Microstructural Imaging of the Human Brain with A ‘Super-Scanner’: 10 Key Advantages of Ultra-Strong Gradients for Diffusion MRI

The PIs and Colleagues of the
National Facility for In Vivo MR Imaging of Human Tissue Microstructure

*Jones DK¹,², Alexander DC³,⁴, Bowtell R⁵, Cercignani M⁶, Dell’Acqua F⁷, McHugh DJ⁸,⁹, Miller KL¹⁰, Palombo M³, Parker GJM⁸,⁹,¹¹, Rudrapatna US¹, Tax CMW¹

AFFILIATIONS
1. Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology, Cardiff University, Maindy Road, Cardiff, CF24 4HQ, UK
2. School of Psychology, Faculty of Health Sciences, Australian Catholic University, Melbourne, Victoria 3065, Australia
3. Centre for Medical Image Computing (CMIC), Department of Computer Science, UCL (University College London), Gower Street, London, UK;
4. Clinical Imaging Research Centre, National University of Singapore, Singapore
5. Sir Peter Mansfield Magnetic Resonance Centre, School of Physics and Astronomy, University of Nottingham, University Park, Nottingham, United Kingdom.
6. Department of Psychiatry, Brighton and Sussex Medical School, Brighton, UK.
7. Natbrainlab, Department of Neuroimaging, King’s College London, London, SE5 8AF, UK
8. Division of Informatics, Imaging and Data Sciences, The University of Manchester, Manchester, UK.
9. CRUK and EPSRC Cancer Imaging Centre in Cambridge and Manchester, Cambridge and Manchester, UK.

*Corresponding author
Email: jonesd27@cf.ac.uk; Phone: +44-29-2087-9412
ABSTRACT

The key component of a microstructural diffusion MRI ‘super-scanner’ is a dedicated high-strength gradient system that enables stronger diffusion weightings per unit time compared to conventional gradient designs. This can, in turn, drastically shorten the time needed for diffusion encoding, increase the signal-to-noise ratio, and facilitate measurements at shorter diffusion times. This review, written from the perspective of the National Facility for In Vivo MR Imaging of Human Tissue Microstructure, an initiative to establish a shared 300 mT/m-gradient facility amongst the microstructural imaging community, describes ten advantages of ultra-strong gradients for microstructural imaging. Specifically, we will discuss how the increase of the accessible measurement space compared to a lower-gradient systems (in terms of Δ, b-value, and TE) can accelerate developments in the areas of 1) axon diameter distribution mapping; 2) microstructural parameter estimation; 3) mapping micro- vs macroscopic anisotropy features with gradient waveforms beyond a single pair of pulsed-gradients; 4) multi-contrast experiments, e.g. diffusion-relaxometry; 5) tractography and high-resolution imaging in vivo and 6) post mortem; 7) diffusion-weighted spectroscopy of metabolites other than water; 8) tumour characterisation; 9) functional diffusion MRI; and 10) quality enhancement of images acquired on lower-gradient systems. We finally discuss practical barriers in the use of ultra-strong gradients, and provide an outlook on the next generation of ‘super-scanners’.

INTRODUCTION

This article reviews the benefits of incorporating an ultra-strong gradient system (Setsompop et al. 2013), optimised for diffusion magnetic resonance imaging (dMRI), into an MRI scanner for imaging the human brain. It is written from the perspective of the National Facility for In Vivo MR Imaging of Human Tissue Microstructure, an initiative in the United Kingdom to establish an MRI system with ultra-strong (300 mT/m) gradients as a shared facility amongst the microstructural imaging community. We have selected ten key areas where we believe ultra-strong gradients can advance the field of microstructural imaging, beginning each section with motivation for advancing the field in each area, reviewing what has been achieved to date with more commonly-available gradient amplitudes, and then discussing the potential benefits and opportunities afforded by having access to ultra-strong gradients. In some cases, this is in providing marked improvements to measurements that are already attainable at lower gradient amplitude, and in others, it facilitates new measurements that are simply impractical at lower gradient amplitude. We
then discuss practical challenges in making use of such as system, including issues of safety and engineering aspects, attempt to highlight which limits are fundamental, and which just require engineering.

The main advantage is, of course, in providing a higher q-value/shorter echo time (TE) for a given b-value, and a higher signal-to-noise ratio (SNR) per unit b-value (see Figure 1a). Shorter diffusion time acquisitions also become more practical as higher b-values can be achieved, and a wider range of b-values can be maintained across all diffusion times (Δ) (Figure 1b).

![Figure 1](image_url)

**Figure 1:** Increase of the accessible parameter space with ultra-strong gradients. (a) b-value vs minimum achievable echo time (TE) for different maximum gradient amplitudes of 40 mT/m, 80 mT/m and 300 mT/m (see also Setsompop et al., 2013). For each point, the colour coding gives the achieved Δ = δ/3 [ms] (colour coding according to Δ results in a similar figure with the colourbar ranging from 19 ms – 97 ms). These curves were obtained by simulation on the system with a Stejskal-Tanner EPI sequence, Voxel size = [2.5 2.5 2.5] mm, no partial Fourier, GRAPPA = 2, Multiband = 1. (b) PGSE b-value (colours) plotted as a function of G and Δ for Gmax = 80 mT/m (left two panels) and 300 mT/m (right panel). Differences in rise times for Gmax = 80 and 300 mT/m (0.4 ms
and 3.6 ms, respectively) mean that lower δ values are achievable with G_max = 80 mT/m (e.g. δ = 5 ms as shown in the left-most panel), resulting in shorter and longer achievable Δ, although b-values are low (note the difference in colour scale between the δ = 5 ms panel and the others). For G_max = 80 mT/m and δ = 11 ms, the range of Δ values matches the range available with the shortest δ possible at 300 mT/m (δ = 8 ms, to the nearest integer). All calculations assume TE = 100 ms.

Making measurements on such kit allows us to develop a truly translational pipeline, i.e., enabling measurements in humans that we could only previously make, in vivo, in animals on preclinical systems. Here, we firstly explore how stronger gradients provide improvements in characterising axon diameter, and microstructural parameters in general. Next we discuss the benefits for getting higher resolution characterisation of brain structure both in vivo and post mortem. We then discuss how, by looking at metabolites other than water, diffusion-weighted spectroscopy can provide deeper insights into tissue microstructure, and how ultra-strong gradients can accelerate developments in this area. While microstructural imaging offers huge potential benefits in a range of diseases, we select as an exemplar application the characterisation of tumours, and explore how stronger gradients may improve characterisation of tumour cellularity and other microstructural features. In addition to characterisation of structure, dMRI has been used to characterise brain function, with the study of rapid temporal changes in the microstructural milieu being an interesting application. Given the disparity of results in the literature on this topic, here we consider how ultra-strong gradients may help to cast further light on the relationship between neural activity and diffusion.

Despite the many potential advantages of ultra-strong gradients, we recognise that access to ultra-strong gradient systems is not likely to be wide-spread in the immediate future, and so we consider how having a small number of such systems, together with machine learning approaches, might still facilitate deeper insight into tissue microstructure beyond the immediate environs of the scanners themselves. We then discuss practical barriers to making full use of such as system, considering the interplay of hardware and physiological safety limits on deployment of such gradient magnitudes, and consider how some of these issues may be addressed in future generations of ‘super-scanners’.
1. MEASURING AXON DIAMETER

THE CHALLENGE AND ITS IMPORTANCE

One of the key features of the diffusion-weighted MR signal is its sensitivity to the size of pores with walls that are either fully restricting, or semi-permeable. The longer the characteristic length scale of the pores, the greater the mean displacement, and the lower the diffusion-weighted signal will be. In the brain, this capability underpins the tantalising prospect of estimating and mapping the distribution of axon diameters in each image voxel, since at least in white matter the axonal walls are the primary geometric structures affecting diffusion in the tissue. The axon diameter distribution is one of the key structural features of tissue related to function, as axon diameter relates directly to transmission speed: larger axons transmit signals more quickly (Hurst 1939; Ritchie, 1995; Richardson, McIntyre, & Grill, 2000). Differences in axon diameter offer a potential early disease indicator in conjunction with loss of axon density, which can arise in a much wider range of conditions. Finally, measurements of axon diameter have potential utility in brain connectivity mapping as a distinguishing feature of particular fascicles with which to resolve ambiguities, such as crossing versus kissing configurations (Girard et al., 2017; Sherbondy, Rowe, & Alexander, 2010), and can be combined into the estimation of fibre trajectories, where uniformity of microstructural parameters can be assumed, to provide more robust estimates of connections in approaches such as COMMIT (Daducci, Dal Palù, Lemkaddem, & Thiran, 2015b).

BACKGROUND

Several groups of researchers have made attempts to estimate and map the axon diameter distribution in nervous tissue. Stanisz et al. (Stanisz, Szafer, Wright, & Henkelman, 1997) designed a mathematical model of the signal from nerve tissue comprising separate signals from axonal water, glial water, and extra-cellular water, and assumed that there is restricted diffusion in an ellipsoidal geometry in the axonal compartment. They acquired a rich spectroscopic data set from fixed nerve-tissue samples with a wide range of b-values and diffusion times. The parameters of this model that best fit the data provide estimates of the volume fractions and dimensions of the various tissue components – in particular, fitting the model to the data identifies dimensions of the ellipsoidal geometry, the minor axis of which is an estimate of the axon diameter. Assaf et al. (Assaf, Blumenfeld-Katzir, Yovel, & Basser, 2008) later presented the AxCaliber technique, which uses a similar model to Stanisz, but assumes cylindrical axonal geometry with a gamma distribution of diameters; these authors also simplify the model by assuming no glial cell contribution and no exchange among compartments. They show a similar spectroscopic experiment to Stanisz using two nerve samples (sciatic, which has relatively large axons, and optic, with smaller diameters) and show a good match between the estimated axon diameter distribution (parameters of the fitted gamma distribution) and the histogram of diameters measured on histology of the same samples. Later work by the same group...
(Barazany, Basser, & Assaf, 2009) combined AxCaliber with imaging to map the axon diameter distribution over the corpus callosum of a live rat and again show good agreement with histological analysis of the axon diameter distribution in sagittal sections of the mid-line (see Figure 2a).

A key assumption in the techniques above is that the fibre orientation is fixed, and perpendicular to the direction of the applied diffusion encoding gradient. While this condition can be satisfied for limited portions of white matter, extending the approach to the whole brain clearly requires a generalisation of the method, to account for:

1. different orientations of a single bundle;
2. multiple distinct fibre bundles passing through the same voxel; and
3. intra-voxel orientational dispersion of a given fibre bundle.

ActiveAx (Alexander et al., 2010) addressed issue (i), by combining high angular resolution diffusion imaging (HARDI) with a simplified model designed to minimise complexity. In this way, they provide an orientationally invariant estimate of a single axon diameter index (approximately a volume-weighted mean (Alexander et al., 2010; Benjamini, Komlosh, Holtzclaw, Nevo, & Basser, 2016); but also see (Burcaw, Fieremans, & Novikov, 2015) for a detailed discussion). Experiments with fixed monkey brains recover the low-high-low trend in axon diameter across the mid-sagittal corpus callosum (known from histology to persist across mammalian species (Aboitiz and Montiel, 2003; Olivares et al. 2001) with high reproducibility, and preliminary results from human volunteers show similar trends albeit more weakly (see Figure 2b). AxCaliber3D (Barazany, Jones and Assaf, 2011) addressed issues (i) and (ii) above (i.e., recovering distinct axon diameter distributions for multiple fibres within the same voxel), by acquiring CHARMED (Assaf & Basser, 2005) data sets at multiple diffusion times. Once the CHARMED model is fitted, it is possible to simulate the diffusion-time dependence of the signal at arbitrary orientations and therefore perpendicular to any fibre component within the voxel. These data are inputted into the AxCaliber framework to recover an apparent axon diameter distribution for that fibre population. Refinements to ActiveAx address issues (ii) and (iii): within-voxel fibre crossings (Zhang, Dyrby, & Alexander, 2011a), as well as within-voxel fibre-orientation dispersion (Zhang, Hubbard, Parker, & Alexander, 2011b). These advances in modeling extend the portion of white matter over which we can obtain sensible axon-diameter indices.

More recent advances in model-based techniques use refined models, which e.g. capture the time-dependence of extra-cellular diffusion (De Santis, Jones, & Roebroeck, 2016; Sepehrband, Alexander, Kurniawan, Reutens, & Yang, 2016) and non-standard pulse sequences, such as oscillating gradients (Kakkar et al., 2017; Siow, Drobnjak, Chatterjee, Lythgoe, & Alexander, 2012; Xu et al., 2014; 2016), double diffusion encoding (Benjamini et al., 2016) and generalized waveforms (Drobnjak, Siow, & Alexander, 2010; Drobnjak, Zhang, Hall, & Alexander, 2011). Alternative estimation approaches include a q-space formulation (Ong et al., 2008; Ong & Wehrli, 2010), convex optimization (Daducci, Dal Palù, Lemkaddem, & Thiran, 2015b), integration with fibre-tracking (Sherbondy, Rowe & Alexander, 2010; Daducci, Dal Palù, Lemkaddem, & Thiran, 2015b; Girard et al., 2017) and dictionary matching (Sepehrband et al., 2016).
THE PROMISE

One of the fundamental limits to axon-diameter estimation is the so-called “resolution limit” of the diffusion experiment. The diffusion-weighted MR signal is sensitive only to a window of restriction lengths and the bounds of that window depend strongly on the available gradient strength. Moreover, most axon diameters lie below the lower bound of the window (the resolution limit) given the available gradient strength in most currently available MRI scanners. Dyrby et al. (Dyrby, Søgaard, Hall, Ptito, & Alexander, 2012; Dyrby et al., 2012) demonstrate empirically the advantages of increasing gradient strength in axon diameter mapping using simulations and ex vivo measurements on a monkey corpus callosum using a small-bore scanner with 400mT/m gradients (see Figure 2c). Drobnjak et al. (Drobnjak, Zhang, Ianuş, Kaden, & Alexander, 2016) identified the resolution limit numerically for combinations of pulsed gradient spin echo (PGSE) and oscillating gradient spin echo (OGSE) measurements, and showed that in the idealised experiment where the gradients are perpendicular to straight, parallel, impermeable cylinders, the standard PGSE experiment gives the greatest sensitivity and that the resolution limit on axon diameter is in the range 4.5-7um (depending on available SNR) with 60mT/m, but reduces to 2-3um with 300mT/m. In the more realistic situation for brain tissue, where the magnetic field gradient is not quite perpendicular to the fibres and/or the fibres have dispersed orientations, the resolution limit increases, although using OGSE instead of PGSE mitigates this increase to some extent. For example, with the gradient 10 degrees from perpendicular, or with moderate orientational dispersion (characterised by a Watson distribution with a dispersion parameter of 16), using OGSE, the resolution limit increases by about 1um (see Figure 2d).
FIGURE 2— Various results estimating axon diameter distribution from the literature (see text). (a) results from Barazany et al. (2009) comparing electron microscopy data and in vivo estimates of axon diameter distribution in the rat brain; (b) Axon diameter index maps in fixed monkey brains showing a characteristic low-high-low trend in axon diameter across the mid-sagittal corpus callosum from Alexander et al 2010; (c) Comparison of axon diameter index maps at increasing (top to bottom) maximum gradient strength using ex vivo measurements on a monkey corpus callosum from a small-bore scanner with 400mT/m gradients from Dyrby et al. (2012); (d) demonstration of benefit of higher SNR and gradient amplitude for minimum resolvable diameter (y-axis) (from Nilsson et al. 2017) both with no fibre orientation dispersion (left) and with fibre orientation dispersion (right).

Nilsson et al. (Nilsson, Lasič, Drobnjak, Topgaard, & Westin, 2017) confirmed these findings analytically. They showed, for perfectly straight parallel cylinders and with the gradient perfectly perpendicular to them, (i.e., with very little attenuation in the intra-axonal water signal), that for a PGSE sequence: $d_{\text{min}} = 
\left(1 / \sqrt{G_{\text{max}}}ight) \left(1 / \delta^2 \sigma D_0 \gamma^2 \Delta \right)^{1/3}$, where $D_0$ is the free diffusivity of the intra-axonal water, $G_{\text{max}}$ is the maximum gradient amplitude, $\delta$ is their pulsed duration, equal to the diffusion time, $\Delta$, and $\sigma$ is the
minimal detectable percentage change in the signal when the gradients are applied (i.e., the equation holds at a particular SNR that is assumed unaffected by, e.g., the choice of \( \delta \)). Critically, the resolution limit depends on \( G_{\text{max}}^{-1/2} \) as well as \( D_0^{1/4} \), for any shape of gradient waveform. In addition to strong gradients, this also calls for a high signal to noise ratio, which may be obtained by more efficient readout schemes (such as spiral read-out to shorten the echo time), more sensitive receiver coils, or indeed by increasing the static (\( B_0 \)) field strength.

The resolution limit raises the question of whether diffusion MRI can provide any useful features of the axon diameter distribution at all, since histological measurements of most axons are 1 um or less – at least a factor of 2 below our best-case resolution limit. However, a few further considerations may renew hope, particularly with the benefits of 300 mT/m gradient systems. First, histological preparation typically causes tissue shrinkage and in-situ and in vivo axons may be somewhat larger in diameter than histological measurements suggest. The degree of shrinkage is hard to establish, but estimates range between 0-65% (Alexander, Dyrby, Nilsson, & Zhang, 2017a). Second, although large axons are a tiny minority in terms of number of axons, they are much more significant in terms of the MRI signal they contribute, because of the relative volume of water they contain. This means that variation in the large-diameter tail of the axon diameter distribution has a disproportionate influence on the MR signal contrast. Figure 3 highlights the practical importance of this observation – although less than 10% of the axons in this example have diameter over 3um, those axons provide 50% of the intra-axonal signal. Moreover, as the resolution limit decreases (due to gradient amplitude increasing from 60mT/m to 300mT/m), the proportion of axons above the resolution limit changes slowly from about 5% to about 20%. However, the proportion of the signal that comes from these axons increases much more – from about 35% to over 70% for the same increase in gradient strength. This suggests that diffusion MR experiments may hold strong sensitivity at least to variability in the existence and size of large axons. It further highlights that 300mT/m gradients offer a major benefit in recovering information from the axon diameter distribution; early results in the brain (with and without multiple sclerosis pathology) (Huang et al., 2015; Huang et al., 2016) and in the spinal cord (Duval et al., 2017) show promise in exploiting these advantages. Given the fact that there is a differential sensitivity of axons, according their diameter, in diseases such as multiple sclerosis (DeLuca et al., 2004), the potential gains from being able to map apparent axon diameter indices are clear.
Figure 3. Illustration of the benefit of gradient strength on recovering features of the axon diameter distribution. Panel (a) shows a histogram of axon diameters from the corpus callosum of a post-mortem human brain, compiled from the data in (Aboitiz, Scheibel, Fisher, & Zaidel, 1992), scaled by the inverse of a typical shrinkage factor of 1.5. The red curve is the cumulative distribution. Panel (b) shows the same histogram as in (a) but now weighted by axonal volume - the histogram of space occupied by axons of each diameter, which corresponds to the proportion of signal contributed by axons in each size bin; again the red line is the cumulative distribution. In panel (c), the graph shows the proportion of axons (in blue) that are above the resolution limit, as a function of available gradient strength, as well as the proportion of intra-axonal signal (in red) that comes from such axons. Error bars show how that proportion varies for histological shrinkage factors over the range: [1, 2]. The simulation accounts for the limitations on rise time imposed by physiological safety constraints (see Section 11).
The resolution limit is not the only challenge for axon diameter mapping. Major challenges remain in identifying models that link the MR signal to underlying tissue parameters sufficiently well to isolate contrast specifically arising from the axon diameter distribution. A range of other effects can confound the estimation. For example, although the original work on ActiveAx (Alexander et al., 2010) showed trends in axon diameter consistent with expectations from histology (Aboitiz et al., 1992), there was nevertheless a substantial overestimation of the axon diameter index in both in vivo and ex vivo experiments. The authors suggest orientation dispersion as one possible explanation. However, even when accounting for orientational dispersion (Zhang et al., 2011b) there is still substantial, albeit reduced, overestimation. This could, in part, be attributable to the definition of ‘orientational dispersion’: Zhang et al. (Zhang et al., 2011b) primarily consider only macroscopic dispersion, i.e., variation in orientation among axons, each of which is assumed to be straight. However, microscopic orientation dispersion, where individual axons vary in orientation over the extent of an image voxel (e.g., axonal undulation), can have a similar confounding effect leading to the characteristic length scale being driven by the amplitude of the undulation as well as axon diameter (Nilsson, Lätt, Ståhlberg, van Westen, & Hagslätt, 2012). Exchange between the intra-axonal and extracellular space due to non-zero cell membrane permeability can also potentially lead to overestimation. While such exchange can be incorporated into the modeling process as a parameter of interest in its own right, see e.g. (Nedjati-Gilani et al 2017), strong gradients resulting in shorter echo times can alternatively help to reduce the influence of exchange. Other effects such as complexity of the extracellular diffusion process (Burcaw et al., 2015) and differences in the relaxation parameters of different water pool, (see Section 4), can also affect the signal significantly, but are often ignored in models used for axon diameter estimation; a study by De Santis et al. (De Santis et al., 2016c) reports a reduction of estimation bias in axon diameter estimates when modeling time-dependence of the extra-axonal diffusivity.

Despite this bias in estimates of features of the axon diameter, Horowitz et al. (Horowitz et al., 2014) showed correlation between estimates of apparent axon diameter distribution and inter-hemispheric apparent conduction velocity, inferred from electrophysiological measurements, suggesting that these measurements might explain individual differences in brain function. This result suggests that even with a lack of accuracy in characterising the absolute axonal distribution it may be possible to define useful imaging biomarkers that characterise disease and relate strongly to axon diameter. Criticism of the work (Innocenti, Caminiti, & Aboitiz, 2015) focuses on the resolution limit and potential bias in estimates of the axon diameter distribution from diffusion MRI. Of course, if the bias (e.g., over-estimation) in axon diameter estimate is approximately uniform, or even monotonic, sufficient correlation may remain to infer apparent conduction velocity from axon diameter distribution estimates. Whether we can predict actual conduction velocity from microstructural measurements in vivo is an altogether different question.
In summary, although it is a compelling application of diffusion MRI, major challenges remain in obtaining trusted measurements and maps of the axon diameter distribution from MRI. However, even having access to apparent axon diameter distributions might give enhanced sensitivity to white matter disease, earlier response to intervention, and increased prognostic value, when compared to other techniques that ignore differences in axon diameter, or treat them as invariant. The range of confounding effects mean that we are yet to be certain about the exact sensitivity to axon diameter distributions, or the range of effects to include in models that enables us to isolate that sensitivity. Nevertheless, in the words of Nilsson et al (Nilsson et al., 2017): “… the diffraction limit in optical microscopy, introduced in 1873, was recently broken. It took 135 years. Perhaps breaking the resolution limit in diffusion MRI can be done a little faster?”

2. PARAMETER ESTIMATION

THE CHALLENGE AND ITS IMPORTANCE
To describe the complex diffusion process in tissue, a variety of approaches have been proposed that relate the signal to diffusion features or tissue microstructure, of which some have been described in the previous section (see: Panagiotaki et al., 2012; Assemal et al., 2011; Tournier et al., 2011; and Novikov et al. 2018) for further discussion). The subsequent step of ‘fitting’ such models or representations to the dMRI data brings up additional challenges. In this process, one tries to find the values of the model parameters that could have generated the observed diffusion weighted signals. Solving this ‘inverse problem’ generally involves the estimation of more than one free parameter. It is the considerable number of parameters associated with most dMRI models, together with the typically low SNR in diffusion MRI, that make robust microstructural parameter estimation quite challenging. Moreover, the non-Gaussian distribution of the RF noise in the dMRI signal (Aja-Fernandez 2016), as well as measurement errors from further signal perturbations such as motion and distortion (Le Bihan et al. 2006; Pierpaoli 2010; Tax et al. 2016), further complicate the recovery of accurate and precise microstructural estimates from models of tissue microstructure. Even in the context of diffusion tensor MRI (DT-MRI) (Basser & Jones, 2002), with its relatively ‘conventional’ b-value (e.g., b ~ 1000 s/mm²) acquisition, parameter estimation is not straightforward: there is a vast body of literature focusing on DT-MRI estimation procedures such as least-squares (Koay, Carew, Alexander, Basser, & Meyerand, 2006;Veraart et al., 2012;Veraart, Sijbers, Sunaert, Leemans, & Jeurissen, 2013) (linear, weighted linear, nonlinear) and maximum likelihood (Fillard, Pennec, Arsigny, & Ayache, 2007; Landman, Bazin, & Prince, 2007; Liu, Gasbarra, & Railavo, 2016); as well as on the incorporation of different constraints, regularizations, re-parameterizations, and outlier rejection
strategies (Chang et al. 2005; Chang et al. 2012; Parker et al. 2013; Collier et al. 2015; Tax et al. 2015; Parker and Jones 2017)

For the parameter estimation of multi-compartment biophysical models, non-linear estimation procedures are commonly employed. Nonlinear fitting methods try to find an optimum by gradually moving around in a ‘nonlinear parameter landscape’, which is time-consuming and prone to getting stuck in local optima.

**BACKGROUND**

To address some of the issues in nonlinear fitting on currently available data, the parameter search space can be reduced by decoupling the estimation of the number and orientation of fascicles, from the microstructural parameters of each fascicle. Daducci et al. (Daducci, Canales-Rodríguez, Zhang, Dyrby, Alexander, et al., 2015a) proposed the AMICO framework that linearizes the problem and drastically speeds up the fitting.

Alternatively, the fitting of multi-compartment parameters can be facilitated by first obtaining some main characteristics of the signal. For example, a rotationally invariant representation of the signal can be derived prior to fitting, e.g. the powder average (also called the orientational average, spherical mean, directional mean, isotropic mean) (Topgaard and Söderman 2002; Kaden, Kruggel, & Alexander, 2016; Jespersen, Lundell, Sønderby, & Dyrby, 2013; Lasić, Szczepankiewicz, Eriksson, Nilsson, & Topgaard, 2014), possibly accompanied by higher order invariants (Mirzaalian et al. 2015; Novikov, Veraart, Jelescu & Fieremans, 2018). From this rotationally-invariant signal, the voxel-averaged microscopic properties can subsequently be derived.

Fieremans et al. (Fieremans, Jensen, & Helpern, 2011) first fit the DKI representation and derive microstructural parameters through analytical relationships with kurtosis parameters. Fick et al. (Fick, Wassermann, Caruyer, & Deriche, 2016) first fit a regularized version of the mean apparent propagator (MAP) method (Özarslan et al. 2013) and then extrapolate the signal as a pre-processing step for multi-compartment model fitting. The advantage is that the multi-compartment estimation can benefit from advanced estimation methods that have already been developed for these other techniques, in addition to the inherent regularisation they provide. Different configurations of multi-compartment models have been evaluated in preclinical acquisitions (Panagiotaki et al., 2012) and in *in vivo* human studies (Bourne et al., 2014; Ferizi et al., 2014). In group studies, multi-compartment models have revealed microstructural changes in a range of neurological and psychiatric disorders, e.g. (De Santis et al., 2017; Kamagata et al., 2017; Schneider et al., 2014; Slattery et al., 2017);

Despite recent efforts in parameter estimation, it has become increasingly apparent that the ultimate aim of being unambiguously specific to the relevant microstructural characteristics can hardly be achieved with
the gradient hardware currently available on clinical MRI systems. Fixing certain microstructural model parameters to a priori values that are assumed to be accurate for the living human brain is a common approach to alleviating the problems with non-linear fitting. However, fixing parameters can generate bias in the other estimated parameters, calling into question their ‘true specificity’ to the biological reality (Jelescu, Veraart, Fieremans, & Novikov, 2016).

The parameter landscape and fitting has been extensively studied for an often used multicompartment dMRI model - sometimes referred to as the ‘standard model’ - in which each compartment can be described by an anisotropic Gaussian function (i.e. a tensor) (Kroenke et al., 2004; Jespersen et al., 2007, 2010), an extension of the model developed by Stanisz et al (Stanisz et al., 1997) and refined by Assaf et al. (Assaf & Basser, 2005). In this model, the measurement is assumed to be beyond the resolution limit described in Section 1 and as such the intra-neurite space is represented as sticks (tensors with zero perpendicular diffusivity). The compartment configuration where the intra-neurite space has perpendicular diffusivity equal to zero, has been used in various approaches such as NODDI (Zhang, Schneider, Wheeler-Kingshott, & Alexander, 2012) and the ball and stick model (Behrens et al., 2003). The extra-neurite space is described by a prolate tensor with nonzero parallel- and perpendicular diffusivities. For reliable fitting of this model, it is often assumed that parameters in the model exhibit the same value (e.g., the parallel intra- and extra-neurite diffusivities are the same), or are fixed to some known value. In the work of Jelescu et al. (Jelescu et al., 2016) all parameters were allowed to be independent of each other. It was found empirically (Jelescu et al., 2016), and analytically (Novikov, Veraart, Jelescu & Fieremans, 2017), that the optimization landscape is very flat, and that the fitting is intrinsically degenerate with two solution ‘branches’ that are equally biologically plausible: parallel intra-neurite diffusivity is larger than parallel extra-neurite diffusivity, or vice versa. For acquisitions up to typical ‘clinical’ b-values (~2500 s/mm^2), it was shown that it impossible to determine which branch is correct, even in the simplest case of straight parallel fibres. Experiments without fixed parameters (Lampinen et al. 2017) indicate that the assumption of equal parallel diffusivities does not seem to hold in WM; neither does the widely employed tortuosity assumption.

THE PROMISE

For the ‘standard’ model described above, high b-value measurements with sufficient SNR hold the potential to determine which is the correct solution branch to adopt, allowing estimation of all the parameters unambiguously, without having to fix them a priori (Jelescu et al., 2016). Preliminary results with b-values up to 10,000 s/mm² from an 80 mT/m gradient system conjecture that the parallel intra-neurite diffusivity is larger in WM, but noise still challenges the estimation and branch selection remains an open question. Compared to standard hardware, ultra-strong gradients allow an increased SNR per unit time can be obtained for high b-values, so that the high b-value parameter space can be studied into more detail and the findings at lower gradient strengths can be validated.
In the case of ambiguities / degeneracies, i.e., where different microstructural configurations lead to the same diffusion-weighted signal (such as the solution ‘branches’ discussed above), ‘orthogonal’ measures that bring new information to the problem are crucial (Novikov et al. 2016; Lampinen et al. 2017). The potential widespread benefit of the super-scanner, therefore, is to provide measurements that are not normally available, but which can inform on parameters in the measurements that we can normally obtain. This provides a more informed strategy for choosing fixed values or priors to stabilize models we fit to routinely-acquired data.

3. MACROSTRUCTURAL vs MICROSTRUCTURAL ANISOTROPY

THE CHALLENGE AND ITS IMPORTANCE

The majority of approaches developed to extract diffusion- or microstructural features from dMRI data, are built on the Stejskal-Tanner sequence (Stejskal & Tanner, 1965) (or ‘single diffusion encoding’, SDE). Diffusion tensor MRI (Basser, Mattiello, & LeBihan, 1994), still being the most commonly applied SDE-based method to showcase sensitivity of dMRI to pathology, merely provides information on the average diffusion process within a voxel. Macroscopic DTI features such as fractional anisotropy (FA) and mean diffusivity (MD) concomitantly give an indication of the average orientation, shape, and size of the diffusion process on the voxel-scale. In reality, however, tissue is more accurately described as being composed of different compartments with distinct characteristics: a distribution of microenvironments (represented by a sum of tensors under certain assumptions (Szczepankiewicz et al., 2015)). See Figure 4a
Figure 4.

a. The difference between microscopic and macroscopic anisotropy, and impact of size variance and orientation coherence on macroscopic anisotropy. $V_I$ = isotropic volume fraction; $V_A$ = anisotropic volume fraction; FA = (macroscopic) fractional anisotropy; $\mu_{FA}$ = microscopic fractional anisotropy. Despite having the same $\mu_{FA}$, the different substrates in the right hand column (for example) have difference macroscopic FA values, purely due to orientational dispersion. (Taken from the Ph.D. thesis of Filip Szczepankiewicz, Lund University, with permission)

b. Comparison of $\mu_{FA}$ and FA in the same healthy (age = 45 years) participant. Note the areas of low FA (dark) where different pathways cross / merge, while the $\mu_{FA}$ is relatively uniform in white matter.

Consequently, the macroscopic FA cannot provide an interpretable marker other than a vague concept of ‘tissue integrity’ (Jones, Knösch, & Turner, 2013); the voxel-scale diffusion tensor entangles the information about the orientation, shape, and size of the microenvironments. It can therefore not distinguish between different configurations of orientation incoherence, microscopic anisotropy, and isotropic heterogeneity.

Many alternative strategies, based on SDE, have been proposed that aim to resolve the limitations of DTI and to distinguish between different configurations. Orientation incoherence can for example be accounted for by explicitly modeling crossing or dispersing microenvironments as for example in spherical deconvolution-based frameworks (e.g.Tournier et al. 2004, 2007; Jeurissen et al. 2014), Lätt et al. 2003) or in NODDI (Zhang et al. 2012), or by factoring out the orientational dependency (Basser & Pierpaoli, 1998; Edén, 2003; Kaden et al., 2016a, 2016b). However, a common underlying assumption is that the kernel (or response function, or single-fiber profile (Tax, Jeurissen, Vos, Viergever, & Leemans, 2014)) is the same along every orientation. Obtaining such information on the average shapes and sizes of the microenvironments that constitute the kernel has already turned out to be challenging based on SDE measurements alone (e.g., an intra- and extra-neurite compartment) – as discussed in Section 2. In addition, the assumption of a single kernel per voxel may not always hold, e.g. in the case of pathology affecting one of multiple crossing fascicles.
SDE-based approaches that aim at quantifying the heterogeneity of diffusivities include statistical approaches (Yablonskiy & Sukstanskii, 2010), q-space approaches (King et al. 1994; Assaf and Cohen 2000; Cohen and Assaf 2002; Assaf et al. 2002; Lätt et al. 2008), and diffusion kurtosis imaging (Jensen, 2015; Szczepankiewicz et al., 2015). Statistical approaches, for example, describe the signal as a sum of signals from spin packets that originate from different positions within a voxel (Scherrer et al., 2016; Yablonskiy & Sukstanskii, 2010). These spin packets interfere with different microenvironments and their contribution to the total signal can be described as a distribution of apparent diffusivities. However, if a voxel with a low macroscopic fractional anisotropy has a broad distribution of apparent diffusivities, it is impossible to tell if this is due to multiple isotropic compartments with variable diffusivity (isotropic heterogeneity), or due to multiple anisotropic compartments with variable orientation (microscopic anisotropy and orientation incoherence), or both.

Taken together, while many strategies have been proposed to overcome the shortcomings of DTI, SDE-based techniques are inherently limited in that they cannot disentangle variations in size, shape, and orientation without prior knowledge or assumptions about the compartmentalisation, that have to be explicitly incorporated as constraints in signal modeling. Alternative acquisition strategies beyond SDE are essential to get a more complete picture of the distribution of microenvironments in the tissue.

**BACKGROUND**

Instead of using a single pair of diffusion pulses to encode the signal as in SDE, multiple periods of diffusion encoding can be employed to establish displacement correlations of spins. Double diffusion encoding (DDE, Mitra, 1995; Cory, Garroway, Miller, 1990; Shemesh and Cohen 2008) and multiple diffusion encoding (MDE) sequences have two or more diffusion encoding periods possibly at an angle (‘angular DDE’), and can distinguish the scenarios of spherical compartments from incoherently oriented anisotropic compartments while relaxing modeling constraints (Lawrenz & Finsterbusch, 2013). As such, it becomes possible to derive microscopic anisotropy measures like the micro-FA that, in analogy to FA (or macro-FA) is related to the variance of the eigenvalues of the microenvironment diffusion tensors. In the absence of orientation incoherence, the micro-FA and macro-FA are the same. DDE and MDE measurements are ideally designed to be rotationally invariant (Jespersen et al., 2013), so that the measured micro-FA does not depend on the orientational configuration of the tissue under investigation.

In addition to pulsed field gradients, more arbitrary time varying gradient waveforms have been proposed (Drobnjak & Alexander, 2011; Lasič et al., 2014; Lundell et al. 2015; Westin et al., 2016) that can have a higher encoding efficiency than DDE to facilitate imaging in a clinical setting. The general time-varying waveforms describe a trajectory in q-space, referred to as Q-space trajectory encoding (QTE) (Westin et al.,
In this framework, the q-space trajectory can be characterised by its B-tensor (or second-order moment). The B-tensor is a natural extension of the b-value (with the b-value being its trace) and has previously been adopted to take into account cross-terms related to the imaging gradients (Karlicek & Lowe, 1980; Mattiello, Basser, & Le Bihan, 1997; Mattiello, Basser, & LeBihan, 1994). SDE and angular DDE acquisitions can in this framework be described by ‘linear’ B-tensors of rank 1 and ‘planar’ B-tensors of rank 2, respectively. ‘Spherical’ B-tensors (or ‘isotropic encodings’) have been used to obtain fast trace-weighted images (Mori & van Zijl, 1995; Eriksson, Lasič, & Topgaard, 2013; Wong, Cox, & Song, 1995).

Microscopic diffusion anisotropy has been measured with clinical-gradient systems using DDE (Hui & Jensen, 2015; Lawrenz, Brassen, & Finsterbusch, 2016) and QTE (Lasič et al., 2014). Several analysis frameworks have been developed that aim to extract additional features from measurements beyond SDE (Hui & Jensen, 2015; Jespersen et al., 2013; Westin et al., 2016) on clinical systems. If it is assumed that the tissue can be represented by a distribution of diffusion tensors, then, by exploiting the cumulant expansion of the QTE signal (Westin et al., 2016), information on the covariance of the microenvironment tensors can be obtained. Significant differences in such features were found between a small population of healthy controls and schizophrenia patients. Alternatively, the DIVIDE framework (Szczepankiewicz et al., 2015) uses a parameterisation of the distribution (e.g. a Gamma distribution) and relates this to the powder-averaged signal obtained from varying B-tensor shapes. In this way, isotropic heterogeneity and microscopic anisotropy could be disentangled and the isotropic variance component could be studied in vivo. This approach has been used to study the heterogeneity in meningiomas and gliomas, and it was shown that the dominant type of variance (isotropic or anisotropic) was different in both tumours, and that this correlated with microscopy findings (Szczepankiewicz et al. 2016).

Measurements beyond SDE have also been used in conjunction with biophysical tissue models that had originally been developed and evaluated based on SDE measurements. DDE and QTE measurements provide invaluable additional information to validate commonly adopted constraints (Lampinen et al, 2017) and to substantiate which compartments to include (Benjamini & Basser, 2016). Spherical tensor encoding has been used to investigate whether a still water compartment (or ‘dot’ compartment), which could represent relatively immobile water in small spherical structures such as cells, indeed explains the better fit in WM previously observed with SDE data.

THE PROMISE

Methods that extract features of the distribution of tensors (assuming to represent the distribution of microenvironments) in the in vivo human brain commonly make some assumptions about the shape of the distribution. In physical chemistry, unconstrained inversion is a method that is commonly employed to find
solutions for the distribution of microenvironments in a variety of samples (Kim, Doyle, Wisnowski, Kim, & Haldar, 2017; Veraart, Novikov, & Fieremans, 2017), e.g. liquid crystals or rock samples. However, these methods are generally ill-conditioned and very sensitive to noise, and require the signal to be sampled over a wide range of b-values with sufficient SNR. On clinical systems with 40 mT/m gradients, performing an unconstrained inversion to obtain solutions for the full tensor distribution is therefore likely less stable.

Using ultra-strong gradients, it becomes feasible to access a larger part of the b-value space. Preliminary results (Tax et al., ISMRM 2018) show that unconstrained inversion yields results in agreement with known anatomy. Even though non-standard encoding with strong gradients raises challenges (e.g. gradient nonlinearities, Section 11), it is without question that with ongoing developments to better understand and reduce the influence of such confounds, the fidelity of the inversion will be improved by the increased SNR and larger b-value range that strong gradients provide. Arbitrary waveforms can be combined with hybrid experiments (Section 4) to obtain a comprehensive description of the chemical composition, density, size, shape, and orientation of in vivo human brain tissue. The unconstrained inversion can be used to explore healthy and pathological tissues in which it is unclear which constituents are present and to find suitable constraints.

DDE and arbitrary waveforms often require a relatively long TE. Ultra-strong gradients in combination with optimised waveforms (Drobnjak et al. 2010; Drobnjak et al. 2011; Sjolund et al. 2015) can significantly reduce the TE in such experiments. For example, the signal fraction of a still water compartment could be underestimated at the long TE obtainable on clinical systems if its T2 is shorter than that of other compartments (Veraart et al., 2017; Wilm et al., 2017). Sampling the signal at shorter TE allows to investigate this possible confound. Moreover, the shorter TE afforded by the stronger gradients results in an improvement in SNR of approximately 50% compared to that achievable on a system with 80 mT/m gradients.

4. MULTI-CONTRAST EXPERIMENTS

THE CHALLENGE AND ITS IMPORTANCE

A variety of MRI contrasts have been used to characterize tissue beyond dMRI, such as relaxometry, susceptibility, and magnetization transfer. Multi-compartment models have become particularly popular in the attempt to provide a suitable description for each of these contrasts. Most commonly, such multi-compartment models are fitted to the signal from each MRI modality separately, and the joint and complementary information between contrasts therefore remains unexploited. In addition, important
challenges in the multi-compartment modeling of each of the contrasts hampers an unequivocal characterization of the tissue microstructure.

There are challenges that are inherent to the mathematical formulation of these models, however: in dMRI and relaxometry, for example, the signal attenuation arising from the putative compartments is commonly described by a sum of exponentials. In Section 2, fitting challenges in dMRI were already highlighted for the case of a commonly used model with Gaussian compartments. The mathematical challenge of decomposing the multi-exponential decay into its constituents in the presence of noise is a fundamental problem that has persistently attracted attention in the past two centuries. Its ill-posed nature (Istratov & Vyvenko, 1999) means that there exist an infinite number of probability distributions of exponential components that could result in the same measured signal decay. To select a single solution from all the possible ones it is necessary to introduce appropriate prior information (i.e., regularisation) into the inversion, such as positivity, smoothness, or a predetermined form of the solution. One example is restricting the number of compartments but this does not always give an unambiguous answer (e.g. the discussion about fitting two compartments in Section 2).

A challenge specific to dMRI in the brain is that while it has marked sensitivity to differences in axon morphology (see Section 1), it has lower sensitivity to myelin (Beaulieu & Allen, 1994) (Beaulieu, 2002). For example, the fractional anisotropy in the corpus callosum of the Shiverer mouse (with no myelination) is only 10% lower than in wild type mice (Nair et al., 2005). Diffusing species within the myelin itself will be largely invisible at the TEs typical of a diffusion-weighted experiment, since the macromolecules have ultra-short T2 (Fischer, Rinck, Van Haverbeke, & Muller, 1990; Fischer, Van Haverbeke, Schmitz-Feuerhake, & Muller, 1989), while water trapped within myelin has T2s lower than ~40ms (MacKay et al., 1994). Overall, the complexity of the white matter is such that one MRI contrast is unlikely to be sufficient to provide a complete characterisation of its microstructure. One might argue that understanding the precise origin of the signal change is not important, as long as we have access to tools that are sensitive to such a change. However, disentangling the contributions from myelin, axons and glial cells is essential to understand the pathophysiology of neurological disorders and therefore to develop treatments. In addition to disease (Bodini & Ciccarelli, 2014; Lim & Helpern, 2002; White & Lim, 2010) (Lim & Helpern, 2002), measuring white matter changes is also relevant for understanding the mechanisms underpinning plastic changes occurring to the brain as a consequence of maturation (Hüppi, 2010; Minati, Grisoli, & Bruzzone, 2007; Sullivan & Pfefferbaum, 2010) and training (Johansen-Berg, 2017) and lifestyle (Marks, Katz, Styner, & Smith, 2011; Ritchie et al., 2017).

The lack of sensitivity to myelin of dMRI can be compensated by combining it with other techniques, such as multi-component relaxometry (MacKay et al., 1994), magnetization transfer (Henkelman, Stanisz, &
Graham, 2001; Sled et al., 2004; Wolff & Balaban, 1989) and susceptibility imaging (Duyn, 2013; 2017; Duyn & Schenck, 2017; Shmueli et al., 2009). While, in principle, the modalities of interest can be acquired independently and combined at the processing stage, the idea of manipulating more than one source of contrast within the same experiments is attractive. This type of “hybrid” experiment could be more time efficient, and minimise any source of bias, such as geometrical differences between acquisitions, and misregistration between voxels in the image (which might compromise the identification of sub-voxel compartments). This approach is also conceptually more straightforward from a processing perspective, because it enables the fitting of a single equation to the data.

The ideal hybrid experiment aims at reconstructing an N-dimensional spectrum, where the N dimensions represent any combination of MR parameters of interest such as (T1, T2), (T2, Diffusion), (T1, Diffusion) in the case of N=2 (Callaghan, Godefroy, & Ryland, 2003), and potentially other combinations (e.g. diffusion and quantitative susceptibility mapping (Kaden et al. 2018). The complementarity of these MRI modalities implies that the separation of distinct environments can be improved by exploring their multidimensional correlation, compared with any of the 1D spectra (Celik, Bouhrara, Reiter, Fishbein, & Spencer, 2013). While several multimodal experiments can be conceived to characterize the brain microstructure, here we will focus on those combining diffusion and relaxometry. This is partly justified by the abundance of examples in the literature, but also by the fact that this type of experiment is likely to benefit substantially from the use of ultra-high gradients. The popularity of the diffusion-relaxometry framework can be explained by the complementary nature of the information available with these techniques.

BACKGROUND
The first examples of “hybrid experiments” in biological tissue date back to the late 90s, when attempts were made to establish the relationship between T2-species and diffusion behaviour (Peled, Cory, Raymond, Kirschner, & Jolesz, 1999; Stanisz et al., 1997). The interest was prompted by the observation of multi-exponential behaviour for T2 and the TE dependency of some diffusion parameters, suggesting the presence of several compartments. Those experiments were mostly performed ex vivo, using some variant of the diffusion-weighted (DW) Carr-Purcell-Meiboom-Gill (CPMG) sequence, to assess whether compartments with shorter T2 (potentially associated with myelin) had diffusion properties consistent with this hypothesis. This acquisition consists of a standard pulsed-gradient spin-echo (PGSE) preparation (van Dusschoten, Moonen, de Jager, & Van As, 1996) followed by a CPMG train of 180° pulses. The T2-decay can thus be measured under variable diffusion-weighting conditions, enabling the apparent diffusion coefficient (ADC) of the water compartments corresponding to each T2 peak to be studied. Alternatively, the TE dependence could point to exchange with myelin water, which would make the TE dependence a potential marker for the degree of myelination (Lin et al., ISMRM 2017).
T1 spectra in single fibre populations have been reported as multi-exponential less consistently than T2, which has often thought to be caused by inter-compartmental mixing of water during the typically long measurement times. Nevertheless, multi-exponential T1 decay has been described e.g. (Does, Beaulieu, Allen, & Snyder, 1998). Hybrid correlation experiments have shown to increase the potential of disentangling unique T1 values per compartment (e.g., T1-T2 relaxometry) (Deoni, Rutt, Arun, Pierpaoli, & Jones, 2008). T1 compartmentalisation was also studied in combination with diffusion by adding inversion recovery (IR) to the DW-CPMG experiment (Andrews, Osborne, & Does, 2006). Recent work by Benjamini et al. (Benjamini & Basser, 2017) enabled the identification of 7 separate peaks by looking at the T2-Diffusion and T1-Diffusion space from ferret’s spinal cord, each associated with a separate type of tissue ranging from glial to axonal, neuronal and myelin water. The difficulty of implementation (hampered by the long scan times and the SAR constraints) limited the pursuit of this type of experiment until recently, when the availability of better hardware and faster sequences, as well as of methods for optimising data sampling (Benjamini & Basser, 2016) have made hybrid experiments feasible in vivo. De Santis et al., (De Santis, Assaf, Jeurissen, Jones, & Roebroeck, 2016a; De Santis, Barazany, Jones, & Assaf, 2016b) combine inversion recovery (IR) with dMRI for the purpose of mapping T1 (exploiting the IR preparation) along specific white matter tracts crossing within a single voxel (exploiting dMRI), assuming a mono-exponential T1 behaviour per fibre population. The feasibility of this approach in vivo was demonstrated by varying the inversion time (TI) for a moderate b-value of 1000 s/mm².

THE PROMISE

One major limiting factor in current in vivo acquisitions is the low SNR. The SNR fundamentally limits the resolution of distinguishing exponentials with different decay rates (Bertero, Brianzi, & Pike, 1999), and the resolution limit can therefore be expected to be worse for in vivo data having typically lower SNR than NMR data. Fortunately, there are several numerical procedures available that have shown to be able to reach the fundamental limit of distinguishing different exponentials (Istratov & Vyvenko, 1999).

The lower SNR of in vivo data further challenges the exponential analysis as the stability of numerical methods and their sensitivity to noise becomes an increasingly important issue. While averaging of signals improves the SNR, this is obviously at the cost of an increased acquisition time. Regularisation is commonly employed in NMR correlation experiments to make the inversion less ill-conditioned, with Tikhonov regularization being a popular method (Venkataramanan, Yi-Qiao Song, & Hurlimann, 2002)). Recent efforts to further optimise and regularise the estimation procedure include the use of compressed sensing (Bai, Cloninger, Czaja, & Basser, 2015) (Cloninger, Czaja, Bai, & Basser, 2014), marginal distributions (Benjamini & Basser, 2016) or spatial smoothness and constrained dictionary-based optimisation (Kim et al., 2017).
For *in vivo* human brain experiments, choosing a predefined form by explicitly modeling the contributions of different compartments is commonly performed to regularise the fit in diffusion microstructural imaging, and an extension of this to hybrid experiments is promising for the disentanglements of these compartments in terms of their diffusion and transverse relaxation properties (Veraart et al., 2017). However, recent findings show a correlation between parameters when the ‘standard’ model for diffusion without constraints (Section 2) was extended and applied in D-T2 hybrid experiments with limited b-values (up to 4000 s/mm$^2$) (Veraart et al., 2017). This raises concerns about the number of independent parameters that can be extracted from the available data (Reisert, Kellner, Dhital, Hennig, & Kiselev, 2017; Lampinen et al. 2017). When performing such D-T2 experiments on clinical systems, higher b-values can only be achieved at the expense of long TEs, thereby significantly limiting the shortest TE and therefore SNR that can be achieved.

Depending on the specific MR contrasts one aims to exploit, ultra-high magnetic field gradients have the ability to provide a better SNR. In T2-D experiments, it is possible to maintain a relatively short TE for a given diffusion sensitisation, improving the SNR. In IR-PGSE sequences for T1-D correlations a similar reasoning applies. At the same time, ultra-strong gradients provide the unique opportunity to sample a much larger part of the multidimensional experimental parameter space than has previously been possible *in vivo*. For T2-diffusion experiments, ultra-high magnetic field gradients can be used to sample higher b-values at shorter TEs (Figures 1 and 5). (Tax et al., 2017; Fan et al., 2017). Higher b-values could resolve the correlations between parameters (Veraart et al., 2017). Preliminary results show that high b-values are beneficial when trying to disentangle compartments (Tax et al., 2017) without fixing parameters to *a priori* values.

Reaching short TE is particularly important when attempting to capture myelin water contributions with very short T2 (MacKay et al., 1994) to the signal. Ultra-strong gradients combined with fast readout techniques, such as those mentioned in the Section 5 on high resolution, provide a promising outlook for characterising myelin water, thereby aiming at a more complete picture than has previously been possible with diffusion experiments. The wider TE range that can be sampled, however, requires a revisit of the signal modeling: if myelin is significantly contributing to the signal it should be included as an extra compartment in the model, and the role of exchange has to be re-examined. In addition, higher b-values with sufficient SNR in IR-PGSE experiments for D-T1, or even D-T1-T2 experiments enables a more detailed investigation of multi-component T1 relaxation per fibre population *in vivo*.

Future work will be directed towards carefully analysing the added information content in this previously inaccessible space, with the ultimate aim of developing joint tissue models. However, such an extensive sampling of the parameter space goes hand in hand with longer acquisition times, which makes it currently...
impractical for clinical translation. Compressive sensing (Bilgic et al., 2012; Doneva et al., 2010; Lustig, Donoho, Santos, & Pauly, 2008; Lustig, Donoho, & Pauly, 2007; Ning et al., 2015), “MR Fingerprinting” (Ma et al., 2013; Ma, Wright, Seiberlich, Gulani, & Griswold, 2014) and other techniques have the potential to capture the essential information while further reducing acquisition times.

**FIGURE 5**: $b$-value vs echo time (TE) experiment for maximum gradient amplitudes of 300 mT/m, allowing for high $b$-values (up to 7000 s/mm$^2$) and low TE (down to 47 ms). As a comparison, for the maximum $b$-value of 7000 s/mm$^2$ used here, the minimum TE would be 75 ms on a 80 mT/m system and 105 ms on a 40 mT/m system, as can be derived from Figure 1. (a) shows the raw signal and (b) the logarithm of the signal, to illustrate that there is still signal at the highest $b$-value and TE.
5. HIGH RESOLUTION DIFFUSION IMAGING AND TRACTOGRAPHY

THE CHALLENGE AND ITS IMPORTANCE

Until now, diffusion imaging methods, applied to the study of the living human brain, have been significantly limited by their own low spatial resolution. However, there are several important applications that could benefit from high resolution acquisitions. E.g. it has been shown that important cortical changes at the macro-scale level can occur in numerous psychiatric, neurodevelopmental and neurodegenerative disorders, (Goldman et al. 2009; Zielinski et al 2014, D. S. Andrews et al., 2017; D’Anna et al. 2016; Ecker et al., 2016), but it has never been possible to show the corresponding micro-structural attributes, and especially in which specific laminae any changes occur. Enabling ultra-high resolution dMRI in this manner could yield new insights into the pathophysiology of a wide range of disorders. At the same time, better spatial resolution not only can benefit the quantification of microstructure diffusion properties, but could also significantly improve the accuracy of tractography reconstructions. By increasing the spatial resolution, different fibre pathways can be better disentangled and new minor white matter connections will emerge besides the major white matter fasciculi (Heidemann, Anwander, Feiweier, Knösche, & Turner, 2012). Intra-lobar connections and u-shaped fibres (Catani et al., 2012; Catani et al. 2017) are essential short range connections that play an important role in distributing information across whole brain networks, but are currently significantly underestimated by most tractography and connectomics approaches.

BACKGROUND

Two of the main reasons for the low resolution of diffusion data are the intrinsic low SNR of MR diffusion data and the prevalence of single-shot acquisitions, selected due to the need to constrain total acquisition time, and to minimise motion artefacts (Turner, 1998; Turner et al., 1990). Although new post processing methods like super-resolution (e.g., Van Steenkiste et al., 2016), data interpolation (e.g., Dyrby et al., 2014) and image quality transfer techniques (see Section 10) are today emerging as alternative ways to mitigate the lack of spatial resolution, improving data quality during acquisition remains of critical importance. Recent attempts have tried to overcome these limitations by combining the benefits of higher spatial resolution and SNR of 7T data with stronger diffusion weightings from 3T data (Sotiropoulos et al., 2016; Sotiropoulos et al., 2017). The use of ultra-strong gradients could potentially allow the same high-resolution high-b-value data to be collected, but using only one scanner.

THE PROMISE

Indeed, with the use of powerful gradient systems it is now possible to achieve a significant reduction of TE per unit b-value (Setsompop et al., 2013) and, if combined with field cameras (Wilm et al., 2017; 2015) and
more echo-time efficient acquisition strategies (e.g. spiral readout), SNR can be further maximised and geometrical distortion significantly reduced. Compared to the use of weaker gradients at higher field (e.g., 70 mT/m at 7 Tesla), stronger gradients at lower field (e.g., 300 mT/m at 3 Tesla) allow the advantages of the longer T2, less T2* blurring and lower susceptibility distortion to be exploited. This in turn allows for higher spatial resolution and more anatomically accurate imaging and tractography reconstructions. Fan et al. (2017) recently proposed an interesting approach, dubbed ‘HIBRID’ (HIgh b-value and high Resolution Integrated Diffusion), that combines high resolution and high b-value data (collected with Connectom gradients), that allowed characterization of diffusion in the cortex without compromising white matter fiber information.

The adoption of new head RF-coils equipped with a high number of transmit and receiver channels can enable the use of high acceleration factors or the combination of multiple, advanced imaging techniques like multi-band (e.g., Larkman et al., 2001), parallel imaging methods (e.g. Larkman and Nunes, 2007), segmented-EPI (e.g. Frost et al., 2014); super-resolution (Van Steenkiste et al., 2016) and zoomed imaging e.g. Heidemann et al. 2012). One example of this new generation of imaging strategies is the Generalised Slice Dither Enhanced Resolution with Simultaneous Multislice (‘gSlider-SMS’) acquisition method (Setsompop, Ning & Rathi, 2016; Setsompop et al., 2017). This technique combines multi-band acceleration with slice-dithering acceleration, as well as parallel and zoomed imaging, which has allowed the acquisition of full brain in vivo diffusion datasets at isotropic resolution of 750um within 50 minutes with a b=1500 s/mm² and 128 directions. At this resolution, cortical regions that are normally heavily affected by partial voluming with CSF or subcortical white matter are better resolved and the whole human cortex can be characterised with a thickness of several voxels. This can allow practical and more meaningful quantification of the cortical diffusion properties compared to previous studies where diffusion measures were inevitably limited by cortical regions of only one or two voxels width. If in the near future this technique will be also combined with microstructure imaging this will potentially enable the analysis in the living human brain of individual cortical layers with a more direct interpretation of neurological and psychiatric disorders.

An example is shown in Figure 6, with isotropic resolution data of 780um, showing both the direction encoded fibre orientation (Pajevic and Pierpaoli, 1999) and the ‘anisotropic power’ obtained by taking the sum of the angular power spectrum of each spherical harmonic of even order \( l \geq 2 \) (Dell’Acqua et al. 2014).
Figure 6: Data obtained with the g-SLIDER technique (Setsompop et al., 2016; Setsompop et al., 2017) at 780 μm. The left panel shows the direction-encoded colour map using standard convention (Pavevic and Pierpaoli 1999), with the lower row showing zoomed-in versions of the areas highlighted in yellow. The right hand panel shows the anisotropic power, processed according to Dell’Acqua et al. (Proc ISMRM 2014; Anisotropic Power Maps: A diffusion contrast to reveal low anisotropy tissues from HARDI data. p.6396), obtained from the sum of the angular power spectrum of each spherical harmonic of even order \( l \geq 2 \), resulting from spherical deconvolution of the HARDI signal.

Furthermore, hybrid acquisitions can also help the integration of high resolution tractography data with microstructure modeling, potentially improving both tractography reconstructions by reducing confounding factor and allowing more refined “tractometry” (Bells et al. 2011) measurements based on-tract specific (Dell’Acqua, Simmons, Williams, & Catani, 2013b; Raffelt et al., 2012) and other microstructure indices.

Together with high-resolution acquisitions, low anatomical distortions can also make anatomical regions normally precluded by susceptibility distortion (e.g. orbital-frontal and temporal regions) accessible, and therefore offer more accurate tractography reconstructions.

We caution that strong gradients \textit{per se} are not a panacea for all the problems of tractography. While tractography today is a very powerful tool to segment the major white matter pathways there are still important limitations that prevent a reliable characterisation of their cortical connectivity. For example, Reveley et al. (2015) have pointed out with high resolution dMRI as much as “50% of the cortical surface was effectively inaccessible for long-range diffusion tracking because of dense white matter zones just beneath the infragranular layers of the cortex”. Stronger gradients neither do they directly solve the issue raised by Maier-Hein et al (2017), that "the geometry of many junctions in the simulated data set is too complex to be resolved by current tractography algorithms, even when given a perfect ground truth field of orientations".
Nevertheless, stronger gradients can indeed simplify the tractography problem. By pushing the spatial resolution of *in vivo* diffusion imaging below 1-mm voxel resolution and boosting SNR per unit time, this will doubtless lead to improvements in the fidelity of tractography reconstructions. Finally, high resolution in-vivo diffusion imaging performed with a ultra gradient system has the potential to generate data that can be directly compared with equivalent high b-value experiments performed, until now, only on animal models or ex-vivo tissues in preclinical systems. E.g. the characterisation of the laminar organization of the cortex or deep gray matter done on ex-vivo samples (Dell’Acqua et al. 2013b, Leuze et al. 2014, Bastiani et al. 2016) could be translated to map the cortical, subcortical and cerebellar organisation of the living human brain and provide a bridge to help the interpretation of clinical and research results obtained on conventional clinical scanner at much lower resolutions.

### 6. POST-MORTEM DIFFUSION IMAGING

**THE CHALLENGE AND ITS IMPORTANCE**

*Post mortem* MRI, and in particular diffusion imaging of whole brain specimens, has attracted interest in recent years for several reasons. Diffusion MRI has tremendous potential to contribute to our knowledge of neuroanatomy, including macroscopic connectivity, local fibre architecture, and patterns of microstructural features across the brain. However, *in vivo* anatomical investigations with dMRI are limited by achievable spatial resolution and diffusion contrast, with improvements in both domains incurring reduced SNR. Post-mortem MRI has potential to overcome this SNR challenge by enabling very long scan times (e.g. overnight or even several days). Indeed, dMRI with spatial resolution as high as 0.5x0.5x0.5mm³ has been demonstrated in whole, *post mortem* human brains (Foxley et al. 2016). A second motivation for scanning *post mortem* brains is the ability to directly compare dMRI data to microscopy in the same tissue. Diffusion modeling generally assumes a relationship between underlying fibre architecture and diffusion signal, but all too often these assumptions are not validated (Dell’Acqua, Bodi, Slater, Catani, & Modo, 2013a).

Microscopy can be used to map these relationships explicitly, including the contribution of specific tissue compartments (Budde and Frank 2012; Schilling et al., 2016; Mollink et al., 2017). *Post-mortem* dMRI has been used in neurosurgical applications, both to validate tractography against histological stains (Mollink et al., 2016) and to support contraindications for surgical procedures (McNab et al., 2009). A more unusual, but no less important, motivation for *post mortem* imaging is the ability to study unusual specimens. Neuroanatomical studies are generally limited to a very narrow range of species (certain rodents and a few primates). Nevertheless, *post mortem* brain collections host a plethora of species that could not conceivably be studied in a laboratory environment, many of which are too rare to dissect. *Post mortem* MRI enables non-invasive study of these specimens not only in terms of gross anatomy, but also in terms of connectivity, which is increasingly of interest in comparative anatomy (Mars et al., 2014). For example, a
recent study demonstrated the ability to reconstruct major white matter pathways from 100-year-old brains of an extinct species, the Tasmanian tiger (Berns & Ashwell, 2017), and another presented novel evidence for the pathways underpinning echolocation in dolphins (Berns et al., 2015).

In order to realize these goals, however, a number of challenges need to be overcome. As with other sections of this article, the primary challenge in realizing high spatial resolution is achieving high SNR. Although one can perform much longer scans than are possible in vivo, this in itself is insufficient to achieve major gains in spatial resolution. As with all MRI, improvements in spatial resolution come at an SNR cost that scales with the cube of voxel dimensions, while SNR scales with the square root of scan time. For example, if we aim to achieve a 0.5x0.5x0.5mm³ resolution without loss in SNR relative to a protocol achieving 1x1x1mm³ spatial resolution in 1 hour, we would have to scan for 64 hours (nearly 3 days).

There are additional challenges presented by post mortem tissue. MR-relevant properties change in unfavourable ways in post-mortem tissue due to death and fixation, with T1, T2 and diffusion coefficients all being significantly reduced (Shepherd, Thelwall, Stanisz, & Blackband, 2009). T2 is reduced to 50-76% of in vivo values (Pfefferbaum et al, 2004; Miller et al., 2011; Foxley et al., 2014), with the loss of T2 generally increasing with field strength from 1.5 to 7T. Reduced T2 is particularly problematic for conventional diffusion-weighted spin-echo (DW-SE) measurements, which require long TEs in order to achieve substantial contrast (b-value) (Miller et al., 2011). In addition, the diffusion coefficient is reduced to 20-50% of in vivo values in perfusion-fixed animal tissue (Sun et al, 2005; D’Arceuil et al., 2007), although measurements down to ~10% of in vivo values have been reported in human brains with long post-mortem intervals (Miller et al., 2011). These reductions in ADC dictate that a higher b-value is needed to obtain contrast equivalent to current in vivo protocols, requiring even longer TEs and reducing signal even more. T1 values are less impacted (40-60% of in vivo values) and in fact can improve image quality by enabling reduced repetition times. In general, these effects are heavily dependent on tissue preparation (Shepherd et al, 2009) and compounded by the long post-mortem interval between death and fixation in human brain specimens (Shepherd et al, 2009; D’Arceuil 2007b). These effects result in a net loss of SNR in post-mortem compared to in-vivo tissue (Fig, 7, left).

BACKGROUND
There are a number of different approaches that have been taken to address these challenges. Small brains and tissue samples can be accommodated in small-bore scanners that are designed for imaging rodents (D’Arceuil, Westmoreland, & de Crespigny, 2007; Dyrby et al., 2011). These systems benefit from stronger gradients than conventional clinical scanners, enabling short TEs to combat reduced T2. Moreover, small tissue samples can be soaked in buffer solution to wash out fixative, which is very effective in raising T2 values (which is predominantly caused by the short T2 of fixative itself) (Shepherd et al., 2009). Additional
benefit can be made by doping the buffer to also alter T1 values (D’Arceuil et al., 2007). Unfortunately, fluid takes weeks to months to fully permeate human brains, and incomplete penetration of buffer leads to problematic hydration boundaries that obscure the intended contrast (Miller et al., 2011). In general, these approaches are only feasible for whole brains up to the size of a macaque (D’Arceuil et al., 2007) or pig (Dyrby et al., 2011). Early proof-of-concept for diffusion imaging of whole post-mortem human brains highlighted the potential to provide a direct validation against histopathology in disease, but generally suffered from modest image quality (Larsson et al 2004; Englund et al 2004; Pfefferbaum et al 2004).

It has been demonstrated that diffusion imaging of whole human brains can be achieved with minor modifications to a DW-SE sequence (in particular, using a 3D segmented readout) (Miller et al., 2011). However, it has also been demonstrated that significant gains in SNR and contrast can be obtained by moving away from conventional DW-SE sequences. One approach that has been explored in depth is a variant of steady-state free precession (SSFP), which has been demonstrated to provide substantial improvements in quality of tractography (Miller, McNab, Jbabdi, & Douaud, 2012). In theory, DW-SSFP achieves greater efficiency than any other technique by re-using the signal created in one excitation over several acquisition blocks, decoupling contrast from TE. Unfortunately, DW-SSFP is also extremely gradient intensive, leading to heating constraints. Moreover, the signal has a complicated dependence on T1, T2, diffusion coefficient and acquisition parameters, and does not have a well-defined diffusion time (McNab & Miller, 2010). Another alternative that has recently been revisited for post mortem imaging is diffusion-weighted stimulated echoes with EPI readout (DW-STE) (Rane, Nair and Duong, 2010; Lundell, Alexander, & Dyrby, 2014; Merboldt, Hänicke, & Frahm, 1991). This sequence has many of the benefits of DW-SSFP, including decoupling of the b-value from the TE, while retaining a well-defined diffusion time and simple signal dependence. The main disadvantage of DW-STE is loss of 50% of the signal in the formation of the echo. Finally, ultra-high field strength (7 Tesla and above) has been demonstrated to provide further improvements for both DW-SSFP (Foxley et al., 2014) and DW-STE (Fritz 2017), although RF transmit inhomogeneity presents a further challenge at these field strengths.

THE PROMISE

As outlined above, the primary need for alternate acquisition approaches in post mortem brain imaging relates to the limitations of current clinical scanners, most specifically the relatively low maximum gradient strength. In pre-clinical (rodent) scanners, the high available gradient strength (typically 300-1000 mT/m) enables the use of DW-SE with high SNR. This enables the acquisition of data that are maximally compatible with in vivo approaches, avoids the 50% signal loss of DW-STE and has a simple signal dependence with well-defined diffusion time. The availability of similar gradient strength in a scanner large enough to accommodate whole, human brains could enable anatomical investigations of unprecedented quality,
providing high spatial resolution across the entire brain and in larger cohorts than can be studied with other methods. Figure 7 shows predicted improvements in SNR for post-mortem dMRI using spin-echo acquisition with 300 mT/m compared to 80 mT/m.

By overcoming the primary challenges associated with SNR, these scanners should also enable investigation of more subtle diffusion-based contrast in post-mortem brains, such as the microstructural features. This enables validation of these features against direct measurements based on microscopy (Mollink et al., 2017), comparison between different microstructural models with a gold standard for truth (Schilling et al., 2018), and critical assessment of the current shortcomings in dMRI modeling (Reveley et al. 2015). Finally, the very nature of such datasets with unusually high quality in multiple dimensions (spatial resolution, contrast and linkage to microscopy) has promise to provide novel findings. In recent years, post-mortem dMRI has enabled new insights into neurology (McNab et al., 2009; Kolasinski et al., 2013; Modo et al., 2016); demonstrated the ability to visualize and segment structures in human brain that are currently inaccessible in vivo (Augustinack et al., 2010; Aggarwal et al., 2013; Augustinack et al., 2014); provided the
ability to detect gray matter microarchitecture features (Leuze et al., 2012; Kleinnijenhuis et al., 2012; Aggarwal et al., 2015; Kleinnijenhuis et al., 2015); and established a unique toolkit for comparative anatomy of a broad range of species (Jbabdi et al., 2013; Mars et al., 2014; Berns et al., 2015; Berns & Ashwell, 2017). With improvements in both hardware and the approaches outlined above, post-mortem MRI could provide unique and rich insight into neuroanatomy.

7. DIFFUSION-WEIGHTED MRS

THE CHALLENGE AND ITS IMPORTANCE

While dMRI has predominantly focused on the diffusion of water molecules in tissue, the problem is that water is ubiquitous in tissue and exists in all compartments (e.g., intra-cellular and extra-cellular). It is this fundamental fact that makes it difficult to disentangle different tissue subcompartments (see Sections 1 and 2). Diffusion-weighted Magnetic Resonance Spectroscopy (DW-MRS) uniquely enables specific characterization of tissues such as the brain in vivo by quantifying the diffusion properties of MR-observable metabolites. Many brain metabolites are predominantly intracellular, and some of them are preferentially localized in specific brain cell populations (e.g., neurons and glia). Since many developmental processes, for instance plasticity and aging, or pathological processes, such as neurological diseases, are characterized by modulations of specific cellular types and their microstructures, and since water signals are not necessarily representative of any specific compartment, metabolite signals can serve as biomarkers with enhanced specificity (Palombo, Shemesh, Ronen, & Valette, 2017b). However, DW-MRS measurements are extremely challenging, from the acquisition to the analysis and quantification stages. In vivo DW-MRS pulse sequences must achieve several goals simultaneously: (1) adequate diffusion weighting; (2) efficient and accurate signal localization to the volume of interest (i.e., MRS voxel, Figure 8-a); and (3) robust acquisition of a free induction decay (FID) to obtain the chemical shift spectral dimension.

Here, we briefly discuss the importance of addressing point (1) by developing ultra-high gradient systems, and therefore the applications that would be made possible in humans in vivo. Since metabolite apparent diffusion coefficient, or ADC, is about 5-10 times smaller than water (0.15-0.25 μm²/ms versus 0.7-1.5 μm²/ms for water in tissue (Ellegood, Hanstock, & Beaulieu, 2011)), larger b-values are required for metabolite signals in order to attenuate the signal (~exp[- b ADC]) to the same extent. For a conventional experiment, like the measurement of the ADCs of metabolites, which requires ‘moderate’ b values, it is necessary to have a gradient strength at least 2-3 times stronger than that typically used for water (i.e. at b~1000 s/mm²). Moreover, recent preclinical studies (Assaf & Cohen, 1999; Cohen & Assaf, 2002; Ellegood, McKay, Hanstock, & Beaulieu, 2007; Kroenke, Ackerman, & Yablonskiy, 2004; Ligneul, Palombo, & Valette,
2017; Palombo, Ligneul, Hernandez-Garzon, & Valette, 2017a; Yablonskiy & Sukstanskii, 2010) have shown that when the ultra-high b-value domain (i.e., $b > 5,000 \text{s/mm}^2$) or ultra-short diffusion time domain (i.e. < 5 ms) are accessible, DW-MRS techniques offer huge potential for the non-invasive estimation of brain cell-specific microstructural properties (e.g., microscopic anisotropy, size and surface-to-volume ratio of neuronal and glial cell projections) that are out of reach of existing MRI techniques that are based solely on water diffusion.

**BACKGROUND**

In the early days of DW-MRS, due to the relatively low maximal gradient strength available on MRI scanners, the stimulated echo-based DW-STEAM pulse sequence (called for simplicity PGSTE) (Moonen, van Zijl, Le Bihan, & DesPres, 1990; Posse, Cuenod, & Le Bihan, 1993) (Figure 8-b) was the most convenient way to reach sufficiently high b-values *in vivo* (Ellegood et al., 2011) without excessive signal loss due to transverse relaxation, since during the mixing time, magnetization relaxes according to $T_1$ (typically 1–1.5 s for metabolites at 3 T (Mlynárik, Gruber, & Moser, 2001)). While this approach is advantageous for many reasons (Cao & Wu, 2017; Nicolay, Braun, Graaf, Dijkhuizen, & Kruiskamp, 2001; Palombo, Shemesh, Ronen, & Valette, 2017b; Ronen & Valette, 2007), the most significant drawback at the moderate magnetic field strength of clinical scanners (typically 3 T) is the loss of half the magnetization (see also Section 6). Therefore, a spin echo sequence might be preferred, as sensitivity is one of the major issues of MRS in general, and of DW-MRS in particular. Nevertheless, in the spin echo sequence, diffusion time and TE are coupled. This, together with signal loss due to $T_2$ relaxation (typically 150-300 ms for metabolites at 3 T (Mlynárik et al., 2001)) occurring during the TE, imposes an upper limit on the longest diffusion time achievable. Hence, in order to translate advanced DW-MRS techniques to *in vivo* measurements in humans using clinical 3 T scanners, and to achieve good sensitivity and very high b-values, it is necessary to have access to a gradient strength that is 5-10 times stronger than that typically used for water.

The first spin echo DW-MRS sequences were based on PRESS (point-resolved spectroscopy), and were performed on preclinical systems with gradient strengths sufficient to reach sufficiently high b-values (Dijkhuizen, de Graaf, Tulleken, & Nicolay, 1999; van der Toorn, Dijkhuizen, Tulleken, & Nicolay, 1996). The sequence was subsequently modified for clinical implementation at 1.5 T by applying a bipolar gradient scheme (Kroenke et al., 2004), thus taking advantage of the twice-refocused spin echo sequence to increase maximal achievable diffusion-weighting. As with DT-MRI, it is also possible to assess the macroscopic anisotropy of metabolite diffusion in the brain tissue by using a minimum of six diffusion gradient directions. This technique is called diffusion tensor spectroscopy (DT-MRS, also known as DTS), and permits the estimation of the elements of the macroscopic diffusion tensor of typical MRS-visible metabolites, such as NAA (N-acetyl-aspartate), tCr (total Creatine) and tCho (total Choline). A DW-PRESS sequence with a specific gradient scheme has also been proposed to provide single-shot isotropic diffusion-
weighting (i.e., weighting by the trace of diffusion tensor) (de Graaf, Braun, & Nicolay, 2001).

Several metabolite ADCs have been measured in the healthy human brain (Ellegood et al., 2011; Ellegood, Hanstock, & Beaulieu, 2006; Posse et al., 1993) with ADC values for NAA, tCr and tCho being around 0.15-0.20 μ m²/μs. DT-MRS has also been used to assess the anisotropy of metabolite diffusion in human brain (Ellegood et al., 2006; Upadhyay, Hallock, Ducros, Kim, & Ronen, 2008). FA values of 0.48–0.72 were reported for NAA (Ellegood et al., 2006; Upadhyay et al., 2008), 0.56–0.73 for tCr (Ellegood et al., 2006) and 0.59–0.74 for tCho (Ellegood et al., 2006) in white matter regions such as the corpus callosum, corticospinal tract and arcuate fasciculus. A first examination of cortical gray matter (GM) yielded an unexpectedly high degree of anisotropy (FA=0.53–0.79) for all three metabolites, similar to those seen in white matter (Ellegood et al., 2006). However, a later study by Ellegood et al. (Ellegood et al., 2011), showed that the high FA measured in GM was an artefact caused by the poor two-point slope estimation of the signal intensity versus b-value when using sub optimal b-values. Using a maximum b-value of ~5 ms/ μ m², instead of the much lower 1.8 ms/ μ m² used in (Ellegood et al., 2006), lower FA values of 0.25, 0.30 and 0.28 for NAA, tCr and tCho, respectively, were indeed estimated in human occipital GM.

Numerous valuable studies have been recently using DW-MRS on clinical scanners (mostly at 3 T and 7 T, and maximum gradient strength of 20-60 mT/m) to measure metabolites ADC and DT, and to provide complementary information to various MRI metrics (DTI metrics, T₁, T₂, MTR) for an improved diagnosis and prognosis of different brain pathologies, (for a more comprehensive review, we refer the reader to (Cao & Wu, 2017; Nicolay et al., 2001; Palombo, Shemesh, Ronen, & Valette, 2017b; Ronen & Valette, 2007)). So far, DW-MRS has been applied to study brain activity (Branzoli, Techawiboonwong, Kan, Webb, & Ronen, 2013), cerebral ischemia (Harada et al., 2002; Zheng, Liu, Fang, Wang, & Zhang, 2012), cerebral tumors (Zheng et al., 2012), normal aging (Zheng et al., 2012) (Branzoli et al., 2013), schizophrenia (Du et al., 2013) and multiple sclerosis (Wood et al., 2017; 2012) in human. For example, metabolite DT-MRS was used together with water-based DT-MRI to investigate axonal damage and to distinguish axonopathy from other processes such as inflammation, edema, demyelination and gliosis in multiple sclerosis (Du et al., 2013). Glial and axonal changes in neuropsychiatric systemic lupus erythematosus were assessed using metabolite DW-MRS, showing that intracellular alterations, and in particular changes in glia, as evidenced by an increase in the ADC of tCho and tCr, correlate with systemic lupus erythematosus activity (Ercan et al., 2016).

THE PROMISE
The possibility of employing stronger gradient systems on human scanners (McNab et al., 2013; Setsompop et al. 2013), would help to improve DW-MRS measurements and enable the translation to humans of cutting-edge techniques that, at the moment, are limited to preclinical applications due to the weaker
gradients available on clinical systems. It has been shown that metabolites ADC and DT metrics can already be estimated in vivo in human brain using clinical scanners with weaker gradients (~40-60 mT/m). However, in order to have robust and reliable estimates, reduced artefacts and good sensitivity at moderate magnetic fields (~3 T), it may be extremely advantageous to have strong gradient systems to reach high b values (> 3000 s/mm\(^2\)), while keeping the TE and the gradient duration as short as possible (Ellegood et al., 2011; Wood et al., 2015). Moreover, stronger gradient systems enable shorter diffusion times, which in turn reduces the variability of the measurement due to fluctuations in the signal due to the heartbeat (i.e., blood flow peaks during the diffusion time), which can become significant for b > 3000 s/mm\(^2\). While cardiac gating could, in principle, be used, it becomes ineffective for very long diffusion times (more than hundreds of ms).

While ADC and DT derived metrics from non-water metabolites can be very useful in some clinical applications, they are only indirect reporters of the underpinning tissue microstructure (see Section 3 for discussion on the lack of specificity). Earlier strategies for more directly quantifying tissue microstructure from metabolite DW-MRS have involved q-space MRS (Assaf & Cohen, 1998; Cohen & Assaf, 2002), the quantification of the average propagator and some of its time-dependent properties. Recent work has described a first attempt to model metabolites diffusion at high q/b values, taking into account the cellular structure, and to estimate the microscopic anisotropy and diameter of cellular projections (Kroenke et al., 2004; Palombo, Ligneul, & Valette, 2016a; Yablonskiy & Sukstanskii, 2010). These preliminary studies involved ultra-high diffusion weightings, obtained by employing ultra-strong diffusion gradient systems, and so to date have only been performed in vivo in preclinical applications. Nevertheless, these pioneering studies at ultra-high q/b values have paved the way for some exciting possibilities enabled by implementing DW-MRS with ultra-strong diffusion gradients. Recent in vivo studies in mouse brain (Ligneul et al. 2016; Ligneul et al. 2017) and numerical simulations (Palombo, Ligneul, Hernandez-Garzon, & Valette, 2017a) showed that measurements of metabolites’ diffusion-weighted signal involving high gradient strength (i.e., PGSTE at b>5,000 s/mm\(^2\) and oscillating-gradients spin echo, OGSE, at high frequencies and high b values: f >100Hz and b>1500. s/mm\(^2\)) help to address some of the limitations of water-based DW-MRI techniques (Figure 8c). These limitations include the lack of cell-specificity, and poor sensitivity to fine cellular microstructure, such as the density of small cell fibres (diameter <3 μm) or neuronal spines. DW-MRS at high gradient strength would be a first necessary step towards the definition of new cell-specific biomarkers of the alteration of tissue microstructure due to neurodegeneration. Preliminary in vivo work using the APP/PS1/tauP301L mouse model of Alzheimer’s disease (Ligneul et al. 2016), and the ciliary neurotrophic factor induced mouse model (Ligneul et al. 2017) showed that it is possible to use DW-MRS at ultra-high b-values to measure changes in the size of neuropsil due to the induced neurodegenerative process in a cell-specific fashion, distinguishing between glial and neuronal cell alterations. Combining high b values PGSTE and high frequency OGSE measurements, it would also be possible to measure changes in
cellular projections’ surface-to-volume ratio and dendritic spine density (Palombo, Ligneul, Hernandez-Garzon, & Valette, 2017a), which may be extremely useful in the diagnosis and prognosis of pathologies such as autism (Hutsler & Zhang, 2010) and essential tremor (Louis et al., 2014). As shown in Figure 8c, the Connectom scanner’s ultra-strong gradient hardware would help filling the gap between preclinical and clinical scanners, opening a unique possibility to start investigating, in vivo in humans, a region at high b values and high frequencies (green areas in Figure 8c), currently inaccessible using clinical scanner and similar PGSTE or OGSE protocols.

Figure 8
Example of advanced DW-MRS experiments currently achievable on preclinical scanners only, due to a necessary maximum gradient strength, $G_{\text{max}} > 100$ mT/m. (a) Typical size of the MRS voxel ($6 \times 2.5 \times 5$ mm$^3$) required to reach adequate signal-to-noise ratio of the diffusion-weighted signal of the relevant brain metabolites, like NAA and Myo-inositol. (b) Advanced DW-MRS experiments require optimized PGSTE and OGSE sequences and ultra-strong gradient hardware in order to reach high b values (> 5,000 s/mm$^2$) and high frequencies ($w > 100$ Hz). Adapted from Reynaud (2017). (c) Examples of DW-MRS signal decays as a function of b and ADC as a function of $w^{1/2}$ for two different metabolites: NAA (mostly compartmentalized in neurons) and Myo-inositol (mostly compartmentalized in glia), obtained by using PGSTE and OGSE sequences with $G_{\text{max}}$=800 mT/m and respectively $TE/b_{\text{max}}$=32 ms / 60,000 s/mm$^2$ and $TE/b_{\text{max}}$=60 ms / 1,200 s/mm$^2$ (adapted from [Palombo et al. (2017a)]. Coloured areas represent the regions accessible by current clinical scanners with $G_{\text{max}}$=80 mT/m (red areas); the Connectom scanner with $G_{\text{max}}$=300 mT/m (green areas) and a preclinical scanner with $G_{\text{max}}$=800 mT/m (blue areas), using the same $TE/b_{\text{max}}$ values reported in (b).

Although much of the development of DW-MRS has been done on preclinical systems, it is clear that the
technique may become a very powerful and useful clinical diagnostic tool. Systems operating at moderate static magnetic field (e.g. 3 T), equipped with strong gradient units (>100 mT/m) are ideal for investigating short-range features of brain tissue microstructure using DW-MRS at intermediate diffusion times (<200 ms) and high b values (>5,000 s/mm$^2$), or ultra-short diffusion times (<5 ms) and moderate b values (b>1500 s/mm$^2$) (Figure 8-c). As previously discussed, this would provide enough sensitivity to quantify changes in cellular features, like the diameter and surface-to-volume ratio of cellular projections, and improve the early diagnosis of neurodegenerative disease (Valette et al. 2018). On the other hand, an important constraint arising from the low concentration of metabolites, compared to water (~10$^4$ times less), is that the spatial resolution i.e., the voxel sizes, and acquisition time are generally much larger than those of DW-MRI (Figure 8-a), representing a major hurdle to the clinical applicability of DW-MRS. Higher static magnetic fields (e.g., 7 T) provide improved spatial resolution and reduced acquisition time, facilitating the characterisation of long-range features of brain tissue microstructure using DW-MRS at long diffusion times (>200 ms) and moderate b values (<5000 s/mm$^2$)(Najac, Branzoli, Ronen, & Valette, 2016; Palombo et al., 2016b). However, the weaker gradients available on these systems (<100 mT/m), and the challenges of achieving a well-shimmed homogenous field, remain an obstacle to performing high b-value investigations to measure changes in the cellular short-range morphology.

Hence, there is still much to be done in terms of acquisition methodology. The critical steps towards the full realisation of the potential of DW-MRS in the clinic are the development of: 1) better and stronger gradient systems; 2) more sensitive RF coils; 3) higher static magnetic fields.

8. CANCER IMAGING / TUMOUR CHARACTERIZATION

THE CHALLENGE AND ITS IMPORTANCE
The dMRI signal is potentially sensitive to a number of features of the tumour microenvironment that are relevant to diagnosis and assessment of treatment response. The promise of specific measurements of such properties is that they provide direct insight into key aspects of pathology that are only indirectly and, therefore potentially ambiguously, reflected using heuristic parameterisations such as ADC or kurtosis (e.g., Bourne, 2015; K. M. O’Connor, Barest, Moritani, Sakai, & Mian, 2013). Getting access to such information over the whole organ of interest, (rather than measurements available from a needle biopsy, for example), would pay huge dividends for diagnosis, longitudinal monitoring and evaluation of treatment response.
**BACKGROUND**

As illustrated in Figure 9, different models and acquisitions have been used to estimate or evaluate the sensitivity to: cell size and intracellular volume fraction (Panagiotaki et al., 2014, 2015; Reynaud et al., 2016; Jiang et al., 2017), nuclear-to-cell volume ratio (Xu et al., 2009), water exchange (Lampinen et al., 2017b), and vascular volume fraction (Panagiotaki et al, 2014, 2015). Intra- and extra-cellular diffusion coefficients have been either estimated (Reynaud et al., 2016; Jiang et al., 2017) or fixed (Panagiotaki et al, 2014, 2015), and the pseudo-diffusion coefficient associated with blood flow has been estimated (Panagiotaki et al, 2014).

![Figure 9. Literature examples of microstructural imaging in cancer. (a) VERDICT yields estimates of compartment volume fractions, cell size, and the pseudo-diffusion coefficient in preclinical colorectal xenografts, demonstrating sensitivity to treatment-induced changes (Panagiotaki et al., 2014). (b) IMPULSED combines PGSE and OGSE measurements to estimate cell size, and compartment volume fractions and diffusivities in preclinical colorectal xenografts, with dMRI-derived cellularities correlating with histology (Jiang et al., 2017). (c) POMACE models the time-dependent diffusivity from PGSE and OGSE data to estimate microstructural properties in preclinical gliomas, with cell sizes consistent with independent measurements (Reynaud et al., 2017). (d) FEXI shows differences in the apparent exchange rate between meningiomas and astrocytomas (Lampinen et al., 2017b).](image-url)
Such additional pathological specificity then has the potential to provide robust biomarkers (O’Connor et al., 2017) for a range of potential applications, including:

- **Distinguishing tumour/non-tumour tissue:** Often tumours exhibit alterations in cell packing that can potentially be identified using dMRI. Vasculature is also typically abnormal in cancer.

- **Distinguishing between tumours of different grades and characterizing tumour extent:** As tumours progress, they exhibit phenomena such as infiltration of surrounding tissue, leading to a mixture of healthy and tumour cell types at the tumour margins, in addition to internal heterogeneity of cellularity and vascularity.

- **Predicting treatment response:** Microstructural characterisation of pre-treatment tumours may aid interpretation of, or offer improved sensitivity over, the observation of higher baseline ADC being linked to poorer treatment response (Tam et al., 2013), for example by providing a means of attributing differences in ADC to differences in necrosis or cellularity.

- **Early assessment of response:** Cell stress occurring early during treatment induces changes in the flux of water across cell membranes, leading to changes in cell size and packing density. As cell death ensues, further changes in these parameters may occur in advance of gross tumour size change.

- **Distinguishing pseudo-progression from true progression:** Apparent tumour growth (typically in the context of brain tumours) following treatment can be associated with changes in contrast enhancement caused by damage to the tumour tissue and potentially to surrounding non-tumourous tissue. This increase in apparent tumour volume would classically be associated with tumour progression but in reality is a sign of response to treatment. The associated cell death, however, is reflected in changes in cell packing, which should be discernible with dMRI.

**THE PROMISE**

The above measurements require the fitting of models of water compartmentalisation and exchange to data obtained at a range of diffusion times and gradient strengths. As noted several times already, current clinical MRI systems lack the gradient performance to enable broad coverage of the $G$-$\Delta$ space (see Figure 1b), leading to model parameter imprecision, poor model fitting and the use of overly simplistic models.

As is the case for all microstructure imaging using dMRI, access to higher gradient strengths enables higher SNR/shorter TE for a given $b$-value, as well as larger $b$-values for shorter diffusion times. Although differences in rise times for 80 mT/m and 300 mT/m gradients mean that a wider range of $\Delta$ values are in principle achievable with the former, the lower $\Delta$ values suffer from very low $b$-values (e.g., 187 s/mm$^2$ for $\delta=5$ ms, $\Delta=18$ ms, $G=80$ mT/m, rise time=0.4 ms; see Figure 1b). 300 mT/m gradients make shorter diffusion times ($\Delta \sim 24$ ms) more practical, as large $b$-values are still achievable (e.g. 8693 s/mm$^2$ for $\delta=8$ ms, $\Delta=24$ ms, $G=300$ mT/m, rise time=3.6 ms; see Figure 9). As the cell sizes in cancer are typically greater than the sizes of axonal features of interest in most neurological dMRI, with the significant exception of increased SNR,
the advantages of higher $G$ are not necessarily as clear for modeling of basic cell dimensions. However, heuristic quantification of the dMRI signal in cancer can also benefit from better gradient performance. For example, recent investigation of diffusion time-dependent ADC in prostate tumours suggests shorter $\Delta$-values yield greater contrast between high- and low-grade tumours (Lemberskiy et al., 2017).

**Cellular and sub-cellular structure**

Combining stronger gradients with OGSE gradient waveforms increases sensitivity to differences in the ratio of nuclear volume to cell volume (Xu, Does, & Gore, 2009), the relevance of which is that nuclear size is related to tumour grade, and can change with therapy. However, the rise time requirements imposed in order to meet safety constraints (see Section 11 below) limits the OGSE frequencies which can be practically achieved (~70 Hz for a 3.6 ms rise time). This suggests that the main benefit of stronger gradients for OGSE will be the ability to achieve higher $b$-values for low frequencies, as opposed to providing access to higher frequencies. Alternative gradient waveforms, including those used for spherical tensor encoding (Eriksson et al., 2013; Lasić et al., 2014) can also provide deeper insight into tissue than DT-MRI, characterizing differences in microscopic anisotropy (Szczepankiewicz et al., 2015; 2016). The relatively long duration of the waveforms required to play out a spherical tensor encoding mean that SNR is compromised on commonly-available hardware, but implementation with stronger gradients reduces the duration for which each ‘lobe’ needs to be played out. In turn, this conveys a reduction in TE and a concomitant boost in SNR (approximately 50% at 300 mT/m compared to 80 mT/m).

**Beyond cellularity – vasculature and other microstructure**

Flow and vascular volume fraction have been estimated using model-based approaches for interrogating the dMRI signal in tumours. The simplest approach is to employ the intra-voxel incoherent motion (IVIM) model (Ilima & Le Bihan, 2016; Koh, Collins, & Orton, 2011; Le Bihan et al., 1986; Marzi, Piludu, & Vidiri, 2013), although more sophisticated approaches such as VERDICT (Panagiotaki et al., 2014) offer the promise of greater specificity of signal characterisation in terms of the relative contribution of different cellular architecture and vasculature. An alternative pragmatic approach is to identify the contribution of the vasculature by combining flow- and non-flow-compensated acquisitions (Wetscherek et al., 2015; Ahlgren et al., 2016). Finally, beyond characterising the intracellular and vascular spaces, dMRI can provide information about extracellular matrix components such as collagen1 (Col1) fibres, which facilitate macromolecular transport within tissue and dissemination of cancer cells. It has already been shown that metrics derived from DT-MRI (FA / MD) are sensitive to differences in Col1 fibres (Kakkad et al., 2016).

**Measurement validation**
Measurement validation is obviously an essential step in the development of imaging biomarkers for cancer (O’Connor et al., 2017). Ultra-strong gradients are not currently available to the clinical cancer imaging community, so methods that work only with access to such systems are unlikely to generate useful biomarkers (although see Section 10 on Image Quality Transfer). As such, the benefits of increased gradient strengths for clinical cancer imaging are currently likely to be indirect, derived from using such high performance systems to provide validation of methods that can be implemented on lower-performance systems. Indeed, this may be the only way to demonstrate equivalence between biomarker information available on human MRI systems, and on preclinical scanners. With preclinical systems, although gradient strength is high, the field strength is typically also larger, and there are differences in coils, relaxation times, and resolution. The ultra-strong gradient system, coupled with a 3T human system, allows us to match other (non-diffusion parameters) to clinical situations, which is useful particularly when histological validation is unavailable or inadequate. Finally, we recognize of course that ultra-strong gradient hardware is not yet widely available, so the immediate benefits for cancer care will be indirect. However, in Section 10, we speculate as to how measurements on such systems can inform more ‘routine’ clinical assessments on more standard hardware.

9. FUNCTIONAL DIFFUSION IMAGING

THE CHALLENGE AND ITS IMPORTANCE

Measuring neuronal activity non-invasively is fundamental to understanding the brain and its dysfunction occurring in disease. Direct measurements of the effects of neuronal current can be achieved with neurophysiological techniques, such as electro-encephalography (EEG) and magneto-encephalography (MEG). These methods, however, suffer from limited spatial resolution, and their sensitivity is typically restricted to cortical activity. Functional MRI (fMRI) based on blood-oxygenation level-dependent (BOLD) contrast on the other hand provides a very indirect estimation of neural activation through related haemodynamic and metabolic changes (neurovascular coupling) (Logothetis & Wandell, 2004). One of the consequences is that the BOLD signal originates from both capillaries close to the site of neuronal activity, and larger vessels or distant draining veins (Lai, Glover, & Haacke, 2000; Olman, Inati, & Heeger, 2007; Turner, 2001; 2002) thus limiting the spatial specificity of BOLD. In addition, the BOLD response will depend on the vascular density, which varies between brain areas (Vigneau-Roy, Bernier, Descoteaux, & Whittingstall, 2013), and might be compromised in conditions such as stroke. The neurovascular coupling also imposes constraints on the temporal resolution of the BOLD signal. The haemodynamic response function (HRF) that defines BOLD signal change is characterised by 1-2 second delay between the onset of neural activity and the time required for blood to travel to the active region, and a time to peak of
approximately 6-12 seconds, with large inter-subject variations. A method able to combine the high spatial resolution of MRI with a more direct measurement of brain activity would offer attractive advantages. As cell swelling has been reported as one of the physiological processes accompanying neuronal activation (Tasaki & Iwasa, 1982b; 1982a), diffusion-weighted (DW) MRI has been proposed as a potential candidate for detecting temporal patterns in cell physiology (Aso et al., 2009; Latour, Hasegawa, Formato, Fisher, & Sotak, 1994; Le Bihan et al., 2012; Le Bihan, Urayama, Aso, Hanakawa, & Fukuyama, 2006).

BACKGROUND

The idea that DW MRI could be sensitive to functional activation has been mooted for several years (le Bihan et al. 2012). Initial attempts to exploit this sensitivity built upon the concept of intravoxel incoherent motion (IVIM) reflecting the displacement of water molecules in the pseudo-randomly orientated capillary network at low b-values (less than 100-200 s/mm$^2$) (Turner et al., 1990). The observation of a transient reduction in ADC of the occipital cortex during visual stimulation at high b value (~1400 s/mm$^2$), however, pointed at a different mechanism (Darquié, Poline, Poupon, Saint-Jalmes, & Le Bihan, 2001). This was initially interpreted as a consequence of cortical cell swelling resulting in increased tortuosity of the extracellular space. In 2006, LeBihan et al (Le Bihan et al., 2006) explored the changes in ADC occurring at a range of b-values, demonstrating that the effect is enlarged at large b-values (>1500 s/mm$^2$) and seems to precede the haemodynamic response measured by blood oxygenation level dependent (BOLD) contrast. Based on this observation they advanced the hypothesis that changes in DW MRI reflect early biophysical events associated with neuronal activation and thus provide a more direct measure than BOLD that, as discussed above, has a large contribution from large and distant draining veins.

Le Bihan et al (2006) proposed a quantitative interpretation by means of the biphasic model of water diffusion in the gray matter: according to such model, the empirically observed slow (SDP) and fast (FDP) diffusion pools correspond to both, extra and intracellular spaces, with SDP originating from molecules hindered by their close proximity to cell membranes (Le Bihan et al., 2006). Swelling of the cell membrane may thus result in the expansion of the SDP volume fraction (and a consequent reduction in the FDP volume fraction) resulting in a decrease in water diffusion. Fitting this model to the data from 6 participants yielded a constant increase of approximately 0.6% in the SDP volume fraction. This model, however, is in conflict with the extremely fast exchange rates expected between these water fractions (Cowan et al., 2005; Jungwirth, 2011).

The hypothesis that the diffusion fMRI signal originates solely from neuronal activity has, been challenged by other authors. Miller et al. (Miller et al., 2007) used hypercapnia, a condition known to induce vascular changes measurable with BOLD, in the absence of neuronal signal. They showed that the amplitude of the change in diffusion fMRI was comparable to that obtained with a visual stimulation, and failed to replicate
the precedence of the diffusion fMRI response to stimulation compared to BOLD. Other studies have also failed to replicate the b-factor dependency (Autio & Roberts, 2014). Finally, a certain degree of anisotropy of the hypercapnia-induced signal changes has been reported, which makes it less likely to be related to cell swelling in the grey matter (Miller et al., 2007). As an alternative explanation, Pampel et al. (Pampel, Jochimsen, & Möller, 2010) have suggested that background gradients resulting from the difference in magnetic susceptibility between water and blood might interact with the diffusion gradients thus causing the observed changes in ADC.

THE PROMISE

The exact source of diffusion fMRI contrast has yet to be determined. Part of the difficulty in designing appropriate experiments lies in the inherently low SNR of the diffusion fMRI signal, which limits the range of feasible experimental manipulations. Such a low SNR is one of the biggest concerns about diffusion fMRI, as even the simplest experiment requires the acquisition of very large datasets (R. J. Williams, Reutens, & Hocking, 2016). The advent of faster sequences such as multiband imaging, aka simultaneous multi-slice (Moeller et al., 2010), can help, but a significant boost in SNR is the only way to make these experiments practical. Moving to higher field strength increases the b=0 signal, but due to the shortening of T2 yields a limited benefit at high b-values. A real improvement can only be achieved with strategies that allow shorter TEs to be used with heavily diffusion-weighted sequences, thus mitigating the signal loss due to T2 decay. Multiple channel receiver coils (e.g., 128-channels), with their better SNR compared to 20- or 32-channel coils, could help, especially if combined with parallel imaging schemes that reduce the TE. However, the acceleration factor achievable is limited (Larkman and Nunes, 2007). The use of ultra-large gradients is thus an efficient way of increasing the SNR in DW-MRI. In terms of clarifying the source of the diffusion fMRI signal change, a high gradient performance-scanner can help by addressing 4 fundamental issues:

1. Is the diffusion signal change associated with functional activation dependent on the b-value?

Le Bihan et al. (Le Bihan et al., 2006) reported a linear increase of the raw signal change with b, for b ranging from 600 to 2400 s/mm². They argued this was compatible with the biphasic model, as higher b values would suppress more of the FDP signal and therefore result in the signal being more heavily weighted by SDP. As discussed in Section 4 on multi-contrast experiments, ultra-high gradients allow much higher b values than possible on standard MRI scanners, while keeping the TE relatively short. Based on Fig 1, the same b value of 4000 s/mm² (which is close to the optimum recommended for diffusion fMRI) can be achieved with a TE of 70 ms on a “super-scanner”, compared to a TE of 120 ms on a system with maximum gradient strength of 80 mTm⁻¹. This results in 2.3 times large SNR for a T2 of 60 ms. Using the theoretical relationship between temporal SNR and number of points needed (Murphy et al., 2007), we can infer that,
in order to detect the expected effect size of 0.16% (Le Bihan et al., 2006), the number of sampling points can be reduced by a factor of 12 when using ultra-strong gradients. The faster gradients also allow efficient implementations of high b-value twice-refocused spin echo sequences (Reese, Heid, Weisskoff, & Wedeen, 2003), which should mitigate the impact of BOLD background gradient contributions (Pampel et al., 2010) to the diffusion fMRI signal.

2. *Is the diffusion signal change associated with functional activation dependent on the TE?*

Another interesting question is whether the change observed with activation in the diffusion fMRI raw signal is caused not by a change in the SDP compartment size, as expected if the effect was related to cell swelling, but by a change in the T2 of the diffusion coefficient of either compartment. Miller et al hypothesised that veins and capillaries might have different T2 BOLD effects (Miller et al., 2007) because of their differing degree of oxygenation, which might explain the changes in the raw signal.

Access to ultra-high magnetic field gradients allows large b-values to be obtained at relatively short TEs, and therefore the effects of mapping the TE dependency of the raw signal (see also Section 4 on multi-contrast experiments).

3. *Is the diffusion signal change associated with functional activation isotropic?*

If the effect observed with diffusion fMRI is caused by swelling of the neuronal cell bodies, it would be expected to be fairly isotropic, i.e., independent of the direction of the diffusion gradients. However, some anisotropy has been observed (Miller et al., 2007) potentially attributable to signal coming from structures with ordered geometry, such as vessels in the cortex. Higher resolution scans, achievable through the SNR enhancement granted by ultra-high gradients (see Section 5) could help to clarify the origin of this observed anisotropy. Realistically, however, the estimated SNR advantage of 2.3 estimated for T2=60ms at b=400 s/mm² allows a reduction in isotropic voxel size of approximately 0.75.

4. *Is the diffusion signal change associated with functional activation more spatially specific than BOLD?*

One of the observations that support the neuronal origin of the diffusion fMRI signal is the more precise localisation of activations within the cortical ribbon, compared to BOLD (Le Bihan et al., 2006). However, due to the extremely low SNR of diffusion fMRI, smoothing is necessary to enable the measurement of any activation at all, thus compromising the spatial resolution (Williams et al., 2016).

As ultra-high magnetic field gradients enable diffusion measurements with short TE, the resulting enhanced SNR could be spent on investigating the spatial specificity of diffusion fMRI. Ideally, the profiles of diffusion and BOLD fMRI could be compared between different cortical layers (Williams et al., 2016).
10. IMAGE QUALITY TRANSFER

THE CHALLENGE AND ITS IMPORTANCE

In this article we have identified clearly the gains to be made by investing in ultra-strong diffusion gradient technology. However, as devices with this technology are not mass-produced, they incur considerable cost and are not currently widely available. A reasonable question, when reading an article such as this would be: “That’s all very well, if you have access to such high performance equipment. However, most clinical/research imaging facilities do not. How then, can such technology be of widespread benefit?”

One approach is to establish regional centres that are willing to accept patients / host research visits from parties outside of the host institution (this is the model employed in establishing the National Facility for In Vivo MR Imaging of Human Tissue Microstructure in the UK). However, this still necessitates the establishment of sufficient regional centres to provide access. Another approach is to ‘transfer’ the benefits of ultra-strong gradient technology to systems with lower gradient performance. The idea of image quality transfer (IQT) is to use modern machine learning techniques to construct a mapping from “low quality” to “high quality” images from matched pairs. We can then use the learned mapping to approximate the high quality equivalent of a low quality image, i.e., estimate the image we would have obtained from more time-expensive protocol or hard-to-access imaging device using only data from easily accessible imaging hardware and software.

Through this technology, we can potentially bridge the gap between rare and experimental systems, such as ultra-strong gradient or ultra-high-field MRI scanners, and widely available commercial systems and bring the power of tomorrow’s imaging technology into today’s clinical applications. The existence of one-off bespoke devices such as the microstructure scanner enables this kind of approach. The ambitious potential is for a new medical-imaging “paradigm exploiting coupled design of a) bespoke high-powered devices to capture databases of high quality images, and b) widely deployed cheap and/or low-power devices designed specifically to exploit the rich information from (a)” via IQT (Alexander et al. 2017).

BACKGROUND

Early demonstrations of the idea show good promise. Alexander et al. (Alexander, Zikic, Ghosh, Tanno, Wottscher, et al., 2017b; Alexander, Zikic, Zhang, Zhang, & Criminisi, 2014) show two potential applications in dMRI: i) resolution enhancement of low-resolution diffusion-parameter maps to high resolution; and ii) estimation of microstructural maps from diffusion tensor images. The resolution enhancement from learned mappings shows substantial improvements over standard techniques, such as interpolation. One key result demonstrates how the operation enables tractography to recover thin white matter fibre
pathways, previously recoverable only at 1.25mm isotropic resolution (unobtainable in standard clinical protocols) but from data acquired at 2.5mm isotropic resolution (see Figure 10a).

Moreover, the operation shows remarkable generalizability beyond training data – even mappings learned from fixed ex vivo monkey brain images enhance in vivo human brain images substantially better than
standard methods. The second operation – estimating microstructural maps from diffusion tensor imaging – is challenging, because microstructural maps, e.g. from NODDI (Zhang et al., 2012) and spherical mean technique (Kaden et al., 2016) typically require multi-shell acquisition protocols whereas DT-MRI requires a much less rich, and cheaper, acquisition protocol over only one shell. The potential of IQT is to estimate these microstructural maps from historical single-shell data sets, from which it would normally be impossible to compute such maps. However, although the estimated microstructural maps in (Alexander, Zikic, Ghosh, Tanno, Wottschel, et al., 2017b) broadly match ground truth from multi-shell data sets, significant detail is lost and the technology likely needs further development before the idea will find useful application. Golkov et al. (Golkov et al., 2016) used a similar idea to learn a prior on the structure of NODDI maps that enables reconstruction from sparse data sets, although still involving multiple b-values. Bahrami et al. (Bahrami et al. 2016) aim to construct a mapping from 3T structural images to 7T equivalents again showing promising results (see Figure 10b). The original implementation of Alexander’s work uses patch-based, random-forest regression; Bahrami’s 7T-3T mappings use canonical correlation analysis. A more recent implementation for diffusion resolution enhancement (Tanno et al., 2017) exploits the power of deep learning to gain substantial performance enhancements. Bahrami’s more recent work also explores deep neural networks for mapping 3T to 7T (Bahrami et al. 2016b).

THE PROMISE

Image quality transfer remains in its infancy and substantial further development and validation is necessary to support widespread adoption. Key challenges in development include: finding matched pairs, and handling and quantifying uncertainty in mappings. One way to find matched pairs is to acquire separate low and high-quality images (Bahrami et al., 2016). However, this approach leads to issues of image registration, as alignment is never perfect, while learning approaches typically assume it is. Future efforts may consider embedding misalignment implicitly in the learning process. Another approach, (Alexander, Zikic, Ghosh, Tanno, Wottschel, et al., 2017b), is to simulate the low quality image from the high quality image through a sub-sampling procedure. While this, by definition, produces pairs of images that are in perfect alignment, the artificial construction of the low-quality image can reduce the accuracy of the learned mapping. Quantification of uncertainty is relatively straightforward with traditional learning algorithms, such as random forests, and Tanno et al. (Tanno et al., 2016) showed compelling results in error prediction from IQT through a Bayesian random forest implementation. Uncertainty estimation with deep neural networks is more challenging. Although efforts to do so show promise (Tanno et al., 2017), this is still an area for significant future research.
11. PRACTICAL ISSUES FOR THE USE OF ULTRA-STRONG GRADIENTS

The advantages of ultra-strong gradients for diffusion-based microstructure imaging have been clearly outlined in the previous sections. Here, we discuss the practical issues that set limits on the exploitation of ultra-strong gradients for different types of microstructure imaging studies. The Connectom scanner’s ultra-strong gradient hardware can reach a maximum gradient amplitude of 300 mT/m (McNab et al., 2013). The maximum achievable slew-rate is 200 mT/m/ms, which is comparable with the maximum slew rate of other commercially available gradient systems that provide lower maximum gradient strengths. However, while these specifications promise vastly improved gradient performance in terms of achievable b-value and q-value per unit TE, it is currently not possible to exploit them to the full hardware limits due to physiology-related constraints.

Physiological limits: Peripheral Nerve Stimulation (PNS) and Cardiac Stimulation:
The first relevant safety constraint concerns the avoidance of peripheral nerve stimulation (PNS). The time-varying magnetic fields produced by fast-switching gradient systems can induce stimulation in peripheral nerves. This is experienced as a tingling or twitching sensation close to the threshold for stimulation, but can become uncomfortable or painful as the rate of change of field is increased significantly above threshold. The threshold for stimulation depends on the spatio-temporal characteristics of the electric fields induced in the conducting tissues of the body and their relationship to the anatomical and physiological characteristics of peripheral nerves. Although this complexity makes it difficult to predict the site and threshold of stimulation at the individual subject level (Davids, Guérin, Malzacher, Schad, & Wald, 2017) the likelihood of stimulation depends upon: (i) the rate of change of gradient amplitude with time; (ii) the peak gradient amplitude; and (iii) the position of the subject within the gradient coils (Chronik & Rutt, 2001). To characterise the average stimulation characteristics of a gradient system, MR scanner manufacturers generally test the PNS limits in experiments on human subjects using controlled gradient waveforms. Using this experimental data, the threshold for stimulation for other gradient waveforms can then be well characterized using the SAFE model (Hebrank & Gebhart, 2000) and thus PNS can be controlled. Rapidly time-varying magnetic field gradients can also potentially cause cardiac stimulation. Based on studies using direct electrical stimulation of the heart, this would be expected to take the form of ectopic beats and induced cardiac arrhythmia at lower levels of stimulation and fibrillation at higher stimulation levels. No such cardiac stimulation due to MRI gradient switching has ever been reported in human subjects, though ectopic beats have been generated in animal experiments using small coils to produce extremely high rates of change of magnetic field (2700 T/s over a ramp duration of 530 µs) (Bourland, Nyenhuis, & Schaefer, 1999; Schaefer, Bourland, & Nyenhuis, 2000). As a result of differences in the electrophysiology of cardiac muscle and peripheral nerves, the probability for cardiac stimulation...
arising from gradient waveforms on clinical scanners is extremely small (of the order of $10^{-9}$), but the higher hardware capabilities of the Connectom scanner means that the potential to cause cardiac stimulation must be considered. In previous work, (Setsompop et al., 2013) it was also found that magnetophosphenes were perceived by some subjects when the eyes were significantly displaced from the scanner’s iso-centre and strong time-varying gradients with long rise times (> 130 mT/m with 5-7ms rise-time) were applied. These flashes of light in the peripheral visual field are thought to result from direct stimulation of the retina, and will need to be given some consideration in studies where the head is not situated close to iso-centre.

The PNS limitations on the Connectom gradients when the subject has their head at iso-centre are shown in Figure 11. It shows stimulation limit plots for Gx.

![Stimulation threshold for a train of trapezoidal pulses](image)

**Figure 11.** Comparison of performance limitations imposed by the system and human physiology. The peripheral nerve and cardiac stimulation thresholds for the Connectom AS302 gradient as a function of gradient strength and rise time are plotted. Up to around 180 mT/m, the PNS limitations dictate the gradient performance. However, above 180 mT/m, the theoretical possibility of cardiac stimulation dominates the gradient performance. These limitations are built into the scanner through the SAFE model (Hebrank & Gebhart, 2000) and hardware gradient safety watchdog. Safe operation of the scanner is always ensured by these mechanisms.

For peak gradient strengths less than ~ 80 mT/m, on a single axis, (the X-axis) it is possible to utilise the maximum, 200 mT/m/ms slew rate allowed by the hardware without causing PNS. For larger gradient strengths, and particularly when using combinations of gradients on 2 or 3 axes, the slew rate must be limited to avoid PNS, with a limiting slew rate of around 150 mT/m/ms for the maximum accessible gradient strength of 300 mT/m. For the single gradient axis in Figure 11, above 150 mT/m the gradient slew rate must be further reduced to conform to the IEC limit for cardiac stimulation, which is based on the
predicted probability of cardiac stimulation. This means that a rise time of \(~3.6\) ms is needed to reach the full gradient amplitude of 300 mT/m, corresponding to a slew rate of around 80 mT/m/ms. While this restriction may not adversely affect classical Stejskal-Tanner diffusion experiments, in which the duration of the pulses are relatively long compared to the rise times, it is clear that this places us far from the ‘narrow pulse’ approximation in q-space imaging, and the achievable spectral bandwidth in temporal diffusion spectroscopy experiments (e.g., (Gore et al., 2010)) will be severely affected when using the highest gradient amplitudes. It is important to note however that although the limits set on gradient waveforms to avoid PNS are based on experimental measurements on human subjects, the cardiac stimulation limit is based on inference and extrapolation from animal experiments, (Bourland et al., 1999; Reilly, 1995; Reilly, 1989; 1991) that have led to the maximum rate of change of magnetic field at the heart being limited to values less than \(20/(1 - e^{-t_{\text{eff}}/3})\) T/s, where \(t_{\text{eff}}\) is the stimulus duration in ms. (Medical electrical equipment Part 2-33: Particular requirements for the safety of magnetic resonance equipment for medical diagnosis IEC 60601-2-33:2002).

This can be translated to limits on the gradient strength and slew rate by estimating the position of the heart within the gradient coil-set and knowledge of the spatial variation of the gradient fields. As the cardiac stimulation limit becomes the main constraint on gradient performance it would be timely to revisit the prior literature and to plan new simulations (Davids et al., 2017) and experiments to test the assumptions on which the limits have been based (though acknowledging that any such experiments are likely to be challenging and complex). These investigations are well worth considering, since were the operating limit set by the PNS, rather than cardiac stimulation, limit, the time needed to ramp to a field gradient of 300 mT/m in the Connectom system could be reduced to around 2 ms.

Modification of the designs of whole-body gradient coil systems of the kind employed in the Connectom system specifically to reduce the likelihood of nerve stimulation in particular body regions may also be considered in the future (Davids et al., 2017; Hidalgo-Tobon, Bencsik, & Bontell, 2011). Use of head-only gradient coils can provide access to higher gradient performance than whole-body sized systems, like the Connectom gradients, at the threshold for stimulation, since the reduced coil size general limits the maximum \(dB/dt\) values produced in the body. For example, a recently described asymmetric head gradient coil system was shown to be able to exploit the full hardware capability in slewing to a maximum gradient of 80 mT/m at a rate of 700 mT/m/ms for the \(y\) (anterior-posterior) gradient direction without causing PNS (in 22/24 subjects scanned) (Lee et al., 2016). However, head-only gradient coils that can generate gradient strengths as high as 300 mT/m have, not yet been demonstrated and it is worth noting that the concomitant field effects (described below) are generally much more severe in asymmetric gradient sets. Moreover, head-only gradient sets limit the available space for radiofrequency coils, which becomes
problematic for systems larger numbers of receiver coils (e.g., more than 32) and for multi-channel transmit systems.

**Duty cycle/heat dissipation:**
The Connectom gradients have been designed to operate with the following duty-cycle limits (assuming Trapezoidal gradient pulses)- at 300 mT/m: 6% on all axes or 17% on one axis, at 210 mT/m: 12% on all axes or 35% on one axis, at 90 mT/m: 50% on all axes or 100% on one axis.

Although the diffusion gradient pulses typically have a short duration, when used at maximum gradient strength, the 6% limit on duty cycle can increase the required TR substantially, especially if the EPI train is short or non-EPI readout techniques are used. Power amplifiers generally operate with maximum and average power specifications. They are designed for average power at 100% duty cycle, but can also be used at peak power with limited duty cycle. The same holds true for gradient coils, for which the limiting factor is power dissipation in the electrical conductors. The most efficient design of ultra-strong gradient kit uses all available space (i.e., wire cross section) in the magnet bore for gradient field-generating conductors and cooling layers. In the Connectom system (McNab et al., 2013) the optimum under these technical boundary conditions requires the use of four gradient amplifiers and results in a nominal (or average) gradient amplitude of ~100mT/m. A further increase under these conditions (i.e., limited space in the magnet bore) would require new cooling technologies for the gradient coil and an increase of the average power of the amplifier. Thinking to the future, one possibility might be to incorporate a superconducting gradient coil either within the magnet cryo vessel or within a separate cryo vessel, keeping in mind the critical field strength for superconductivity which depends on the static B₀ field and the time varying magnetic field of the gradient field conductors. Although this is a challenge given that the currently available wire materials show significant heat production for alternating currents in the 1-5KHz range even at cryogenic temperatures, latest research seem to indicate that this is technically achievable (see: Yuan and Shen, 2007 and patent number GB2433600)

**Gradient non-linearity:**
A current trade-off in achieving ultra-high gradient amplitudes has been in the region over which the gradient system behaves linearly. A strict boundary condition for linearity would lead to suboptimal gradient efficiency and lower the PNS limit. Thus, the system was designed to give linearity over the typical volume of a head, with homogeneity of approximately 6% over a 20cm sphere and 17% over a 40cm sphere. However, due to the nature of the gradient design (a symmetric, whole body system), such a system can be applied to other parts of the human body, but the reduced linearity over the extended volume require improved strategies for handling gradient non-linearity. The impact of non-linearity in
terms of image distortion is well documented (Jovicich et al., 2006) and scanner vendors typically correct these distortions using well-established unwarping algorithms. Since the non-linearities in typical whole-body scanners in a volume of the size of human head are very small, the effects of such non-linearities are very minor. The gradient non-linearities can have a significant impact on diffusion studies in ways other than simple image distortion, if positioning is not well controlled. Depending on the non-linearity, the b-matrices at each voxel location in the brain will be spatially dependent and differ from the nominal values (Bammer et al., 2003; Borkowski, Kłodowski, Figiel, & Krzyżak, 2017; Malyarenko, Ross, & Chenevert, 2014; Mohammadi et al., 2012). Besides, gradient non-linearity distorts the true image in concert with $B_0$ inhomogeneity, making it harder to disentangle the two effects in a sequential manner at the post-processing stage. This complication has led to the development of diverse processing pipelines (Fan et al., 2016; Glasser et al., 2013; Setsompop et al., 2013; Sotiropoulos et al., 2013). Moreover, given that the gradient non-linearities need not be rotationally symmetric, head movement can complicate the b-matrix temporal evolution at different voxel locations significantly, requiring one to keep a temporal history of b-matrix evolution at each location to remove unwanted biases in measured diffusion parameters. The effect of each of these issues on the final microstructure parameters obtained should be carefully assessed and where necessary, dedicated processing pipelines that address gradient non-linearity should be implemented to make the best use of these novel hardware platforms.

Concomitant Fields

The use of high gradient strengths increases the concomitant field effects (Bernstein et al. 1998; Baron, Lebel, Wilman, & Beaulieu, 2012; Meier et al. 2008; Norris & Hutchison, 1990) that can confound diffusion measurements. Concomitant fields arise because the vector magnetic field generated by a gradient coil has to satisfy Maxwell’s equations. Hence, to ensure that the curl and divergence of the field are both zero in current free regions of space, the desired spatial variation of the $z$-component of the magnetic field must always be accompanied by a spatially varying $x$- and/or $y$-component of the field. Since these concomitant fields are orthogonal to (and much smaller in magnitude than) $B_0$, their effect on the Larmor frequency can usually be neglected, but in fact there is always a net change to the Larmor frequency $\omega_c = \gamma B_c$ due to these concomitant fields which is given by

$$B_c = \frac{(G_x^2 + G_y^2)z^2}{2B_0} + \frac{G_x^2(x^2 + y^2)}{8B_0} + \frac{G_{xz}}{2B_0} - \frac{G_{yz}yz}{2B_0}$$ \[1\]

when gradients $G_x$, $G_y$ and $G_z$ are applied simultaneously. Equation [1] shows that concomitant fields scale quadratically with the applied gradient strength and so become more than 56 times larger when the gradient strength is increased from 40 mT/m (gradients on a conventional clinical scanner) to 300 mT/m whilst keeping $B_0$ fixed. Note that this problem would be even more pronounced for asymmetric gradient designs (e.g., head-only gradient systems), the extra complication arising from the presence of odd-order terms to the concomitant field correction. Techniques to circumvent such problems have been proposed,
including manipulating the pre-emphasis (Tao et al. MRM 2017), real-time adjustment of the centre (B₀) frequency (Weavers et al. MRM 2018). These concomitant field effects are well known to cause image distortion, which particularly affects echo planar imaging since the concomitant fields due to the switched read-out gradient produce a progressively increasing position-dependent phase accumulation down the echo train (Y. P. Du, Joe Zhou, & Bernstein, 2002; Norris & Hutchison, 1990). Fortunately, with knowledge of the applied gradients this distortion can be corrected (Du et al., 2002). When the Connectom gradient system is applied in diffusion imaging, the peak gradient strength during the EPI readout is similar to that used on a conventional scanner, so the problem of image distortion due to concomitant fields is not particularly exacerbated, but the effects of the concomitant fields due to the diffusion encoding gradient pulses can be challenging. In particular, the quadratic dependence of the effective field, B₀, on the gradient strength means that phase accumulation due to concomitant fields is identical for positive and negative gradient lobes of the same amplitude and duration. There is consequently no net phase accumulation when unipolar gradients are applied in conjunction with a 180° RF pulse in a Stejskal-Tanner sequence (Stejskal & Tanner, 1965), but the concomitant fields produce a spatial phase variation over the imaging region when balanced bipolar gradient pulses are applied. This is problematic for diffusion encoding sequences employing bipolar gradients (Setsompop et al., 2013) such as the twice-refocused spin echo sequence (Reese et al., 2003) that is commonly used to reduce eddy current effects. With such sequences, through-slice variation of the accrued phase can produce a spatially varying image attenuation which modulates the effect of diffusion, while the variation of phase within the image plane can shift the time of acquisition of the centre of k-space by different amounts for different regions causing image blurring and signal loss (Baron et al., 2012). The terms in Eq. [2] varying as xz and yz can also produce a linear phase shift across the whole image, which may shift the centre of k-space outside of the acquisition window, thus causing complete signal loss (Baron et al., 2012). For example if a bipolar x-gradient of 300 mT/m strength is applied with lobes of 10 ms duration, the phase variation across a 2 mm-thick axial slice displaced by 2.5 cm from isocentre is around 8π, while the linear phase variation produced in the same axial slice when bipolar gradient encoding of similar strength and duration is applied with the gradient oriented at 45° to x and z would be around 320π m⁻¹ which would be sufficient to shift the centre of k-space outside the acquisition window in 3 mm resolution EPI data. Both these values increase linearly with the displacement of the axial slice from isocentre. While these effects mean that there are strong incentives to stick with ST sequences for diffusion encoding when using ultra-strong gradients, approaches for ameliorating the effects of concomitant fields in sequences using bipolar gradients have been developed (Baron et al., 2012; Meier, Zwanger, Feiweier, & Porter, 2008; Baron et al., 2012). These approaches, which involve applying compensating pre-phasing gradient pulses to eliminate the spatially varying phase produced by the concomitant fields during diffusion encoding, currently allow localised reduction of confounding effects, but their efficacy is yet to be evaluated with ultra-high gradient strengths. Full exploitation of the power of
Connectom-like gradients in experiments using bipolar gradients will require the development of new approaches that can cancel the phase variation due to concomitant fields over larger image volumes.

**CONCLUSION**

Even though diffusion encoding with strong gradients raises challenges, it is without question that with ongoing developments to better understand such issues, a microstructural ‘super-scanner’ provides considerable improvements to ‘conventional’ diffusion measurements and facilitates new measurements that are impractical at lower gradient amplitudes. New insights arising from this extended measurement space, together with practical experience on the challenges related to strong gradients, will provide a sound foundation to develop the next generation of microstructural super-scanners.

**ACRONYMS USED IN THIS WORK:**

ADC = apparent diffusion coefficient  
AMICO = Accelerated Microstructure Imaging via Convex Optimization  
BOLD = blood-oxygenation level-dependent  
CHARMED = composite hindered and restricted model of diffusion  
CPMG = Carr-Purcell-Meiboom-Gill  
DDE = double diffusion encoding  
DIVIDE = diffusional variance decomposition  
DKI = diffusion kurtosis imaging  
dMRI = diffusion MRI  
DTI = diffusion tensor imaging  
DW-MRS = Diffusion-weighted Magnetic Resonance Spectroscopy  
DW-SSFP = diffusion weighted steady-state free precession;  
DW-STE = diffusion weighted stimulated echo  
EPSRC = Engineering and Physical Sciences Research Council  
FA = fractional anisotropy  
FEXI = Filter Exchange Imaging  
gSlider-SMS = Generalised Slice Dither Enhanced Resolution with Simultaneous Multislice  
HARDI = high angular resolution diffusion imaging  
IMPULSED = imaging microstructural parameters using limited spectrally edited diffusion  
IQT = image quality transfer  
IR-PGSE = inversion-recovery prepared pulsed-gradient spin echo  
MD = mean diffusivity  
MDE = multiple diffusion encoding
NAA = N-acetyl-aspartate
NODDI = neurite orientation dispersion and density imaging
OGSE = oscillating gradient spin echo
POMACE = Pulsed and oscillating gradient MRI for assessment of cell size and extracellular space
PGSE = pulsed gradient spin echo
PGSTE = pulsed-gradient stimulated echo
PNS = peripheral nerve stimulation
PRESS = point resolved spectroscopy
QTE = Q-space trajectory encoding
SAFE = Stimulation approximation by filtering and evaluation
SDE = single diffusion imaging
SNR = Signal to noise ratio
TE = echo time
tCr = total Creatine
tCho = total Choline
VERDICT = Vascular, Extracellular, and Restricted Diffusion for Cytometry in Tumours

**ACKNOWLEDGEMENTS:**

The UK National Facility for In Vivo MR Imaging of Human Tissue Microstructure was generously funded by the Engineering and Physical Sciences Research Council (EPSRC) (grant EP/M029778/1), and The Wolfson Foundation, to whom the authors give their thanks.

DKJ is supported by a Wellcome Trust Investigator Award (096646/Z/11/Z) and a Wellcome Trust Strategic Award (104943/Z/14/Z). CMWT is supported by a Rubicon grant (680-50-1527) from the Netherlands Organisation for Scientific Research (NWO) and Wellcome Trust grant (096646/Z/11/Z). MP is supported by EPSRC (EP/G007748, EP/I027084/01, EP/L022680/1, EP/M020533/1, N018702). EPSRC grants EP/G007748 EP/L022680/1 EP/M00855X/1 EP/M020533/1 EP/N018702/1 EP/R014019/1 and MRC grant MR/M009106/1 support DCA’s work on this topic.

The authors are indebted to Dr. Ralph Kimmlingen, Siemens Healthcare GmbH, Erlangen, Germany for his helpful comments and contributions to Section 11 of this article, and to Dr Kawin Setsompop, Athinoula A. Martinos Center for Biomedical Imaging, for supplying the diffusion-weighted data used to create Figure 6.
REFERENCES


Assaf Y, Ben-Bashat D, Chapman J, Peled S, Biton IE, Kafri M, Segev Y, Hendler T, Korczyn AD, Graif M,


Barazany D, Jones DK, Assaf Y. 2011. AxCaliber 3D. In “Proc. ISMRM 19th Annual Meeting, Montreal, Canada”.


Comparison to the Tasmanian Devil. PloS One, 12(1), e0168993.  
http://doi.org/10.1371/journal.pone.0168993


http://doi.org/10.1016/j.jmr.2013.07.008


Fan Q, Huang SY, Nummenmaa A, Witzel T, and Wald LL T2 relaxation rates of the fast and slow bi-exponential diffusion components in the in vivo corpus callosum

In International Society for Magnetic Resonance in Medicine, 25th Scientific Meeting, Honolulu, Hawai‘i. p. 1757 838.


Larsson EM, Englund E, Sjöbeck M, Latt J, Brockstedt S. MRI with Diffusion Tensor Imaging Post-Mortem at 3.0 T in a Patient with Frontotemporal Dementia. Dementia and Geriatric Cognitive Disorders. 2004; 17:316-319


http://doi.org/10.3389/fphy.2014.00011


Parker, Greg D.; Marshall, David; Rosin, Paul; Drage, Nicholas; Richmond, Stephen; Jones, Derek K. 2013. RESDORE: Robust Estimation in Spherical Deconvolution by Outlier Rejection In International Society for Magnetic Resonance in Medicine, 21st Scientific Meeting, Salt Lake City, USA. p. 3148.


Pierpaoli C. Artifacts in Diffusion MRI, in ‘Diffusion MRI: Theory, Methods and Application’ (Jones, D.K., Oxford University Press, 2010).


http://doi.org/10.1016/j.neuroimage.2013.07.067


http://doi.org/10.1002/mrm.24529


