Figs. 9 and 10), it seems that for both patients the ODF derived score is higher at M_0 when compared to the value in the normal appearing white matter. These score differences do not seem to be sufficient to have a definite large abnormality detection at M_0 which may be explained by the fact that the lesion has not yet developed. However, the scores seem to be a good indicator of lesions appearing three months later.

To further evaluate ODF scores as a marker of future lesion apparition, we have performed a quantitative study based on the ODF scores obtained

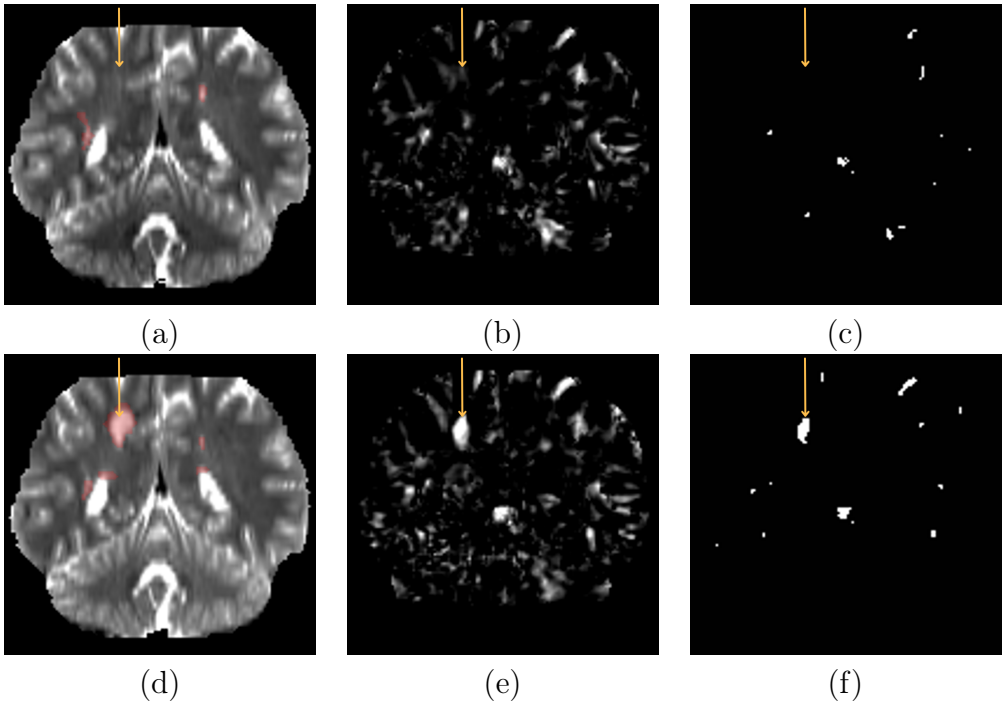


Figure 10: **ODF-Based Difference Scores on a Second MS Patient with Appearing Lesions**. Illustration of the ODF scores at the two time points for a second patient with appearing lesions. First row: at M_0, second row: at M_3. (a,d) are coronal views of T2w images overlaid with the manual lesion delineation from the expert. (b,e) correspond to the same views of the ODF scores as computed from our pipeline. (c,f) correspond to the abnormalities with respect to normal anatomy detected (at $p < 0.05$ after correction). Although the ODF scores at M_0 do not always highlight a significant difference from controls, these scores are higher when a lesion will appear at M_3 (see arrows). Legend: vertical axis: Mahalanobis distance.

for each patient. Thanks to the manual delineations at both M_0 and M_3 time points, we have computed the average for all patients of ODF-based scores in four regions: normal appearing white matter at both time points, regions of appearing, staying and disappearing lesions at M_3. We report the Box and Whiskers plot of the ODF scores for all 15 patients in Fig. 11.

These quantitative comparisons bring up several important results. First, normal appearing white matter (NAWM) does not significantly change of score between M_0 and M_3 (see NAWM in Fig. 11, paired t-test, $p = 0.472$). As expected, appearing and disappearing lesions have significantly different scores in between M_0 and M_3 (see appearing and disappearing in Fig. 11, paired t-tests, $p = 0.017$ for appearing lesions and $p = 0.009$ for disappearing lesions). Interestingly, lesions that were already present at M_0 and that stay at M_3 also have significantly different scores (see Staying in Fig. 11, paired

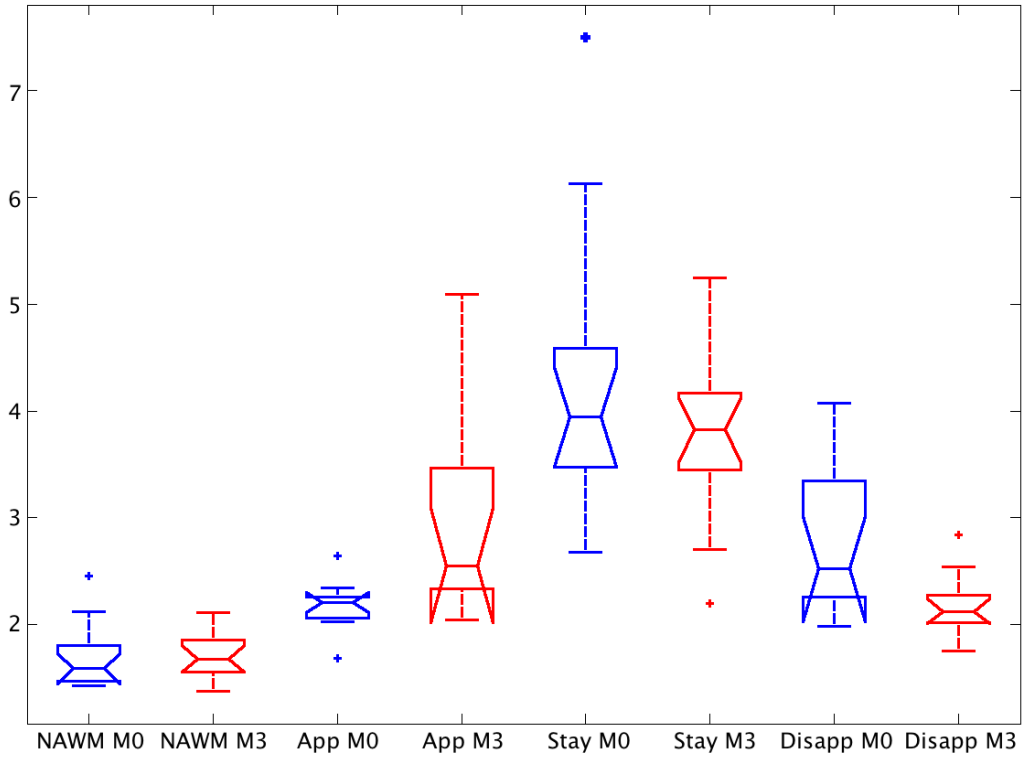


Figure 11: **ODF-Based Difference Scores on 15 MS Patients**. Illustration of the average difference scores computed from ODFs over the 15 MS patients, respectively on NAWM, appearing, staying and disappearing lesions. Scores in lesions appearing at M_3 are significantly different at M_0 from normal appearing white matter.

t-test, $p = 0.021$), which highlights the fact that, although these lesions stay, their internal structure may also change in time.

Another important result is that the NAWM score at M_0 is significantly lower than the score at M_0 for those tissues where lesions will appear at M_3 (NAWM vs appearing lesions taken at M_0, one-way ANOVA, $p < 0.001$). These results confirm the visual results on ODF score maps, presented in Figs. 9 and 10, that showed different scores in regions of future appearing lesions. In addition, staying lesions have a significantly higher score at M_0 than lesions that will disappear at M_3 (staying vs disappearing taken at M_0, one-way ANOVA, $p = 0.0015$). Both these results clearly show that ODF-based scores are able to show in advance some regions where lesions will appear or disappear.

4. Conclusion

We have presented a new method for the individual comparison of a patient diffusion MRI with respect to a population of healthy subjects. It is based on a generic framework to compare high dimensional models of diffusion such as Orientation Distribution Functions expressed in a modified spherical harmonics basis. We have demonstrated how such models allow for a better detection of differences with respect to simpler diffusion models as the classical diffusion tensor on simulated datasets.

We have applied this framework for the detection of white matter structure abnormalities at a very early stage of multiple sclerosis. This study performed on two time points right after the first CIS event, separated by three months, showed significant differences inside the lesions but also significant differences in ODF-related scores in regions where lesions were set to appear or disappear at M_3. This is a very significant result illustrating the potential prediction power of diffusion imaging for multiple sclerosis evolution, at the individual level, and could therefore lead to adapted treatments in the future based on single time points evaluation of the lesions.

Future works will further study this predictive power, namely by integrating evaluation of false positive detections among the results. Furthermore, other modalities may be of great interest in addition to diffusion imaging to characterize white matter focal and diffuse lesions, such as magnetization transfer imaging or relaxometry. We will further investigate how these modalities are providing complementary results and how to integrate them into one framework for the detection of abnormalities both at the patient and population level. In addition, these modalities could provide a further insight in white matter diffusion abnormalities, thereby helping to define a white matter degeneracy ground truth for the validation of our detection technique. Although not developed for MS lesion segmentation, our algorithm could also be coupled to other conventional modalities to help traditional MS lesion segmentation algorithms.

At a more technical level, we will also investigate more time points, and specifically designed ODF metrics such as the ones proposed by Goh et al. (2011). The influence of registration and normal variability among controls on the detection power could also be further studied, for example by adding typical transformations to each simulated database image, or by varying the simulated DWI intensities. Such an extensive comparison would bring further insight into our approach. Our comparison framework is also generic and could be applied to multi-compartment models such as multi-tensor (Scherrer and Warfield (2012)) or DDI (Stamm et al. (2012)), which may be more adapted to low angular acquisitions and do not consider only the directional

component of the diffusion.

Acknowledgments

The authors acknowledge Neurinfo MRI research facility for logistic support and MRI data acquisition. This work was supported by ARSEP, France (Association pour la Recherche de la Sclérose en Plaques).

References

Arsigny, V., Commowick, O., Pennec, X., Ayache, N., 2006a. A Log-Euclidean Framework for Statistics on Diffeomorphisms, in: Proc. of the 9th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'06), Part I, pp. 924–931.

Arsigny, V., Fillard, P., Pennec, X., Ayache, N., 2006b. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. Magnetic Resonance in Medicine 56, 411–421.

Assaf, Y., Basser, P., 2005. Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. Neuroimage 27, 48–58.

Blanco, M.A., Florez, M., Bermejo, M., 1997. Evaluation of the rotation matrices in the basis of real spherical harmonics. Journal of Molecular Structure 419, 19–27.

Commowick, O., Fillard, P., Clatz, O., Warfield, S.K., 2008. Detection of DTI white matter abnormalities in multiple sclerosis patients, in: Proceedings of the 11th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'08), Part I, pp. 975–982.

Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R., 2007. Regularized, Fast, and Robust Analytical Q-Ball Imaging. Magnetic Resonance in Medicine 58, 497–510.

Filippi, M., Cercignani, M., Inglese, M., Comi, M.H.G., 2001. Diffusion tensor magnetic resonance imaging in multiple sclerosis. Neurology 56, 304–311.

Geng, X., Ross, T.J., Zhan, W., Gu, H., Chao, Y.P., Lin, C.P., Christensen, G.E., Schuff, N., Yang, Y., 2009. Diffusion MRI Registration Using Orientation Distribution Functions, in: IPMI '09, pp. 626–637.

Giorgio, A., Palace, J., Johansen-Berg, H., Smith, S.M., Ropele, S., Fuchs, S., Wallner-Blazek, M., Enzinger, C., Fazekas, F., 2010. Relationships of brain white matter microstructure with clinical and MR measures in relapsing-remitting multiple sclerosis. Journal of Magnetic Resonance Imaging 31, 309–316.

Goh, A., Lenglet, C., Thompson, P.M., Vidal, R., 2011. A nonparametric Riemannian framework for processing high angular resolution diffusion images and its applications to ODF-based morphometry. Neuroimage 56, 1181–1201.

Grigis, A., Noblet, V., Heitz, F., Blanc, F., de Sèze, J., Kremer, S., Rumbach, L., Armspach, J.P., 2012. Longitudinal change detection in diffusion MRI using multivariate statistical testing on tensors. Neuroimage 60, 2206–2221.

Guimond, A., Meunier, J., Thirion, J., 2000. Average brain models: A convergence study. Computer Vision and Image Understanding 77, 192–210.

Jbabdi, S., Behrens, T.E., Smith, S.M., 2010. Crossing fibres in tract-based spatial statistics. Neuroimage 49, 249–256.

Lepore, N., Brun, C.A., Chou, Y.Y., Chiang, M.C., et al., 2008. Generalized tensor-based morphometry of HIV/AIDS using multivariate statistics on deformation tensors. IEEE Transactions on Medical Imaging 27, 129–141.

Liu, Y., Duan, Y., He, Y., Yu, C., Wang, J., Huang, J., Ye, J., Parizel, P.M., Li, K., Shu, N., 2012. Whole brain white matter changes revealed by multiple diffusion metrics in multiple sclerosis: A TBSS study. European Journal of Radiology 81, 2826–2832.

Mani, M., Srivastava, A., Barillot, C., 2013. Morphological changes in the corpus callosum: A study using a joint Riemannian feature spaces, in: SPIE Medical Imaging, p. 866908.

Ourselin, S., Roche, A., Prima, S., et al., 2000. Block matching: A general framework to improve robustness of rigid registration of medical images, in: Med Image Comput Comput Assist Interv, Springer. pp. 557–566.

Ritchie, D.W., Kemp, G.J., 1999. Fast computation, rotation and comparison of low resolution spherical harmonic molecular surfaces. Journal of Computational Chemistry 20, 383–395.

Rocca, M., Cercignani, M., Iannucci, G., Comi, G., Filippi, M., 2000. Weekly diffusion-weighted imaging of normal-appearing white matter in MS. Neurology 55, 882–4.

Rueckert, D., Sonoda, L., Hayes, C., et al., 1999. Nonrigid registration using free-form deformations: Application to breast MR images. IEEE Transactions on Medical Imaging 18, 712–721.

Scherrer, B., Warfield, S.K., 2012. Parametric Representation of Multiple White Matter Fascicles from Cube and Sphere Diffusion MRI. PLoS ONE 7, e48232.

Stamm, A., Pérez, P., Barillot, C., 2012. A new multi-fiber model for low angular resolution diffusion MRI, in: 9th IEEE International Symposium on Biomedical Imaging (ISBI), pp. 936–939.

Suarez, O., R., Commowick, O., Prabhu, P., S., Warfield, K., S., 2012. Automated delineation of white matter fiber tracts with a multiple region-of-interest approach. Neuroimage 59, 3690–3700.

Taquet, M., Scherrer, B., Commowick, O., Peters, J., Sahin, M., Macq, B., Warfield, K., S., 2014. A Mathematical Framework for the Registration and Analysis of Multi-Fascicle Models for Population Studies of the Brain Microstructure. IEEE Transactions on Medical Imaging 33, 504–517.

Werring, D.J., Brassat, D., Droogan, A.G., Clark, C.A., Symms, M.R., Barker, G.J., MacManus, D.G., Thompson, A.J., Miller, D.H., 2000. The pathogenesis of lesions and normal-appearing white matter changes in multiple sclerosis: a serial diffusion MRI study. Brain 123 (Pt 8), 1667–1676.

Whitcher, B., Wisco, J.J., Hadjikhani, N., Tuch, D.S., 2007. Statistical group comparison of diffusion tensors via multivariate hypothesis testing. Magnetic Resonance in Medicine 57, 1065–1074.

Zhang, H., Schneider, T., Wheeler-Kingshott, C., Alexander, D.C., 2012. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. Neuroimage 61, 1000–16.