Review Article
Diffusion-tensor MRI: theory, experimental design and data analysis – a technical review

Peter J. Basser\(^1\) and Derek K. Jones\(^2\)

\(^1\)Section on Tissue Biophysics and Biomimetics, NICHD, National Institutes of Health, Bethesda, MD, USA
\(^2\)Section of Old Psychiatry, Institute of Psychiatry, London, UK

Accepted 16 November 2001; Revised 8 January 2002; Accepted 9 January 2002

ABSTRACT: This article treats the theoretical underpinnings of diffusion-tensor magnetic resonance imaging (DT-MRI), as well as experimental design and data analysis issues. We review the mathematical model underlying DT-MRI, discuss the quantitative parameters that are derived from the measured effective diffusion tensor, and describe artifacts that arise in typical DT-MRI acquisitions. We also discuss difficulties in identifying appropriate models to describe water diffusion in heterogeneous tissues, as well as in interpreting experimental data obtained in such issues. Finally, we describe new statistical methods that have been developed to analyse DT-MRI data, and their potential uses in clinical and multi-site studies. Copyright © 2002 John Wiley & Sons, Ltd.

KEYWORDS: diffusion; tensor; MRI; methods

INTRODUCTION

It is now well established that the MR measurement of an effective diffusion tensor of water in tissues can provide unique biologically and clinically relevant information that is not available from other imaging modalities. This information includes parameters that help characterize tissue composition, the physical properties of tissue constituents, tissue microstructure and its architectural organization. Moreover, this assessment is obtained non-invasively, without requiring exogenous contrast agents.

This article describes methodological issues related to the estimation of the effective diffusion tensor of water in tissue. It also tries to review and showcase new methods that have been proposed to advance the state of the art in this burgeoning field. We focus primarily upon diffusion tensor MRI (DT-MRI) data acquisition, experimental design, artifacts and post-processing issues. The definition and physical interpretation of useful MR parameters derived from the effective diffusion tensor, such as the Trace as well as measures of diffusion anisotropy, have been reviewed elsewhere.\(^1\) Moreover, several review articles and book chapters have covered many different aspects of diffusion tensor MRI.\(^2\)–\(^5\) This article also tries to identify unresolved issues in DT-MRI data acquisition and analysis, with the hope of interesting readers to pursue and address some of these problems.

CHARACTERIZING DIFFUSION IN BIOLOGICAL SYSTEMS

In tissues, such as brain gray matter, where the measured apparent diffusivity is largely independent of the orientation of the tissue (i.e. isotropic), it is usually sufficient to characterize the diffusion characteristics with a single (scalar) apparent diffusion coefficient (ADC). However, in anisotropic media, such as skeletal and cardiac muscle\(^6\)–\(^8\) and in white matter,\(^9\)–\(^11\) where the measured diffusivity is known to depend upon the orientation of the tissue, no single ADC can characterize the orientation-dependent water mobility in these tissues. The next most complex model of diffusion that can describe anisotropic diffusion is to replace the scalar diffusion coefficient with a symmetric effective or apparent diffusion tensor of water, \(\mathbf{D}\).\(^12\)

PHYSICAL UNDERPINNINGS OF DIFFUSION TENSOR NMR AND MRI

Torrey\(^13\) first incorporated anisotropic translational diffusion in the Bloch (magnetization transport) equations.\(^14\)
which could lead to additional attenuation of the NMR signal. Analytical solutions to this equation followed for freely diffusing species during a spin echo experiment\textsuperscript{15} and, later, for diffusion in restricted geometries.\textsuperscript{16–18} About a decade after its introduction, Stejskal and Tanner solved the Bloch–Torrey equation\textsuperscript{19} for the case of free, anisotropic diffusion in the principal frame of reference. However, the Stejskal–Tanner formula is not generally useable to measure an effective diffusion tensor using NMR or MRI methods for several reasons. First, this formula relates a time-dependent diffusion tensor, to the NMR signal, so one must establish a relationship between the time-dependent diffusion tensor and an effective diffusion tensor. Second, in the pre-MRI era in which the Stejskal–Tanner formula was derived, it was always tacitly assumed that a homogeneous anisotropic sample could be physically reoriented within the magnet so that its principal axes could be aligned with the laboratory coordinate system. After the development of MRI, however, this assumption was no longer tenable. Materials under study were often heterogeneous media whose ‘fibre’ or principal axes were generally not known \textit{a priori} and could vary from place to place within the sample. Thus, a general scheme had to be developed to measure the entire diffusion tensor (both its diagonal and off-diagonal elements) in the laboratory frame of reference.\textsuperscript{20}

The NMR measurement of the effective diffusion tensor\textsuperscript{20} and the analysis, and display of the information it contains in each voxel, is called diffusion tensor MRI (DT-MRI).\textsuperscript{21} The effective diffusion tensor, $\mathbf{D}$, (or functions of it) is estimated from a series of diffusion-weighted images (DWI) using a relationship between the measured echo attenuation in each voxel and the applied magnetic field gradient sequence.\textsuperscript{20} Just as in diffusion imaging (DI) where a scalar $b$-factor is calculated for each DWI, in DT-MRI a symmetric $b$-matrix is calculated for each DWI.\textsuperscript{22} Whereas the $b$-value summarizes the attenuating effect on the MR signal of all diffusion and imaging gradients in one direction,\textsuperscript{23} the $b$-matrix summarizes the attenuating effect of all gradient waveforms applied in all three directions, $x$, $y$ and $z$.\textsuperscript{22,24,25} In isotropic media, gradients applied in orthogonal directions do not result in cross-terms; in anisotropic media, however, they can.

Finally, it is easy to see that DT-MRI subsumes DI. If the medium is isotropic, then $D_{xx} = D_{yy} = D_{zz} = D$, and $D_{xy} = D_{xz} = D_{yz} = 0$. Then, eqn (1) reduces to a model of isotropic diffusion with $b = b_{xx} + b_{yy} + b_{zz} = \text{Trace}(\mathbf{b})$.

**QUANTITATIVE PARAMETERS OBTAINED BY DT-MRI**

Quantitative parameters provided by diffusion-tensor MRI can be obtained and explained using a geometric approach. Intrinsic quantities can be found that characterize different unique features, for example, describing the size, shape and orientation of the root mean square (rms) displacement profiles within an imaging volume, which can be represented as diffusion ellipsoids. Scalar parameters, functionally related to the diagonal \textit{and} off-diagonal elements of $\mathbf{D}(x, y, z)$, can also be displayed as an image, revealing ways in which the tensor field varies from place to place within the imaging volume.\textsuperscript{28} These quantities are rotationally invariant, i.e. independent of the orientation of the tissue structures, the patient’s body within the MR magnet, the applied diffusion sensitizing gradients, and the choice of the laboratory coordinate system in which the components of the diffusion tensor

\begin{equation}
\ln \left( \frac{A(\mathbf{b})}{A(\mathbf{b} = 0)} \right) = -\sum_{i=1}^{3} \sum_{j=1}^{3} b_{ij}D_{ij} = -(b_{xx}D_{xx} + 2b_{xy}D_{xy} + 2b_{xz}D_{xz} + b_{yy}D_{yz}) + 2b_{yz}D_{yz} + b_{zz}D_{zz}) = -\text{Trace}(\mathbf{bD})
\end{equation}

where $A(\mathbf{b})$ and $A(\mathbf{b} = 0)$ are the echo magnitudes of the diffusion weighted and non-diffusion weighted signals respectively, and $b_{ij}$ is a component of the symmetric $b$-matrix, $\mathbf{b}$. In DT-MRI a symmetric $b$-matrix is calculated for each DWI.\textsuperscript{22} The $b$-matrix summarizes the attenuating effect of all gradient waveforms applied in all three directions, $x$, $y$ and $z$.\textsuperscript{22,24,25} We then use each DWI and its corresponding $b$-matrix to estimate $\mathbf{D}$ using multivariate linear regression* of eqn (1) as in Basser \textit{et al.}\textsuperscript{20}

There are two important distinctions between DI and DT-MRI. First, DI is inherently a \textit{one-dimensional} technique, i.e. it is used to measure the projection of all molecular displacements along one direction at a time. Therefore, it is sufficient to apply diffusion gradients along only one direction. DT-MRI is inherently \textit{three-dimensional}; one must apply diffusion gradients along at least six noncollinear, non-coplanar directions in order to provide enough information to estimate the six independent elements of the diffusion tensor from eqn (1).\textsuperscript{20}

Second, the $b$-matrix formalism forces us to expand the notion of ‘cross-terms’ between imaging and diffusion gradients, to account for possible interactions between imaging and diffusion gradients that are applied in orthogonal directions, and even between imaging gradients that are applied in orthogonal directions.\textsuperscript{22,24,25} In isotropic media, gradients applied in orthogonal directions do not result in cross-terms; in anisotropic media, however, they can.
and magnet field gradients are measured. Some examples are given below.

**First moment of the diffusion tensor field**

The first moment of the diffusion tensor field, or the orientationally averaged value of the diffusion tensor field can be calculated at each point within an imaging volume:

\[
(D) = \text{Trace}(\mathbf{D})/3 = (D_{xx} + D_{yy} + D_{zz})/3 = (\lambda_1 + \lambda_2 + \lambda_3)/3 = \langle \lambda \rangle
\]

where \(\lambda_1, \lambda_2, \lambda_3\) are the three eigenvalues and \(\langle \lambda \rangle\) their mean.

Physically, an estimate of \(\langle D \rangle\) can be obtained by taking the arithmetic average of ADCs acquired in all possible directions. By integrating over all directions uniformly, we obtain an intrinsic property of the tissue, which is independent of fiber orientation, gradient directions, etc. Recently, terms like `trace-ADC`, `mean trace`, `trace mean`, etc. have been used to signify \(\langle D \rangle\), however these terms are not meaningful. We suggest, as an alternative, the term `bulk mean diffusivity`.

Several interesting issues remain unresolved about the distribution of \(\text{Trace}(\mathbf{D})\) within tissue. For example, why is it so uniform within normal adult brain parenchyma? In particular, why is its value so similar in normal white and gray matter, even though these tissues are so different histologically? This spatial uniformity has contributed to the increasing clinical utility of \(\text{Trace}(\mathbf{D})\) in disease assessment and monitoring since it makes diseased regions more conspicuous when juxtaposed against the homogeneous background of normal parenchyma. A second reason that makes \(\text{Trace}(\mathbf{D})\) useful is that it appears so similar between and among normal human subjects. In fact, it appears to be quite similar across a range of normal mammalian brains including mice, rats, cats, monkeys and humans. It is worth considering whether mammalian brains are `designed' to force \(\text{Trace}(\mathbf{D})\) to lie within such a narrow range of values and, if so, what these optimal design criteria are.

As an aside, `trace-weighted' or `isotropically weighted' DWIs have become a popular means of depicting regions in which the diffusivity has changed (particularly dropped) with respect to the surrounding tissue. Some isotropically weighted sequences have already been implemented commercially. In a Trace-weighted DWI, image intensity is brighter in regions of low diffusivity, making them more conspicuous. One way to construct a Trace-weighted image is to take a geometric mean of \(N\) DWIs, which we designate by \(A(\mathbf{b}^i)\), so that the trace-weighted intensity (TWI) becomes:

\[
\text{TWI} = \sqrt[N]{\prod_{i=1}^{N} A(\mathbf{b}^i)}
\]

If the DWI signal attenuation is given by:

\[
A(\mathbf{b}^i) = A(0)e^{-(\mathbf{b}^i \cdot \mathbf{D}^\text{TWI})}
\]

then the conditions for producing a Trace-weighted DWI are:

\[
\sum_{i=1}^{N} b^i_{xx} = \sum_{i=1}^{N} b^i_{yy} = \sum_{i=1}^{N} b^i_{zz} = N/3 > 0
\]

i.e. that the total diffusion weighting along the \(x\), \(y\) and \(z\) directions is the same, and

\[
\sum_{i=1}^{N} b^i_{xy} = \sum_{i=1}^{N} b^i_{yx} = \sum_{i=1}^{N} b^i_{zx} = 0
\]

i.e. that the sum of each of the off-diagonal elements of the \(b\)-matrix is zero. In this way

\[
\text{TWI} = A(0) e^{-\beta \text{Trace}(\mathbf{D})}
\]

which results in an image whose intensity is `weighted' by \(\text{Trace}(\mathbf{D})\).

**Higher moments of the diffusion tensor field in measures of diffusion anisotropy**

The second and higher moments of \(\mathbf{D}\) have been proposed for use as diffusion anisotropy measures because they characterize different ways in which the diffusion tensor field deviates from being isotropic. This has resulted in a number of diffusion anisotropy measures based upon the second moment of the distribution of the eigenvalues of \(\mathbf{D}\): \((\lambda_1 - \langle \lambda \rangle)^2 + (\lambda_2 - \langle \lambda \rangle)^2 + (\lambda_3 - \langle \lambda \rangle)^2\), such as the relative anisotropy (RA), and the fractional anisotropy (FA), which characterize the eccentricity of the diffusion ellipsoid. The RA is just the coefficient of variation of the eigenvalues, which has been previously used in crystallography as an `aspherism coefficient'. Anisotropy measures based upon the higher moments of the diffusion tensor or the distribution of eigenvalues of \(\mathbf{D}\), such as the skewness or kurtosis, could potentially be used to characterize diffusion anisotropy more completely, but the effects of noise make such measures unreliable (see below).

**Other diffusion anisotropy measures**

Novel anisotropy measures have been proposed that are based upon a `barycentric' representation of the diffusion tensor, in which it is decomposed into line-like, plane-like, and sphere-like tensors corresponding to diffusion ellipsoids that are prolate, oblate and spherical, respectively. The information provided by this interesting approach should be compared systematically with the
information contained in the first three moments of \( D \), the mean, variance and skewness. One issue that should be examined is the sensitivity of the barycentric representation to the order in which the eigenvalues of \( D \) are sorted. Whereas the moments of \( D \) given above are insensitive to the order of the eigenvalues, dependence on their order renders quantities susceptible to a statistical bias caused when these eigenvalues are sorted.34

Another novel anisotropy measure has recently been proposed by Frank42 to treat cases in which two or more distinct fiber populations may occupy a voxel. When this occurs, the diffusion tensor measured using the single tensor model represents only a powder average of the underlying tensors. This always results in a reduction in the measured diffusion anisotropy. Frank’s method is to measure ADCs in many non-collinear directions, and to calculate the variance of these ADC measurements about their mean value, which he defines as a new anisotropy measure. The sensitivity of this measure to the SNR of the acquisition, the degree of diffusion sensitization, and other features of the experimental design should be further investigated.

Characterizing orientational properties of the diffusion tensor field

Another important development in DT-MRI is the introduction of quantities that reveal architectural features of anisotropic structures, such as nerve fiber tracts in brain. Useful information can be gleaned from the directional pattern of diffusion ellipsoids within an imaging volume. Early on, it was proposed that, in ordered fibrous tissues, the eigenvector associated with the largest eigenvalue within a voxel is parallel to the local fiber orientation.21 Imaging methods that apply this idea include direction field mapping (in which the local fiber direction is displayed as a vector in each voxel) and fiber-tract color mapping (in which a color, assigned to a voxel containing anisotropic tissue, is used to signify the local fiber-tract direction43–46).

DT-MRI fiber tractography

DT-MRI fiber tractography21,47–60 is a method for following fiber-tract trajectories within the brain and other fibrous tissues. Here, fiber-tract trajectories are generated from the fiber-tract direction field in much the same way that fluid streamlines are generated from a fluid velocity field. Many unexpected artifacts in DT-MRI fiber tractography can arise when discrete, coarsely sampled, noisy, voxel-averaged direction field data48 are used, or when one attempts to follow incoherently organized nerve pathways.51,62 These artifacts could suggest ‘phantom’ connections between different brain regions that do not exist anatomically. Therefore, great care must be exercised both in obtaining and in interpreting such ‘connectivity’ data.47,53

Differential geometry and algebraic measures

A less intuitive, but powerful method of motivating and developing quantitative imaging parameters from DT-MRI data is by considering the differential geometry and algebraic properties of the diffusion tensor field itself, whose local features are sampled discretely in a DT-MRI experiment. Until recently, this approach was only of academic interest since there was no practical method to obtain a continuous representation of a diffusion tensor field from the noisy, voxel-averaged, discrete diffusion tensor data. However, this situation has changed with the advent of methods to construct such tensor field representations.57,63,64 For instance, this approach has lead to new applications such as DT-MRI tractography, hyperstreamline and hyperstreamsurfacing imaging,65 connectivity analysis,66 and should lead to other innovations that were not previously feasible.

For example, in structurally complex anisotropic media, such as the heart, which has a laminar architecture, one can also attempt to describe the deformation (curving, twisting, and bending) of the normal, rectifying, and oscillating ‘sheets’ formed by muscle and connective tissue. To do this, we can construct surfaces from the diffusion tensor field, which can be parameterized by two variables. Concepts of the differential geometry of surfaces67 can then be used to determine additional geometric features of sheet shape that can be calculated and displayed as intrinsic MRI parameters. These include the First and Second Fundamental Forms, \( I \) and \( II \), and the normal, gaussian and mean curvatures.67 These parameters are intrinsic because they characterize different features of the local shape of the lamina, independent of the coordinate frame of reference, and constitute new parameters.

ARTIFACTS IN DT-MRI

Subject motion

Subject motion during data acquisition can cause ghosting or artifactual redistribution of signal intensities within DWIs. Artifacts resulting from rigid body motion are the easiest to correct for, since this involves applying a uniform phase correction to an entire image. This problem has been only partially addressed by incorporating navigator echoes in the DWI pulse sequence.68,69 However, artifacts due to other physiological motion, for example, eye movements or pulsation of cerebrospinal fluid, are more difficult to correct. While these artifacts are mitigated by the use of fast echo-planar DWI sequences and cardiac gating, no general theoretical
Post-processing strategies generally aim to warp each DWI to a common template. An interesting approach is to use a mutual information criterion to determine a warp that maximizes the overlap within a series of diffusion-weighted images. Other post-processing methods that have been proposed include correcting the phase map using a model of the effects of the eddy current on it, and mapping the eddy current induced fields directly.

**Magnetic susceptibility effects**

Large discontinuities in bulk magnetic susceptibility, such as those occurring at tissue–air interfaces, produce local magnetic field gradients that notoriously degrade and distort DWIs, particularly during echo-planar imaging. In addition to image distortion, susceptibility variations within the brain adversely affect DWIs because the additional local gradients act like diffusion gradients causing the $b$-matrix to be spatially varying. This problem is partly compensated for by the use of the logarithm of the ratio of the diffusion-weighted to non-diffusion-weighted intensity [see eqn (1)], in which case the effect of these susceptibility-induced gradients, which are present during both diffusion-weighted and non-diffusion-weighted images, is cancelled.

Susceptibility effects are particularly acute in the brain in regions adjacent to the sinuses. As DWIs are being acquired increasingly on higher field strength magnets (3 T and above), these problems will become more severe. Figure 2 illustrates a typical example of susceptibility-induced artifacts at 3 T using a single-shot EPI sequence. It is clear that, at higher field strengths,
different strategies to those used for imaging at 1.5 T are needed. While the distortions in Figure 2 are typical of distortions using single-shot EPI at 3T, diffusion weighted SENSE acquisitions appear to mitigate these artifacts.\textsuperscript{76}

**Image noise**

While at low levels of diffusion weighting the logarithm of the signal attenuation decreases linearly with increasing $b$-value, in common with all MRI acquisitions, background noise causes the DWI intensity to approach a baseline ‘noise floor’ as one progressively increases the degree of diffusion weighting. Even above this baseline, noise in DWIs can introduce significant bias in the estimates of the eigenvalues, which makes isotropic media appear anisotropic, and anisotropic media appear more anisotropic.\textsuperscript{34} Pierpaoli \textit{et al.} have proposed the Lattice Anisotropy index which combines pure measures of anisotropy with measures of inter-voxel coherence of eigenvectors in neighboring voxels to try to ameliorate this problem.\textsuperscript{34} This bias often makes higher order measures, such as the skewness or kurtosis, too inaccurate to be used reliably.

More recently, it was found that RF noise also biases the mean and variance of the eigenvectors of $\mathbf{D}$.\textsuperscript{80} Unfortunately, the current understanding of the deleterious effects of this artifact is more advanced than our understanding of useful remedies to correct for it.\textsuperscript{80,81}

In the presence of noise, it is possible to estimate negative eigenvalues of the diffusion tensor using the regression methods described above. Physically, each of the principal diffusivities of the diffusion tensor must be non-negative (i.e. the tensor is positive semi-definite). To ensure this condition, some groups have imposed the constraint of positive definiteness on the diffusion tensor explicitly.\textsuperscript{82} This can be done, for example, by minimizing the error norm or $\chi^2$ value subject to the constraints that the three principal diffusivities are positive.

**Hardware issues**

Background gradients can be present if the magnetic field is improperly shimmed, leading to additional signal attenuation if not properly compensated for. This problem can often be remedied by measuring the background gradients directly\textsuperscript{83} and incorporating them explicitly in the $b$-matrix.

Gradient non-linearity and miscalibration can lead to small but significant errors in the calculation of the diffusion coefficient or diffusion tensor elements. The signal attenuation and the gradient pulse sequence must be known for each DWI. If the gradients are not well calibrated, or they are not linear, signal attenuation attributed to the diffusion processes in the sample could be miscalculated using eqn (1). If the gradients in the $x$, $y$ and $z$ directions are coupled to one-another (i.e. there is cross-talk) because of misalignment of the gradient coils, then gradients applied in logical directions may have

---

**Figure 2.** Susceptibility artifacts in DT-MRI images acquired at 3 T with single-shot EPI. The phase encode direction is left–right. (a) $T_2$-weighted image intensity. (b) Fractional anisotropy map
components in other directions, which could be particularly problematic in DT-MRI.\textsuperscript{84,85}

Flipping the signs of the diffusion gradients on alternate averages has been suggested as an aid in removing cross-terms. However, owing to noise, there is a question whether such averaging should be done on the logarithm of the intensity or whether the geometric mean should be performed on the measured intensities. Clearly, this method would ameliorate cross-terms arising from specific interactions between imaging and diffusion gradients applied in parallel or orthogonal directions, but not cross-terms arising between imaging gradients, which, albeit small in MRI clinical applications, would not be corrected. The advantage of this method is that it could simplify the analysis of DWI data. In new single-shot acquisition schemes that acquire DWIs with gradients applied in many directions,\textsuperscript{84,86} this strategy has the disadvantage that twice as many DWIs would have to be acquired. Moreover, an additional parameter is required per voxel as a free parameter in the estimation of the diffusion tensor to determine the cross-terms due to the imaging gradients alone.

**MODEL SELECTION**

One of the advantages of using the single effective diffusion tensor formalism in tissue is that it provides a great deal of new information without making many explicit assumptions about the underlying tissue architecture and microstructure. The only explicit assumption is that the diffusion characteristics can be represented by a single symmetric gaussian displacement distribution. We can appreciate its conceptual economy by considering the large complexity of minimal models one could propose to represent the actual known biological compartments with tissues.

Inferring the microstructure and the underlying architectural organization of tissue using diffusion imaging data is complicated by several factors. First, homogeneity within each voxel cannot be assured. Numerous microscopic compartments exist within brain parenchyma. \textit{A priori}, we must assume that gray matter, white matter and cerebrospinal fluid (CSF) could occupy the same macroscopic voxel. The crudest model one could pose, then, for these three compartments is:

\[
\frac{A}{A_0} = f_1 e^{-\text{Trace}(bD_{wm})} + f_2 e^{-\text{Trace}(bD_{gm})} + f_3 e^{-\text{Trace}(bD_{ot})}
\]

where $D_{wm}$ represents the diffusion tensor for white matter, and $D_{gm}$ and $D_{ot}$ represent the apparent diffusion coefficients for gray matter and for CSF, respectively, which are assumed to be isotropic. In this model, the three compartments are not assumed to be exchanging. We could further simplify this model by requiring that the $f$-coefficients sum to one. This would implicitly involve assuming that the $T_1$ and $T_2$ in each compartment were the same. Even so, this model contains $1 + 7 + 2 + 1 = 11$ parameters to estimate in each voxel.

If we were to assume that there is an additional white matter compartment, described by two individual diffusion tensors, $D_{wm}^1$ and $D_{wm}^2$, such as in regions where white matter fibers cross,\textsuperscript{34} the model would be further complicated, involving 18 free parameters:

\[
\frac{A}{A_0} = f_1 e^{-\text{Trace}(bD_{wm}^1)} + f_2 e^{-\text{Trace}(bD_{wm}^2)} + f_3 e^{-\text{Trace}(bD_{gm})} + f_4 e^{-\text{Trace}(bD_{ot})}
\]

If the diffusion processes within the white matter fiber compartments were cylindrically symmetric, either described by cylindrically symmetric prolate or oblate ellipsoids, then we could impose additional constraints on the diffusion tensors, $D_{wm}^1$ and $D_{wm}^2$, which would reduce the number of free parameters to estimate each tensor from 6 to 4.\textsuperscript{87} This is because only two parameters are needed to specify the orientation of the ellipsoid in space, and two parameters to specify its principal diffusivities.

Additionally, intra and extracellular compartments without exchange can further complicate the model. Returning to the model in eqn (8), we now have to add two additional compartments for isotropic gray and anisotropic white matter, although obviously not for CSF:

\[
\frac{A}{A_0} = f_1 e^{-\text{Trace}(bD_{wm}^1)} + f_2 e^{-\text{Trace}(bD_{wm}^2)} + f_3 e^{-\text{Trace}(bD_{gm})} + f_4 e^{-\text{Trace}(bD_{ot})}
\]

Already, we have accumulated $1 + 7 + 6 + 2 + 1 + 1 = 18$ free parameters for this model. If we use the rule of thumb that one would like to obtain at least four times as many DWIs as the number of free parameters to estimate them, we require now at least 72 DWIs. Generally, however, the number of these distinct tissue types and their distribution within the voxel is unknown. At an ultrastructural level, gray and white matter are themselves generally quite heterogeneous, having a distribution of macromolecular structures of varying size, shape, composition and physical properties (such as $T_2$, $D$). Thus, even the model in eqn (10) is quite naive.

Several groups have recently tried to address the problem of multiple white matter compartments in the brain. Their approach was to use a two-compartment model with non-exchanging spins, in which each compartment is characterized by its own diffusion tensor.\textsuperscript{88,89}

\[
\frac{A}{A_0} = f_1 e^{-\text{Trace}(bD_1)} + f_2 e^{-\text{Trace}(bD_2)}
\]

Even with this simple model, there are 14 free parameters to estimate. As each white matter compartment is added,
another six parameters are needed to prescribe the diffusion tensor. This leads to a proliferation of free parameters to estimate. It also taxes experimental resources and one’s ability to design an efficient experiment. Which diffusion gradient directions and gradient strengths are optimal? How many DWIs are needed? How can we tell if this model is adequate? These become complicated multifarious problems to consider. At present, a general framework for assessing the adequacy of different models to describe water diffusion within a voxel has yet to be proposed and implemented.

Other complexities

Differences in relaxation parameters can lead to different rates of echo attenuation in each compartment, making it more difficult to explain the cause of signal loss within a voxel. There are also irregular boundaries between macromolecular and microscopic-scale compartments. Different macromolecular structures comprising these boundaries may affect the displacement distribution of water molecules differently, necessitating even more complex models. Water molecular motion may be restricted or hindered. Even within a compartment, some water will be associated with certain macromolecules while some will be free to diffuse.

Another unknown is whether there is exchange between compartments, which can also affect the relaxation rates of the spin system. How water moves within and between compartments is still not well understood. Owing to differences in blood flow and thermal conductivity, temperature cannot be assumed to be uniform throughout a tissue sample. It is well known that temperature affects the measured diffusivity (~1.5% per 1 °C)90,92 and is predicted to have the same effect on all diffusion tensor components.

For all these reasons, the underlying cause of diffusion anisotropy has not been fully elucidated in brain parenchyma, although most investigators ascribe it to ordered, heterogeneous structures, such as large oriented extracellular and intracellular macromolecules, supermacromolecular structures, organelles, and membranes. In the central nervous system (CNS), diffusion anisotropy is not simply caused by myelin in white matter, since several studies have shown that, even before myelin is deposited, diffusion anisotropy can be measured using MRI.93-96 Thus, despite the fact that increases in myelin are temporally correlated with increases in diffusion anisotropy, structures other than the myelin sheath must be contributing to diffusion anisotropy.97 This is an important point, because there is a common misconception that the degree of diffusion anisotropy can be used as a quantitative measure or ‘stain’ of myelin content, when in reality no such simple relationship exists.

WHAT DOES D TELL US ABOUT TISSUE AT A LOW b-VALUE?

Suppose, for generality, we have $N$ non-interacting compartments, each described by its own diffusion tensor, but each with the same $T_1$ and $T_2$. Expanding the model describing the diffusion attenuation about $\text{Trace}(b\mathbf{D}) = 0$ gives

$$
\frac{A}{A_0} = \sum_{i=1}^{N} f_i e^{-\text{Trace}(b\mathbf{D}_i)}
$$

$$
\approx \sum_{i=1}^{N} f_i \left[ 1 - \text{Trace}(b\mathbf{D}_i) + \frac{1}{2} \left( \text{Trace}(b\mathbf{D}_i) \right)^2 \right]
$$

This can be rewritten as:

$$
\frac{A}{A_0} \approx \sum_{i=1}^{N} f_i - \text{Trace} \left( b \sum_{i=1}^{N} f_i \mathbf{D}_i \right) + \ldots = 1 - \text{Trace}(b\mathbf{D}_{\text{eff}}) + \ldots
$$

In the vicinity of $\text{Trace}(b\mathbf{D}) = 0$, eqn (13) can be well approximated by the following expression:

$$
\frac{A}{A_0} \approx e^{-\text{Trace}(b\mathbf{D}_{\text{eff}})} \quad \text{where} \quad \mathbf{D}_{\text{eff}} = \sum_{i=1}^{N} f_i \mathbf{D}_i
$$

Thus, in this case, the single effective diffusion tensor model produces a quantity, $\mathbf{D}_{\text{eff}}$, which is just the weighted average of each of the diffusion tensors within the volume. The expression in eqn (14) shows that, in the vicinity of $\text{Trace}(b\mathbf{D}) = 0$, the attenuation due to two compartments that are rapidly exchanging and due to two compartments that are non-exchanging are the same. Figure 3 illustrates the diffusion-weighted intensities plotted as a function of $b$ for a two-compartment system under both the rapid-exchange and no-exchange models. It has been assumed that compartment 1 has a volume fraction, $f_1 = 0.70$ and $\text{ADC}_1 = 1.02 \times 10^{-3} \text{mm}^2 \text{s}^{-1}$, and compartment 2 has a volume fraction, $f_2 = (0.3 = 1 - 0.70)$, and $\text{ADC}_2 = 0.18 \times 10^{-3} \text{mm}^2 \text{s}^{-1}$. (These values were suggested for the intra- and extracellular tissue compartments of white matter by Clark et al.98) Note that, in the figure, at low $b$-factors the difference between the non-exchanging and rapidly exchanging intensities is very small.

Hence, while DT-MRI produces a well-defined effective diffusion tensor in the linearly decaying regime of the relationship between the logarithm of the signal attenuation and the elements of the $b$-matrix, it provides no information about features of the individual compartments. The interpretation of the effective diffusion tensor is not necessarily clear in a more complex tissue at higher $b$-values.
DT-MRI FINDINGS AT ‘HIGH b-VALUES’

Recently, with the advent of stronger magnetic field gradients, several groups have reported multi-exponential decay of the MR signal intensity as a function of $b$-value.\textsuperscript{98,99} Some have inferred from this data that properties of distinct tissue compartments can be meaningfully observed using DT-MRI. Clearly, there is interesting biological information to be gleaned in the non-linear regime that may help to resolve some of these issues. High $b$-value acquisitions are being treated by others in this issue, so only a few points will be made below that pertain to the DT-MRI aspects.

Putting aside the complexities of obtaining stable estimates of discrete exponentials (i.e. diffusion relaxation), numerous microstructural and architectural configurations could produce the same multi-exponential relaxation data. For example, Peled recently showed that a system of impermeable tubes with a distribution of diameters consistent with those found in histological brain slices could give rise to multi-exponential decay.\textsuperscript{100} Similar behavior is expected when there is a statistical distribution of any relevant physical property or microstructural dimension within a voxel. With the exception of CSF, it is unlikely that a particular exponential can ever meaningfully be assigned to a particular and distinct tissue compartment. Clearly, without invoking additional \textit{a priori} information about tissue structure, tissue composition, the physical properties of the different compartments and their spatial distribution, determining tissue microstructural and architectural features from the NMR signal is an ill-posed intractable inverse problem.

Another approach to analyzing tissues with multiple compartments is to use three-dimensional $q$-space techniques, originally developed by Callaghan.\textsuperscript{101,102} Most recently, this strategy has been suggested by Tuch \textit{et al.} for use in clinical scanners, using conventional diffusion-weighted images.\textsuperscript{99,103} In principle, the potential benefits of using such an approach are numerous: $q$-space MRI would provide a probability displacement distribution within each voxel in a model-independent way. This method would also allow one to vary the diffusion time and the length scale probed independently and systematically.

However, unlike ‘$q$-space’ MRI in which ‘high’ and ‘low’ $q$-values provide an unambiguous physical interpretation of the length scale probed and the diffusion time of the experiment, in clinical DWI acquisitions, neither of these experimental quantities is well defined. True three-dimensional $q$-space MRI requires ultrashort ($\sim 1 \text{ ms}$) and ultrastrong ($\sim 500 \text{ G cm}^{-1}$) gradient pulses. Producing these using a whole body or even a head gradient set would clearly cause a severe shock to the patient. If one attempts to use DWIs with longer duration and weaker diffusion gradient pulses, one loses the ability to infer the underlying displacement distribution, the diffusion time and the length scale probed.
STATISTICAL ASPECTS OF DT-MRI ACQUISITION

Since the diffusion tensor data is statistically estimated in each voxel, we must treat it as a set of random variables. However, the statistical analysis of DT-MRI data is complicated by several factors. Although in an ideal DT-MRI experiment D has been shown to be distributed according to a multivariate normal distribution, and thus Trace (D) has been shown to be normally distributed, the parametric distribution of many other derived DT-MRI parameters is either unknown or known not to be normal. In these cases, we are precluded from using statistical hypothesis testing methods that assume an underlying normal distribution to determine whether an observed difference between different regions of interest (ROI) is statistically significant. In such cases, empirical methods like the bootstrap—which allows determination of the distribution of a statistical parameter empirically without knowledge of the form of its distribution a priori—show great promise in diffusion imaging studies. This is particularly true now that many single-shot DWIs can be acquired during a single scanning session. Bootstrap methods become more practicable and accurate as more raw data are available. One promising application of this empirical statistical method is to assure data quality and to test for systematic artifacts that might be present during the acquisition. In this way, statistical properties of measured DT-MRI parameters can be studied meaningfully on a voxel by voxel basis.

MULTI-SITE AND MULTI-SUBJECT STUDIES

In performing multi-site or longitudinal diffusion tensor imaging studies, several additional issues arise. The most basic is how to compare high-dimensional diffusion tensor data from different subjects or from the same subject acquired at different time points. Applying warping transformations developed for scalar images, can produce nonsensical results when applied to DT-MRI data without taking appropriate precautions. Our understanding of admissible transformations that can be applied to warp and register diffusion tensor field data is still limited.

A second issue to consider is the proliferation of measures derived from the diffusion tensor to characterize different features of the isotropic and anisotropic diffusion. Consistent definitions of quantities such as the orientationally averaged diffusivity, RA, FA, etc. should be employed. It is advisable to use the same imaging acquisition hardware, reconstruction software and post-processing routines to control for unnecessary variability. All sites should use a well-characterized phantom, even if it is an isotropic phantom, to ensure that no systematic artifacts occur, and that the DWI acquisition is stable in time and across platforms.

CONCLUDING REMARKS

DT-MRI provides new means to probe tissue structure at different levels of architectural organization. While experimental diffusion times are associated with water molecule displacements on the order of microns, these molecular motions are ensemble-averaged within a voxel, and then subsequently assembled into multi-slice or three-dimensional images of tissues and organs. Thus, this imaging modality permits us to study and elucidate complex structural features spanning length scales ranging from the macromolecular to the macroscopic, without the use of exogenous contrast agents.

New structural and functional parameters provided by DT-MRI, such as maps of the eigenvalues of the diffusion tensor, its Trace, measures of the degree of diffusion anisotropy and organization and estimates of fiber direction all advance our understanding of nerve pathways, fiber continuity and, potentially, functional connectivity in the CNS.

Acknowledgement

The Wellcome Trust supports DKJ.

REFERENCES

11. Moseley ME, Kucharczyk J, Asgari HS, Norman D. Anisotropy
DIFFUSION-TENSOR MRI


90. Basser PJ. A sensitive method to calibrate magnetic field gradients using the diffusion tensor. Third Meeting of the SIRM, Nice, 1995; 308.


