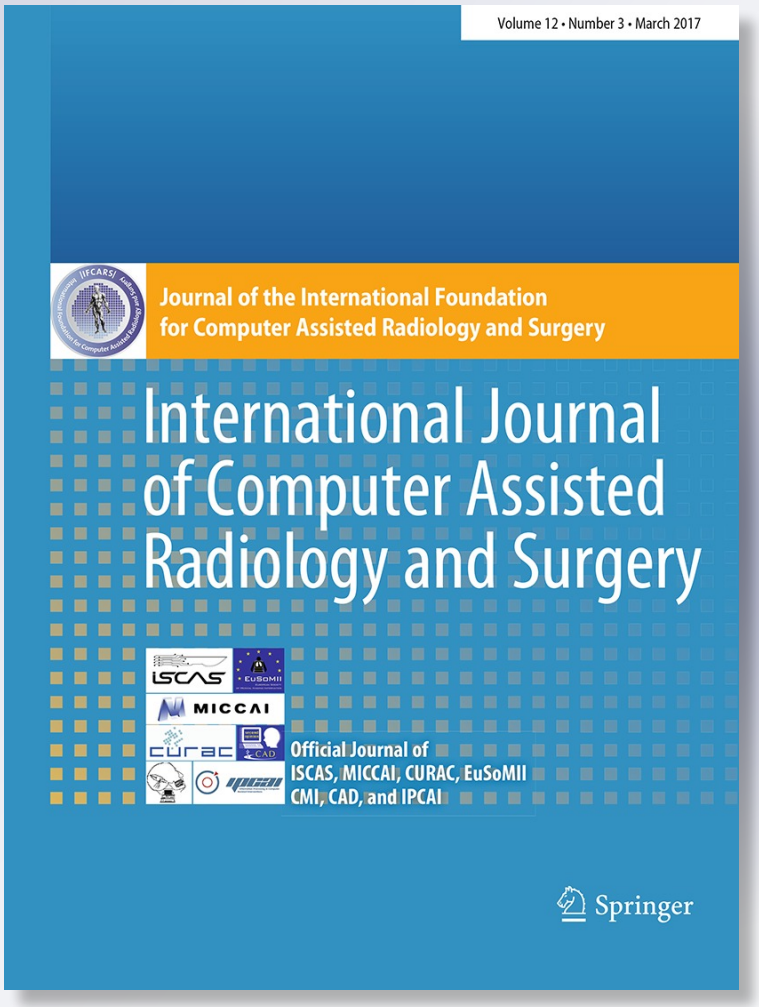
Local texture descriptors for the assessment of differences in diffusion magnetic resonance imaging of the brain

Felix Sebastian Leo Thomsen, Claudio Augusto Delrieux & Rodrigo de Luis-García

International Journal of Computer Assisted Radiology and Surgery
A journal for interdisciplinary research, development and applications of image guided diagnosis and therapy

ISSN 1861-6410
Volume 12
Number 3

Int J CARS (2017) 12:389-398
DOI 10.1007/s11548-016-1505-1



Springer

Int J CARS (2017) 12:389–398
DOI 10.1007/s11548-016-1505-1

ORIGINAL ARTICLE

Local texture descriptors for the assessment of differences in diffusion magnetic resonance imaging of the brain

Felix Sebastian Leo Thomsen[1] · **Claudio Augusto Delrieux**[1] · **Rodrigo de Luis-García**[2]

Received: 1 June 2016 / Accepted: 11 November 2016 / Published online: 21 November 2016

Abstract
Purpose Descriptors extracted from magnetic resonance imaging (MRI) of the brain can be employed to locate and characterize a wide range of pathologies. Scalar measures are typically derived within a single-voxel unit, but neighborhood-based texture measures can also be applied. In this work, we propose a new set of descriptors to compute local texture characteristics from scalar measures of diffusion tensor imaging (DTI), such as mean and radial diffusivity, and fractional anisotropy.
Methods We employ weighted rotational invariant local operators, namely standard deviation, inter-quartile range, coefficient of variation, quartile coefficient of variation and skewness. Sensitivity and specificity of those texture descriptors were analyzed with tract-based spatial statistics of the white matter on a diffusion MRI group study of elderly healthy controls, patients with mild cognitive impairment (MCI), and mild or moderate Alzheimer's disease (AD). In addition, robustness against noise has been assessed with a realistic diffusion-weighted imaging phantom and the contamination of the local neighborhood with gray matter has been measured.
Results The new texture operators showed an increased ability for finding formerly undetected differences between groups compared to conventional DTI methods. In particular, the coefficient of variation, quartile coefficient of variation, standard deviation and inter-quartile range of the mean and radial diffusivity detected significant differences even between previously not significantly discernible groups, such as MCI versus moderate AD and mild versus moderate AD. The analysis provided evidence of low contamination of the local neighborhood with gray matter and high robustness against noise.
Conclusions The local operators applied here enhance the identification and localization of areas of the brain where cognitive impairment takes place and thus indicate them as promising extensions in diffusion MRI group studies.

Keywords Local texture · Diffusion tensor imaging · Alzheimer's disease · White matter

✉ Felix Sebastian Leo Thomsen
Felix.Thomsen@uns.edu.ar

Claudio Augusto Delrieux
cad@uns.edu.ar

Rodrigo de Luis-García
RodLui@tel.uva.es

[1] CONICET and Departamento de Ingeniería Eléctrica y Computadoras, Universidad Nacional del Sur, Avenida Alem 1253, 8000 Bahía Blanca, Buenos Aires, Argentina

[2] ETSI Telecomunicación, Universidad de Valladolid, Campus Miguel Delibes, s/n, 47011 Valladolid, Spain

Introduction

Diffusion MRI is predominantly interpreted as a modality to estimate the spatial distribution of the diffusion of water molecules [2]. One common diffusion MRI model is diffusion tensor imaging (DTI), which describes the tensor of the diffusion directions. DTI found numerous applications in the study of neurosurgery [19], psychiatric [18] and neurological disorders [32] and is most commonly applied on the white matter (WM) of the brain [12] where water tends to diffuse preferentially parallel to fiber tracts because the myelin sheath and cell membranes restrict the diffusion perpendicular to the direction of the axons. Given that the diffusion tensor is a mathematical entity whose comparison and interpretation are not straightforward, most DTI group studies

rely on scalar measures which encode different rotationally invariant properties of the original diffusion tensor.

The most commonly employed measures are fractional anisotropy (FA) and mean diffusivity (MD) [29]. Also useful for describing the diffusion are radial and axial diffusivity [6], linear measure, tensor mode [20], and others [3,8,22]. The choice of suitable scalar measures is a critical step for DTI analyses, as their robustness to noise, or sensitivity and specificity to the pathological condition (i.e., ability to discriminate healthy from diseased patients, or between patients with different disease states) have a key role in the performance of the studies.

Most of the literature related to the derivation of new DTI scalar measures has focused on the behavior of different anisotropy measures with respect to noise. It is worth noting that most of the recently proposed anisotropy indices have similar contrast-to-noise ratio than FA [7,15]. Only few DTI methods, which incorporate a local neighborhood, have been described. The lattice index [26] measures the similarity of the predominant diffusion directions on a local weighted neighborhood in 2D. The inter-voxel local diffusion homogeneity [10] uses an unweighted 3D neighborhood. The local diffusion homogeneity might be sensitive to age-related changes in WM but less robust than MD or FA against motion artifacts in the MRI acquisition process [17]. Neither of these methods is rotational invariant.

The aim of this work is to present new DTI descriptors to improve the identification of subtle differences of the integrity of the white matter architecture between healthy and diseased brains (dementia, multiple sclerosis, stroke, aging, etc.). Our hypothesis is that well-defined texture operators, attached to existing scalar measures, can improve the identification of differences of the white matter architecture between patients with different levels of cognitive impairment. We propose a new set of image filtering methods directly on traditional scalar measures using rotationally invariant weighted local descriptors (average, standard deviation, coefficient of variation, skewness, etc.) to improve the discriminative power and robustness against noise. These statistical descriptors are inter-voxel in nature and may extract additional texture information that was not revealed by the original scalar measures, allowing to localize formerly undetected areas and opening a door to new viewpoints for the analysis of neuronal pathologies.

For the evaluation of the texture operators, we directly analyze the performance with a real-world DTI group study, instead of indirectly assessing the sensitivity and specificity to disease by means of the ability to distinguish between different types of brain tissue. We choose an Alzheimer's disease (AD) study with four distinct cohorts and restrict the final analysis to three input scalar measures MD, RD and FA, which were shown to differentiate well on AD [29]. Tract-based spatial statistics (TBSS) are employed as they have become widely accepted for such analyses [30]. In addition, we analyze the overlap of the neighborhood with gray matter and examine the performance of the local operators with respect to Rician noise.

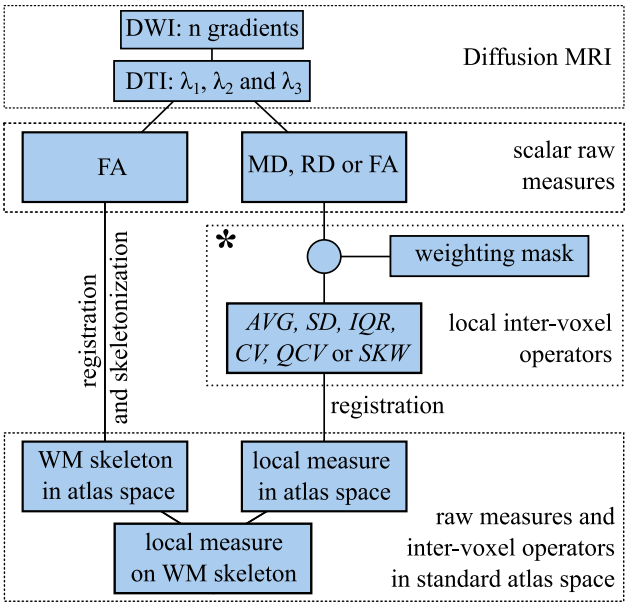


Fig. 1 DTI scalar measures (MD, RD or FA) are filtered with weighted local texture operators (∗). Then, they are registered to atlas space and projected to the WM skeleton, allowing to derive statistical comparisons between patients

Materials and methods

For the separate evaluation of each anatomical entity with TBSS, the following image processing steps are performed (Fig. 1). DTI is estimated from DWI volumes, and thereafter, the traditional single-voxel scalar measures (MD, RD or FA) are derived. Now, the texture descriptors are computed from the scalar measures using a local weighting mask. The final steps are the registration of the texture and non-texture measures to a standard atlas space and the projection to the registered WM skeleton, which is derived from FA. The following subsections review the traditional scalar measures in "Scalar measures" section, the proposed texture operators "Local texture operators" section and the conducted experiments (see "Analysis of diagnostic performance" and "Analysis of robustness" sections).

Scalar measures

The definitions of the scalar measures that were employed in this work are briefly reviewed. Given a diffusion tensor and its three eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$, the mean diffusivity is a measure of the total amount of diffusivity

$$\mathrm{MD} = \bar{\lambda} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}. \tag{1}$$

The radial diffusivity, on the other hand, measures how much diffusivity remains perpendicular to the main diffusion direction

$$\mathrm{RD} = \frac{\lambda_2 + \lambda_3}{2}, \tag{2}$$

and the fractional anisotropy measures how much the diffusion deviates from isotropy

$$\mathrm{FA} = \sqrt{\frac{3}{2}} \frac{\sqrt{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}. \tag{3}$$

The following additional scalar measures were computed, but finally discarded as they did not add much information with respect to the ones described in this work: linear, planar and spherical measure [37], anisotropy and elongation index [5] and an operator describing the DTI ellipsoid in terms of being oblate or prolate [36].

Local texture operators

Linear and nonlinear local texture operators are computed. Both incorporate a Gaussian weighting mask $\mathcal{N}_R$ and a binary volume of interest VOI. The weighting mask $\mathcal{N}_R$ is rotationally invariant and monotonic decreasing from the center voxel and thus minimizes remote influences of potential gray matter voxels within the neighborhood. It reads:

$$\mathcal{N}_R(\mathbf{r}) = \begin{cases} \exp\left(\frac{\|\mathbf{r}\|^2 (\Phi^{-1}(0.99))^2}{-2R^2}\right), & \text{if } \|\mathbf{r}\| \leq R \\ 0, & \text{otherwise,} \end{cases} \tag{4}$$

where $\Phi^{-1}(0.99)$ is the evaluation of the probit function at $p = 0.99$, $\mathbf{r}$ the 3D offset relative to voxel $\mathbf{x}$, $\|\cdot\|$ the Euclidean norm and R the isotropic radius. To compute meaningful 3D information, the mask must contain at least three non-trivial voxels in each direction, but must be small enough not to contaminate the texture information with gray matter voxels. Formulas for the linear operators are based on weighted cumulants [27]. First, the following parameters are defined:

$$V_i(\mathbf{x}) = (\mathrm{VOI} * \mathcal{N}_R^i)(\mathbf{x}), \text{ and} \tag{5}$$

$$X_i(\mathbf{x}) = ((S^i \cdot \mathrm{VOI}) * \mathcal{N}_R)(\mathbf{x}), \tag{6}$$

where VOI contains ones inside and zeros outside the brain, S is the input scalar measure, $\cdot$ is the point-wise multiplication, $*$ is the convolution operator and $i \geq 1$ is an integer number. Now, the weighted cumulants read

$$K_1(\mathbf{x}) = (X_1 / V_1)(\mathbf{x}), \tag{7}$$

$$K_2(\mathbf{x}) = \left(\frac{X_2 V_1 - X_1}{V_1^2 - V_2}\right)(\mathbf{x}), \text{ and} \tag{8}$$

$$K_3(\mathbf{x}) = \left(\frac{X_3 V_1^2 - 3X_1 X_2 V_1 - 2X_1^3}{V_1^3 - 3V_1 V_2 + 2V_3}\right)(\mathbf{x}). \tag{9}$$

The nonlinear operators are computed from weighted quantiles. For each voxel $\mathbf{x}$, let $\mathbf{r}_1, \ldots, \mathbf{r}_M$ be the labeling of all 3D offsets of $\mathcal{N}_R$ which induces the ascending sorting of the values of S: $\forall i \in \{1, \ldots, M-1\} : S(\mathbf{x} + \mathbf{r}_i) \leq S(\mathbf{x} + \mathbf{r}_{i+1})$. The weighted quantile at a given percentile $Pr \in [0, 1]$ reads now:

$$Q_{Pr}(\mathbf{x}) = S(\mathbf{x} + \mathbf{r}_j) : \sum_{i=1}^{j-1} \mathcal{N}_R^*(\mathbf{r}_i) < Pr \leq \sum_{i=1}^{j} \mathcal{N}_R^*(\mathbf{r}_i) \tag{10}$$

with $\mathcal{N}_R^*$ the normalized weighting mask ($\sum_{i=1}^{M} \mathcal{N}_R^*(\mathbf{r}_i) = 1$). From the definitions above, six scalar measures are defined; the weighted average is only included for completeness, though not new in DTI imaging:

- the weighted average

$$\mathrm{AVG}(\mathbf{x}) = K_1(\mathbf{x}), \tag{11}$$

- the weighted standard deviation

$$\mathrm{SD}(\mathbf{x}) = \sqrt{K_2(\mathbf{x})}, \tag{12}$$

- the weighted coefficient of variation

$$\mathrm{CV}(\mathbf{x}) = (\sqrt{K_2} / K_1)(\mathbf{x}), \tag{13}$$

- the weighted standardized skewness

$$\mathrm{SKW}(\mathbf{x}) = (K_3 / K_2^{(3/2)})(\mathbf{x}), \tag{14}$$

- the weighted inter-quartile range

$$\mathrm{IQR}(\mathbf{x}) = (Q_{75\%} - Q_{25\%})(\mathbf{x}), \tag{15}$$

- and the weighted quartile coefficient of variation

$$\mathrm{QCV}(\mathbf{x}) = \left(\frac{Q_{75\%} - Q_{25\%}}{Q_{25\%} + Q_{75\%}}\right)(\mathbf{x}). \tag{16}$$

Figure 2 shows axial views of all scalar (RAW) and neighborhood-based operators for a certain sample case. AVG contains the smoothed information of RAW. The parameters IQR and QCV represent nonlinear counterparts of

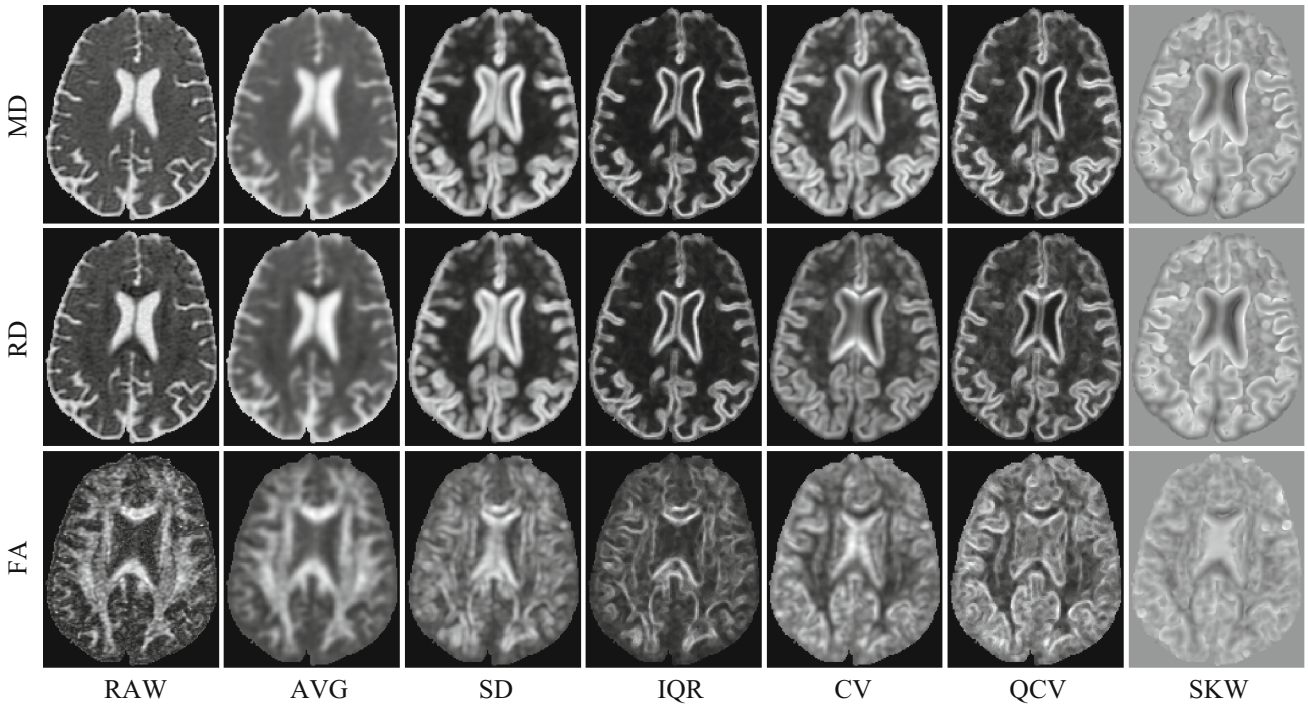


Fig. 2 DTI measures obtained from a healthy patient (voxel size $1.015 \times 1.015 \times 3\,\text{mm}^3$, matrix is cropped to 131×162 voxels) obtained with $\mathcal{N}_{4\text{mm}}$ (mask size $7 \times 7 \times 3$). *Rows* MD, RD and FA. *Columns* Single-voxel measure (RAW), average (AVG), texture operators standard deviation (SD), inter-quartile range (IQR), coefficient of variation (CV), quartile coefficient of variation (QCV) and skewness (SKW)

SD and CV. SD and IQR emphasize areas with a high local gradient similar to their scale invariant counterparts CV and QCV. The third moment SKW reveals different information than the other operators, which becomes visually apparent in the case of its application to MD and RD. However, in the case of FA, SKW equals to a certain extent CV and QCV. The weighted standardized kurtosis was derived in a similar manner to the other linear operators, but later excluded due to its weak sensitivity and specificity between the tested cohorts.

Analysis of diagnostic performance

A DTI group study with four age-adjusted cohorts was analyzed ($N = 58$). The study contained healthy controls (HC: $N = 19$, age $= 74.05$ years, SD $= 0.88$ years), patients with mild cognitive impairment (MCI: $N = 12$, age $= 76.33$ years, SD $= 1.11$ years), patients with mild Alzheimer's disease (Mild AD: $N = 20$, age $= 76.30$ years, SD $= 0.86$ years) and patients with moderate Alzheimer's disease (Mod AD: $N = 7$, age $= 76.57$ years, SD $= 1.45$ years). Raw diffusion-weighted images (DWIs) were acquired with a GE Signa 1.5 T MRI unit at QDiagnóstica radiological facilities in Valladolid, Spain. The parameters of the acquisition protocol were the following: 25 gradient directions, $b = 1000\,\text{s/mm}^2$, one baseline volume, 128×128 acquisition matrix zero-padded to 256×256 prior to reconstruction, reconstructed voxel size $1.015 \times 1.015 \times 3\,\text{mm}^3$, TR $= 13,000$ ms, TE $= 85.5$ ms, NEX $= 2$ and 39 slices covering the entire brain.

The DWIs were linearly registered to the baseline volume to account for possible minor motion artifacts. Then, they were processed with an algorithm based on the Otsu threshold [24] to remove the image background as well as non-brain structures such as the skull. Diffusion tensors were estimated using a weighted least squares method [28], and the image quality was individually checked on the tensor volumes using color by orientation maps. From the tensor volume, the scalar measures and local operators were computed ($\mathcal{N}_{4\text{mm}}$, mask size $7 \times 7 \times 3$) and TBSS analysis [30] was performed using FSL 4.1 (FMRIB software library, http://www.fmrib.ox.ac.uk/fsl [13]). By using this methodology, FA volumes were non-linearly registered to the MNI 152 standard space (voxel size $1 \times 1 \times 1\,\text{mm}^3$ [11]). Then, a WM skeleton was created from the thresholded mean FA map (FA ≥ 0.2). Finally, all original scalar measures and the ones derived from local operators were projected onto the WM skeleton for statistical comparisons.

Statistical analyses for TBSS were carried out using RANDOMISE, an FSL tool that performs permutations for inference on statistical maps when the null distribution is unknown [21]. This allows to measure the precision and

specificity, similar to conventional tests of receiver operator characteristics, but with increased sensitivity to differences in a specific direction [40]. Corrections for family-wise errors and multiple comparisons with threshold-free cluster enhancement [31] were performed. Pair-wise comparisons were made for each combination of different cohorts and for all measures (original measures and the ones proposed in this paper) in order to obtain the number of voxels for which significant differences ($p < 0.05$; $p < 0.1$) were found among groups. The significant levels of every voxel in atlas space have been separately adjusted for multiple comparisons with the Benjamini–Hochberg procedure [4]. This correction was applied for $M = 21$ comparisons (MD, RD, FA in combination with the 7 operators RAW, AVG, STD, etc.), while the 6 combinations of cohorts (HC versus MCI, HC versus Mild AD etc.) were analyzed separately and thus not included in the Benjamini–Hochberg correction.

Analysis of robustness

In order to evaluate the performance of the proposed new scalar measures with respect to noise, noise-controlled DTI volumes were employed based on a realistic phantom. Following the approach described in [33], a real dataset was denoised [34] to create a ground truth with 16 gradient directions, $b = 1200\,\mathrm{s/mm^2}$, one baseline volume, reconstructed voxel size $0.9375 \times 0.9375 \times 1.7\,\mathrm{mm^3}$, 128×128 acquisition matrix zero-padded to 256×256 prior to reconstruction and 81 slices covering the entire brain. Then, the noiseless DWI (D_∞) was interfered with Rician noise of eleven different signal-to-noise ratios ranging in amplitude scale from SNR = 3.16 to SNR = 56.23. The SNR's have been chosen to be uni-distant in (logarithmic) decibel scale with a spacing of 2.5 dB and covering the interval [10 dB, 35 dB]. The noisy DWI signal D_{SNR} read:

$$D_{\mathrm{SNR}} = \sqrt{(D_\infty + \eta(E/\mathrm{SNR}))^2 + \eta'(E/\mathrm{SNR})^2}, \qquad (17)$$

where E and SNR were both expressed in amplitude scale and the expected value E was the mean value of the noiseless DWIs. The functions $\eta(\sigma)$ and $\eta'(\sigma)$ were two Gaussian random variables with zero mean and standard deviation σ. The traditional and novel scalar measures ($\mathcal{N}_{4\mathrm{mm}}, 9 \times 9 \times 5$ voxels) were obtained from the resulting DTI volumes, and TBSS was applied, thus obtaining a thresholded FA skeleton (noiseless FA ≥ 0.2), over which the evolution in noisy scenarios was analyzed. Quantitative comparisons between operators and SNRs were derived from the normalized root mean square error

$$\mathrm{NRMSE}(X_{\mathrm{SNR}}, X_\infty) = \frac{\mathrm{RMSE}(X_{\mathrm{SNR}}, X_\infty)}{\mathrm{SD}(X_\infty)}, \qquad (18)$$

with $\mathrm{RMSE}(X_{\mathrm{SNR}}, X_\infty)$ the root mean square error between the noisy X_{SNR} and the noiseless signal X_∞, and $\mathrm{SD}(X_\infty)$ the standard deviation of the noiseless signal. The normalization with the standard deviation allows for the optimum comparison between the different tested operators: it does not require ratio variables (in contrast to a normalization with the mean) and is robust against outliers (in contrast to a normalization with the range). In particular, the variables RAW and AVG of FA on the white matter and SKW require this kind of normalization. Since the measures SNR and RMSE express amplitude ratios, the transformation to dB scale was performed with the formula for root power quantities.

The analysis of the contamination with gray matter was obtained from all 58 patients of the AD study. A histogram was computed of the minimum amount of white matter in the neighborhoods of the TBSS. The proposed mask $\mathcal{N}_{4\mathrm{mm}}$ was compared with three different unweighted $3 \times 3 \times 3$ masks used for the computation of the inter-voxel local diffusion homogeneity [10]: the neighborhood N_{27} contained all 27 voxels, while the other contained only voxels sharing an edge (N_{19}) or side (N_7) with the center voxel.

Results

Analysis of diagnostic performance

No significant differences ($p < 0.1$) were found between HC versus MCI or between MCI versus Mild AD. Except for FA RAW and AVG, parameters between other cohorts increased with progressing dementia: HC < Mild AD < Mod AD and MCI < Mod AD. None of the parameters contained significantly increased and decreased areas at the same moment. The traditional measures (RAW or AVG) found significant differences only between healthy patients and those with AD, while the texture operators additionally located change at MCI < Mod AD and between the two AD cohorts (Mild AD < Mod AD), as shown in Fig. 3. Note that although MD and RD consistently increased with disease severity and FA decreased, less voxels with significant differences were found on the RAW and AVG operators in the comparison between HC versus Mod AD compared to HC versus Mild AD. This might have been caused by the limited statistical power of this comparison due to the small number of patients in Mod AD ($N = 7$).

The application of the operators to MD and RD followed a widely similar pattern, both, in regard to the number and in regard to their spatial distribution. With regard to MD, SD, IQR, CV and QCV discriminated on four pairs, while the other operators only at two (SKW and AVG) or one pair of groups (RAW). MD SD and CV detected similar locations; however, MD CV revealed stronger indications of impairments between Mild AD and Mod AD on the sensory cortex

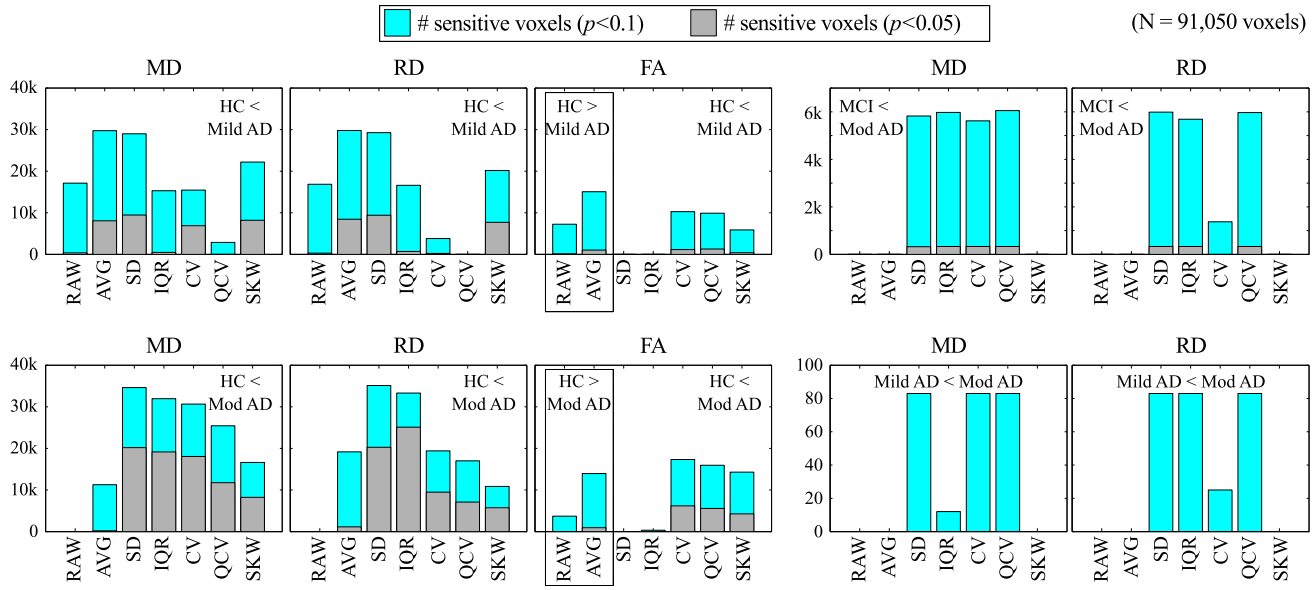


Fig. 3 Number of significant different TBSS voxels ($N = 91,050$, $p < 0.1$ and $p < 0.05$, corrected for multiple comparisons) between HC versus Mild AD (*top left*), HC versus Mod AD (*bottom left*), MCI versus Mod AD (*top right*) and Mild AD versus Mod AD (*bottom right*). The subset of voxels which were significant with $p < 0.05$ are shaded in *gray*

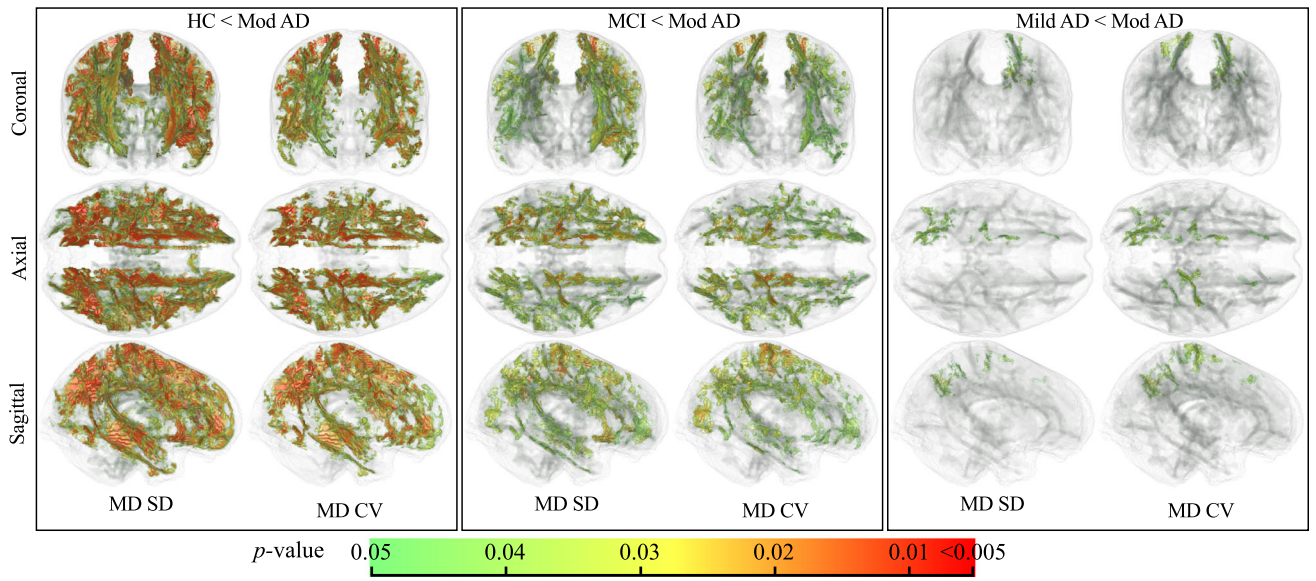


Fig. 4 Spatial distribution of significant differences of MD SD and MD CV, p Values have been adjusted for multiple comparison correction with threshold-free cluster enhancement and family-wise errors. Groups are HC < Mod AD, MCI < Mod AD and Mild AD < Mod AD where MD RAW was not discriminative ($p < 0.05$)

and parietal lobe, as shown in Fig. 4. The application of higher order moments (SD, CV, SKW) on FA did not improve the diagnostic performance of FA AVG.

Analysis of robustness

The most of the local operators had higher robustness against noise than the single-voxel RAW operator, as shown in Fig. 5. The noise characteristics obtained on MD and RD were very similar. Applied to MD or RD, AVG increased the robustness against noise constantly by a factor of 1.8, while FA AVG increased the robustness depending on the SNR by up to factor of 4, compared to RAW. Also the local texture operators of MD and RD were more robust than RAW, with the exception of SD for high SNRs and SKW for low SNRs. The

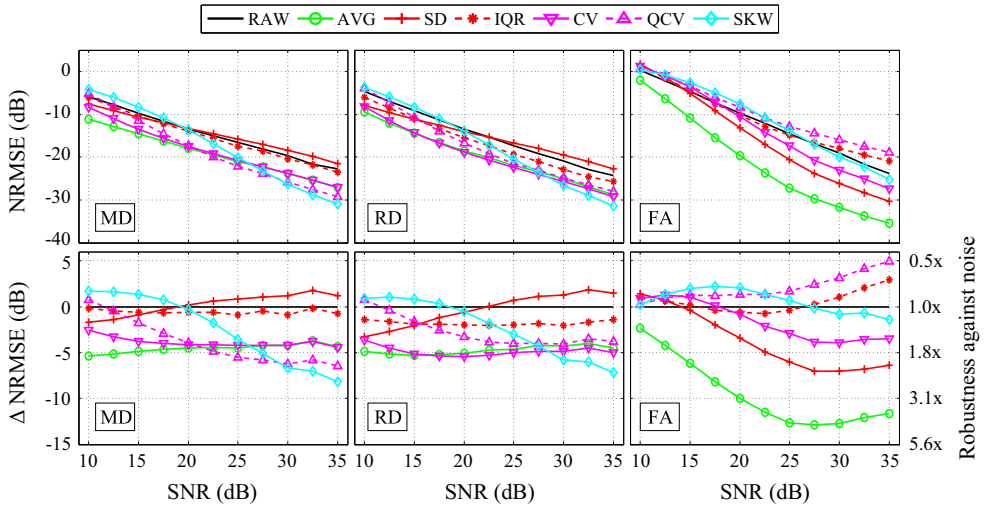


Fig. 5 Noise analysis with NRMSE versus eleven SNRs $\in$ [10 dB, 35 dB]. *1. Row* NRMSE(dB) of MD, RD and FA. *2. Row* Δ NRMSE as the difference of NRMSE(dB) of the texture operators to the NRMSE(dB) of the non-texture operator RAW

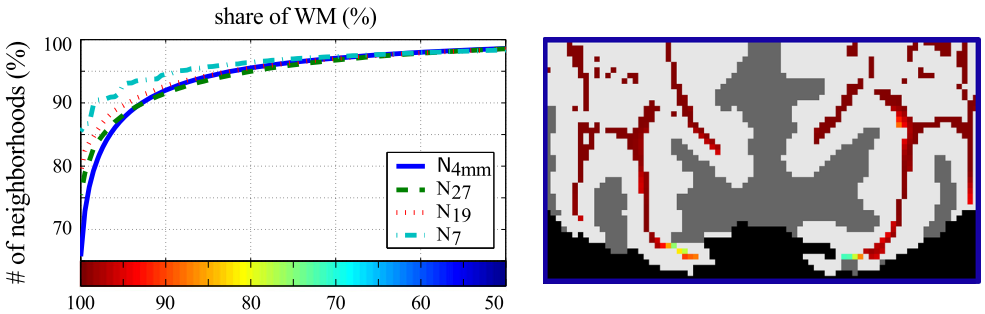


Fig. 6 Percentage of WM voxels inside the evaluated neighborhoods, color-coded from 50 to 100%. *Left* cumulative histogram of $\mathcal{N}_{4\text{mm}}$, mask size ($9 \times 9 \times 5$) in comparison with N_{27}, N_{19} and N_7 (mask size $3 \times 3 \times 3$). *Right* Spatial distribution of $\mathcal{N}_{4\text{mm}}$ on the posterior of the center axial slice

robustness against noise of MD/RD CV, *QCV* and for high SNRs of SKW increased by a factor 1.8 to 2.2, exceeding also the robustness of MD/RD AVG. If applied to FA, SD and CV were only robust for high SNRs (2.2 and, respectively, 1.6 times of RAW), while IQR, QCV and SKW were rather sensitive to noise.

The comparison of $\mathcal{N}_{4\text{mm}}$ with unweighted $3 \times 3 \times 3$ neighborhoods N_{27}, N_{19} and N_7 is shown in Fig. 6. Although the spatial dimension of $\mathcal{N}_{4\text{mm}}$ exceeded the one of the unweighted neighborhoods, its contamination with gray matter was still low: approximately 92% of all neighborhoods were at least to 90% in WM. Furthermore, $\mathcal{N}_{4\text{mm}}$ contained more neighborhoods with a minimum white matter share of 93% (N_{27}), 80% (N_{19}) and 65% (N_7), respectively.

Discussion

The experimental work performed here focused on the analysis of two important aspects: the ability to discriminate between pathological conditions in a group study and the evolution with varying SNR. With regard to the first one, results showed that some of the proposed local texture operators highly increased the ability to differentiate between groups. Concerning the second one, most parameters obtained high robustness against noise, some texture operators even exceeded the robustness of AVG.

A DTI data set with multiple stages of Alzheimer's disease was selected as an experimental benchmark because it allowed to quantify the potential gain of the new operators. Indeed, new differences between groups arose when employing the proposed approach, which is a strong advantage over traditional measures. In particular, the texture parameters SD, CV, IQR and QCV, if applied on MD or RD, exceeded the performance of RAW and even AVG. For instance, CV and QCV obtained generally higher robustness to noise and higher discriminative potential than AVG. When considering noise and discriminative analysis together, local texture parameters of MD and RD were more adequate for the conducted AD group studies than those based on FA. AVG was

the only neighborhood-based standard DTI operator among the proposed ones, since weighted averaging is commonly employed in voxel-based morphometry [1] and TBSS [30]; it obtained high robustness to noise, but was, unlike other texture parameters, unable to reveal differences between MCI versus Mod AD or Mild AD versus Mod AD.

In general, also alternative statistical analyses could be performed, such as the general voxel-based (VBA) or atlas-based analysis (ABA), which do not require skeletonization of the WM [23]. However, TBSS analysis was used in this work because of two reasons. First, as TBSS is a widely employed DTI analysis tool [38], it facilitates the interpretation and comparison of the results among other approaches. Second, since TBSS centers the neighborhoods on the WM skeleton, it minimizes any contamination with gray matter.

Still, the mask size must be selected carefully according to the spatial resolution. If the mask exceeds noticeably the distance from the WM skeleton to gray matter voxels, local features become mainly dependent on changes in local morphometry, which decreases the discriminative power particularly at the boundaries of the WM. In terms of contamination with gray matter, $\mathcal{N}_{4\text{mm}}$ contained similar characteristics than the unweighted reference masks, while being advantageous in terms of sample size and isotropy: in contrast to $\mathcal{N}_{4\text{mm}}$, the reference masks N_{27}, N_{19} and N_7 are anisotropic ($3 \times 3 \times 9\,\text{mm}^3$) and thus favor longitudinal fiber tracts over transversal or sagittal ones.

The interpretation of the results may be difficult compared to traditional measures, which is a common issue of texture operators [10]. Differences in traditional scalar measures, such as FA, have been often related to physiological changes in the WM tissue (demyelination, lower packing density or different membrane permeability, among others) [16]. However, changes in FA, MD or RD can be caused by very different mechanisms within the tissue, which makes bold interpretations of changes in the WM "integrity," although very common, possibly flawed [14]. The use of the proposed local operators, on the other hand, can open a door to new viewpoints for the analysis of pathologies. For instance, MD CV discovered a progressing impairment from mild to moderate AD of brain areas, which are commonly related to sensations (sensory cortex), perception, spelling and arithmetic (parietal lobe). These impairments remained hidden on standard parameters (RAW and AVG).

Head motion during diffusion MRI acquisition is a known source of image artifacts. As indicated in "Analysis of diagnostic performance" section, the DWIs were linearly registered to the baseline volume before tensor estimation, and image quality was individually checked on the tensor volumes using color by orientation maps. However, undesired attenuation in single DWI slices due to motion during the diffusion-encoding gradient pulse cannot be completely excluded. In principle, these motion effects could lead to artifacts in the scalar measures obtained from the diffusion tensor, such as an underestimation of anisotropy indices. Even though previous studies in the literature have reported possible false positive findings in group studies due to this effect and have even described how these artifacts can be more prominent in some neural pathways [39], there is no indication that the proposed measures could be more affected by motion artifacts than the original RAW measures. As these motion artifacts can be expected to behave in a somehow smooth manner across the brain, it is even possible that the measures based on local variability are more robust to motion artifacts than the original measures, but a detailed analysis of this effect would be needed to elucidate this.

The application of the proposed texture operators is not limited to DTI, but could be applied to any scalar measure. In particular, the application on Q-Ball scalar measures from high angular resolution diffusion imaging might be investigated, for instance the generalized anisotropy [25], generalized fractional anisotropy [35] or fractional multi-fiber index [9]. Studies with higher spatial resolution than the one employed here could achieve improvements regarding the robustness and localization of the changed areas. The MATLAB code with a sample script that computes the local operators is available at members.imaglabs.org/felix.thomsen/LocalOperatorsMRI/Matlab.zip.

Conclusion

In this work, we presented new image processing methods for the identification and characterization of changes in the white matter of the brain based on DTI. Instead of deriving new scalar measures directly from DWI, we further processed existing DTI scalar measures such as FA, MD and RD with rotational invariant texture operators. Each combination of existing scalar measures and texture operators provided a novel measure with new and distinct properties, which was in many cases more suitable than the original single-voxel measure for particular discrimination problems of DTI studies. Compound sensitivity and specificity of the operators were compared with the one of the traditional measures on a group study of Alzheimer's disease. Additionally, noise characteristics were measured with a synthetic DWI phantom and the overlap of the local operators with gray matter was analyzed.

Texture parameters based on MD and RD yielded the highest discriminative power and robustness against noise and were able to identify formerly undetected changes even between previously insignificant pairs of groups Mild AD versus Mod AD. Thus, the results indicate that the methods are a promising extension in DTI studies. Future work might investigate the application of the proposed methods on group

studies of different alterations of the brain (multiple sclerosis, schizophrenia, etc.) to improve the detection and localization of the changes between groups.

Acknowledgements The authors acknowledge the company QDiagnóstica, Valladolid, Spain, whose facility has been used for data acquisition purposes and thank Dr. Miguel Angel Tola-Arribas for his valuable help with the recruitment and diagnosis of patients.

Funding Felix Sebastian Leo Thomsen received a doctoral fellowship from Consejo Nacional de Investigaciones Científicas y Técnicas of Argentina (CONICET). This work was partially funded by the Universidad Nacional del Sur (PGI 24/K061). Rodrigo de Luis García has received research grants from Ministerio de Ciencia e Innovación of Spain (TEC 2013-44194-P), Fondo de Investigaciones Sanitarias (PI 11-01492) the Consejería de Sanidad de Castilla y León (BIO/VA30/14).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Human and animals right All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Ashburner J, Friston KJ (2000) Voxel-based morphometry—the methods. Neuroimage 11(6):805–821
2. Basser PJ, Mattiello J, LeBihan D (1994) Estimation of the effective self-diffusion tensor from the NMR spin echo. J Magn Reson Ser B 103(3):247–254
3. Basser PJ, Pierpaoli C (1996) Microstructural and physiological features of tissues elucidated by quantitative-diffusion tensor MRI. J Magn Reson Ser B 111:209–219
4. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 57(1):289–300
5. Benn DI (1994) Fabric shape and the interpretation of sedimentary fabric data. J Sediment Res 64(4):910–915
6. Budde MD, Xie M, Cross AH, Song SK (2009) Axial diffusivity is the primary correlate of axonal injury in the experimental autoimmune encephalomyelitis spinal cord: a quantitative pixelwise analysis. J Neurosci 29(9):2805–2813
7. Chao Wang JJ, Tc Wai YY, Hsu Y (2006) Novel diffusion anisotropy indices: an evaluation. J Magn Reson Imaging 24(1):211–217
8. Ennis DB, Kindlmann G (2006) Orthogonal tensor invariants and the analysis of diffusion tensor magnetic resonance images. Magn Reson Med 55(1):136–146
9. Frank LR (2002) Characterization of anisotropy in high angular resolution diffusion-weighted MRI. Magn Reson Med 47(6):1083–1099
10. Gong G (2013) Local diffusion homogeneity (LDH): an inter-voxel diffusion MRI metric for assessing inter-subject white matter variability. PLoS ONE 8(6):1–11
11. Grabner G, Janke AL, Budge MM, Smith D, Pruessner J, Collins DL(2006) Symmetric atlasing and model based segmentation: an application to the hippocampus in older adults. In: Medical image computing and computer-assisted intervention—MICCAI 2006. Springer, pp 58–66
12. Horsfield MA, Jones DK (2002) Applications of diffusion-weighted and diffusion tensor MRI to white matter diseases—a review. NMR Biomed 15(7–8):570–577
13. Jenkinson M, Beckmann C, Behrens T, Woolrich M, Smith S (2012) FSL. Neuroimage 62(2):782–790
14. Jones DK, Knösche TR, Turner R (2013) White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. Neuroimage 73:239–254
15. Kingsley PB, Monahan WG (2005) Contrast-to-noise ratios of diffusion anisotropy indices. Magn Reson Med 53(4):911–918
16. Kochunov P, Thompson P, Lancaster J, Bartzokis G, Smith S, Coyle T, Royall D, Laird A, Fox P (2007) Relationship between white matter fractional anisotropy and other indices of cerebral health in normal aging: Tract-based spatial statistics study of aging. Neuroimage 35(2):478–487
17. Kong XZ (2014) Association between in-scanner head motion with cerebral white matter microstructure: a multiband diffusion-weighted MRI study. Peer J 2:e366
18. Kubicki M, McCarley R, Westin CF, Park HJ, Maier S, Kikinis R, Jolesz FA, Shenton ME (2007) A review of diffusion tensor imaging studies in schizophrenia. J Psychiatr Res 41(1):15–30
19. Lerner A, Mogensen MA, Kim PE, Shiroishi MS, Hwang DH, Law M (2014) Clinical applications of diffusion tensor imaging. World Neurosurg 82(1):96–109
20. de Luis-García R, Cabús-Piñol G, Imaz-Roncero C, Argibay-Quiñones D, Barrio-Arranz G, Aja-Fernández S, Alberola López C (2015) Attention deficit/hyperactivity disorder and medication with stimulants in young children: a DTI study. Prog Neuro-Psychopharmacol Biol Psychiatry 57:176–184
21. Nichols TE, Holmes AP (2002) Nonparametric permutation tests for functional neuroimaging: a primer with examples. Human Brain Mapp 15(1):1–25
22. O'Donnell LJ, Westin CF (2011) An introduction to diffusion tensor image analysis. Neurosurg Clin N Am 22(2):185–196
23. Oishi K, Faria A, Jiang H, Li X, Akhter K, Zhang J, Hsu JT, Miller MI, van Zijl PC, Albert M, Lyketsos CG, Woods R, Toga AW, Pike GB, Rosa-Neto P, Evans A, Mazziotta J, Mori S (2009) Atlas-based whole brain white matter analysis using large deformation diffeomorphic metric mapping: application to normal elderly and Alzheimer's disease participants. Neuroimage 46(2):486–499
24. Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 9(1):62–66
25. Özarslan E, Vemuri BC, Mareci TH (2005) Generalized scalar measures for diffusion mri using trace, variance, and entropy. Magn Reson Med 53(4):866–876
26. Pierpaoli C, Basser PJ (1996) Toward a quantitative assessment of diffusion anisotropy. Magn Reson Medic 36(6):893–906
27. Rimoldini L (2014) Weighted skewness and kurtosis unbiased by sample size and Gaussian uncertainties. Astron Comput 5:1–8
28. Salvador R, Peña A, Menon DK, Carpenter TA, Pickard JD, Bullmore ET (2005) Formal characterization and extension of the linearized diffusion tensor model. Human Brain Mapp 24(2):144–155
29. Sexton CE, Kalu UG, Filippini N, Mackay CE, Ebmeier KP (2011) A meta-analysis of diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease. Neurobiol Aging 32(12):2322.e5–2322.e18
30. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM, Behrens TEJ (2006) Tract-based spatial statistics: voxelwise

analysis of multi-subject diffusion data. Neuroimage 31(4):1487–1505
31. Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44(1):83–98
32. Stebbins G, Murphy C (2009) Diffusion tensor imaging in Alzheimer's disease and mild cognitive impairment. Behav Neurol 21(1–2):39–49
33. Tristán-Vega A, Aja-Fernández S (2009) Design and construction of a realistic DWI phantom for filtering performance assessment. In: Medical image computing and computer-assisted intervention—MICCAI 2009. Springer, pp 951–958
34. Tristán-Vega A, Aja-Fernández S (2010) DWI filtering using joint information for DTI and HARDI. Med Image Anal 14(2):205–218
35. Tuch DS (2004) Q-ball imaging. Magn Reson Med 52(6):1358–1372
36. Vasilić B, Rajapakse CS, Wehrli FW (2009) Classification of trabeculae into three-dimensional rodlike and platelike structures via local inertial anisotropy. Med Phys 36(7):3280–3291
37. Westin CF, Maier SE, Mamata H, Nabavi A, Jolesz FA, Kikinis R (2002) Processing and visualization for diffusion tensor MRI. Med Image Anal 6(2):93–108
38. Wise T, Radua J, Nortje G, Cleare AJ, Young AH, Arnone D (2016) Voxel-based meta-analytical evidence of structural disconnectivity in major depression and bipolar disorder. Biol Psychiatry 79(4):293–302
39. Yendiki A, Koldewyn K, Kakunoori S, Kanwisher N, Fischl B (2014) Spurious group differences due to head motion in a diffusion MRI study. Neuroimage 88:79–90
40. Zou KH, Liu A, Bandos AI, Ohno-Machado L, Rockette HE (2011) Statistical evaluation of diagnostic performance: topics in ROC analysis. CRC Press, Boca Raton