Magn Reson Med Sci 2022; 21; 132–147
doi:10.2463/mrms.rev.2021-0005
Published Online: May 21, 2021

REVIEW

An Invited Review for the Special 20th Anniversary Issue of MRMS

# Recent Advances in Parameter Inference for Diffusion MRI Signal Models

Yoshitaka Masutani[1*]

In this paper, fundamentals and recent progress for obtaining biological features quantitatively by using diffusion MRI are reviewed. First, a brief description of diffusion MRI history, application, and development was presented. Then, well-known parametric models including diffusion tensor imaging (DTI), diffusional kurtosis imaging (DKI), and neurite orientation dispersion diffusion imaging (NODDI) are introduced with several classifications in various viewpoints with other modeling schemes. In addition, this review covers mathematical generalization and examples of methodologies for the model parameter inference from conventional fitting to recent machine learning approaches, which is called Q-space learning (QSL). Finally, future perspectives on diffusion MRI parameter inference are discussed with the aspects of imaging modeling and simulation.

**Keywords:** *diffusion magnetic resonance imaging, signal models, parameter inference, Q-space learning*

## Introduction

Since the pioneer work by Le Bihan et al.[1] on clinical application of diffusion weighted imaging (DWI) in 1986, diffusion MRI (dMRI) remains still as a developing research field. It covers a wide spectrum of related areas including not only MR physics and medical science but also applied mathematics and computer science.[2–4] In addition, the successful applications in clinical practice are remarkable, thanks to progress in both hardware and software aspects. The former can be represented by hardware-related development for faster imaging such as echo planar imaging (EPI).[5] Also, related imaging sequences with physical and mathematical models for DWI signal, a focus of this review, are included in the latter. After a certain utility of DWI was clinically confirmed for the acute stroke[6] in the 1990s, the dMRI application, mainly in neuro-brain area, was explosively spread by the diffusion tensor imaging (DTI)[7] in the 2000s. The dMRI brought new information on both structural and functional aspects to the brain science.[8] For example, the fMRI analysis associated with the white matter fiber tract network opened new horizon of brain function assessment.[9] Also, recent research on dMRI-based fMRI[10] may contribute further development of structural and functional analysis of the brain. Clinical applications are also in a wide range in both diagnostic and therapeutic purposes. For the diagnostic purpose, the researches on dMRI-based detection and staging of diseases and disorders cover various targets, such as stroke,[11] infarction,[12] tumors,[13] multiple sclerosis,[14] Parkinson's disease,[15] Alzheimer disease,[16] Creutzfeldt-Jakob disease,[17] and so on. To realize minimally invasive surgery and radiotherapy, information on the white matter fiber tracts helps planning and navigation. For instance, the corticospinal tract extracted from dMRI is often utilized as an object that should not be damaged during therapy.[18,19] In addition, the target organs and tissues of dMRI have been extended to other areas, not only the spine[20] and the prostate[21] but also whole-body scan[22] for cancer detection.

In another viewpoint, the dMRI also covers a wide range of spatial scale of anatomy, for example, in neuro-brain area, from global structure of white matter fiber tracts to microstructure of axons. For global structure analysis, tractography[23–25] is an epoch-making tool to visualize the white mater fiber tracts, which had been invisible with conventional imaging modalities. Also, as described later, several parametric dMRI signal models can quantify the fine structures in micrometer order such as axon diameter. One of the most significant features of dMRI can be its quantitativeness to characterize local properties of organs and tissues. A single DWI can provide contrast between tissues based on the difference of water diffusion property. In addition to such visual information, a lot of kinds of quantities can be obtained by multiple acquisition of DWI. The available quantities depend on the

[1]Graduate School of Information Sciences, Hiroshima City University, Hiroshima, Japan

*Corresponding author: Graduate School of Information Sciences, Hiroshima City University, 3-4-1, Ozukahigashi, Asaminami-ku, Hiroshima, Hiroshima 731-3194, Japan, Phone/Fax: +81-82-830-1774, E-mail: masutani@hiroshima-cu.ac.jp



Received: January 12, 2021 | Accepted: March 4, 2021

models used. For example, DTI provides diffusion tensor, which can be converted to mean diffusivity and diffusion anisotropy. These quantities are obtained voxel by voxel to provide dMRI parameter maps, which can play many roles of new image modalities. Those parameters at arbitrary location can also provide tract-specific information with the additional information of tractctography, that is, quantitative evaluation in tract-by-tract manner is available. The analysis is called tract-specific analysis, which is used for the evaluation of Alzheimer disease,[26] autism,[27] and so on. An essential and unique feature of dMRI is that it provides not only functional information related to phenomenon of water diffusion but also morphological information of biostructure that constrains the diffusion. Wide varieties of all those features are provided not only by imaging sequences to acquire DWI but also by additional analysis models. That is, how the DWI signal is formed by structures that affect water diffusion.

Recent progress in dMRI can be categorized in the following three aspects: DWI acquisition, analysis model, and analysis methodologies. One of the major and important updates in recent DWI acquisition is diffusion encoding gradients, also called as motion probing gradient (MPG). In the standard acquisition of DWI by using spin echo, a single pair of gradient pulses in both sides of the 180 degrees for phase shift and refocusing is employed. This technique is categorized in pulsed gradient spin echo (PGSE) by single diffusion encoding (SDE). One of the drawbacks of this technique is difficulty in shortening diffusion time for measurement, which is necessary to evaluate the compartment sizes of microstructures restricting water diffusion. In the recent techniques, more complex MPG is used. First, multi-diffusion encoding (MDE), including double diffusion encoding (DDE),[28] was introduced. Namely, MDE applies MPGs in more than two directions. It allows quantification of microscopic diffusion anisotropy.[29–31] In addition, non-pulse gradient techniques called oscillating gradient spin echo (OGSE)[32] were also introduced for shorter diffusion times. These techniques lead us to more detailed information of microstructures. According to the rapid development of MRI acquisition including hardware progress, more detailed dMRI data have become available. The set of dense samples in Q-space leads to more detailed analysis by using more complex models. For detailed description of the physical phenomenon, it has the affinity with the field of the applied mathematics. Such models consider more parameters or features by embedding anatomical structures in the model. Typical models are reviewed in this manuscript later. In addition to the models, analysis methodologies have been also developing. One of the recent important factors in dMRI analysis methodologies is machine learning, which provides empirical solutions for complex problems. For example, machine learning techniques were proved to be feasible in various dMRI problems, including inference of signal model parameters.

Thus, based on the recent progress of dMRI described above, fundamentals and recent progress for obtaining biological features quantitatively by using dMRI are reviewed in this paper. Especially, the important focus is twofold. One is parametric signal models, in which features are clearly defined as parameters. Several classifications in a few viewpoints are also provided. The other is methodologies for model parameter inference. To discuss the methodologies, mathematical generalization of the problem is presented. In the following section, Models for dMRI Analysis, reviews of signal models with several classifications are shown. Among the analysis models, the focus of this review paper is parametric models. Next in section, Parameter Inference for Signal Models, reviews of parameter inference methodologies for parametric signal models are described with a mathematical formulation of the problem. Then, in section, Q-space Learning (QSL) and Examples, recent approaches for parameter inference by machine learning, namely, QSL, are introduced. Especially, details on synthetic approaches for QSL are provided, which consist of the simulation of DWI signal formation with noise. Also, the examples on several signal models are presented by comparing with conventional methods for parameter inference. Finally, in section, Perspectives and Discussion, discussion and perspectives of dMRI analysis are shown for closing.

## Models for dMRI Analysis

In this section, various models for dMRI analysis, especially for signal modelling, are introduced.

Since Stejskal–Tanner spin echo signal model[33] in MR spectroscopy (MRS) has introduced to DWI, various signal models have been proposed. Those models could be categorized in several viewpoints, such as dimension and diffusion Gaussianity. Among them, one of the important categorizations is parametric or non-parametric. Therefore, the following part of this section is threefold. One is on parametric models, in which quantities of physical and/or biological features are included as explicit parameters. Next is non-parametric, in which diffusion-related functions such as probability density function (PDF) or orientation distribution function (ODF) are defined. The last one is based on new concepts including geometric modelling of microstructures and diffusion simulation.

### *Parametric signal models*

The parametric signal models of dMRI are approximations on how the DWI signals are formed based on local properties of tissues as model parameters and image acquisition setting values. Basically, the models describe a single DWI value that is determined by local properties at a single location. There exist really a lot of signal models for dMRI. For better understanding, those signal models are classified in several viewpoints. Below, three major

classifications are described. The simplest classification of dMRI signal models is dimension, that is, 1D or 3D. The former is only for a single orientation of diffusion measurement, while the latter is for whole 3D space. In other words, it is also a dimension of Q-space,[34] in which measured signal values are located. The next is diffusion type, that is, whether the observed diffusion phenomenon is regarded as Gaussian or not. More precisely, the Gaussianity is for the spatial distribution of water molecules, after certain period of time, which are initially at an identical location. If the assumption is true, only a single parameter of diffusion coefficient can represent the phenomenon in 1D, or a single diffusion tensor can in 3D. Moreover, several models assume a mixture of Gaussian, that is, the observed dMRI signal consists of several compartments in different diffusion speed. Basically, in biological structures excepting for limited space, water diffusion is restricted somehow by those structures. In such environments of non-Gaussian diffusion, signal models need extra parameters. Also, consideration of such relation between water diffusion and biostructures made dMRI signal models developed. In the early phase, the models are purely generic to represent environments of physical phenomenon of diffusion. In the next generation, some models consider anatomical structures explicitly. Such models are dedicated to specific analysis; parameters related to biostructures such as fiber geometry are embedded in the model equations. In this section, typical and well-known dMRI signal models are introduced with the classification in the viewpoints above and also summarized in Table 1.

**Stejskal–Tanner (ST) model**

The first model is the 1D Gaussian diffusion model by Stejskal and Tanner, which is originally proposed in MRS.[33] This could be the simplest dMRI model without assumption for any anatomical structures. The DWI signal $S$ is expressed by a single exponential form:

$$S = S_0 \cdot \exp(-bD) \tag{1}$$

where $S_0$ is the baseline signal without MPG, $b$ is the so-called b-value representing the strength of MPG, and $D$ is the diffusion coefficient as the model parameter. The absolute magnitude of $S$ depends on amplitude setting of signal processing in MR scanners, normalized signal (or signal decay) $E$ is often used instead as follows:

$$E = \frac{S}{S_0} = \exp(-bD) \tag{2}$$

$S_0$ is often replaced by measured signal by $b = 0$, but the true value is also a model parameter included in the model. Hereafter in this article, $E$ is used for the equations of various signal models. Except for $S_0$, diffusion environment is represented by a single parameter of $D$ in this model. It is simple but useful for approximating total estimation of diffusivity at a voxel.

**intra-voxel incoherent motion (IVIM)**

The next is the IVIM model,[35] which is often referred to as bi-exponential model. It considers that diffusion signal contains not only diffusion but also perfusion by randomly oriented micro-circulation of blood flow. The model equation is in a bi-exponential form.

$$E = \frac{S}{S_0} = f_p \cdot \exp\{-b(D + D^*)\} + \left(1 - f_p\right) \cdot \exp(-bD) \tag{3}$$

where $f_p$ represents the fraction of perfusion, $D$ is the apparent diffusion coefficient, and $D^*$ is the equivalent diffusion coefficient by regarding random perfusion as diffusion process. As the $D^*$ is derived from blood flow, it is much higher than $D$. Therefore, the coefficients are often referred to as $D_{slow}$ and $D_{fast}$. The design of IVIM model introducing $f_p$ and $D^*$ is aimed at reproduction of certain environment of bio-structures. However, it can also be regarded as just a mixture of two different Gaussian diffusion phenomena, where explicit anatomy is not necessarily expressed.

**Diffusional kurtosis imaging (DKI)**

The DKI signal model is proposed by Jensen and colleagues[36] for quantifying non-Gaussianity of 1D diffusion. Originally, diffusional kurtosis is given for the shape of the PDF of water molecule displacement, which is obtained by Q-space imaging (QSI).[37] The DKI signal model is designed to obtain the kurtosis parameter directly from DWI signals without PDF. The model equation is in an extended form of ST model as follows:

$$E = \frac{S}{S_0} = \exp\left(-bD + \frac{1}{6}b^2D^2K\right) \tag{4}$$

where $K$ represents the diffusional kurtosis of the additional parameter. The model is intended to evaluate diffusion environment restricted by biostructures. Similar to IVIM, however, anatomical consideration is also implicit in the DKI model.

**DTI**

The DTI signal model[7] also assumes Gaussian diffusion, but in 3D space, where diffusion anisotropy is allowed. Basically, the DTI model is similar to the ST model,[33] in which diffusion in 1D is quantified by a single parameter of diffusion coefficient. To express anisotropy, diffusion tensor is introduced, and the quadratic form of the tensor and orientation vectors shows the diffusion coefficient in the orientation. The diffusion tensor is a second-order tensor in the form of symmetric and positive definite 3 × 3 matrix with six independent components. In the DTI model, normalized signal $E$ in direction $\mathbf{d} = \left(d_x, d_y, d_z\right)^t$ is expressed by the equation including diffusion tensor $\mathbf{D}$:

$$E(\mathbf{d}) = \frac{S}{S_0} = \exp\left(-b\mathbf{d}^T\mathbf{D}\mathbf{d}\right) \tag{5}$$

**Table 1** dMRI parametric signal models and attributes

| Model name | Dimension | Number of compartments | Parameters | Anatomical Parameters | Inference methods |
|---|---|---|---|---|---|
| Stejskal and Tanner | 1D | 1 | $D$ | none | Closed-form, least squares |
| IVIM | 1D | 2 | $D, D_p, f_p$ | fractions for compartments | non-linear least squares |
| DKI | 1D | 1 | $D, K$ | none | LM[32], closed-form only for three measurements[55], general closed-form[56] |
| DTI | 3D | 1 | $\mathbf{D}$ | none | Multivariate linear regression by linear solution[6] |
| DKTI | 3D | 1 | $\mathbf{D}, \mathbf{W}$ | none | LM[32] |
| Ball and Stick | 3D | 2 (isotropic and anisotropic) | $f_i, f_a(=1-f_i), D, \theta_{(i)}, \varphi_{(i)}$ | fiber orientation(s), fractions for compartments | Bayesian estimation with shrinkage priors and MH-MCMC sampling[37] |
| CHARMED | 3D | 2 (hindered and restricted) | $f_{r(i)}, f_h(=1-\sum f_{r(i)}), R, D_{\parallel(r)}, D_\perp, \mathbf{D}_h$ | radius and orientation (s) of fiber, fractions for compartments | LM[38,39] |
| AxCalibar | 1D | 2 (hindered and restricted) | $f_{r(i)}, f_h(=1-\sum f_{r(i)}), R_{(i)}, D_{\parallel(r)}, D_\perp$ | (fiber orientation and) radii of fibers, fractions for compartments | LM[40] |
| Active-Ax | 3D | 4 (hindered, restricted, stationary water and CSF) | $f_h, f_r, f_{sw}\ (=1-f_h-f_r), f_{CSF}, \theta_{(i)}, \varphi_{(i)}, D_t$ | radius and orientation of fiber, fractions for compartments | MCMC[42], AMICO[54] |
| NODDI (Bingham-NODDI) | 3D | 3 (isotropic, hindered and restricted) | $f_i, f_r, , f_h(=1-f_r), \boldsymbol{\mu}, OD_{(p/s)}$ | fiber orientation distribution, fractions for compartments | gradient descent[43], AMICO[54], MLP[59,64,67] |
| SANDI | 1D (averaged) | 3 (extracellular, intra-neurite, intra-soma) | $f_h, f_{in}, f_{is}(=1-f_{in}), D_{ec}, D_{in}, R_s$ | Radius of soma, fractions for compartments | random forest regression[45] |
| VERDICT | 3D | 3 (vascular, hindered and restricted) | $D, D_p, f_v, f_r, f_h(=1-f_v-f_r), \theta_{(i)}, \varphi_{(i)}, R$ | radius and orientation of fiber, fractions for compartments | LM[46] |
| Reisert (3-compartment model) | 1D | 3 (intra-axonal, extra-axonal, and free water) | $f_r, f_h, f_i(=1-f_r-f_h), R, D_{\parallel(r)}, D_{\parallel(h)}, D_\perp, \mathbf{D}_h$ (with given fiber orientation estimated) | (fiber orientation and) fractions for compartments | Bayesian polynomial regression[58] |

$\boldsymbol{\mu}$: mean orientation of fibers in Watson/Bingham distribution. $\theta_{(i)}, \varphi_{(i)}$: fiber orientation angle(s). $\mathbf{D}$: diffusion tensor (2nd order) with 6 independent components. $D$: apparent diffusion coefficient (or diffusivity). $D^*$: pseudo-diffusion coefficient for perfusion. $D_{\parallel(r/h)}$: axon-parallel (longitudinal) diffusivity in restricted/hindered compartment. $D_\perp$: axon-radial (transverse) diffusivity in restricted compartment. $D_{ec}$: effective extracellular diffusion coefficient (isotropic). $\mathbf{D}_h$: hindered diffusion tensor with only 2 eigenvalues. $D_{in}$: longitudinal apparent diffusion coefficient in neurite. $D_t$: tortuous diffusivity for axon-radial orientation. $f_a$: signal volume fraction for anisotropic diffusion (stick). $f_{CSF}$: diffusivity for CSF. $f_i$: signal volume fraction for isotropic diffusion (ball). $f_h$: signal volume fraction for hindered diffusion (extra-axonal). $f_{in}$: signal volume fraction for intra-neurite compartment. $f_{is}$: signal volume fraction for intra-soma compartment. $f_p$: signal volume fraction for perfusion. $f_{r(i)}$: signal volume fraction(s) for restricted diffusion (intra-axonal). $f_{sw}$: diffusivity for stationary water. $f_v$: signal volume fraction for vascular compartment. $K$: diffusional kurtosis. $OD_{(p/s)}$: fiber orientation dispersion (in primary/secondary orientation). $R_{(i)}$: radius/radii of axon. $R_s$ radius of soma. $\mathbf{W}$: diffusion kurtosis tensor (4th order) with 15 independent components. MCMC, Markov Chain Monte Carlo; MH, Metropolis Hastings.

That is, diffusion coefficient in an arbitrary direction can be obtained by the quadratic form: $\mathbf{d}^T\mathbf{D}\mathbf{d}$. Each element in $\mathbf{D}$ is difficult to understand intuitively, but eigenvalues and eigenvectors of $\mathbf{D}$ obtained by matrix diagonalization represent features that are easier to recognize. The eigenvalues show the orientational maxima of diffusion coefficient, and the eigenvalues indicate their orientations. The well-known quantities of DTI, mean diffusivity (MD), and fractional anisotropy (FA), are derived from those eigenvalues. In application of the DTI model, the eigenvalue for the maximum eigenvalue is often approximated as orientation of fibers as the simple tractgraphy algorithms assume.[23–25] This is also an implicit consideration of anatomy. In addition, similar to IVIM for ST model, the multi-tensor model (MTM)[38] was also introduced for the DTI model. The MTM is also a mixture of Gaussian diffusion and expressed in the multi-exponential form:

$$E(\mathbf{d}) = \frac{S}{S_0} = \sum_{i=1}^{N} f_i \cdot \exp\left(-b\mathbf{d}^T\mathbf{D}_i\mathbf{d}\right) \tag{6}$$

Its important advantage is expression of complex fiber structures such as fiber crossings.[39,40] This concept is succeeded to more anatomy-conscious models described later.

**Diffusional kurtosis tensor imaging (DKTI)**

Similar to dimension upgrade from ST model to DTI model, DKTI model was also designed to obtain 3D description of non-Gaussian diffusion by the fourth-order tensor.[36] The elements of the kurtosis tensor are often expressed as $W_{ijkl}$ where *i, j, k,* and *l* are one of *x, y,* and *z*. The kurtosis tensor is also symmetric and therefore consists of 15 independent components. As the DTI model provides diffusion coefficient in an arbitrary direction by the quadratic form, the DKTI model can give us kurtosis in an arbitrary direction $\mathbf{d} = \left(d_x, d_y, d_z\right)^t$ by the equation:

$$K(\mathbf{d}) = \frac{\bar{D}}{D(\mathbf{d})} \sum_{i\in x,y,z} \sum_{j\in x,y,z} \sum_{k\in x,y,z} \sum_{l\in x,y,z} d_i d_j d_k d_l W_{ijkl} \tag{7}$$

where $\bar{D}$ is the average diffusion coefficient in all orientation, $D(\mathbf{d})$ is the diffusion coefficient in the direction $\mathbf{d}$.

**Ball and stick(s) model (BSM)**

The BSM is originally designed for the probabilistic tractography.[41] It explicitly assumes free water and fibers in one orientation or multiple orientations. The ball means the former, that is, isotropic diffusion tensor in the shape of sphere, while the latter is expressed as stick(s), which means diffusion tensor(s) similar to the DTI model, but of the highest anisotropy, in which two eigenvalues are equal to zero. In this model, the signal $E$ in direction $\mathbf{d}$ for given number of fiber orientations $N$ is expressed as:

$$E = \frac{S}{S_0} = \left(1 - \sum_{i=1}^{N} f_i\right) \cdot \exp(-bD_{iso}) + \sum_{i=1}^{N} f_i \cdot \exp\left(-b\mathbf{d}^T\mathbf{D}_i\mathbf{d}\right) \tag{8}$$

where $D_{iso}$ is the diffusion coefficient by the isotropic diffusion of free water; $f_i$ and $\mathbf{D}_i$ are the volume fraction and diffusion tensor of the *i*th orientation.

Hereafter, several models with anatomical parameters are introduced. Because the definition of signal model equations is much more complicated, the detailed definitions of the formulas are omitted in this review, which can be accessed in the original literature. Rather, the main structures of compartments are emphasized in this article to aim at the general understanding of the models.

**Composite hindered and restricted model of diffusion (CHARMED) and AxCalibar**

Similar to the BSM, the CHARMED was proposed.[42,43] It assumes that the signal consists of two compartments of different diffusion nature. One is for a hindered or extra-axonal compartment, while the other is for a restricted or intra-axonal one. For the former, diffusion properties are expressed by a diffusion tensor with certain anisotropy. For the latter, a restricted diffusion is modelled as water molecules within cylinders are employed for characterizing diffusion properties. The equation for the single hindered and restricted compartment model seems very similar to the BSM and is expressed as:

$$E = \frac{S}{S_0} = f_h \cdot E_h + (1 - f_h)E_r \tag{9}$$

where the pairs of *f* and *E* are volume fraction and signal decay for hindered and restricted diffusion. In CHARMED, the hindered diffusion term can be a summation of anisotropic diffusion tensor unlike the BSM. Some parameters in CHARMED are common with the BSM, including the fiber orientations, volume fractions of extra- and intra-axonal compartments, and principal diffusivities. A new parameter in CHARMED is axon size for the restricted diffusion, which is defined with multiple size distribution in the multiple compartment model of CHARMED. Another unique feature of CHARMED is that it is assumed that hindered diffusion explains Gaussian signal attenuation observed at low b-value, while the restricted diffusion describes non-Gaussian at high b-value.

Apparently, the AxCaliber model[44] is a derivative of CHARMED by simplifying it in 1D. The model is designed only for describing the diffusion perpendicular to the fiber. It also includes parameters of axon radius with an extension of population with multiple size. The equation is identical to that of CHARMED (eq.9), but in a 1D form. As a 1D model in fiber-radial orientation similar to the AxCaliber model, Zhou and Laidlaw[45] presented a model including axon size as a parameter with new features of water exchange time among intra-axon and extra-axon spaces.

Note that those parameters of microstructures are apparently beyond the image resolution of standard MR imaging, of which

pixel size is in millimeter order. That is, the parameters such as axon diameter are averaged ones in a single pixel. Recent researches[46, 47] focusing on such partial volume averaging of microstructure parameters are expected to resolve the limitation.

**ActiveAx**

Another fiber size model ActiveAx was proposed by Alexander et al. for parameterizing microstructure including axon radius.[48] Similar to the CHARMED, it assumes that DWI signal is intra- and extra-axonal diffusion and cerebrospinal fluid (CSF). In addition, it adopts another assumption of diffusion environment such as fixed diffusivity and existence of stationary water.

**Neurite orientation dispersion diffusion imaging (NODDI) and Bingham NODDI**

The NODDI model[49] was proposed to quantify the fiber orientations in low coherency. The model equation includes three compartments: isotropic diffusion such as in CSF, and intra- and extra-cellular diffusion as follows:

$$E = \frac{S}{S_0} = f_{iso} \cdot E_{iso} + (1 - f_{iso}) \cdot \{f_{ic} \cdot E_{ic} + (1 - f_{ic}) \cdot E_{ec}\} \tag{10}$$

where $f$ and $E$ represent the signal fraction and the normalized signal of the three elements, respectively. The partial signal $E_{ic}$ consists of the degree of angular variation of neurites approximated by Watson distribution and its central orientation. While the Watson distribution assumes isotropic dispersion of orientations, the Bingham distribution assumes anisotropic dispersion, which is employed in the Bingham NODDI model.[50]

**Soma and neurite density imaging (SANDI)**

As a model similar to NODDI, the SANDI model[51] was proposed for diffusion environment in the grey matter (GM), where water molecules in the extra-cellular space with those in soma are combined. The equation consists of weighted signals of three compartments:

$$E = \frac{S}{S_0} = f_{ic} \cdot (f_{in} \cdot E_{in} + f_{is} \cdot E_{is}) + f_{ec} \cdot E_{ec} \tag{11}$$

where $f_{is}$ and $E_{is}$ are volume fraction for soma and normalized signal, respectively. In the $E_{is}$, soma radius is included as a parameter. Another uniqueness is that it employs MPG orientation-averaged signals for a practical solution.

**Vascular extracellular and restricted diffusion for cytometry in tumors (VERDICT)**

The diffusion MRI-based modeling of microstructure is describing not only for healthy anatomy but also for pathological structures. The model: VERDICT considers microstructures around tumors with blood vessels.[52] The VERDICT parameters include volume fraction of blood vessels and fractions of intra- and extra-cellular axon diffusion. It is unique that it quantifies blood flow by pseudo-diffusivity.

### *Non-parametric models*

Unlike parametric models in which some parameters represent diffusion environment, non-parametric model extract information of diffusion environment in less simplified form. Several functions such as diffusion PDF and ODF are used. In this category, the models do not express DWI signals but rather express analysis manner strongly associated with imaging technique. For acquisition of diffusion PDF, QSI[37] and diffusion spectrum imaging (DSI)[53] are well known. These non-parametric models can provide more detail or raw information on diffusion environment by function forms. However, when it is applied to specific purpose, several quantities are extracted from the function. For example, for 1D QSI, maximum probability or full width at half maximum (FWHM) in PDF is used for characterizing local diffusion properties. Diffusional kurtosis is also available from PDF as mentioned above. Generally, PDF is obtained by the Fourier transform of the Q-space data. Therefore, certain sample density in the Q-space is needed, which elongate imaging time.

The ODF represents orientation probabilities for the existence of fiber structures, of which local maxima is often regarded as fiber orientations. Unlike discrete models such as the CHAMED and the BSM, orientational probability of fiber existence is continuously expressed. Basically, the ODF is obtained as a radial projection of PDF in Q-space. However, Q-ball imaging (QBI)[54] model simplifies acquisition of ODF without PDF, in which a single-shell Q-space data can be converted to ODF by Funk-Radon transform. The application of ODF is mainly for fiber tracking in both deterministic and probabilistic manners. For local maxima detection of ODF, preprocessing in the frequency domain is often performed by spherical harmonics and constrained spherical deconvolution (CSD).[55]

### *New approaches in modeling*

As the last category following the parametric and non-parametric models, a new approach in dMRI modeling is introduced. It is initiated by fine geometric modeling of microstructures to construct a virtual living body, which constrains water diffusion. In the neuro-brain area, the constructed structures are mainly neural cells with axons, soma, and neurites, in addition to surrounding structures such as micro-blood vessels. Then, the physical phenomenon of diffusion is simulated by using random walk process within the virtual environment.[56,57] The blood flow can be simulated as pseudo- (or apparent) diffusion of incoherent motion. Finally, DWI acquisition is simulated for a certain pulse sequence considering the fundamental process of phase shift and refocusing of spins. In this approach, the objective for modeling is not approximation of signal formation but for the whole process related to diffusion from microscopic random-walk to image acquisition. Such simulation-based modeling may open a new horizon of dMRI leading to new signal models as discussed in the last section of this article.

## Parameter Inference for Signal Models

Real DWI signals are influenced by many factors including biological phenomenon and measurement system of MR scanner. The dMRI signal models are quite simplified by neglecting several factors of details. Therefore, even in an ideal measurement environment, there exists systematic error due to model simplification. On the other hand, real measurement system cannot avoid random noise in imaging, which also yields difference among values of real signal and model signal. The inference methods of dMRI signal model parameters premise those two factors. In this section, the methodologies of parameter inference in dMRI signal models are described. For a set of given measured signals and dMRI signal model, parameter inference is a mathematical problem. First, mathematical foundation is introduced by generalizing parameter inference problem. Then, three categories of methodologies are described.

### *Generalization as mathematical problem*

In dMRI signal models, DWI signal value at location $\mathbf{x}$, $S$ can be expressed in the following form:

$$S(\mathbf{x}, \mathbf{c}) \cong F(\mathbf{p}(\mathbf{x}), \mathbf{c}) \tag{12}$$

where $\mathbf{p}(\mathbf{x}) = (p_1, p_2, ..p_N)$ represents model parameters consisting of $N$ values, i.e. local physical and/or biological features at $\mathbf{x}$, and c is the set of acquisition setting values constants, such as MPG strength and direction. Often, normalized signal $E(\mathbf{x})$ is also used for expression as follows:

$$E(\mathbf{x}, \mathbf{c}) \cong \frac{S(\mathbf{x},\mathbf{c})}{S_0} = G(\mathbf{p}(\mathbf{x}), \mathbf{c}) \tag{13}$$

In this form, $S_0$ is also a parameter that represents ideal signal value when MPG is off. In practice, measured $S$ without MPG is used alternatively. Thus, signal models assume that DWI signals are determined by two factors: local tissue properties (model parameters) and image acquisition setting values. Note that the former is location-variant, while the latter is location-invariant and common in a single acquisition. On the other hand, real signal values of DWI measurements are also obtained as: $S_m(\mathbf{x}, i)$ $(i = 1,\ 2, \ldots M)$ where the number of measurements is $M$, and the index $i$ is for the single acquisition with setting $\mathbf{c}(i)$. Note that those measured signal values are contaminated by noise. Generally, a dMRI dataset consists of multiple DWI measurements by different acquisition setting values. For example, at DTI, $\mathbf{p}(\mathbf{x})$ consists of the baseline signal $S_0$ and six independent elements in diffusion tensor; $\mathbf{c}$ is a set of MPG direction vectors and b-values. In summary, parameter inference is the determination of $\mathrm{p}(\mathbf{x})$ for each location $\mathbf{x}$, with the cue of $M$ sets of measurement values $S_m(\mathbf{x}, i)$ with settings $\mathbf{c}(i)$ $(i = 1,\ 2, \ldots M)$ and model equation $S(\mathbf{x}, \mathbf{c})$.

Mathematically, there are several solution strategies depending on the state of the problem. When the number of parameters is unknown, $N$ is equal to the number of acquisitions; $M$, the problem is often mathematically even-determined and the solution could be often expressed in a closed form. If $M > N$, it is called over-determined, while $M < N$ cases are called as under-determined. Here, please note that not only the number of measurements matters for $M$ but also the locations at Q-space. For instance, even if we repeat measurement hundred times at an identical Q-space location, we cannot obtain diffusion tensor, which needs six or more non-collinear directions of MPG. Also, diffusional kurtosis cannot be estimated with single shell data. Thus, $M$ means the effective number of measurements that matches the requirement of signal models, including the numbers of Q-space locations and shells. Generally, most of the dMRI measurements provide over-determined situation and are conventionally solved by fitting as described in the next section. Under-determined situation is rare and not desirable under existence of noise, but when strong constraints exist, it might become solvable. In most of the dMRI parameter inference, one of the important factors is countermeasure for noise, which is one of the important reasons for making over-determined situation in dMRI measurements. In real DWI measurement with limited acquisition time, it is hard to increase the measurement, that is, increase the number of $M$. To tackle with this problem, two approaches are available. One is the spatial regularization, which is based on the fact that biological features are spatially smooth, and parameters in neighborhood are similar. This scheme is easily embedded in parameter inference by fitting described later. The other is noise reduction as preprocessing. It is also called as image filters for denoising, which is often used in general image processing area. This aspect is also discussed in the last section of this review. The following three sections introduce the categories of parameter inference, which are fitting, closed-form, and regression.

### *Fitting*

So far, most of the dMRI parameters are inferred by fitting technique due to their over-determined situation. Fitting is performed by minimizing value of an error function between model signal value: $S$ and measured signal value: $S_m$ by searching the set of parameters: $\mathbf{p}$ to minimize the function. That is, inferred parameters are defined as follows:

$$\hat{\mathbf{p}} = \arg\min_{\mathbf{p}} J(S, S_m) \tag{14}$$

where $J$ represents an error function between $S$ and $S_m$. The function is basically a summation of errors for each acquisition setting value $\mathbf{c}(i)$. For example, when error is defined as the square of difference of $S$ and $S_m$, $J$ is expressed as follows:

$$J = \sum_{i=1}^{M} |S(\mathbf{x}, i) - S_m(\mathbf{x}, i)|^2 \quad (15)$$

The fitting by minimizing this is called as least-square method (LSM).[58] The minimization of error function is a mathematical problem to search optimal parameters, for which a lot of numerical algorithms exist,[59] such as gradient descent (GD), Gauss–Newton (GN), and Levenberg–Marquardt (LM). In those methods, initial parameter is set first, and then parameters are iteratively updated to obtain less error values. This search process in the parameter space is repeated until a certain convergence is recognized or by fixed number of iteration. The solution depends on the complexity of signal models, but local minima could be a problem, similar to general optimization problems.

As a variation of fitting, an approach combined with dictionary-based solution and spatial regularization was proposed. In this approach, called as accelerated microstructure imaging via convex optimization (AMICO),[60] measurements are approximated by a linear combination of inferred parameters or their reformatted values. The optimized parameter is obtained by:

$$\hat{\mathbf{p}} = \arg\min_{\mathbf{p}} |\mathbf{Qp} - \mathbf{S_m}| + R(\mathbf{p}) \quad (16)$$

where $\mathbf{Q}$ is the linear operator of matrix for $\mathbf{p}$, which is called dictionary, and $R$ is the spatial regularization term, for which L1-norm of the parameters is used. The elements of the dictionary are given based on signal model equation to be fitted. An important feature of this approach is the use of convex optimization, which is never bothered by the local minima problem. Also, thanks to convex optimization with linear formulation by dictionary, the solution is obtained very fast with guaranteeing spatial smoothness. Thus, fitting is the standard technique of dMRI parameter inference though several optional updates have been proposed.

### *Closed form*

Closed-form solution is available only for special cases of signal models in simple formulation. It is applicable when the problem is even-determined. Generally, in this solution, parameters are explicitly obtained as:

$$p_i = H_i(S_m(1), \ldots, S_m(M), c(1), \ldots, c(M)) \quad (17)$$

That is, parameters are explicitly expressed by combination of measured signal values and acquisition setting values. For example, the diffusion coefficient: $D$ in ST-model can be simply obtained by two measurements as follows:

$$D = \frac{(\log S_1 - \log S_2)}{b_2 - b_1} \quad (18)$$

where $S_1$ is the measured signal value with b-value of $b_1$, and $S_2$ is by $b_2$. In addition to the ST model, the DKI model with three measurements has closed-form solution[61] for the parameters as follows:

$$D = \frac{(b_3+b_1)D^{(12)} - (b_2+b_1)D^{(13)}}{b_3 - b_2}, K = 6\frac{D^{(12)} - D^{(13)}}{(b_3 - b_2)D} \quad (19)$$

where $D^{(ij)}$ denotes the value $\ln\big(S(b_i)/S\big(b_j\big)\big)/\big(b_j - b_i\big)$. The closed-form solutions in dMRI parameter inference are, however, strongly influenced by noise because it assumes no noise. Therefore, only when the signal-to-noise ratio (SNR) of dMRI datasets is quite high, it can be an option. Otherwise, the obtained map is in very low quality for noise-sensitive parameters such as diffusional kurtosis. Also, in a few over-determined cases, solutions of fitting can be obtained in a closed form. For instance, DKI parameters can be expressed in closed forms for more than three measurements.[62]

### *Regression*

For given signal model function: $S$, regression provides empirical determination of pseudo-inverse function: $S^{-1}$ by combinations of various functions including linear and non-linear ones. This is realized by machine learning techniques based on sample (or training) data to approximate mapping from inputs: $S_m$ to outputs: $\mathbf{p}$. The regression-based parameter inference is also referred to as QSL due to its input data in Q-space. It is apparently similar to the closed-form solution in which parameters are calculated from the measured values: $S_m$ and acquisition setting values: $\mathbf{c}$. In QSL, inputs of pseudo-inverse function are often limited to only measured signals or its converted data as discussed in the next section. While closed-form solution is applicable only for few simple models, regression can be performed at any form of signal equations as far as input and output samples are given. For regression, a lot of pairs of measured signals and corresponding set of parameters are required for training. The inferred parameters are from various models, such as fractional anisotropy for DTI, diffusional kurtosis for DKI, and orientation dispersion for NODDI. Also, fiber orientations are also determined by machine learning for the purpose of tractography.[63] In the next section, Q-space learning is described with sample results.

## QSL and Examples

### *Overview*

In this section, a recent approach of dMRI parameter inference categorized in QSL is introduced with sample results. For dMRI parameter inference by machine learning-based regression, there have been proposed several ways for construction of the mapping function as a regressor by machine learning. Reisert and colleagues[64] reported regression by polynomial function with Bayesian inference scheme for a three-compartment model. Golkov et al.[65] employed neural networks of multi-layer perceptron (MLP) for DTI, DKI, and NODDI. Nedjati-Gilani et al showed that random forest has advantages in regression in comparison with other regressors.[66] In addition to those applied parameters, there exist some variations of QSL for training dataset as described in the next section.

### *QSL variations*

Input values for QSL regressor include DWI signal values similar to the closed-form solution. Those can be raw signal, normalized signal decay, or logarithm of them. However, unlike the closed-form solution, the pseudo-inverse function: $S^{-1}$ learned from training data is acquisition-specific. In other words, fixed acquisition setting is assumed. In principle, it is possible to construct a mapping function including acquisition setting values as input. However, it requires further wide range and vast quantity of training samples, and therefore it is not practical at present. It implies that a QSL regressor requires the Q-space sample locations identical to those of training data. It is not convenient in the viewpoint of compatibility. As alternatives to make QSL input independent of acquisition setting, there exist two approaches. One is conversion of Q-space data to frequency domain data, for which spherical harmonics coefficients[67] are often used. The other technique is Q-space resampling[62] with interpolation to match sample locations with those of training data. As the interpolator, the family of radial basis functions[68,69] including thin-plate spline is often employed.

Output values in QSL training data are basically provided by the conventional inference methods including fitting. Instead of real data, training data can be synthesized by a simulation process of dMRI acquisition. That is, a set of parameters in a dMRI signal model can yield corresponding DWI signal values by using its model equation. This can be called as synthetic QSL (synQSL).[70,71] To simulate more realistic signals, Rician noise is often added to signal values.[72] A research on the synQSL reported an important relationship among noise amounts of training data and test data.[73] That is, when the two noise levels are in the similar level, inference robustness is highest. There are several characteristics of synQSL in comparison with QSL by real data. Generally, machine learning methods suffer from shortage of training data, and data augmentation is often employed. Also, less variation of training data may cause overtraining and cannot cope with parameters out of trained range. The synQSL can solve these two problems, i.e. quantity and parameter range of training data are fully controllable. On the other hand, it is often difficult to simulate the noise in real data due to recent acquisition techniques such as parallel imaging, by which SNR is not uniform even in a single image. This problem is discussed in the last section.

### *synQSL example*

In this section, examples of dMRI parameter inference by synQSL are shown, which include parameter inference for DTI, DKI, and NODDI models.[67] As the regressor, a MLP with five layers is used, for which the logarithm values of the normalized signals are input. It assumes 14 diffusion measurements for DTI, 5 for DKI, and 73 for NODDI including the baseline signal: $S_0$, that is, the numbers of MLP inputs were 13 for DTI, 4 for DKI (1D), and 72 for NODDI. The hidden layers were trained without dropout, and the last layer was designed to output a single parameter in all models. Therefore, the numbers of units are in 13-256-256-256-1 for DTI, 4-128-128-128-1 for DKI, and 72-400-400-400-1 for NODDI. For activation, the rectified linear unit (ReLU) function was used. The training parameters were *Normal* for initialization and *Adam* for optimization, with 100 epochs, a batch size of 10000, and other default settings of the Keras/Tensorflow regressor. As training datasets for each signal model, $10^3$~$10^6$ samples were prepared for each noise level. To simulate the acquisition of dMRI signals, the signal model parameters were randomly generated within a predetermined range. After generating the model parameters, MRI signals without noise were obtained using the model equations. The measurement of dMRI signal sets was simulated for each model. For the DTI model, a single b-value is 1000 $s/mm^2$, with 13 directions of MPG vectors, and the baseline signal $S_0$ is within the range 500~8000. Similarly, four signals with b-values of 0, 311, 1244, and 2800 ($s/mm^2$) were generated for the DKI, and two shell data with b = 700 in 24 MPG directions and b = 2000 in 48 directions were simulated for NODDI. These settings were identical to those of the real MRI for testing. Finally, the generated signals were contaminated with Rician noise, which is simply given by the equation[36]:

$$S' = \sqrt{S^2 + \eta^2} \tag{20}$$

where $S'$ and $S$ represent signal values after and before noise contamination, while $\eta$ denotes the zero-mean Gaussian noise term with standard deviation $\sigma$. Four levels of SNR were configured for training by the noise ratio of $\sigma/S_0 = 0.0$, 0.005, 0.05, and 0.5.

Figure 1 shows the inference results by the conventional fitting method[68] and by the synQSL for the MD and FA values in DTI model. No significant difference between the two methods is observed. Moreover, quantitative comparison in Fig. 2 implies alight but interesting difference in the range of inferred values in both MD and FA. That is, the inferred values by the synQSL remain within reasonable ranges, in which the MD and FA values are positive, and also most of the FA values are lower than 1.0, while the results by fitting are not. This is due to the range of training data for synQSL, which were trained in the range of physically possible values.[73] The diffusional kurtosis inference results for the DKI model are shown in Fig. 3. Figure 3a shows the result by a conventional fitting,[62] while Fig. 3b-3d is by training with noise levels of $\sigma/S_0$ = 0.005, 0.05, and 0.5 respectively and with $10^5$ training samples. The noise level for training was 0.005, as observed in Fig. 3a; the diffusional kurtosis map contained a large number of black dots due to errors from negative values, which are also observed in Fig. 3a. On the contrary, Fig. 3d shows the result by $\sigma/S_0 = 0.5$, in which the inferred parameters converged to a virtually uniform value. This is often observed in synQSL results when training noise level is excessively high. The parameters seem robustly inferred in Fig. 3c, in which noise level is supposed to be matched between training and test

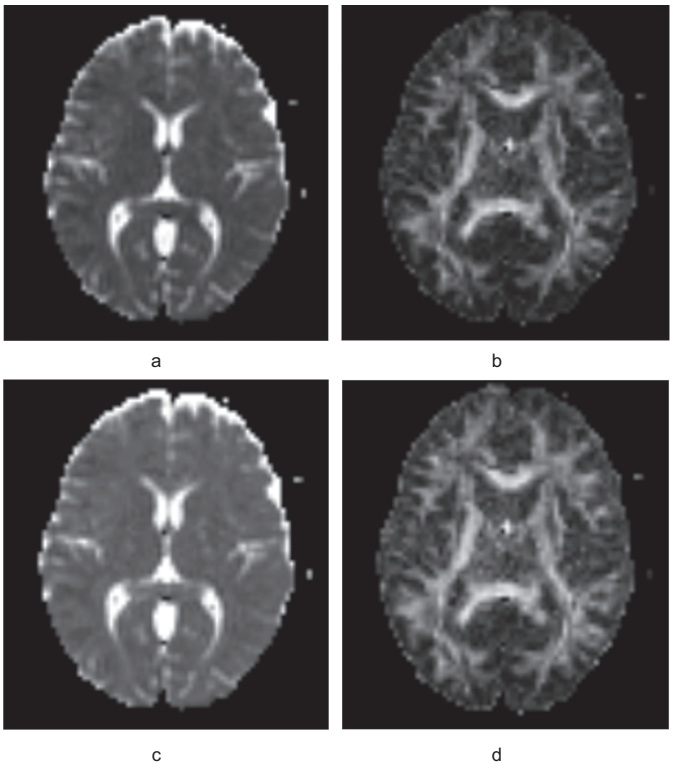


**Fig. 1** DTI parameter estimation results by conventional method (LSF) and synQSL. (**a**) MD by LSF. (**b**) Fractional Anisotropy by LSF. (**c**) MD by synQSL. (**d**) FA by synQSL. LSF, least-squares fitting.

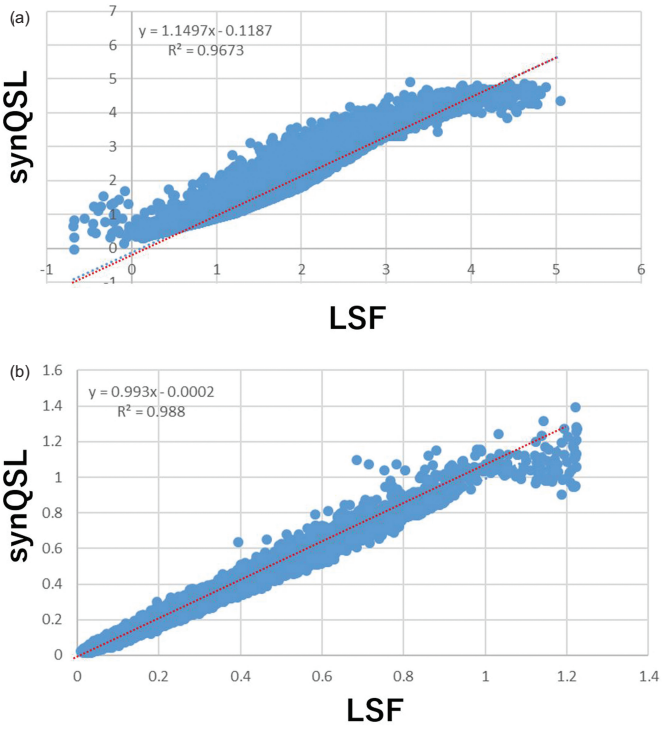


**Fig. 2** Correlation between estimated DTI parameters by conventional method (LSF) and synQSL. (**a**) MD. (**b**) FA.

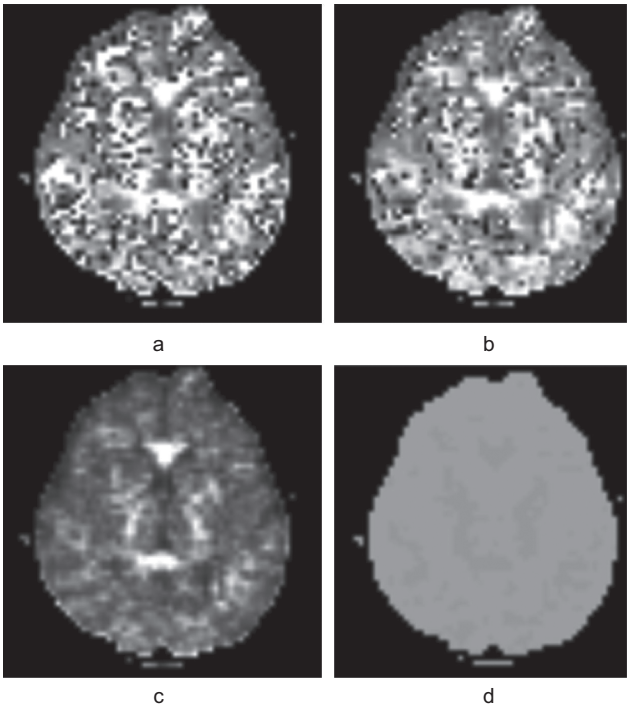


**Fig. 3** Diffusional kurtosis estimation results, (**a**) by LSF, (**b**) by synQSL (trained with lower noise level), (**c**) by synQSL (trained with optimal noise level), and (**d**) by synQSL (trained with higher noise level).

datasets. The parameter maps of the NODDI model parameters inferred by the fitting and the synQSL are shown in Fig. 4, which are inference results of isotropic and intracellular volume fractions, and fiber orientation dispersion. The synQSL results were obtained with the noise level of 0.05 and $10^6$ training samples. As seen in the figure, two methods show similar results visually. Based on synthetic data experiments as shown in Fig. 5, the importance of the noise level matching[73] was proved for minimizing inference error. In addition, Fig. 6 shows the quantitative comparison results in comparison with the results by two fitting methods: GD-based NODDI in the original paper[49] and dictionary-based AMICO-NODDI,[60] which was provided by Fukunaga et al.[75] In addition, the correlation is quite high, while the processing time is much shorter in synQSL.

### *Software implementation (diMaRIA)*

The author has developed a software tool open to public, which is based on the synQSL, and was named diffusion MR image analyzer (diMaRIA).[76] The current version of diMaRIA19b provides dMRI parameter maps of DKI and NODDI models, that is, radial and axial components of diffusion coefficient and diffusional kurtosis, and isotropic and intra-cellular volume fractions, and fiber orientation dispersion. Though the training was performed intensively on a GPU workstation for several hours per parameter, it can perform parameter inference very fast thanks to parallel processing based on multi-thread computation with the operations by single instruction, multi-data (SIMD) at multi-core CPU. Also, the software succeeded to the former software dTV,[74] in which main functions are tractography, tract-specific analysis, and DTI parameter maps. An important feature of the diMaRIA software is that it allows arbitrary Q-space sampling locations, i.e. combinations of b-values and MPG directions. In the case of Q-space sampling locations different from those of training, Q-space interpolation and resampling are automatically performed.[68] Note that the examples of synQSL result shown in the last section were obtained by a prototype of the diMaRIA software.

## Perspectives and Discussion

In this section, future perspectives of dMRI parametric models in three aspects are described and discussed, which are imaging, modeling, and inference methodology.

Regarding acquisition of dMRI data, two major streams can be expected. One is the direct imaging for specific features. Unlike post-processing-based acquisition as dMRI model parameters, some feature maps would be obtained by complex combination of diffusion encoding. Westin et al.[77] proposed a new framework of dMRI: Q-space trajectory imaging (QTI). It is based on an extended concept of OGSE, to obtain information of microstructures by diffusion tensor distribution within tissue. The QTI provides new features such as microscopic orientation coherence measure in addition to some features equivalent to those by the conventional approach such as FA. A unique characteristic of these approaches is that simultaneous development of signal model and acquisition sequence are performed, which will bring potential advantage in clinical application. On the other hand, naturally, such direct approaches of specific parameter acquisition may be lack of generalization. In that sense, the other major streams of dMRI data acquisition could be for more generic DWI scheme so that arbitrary model can be applied, in addition to the prediction ability of a particular model for unseen data.[78] The second stream includes new sampling scheme, reconstruction from less sample data, and optimization of sample locations in Q-space. Those approaches are focused on faster and high-quality imaging.

Another new aspect of future dMRI development is modeling with simulation. As introduced in the previous section the recent dMRI models are more conscious of specific structures of micro anatomy. Such anatomy-oriented modeling will lead to more precise representation of micro-environment for water diffusion. Recent modeling tries to consider more factors of water displacement. The key factors might include the intra- and extra-cellular exchange,[79] the micro-blood flow,[80] and the glymphatic system.[81] Simultaneously, it is needed to consider yet diversity of microstructures as

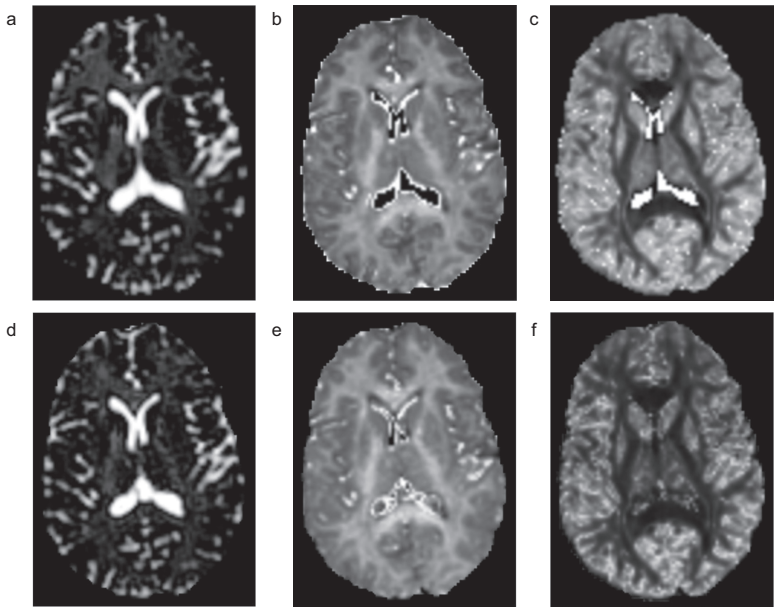


**Fig. 4** NODDI parameter estimation results by conventional method (GD) and synQSL. (**a**) $f_{iso}$ by GD. (**b**) $f_{ic}$ by GD. (**c**) OD by GD. (**d**) $f_{iso}$ by synQSL. (**e**) $f_{ic}$ by synQSL. (**f**) OD by synQSL. *fic*, volume fraction for intracellular diffusion; *fiso*, volume fraction for isotropic diffusion; OD, orientation dispersion.

the dominant factor of dMRI. For example, NODDI considers finer orientation dispersion, and QTI also assumes diffusion tensor distribution. Unlike simple models assume uniform or coherent microstructure within a voxel, new models will introduce such diversity. To realize it more precise, it is necessary to model shape and structures of micro-anatomy, that is, fine geometric modeling of microstructures. Such researches cover generating realistic axon geometries[82,83] and diffusion simulation-based on geometric axon models.[84–86] For example, Palombo and colleagues modeled neural cells including tree structures with variations.[86] Of course, it seems virtually impossible to infer all the precise structure within voxel beyond image resolution. But their work contributes to prove the evidence for that small structural change could cause DWI signal difference. Also, Rensonnet et al. proposed a new approach of microstructure fingerprinting, that is, local tissue property inference based on dictionary constructed based on diffusion simulation.[87] Generally, more complicated modeling makes its validation more difficult. In that sense, simulation of diffusion phenomenon and DWI acquisition[88] will get more important.

As far as various dMRI signal models are effective on specific diagnostic purposes and no better alternative is given, their parameter inference is still an important problem. Instead of the fitting approaches, the QSL by machine learning would play more important roles in dMRI parameter inference. As introduced in the previous sections, the regression including the QSL can be understood as the empirical formation of pseudo-inverse function of signal model equation. The complexity of the function apparently depends on the signal model. As well as other machine learning problems, the optimal types and sizes of regressors for dMRI parameter inference have been determined experimentally. The expressive power,[89] that is, how well regressors can express (or approximate) the pseudo-inverse function of signal model, is an important key for parameter inference by machine learning. Each regressor is based on different types of function elements and structures. For example, neural networks are based on linear combinations of nonlinear activation function. Ideally, sufficient flexibility to express the function should be expected. However, for training with noisy and limited amount of data, too flexible power may cause over-fitting. This is a kind of tradeoff issues depending on the signal model equations. Therefore, in addition to experimental works, theoretical and mathematical aspects should be considered to establish the criteria to choose the best type or setting of regressor for each signal model. Another important issue at dMRI parameter inference is image noise, which causes inference errors and biases. Especially, the spatially variant noise amount due to parallel imaging[90] seems hard to be solved with simple countermeasures. Currently, we can consider two options: noise reduction in DWI before inference and regressors trained with noise. Furthermore, the former can be performed in two phases: during image reconstruction[91] and post-reconstruction.[92] Unlike other images for visual diagnosis, however, quantitative images including dMRI might require high standard of noise reduction not to give new biases on the inferred parameter values. For that purpose, simulation studies might get more important including signal collection and reconstruction. Also for the regressor training with noise, two ways for noise level

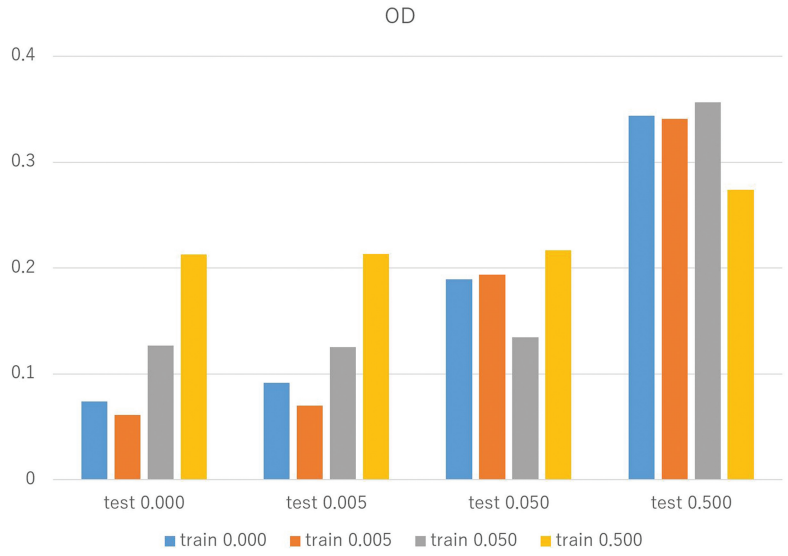


**Fig. 5** Effect of noise level matching between training data and test data in synthetic data experiments: RMS errors in various combinations of noise levels for OD estimation. RMS, root mean square.

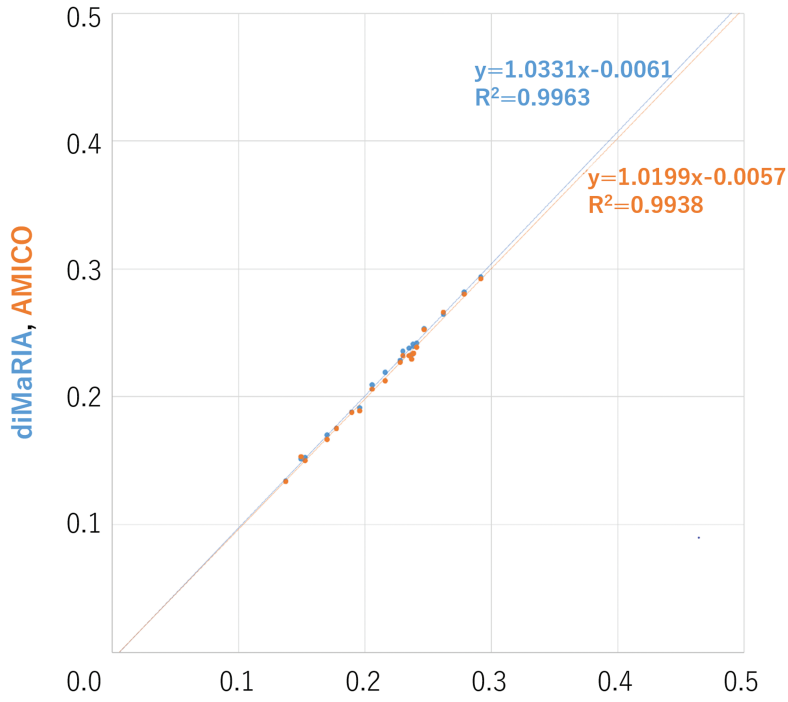


**Fig. 6** Comparison of OD estimation results by original NODDI toolbox (GD), AMICO, and synQSL: dots represent average value in various ROIs included in atlas.

matching during training can be considered: local/voxel-wise[64] approach and global one.[70,73] For both approaches, noise level estimation in DWIs plays a key role.

By putting them all together, dMRI elements of imaging, modeling, simulation, and parameter inference will be fused in a denser relationship. With certain accumulation of experimental facts with insights and assumptions on biological structures, construction of new models is expected, which include certain parameters that contribute as new diagnostic biomarkers.

## Acknowledgment

This research was partially supported by JST CREST Grant Number JPMJCR15D1. The author is deeply grateful to all the co-researchers for their valuable comments and advices.

## Conflicts of Interest

The author declares no conflicts of interest.

## References


1. Le Bihan D, Breton E, Lallemand D, et al. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. Radiology 1986; 161:401–407.
2. Jones DK. Diffusion MRI theory, methods, and applications. Oxford:Oxford University Press, 2011.
3. Johansen-Berg H, Behrens TEJ. Diffusion MRI. From quantitative measurement to *in vivo* neuroanatomy, 2nd ed. Oxford: Elsevier, 2013.
4. Aoki S, Abe O, Masutani Y, et al. Korede Wakaru Kakusan MRI, the 3rd edition. Tokyo:Gakken Medical Shujunsha, 2013. (in Japanese)
5. Mansfield P. Multi-planar image formation using NMR spin echoes. J Phys C: Solid State Phys 1977; 10:L55–L58.
6. Moseley ME, Kucharczyk J, Mintorovitch J, et al. Diffusion-weighted MR imaging of acute stroke: correlation with T2-weighted and magnetic susceptibility-enhanced MR imaging in cats. AJNR Am J Neuroradiol 1990; 11:423–429.
7. Basser PJ, Mattiello J, LeBihan D. Estimation of the effective self-diffusion tensor from the NMR spin echo. J Magn Reson B 1994; 103:247–254.
8. Toga AW, Clark KA, Thompson PM, et al. Mapping the human connectome. Neurosurgery 2012; 71:1–5.
9. Guye M, Parker GJ, Symms M, et al. Combined functional MRI and tractography to demonstrate the connectivity of the human primary motor cortex in vivo. Neuroimage 2003; 19:1349–1360.
10. Le Bihan D, Urayama S, Aso T, et al. Direct and fast detection of neuronal activation in the human brain with diffusion MRI. Proc Natl Acad Sci U S A 2006; 103:8263–8268.
11. Wessels T, Wessels C, Ellsiepen A, et al. Contribution of diffusion-weighted imaging in determination of stroke etiology. AJNR Am J Neuroradiol 2006; 27:35–39.
12. Wen HM, Lam WW, Rainer T, et al. Multiple acute cerebral infarcts on diffusion-weighted imaging and risk of recurrent stroke. Neurology 2004; 63:1317–1319.
13. Maier SE, Sun Y, Mulkern RV. Diffusion imaging of brain tumors. NMR Biomed 2010; 23:849–864.
14. Rovaris M, Gass A, Bammer R, et al. Diffusion MRI in multiple sclerosis. Neurology 2005; 65:1526–1532.
15. Atkinson-Clement C, Pinto S, Eusebio A, et al. Diffusion tensor imaging in Parkinson's disease: Review and meta-analysis. Neuroimage Clin 2017; 16:98–110.
16. Stebbins GT, Murphy CM. Diffusion tensor imaging in Alzheimer's disease and mild cognitive impairment. Behav Neurol 2009; 21:39–49.
17. Kumaran SP, Gupta K, Pushpa B, et al. Diffusion-weighted imaging: As the first diagnostic clue to Creutzfeldt Jacob disease. J Neurosci Rural Pract 2012; 3:408–410.
18. Wende T, Hoffmann KT, Meixensberger J. Tractography in neurosurgery: a systematic review of current application. J Neurol Surg A Cent Eur Neurosurg 2020; 81:442–455.
19. Witulla B, Goerig N, Putz F, et al. On PTV definition for glioblastoma based on fiber tracking of diffusion tensor imaging data. PLoS One 2020; 15:e0227146.
20. Vedantam A, Jirjis MB, Schmit BD, et al. Diffusion tensor imaging of the spinal cord: insights from animal and human studies. Neurosurg 2014; 74:1–8.
21. Maurer MH, Heverhagen JT. Diffusion weighted imaging of the prostate-principles, application, and advances. Transl Androl Urol 2017; 6:490–498.
22. Kwee TC, Takahara T, Ochiai R, et al. Diffusion-weighted whole-body imaging with background body signal suppression (DWIBS): features and potential applications in oncology. Eur Radiol 2008; 18:1937–1952.
23. Mori S, Crain BJ, Chacko VP, et al. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Ann Neurol 1999; 45:265–269.
24. Conturo TE, Lori NF, Cull TS, et al. Tracking neuronal fiber pathways in the living human brain. Proc Natl Acad Sci U S A 1999; 96:10422–10427.
25. Basser PJ, Pajevic S, Pierpaoli C, et al. *In vivo* fiber tractography using DT-MRI data. Magn Reson Med 2000; 44:625–632.
26. Nakata Y, Aoki S, Sato N, et al. Tract-specific analysis for investigation of Alzheimer disease: a brief review. Jpn J Radiol 2010; 28:494–501.
27. Shukla DK, Keehn B, Müller RA. Tract-specific analyses of diffusion tensor imaging show widespread white matter compromise in autism spectrum disorder. J Child Psychol Psychiatry 2011; 52:286–295.
28. Shemesh N, Jespersen SN, Alexander DC, et al. Conventions and nomenclature for double diffusion encoding NMR and MRI. Magn Reson Med 2016; 75:82–87.
29. Lawrenz M, Finsterbusch J. Mapping measures of microscopic diffusion anisotropy in human brain white matter in vivo with double-wave-vector diffusion-weighted imaging. Magn Reson Med 2015; 73:773–783.
30. Lawrenz M, Brassen S, Finsterbusch J. Microscopic diffusion anisotropy in the human brain: Age-related changes. Neuroimage 2016; 141:313–325.
31. Yang G, Tian Q, Leuze C, et al. Double diffusion encoding MRI for the clinic. Magn Reson Med 2018; 80:507–520.
32. Schachter M, Does MD, Anderson AW, et al. Measurements of restricted diffusion using an oscillating gradient spin-echo sequence. J Magn Reson 2000; 147:232–237.
33. Tanner JE, Stejskal EO. Restricted self-diffusion of protons in colloidal systems by the pulsed-gradient, spin-echo method. J Chem Phys 1968; 49:1768–1777.
34. Callaghan PT. PGSE-MASSEY, a sequence for overcoming phase instability in very-high-gradient spin-echo NMR. J Magn Reason 1990; 88:493–500.
35. Le Bihan D, Breton E, Lallemand D, et al. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. Radiology 1988; 168:497–505.
36. Jensen JH, Helpern JA, Ramani A, et al. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. Magn Reson Med 2005; 53:1432–1440.
37. King MD, Houseman J, Roussel SA, et al. q-Space imaging of the brain. Magn Reson Med 1994; 32:707–713.
38. Tuch DS, Reese TG, Wiegell MR, et al. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. Magn Reson Med 2002; 48:577–582.
39. Wiegell MR, Larsson HB, Wedeen VJ. Fiber crossing in human brain depicted with diffusion tensor MR imaging. Radiology 2000; 217:897–903.

40. Tuch DS, Reese TG, Wiegell MR, et al. Diffusion MRI of complex neural architecture. Neuron 2003; 40:885–895.
41. Behrens TE, Berg HJ, Jbabdi S, et al. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? Neuroimage 2007; 34:144–155.
42. Assaf Y, Freidlin RZ, Rohde GK, et al. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. Magn Reson Med 2004; 52:965–978.
43. Assaf Y, Basser PJ. Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. Neuroimage 2005; 27:48–58.
44. Assaf Y, Blumenfeld-Katzir T, Yovel Y, et al. AxCaliber: a method for measuring axon diameter distribution from diffusion MRI. Magn Reson Med 2008; 59:1347–1354.
45. Zhou W, Laidlaw DH. An analytical model of diffusion and exchange of water in white matter from diffusion-MRI and its application in measuring axon radii. Proceedings of the 17th Annual Meeting of ISMRM, Honolulu, 2009; 17:263.
46. Christiaens D, Veraart J, Cordero-Grande L, et al. On the need for bundle-specific microstructure kernels in diffusion MRI. Neuroimage 2020; 208:116460.
47. Barakovic M, Tax CMW, Rudrapatna U, et al. Resolving bundle-specific intra-axonal T2 values within a voxel using diffusion-relaxation tract-based estimation. Neuroimage 2021; 227:117617.
48. Alexander DC, Hubbard PL, Hall MG, et al. Orientationally invariant indices of axon diameter and density from diffusion MRI. Neuroimage 2010; 52:1374–1389.
49. Zhang H, Schneider T, Wheeler-Kingshott CA, et al. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. Neuroimage 2012; 61:1000–1016.
50. Tariq M, Schneider T, Alexander DC, et al. Bingham-NODDI: Mapping anisotropic orientation dispersion of neurites using diffusion MRI. Neuroimage 2016; 133:207–223.
51. Palombo M, Ianus A, Guerreri M, et al. SANDI: a compartment-based model for non-invasive apparent soma and neurite imaging by diffusion MRI. Neuroimage 2020; 215:116835.
52. Panagiotaki E, Walker-Samuel S, Siow B, et al. Noninvasive quantification of solid tumor microstructure using VERDICT MRI. Cancer Res 2014; 74:1902–1912.
53. Wedeen VJ, Hagmann P, Tseng WY, et al. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magn Reson Med 2005; 54:1377–1386.
54. Tuch DS. Q-ball imaging. Magn Reson Med 2004; 52:1358–1372.
55. Tournier JD, Calamante F, Connelly A. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. Neuroimage 2007; 35:1459–1472.
56. Yeh CH, Schmitt B, Le Bihan D, et al. Diffusion microscopist simulator: a general Monte Carlo simulation system for diffusion magnetic resonance imaging. PLoS One 2013; 8:e76626.
57. Tachibana Y, Duval T, Obata T. Monte Carlo simulator for diffusion-weighted imaging sequences. Magn Reson Med Sci. 2021; 20:222–226.
58. Wolberg J. Data analysis using the method of least squares: Extracting the most information from experiments. Heidelberg: Springer, 2006.
59. Gill PE, Murray W. Algorithms for the solution of the nonlinear least-squares problem. SIAM J Numer Anal 1978; 15:977–992.
60. Daducci A, Canales-Rodríguez EJ, Zhang H, et al. Accelerated Microstructure Imaging via Convex Optimization (AMICO) from diffusion MRI data. Neuroimage 2015; 105:32–44.
61. Jensen JH, Hu C, Helpern JA. Rapid data acquisition and post-processing for diffusional kurtosis imaging. Proceedings of the 17th Annual Meeting of ISMRM, Honolulu, 2009; 17:1403.
62. Masutani Y, Aoki S. Fast and robust estimation of diffusional kurtosis imaging (DKI) parameters by general closed-form expressions and their extensions. Magn Reson Med Sci 2014; 13:97–115.
63. Neher PF, Götz M, Norajitra T, et al. A Machine learning-based approach to fiber tractography using classifier voting. Proceedings of the 18th International Conference of MICCAI, Munich, 2015; 9349:45–52.
64. Reisert M, Kellner E, Dhital B, et al. Disentangling micro from mesostructure by diffusion MRI: a Bayesian approach. Neuroimage 2017; 147:964–975.
65. Golkov V, Dosovitskiy A, Sperl JI, et al. q-Space deep learning: twelve-fold shorter and model-free diffusion MRI scans. IEEE Trans Med Img 2016; 35:1344–1351.
66. Nedjati-Gilani GL, Schneider T, Hall MG, et al. Machine learning based compartment models with permeability for white matter microstructure imaging. Neuroimage 2017; 150:119–135.
67. Healy DM, Hendriks H, Kim PT. Spherical deconvolution. J Multivariate Anal 1998; 67:1–22.
68. Masutani Y, Hori M, Aoki S. Unstructured sampling and RBF-based ODF reconstruction in Q-space for diffusion MR tractography. Int J Comput Assist Radiol Surg 2015; 10:S192–S193.
69. Buhmann MD. Radial basis functions: theory and implementations. Cambridge:Cambridge University Press, 2003.
70. Masutani Y, Sasaki K. Generation and extension of training data for deep regression-based inference of various model parameters in diffusion MRI. Proceedings of Meeting on Artificial Intelligence Applied to Medical Images (MAIAMI) 2018, Tokyo, 2018.
71. Ye C, Cui Y, Li X. Q-space learning with synthesized training data. Proceeding of the CDMRI 2018, Shenzhen, 2018.
72. Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. Magn Reson Med 1995; 34:910–914.
73. Masutani Y. Noise level matching improves robustness of diffusion MRI parameter inference by synthetic q-space learning. Proceedings of the 16th International Symposium on Biomedical Imaging, Venice, 2019; 139–142.
74. Masutani Y, Aoki S, Abe O, et al. MR diffusion tensor imaging: recent advance and new techniques for diffusion tensor visualization. Eur J Radiol 2003; 46:53–66.
75. Fukunaga I, Hori M, Masutani Y, et al. Comparison of diffusion Magnetic Resonance Image Analyzer (diMaRIA) NODDI and AMICO NODDI using 2-shell dMRI data. Proceedings of the 47th Annual Meeting of JSMRM, Kumamoto, 2019.
76. Masutani Y. diffusion Magnetic Resonance Image Analyzer (diMaRIA) page. http://www.medimg.info.hiroshima-cu.ac.jp/diMaRIA/diMaRIA.htm (Accessed: Jan 1, 2021)

77. Westin CF, Knutsson H, Pasternak O, et al. Q-space trajectory imaging for multidimensional diffusion MRI of the human brain. Neuroimage 2016; 135:345–362.
78 Ning L, Bonet-Carne E, Grussu F, et al. Cross-scanner and cross-protocol multi-shell diffusion MRI data harmonization: Algorithms and results. Neuroimage 2020; 221:117128.
79. Nilsson M, van Westen D, Ståhlberg F, et al. The role of tissue microstructure and water exchange in biophysical modelling of diffusion in white matter. MAGMA 2013; 26:345–370.
80. Wells JA, Thomas DL, Saga T, et al. MRI of cerebral microvascular flow patterns: A multi-direction diffusion-weighted ASL approach. J Cereb Blood Flow Metab 2017; 37:2076–2083.
81. Plog BA, Nedergaard M. The glymphatic system in central nervous system health and disease: past, present, and future. Annu Rev Pathol 2018; 13:379–394.
82. Callaghan R, Alexander DC, Palombo M, et al. ConFiG: contextual fibre growth to generate realistic axonal packing for diffusion MRI simulation. Neuroimage 2020; 220:117107.
83. Ginsburger K, Poupon F, Beaujoin J, et al. Improving the realism of white matter numerical phantoms: a step toward a better understanding of the influence of structural disorders in diffusion MRI. Front Phys 2018; 6:12.
84. Lee HH, Jespersen SN, Fieremans E, et al. The impact of realistic axonal shape on axon diameter estimation using diffusion MRI. Neuroimage 2020; 223:117228.
85. Andersson M, Kjer HM, Rafael-Patino J, et al. Axon morphology is modulated by the local environment and impacts the non-invasive investigation of its structure-function relationship. Proc Natl Acad Sci U S A 2020; 117:33649–33659.
86. Palombo M, Alexander DC, Zhang H. A generative model of realistic brain cells with application to numerical simulation of the diffusion-weighted MR signal. Neuroimage 2019; 188:391–402.
87. Rensonnet G, Scherrer B, Girard G, et al. Towards microstructure fingerprinting: estimation of tissue properties from a dictionary of Monte Carlo diffusion MRI simulations. Neuroimage 2019; 184:964–980.
88. Fang C, Nguyen VD, Wassermann D, et al. Diffusion MRI simulation of realistic neurons with SpinDoctor and the Neuron Module. Neuroimage 2020; 222:117198.
89. Raghu M, Poole B, Kleinberg J, et al. On the expressive power of deep neural networks. Proceedings of the 34th International Conference on Machine Learning, PMLR 2017; 70:2847–2854.
90. Reeder SB. Measurement of signal-to-noise ratio and parallel imaging, In: Schoenberg SO, Dietrich O, Reiser MF eds Parallel Imaging in Clinical MR Applications. Heidelberg: Springer, 2007; 49–61.
91. Kidoh M, Shinoda K, Kitajima M, et al. Deep learning based noise reduction for brain MR imaging: tests on phantoms and healthy volunteers. Magn Reson Med Sci 2020; 19:195–206.
92. Chen G, Wu Y, Shen D, et al. Noise reduction in diffusion MRI using non-local self-similar information in joint x-q space. Med Image Anal 2019; 53:79–94.