More Accurate Estimation of Diffusion Tensor Parameters Using Diffusion Kurtosis Imaging

Jelle Veraart,[1*] Dirk H. J. Poot,[1,2] Wim Van Hecke,[3,4] Ines Blockx,[5] Annemie Van der Linden,[5] Marleen Verhoye,[5] and Jan Sijbers[1]

With diffusion tensor imaging, the diffusion of water molecules through brain structures is quantified by parameters, which are estimated assuming monoexponential diffusion-weighted signal attenuation. The estimated diffusion parameters, however, depend on the diffusion weighting strength, the *b*-value, which hampers the interpretation and comparison of various diffusion tensor imaging studies. In this study, a likelihood ratio test is used to show that the diffusion kurtosis imaging model provides a more accurate parameterization of both the Gaussian and non-Gaussian diffusion component compared with diffusion tensor imaging. As a result, the diffusion kurtosis imaging model provides a *b*-value-independent estimation of the widely used diffusion tensor parameters as demonstrated with diffusion-weighted rat data, which was acquired with eight different *b*-values, uniformly distributed in a range of [0,2800 sec/mm^2]. In addition, the diffusion parameter values are significantly increased in comparison to the values estimated with the diffusion tensor imaging model in all major rat brain structures. As incorrectly assuming additive Gaussian noise on the diffusion-weighted data will result in an overestimated degree of non-Gaussian diffusion and a *b*-value-dependent underestimation of diffusivity measures, a Rician noise model was used in this study. Magn Reson Med 65:138–145, 2011. © 2010 Wiley-Liss, Inc.

Key words: DKI; likelihood ratio test; *b*-value dependency; parameter estimation

Diffusion tensor magnetic resonance imaging (DTI) is an important medical imaging modality in neuroscience research, because it allows the study of the complex network of myelinated axons, in vivo and noninvasively (1,2). In DTI, the diffusion of water molecules through brain structures is mathematically described by a second order 3D diffusion tensor (DT). It is generally accepted that the first eigenvector of the tensor, corresponding to the direction of maximal diffusion, is aligned with the underlying fiber structures. Furthermore, the diffusion is often quantified with diffusion parameters (i.e., fractional anisotropy (FA) and mean (MD), radial ($D_\perp$) and axial ($D_\parallel$) diffusivity), which provide insight in the organization, structural integrity, and development of white matter (WM) structures of the normal and pathological brain (3–9).

In DTI, the diffusion of water molecules along a certain gradient direction is assumed to occur in an unrestricted environment. Consequently, the molecules' probability of diffusing from one location to another in a given time is described by a Gaussian distribution of which the standard deviation relates to the apparent diffusion coefficient (ADC). As a result, the normalized diffusion-weighted signal that is measured along a certain axis can be described by a monoexponential function; the exponent equals the ADC, weighted by the diffusion weighting strength that is given by the *b*-value. Several DTI studies, however, reported that the estimation of diffusion parameters depends on the *b*-value that is used during data acquisition. Therefore, the comparison and interpretation of various DTI studies are hampered. Jones and Basser (10) and Andersson (11) attributed the *b*-value dependency to the use of an inaccurate, Gaussian noise model while estimating diffusion parameters with a (weighted) least squares estimator, as MR images are corrupted with Rician noise (12,13). Other related work addressed the *b*-value dependency of the quantification of DTI measures in biological tissue to the complex relation between the diffusion-weighted signal and the *b*-value due to factors such as cerebral perfusion, restricted diffusion, membrane permeability, and extra- and intracellular water compartments (8,14, 15). As a result, the diffusion will appear non-Gaussian and hence cannot be approximated accurately by a DTI model (8,16,17).

Recently, Jensen et al. (17) and Lu et al. (18) introduced diffusion kurtosis imaging (DKI), a higher order diffusion model that is a straightforward extension of the DTI model. DKI approximates the diffusion-weighted signal attenuation more accurately by quantifying the degree of non-Gaussian diffusion. To this end, the exponent of the DTI model is extended with a quadratic term in the *b*-value. The coefficient of the additional term relates to the apparent excess kurtosis (AKC), a dimensionless metric quantifying the non-gaussianity. By measuring the AKC in at least 15 different gradient directions, a fourth-order 3D, fully symmetric tensor—the diffusion kurtosis tensor (DKT)—can be calculated in addition to the DT. As the DKI model is parameterized by 22 elements: nondiffusion-weighted signal b_0, six independent DT elements, and 15 independent DKT elements, it requires at least diffusion weighting along 15 noncollinear gradient directions with one or two nonzero *b*-values in such a way that a total of 22 diffusion-weighted

[1]Visionlab, Department of Physics, University of Antwerp, Wilrijk, Antwerp, Belgium
[2]Biomedical Imaging Group Rotterdam, Erasmus MC, Rotterdam, The Netherlands
[3]Department of Radiology, University Hospital Antwerp, University of Antwerp, Edegem, Antwerp, Belgium
[4]Department of Radiology, University Hospitals of the Catholic University of Leuven, Leuven, Belgium
[5]Bio-Imaging Lab, Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium

Grant sponsor: SBO "Quantiviam" of IWT; Grant number: 060819
Grant sponsor: IAP-Grant of the Belgian Science Policy; Grant number: P6/38

*Correspondence to: Jelle Veraart, Visionlab, M.S., Department of Physics, University of Antwerp, Universiteitsplein 1, N 1.16, B-2610 Wilrijk, Belgium. E-mail: jelle.veraart@ua.ac.be

Received 5 March 2010; revised 15 July 2010; accepted 20 July 2010.
DOI 10.1002/mrm.22603
Published online 27 September 2010 in Wiley Online Library (wileyonlinelibrary.com).

images (DWIs) are acquired. Therefore, the DKI model is still feasible for clinical studies.

As DTI and DKI both provide a parameterization of the diffusion-weighted signal, we will evaluate the relevance for extending the DTI model to the DKI model using a likelihood ratio test in which the Rician distribution of the diffusion-weighted data is taken into account. Next, we will quantify the effect of the DKI model on the calculation of the conventional, and widely used, DT parameters in the structures of the healthy rat brain. In the analysis, the *b*-value dependency of the estimated diffusion parameters will be of particular interest.

MATERIALS AND METHODS

Data Acquisition

Diffusion-weighted data were acquired from nine adult, normal Sprague Dawley rat brains on a 9.4 T Bruker Biospec scanner (Ettlingen, Germany) using a spin echo EPI sequence with an encoding scheme of 30 diffusion-weighted gradient directions (19) using TR/TE = 6500/24 msec, $\delta = 5$ msec/$\Delta = 12$ msec, an acquisition matrix $= 96 \times 64$ with resolution $= 0.3 \times 0.3\,\text{mm}^2$, and 35 slices with slice thickness = 0.6 mm. For each of the 30 gradients, seven diffusion gradient strengths were applied, corresponding to the following *b*-values: 400, 800, 1200, 1600, 2000, 2400, and 2800 sec/mm^2. In addition, seven images without diffusion weighting (b_0) were acquired. This sequence was repeated four times. Hence, in total, 868 DWIs were acquired for each data set.

Estimation of Diffusion Parameters

Diffusion Models

The natural logarithm of the magnitude of the noise-free diffusion-weighted signal, $\mathcal{S}(b, \boldsymbol{g})$, as a function of the diffusion weighting strength b and gradient direction $\boldsymbol{g}$ can be approximated by a Taylor expansion around $b = 0$. The second-order approximation is given by:

$$\ln \mathcal{S}(b,\boldsymbol{g}) = \ln S(0) - b\sum_{i,j=1}^{3} g_i g_j D_{ij} + \frac{b^2}{6}\left(\sum_{i=1}^{3}\frac{D_{ii}}{3}\right)^2 \sum_{i,j,k,l=1}^{3} g_i g_j g_k g_l W_{ijkl} + O(b^3), \quad [1]$$

with g_i the ith component of $\boldsymbol{g}$ and $S(0)$ the nondiffusion-weighted signal intensity. D_{ij} is the ijth element of the fully symmetric DT $\boldsymbol{D}$, characterized by six independent elements: $\boldsymbol{\theta}_D = \{D_{ij}\}_{i\le j\le 3} = \{D_{11}, D_{12}, D_{22}, D_{13}, D_{23}, D_{33}\}$, and W_{ijkl} denotes an element of the diffusion kurtosis tensor $\boldsymbol{W}$, which is fully parameterized by 15 independent elements $\boldsymbol{\theta}_K = \{W_{ijkl}\}_{i\le j\le k\le l\le 3}$.

The frequently used DTI model is simply given by the first two terms of Eq. 1:

$$S(b,\boldsymbol{g};\boldsymbol{\theta}^{\text{DTI}}) = S(0)\exp\left(-b\sum_{i,j=1}^{3} g_i g_j D_{ij}^{\text{APP}}\right). \quad [2]$$

In Eq. 2, D_{ij}^{APP} is the ijth element of $\boldsymbol{D}^{\text{APP}}$, which is the apparent DT. The parameter vector $\boldsymbol{\theta}^{\text{DTI}}$ is the concatenation of $S(0)$ and the six-element vector $\boldsymbol{\theta}_D^{\text{APP}}$, which describes the six independent elements of $\boldsymbol{D}^{\text{APP}}$. The DKI model additionally includes the second-order term of Eq. 1:

$$S(b,\boldsymbol{g};\boldsymbol{\theta}^{\text{DKI}}) = S(0)\exp\left[-b\sum_{i,j=1}^{3} g_i g_j D_{ij}^{\text{APP}} + \frac{b^2}{6}\left(\sum_{i=1}^{3}\frac{D_{ii}^{\text{APP}}}{3}\right)^2 \sum_{i,j,k,l=1}^{3} g_i g_j g_k g_l W_{ijkl}^{\text{APP}},\right] \quad [3]$$

with W_{ijkl}^{APP} the $ijkl$th element of $\boldsymbol{W}^{\text{APP}}$, the apparent kurtosis tensor. The DKI model is parameterized by $\boldsymbol{\theta}^{\text{DKI}}$, which includes $S(0)$, $\boldsymbol{\theta}_D^{\text{APP}}$, and $\boldsymbol{\theta}_K^{\text{APP}}$. The latter parametrizes the 15 independent elements of $\boldsymbol{W}^{\text{APP}}$.

As fitting one of both models to the diffusion-weighted data results in an estimate for the diffusion coefficient, $\boldsymbol{\theta}_D^{\text{APP}}$, the diffusion parameters (FA, MD, $D_\perp$, and $D_\parallel$) can be calculated for each voxel by the eigenvalue decomposition of the related DTs (2). Note that $\boldsymbol{D}^{\text{APP}}$ of Eqs. 2 and 3 is not quantitatively equal.

Parameter Estimation

Two algorithms were used to estimate the parameter vector $\boldsymbol{\theta}_D$ with either of the diffusion models. The first algorithm estimated the parameters $\boldsymbol{\theta}^{\text{DTI}}$ and $\boldsymbol{\theta}^{\text{DKI}}$ with the weighted least squares (WLS) algorithm. The second algorithm estimated the parameters with maximum likelihood (ML) estimator. ML estimators are known to be asymptotically unbiased and to exploit the a priori knowledge of the data statistics in an optimal way. Nevertheless, although theoretical properties of the WLS are less favorable, WLS estimation is more commonly used because of ease of implementation and lower computational cost (11).

Weighted Least Squares Estimation. In this approach, the tensors were fitted linearly to the log-transformed diffusion-weighted data, such that the sum of the weighted squared differences was minimized (20). The use and implementation of the WLS estimator is fully elaborated in the appendix.

Maximum Likelihood Estimation. As the magnitude diffusion-weighted data are independently Rice distributed, the actual probability density function (PDF) of the magnitude of the observed diffusion-weighted signal is given by (13):

$$p(y_n|S(b_n,\boldsymbol{g}_n;\boldsymbol{\theta}),\sigma) = \frac{y_n}{\sigma^2} e^{\left(-\frac{y_n^2+(S(b_n,\boldsymbol{g}_n;\boldsymbol{\theta}))^2}{2\sigma^2}\right)} I_0\left(\frac{y_n S(b_n,\boldsymbol{g}_n;\boldsymbol{\theta})}{\sigma^2}\right), \quad [4]$$

with y_n the nth observation after applying diffusion weighting with strength b_n and gradient direction $\boldsymbol{g}_n$. The underlying magnitude signal, $S(b_n, \boldsymbol{g}_n; \boldsymbol{\theta})$, is given by Eqs. 2 and

3 for DTI and DKI, respectively. The noise level σ was estimated from the histogram mode of the image background (21). Furthermore, I_0 is the order zero modified Bessel function of the first kind. The parameter vector $\boldsymbol{\theta}$ was estimated from the N independent DWIs with a ML estimator in each voxel by substituting the observed values for the stochastic variables and maximizing over the parameters:

$$\hat{\boldsymbol{\theta}}_{\mathrm{ML}} = \arg\max_{\boldsymbol{\theta}} \sum_{n=1}^{N} \ln p(y_n | S(b_n, \boldsymbol{g}_n; \boldsymbol{\theta}), \sigma). \quad [5]$$

An ML estimator is asymptotically consistent, meaning that for $N \to \infty$, the bias tends to zero, while the variance of the parameter estimator converges to the Cramér-Rao lower bound (CRLB).

Cramér-Rao Lower Bound. As the CRLB is the theoretical lower bound on the variance of an unbiased estimator (22), the precision of the ML estimator based on both the DTI and the DKI model was evaluated with the CRLB. The CRLB states that the covariance matrix of an unbiased estimator of the model parameter $\boldsymbol{\theta}$ is equal to or greater than the inverse of the Fisher information matrix, $\boldsymbol{I}$:

$$\mathrm{cov}(\boldsymbol{\theta}) \geq \boldsymbol{I}^{-1}, \quad [6]$$

with

$$I_{ij} = -\mathbb{E}\left[\frac{\sum_{n=1}^{N} \partial^2 \ln p(y_n | S(b_n, \boldsymbol{g}_n; \hat{\boldsymbol{\theta}}))}{\partial\theta_i \partial\theta_j} \right], \quad [7]$$

in which $\mathbb{E}[]$ denotes the expectation operator (23).

Experiments

Evaluation of Goodness-of-Fit

Likelihood Ratio Test. As the DTI and DKI model are the first- and second-order Taylor expansion of the natural logarithm of the true magnitude of the diffusion-weighted signal, both diffusion models are nested. Therefore, one can use a likelihood ratio test to evaluate whether the DKI model fits the experimental data significantly better and should thus be preferred over the DTI model. The null hypothesis, H_0, stating that increasing the number of model parameters does not significantly improve the description of the experimental data is rejected if the probability of the likelihood ratio Λ exceeds a given threshold (24). The rejection of H_0 is only expected if the apparent non-Gaussian diffusion component is inherent to the diffusion process and cannot entirely be regarded as a noise artifact (10). Λ was calculated in each voxel as:

$$\Lambda = -2 \ln \frac{\prod_{n=1}^{N} p(y_n | S(b_n, \boldsymbol{g}_n; \hat{\boldsymbol{\theta}}^{\mathrm{DTI}}))}{\prod_{n=1}^{N} p(y_n | S(b_n, \boldsymbol{g}_n; \hat{\boldsymbol{\theta}}^{\mathrm{DKI}}))}, \quad [8]$$

with $\hat{\boldsymbol{\theta}}^{\mathrm{DTI}}$ and $\hat{\boldsymbol{\theta}}^{\mathrm{DKI}}$ the parameter vectors estimated with the DTI and DKI model, respectively. These parameters were estimated by fitting either of the diffusion models to all 868 DWIs with the asymptotically unbiased ML estimator, as it was observed that the nonlinear fitting algorithm provides the most accurate description of the true diffusion profile (10,11). The threshold on Λ to reject H_0 is derived from the χ^2-distribution with j degrees of freedom. The parameter j denotes the number of additional parameter of the DKI model with respect to the DTI model; in this analysis, j equals 15. Furthermore, the significance level was set to 1%, resulting in a threshold on Λ of 30.8.

Validation of the Distribution of Λ. The critical value for rejecting a null hypothesis was computed using the theoretical asymptotic distribution of the likelihood ratio test statistics, the χ^2_j-distribution. However, because of the limited number of data points for every voxel in this study, this may lead to an inaccurate false-positive rate (P_{f}), the probability of falsely rejecting H_0. Therefore, the validity of the χ^2_j-distribution was evaluated by a simulation experiment.

First, a diffusion-weighted data set, associated to purely Gaussian diffusion, was synthesized by recalculating the diffusion-weighted signal in each voxel using Eq. 2. $S(0)$ and $\boldsymbol{D}^{\mathrm{APP}}$ corresponded to the parameter vector $\hat{\boldsymbol{\theta}}^{\mathrm{DTI}}$. $\hat{\boldsymbol{\theta}}^{\mathrm{DTI}}$ was estimated with the ML algorithm by fitting the DTI model to each voxel of the experimental data set. The gradient directions $\boldsymbol{g}$ and the b-values used to recalculate DWIs with Eq. 2 correspond to the ones used to acquire the experimental data (see section "Data Acquisition"). Next, the simulated diffusion-weighted data were corrupted with Rician noise with a level equal to the estimated noise level in the experimental data; the SNR of the simulated nondiffusion-weighted images was 20.3. Finally, $\hat{\boldsymbol{\theta}}^{\mathrm{DTI}}$ and $\hat{\boldsymbol{\theta}}^{\mathrm{DKI}}$ were estimated by fitting, respectively, the DTI and DKI model to the simulated diffusion-weighted data.

Because of the simulation setup, all kurtosis tensor elements were expected to be zero in each voxel. As a result, the validity of the theoretical χ^2_{15}-distribution could be evaluated by comparing χ^2_{15} with the observed distribution of the Λs over the entire volume.

b-Value Dependency of Diffusion Parameters Using the DTI and DKI Model

The dependency of the DT parameter on the b-value was evaluated for both the DTI model and the DKI model. Therefore, both diffusion models were fitted to several subsets of the diffusion-weighted data that were selected in such a way that each subset contained the images corresponding to a unique (set of) b-value(s). An overview of the analyzed models and their associated DWI subsets is given by:

i. The DTs and its related parameters MD, FA, $D_\perp$, and $D_\parallel$ were estimated by fitting the DTI model to seven different subsets, each including 127 DWIs: seven b_0 images and all images of the four repetitions associated with a unique nonzero b-value. This DTI approach with one nonzero b-value is called single b-valued DTI in the remainder of this work.
ii. As the DKI model requires DWIs that are at least acquired with two different nonzero b-values, the subsets of (i) were inadequate. Therefore, 21 different subsets were defined by selecting, in addition to seven

Table 1
An Overview of the Different Diffusion Models and Their Corresponding DWI Subsets That Were Used to Estimate the Diffusion Parameters in the Rat Brain Structures with a Comparable Precision, Quantified by the CRLB

	Model	*b*-Value(s)	# Grad. dirs	NEX	*N*
(a)	DTI	1200	30	2	62
(b)	DKI	800, 2400	30	4	248
(c)	DTI	400, 2400	30	2	124

b_0 images, the images of only the first two repetitions associated to one out of all two-combinations of the nonzero *b*-values. So, the total number of DWIs was 127 for each set as well.

iii. Finally, the DTI model was fitted to the same 21 subsets as defined in (ii). This analysis will be referred to as double *b*-valued DTI.

For every model and subset, the distributions of MD, FA, $D_\perp$, and $D_\parallel$ over all WM voxels were calculated. Note that the set of WM voxels included every voxel exceeding an FA threshold of 0.25. The FA map used to mask the WM was estimated by fitting the DTI model to all DWIs that were acquired with a *b*-value of 1200 sec/mm^2, and this single mask was used for all analyses.

Diffusion Parameter Estimation with DKI vs. DTI Model in the Rat Brain Structures

Regions of interest (ROI) were manually segmented by an expert and confirmed by a second one. The delineation of some gray matter [GM; caudate putamen (CPu) and the hippocampus (HP)] and WM structures [corpus callosum (CC) and external capsule (EC)] was based on a rat brain atlas (25). Nondiffusion-weighted images and DT parametric maps were used to guide the segmentation of the ROIs, using AMIRA software (Visage Imaging, Berlin, Germany). First, the DT parameter values (FA, MD, $D_\perp$, and $D_\parallel$) were estimated by fitting (a) the single *b*-valued DTI, (b) the DKI, and (c) the double *b*-valued DTI model to new, well chosen, subsets of the diffusion-weighted data set with the ML estimator. Next, the parameter distributions within the delineated anatomical structures were calculated for (a), (b), and (c). Finally, for each structure, all distributions were statistically compared with a paired Wilcoxon signed rank test.

During this experiment, DWI subsets for (a), (b), and (c) were selected in such a way that the precision of the parameter estimation was similar for all three approaches. As the asymptotically unbiased ML estimator was used in this experiment, the variance of the estimator tended to the CRLB. The selected subsets (see Table 1) consisted out of one or two shells of which the strength was indicated by the *b*-value(s). For each shell, DWIs were acquired along the same 30 diffusion gradient directions with multiple repetitions (NEX). The subset that was used for the single *b*-valued DTI (a) analysis was selected so that its set of *b*-values and gradient directions was in agreement with conventional DTI studies. The other subsets were selected in such a way that the associated CRLB distributions of MD over all voxels were not significantly different to the CRLB distribution of MD that corresponds to (a). Significant ($P < 0.1$) differences between distributions were shown with the Wilcoxon signed rank test.

Note that number of DWIs (N) for the DKI model (b) was substantially increased, compared with the DTI model, to obtain a similar precision of the DT parameters.

RESULTS

Noise Levels

The estimated SNR, defined as the ratio of the median voxel value of the b_0 image to the noise level, was 20.3. As the images were acquired with a single coil receiver, the noise was homogeneously distributed over the entire volume (26).

Likelihood Ratio Test

Validation of Distribution of Λ

Figure 1 shows the logarithmic plot of the observed false-positive rate (P_f) as a function of the P_f, computed from the χ^2_{15}-distribution, which is the asymptotic distribution for the likelihood ratio test statistics. The diagonal corresponds to the asymptotic distribution. The shaded areas indicate the 95% confidence region of the observed P_f as computed by the binomial counting statistics.

As the diagonal is within the 95% confidence bound, one can conclude that the distribution of Λ over all simulated voxels is well approximated by χ^2_{15}.

Evaluation of Goodness-of-Fit

In Fig. 2a, the likelihood ratio value, Λ, is visualized in each voxel of a single axial slice of the rat brain. In the majority

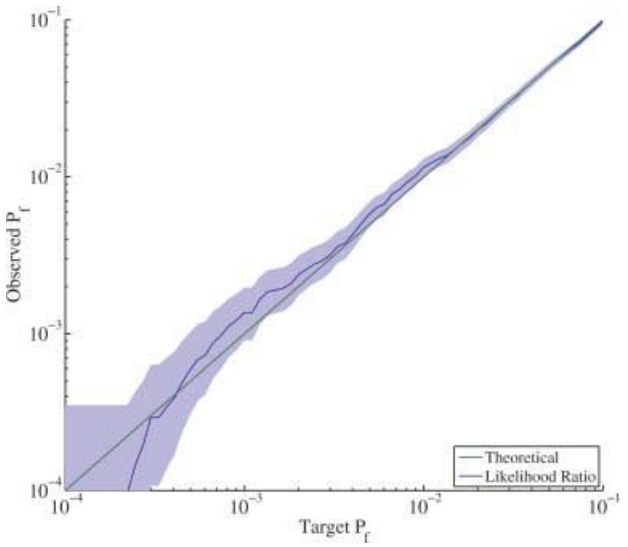

FIG. 1. Logarithmic plot of P_f of (blue) the observed Λ distribution and (green) the theoretical χ^2_{15}-distribution as a function of the P_f computed from the χ^2_{15}-distribution, which is the asymptotic distribution for the likelihood ratio test statistics. The shaded areas indicate the 95% confidence regions of the P_f of the Λ distribution as computed by the binomial counting statistics. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

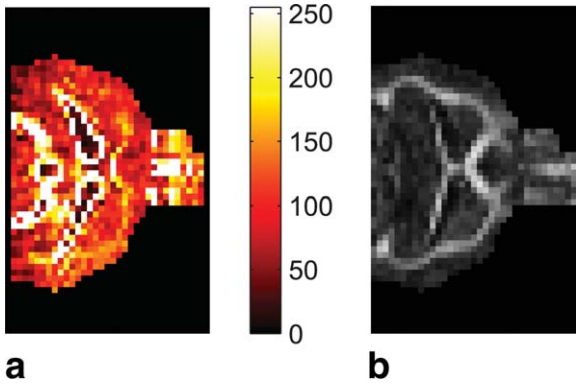

FIG. 2. **a**: The likelihood ratio Λ is visualized in each voxel of one axial slice of the rat brain. H_0 was rejected in each voxel for which Λ exceeded the critical value of 30.8. **b**: The corresponding FA map is shown as anatomical reference.

(>95%) of all voxels, Λ widely exceeds the critical value of 30.8, corresponding to a significance level of 1% in a χ^2_{15}-distribution. Only in the cerebrospinal fluid (CSF) of the rat brain, in which Gaussian diffusion is expected, Λ is lower than the critical value. Hence, the null hypothesis H_0, which states that DKI does not significantly improve the description of diffusion-weighted data, compared with DTI, is rejected brain for all areas with exception of the CSF. The FA map of the same axial slice is shown as anatomical reference (see Fig. 2b).

b-Value Dependency of Diffusion Parameters Using the DTI and DKI Model

Figure 3 shows the effect of the selected *b*-value(s) on the estimation of diffusion parameters. This effect differs for the diffusion models: (i) single *b*-valued DTI, (ii) DKI, and (iii) double *b*-valued DTI. The height of the bars relates to the median of distribution of the diffusion parameters corresponding to either of the diffusion model and a *b*-value-dependent subset, calculated over all WM voxels. The lower and upper edges of the error bars indicate the 25th and 75th percentile of the distributions. With single *b*-valued DTI (i), the diffusion parameters MD, $D_\perp$, $D_\parallel$, and, although less pronounced, FA significantly decrease with increasing *b*-value. Even in the range of low *b*-values, $400-1200\,\mathrm{sec/mm^2}$, the shifts in median values are clearly visible. The *b*-value dependency is reduced when fitting the DTI model to subsets that contained data acquired with more than one *b*-value (iii). However, even then, one can easily observe that the parameters MD, $D_\perp$, and $D_\parallel$ vary more with the *b*-value of the DWI from which they were estimated than the same parameters estimated with the DKI model (ii).

Diffusion Parameter Estimation with DKI vs. DTI Model in the Rat Brain Structures

Figure 4 shows the ROI measurements of various diffusion parameters that were computed by different diffusion models in multiple anatomical structures of the rat brain. The median values of the diffusion parameters, calculated over all voxels in each anatomical structure, are indicated by the height of the bars, while the edges of the error bars are the 25th and 75th percentiles. These results suggest that the quantification of DT parameters strongly depends on the diffusion model selection. A statistically significant ($P < 0.001$) increase of MD, $D_\perp$, and $D_\parallel$ was noticed when estimating the parameters with the DKI model compared with the DTI model. The estimated parameter values also depend on the number of different *b*-values included in the

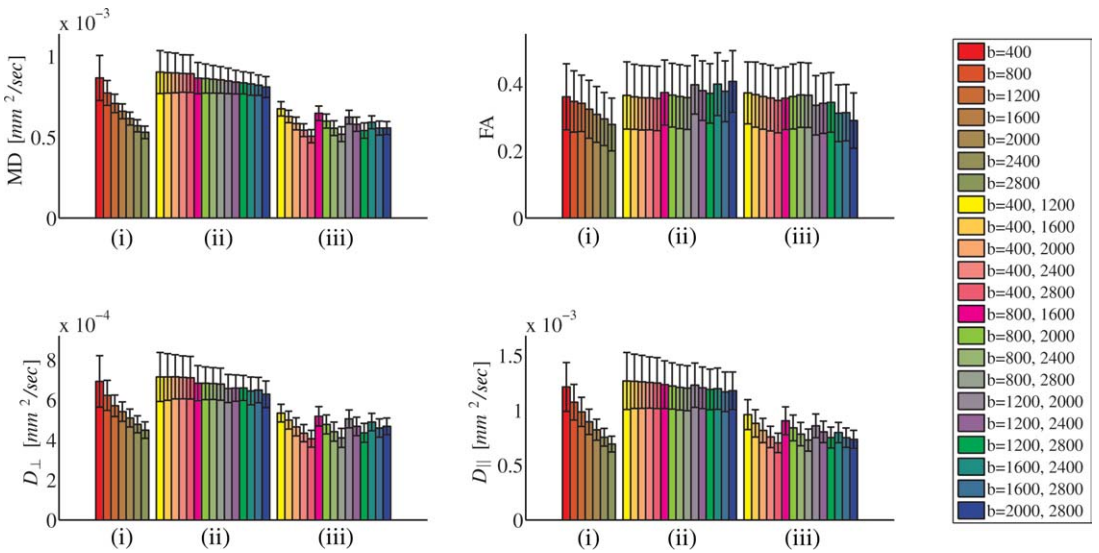

FIG. 3. The height of the bars represents the median of the distribution of the associated DTI parameter over all WM voxels, whereas the edges of the error bars correspond to the 25th and 75th percentile. One can observe that the parameter estimation with DKI (ii) is less *b*-value dependent, compared with the single and double *b*-valued DTI model, (i) and (iii), respectively. The *b*-value dependency was evaluated by fitting the DTI and DKI models to well-chosen subsets which were characterized by a unique (set of) *b*-value(s). The selected subsets were summarized in the legend by means of their associated nonzero *b*-value(s) [$\mathrm{sec^2/mm}$].

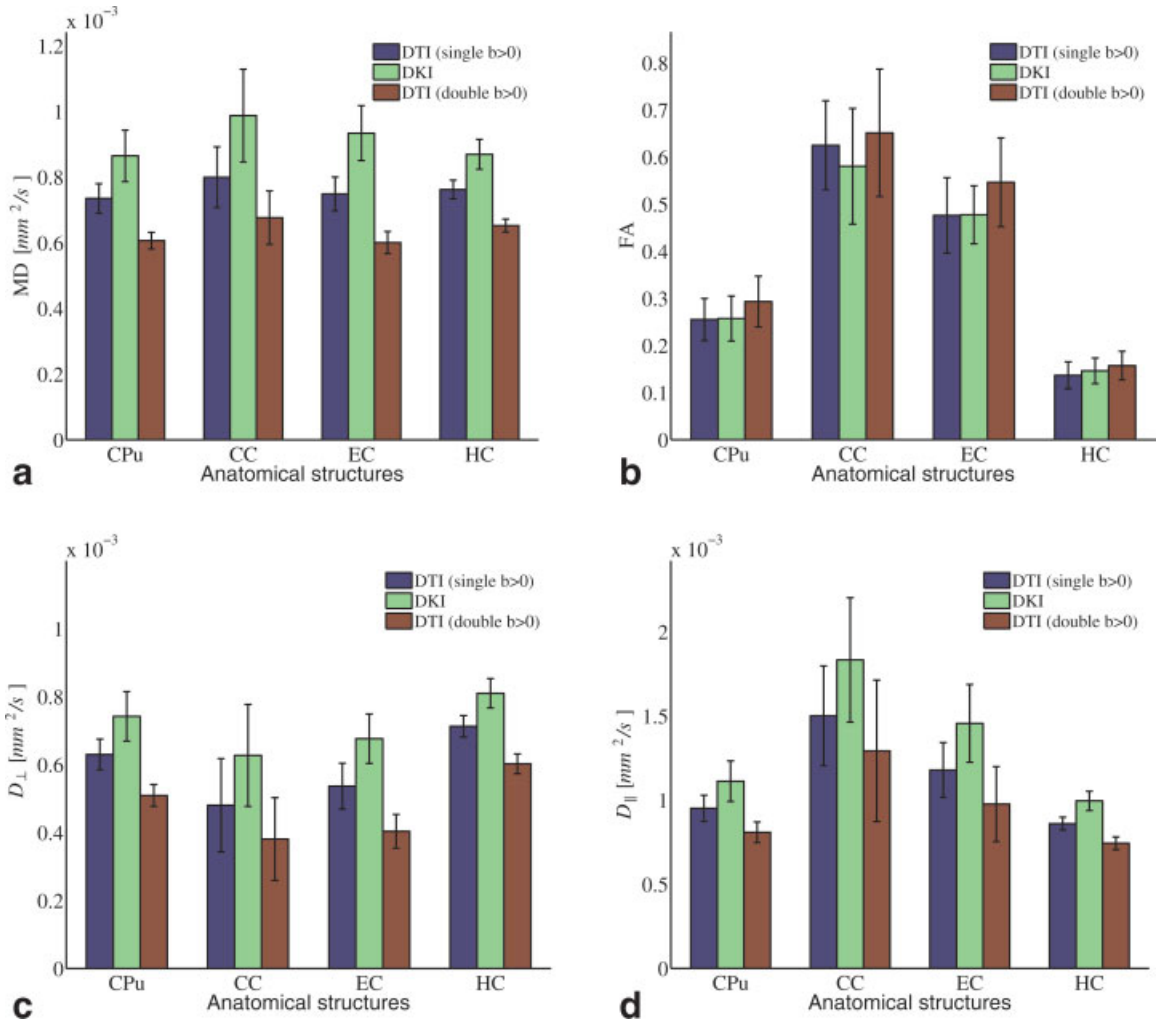

FIG. 4. The average diffusion parameter values in typical rat brain structures [caudate putamen (CPu), corpus callosum (CC), external capsule (EC) and hippocampus (HP)] calculated using different strategies: single *b*-valued DTI, DKI and double *b*-valued DTI. One can notice a significant increase of MD, $D_{\perp}$, and $D_{\parallel}$ when performing a DKI analysis compared with the DTI approaches. The estimated values of those diffusion parameters also depend on the number of different *b*-values included in the DTI model. Adding an additional higher *b*-value to the acquisition scheme results in significant decreased diffusion parameter values when performing a DTI analysis. No statistical significant differences were observed when comparing the FA values estimated with single *b*-valued DTI and the DKI model. In addition, double *b*-valued DTI resulted in significant increased FA values compared with single *b*-valued DTI and DKI. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

DTI analysis. Adding a higher *b*-value to the acquisition scheme results in significantly decreased parameter values when performing a DTI analysis ($P < 0.001$). No statistical significant differences ($P > 0.05$) were observed when comparing the FA values estimated with single *b*-valued DTI and the DKI model. In addition, double *b*-valued DTI resulted in significantly increased FA values ($P < 0.01$) compared with single *b*-valued DTI and DKI.

As the biased WLS estimator is likely to be used more often in practice because of its simplicity, it is briefly suggested that using a WLS estimator instead of a ML estimator would slightly increase the discrepancies between DT measures estimated using (a) single *b*-valued DTI, (b) DKI, or (c) double *b*-valued DTI. In Table 2, the median values of the MD, calculated over all voxels in each anatomical structure, estimated with both the WLS and ML estimator are summarized for all diffusion models. In general, one can note a decrease of MD with the WLS estimator instead of the ML estimator. However, the MD decrease is slightly larger with the DTI models compared with the DKI model.

Table 2
The Median Values of the MD, Over All Voxels in the Studied Anatomical Structures, Estimated with the WLS and ML for (a) Single *b*-Valued DTI, (b) DKI, or (c) Double *b*-Valued DTI

	MD ($\times 10^{-4}$ mm^2/sec)							
	ML				WLS			
	CPu	CC	EC	HC	CPu	CC	EC	HC
(a)	7.35	7.99	7.49	7.62	7.07	7.44	7.04	7.41
(b)	8.65	9.87	9.34	8.70	8.41	9.32	8.90	8.55
(c)	6.06	6.77	6.01	6.52	5.57	5.61	5.23	6.07

In general, a decrease of MD is noticed when using the WLS estimator instead of the ML estimator. However, the MD decrease is slightly more expressed with the DTI models compared with the DKI model.

DISCUSSION

Diffusion of water molecules is a physical property of the tissue being measured, and, thus, its estimated coefficient should not depend on scanner settings or properties, such as the *b*-value (27). However, because of the nonlinear relation between the natural logarithm of the diffusion-weighted signal and the *b*-value, the DTI model results in an inaccurate and *b*-value-dependent parameterization of the diffusion profile.

In previous studies, several new diffusion models, such as the biexponential and the cumulant expansion model, were presented to improve the description of nonmonoexponential diffusion-weighted signal attenuation as observed in biological tissue (17,28,29). The cumulant expansion model can be interpreted as a truncated Taylor expansion of the log-transformed diffusion-weighted signal in terms of the *b*-value. The model provides a good approximation to the true signal within a certain range of *b*-values; for the second-order cumulant expansion, the *b*-values should be limited to $\sim 2500\,\mathrm{sec/mm^2}$ (17,30). The coefficient of the second-order term is associated to the apparent excess kurtosis probing the degree of non-Gaussian diffusion due to microscopic restrictions in biological tissue (17). Therefore, the cumulant expansion supports another interpretation of nonmonoexponential diffusion-weighted signal decay compared with the biexponential model that suggests the presence of two compartments in the brain tissue. However, Kisiliv and Il'Yasov stated that the good accuracy of data fitting with the biexponential model does not prove the presence of distinct compartments in the human brain's GM (30). In addition, they demonstrated that with isotropic diffusion, the second-order cumulant expansion describes the experimental data equally well without the artificial appearance of two compartments.

Recently, Jensen et al. (17) proposed the DKI model as a 3D generalization of the second-order cumulant expansion that can be used to study the (non-)Gaussian diffusion in GM and WM on a direction-dependent basis. It is a straightforward extension of the DTI model that allows estimating a higher order DKT in addition to the DT with only minor changes in data acquisition and processing. The physical interpretation and the clinical relevance of the additional kurtosis parameters are currently studied (31,32). In this study, however, we focused on the quantification of the widely used DT parameters of which the estimated values clearly depend on the diffusion model as demonstrated in our experiments. As the DTI and DKI model could both be used to estimate the DT elements and its associated parameters, we studied the relevance of extending the DTI model to parameterize the diffusion-weighted signal decay with a likelihood ratio test in which we included a Rician noise model. Our results suggest that DKI provides a significantly improved fit to the experimental diffusion-weighted signal, resulting in a more plausible parameterization in all voxels of the GM and WM of the rat brain. Only in the brain CSF, we concluded that the DTI and DKI models describe the data equally well, which is in agreement with the assumption of Gaussian diffusion in the CSF.

A decreased precision of the DT parameter estimation is inherent to the DKI model because of the additional model parameters compared with the DTI model. Therefore, previous DKI studies increased the voxel sizes and/or the number of acquired DWIs. In this work, an equal or, at least, a similar precision of DT parameter estimation as obtained in DTI was achieved by substantially increasing the number of data points (see Table 1) as a result of which the clinical feasibility might be questioned. However, recently, Poot et al. proposed a procedure to optimize the acquisition scheme to achieve a minimal CRLB of the diffusion parameters of interest (33). The procedure allows reducing the time needed to acquire a DKI data set. Furthermore, an improved and robust computational framework that allows fewer DWIs to estimate both diffusion tensors with DKI was recently proposed by Tabesh et al. (34).

As DKI offers a more probable parameterization of the diffusion parameters, including the six independent DT elements, it was worth studying the properties of the DT estimation with DKI. Our main findings are as follows:

a. The estimation of diffusion parameters with the DKI model is less dependent on the *b*-value, compared with the more commonly used DTI model. The *b*-value dependency of the DTI model can also be reduced by fitting the model to DWIs acquired with multiple nonzero *b*-values, but not to the degree obtained with the DKI model.
b. The diffusion indices that are linearly related to the ADC are significantly increased over the entire rat brain when fitting the DKI model to the diffusion-weighted data. The observed FA values are less sensitive to the model selection.

The second finding was already suggested by Hui et al. (35) and Cheung et al. (8). However, both studies lack unbiased quantitative results because of intersubject and ROI differences, a Gaussian noise model, and the use of high *b*-values in the DTI analysis. Both the Gaussian noise model and the high *b*-values resulted in underestimated diffusivity measures within their DTI analysis.

In this study, we stressed on the use of a proper noise model in quantitative studies to avoid biased results. However, one should note that in case of parallel imaging, which is commonly used in clinical studies nowadays, the noise amplitude is spatially dependent and can be described by a Rician or a χ-distribution according to the used reconstruction technique (26). As methods dedicated to those specific noise characteristics are currently studied, the WLS estimator is still widely used to estimate diffusion parameters in clinical studies. We would like to point out that with the WLS estimator, the *b*-value dependency of the DTI model and the decreased diffusion parameters is even more pronounced as a result of which DKI becomes more preferable in such studies.

ACKNOWLEDGMENT

This work is supported by the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen).

APPENDIX

The tensors are fitted to each voxel of the diffusion-weighted data set such that the sum of weighted squared

differences is minimized (20). The WLS estimator of seven unknown DTI coefficients, $\boldsymbol{\theta} = [\ln S(0), \boldsymbol{\theta}_D]^T$ is

$$\hat{\boldsymbol{\theta}}_{\mathrm{WLS}} = \left(\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \boldsymbol{W} \ln \boldsymbol{y}, \qquad [A1]$$

with the nth element of $\boldsymbol{y}$, $y^{(n)}$, the observation after applying diffusion weighting with strength $b^{(n)}$ and gradient direction $\boldsymbol{g}^{(n)}$. The nth row of $\boldsymbol{X}$ $[N \times 7]$ is defined as follows:

$$X_{n,*} = [-b_n, \boldsymbol{\alpha} \circ \{-b_n g_{n_i} g_{n_j}\}_{i \le j \le 3}] \qquad [A2]$$

with $\boldsymbol{\alpha} = [1, 2, 1, 2, 2, 1]$ and $\circ$ denotes the element-wise product.

Furthermore, $\boldsymbol{W}$ is the diagonal matrix for which the elements are defined as follows:

$$W_{nn} = y_n^2, \qquad [A3]$$

which is motivated by assumption of a Gaussian noise model as a result of which y_n^2 is proportional to the inverse variance of

$$\epsilon_n = y_n - X_{n,*} \hat{\boldsymbol{\theta}}_{\mathrm{WLS}}. \qquad [A4]$$

An extended algorithm is applied to fit the DKI model, given by Eq. 3, to the DWIs. To estimate $\hat{\boldsymbol{\theta}}^{\mathrm{DKI}}$ with the WLS estimator, a modified matrix $\boldsymbol{X}$ was constructed (18). The nth row is given by:

$$X_{n,*} = \left[-b_n, \boldsymbol{\alpha} \circ \{-b_n g_{n_i} g_{n_j}\}_{i \le j \le 3}, \; \boldsymbol{\beta} \circ \left\{ \frac{b^{(n)^2}}{6} g_{n_i} g_{n_j} g_{n_k} g_{n_l} \right\}_{i \le j \le k \le l \le 3} \right] \qquad [A5]$$

with $\boldsymbol{\beta} = [1, 4, 6, 4, 1, 4, 12, 12, 4, 6, 12, 6, 4, 4, 1]$.

REFERENCES

1. Basser PJ, Mattiello J, Le Bihan D. MR diffusion tensor spectroscopy and imaging. Biophys J 1994;66:259–267.
2. Le Bihan D, Mangin JF, Poupon C, Clark C, Pappata S, Molko N, Chabriat H. Diffusion tensor imaging: concepts and applications. J Magn Reson Imaging 2001;13:534–546.
3. Van Camp N, Blockx I, Verhoye M, Casteels C, Coun F, Leemans A, Sijbers J, Baekelandt V, van Laere K, Van Der Linden AM. Diffusion tensor imaging in a rat model of parkinson's disease after lesioning of the nigrostriatal tract. NMR Biomed 2009;22:697–706.
4. Chahboune H, Mishra A, DeSalvo M, Staib L, Purcaro M, Scheinost D, Papademetris X, Fyson S, Lorincz M, Crunelli V, Hyder F, Blumenfeld H. DTI abnormalities in anterior corpus callosum of rats with spike-wave epilepsy. Neuroimage 2009;47:459–466.
5. Song SK, Sun SW, Ju WK, Lin SJ, Cross AH, Neufeld AH. Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia. Neuroimage 2003;20:1714–1722.
6. Song SK, Yoshino J, Le T, Lin SJ, Sun SW, Cross A, Armstrong R. Demyelination increases radial diffusivity in corpus callosum of mouse brain. Neuroimage 2005;26:132–140.
7. Narayana PA, Ahobila-Vajjula P, Ramu J, Herrera J, Steinberg JL, Moeller FG. Diffusion tensor imaging of cocaine-treated rodents. Psychiatr Res Neuroimag 2009;171:242–251.
8. Cheung MM, Hui ES, Chan KC, Helpern JA, Qi L, Wu EX. Does diffusion kurtosis imaging lead to better neural tissue characterization? A rodent brain maturation study. Neuroimage 2009;45:386–392.
9. Donald CM, Dikranian K, Song S, Bayly P, Holtzman D, Brody D. Detection of traumatic axonal injury with diffusion tensor imaging in a mouse model of traumatic brain injury. Exp Neurol 2007;205:116–131.
10. Jones DK, Basser PJ. Squashing peanuts and smashing pumpkins: how noise distorts diffusion-weighted MR data. Magn Reson Med 2004;52:979–993.
11. Andersson J. Maximum a posteriori estimation of diffusion tensor parameters using a Rician noise model: why, how and but? Neuroimage 2008;42:1340–1356.
12. Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. Magn Reson Med 1995;34:910–914.
13. Sijbers J, den Dekker AJ, Scheunders P, Van Dyck D. Maximum likelihood estimation of Rician distribution parameters. IEEE Trans Med Imaging 1998;17:357–361.
14. Melhem ER, Itho R, Jones L, Barker PB. Diffusion tensor MR imaging of the brain: effect of diffusion weighting on trace and anisotropy measurement. AJNR Am J Neuroradiol 2000;21:1813–1820.
15. Hui ES, Cheung MM, Chan KC, Wu EX. B-value dependence of DTI quantitation and sensitivity in detecting neural tissue changes. Neuroimage 2010;49:2366–2374.
16. Tuch D, Reese T, Wiegell M, Wedeen VJ. Diffusion MRI of complex neural architecture. Neuron 2003;40:885–895.
17. Jensen JH, Helpern JA, Ramani A, Lu H, Kaczynski K. Diffusional kurtosis imaging: the quantification of non-Gaussian water diffusion by means of magnetic resonance imaging. Magn Reson Med 2005;53:1432–1440.
18. Lu H, Jensen JH, Ramani A, Helpern JA. Three dimensional characterization of non-Gaussian water diffusion in humans using diffusion kurtosis imaging. NMR Biomed 2006;19:236–247.
19. Jones D, Horsfield M, Simmons A. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. Magn Reson Med 1999;42:515–525.
20. Chung S, Lu Y, Henry RG. Comparison of bootstrap approaches for estimation of uncertainties of DTI parameters. Neuroimage 2006;33:531–541.
21. Sijbers J, Poot DHJ, den Dekker AJ, Pintjens W. Automatic estimation of the noise variance from the histogram of a magnetic resonance image. Phys Med Biol 2007;52:1335–1348.
22. Sijbers J, den Dekker AJ, Raman E, Van Dyck D. Parameter estimation from magnitude MR images. Int J Imag Syst Tech 1999;10:109–114.
23. Kay SM. Fundamentals of statistical signal processing, Vol. 1: Estimation theory. Upper Saddle, NJ: Prentice Hall PTR; 1993.
24. Kay SM. Fundamentals of statistical signal processing, Vol. 2: Detection theory. New Jersey: Prentice Hall PTR; 1998.
25. Paxinos G. Watson C. The rat brain in stereotaxic coordinates, 6th ed. Academic Press: San Diego; 2006.
26. Dietrich O, Raya JG, Reeder SB, Ingrisch M, Reiser MF, Schoenberg SO. Influence of multichannel combination, parallel imaging and other reconstruction techniques on MRI noise characteristics. Magn Reson Imaging 2008;26:754–762.
27. Horsfield MA. Using diffusion-weighted MRI in multicenter clinical trials for multiple sclerosis. J Neurol Sci 2001;186:S51–S54.
28. Niendort T, Dijkhuizen RM, Norris DG, van Lookeren Campagne M, Nicolay K. Biexponential diffusion attenuation in various states of brain tissue: implications for diffusion-weighted imaging. Magn Reson Med 1996;36:847–857.
29. Frøhlich AF, Østergaard L, Kiselev VG. Effect of impermeable boundaries on diffusion-attenuated MR signal. J Magn Reson 2006;179:223–233.
30. Kiselev VG, Il'yasov KA. Is the biexponential diffusion biexponential. Magn Reson Imaging 2007;57:464–469.
31. Fieremans E, Jensen J, Tabesh A, Hu C, Helpern J. White matter model for diffusional kurtosis imaging. In: Proceedings of the 18th Annual Meeting of the International Society for Magnetic Resonance in Medicine, Stockholm, Sweden, 2010. p 1596.
32. Fieremans E, Novikov D, Jensen J, Helpern J. Monte carlo study of a two-compartment exchange model of diffusion. In: Proceedings of the 18th Annual Meeting of the International Society for Magnetic Resonance in Medicine, 2010. p 302.
33. Poot D, den Dekker A, Achten R, Verhoye M, Sijbers J. Optimal experimental design for diffusion kurtosis imaging. IEEE Trans Med Imaging 2010;29:819–829.
34. Tabesh A, Jensen J, Ardekani B, Helpern J. Robust estimation of kurtosis and diffusion tensors in diffusional kurtosis imaging. In: Proceedings of the 18th Annual Meeting of the International Society for Magnetic Resonance in Medicine, 2010. p 3974.
35. Hui ES, Cheung MM, Qi L, Wu EX. Towards better MR characterization of neural tissues using directional diffusion kurtosis analysis. Neuroimage 2008;42:122–134.