

Microstructural parameter estimation *in vivo* using diffusion MRI and structured prior information

Jonathan D. Clayden^{1†‡}, Zoltan Nagy^{2,3‡}, Nikolaus Weiskopf², Daniel C. Alexander⁴ and Chris A. Clark¹

¹UCL Institute of Child Health, University College London, London, UK

²Wellcome Trust Centre for Neuroimaging, UCL Institute of Neurology, University College London, London, UK

³Laboratory for Social and Neural Systems Research, University of Zurich, Zurich, Switzerland

⁴Centre for Medical Image Computing, University College London, London, UK

[†]Corresponding author. E-mail: j.clayden@ucl.ac.uk

[‡]These authors contributed equally to this work

Purpose: Diffusion MRI has recently been used with detailed models to probe tissue microstructure. Much of this work has been performed *ex vivo* with powerful scanner hardware, to gain sensitivity to parameters such as axon radius. By contrast, performing microstructure imaging on clinical scanners is extremely challenging.

Methods: We use an optimised dual spin-echo diffusion protocol, and a Bayesian fitting approach, to obtain reproducible contrast (histogram overlap of up to 92%) in estimated maps of axon radius index in healthy adults at a modest, widely-available gradient strength (35 mT m⁻¹). A key innovation is the use of influential priors.

Results: We demonstrate that our priors can improve precision in axon radius estimates—a sevenfold reduction in voxelwise coefficient of variation *in vivo*—without significant bias. Our results may reflect true axon radius differences between white matter regions, but this interpretation should be treated with caution due to the complexity of the tissue relative to our model.

Conclusions: Some sensitivity to relatively large axons (3–15 μm) may be available at clinical field and gradient strengths. Future applications at higher gradient strength will benefit from the favourable eddy current properties of the dual spin-echo sequence, and greater precision available with suitable priors.

Introduction

Diffusion-weighted magnetic resonance imaging (dMRI) uses the random self-diffusion of water molecules as the basis for an endogenous contrast in biological tissues (Le Bihan, 2003). A greater diffusivity is associated with greater attenuation in the MR signal due to the dispersion of “labelled” molecules during the course of an imaging experiment. Diffusion tensor imaging takes advantage of the orientational dependence of diffusivity in tissue to infer the arrangement of structures such as white matter tracts (Basser et al., 1994), and varying the time during which molecules are allowed to diffuse allows further properties of tissue architecture to be inferred (Callaghan, 1991).

A recent trend has been to use dMRI in combination with detailed models of tissue microstructure to try to estimate characteristics which are generally more associated with invasive histology than clinical imaging, such as axon radius. The tissue models are typically built up from simple geometric shapes such as cylinders and

spheres, but despite their simplicity they may be able to provide more direct tissue microstructure parameters than can be obtained from traditional dMRI analysis (Alexander et al., 2010; Assaf & Basser, 2005; Stanisz et al., 1997). These model parameters may in turn offer greater interpretability and sensitivity as biomarkers. In addition to the pioneering work by Stanisz et al., the “AxCaliber” technique has demonstrated the feasibility of recovering axon radius information from MR images of nervous tissue (Assaf et al., 2008). Subsequently, the “ActiveAx” approach has developed the area towards feasibility *in vivo* by using orientationally invariant protocols, to allow axon radii to be estimated throughout the brain, and optimised pulses sequences, to make acquisition times feasible (Alexander et al., 2010). The strength of the magnetic gradients available has been shown to be a key limiting factor for these applications (Dyrby et al., 2013; Huang et al., 2015), and the lack of strong gradients at most sites is a major barrier to their widespread uptake.

The best choice of diffusion-weighted pulse se-

quence for these applications is the focus of ongoing discussion in the literature. In addition to the original pulsed-gradient spin-echo (PGSE) sequence developed by [Stejskal & Tanner \(1965\)](#), other diffusion-weighted sequences have been suggested as candidates for more effective tissue microstructure imaging. For example, oscillating gradient sequences allow very short diffusion times to be achieved, and therefore may have better sensitivity to small axon radii ([Li et al., 2014](#); [Siow et al., 2013](#)). Multiple wave-vector protocols can help to distinguish between signals from compartments with different shapes ([Shemesh & Cohen, 2011](#)), and several authors have proposed that they may provide additional sensitivity, beyond that of PGSE, for axon radius estimation ([Koch & Finsterbusch, 2008](#); [Lawrenz & Finsterbusch, 2010](#); [Shemesh et al., 2009](#)). One can even use a generalised gradient waveform, enforcing only realistic slew rates and balance to ensure that refocussing and a main echo occur ([Drobnjak et al., 2010](#); [Siow et al., 2012](#)).

For routine diffusion-weighted imaging on clinical scanners, the dual spin-echo sequence (DSE; [Feinberg & Jakab, 1990](#); [Reese et al., 2003](#)) is very popular because it reduces eddy currents at the time of readout, and hence the image distortions caused by them. It differs from the standard Stejskal–Tanner sequence in that refocussing is applied twice, with four diffusion-sensitising gradients appearing around the refocussing pulses (see Fig. 1). It may offer additional benefits for microstructural imaging in terms of sensitivity, due to its allowance for gradient pulses placed next to each other, which produces relatively short effective diffusion times. (With PGSE on clinical hardware, by contrast, a lower bound is imposed by the time required for the radio-frequency refocussing pulse.) The trade-off is that a longer echo time is needed for DSE, reducing the available signal. We have previously derived an expression for the signal expected from this sequence within impermeable cylinders ([Clayden et al., 2009](#)), using the Gaussian phase distribution approximation ([Price, 1997](#)). We further adapted the experimental design optimisation from [Alexander \(2008\)](#) for the DSE sequence using this signal model, and demonstrated the sequence’s potential advantages for estimating small axon radii, in particular, using simulations ([Clayden et al., 2009](#)).

In this study our aim is to investigate whether meaningful axon radius information can be obtained in practice using standard scanner hardware and a widely-available pulse sequence. Specifically, we apply an optimised DSE sequence on a standard 3 T clinical scanner, using a maximum gradient strength of 35 mT m⁻¹. We consider the corpus callosum in the human brain, where the distribution of axon radii is well characterised by post-mortem histology ([Aboitiz et al., 1992](#)). This structure has been well studied in the broader imaging literature, since damage to it has been shown to have a role in a number of disease processes. However, microstruc-

tural measurements at the modest gradient strength applied here are sensitive to radii only at the very upper limit of those observed in human callosal tissue. To improve precision we use a standard model together with a new parameter estimation algorithm that incorporates prior information about plausible radii. We use simulations to demonstrate that the algorithm provides contrast between large and small radii, without substantial bias, under idealised conditions. In brain data acquired from adult volunteers, axon radius index maps consistently indicate the presence of large axons in the same regions suggested by histology, in particular in the anterior mid-body of the corpus callosum.

Methods

We begin by outlining our signal model, and the processes we applied for optimising the DSE sequence, acquiring data and fitting tissue model parameters.

Signal model

The diffusion-weighted signal in white matter, S , is modelled as a weighted sum of signal contributions from three compartments: an isotropic compartment representing cerebrospinal fluid contamination, a restricted “intracellular” compartment, and a hindered “extracellular” compartment (cf. [Alexander et al., 2010](#); [Barazany et al., 2009](#)). We denote the isotropic volume fraction with f_i and the restricted volume fraction with f_r , subject to $0 \leq f_i + f_r \leq 1$. Then,

$$\frac{S(\Theta, f_i, f_r)}{S_0} = f_i S_i(\Theta) + f_r S_r(\Theta) + (1 - f_i - f_r) S_h(\Theta), \quad (1)$$

where S_i , S_r and S_h are the signals from the isotropic, restricted and hindered compartments respectively, S_0 is the signal without diffusion weighting, and Θ is a set of additional parameters. The isotropic signal component is a simple function of the standard diffusion b -value and the diffusivity, D_i , of free water, viz. $S_i = \exp(-bD_i)$.

The white matter tissue of interest is modelled as a coherent bundle of parallel, impermeable, hollow cylinders of fixed radius, R . The extracellular, hindered compartment is assumed to be homogeneous, and diffusion is assumed to follow a cylindrically symmetric 3D Gaussian distribution, viz.

$$S_h(\Theta) = \exp\left(-b\left(\cos^2 \alpha (D_{\parallel} - D_{\perp}) + D_{\perp}\right)\right), \quad (2)$$

where α is the angle between the gradient direction and the orientation of the white matter bundle, D_{\parallel} is the diffusivity parallel to the cylinders, and D_{\perp} is the diffusivity perpendicular to them.

Following [Neuman \(1974\)](#) and [van Gelderen et al. \(1994\)](#), we have previously derived an expression for

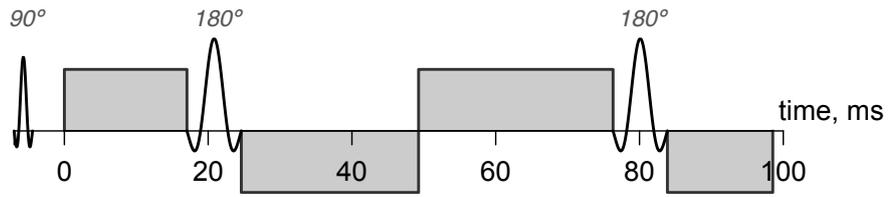


Figure 1: Pulse timing diagram for the standard dual spin-echo sequence. Only RF pulses and the four diffusion-weighting gradient pulses are shown for simplicity. Time zero is the earliest time at which a diffusion gradient can first be applied, allowing for the time required for the 90° excitation RF pulse and other preparatory gradient pulses.

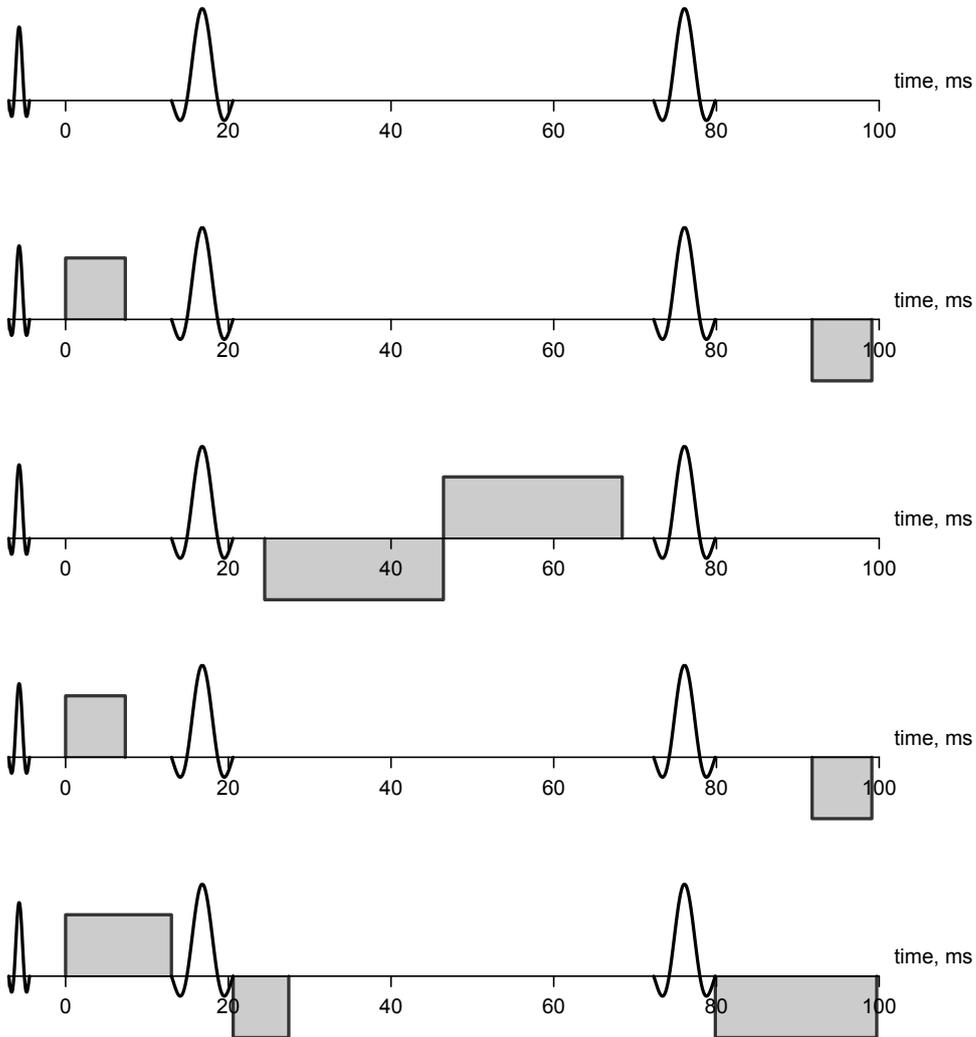


Figure 2: Optimised series of five pulse arrangements, with b -values of 0, 422, 620, 422 and 2378 s mm^{-2} . Note that all four standard pulses do not actually exist in any one arrangement, but all are properly balanced. The second and fourth arrangements are in fact identical.

	Pulse arrangement				
	1	2	3	4	5
Onset 1		0.00		0.00	0.00
Length 1		7.34		7.34	13.01
Onset 2			24.47		20.60
Length 2			21.97		6.83
Onset 3			46.45		
Length 3			21.97		
Onset 4		91.76		91.76	79.86
Length 4		7.34		7.34	19.84
<i>b</i>-value, s mm⁻²	0	422	620	422	2378

the restricted diffusion signal within cylinders for the DSE sequence (Clayden et al., 2009), using the Gaussian phase distribution approximation (Price, 1997). That result is used here for S_r , unmodified. It depends on the cylinder radius, R ; the intracellular diffusivity, which we take as equal to $D_{||}$; and the orientation of the cylinder, which we parameterise using the spherical coordinate angles, θ and ϕ .

The full parameter set is therefore $\Phi = \{S_0, f_i, f_r, D_i, D_{||}, D_{\perp}, R, \theta, \phi\}$.

We note that this model does not take into account differences in T_1 or T_2 relaxation times in the different compartments, whereas in practice such differences will exist. However, the isotropic volume fraction is expected to be very small in most voxels, and so the influence of this limitation on parameter estimates will be minimal. Moreover, f_i and D_i are treated as nuisance parameters of little interest in this study.

Sequence optimisation

The experiment design optimisation framework developed by Alexander (2008) was used to optimise a DSE imaging protocol for estimating the tissue parameters of interest. This framework aims to identify combinations of pulse arrangements, within the constraints imposed by the sequence and the performance of the scanner, which will maximise the expected precision of the tissue parameters, using the formalism of the Cramér–Rao lower bound.

The generative parameters used for the optimisation were: $S_0 = 1$, $f_i = 0$, $f_r = 0.7$, $D_i = 3 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$, $D_{||} = 1.7 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$ and $D_{\perp} = 1.2 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$, following Clayden et al. (2009). Generative radii were $R \in \{5, 10, 20\} \mu\text{m}$, as in Alexander et al. (2010), with the optimisation seeking a combination of pulse arrangements that jointly minimise average expected variance in the model parameters across these three values. (Although large relative to the expected radii in most tissue, they represent the domain where we expect some sensitivity at low gradient strength.) No assumptions were made about the orientations of the axon bundles within

Table 1: Pulse timings, in milliseconds, as implemented in the final protocol shown in Fig. 2. Time zero is the earliest time at which a diffusion gradient can first be applied, allowing for the time required for the 90° excitation RF pulse and other preparatory gradient pulses. Missing values correspond to omitted pulses. Echo time is 118.54 ms in all five pulse arrangements. The time required for a 90° RF pulse was 2.56 ms, and 7.56 ms was required for a 180° pulse and its associated crusher gradients.

the voxel, and gradient directions for each arrangement were therefore uniformly spread over the sphere. For the purposes of estimating the noise properties of the protocol, the spin–spin relaxation constant, T_2 , was taken to be 0.07 s. The number of separate diffusion-weighted pulse arrangements was fixed to four, and the number of gradient directions per arrangement was fixed to 90. A $b = 0$ arrangement was also included.

Eddy currents with a time constant of $0.7/\tilde{T}$ were nulled, with \tilde{T} the maximal sum of all diffusion-encoding gradient pulse lengths across the five arrangements, as proposed by Heid (2000). Off-design eddy current effects are also reduced by this process, which removes one degree of freedom from the optimisation.

The optimised pulse arrangements are shown in Fig. 2, and precise gradient timings are given in Table 1. The diffusion b -values corresponding to the five pulse arrangements were 0, 422, 620, 422 and 2378 s mm^{-2} . Gradient amplitude in each case, except where $b = 0$, was the maximum allowed, at 35 mT m^{-1} . The echo time was included in the optimisation but not allowed to vary across arrangements; its final value was 118.54 ms. In general, the DSE sequence has no specific diffusion time associated with it, but for the simple arrangements in Fig. 2, we can compute an effective diffusion time in a similar way to the PGSE sequence. On this basis, arrangements 2, 4 and 5 have long diffusion times (80–92 ms), while arrangement 3 has a short diffusion time (22 ms).

Synthetic data

Synthetic data were obtained using Monte Carlo simulation, as implemented in Camino (Cook et al., 2006; Hall & Alexander, 2009). The simulation tracks the phase of 10,000 spins over 1000 time steps during the optimised pulse arrangement, to calculate the final signal. The simulated tissue geometry consisted of hexagonally packed impermeable cylinders of fixed radius, with a universal diffusivity of $1.7 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$. Run time was approximately five minutes on a standard iMac desktop computer.

The simulation was carried out for axon radii of 1, 3, 5, 10, 15 and 20 μm , with the centre-to-centre cylinder separation fixed at 2.3 times the radius, to maintain an intracellular volume fraction of 0.69 in all cases. Ten different axon orientations were used for each radius, equally distributed on the sphere. Rician noise was added to the simulated signals, based on an SNR of 19 at $b = 0$ with the optimised echo time, matching the estimated noise characteristics of the scanner.

In addition, synthetic substrates containing a range of axon radii, distributed according to gamma distributions with means of 3, 5 and 8 μm and variances of 2.5 μm were also generated and fitted using our model. In this case cylinders were placed at random, but intracellular volume fractions were always between 0.65 and 0.70. This experiment was intended to test the influence of a particular type of mismatch between the data and the model.

Finally, to investigate axon radius estimation at different gradient strengths, the protocol was reoptimised using a maximum gradient strength of 300 mT m^{-1} , in line with the top end of what is currently available on human scanners (McNab et al., 2013). In this case synthetic data were generated using a single axon radius of 3 μm .

Data acquisition and preprocessing

Although the gradient strength capabilities of the scanner, as well as the time needed for RF pulses, preparation pulses and readout are taken into account by the optimisation, it does not incorporate slew rate information. It was therefore necessary to remove some very short pulses from the optimised arrangements to make them realisable in practice. The other pulse lengths were adjusted as necessary to restore balance, and the pulse diagrams in Fig. 2 incorporate these edits.

Three individuals—a 23 year-old female, a 32 year-old male, and a 31 year-old female—were each scanned on two separate occasions on a Siemens Trio 3 T clinical scanner, using a body transmit coil and vendor-supplied 32 channel receive-only head coil, as well as a standard gradient coil set ($G_{\text{max}} = 35 \text{ mT m}^{-1}$). The optimised protocol was applied, consisting of ten $b = 0$ images followed by sets of 90 identical diffusion sensitising directions for each of pulse arrangements 2–5, shown in Fig. 2. Data were acquired from a series of contiguous axial slices covering the corpus callosum, at 2.3 mm isotropic resolution. Scan time was approximately one hour, but cardiac pulse triggering was used, and so the exact time depended on each subject’s heart rate. Three slices were imaged per heartbeat.

DICOM files were converted to NIfTI format using the TractoR software package (Clayden et al., 2011). The first $b = 0$ volume was treated as a dummy and removed from each data set. Due to the favourable eddy current properties of the DSE sequence, and a high de-

gree of cooperation from scan subjects, the image volumes were observed to be well aligned. Therefore no coregistration was performed, to avoid introducing spurious motion and blurring in the data. Diffusion tensors were fitted to the data using ordinary least squares and a mask of the corpus callosum was drawn by hand by a single observer (JDC) on one midsagittal slice in each data set, using the fractional anisotropy map as a guide. Parameter fits were then performed within this mask.

Parameter fitting

Equation (1) gives an expression for the expected signal from our simple tissue substrate given known tissue parameters. Given a set of measured signal values, we need to solve the inverse problem, finding the set of parameters, $\hat{\Phi}$, which best explain the measurements. Since we have some expectations regarding the regime of values for many of the parameters, we would also like to make use of that information. We therefore take a Bayesian approach, estimating a posterior distribution over the parameters using Markov chain Monte Carlo (MCMC).

We assume that measurements of the signal, $x_{(k)}$, are drawn from a Rician distribution around the modelled value from Eq. (1), $S_{(k)}$, viz.

$$P(x_{(k)} | \Phi, \sigma) = \frac{x_{(k)}}{\sigma^2} \exp\left(-\frac{x_{(k)}^2 + S_{(k)}^2}{2\sigma^2}\right) I_0\left(\frac{x_{(k)}S_{(k)}}{\sigma^2}\right), \quad (3)$$

where σ controls the noise level. $I_0(\cdot)$ is the modified Bessel function of the first kind, order zero. If the noise associated with each measurement can be considered to be independent and identically distributed, then we can write down the joint distribution of the full set of measurements, $\mathbf{X} = (x_{(k)})$, as

$$P(\mathbf{X} | \Phi, \sigma) = \prod_k P(x_{(k)} | \Phi, \sigma). \quad (4)$$

Bayes’ rule then allows us to calculate a posterior distribution over the parameters as

$$P(\Phi, \sigma | \mathbf{X}) = \frac{P(\mathbf{X} | \Phi, \sigma)P(\Phi)P(\sigma)}{\iint P(\mathbf{X} | \Phi, \sigma)P(\Phi)P(\sigma) d\Phi d\sigma}, \quad (5)$$

where $P(\Phi)$ represents the prior information available regarding Φ , and $P(\sigma)$ likewise for σ . (We assume prior independence of Φ and σ .)

At this point we could simply choose “uninformative” priors for each parameter, ensuring only that physical requirements such as positivity are met by the estimates. However, estimating our tissue model parameters using data from a scanner with clinical gradient strengths is expected to yield relatively low precision, and by choosing more influential priors we regularise the problem. To this end, the following informative priors were used across all model fits:

$$\begin{aligned}
f_i &\sim \text{Beta}(1.2, 1.2), \\
\frac{f_r}{1-f_i} &\sim \text{Beta}(5, 5), \\
D_{\parallel} &\sim \text{Log-}\mathcal{N}(-20.69, 1), \\
D_{\perp} &\sim \text{Log-}\mathcal{N}(-21.04, 1), \\
R &\sim \text{Gamma}(3.562, 1.404 \times 10^{-6}). \quad (6)
\end{aligned}$$

These values are based on SI length and time units, i.e. metres and seconds. We use $f_r/(1-f_i)$ rather than f_r directly to give the parameter fixed bounds of 0 to 1. The two volume fractions have nonuniform distributions to regularise the posterior away from the extremes. The log-normal distribution has been used before as a prior for diffusivity parameters, for example by [Anderson \(2008\)](#), and the means of the distributions given above correspond to the values used for optimisation, i.e. $1.7 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$ for D_{\parallel} and $1.2 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$ for D_{\perp} . The gamma distribution—parameterised here using shape and scale parameters—has likewise been used to represent distributions over axon radii (e.g. [Assaf et al., 2008](#)), and our values are chosen to give the distribution a mean of $5 \mu\text{m}$, approximately 4–5 times the mean observed across the corpus callosum in histological studies such as [Aboitiz et al. \(1992\)](#), in line with previous findings of overestimation of the histological mean axon radius by the estimated axon radius index from dMRI ([Alexander et al., 2010](#); [Dyrby et al., 2013](#)).

Despite our use of informative prior distributions for the parameters described above, we do not explicitly impose any distributional assumptions on the posterior distributions. These are represented by the empirical distributions of the values sampled from the MCMC algorithm.

A series of samples were drawn from the posterior distribution, Eq. (5), using a blocked Metropolis–Hastings algorithm, with a tuned multivariate Gaussian proposal distribution for each block. The parameter blocks in this case were (S_0) , (f_i, f_r) , $(D_{\parallel}, D_{\perp}, R)$, (θ, ϕ) and (σ) . The isotropic diffusivity, D_i , was fixed at $3.0 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$ to represent free water, and not sampled. For the synthetic data only, f_i was also fixed at zero.

A least-squares tensor fit was first performed for each voxel, yielding principal diffusivities $\lambda_1 \geq \lambda_2 \geq \lambda_3$. MCMC chains were then initialised according to

$$\begin{aligned}
\tilde{f}_i &= \frac{\lambda_3^2}{\lambda_1 \lambda_2} & \tilde{f}_r &= 0.7(1 - \tilde{f}_i) \\
\tilde{D}_{\parallel} &= \lambda_1 & \tilde{D}_{\perp} &= \frac{\lambda_2 + \lambda_3}{2} & \tilde{R} &= 5 \mu\text{m}. \quad (7)
\end{aligned}$$

An iterative process was then used to tune the covariance matrix for each proposal distribution, bringing

each block acceptance rate close to the theoretical optimum of around 23% so that the parameter space would be explored efficiently (cf. [Roberts et al., 1997](#)). Afterwards, the chain was run for a burn-in period of 50,000 steps, followed by a sampling phase gathering 50 samples of each parameter, with each sample separated by 100 steps. In practice, samples were actually of the logit or logarithm of parameters with bounds, to ensure that the sample space was infinite.

The algorithm was implemented on the TractoR platform in R and C++, using the “Rcpp” and “RcppArmadillo” libraries for the R statistical language and environment ([Clayden et al., 2011](#); [Eddelbuettel & François, 2011](#); [Eddelbuettel & Sanderson, 2014](#); [R Core Team, 2014](#)).

The consistency of the MCMC over multiple chains was assessed, and samples were checked visually in several voxels for evidence of good mixing and convergence to the stationary posterior distribution.

Histogram similarity

The histogram intersection measure introduced by [Swain & Ballard \(1991\)](#) was used to quantify the similarity of axon radius distributions across scans. This normalised measure is based on the observed probability density in each bin, j , of the two histograms, h_1 and h_2 . Specifically,

$$H = \frac{\sum_j \min\{h_1^j, h_2^j\}}{\sum_j h_1^j}. \quad (8)$$

Intuitively, H represents the proportion of the histogram which is common to both data sets, and it therefore has a range of 0 to 1. Since the denominator will evaluate to the inverse of the bin width, this measure is symmetric as long as the bins are the same in each histogram.

Results

Figure 3 shows histograms of sampled axon radii from the synthetic data. At the low end of the scale, precision is too poor to distinguish radii of 1 and $3 \mu\text{m}$ ($H = 0.84$, using bins of width $0.25 \mu\text{m}$); and at the high end, the influence of the prior forces underestimation of very large radii of above $15 \mu\text{m}$. However, radii of 3, 5, 10 and $15 \mu\text{m}$ are clearly discriminable, with sample medians closely approximating the true, generative radius in each of these cases ($H = 0.41$ between 3 and $5 \mu\text{m}$; $H = 0.12$ between 5 and $10 \mu\text{m}$; $H = 0.44$ between 10 and $15 \mu\text{m}$). When the synthetic substrate contains a gamma distribution of radii, as shown in the bottom part of the figure, there is a bias in the posterior distributions, but it remains possible to distinguish larger radii from smaller ones. Fig. 4 demonstrates that much greater precision is available at $3 \mu\text{m}$ when gradient strengths increase to

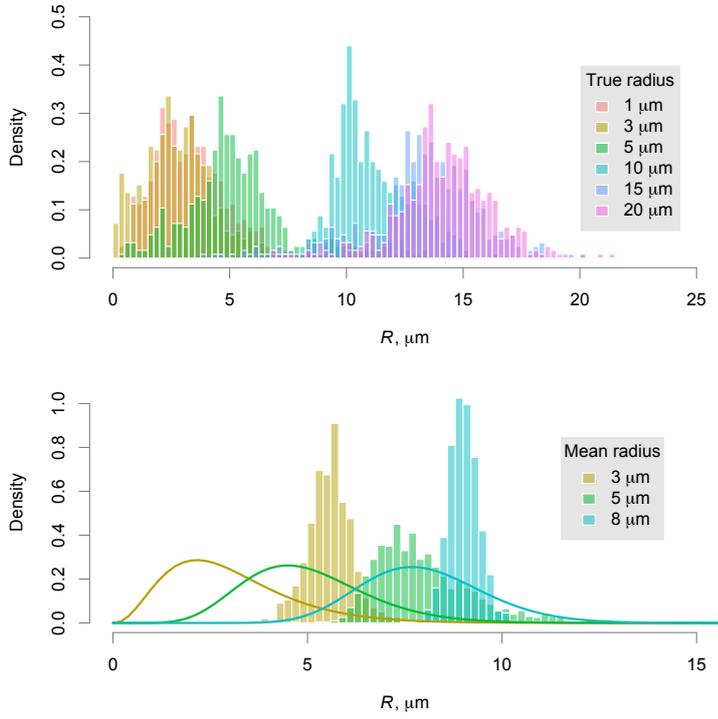


Figure 3: Histograms of sampled posterior axon radius values from our synthetic data, using a single fixed cylinder radius (top) or a gamma distribution of radii (bottom). For the former, samples cover ten axon orientations at each radius. For the latter, the generative distributions (shown) have means of 3, 5 and 8 μm , and variances of 2.5 μm .

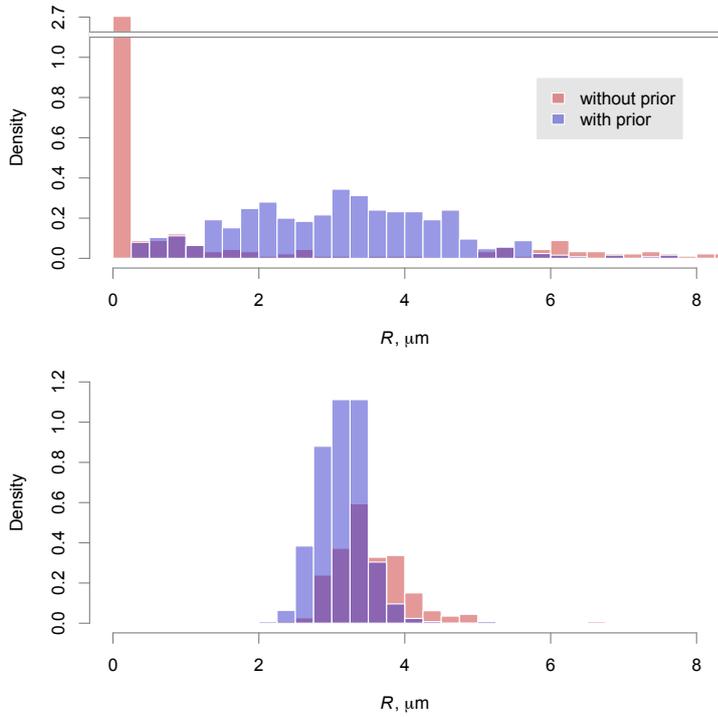


Figure 4: Histograms of sampled posterior axon radius from synthetic data, based on scan protocols optimised for maximum gradient strengths of 35 mT m^{-1} (top) and 300 mT m^{-1} (bottom). In each case, samples are shown from fits performed both with (blue) and without (red) applying the gamma prior distribution on R . The true generative axon radius was 3 μm , and samples cover ten axon orientations. Only samples below 8 μm are shown. Note that the y-axis in the upper plot is broken to accommodate a very tall bar.

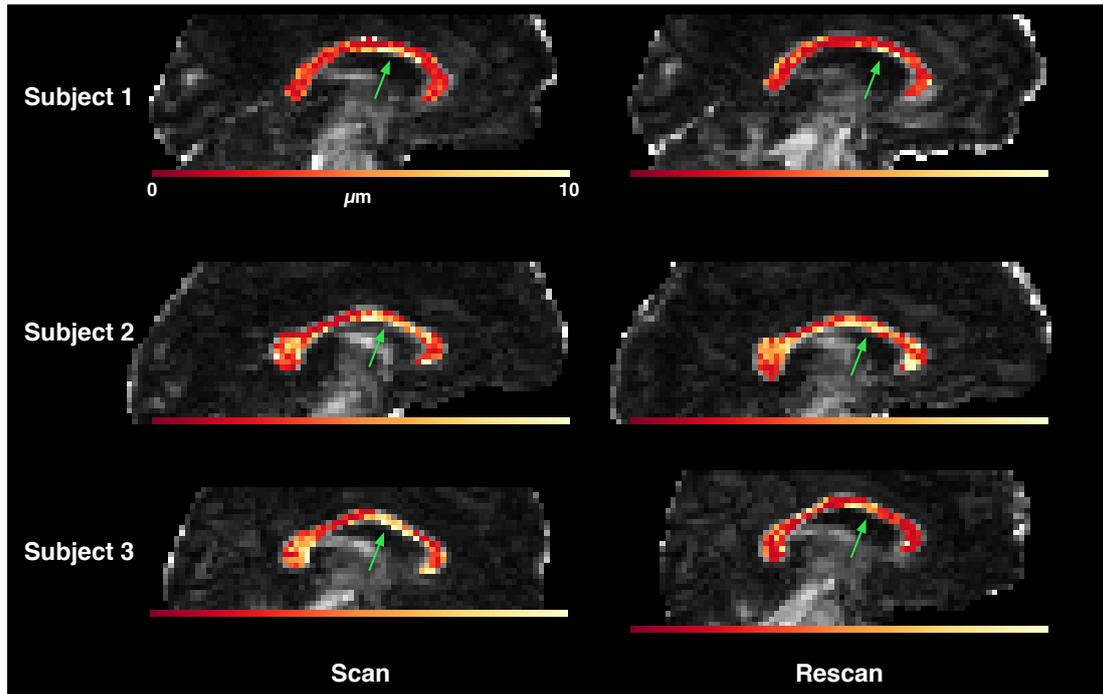


Figure 5: Maps of axon radius index, estimated as the median of the posterior sampled values in each voxel. An area of relatively large axons is consistently observed in the anterior mid-body of the corpus callosum (green arrows). Subjects are in rows, and scans in columns. The underlying greyscale map is fractional anisotropy, and the colour scale is the same in all subfigures.

300 mT m⁻¹, but it improves still further when prior information is used.

Maps of axon radius index, the median sampled value of R in each voxel, are shown for *in vivo* data in Fig. 5. A fair degree of consistency was observed, both between the first and second scans for each subject, and between subjects. The maximal radius index was generally observed in the anterior mid-body region and, in subjects 2 and 3, the splenium. Histograms across the whole corpus callosum showed substantial overlap between each pair of scans ($H = 0.92$ for subject 1, 0.89 for subject 2, and 0.76 for subject 3, using bins of width $0.5 \mu\text{m}$), and a similar level of overlap between subjects (H between 0.78 and 0.91).

To further examine variation along the corpus callosum, we divided each subject’s segmented region of interest into five subregions, separated by equally spaced coronal planes, similar to Aboitiz et al. (1992). Boxplots of sampled axon radius index in each subregion for each subject are shown in Fig. 6. The anterior mid-body subregion (shown in blue) has the highest radii on average in every subject, although the degree of variation across subregions is small in subject 1. The scan–rescan relative absolute difference, viz. $|a - b|/(a + b)$, in the median axon radius on a subregion-by-subregion basis varied from 11% (posterior mid-body) to 24% (anterior mid-body).

The model was a good fit to the acquired data, in terms of the proportion of total signal variance ex-

plained. Voxelwise coefficients of determination (R^2) averaged between 0.84 and 0.93 for each of the six scans.

To illustrate the influence of the priors used in our fit, Eq. (6), we show in Fig. 7 all of the sampled values for the axon radius parameter for one scan, both with and without the prior in use. (Priors on all other parameters were retained in the latter case.) We can observe that the prior is strongly influential, since there is an extremely wide range of sampled radii in its absence. The influence of the data, encapsulated in the Bayesian likelihood term, may therefore be considered to be relatively weak, which is to be expected at clinical gradient strengths. However, Fig. 8 shows that in regions of substantial probability mass with respect to f_r and D_{\perp} , the likelihood does show a substantial peak in the micron range of axon radius. It also illustrates that the prior influences the chains away from the very small radii observed without it, in Fig. 7.

Marginal posterior distributions for all other parameters with priors are shown in Fig. 9, and in each case there are substantial differences between the prior and posterior distributions, indicating the influence of the data.

Axon radius index maps obtained without use of the prior on R are shown in Supporting Fig. S1. Although some of the same areas of higher radius are still just about visible, the maps appear more noisy and demonstrate less consistency and smoothness. Reproducibility is also substantially lower in this case: in subject 1, for

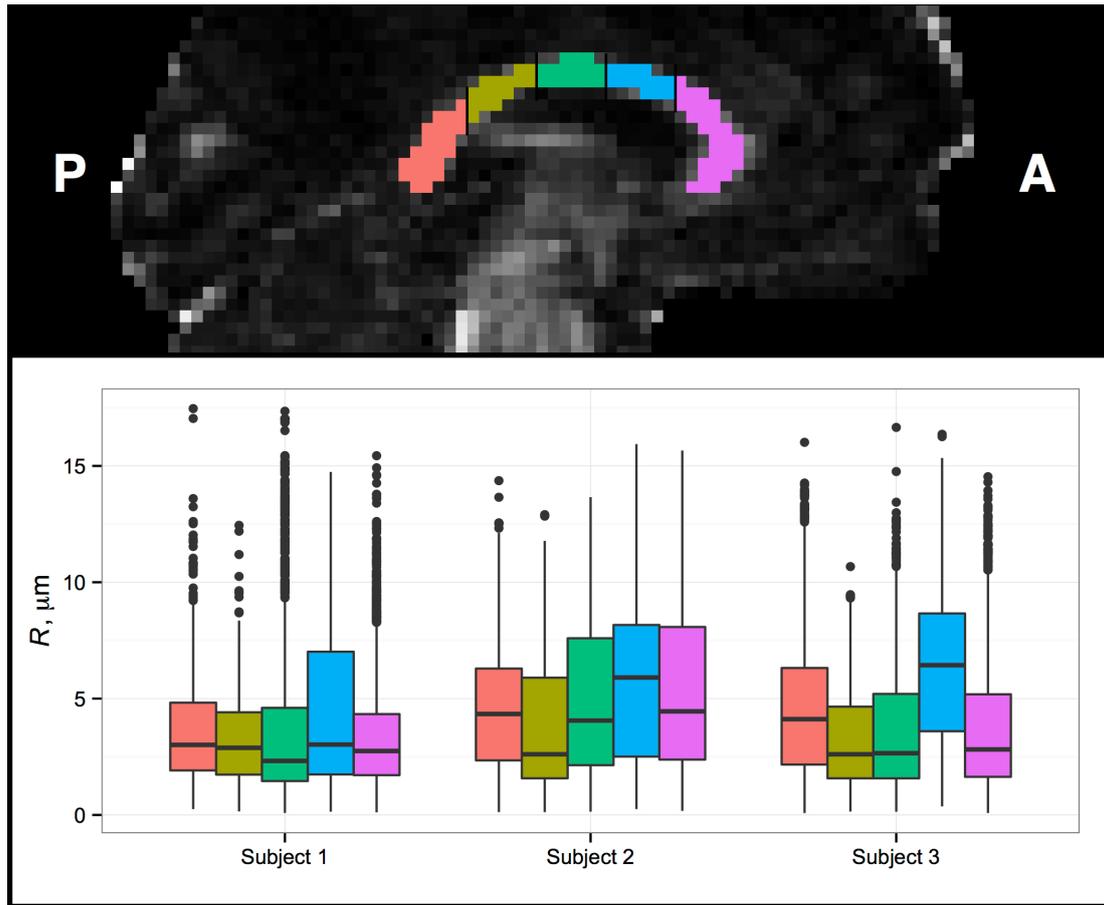


Figure 6: Boxplots of sampled posterior axon radius index in five subregions of the corpus callosum, across both scans, in each subject. In each case median axon radius index was highest in the anterior mid-body (blue). Filled circles represent outliers more than 1.5 interquartile ranges above the third quartile.

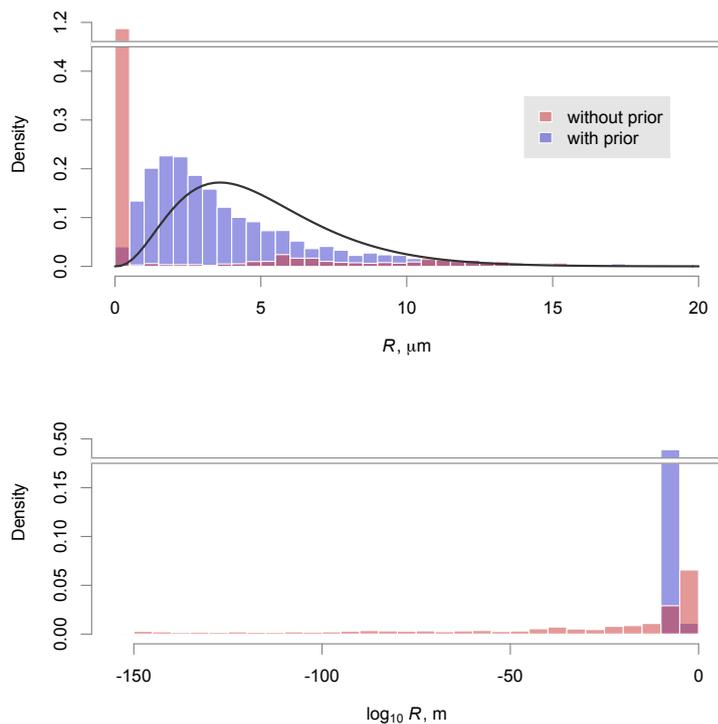


Figure 7: Histograms of sampled posterior axon radius index values across the entire corpus callosum, in the first scan of the first subject. Samples are shown from fits performed both with (blue) and without (red) applying the gamma prior distribution (black curve). The upper figure uses a standard x -axis and shows just those samples below $20 \mu\text{m}$, while the bottom figure shows all samples on a logarithmic x -axis. Note that both y -axes are broken to accommodate very tall bars.

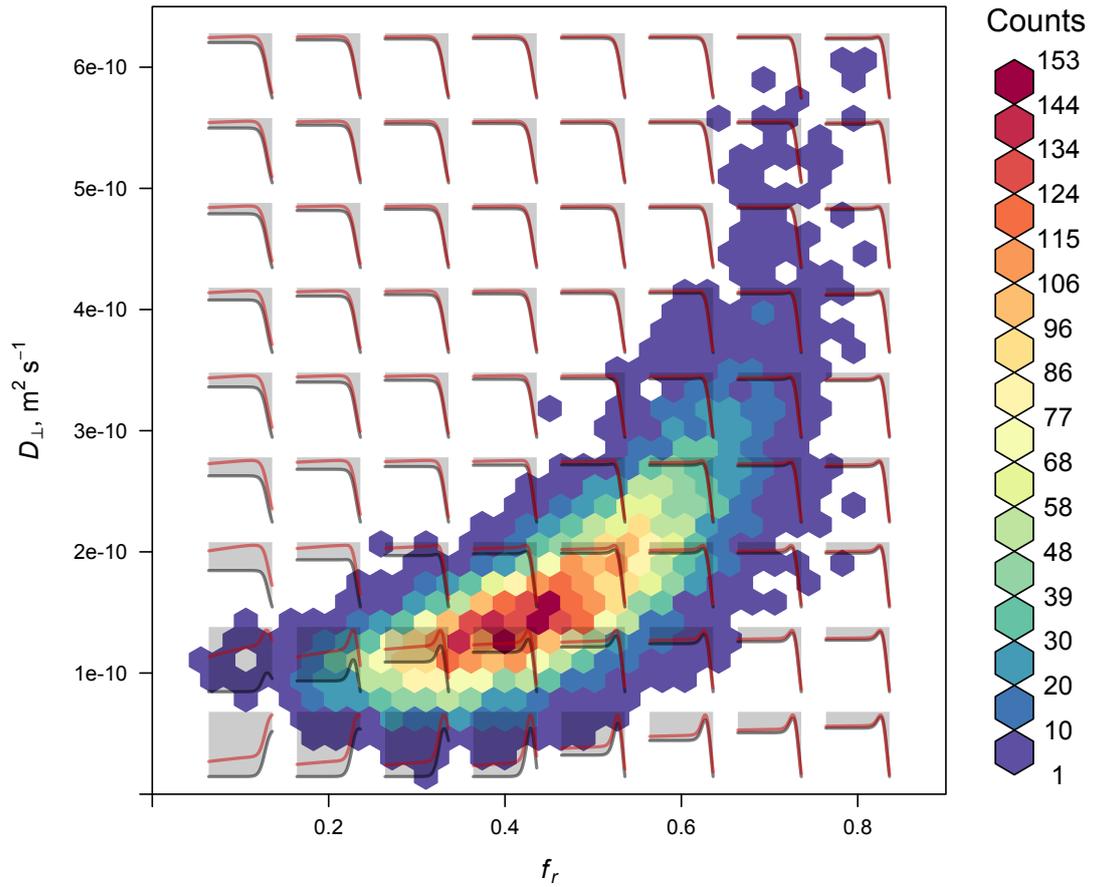


Figure 8: Joint histogram of the marginal posterior distribution over f_r and D_{\perp} in one voxel in the centre of the splenium in subject 2, scan 1, based on 10,000 samples. Overlaid are plots of log likelihood (black) and log posterior (red) as functions of log axon radius, over the range 0.02 to 20 μm , conditioned on the values of f_r and D_{\perp} corresponding the relevant location in the histogram. These overlaid plots are not to scale with one another, and are instead intended to show the shapes of the relevant functions.

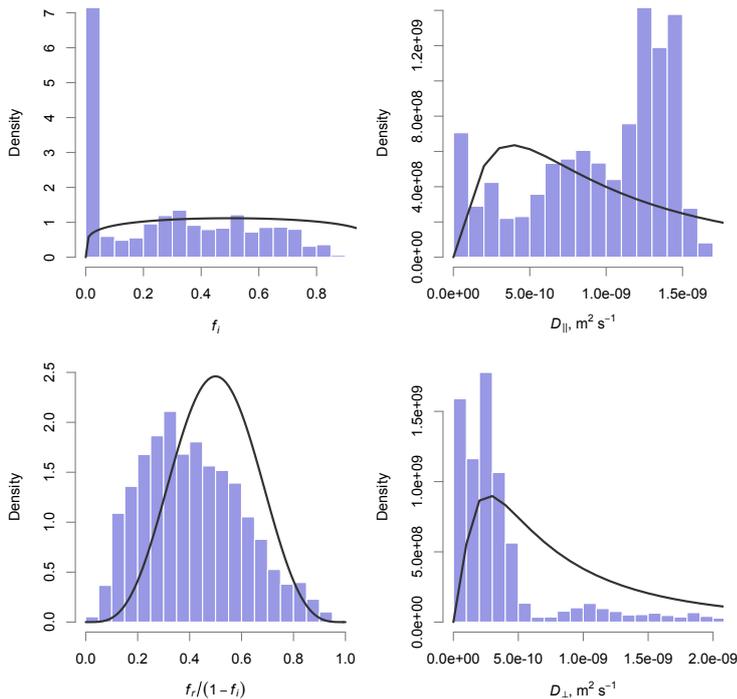


Figure 9: Histograms of sampled posterior values for each parameter with a prior distribution, except axon radius index which is shown in Fig. 7. The prior itself is shown as a black curve. The posterior distributions differ substantially from the priors in each case, demonstrating the influence of the data.

example, H decreases from 0.92 to 0.69. Voxelwise coefficients of variation averaged 311% without the prior, across all subjects, compared to 42% with the prior.

Discussion

Our key contribution in this paper has been to demonstrate that structured prior information can improve precision in microstructural parameter estimates. Using this platform, we have explored the application of the ActiveAx approach to *in vivo* axon radius imaging using the popular DSE pulse sequence on a clinical scanner. Such a standard MRI system is suboptimal for this kind of work, but commonly available, and our results inform the feasibility of “histological” imaging on typical hardware and with a standard pulse sequence, albeit one which has been optimised for the task. We used Bayesian MCMC simulation to estimate the tissue model parameters of interest, formally accounting for both prior expectations and the information available in the data.

Given the challenges associated with microstructure imaging on this platform, we found that it was necessary to regularise the problem by using informative prior distributions for several parameters, and particularly for the axon radius itself. We showed that this information improves precision at intermediate values of axon radius index (Fig. 7), without introducing significant bias when the assumption of a single radius holds (Fig. 3). Synthetic data based on a substrate containing a distribution of radii did lead to bias, most likely because of

the greater signal contribution from within wider cylinders. Nevertheless, it was still possible to tell distributions with different means apart *a posteriori*, with their ordering preserved.

Because of the strong influence of the priors, it cannot be said that we are truly estimating axon radius, since the range of acceptable values is effectively imposed by the prior. Moreover, we cannot demonstrate any improvement in accuracy, since ground-truth data are not available *in vivo*. Nevertheless, it is reasonable to interpret the relative differences from voxel to voxel as corresponding to meaningful variation across, in this case, the corpus callosum—and likewise, results are comparable across subjects since the priors are the same in each case. Scan–rescan consistency was seen to be reasonable, both visually and in terms of a quantitative histogram intersection measure. Subject 3 was the least consistent, possibly due to small within-volume movements during scanning or other external factors.

Our experiment with synthetic data suggested that our fitting process should provide sensitivity to axon radii of around 3–15 μm (Fig. 3). The contrast in the *in vivo* results may therefore reflect the detection of relatively large axon radii in certain parts of the corpus callosum—although our fits from substrates containing gamma distributions of radii suggest that while qualitatively meaningful, our axon radius index values are likely to be overestimates. Nevertheless, while Alexander et al. (2010) and Dyrby et al. (2013) have suggested that only very weak sensitivity to axon radius index is available at clinical gradient strengths, we have shown that using suitable priors can ameliorate the situation.

This advantage applies even at the upper end of what is currently achievable on human systems (cf. Fig. 4, where precision is shown to be higher with the prior than without it). Moreover, since eddy current distortions are worse at higher gradient strengths, the favourable eddy current properties of the DSE sequence will also be particularly useful in this regime. The combination of the DSE sequence and informed fitting process may therefore be a powerful one for future microstructural imaging studies on the next generation of scanner hardware—although there are other practical challenges to overcome, such as signal loss due to the greater influence of concomitant fields [Setsompop et al. \(2013\)](#). It will also be important for future work to compare DSE against oscillating gradient sequences and other alternatives, in combination with suitable priors, to fully evaluate their relative merits.

Although the trends we observe suggest that our technique may act as a detector for voxels containing larger axons, the results must be interpreted with caution because of the simplicity of the tissue model used. The three compartment model described by Eq. (1) ignores a great deal of the complexity in real tissue, even in relatively coherent and homogeneous areas of white matter such as the corpus callosum. There has been recent work adding characteristics such as membrane permeability and fibre dispersion and crossing to similar models ([Nilsson et al., 2009](#); [Zhang et al., 2011a,b](#)), and distributions of axon radii were included in the Ax-Caliber model ([Assaf et al., 2008](#)). Other factors which may need to be taken into account include the possibility of different T_2 relaxation constants in different tissue compartments (cf. [Szafer et al., 1995](#); [Xu et al., 2011](#)). Hence, further investigation using more complex models will be required to clarify the source of the contrast observed in our axon radius index maps (Fig. 5). There are also some limitations to our sequence parameterisation, such as the assumption of rectangular—rather than trapezoidal—pulses, which were not addressed so as to keep the theoretical signal model analytically tractable. The latter may, however, have only a small influence ([İanuş et al., 2013](#)). Finally, we optimised the sequence for axon radii larger than those actually expected in tissue, to make the optimisation stable—although it is unlikely that sensitivity to smaller radii could be improved substantially using the same scanner hardware.

In conclusion, we have found that some sensitivity to relatively large axons may be available using the standard DSE pulse sequence at clinical field and gradient strengths, if the scan is well set up and parameter estimation is performed carefully, with the use of suitable prior information. We have demonstrated a consistent trend in axon radius parameter maps which is broadly in keeping with known tissue characteristics, although caution is required in the interpretation of this finding. It may be that contrast arises from large axons of radius around $3 \mu\text{m}$ and above, although it may also arise from

variations in microscopic or macroscopic fibre dispersion (cf. [Nilsson et al., 2012](#)), axonal undulation ([Dyrby et al., 2014](#)), or from differences in other tissue properties such as membrane permeability. *In vivo* axon radius imaging on clinical scanners should therefore be treated cautiously at present, but as stronger gradients become available on these scanners—a tendency which is beginning to become reality—results will improve, and inferring tissue characteristics should become more practical.

Acknowledgments

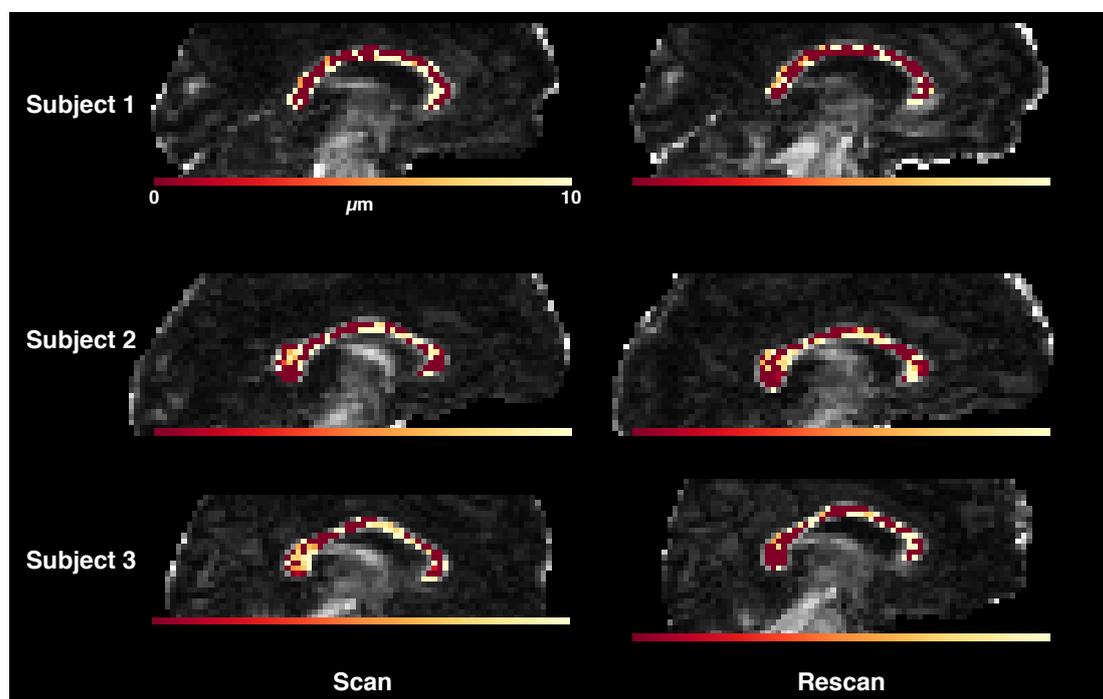
This work was supported by EPSRC grants EP/C536851/1 and EP/E007748. The Wellcome Trust Centre for Neuroimaging is supported by Wellcome Trust grant 079866/Z/06/Z. The authors are grateful to Dr Martin King and Dr Matt Hall for useful discussions on aspects of this work.

References

- Aboitiz F., Scheibel A.B., Fisher R.S. & Zaidel E. (1992). Fiber composition of the human corpus callosum. *Brain Research* **598**(1-2):143–153.
- Alexander D.C. (2008). A general framework for experiment design in diffusion MRI and its application to measuring direct tissue-microstructure features. *Magnetic Resonance in Medicine* **60**(2):439–448.
- Alexander D.C., Hubbard P.L., Hall M.G., Moore E.A., Ptito M., Parker G.J. & Dyrby T.B. (2010). Orientationally invariant indices of axon diameter and density from diffusion MRI. *NeuroImage* **52**(4):1374–1389.
- Andersson J.L.R. (2008). Maximum a posteriori estimation of diffusion tensor parameters using a Rician noise model: Why, how and but. *NeuroImage* **42**(4):1340–1356.
- Assaf Y. & Basser P.J. (2005). Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. *NeuroImage* **27**(1):48–58.
- Assaf Y., Blumenfeld-Katzir T., Yovel Y. & Basser P.J. (2008). Axcaliber: A method for measuring axon diameter distribution from diffusion MRI. *Magnetic Resonance in Medicine* **59**(6):1347–1354.
- Barazany D., Basser P.J. & Assaf Y. (2009). In vivo measurement of axon diameter distribution in the corpus callosum of rat brain. *Brain* **132**(5):1210–1220.
- Basser P.J., Mattiello J. & Le Bihan D. (1994). Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance, Series B* **103**(3):247–254.

- Callaghan P.T. (1991). *Principles of nuclear magnetic resonance microscopy*. Oxford University Press, UK.
- Clayden J.D., Muñoz Maniega S., Storkey A.J., King M.D., Bastin M.E. & Clark C.A. (2011). TractoR: Magnetic resonance imaging and tractography with R. *Journal of Statistical Software* **44**(8):1–18.
- Clayden J.D., Nagy Z., Hall M.G., Clark C.A. & Alexander D.C. (2009). Active imaging with dual spin-echo diffusion MRI. In J.L. Prince, D.L. Pham & K.J. Myers (Eds.), *Information Processing in Medical Imaging*, vol. 5636 of *Lecture Notes in Computer Science*, pp. 264–275. Springer-Verlag.
- Cook P.A., Bai Y., Nedjati-Gilani S., Seunarine K.K., Hall M.G., Parker G.J.M. & Alexander D.C. (2006). Camino: Open-source diffusion-MRI reconstruction and processing. In *Proceedings of the ISMRM 14th Scientific Meeting & Exhibition*, p. 2759. International Society for Magnetic Resonance in Medicine.
- Drobnjak I., Siow B. & Alexander D.C. (2010). Optimizing gradient waveforms for microstructure sensitivity in diffusion-weighted MR. *Journal of Magnetic Resonance* **206**(1):41–51.
- Dyrby T.B., Burke M., Alexander D.C. & Ptito M. (2014). Undulating and crossing axons in the corpus callosum may explain the overestimation of axon diameters with ActiveAx. In *Proceedings of the ISMRM-ESMRMB Joint Annual Meeting*, p. 2619. International Society for Magnetic Resonance in Medicine.
- Dyrby T.B., Søgaard L.V., Hall M.G., Ptito M. & Alexander D.C. (2013). Contrast and stability of the axon diameter index from microstructure imaging with diffusion MRI. *Magnetic Resonance in Medicine* **70**(3):711–721.
- Eddelbuettel D. & François R. (2011). Rcpp: Seamless R and C++ Integration. *Journal of Statistical Software* **40**(8):1–18.
- Eddelbuettel D. & Sanderson C. (2014). RcppArmadillo: Accelerating R with high-performance C++ linear algebra. *Computational Statistics and Data Analysis* **71**:1054–1063.
- Feinberg D.A. & Jakob P.D. (1990). Tissue perfusion in humans studied by Fourier velocity distribution, line scan, and echo-planar imaging. *Magnetic Resonance in Medicine* **16**(2):280–293.
- Hall M.G. & Alexander D.C. (2009). Convergence and parameter choice for Monte-Carlo simulations of diffusion MRI. *IEEE Transactions on Medical Imaging* **28**(9):1354–1364.
- Heid O. (2000). Eddy current-nulled diffusion weighting. In *Proceedings of the ISMRM 8th Scientific Meeting & Exhibition*, p. 799. International Society for Magnetic Resonance in Medicine.
- Huang S.Y., Nummenmaa A., Witzel T., Duval T., Cohen-Adad J., Wald L.L. & McNab J.A. (2015). The impact of gradient strength on in vivo diffusion MRI estimates of axon diameter. *NeuroImage* **106**:464–472.
- Ianuș A., Siow B., Drobnjak I., Zhang H. & Alexander D.C. (2013). Gaussian phase distribution approximations for oscillating gradient spin echo diffusion MRI. *Journal of Magnetic Resonance* **227**:25–34.
- Koch M.A. & Finsterbusch J. (2008). Compartment size estimation with double wave vector diffusion-weighted imaging. *Magnetic Resonance in Medicine* **60**(1):90–101.
- Lawrenz M. & Finsterbusch J. (2010). Double-wave-vector diffusion-weighting experiments with multiple concatenations at long mixing times. *Journal of Magnetic Resonance* **206**(1):112–119.
- Le Bihan D. (2003). Looking into the functional architecture of the brain with diffusion MRI. *Nature Reviews Neuroscience* **4**(6):469–480.
- Li H., Gore J.C. & Xu J. (2014). Fast and robust measurement of microstructural dimensions using temporal diffusion spectroscopy. *Journal of Magnetic Resonance* **242**:4–9.
- McNab J.A., Edlow B.L., Witzel T., Huang S.Y., Bhat H., Heberlein K., Feiweier T., Liu K., Keil B., Cohen-Adad J., Tisdall M.D., Folkerth R.D., Kinney H.C. & Wald L.L. (2013). The Human Connectome Project and beyond: initial applications of 300 mT/m gradients. *NeuroImage* **80**:234–245.
- Neuman C.H. (1974). Spin echo of spins diffusing in a bounded medium. *Journal of Chemical Physics* **60**(11):4508–4511.
- Nilsson M., Lätt J., Nordh E., Wirestam R., Ståhlberg F. & Brockstedt S. (2009). On the effects of a varied diffusion time in vivo: Is the diffusion in white matter restricted? *Magnetic Resonance Imaging* **27**:176–187.
- Nilsson M., Lätt J., Ståhlberg F., van Westen D. & Hagglätt H. (2012). The importance of axonal undulation in diffusion MR measurements: A Monte Carlo simulation study. *NMR in Biomedicine* **25**(5):795–805.
- Price W.S. (1997). Pulsed-field gradient nuclear magnetic resonance as a tool for studying translational diffusion: Part 1. Basic theory. *Concepts in Magnetic Resonance* **9A**(5):299–336.

- R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>. Accessed March 17, 2015.
- Reese T.G., Heid O., Weisskoff R.M. & Wedeen V.J. (2003). Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. *Magnetic Resonance in Medicine* **49**(1):177–182.
- Roberts G.O., Gelman A. & Gilks W.R. (1997). Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability* **7**(1):110–120.
- Setsompop K., Kimmlingen R., Eberlein E., Witzel T., Cohen-Adad J., McNab J., Keil B., Tisdall M., Hoecht P., Dietz P., Cauley S., Tountcheva V., Matschl V., Lenz V., Heberlein K., Potthast A., Thein H., Van Horn J., Toga A., Schmitt F., Lehne D., Rosen B., Wedeen V. & Wald L. (2013). Pushing the limits of in vivo diffusion MRI for the Human Connectome Project. *NeuroImage* **80**:220–233.
- Shemesh N. & Cohen Y. (2011). Microscopic and compartment shape anisotropies in gray and white matter revealed by angular bipolar double-PFG MR. *Magnetic Resonance in Medicine* **65**(5):1216–1227.
- Shemesh N., Ozarslan E., Basser P.J. & Cohen Y. (2009). Measuring small compartmental dimensions with low-q angular double-PGSE NMR: The effect of experimental parameters on signal decay. *Journal of Magnetic Resonance* **198**(1):15–23.
- Siow B., Drobnyak I., Chatterjee A., Lythgoe M.F. & Alexander D.C. (2012). Estimation of pore size in a microstructure phantom using the optimised gradient waveform diffusion weighted NMR sequence. *Journal of Magnetic Resonance* **214**(1):51–60.
- Siow B., Drobnyak I., Ianuș A., Christie I.N., Lythgoe M.F. & Alexander D.C. (2013). Axon radius estimation with oscillating gradient spin echo (OGSE) diffusion MRI. *Diffusion Fundamentals* **18**(1):1–6.
- Stanisz G.J., Szafer A., Wright G.A. & Henkelman R.M. (1997). An analytical model of restricted diffusion in bovine optic nerve. *Magnetic Resonance in Medicine* **37**:103–111.
- Stejskal E.O. & Tanner J.E. (1965). Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *Journal of Chemical Physics* **42**(1):288–292.
- Swain M.J. & Ballard D.H. (1991). Color indexing. *International Journal of Computer Vision* **7**(1):11–32.
- Szafer A., Zhong J. & Gore J.C. (1995). Theoretical model for water diffusion in tissues. *Magnetic Resonance in Medicine* **33**(5):697–712.
- van Gelderen P., DesPres D., van Zijl P.C.M. & Moonen C.T. (1994). Evaluation of restricted diffusion in cylinders. Phosphocreatine in rabbit leg muscle. *Journal of Magnetic Resonance, Series B* **103**:255–260.
- Xu J., Does M.D. & Gore J.C. (2011). Dependence of temporal diffusion spectra on microstructural properties of biological tissues. *Magnetic Resonance Imaging* **29**(3):380–390.
- Zhang H., Dyrby T.B. & Alexander D.C. (2011a). Axon diameter mapping in crossing fibers with diffusion MRI. In G. Fichtinger, A. Martel & T. Peters (Eds.), *Medical Image Computing and Computer-Assisted Intervention*, vol. 6892 of *Lecture Notes in Computer Science*, pp. 82–89. Springer Berlin Heidelberg.
- Zhang H., Hubbard P.L., Parker G.J.M. & Alexander D.C. (2011b). Axon diameter mapping in the presence of orientation dispersion with diffusion MRI. *NeuroImage* **56**:1301–1315.



Supporting Figure S1: Maps of axon radius index, estimated in the absence of the prior on R . Compared to Fig. 5, the maps are noisier and show less scan–rescan consistency.