Chapter

18 To BOLDly Go: fMRI, Perfusion and Diffusion

18.1 Introduction

We turn now to some of the more advanced areas of MR straight from the cutting edge between research and clinical practice. The techniques in this chapter all relate to what is sometimes termed 'micro-contrast' mechanisms. This term is unfortunate because basic T_1 and T_2 relaxation depend upon the microscopic (at molecular level) interactions of spins. Nevertheless, these new techniques give specific information about aspects of tissue not readily available through conventional contrast mechanisms. In this chapter you will learn that:

- diffusion-weighted imaging relates to the mobility of water molecules, particularly when mobility is restricted to specific orientations, as in white-matter tracts;
- perfusion-weighted imaging relates to the delivery of blood to tissues, in terms of relative or absolute concentrations, or in terms of rates of delivery or mean transit times;
- perfusion-weighted imaging can use exogenous contrast (i.e. a bolus of contrast agent) or endogenous contrast, by the 'spin labelling' of arterial blood;
- dynamic contrast-enhanced MRI, also called permeability imaging, relates to the leakiness of blood vessels;
- BOLD (Blood Oxygenation Level Dependent) imaging, sometimes called fMRI (functional MRI – note the small 'f'), is used to investigate regional brain activation; in lay terms, to observe the brain working.

Echo planar imaging (EPI), which was explained in Chapters 12 and 13, is the imaging sequence of choice for these applications on account of its extreme speed. The compromise for speed in EPI is extreme sensitivity to susceptibility-induced image distortion, low resolution and higher artefact levels.

18.2 Diffusion Imaging

The sensitivity of MR signals to random molecular motions, self-diffusion, has been known since the original pioneering work of Hahn, Carr and Purcell over half a century ago. More recently diffusion-weighted MRI (DW-MRI) has established itself as an important method in the assessment and diagnosis of acute stroke, revealing dramatic signal differences within an hour of onset.

18.2.1 A Trip to the Mall: Molecular Motion

Figure 18.1 shows the types of molecular motions that occur in tissues: bulk flow (e.g. along vessels), isotropic diffusion (which occurs in pure fluids) and restricted diffusion where cell membranes restrict the movement of molecules in one or more directions. It is important to understand that diffusion is different from flow. In bulk flow there is a net transport of molecules from one location to another, like cars driving down the highway from one city to another. In diffusion there is no overall shift in the average location of the molecules – just increased agitation or energy. Their movement is similar to that of the shoppers in a mall – no one is actually getting anywhere – but there's loads of activity all in seemingly random directions! Figure 18.2 illustrates the 'random walk' of molecules under diffusion.

18.2.2 Pulsed Gradient Spin Echo

The most commonly applied method for producing diffusion-weighted contrast is the **P**ulsed **G**radient **S**pin **E**cho (PGSE) method, sometimes called Stejskal and Tanner after its inventors. It consists of a 90°–180° pair of RF pulses with large and equal sized gradients placed on either side of the 180° pulse (Figure 18.3). By manipulating the gradients we can control the degree of diffusion weighting or *b*-factor. This involves the scanner changing the strength of the diffusion sensitizing gradients (G) and their timing as described in Box 'To *b* or not to *b*'.

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

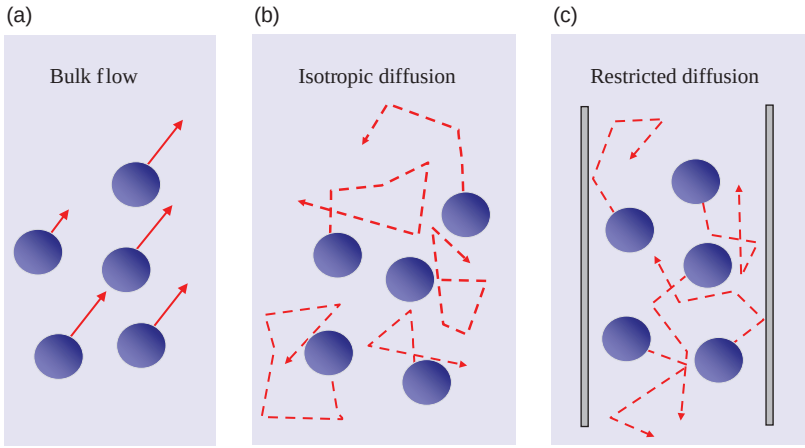

Figure 18.1 Three types of molecular motion (indicated by arrows) which may occur in tissue in (a) bulk flow, (b) isotropic diffusion where molecular motion is random and (c) restricted diffusion where random motion is constrained by physical barriers, e.g. by cell membranes.

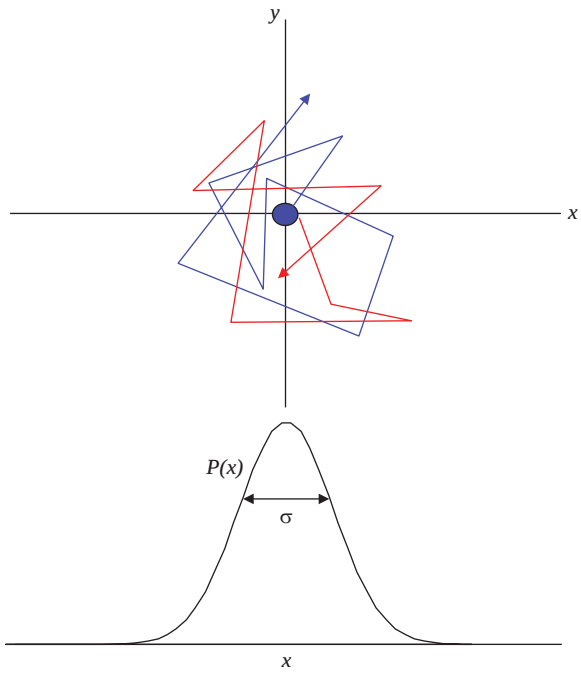

Figure 18.2 Random walk of molecules or 'Brownian motion'. More diffusion results in a greater spread of the molecules, but the average location remains constant. The probability of a molecule's location follows a Gaussian bell-shaped distribution with standard deviation σ.

DW contrast behaves rather like inverse T_2 weighting, in that watery tissues that have very mobile molecules give lower signal intensity, while more solid and static tissues give a stronger signal. The signal strength is described by the equation

$$S(b) = S_0 \exp(-bD)$$

where $S(b)$ is the signal for a particular b-value (see Box 'To b or not to b') and D is the self-diffusion constant of the tissue. In MRI we use the term *apparent diffusion coefficient* (ADC) (see Box 'The Diffusion Coefficient – Apparently'). This can be calculated from two or more images with different b-values and can be displayed as an ADC map. The diffusion coefficient D is measured in units of $mm^2\ s^{-1}$. At room temperature pure water has a value of D approximately $2.2 \times 10^{-3}\ mm^2\ s^{-1}$. With a b of 1000, the water signal will be reduced to 11% of its unweighted value. In cancerous tissues there is often a proliferation of cellularity and a subsequent reduction in the intercellular space. This can result in a reduction in the free diffusion of extracellular water molecules and hence a reduction in ADC values. Thus the use of ADC values can help in the detection of cancer and in differentiating between benign and malignant lesions. Typical values are given in Table 18.1.

Another reason for calculating ADC maps is to deal with the issue of 'T_2 shine-through'. As DWI images necessarily involve a long TE, they are T_2 weighted in addition to being diffusion weighted. Areas which have an elevated T_2 may be indistinguishable on a DW-image from those with low diffusion (Figure 18.4). The ADC map removes this ambiguity.

To *b* or not to *b*

In the PGSE sequence the value of b is given by

$$b = \gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3}\right)$$

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

and is determined by the gradient amplitude G and duration δ ('little delta') and trailing-to-leading edge separation ($\Delta - \delta$) where 'big delta' is the centre-to-centre separation. b has units of s mm^{-2}. We most often use a b-factor of 1000 s mm^{-2} to obtain good contrast. Large gradient amplitudes (of up to 50 mT m^{-1}) are very useful because they allow the timing parameters to be minimized, thus reducing the TE. Even so a DW acquisition will always have some T_2 weighting which can result in the so-called 'T_2 shine-through'.

The quantity ($\Delta - \delta/3$) is known as the diffusion time τ and is related to molecular motion through the Einstein equation

$$\langle R^2 \rangle = 6D\tau$$

where $\langle R^2 \rangle$ is the mean square displacement of a collection (or ensemble) of molecules.

The PGSE gradient scheme works like the velocity-encoding gradients used in MR angiography.

Table 18.1 Typical apparent diffusion coefficient (ADC) values for human tissues

	ADC ($\times 10^{-3}$ mm^2 s^{-1})	Relative signal (at b = 1000)
CSF	2.94	0.05
Grey matter	0.76	0.47
White matter	0.45	0.63
White matter parallel to fibres	0.95	0.39
Liver	1.8	0.17
Liver benign (cysts, haemangioma)	2.5	0.08
Liver (metastases, HCC)	1.1	0.33

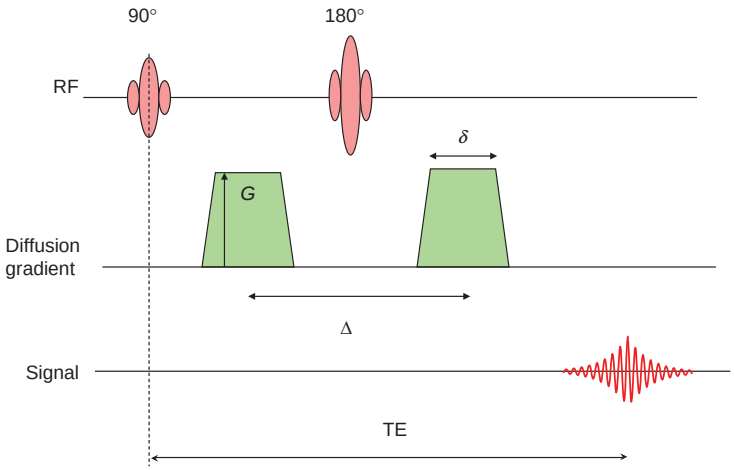

Figure 18.3 Basic pulsed gradient spin echo (PGSE) sequence for diffusion weighting. For imaging this precedes a spin-echo EPI acquisition. δ denotes the pulse width and Δ the centre-to-centre spacing. G is the magnitude of the diffusion-weighting gradient.

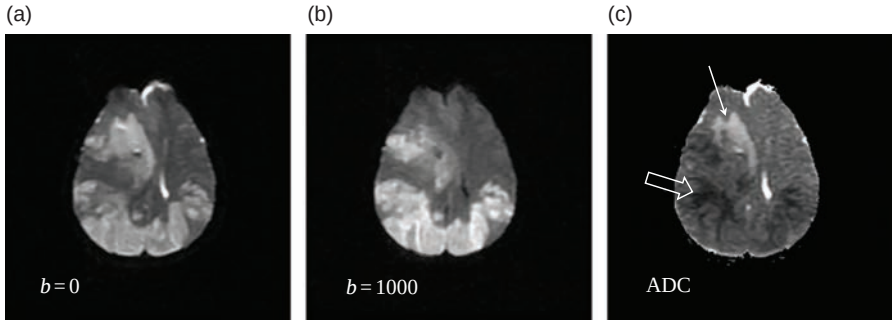

Figure 18.4 DW imaging of a stroke patient: (a) EPI image with b = 0, (b) b = 1000, (c) calculated ADC map. Acute infarcts have reduced ADC (open arrow) while chronic infarcts have elevated ADC (arrow). T_2 shine-through is evident where the DW image is bright, but the corresponding ADC is unchanged.

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

Bulk flow will have its phase changed according to its velocity, but the phase angle will be many times 2π. With self-diffusion the molecules are moving about randomly, changing direction many times. DW imaging works because all the spins accrue random and unique phase changes as they move about within the gradient. This results in a net loss of signal within each voxel, provided sufficient DW is applied. Restricted diffusion applies where physical barriers, e.g. cell walls, prevent or limit this motion. A high value of D (the self-diffusion constant of the tissue) or ADC implies high motion and therefore low signal in DW-MRI. The corresponding ADC map will be bright.

The Diffusion Coefficient – Apparently

Diffusion weighting in an image is actually measured indirectly, via the relative dephasing of signals where diffusion occurs. However, dephasing can be caused by several other physiological motions, e.g. CSF or blood flow. In a simple diffusion scan we cannot distinguish between these effects, so we usually refer to the apparent diffusion coefficient or ADC, which can be calculated from

$$\text{ADC image} = -\frac{1}{b}\ln\left(\frac{\text{DW image}}{T_2\text{w image}}\right)$$

for cases where the same TE is used for both the weighted and unweighted images. Since SE-EPI has geometric distortions, it is normal to use SE-EPI for the T_2w image. All manufacturers provide this automatically in the pulse sequence, so that the scan contains a $b = 0$ image together with one or more b factors for diffusion weighting. b-values of 500 and 1000 s mm^{-2} are common in DWI of the brain, while slightly lower values are common in body applications, e.g. prostate.

Instead of using just two b-values (0 and a higher value), a range of b-values can be applied and a 'least-squares' fit can be performed on a plot of $\ln(S_b)$ against b. The gradient of this line will give a more accurate value of ADC. See Chapter 19 for more sophisticated quantification methods.

Clinical Applications: Neuro

Neuro DWI is by far the most common clinical application. DWI is very sensitive for early changes in ischaemic stroke and in focal brain lesions. Ischaemic stroke, arising from thrombotic or embolic occlusion of a cerebral artery, leads to cell swelling or cytotoxic oedema, which results in a depressed ADC. Acute stroke thus appears hyperintense (bright) on DW imaging (Figure 18.4). These changes are apparent in under an hour on DW imaging (compared with over 6 h on CT or 12 h on T_2-weighted MRI). It can also help to distinguish between acute stroke and other conditions with acute neurological deficits such as transient ischaemic attacks and atypical migraine, which show no abnormality in DW imaging. DWI can also help in differentiating solid tumours from cystic ones, and for trauma, haemorrhage and infections.

18.2.3 Anisotropy and Diffusion Tensor Imaging (DTI)

Pure water has isotropic diffusion properties, meaning that the molecules are equally likely to wander off in any direction. So the diffusion gradients can be applied in any physical direction and the effect on the MR signal should be the same. In many biological tissues (and especially white-matter tracts) diffusion is restricted by the presence of cell membranes and there may be a preferential direction, e.g. along nerve fibres.

This type of diffusion behaviour is called anisotropic. The measurement of diffusion anisotropy can yield useful biological information on the tissue's microstructure, e.g. degree of myelination. To measure anisotropy it is necessary to apply diffusion weighting on a number of different directions, known as **D**iffusion **T**ensor **I**maging (DTI). We define a new parameter, called 'fractional anisotropy' (FA) which has values from 0 to 1. For pure isotropic tissues FA = 0, while for tissues which have very strong anisotropy in one direction FA = 1. Figure 18.5a shows an FA map of the brain. There have been many clinical studies showing that reduced FA is a sensitive (but non-specific) indicator of a disease process. Colour maps (Figure 18.5b) for fractional anisotropy provide a way of displaying two types of information simultaneously: the hue represents the principal direction of the anisotropy and brightness represents how strongly anisotropic the diffusion is within that voxel.

For anisotropic tissues, the physical orientation of the tissue (e.g. fibre direction) in conjunction with the applied gradient direction will determine the signal intensity. If these two directions happen to be co-linear, we can accurately measure the diffusion

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

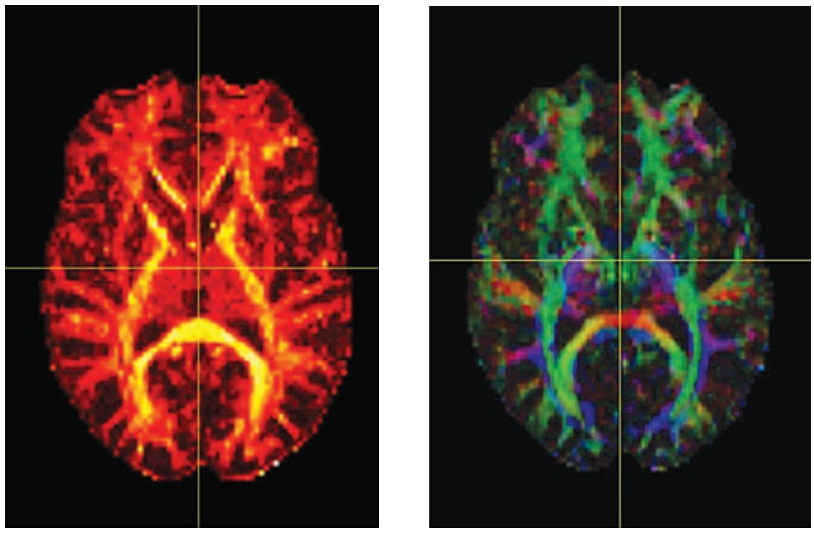

Figure 18.5 (a) FA map of the brain. Note how it resembles a strongly T_1-weighted image. (b) Colour FA map from the same subject. Green represents the principal eigenvector of diffusion in the AP direction, red is LR, blue is SI.

coefficient in that direction. However, that's not normally the case and we must generalize using a tensor. The diffusion tensor is a kind of two-dimensional vector and so has nine elements or values, each corresponding to a gradient orientation and a cell orientation as described further in Box 'Diffusion Tensor Maths'. In practice we don't know before the scan what the cell orientation values will be, so many diffusion gradient directions are applied, typically 30–60, or even more in research studies. Figure 18.6 shows the elements of the diffusion tensor.

Once the tensor terms have been found, we can define several new diffusion parameters: the 'trace-ADC' (see Box 'Diffusion Tensor Maths') always comes out as the same number no matter the direction – so is a very useful parameter in stroke or oncology. From this average ADC can be computed.

Once we have the diffusion tensor for every voxel in the image, we can also create a 'tractography' image, which shows the water diffusion pathways. Since white-matter bundles are the main biological cause of anisotropy, these pathways tend to follow the direction of the white-matter fibres. The resulting tractography images are remarkably similar to anatomical dissections showing WM bundles (Figure 18.7). Although there are many caveats to interpreting tractography (see Box 'Tractography Processing'), they are hugely popular with neuroradiologists and neurosurgeons, as they can reveal important relationships between brain lesions and surrounding normal tissue.

Diffusion Tensor Maths

The diffusion tensor is defined as:

$$DT = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix}$$

The first subscript (x, y, z) refers to the 'natural' orientation of the cells or tissue, and the second refers to the gradient orientation. The diagonal elements D_{xx}, D_{yy} and D_{zz} correspond to the simple three-direction measurements found in commercial scanners. To make a full measurement of ADC in an anisotropic tissue all these components plus a $b = 0$ unweighted image is required. In practice there is a degree of redundancy (because D_{xy} is the same as D_{yx}, etc.) and only seven measurements are required, six diffusion directions and an unweighted ($b = 0$) image as in Figure 18.6.

The trace diffusion constant, which is a so-called scalar invariant, can be computed as

$$\text{Trace}(D) = D_{xx} + D_{yy} + D_{zz}$$

from which an average ADC is obtained, equal to

$$D_{ave} = \frac{1}{3}\text{Trace}(D)$$

The fractional anisotropy FA is calculated from the principal eigenvalues λ of the tensor:

$$\text{FA} = \sqrt{\frac{3}{2}} \cdot \sqrt{\frac{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$$

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

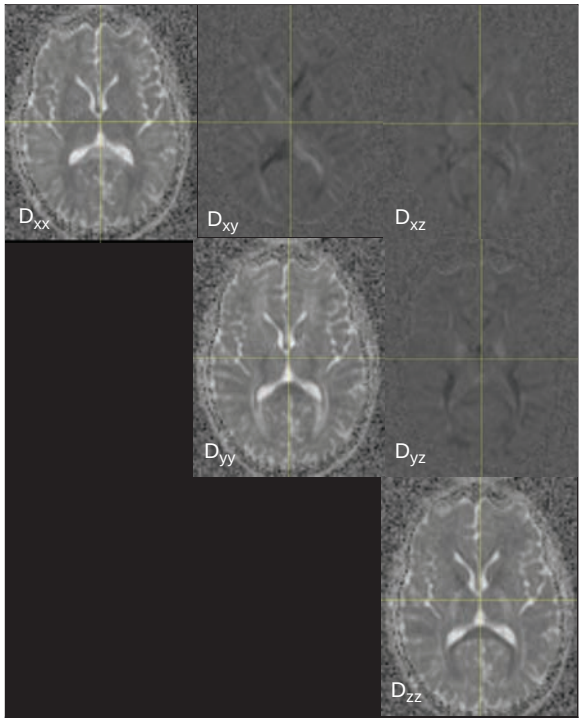

Figure 18.6 Elements of the diffusion tensor determined from a multi-direction DTI acquisition.

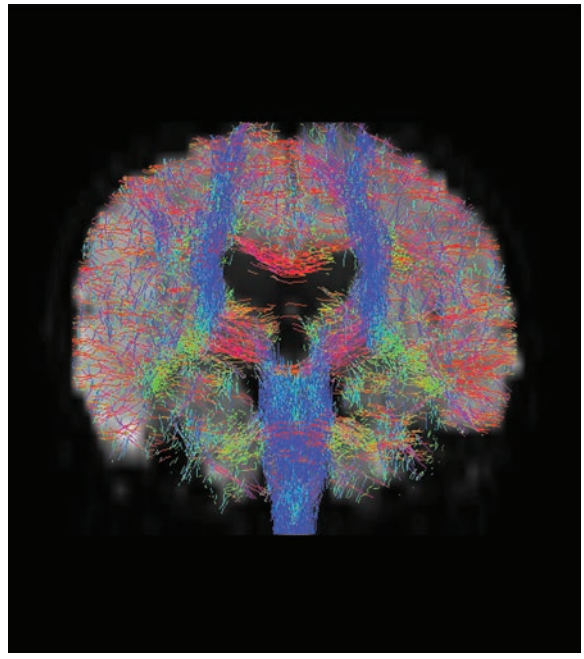

Figure 18.7 Tractography using constrained spherical deconvolution. Tract directions are blue – SI, green – AP, red – LR.

Tractography Processing

Conventional tractography requires the selection from a seed (or starting point from which to draw the tracts). It is essential that this seed point is chosen in an anatomically meaningful way and that a neurology or neuro-radiology specialist reviews the tractography images to ensure they do not contain artefactual information. One problem with DTI tractography is how to resolve fibres which cross paths within a voxel. Several approaches have been tried, including the use of very high b-values, 'Q-ball' and 'q-space' techniques.

Constrained **S**pherical **D**econvolution (CSD) is a tractography method that can account for crossing fibres and does not require a user-defined seed point. CSD works by estimating a fibre response function. To do this a normal DTI acquisition is made using at least 30 directions. From these a tensor image is derived (Figure 18.6) and corresponding FA maps. From the FA maps voxels which only contain single direction fibres (e.g. from the corpus callosum) are selected (Figure 18.8a). From these a response function for a single fibre direction can be calculated (Figure 18.8b). Deconvolution of this with the tensor image results in an orientation map, where crossing fibres are unambiguously depicted (Figure 18.8c). A tractography map can be computed by plotting the streamlines from the orientation plots (Figure 18.7).

18.2.4 Diffusion Sequences

Spin-echo EPI is the sequence of choice for DW imaging. Its very rapid acquisition time means that very little motion-induced artefact is encountered. The PGSE portion is appended to the front of the sequence as a preparation phase. Often sequences will be arranged to provide multiple images with a range of diffusion directions and *b*-values, and sometimes the scanner software will calculate an ADC map. Because of the need to have a reasonable diffusion time in the PGSE preparation, the TE value is normally quite high, greater than 100 ms. However, parallel imaging should be used to reduce the EPI echo train, which not only reduces the TE to a more suitable 60–70 ms, but also reduces the amount of geometric distortion in the images and reduces T_2 shine-through. Since SE-EPI is a multi-slice technique, DWI is typically performed with slice thicknesses of 5 mm. For DTI, it is desirable to have isotropic voxel resolution – typically 2–3 mm

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

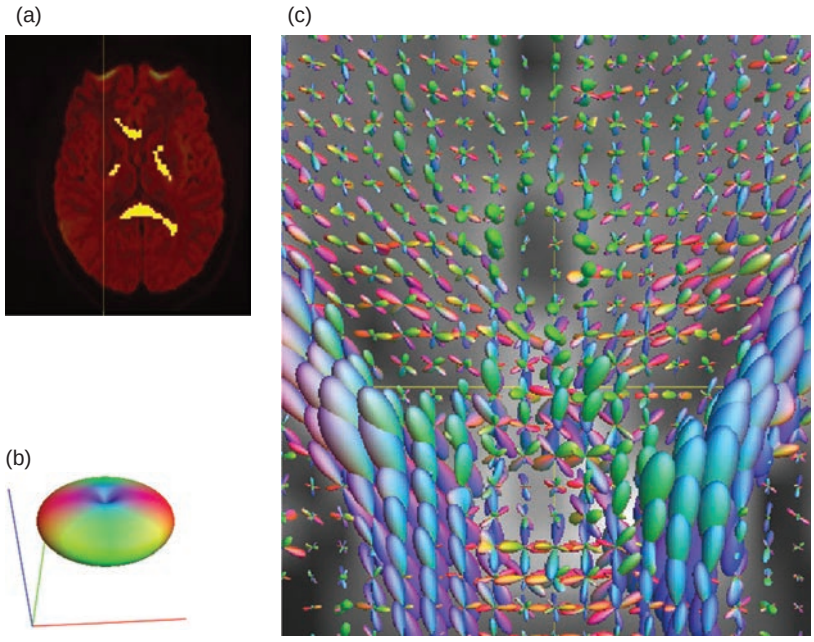

Figure 18.8 (a) Using an FA map, voxels containing single fibre orientations are selected. (b) A direction response function is generated. (c) Deconvolution of the tensor with the response functions resolves crossing fibres, displayed in an ellipsoid map, where the size of each 'bubble' represents the ADC value, with colours as before.

slice thickness is used, and so the number of slices to cover the brain is 80–100.

Conventional sequences such as (turbo) spin echo can also be used for DW imaging. However, their long scan times render the images very sensitive to bulk motion. DW-TSE is typically used for non-neuro diffusion imaging, e.g. the inner ear or head–neck region, where the long scan times are acceptable because the images have no EPI distortion. However, for the vast majority of DWI and DTI, spin-echo EPI is still the best choice.

Clinical Applications: Non-neuro

Outside of the brain, DWI is also very important, and is becoming a mandatory imaging sequence in many examinations. The main role is to localize tumours, e.g. in breast, prostate, liver and bone cancers. In all these areas, the lesion has a lower diffusion coefficient (more restricted diffusion), so it shows as a bright area on the DWI.

The key difficulty with non-neuro DWI is that SE-EPI is sensitive to B_0 inhomogeneity and the images can be very distorted. As well as making it difficult to recognize the anatomy, the distortion also leads to signal distortions, and especially signal pile-up which can be misinterpreted as a lesion. There are ways to avoid the distortion, e.g. using turbo spin echo instead of EPI, or using multi-shot EPI, but these all extend the scan time and are sensitive to patient motion.

When quantifying diffusion in non-neuro applications, it is important to separate perfusion (e.g. capillary flow) from diffusion. In order to do this, several *b*-values are acquired, and the signal curve is fitted using a bi-exponential model. The shorter exponential represents the perfusion fraction, while the longer one is the true diffusion coefficient.

Whole-Body Diffusion

The technique of whole-body diffusion, sometimes called DWIBS (**D**iffusion-weighted **W**hole-body **I**maging with **B**ackground **S**uppression) can produce 'PET-like' images useful for evaluating metastases. DWIBS combines DW-EPI with a STIR-based fat suppression using a high *b*-value (typically 1000). The scans are acquired with free-breathing with multiple NSA used to attenuate any movement artefact. Images are acquired from multiple positions (stations) usually in the axial plane (Figure 18.10a) and are reformatted (Figure 18.10b) and stitched together to make a whole-body image. They are often viewed as Maximum Intensity Projections (MIPs) with a contrast reversal (Figure 18.10c). This renders them similar in appearance to PET images – although it must be remembered that DWIBS is measuring molecular motion while FDG-PET is sensitive to glucose metabolism.

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

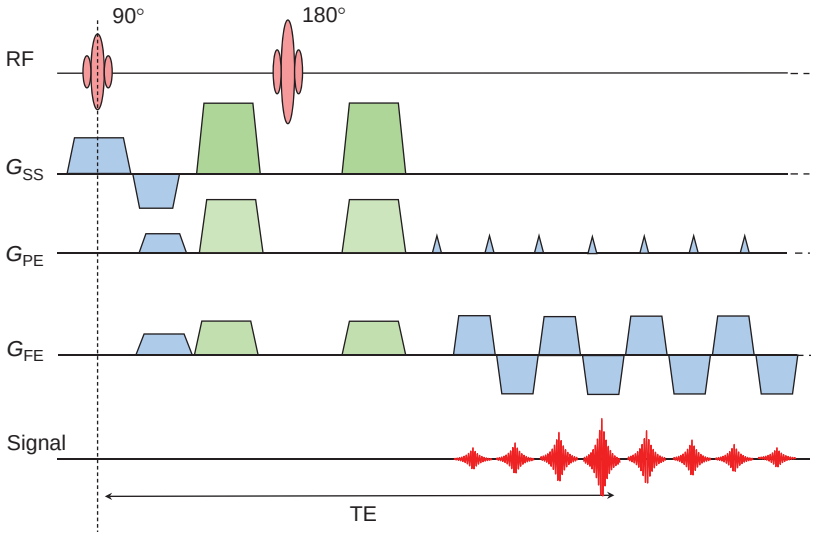

Figure 18.9 Diffusion-weighted EPI sequence. The diffusion gradients are shown in green and their relative amplitudes are varied to give the different directions.

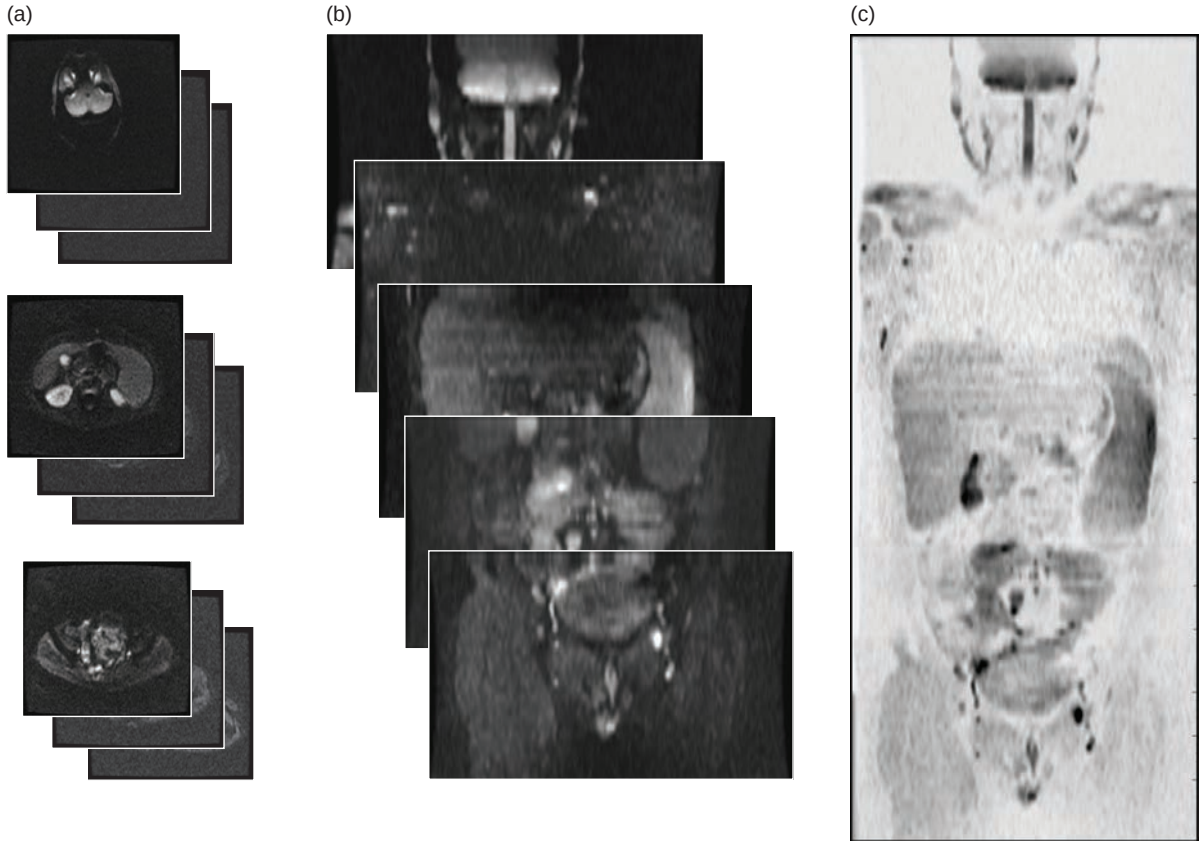

Figure 18.10 Whole-body diffusion – axial DWI images are acquired (a), then reformatted (b) and stitched together with contrast reversed (c).

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

18.3 Perfusion Imaging

Perfusion is a term that has different meanings to different professionals. Biotechnologists use it to mean the process of keeping tissues alive in a solution containing all the vital nutrients. In MRI it has a much more specific meaning, and refers to the capillary blood supply to a tissue, measured in ml min^{-1} g^{-1}. In the brain this is usually called the cerebral blood flow (CBF) or simply *f*. Two other measures are commonly used in perfusion MRI, the cerebral blood volume (CBV) and mean transit time (MTT). Often these terms are prefixed with 'r' (rCBF, rCBV and rMTT) meaning 'relative', since it is difficult to quantify them absolutely and it is usually preferable to find a ratio between the ipsi- and contralateral sides.

Two MRI techniques for imaging perfusion have been developed in the last few years to distinguish perfused tissue from unperfused, primarily for neurological applications (see Chapter 16 for cardiac perfusion). The quicker method is to image the required slices rapidly and repeatedly during the first pass of a gadolinium bolus, sometimes referred to as '**D**ynamic **S**usceptibility **C**ontrast **MRI**' (DSC-MRI). The other technique tags protons in the arterial blood supply with a magnetic 'label', then images the required slices with and without the labelling. Although a slower technique with poor signal-to-noise ratio (SNR), it is completely non-invasive and can be repeated as often as required (without having to wait for excretion of the gadolinium). It is known as 'arterial spin labelling' or 'arterial spin tagging', and there are several variations on this basic theme (which produce another rash of acronyms!). We will start by considering DSC-MRI, which to date has been more popular in clinical applications.

18.3.1 Dynamic Susceptibility Contrast MRI

In DSC-MRI a volume of tissue is imaged repeatedly using an EPI sequence. After a few images have been collected as a baseline, a bolus of gadolinium is injected as fast as possible. During the first pass through the intracranial circulation the gadolinium remains in the vasculature, causing a reduction of T_2 and T_2^*, which is seen as a dramatic drop in signal intensity on T_2-weighted or T_2^*-weighted images (Figure 18.11). The second pass may also be detected as a slight drop in intensity, before the signal returns to baseline. The whole imaging sequence takes no more than 2–3 min.

Originally gradient-echo EPI sequences were used for DSC-MRI since they are most sensitive to changes in T_2^*. However, it is preperable to use spin-echo EPI and measure changes in T_2, which has been shown to be due only to the microcirculation, excluding the confounding signals from larger arteries and veins. A TE of 35–60 ms is used to detect the T_2 changes, with a TR of no more than 1500 ms in order to maintain reasonable temporal resolution. This rather short TR usually restricts the volume coverage even on a system with high-power gradients, or requires rather thick slices. The exception is on Philips scanners with a 3D perfusion sequence called PRESTO, which uses a time-reversed gradient-echo sequence to produce a Hahn echo collected by an EPI acquisition. There should be at least five images (time-points) in the baseline section for analysis purposes; since a bolus injection in the antecubital vein typically takes 8–10 s to reach the brain, the injection should be started soon after the start of the imaging sequence. Many groups regard a power injector as essential, as an injection rate of 3–5 ml s^{-1} is necessary to achieve a good bolus. (Compare this with the requirements for contrast-enhanced MRA, where the timing of the injection is more critical relative to the image acquisition and the bolus shape should be rather longer.) If the injection is performed by hand, large-gauge IV tubing with a minimum of connections and a mechanism for rapidly switching to saline flush should be used. It is also possible to make rapid hand injection easier by warming the gadolinium to body temperature, which reduces the viscosity by a factor of two. Repeated imaging of the volume should continue for a total of 2–3 min.

Analysis of all these images in the clinical setting is best done with proprietary software on a workstation, although for quantitative results or research work it may be better to use home-written software which can be well controlled. Most software produces pixel-by-pixel maps of the required parameters and often colour scales are used for display (Figure 18.12), giving a similar 'look' to nuclear medicine or PET scans. At a simple level, so-called summary parameters can be measured directly from the signal intensity curves; for example, the area under the curve (or 'negative enhancement integral'), the bolus arrival time (t_{arr}) and the time to peak (t_p), and the peak height (C_p). These are roughly related to rCBV, rMTT

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

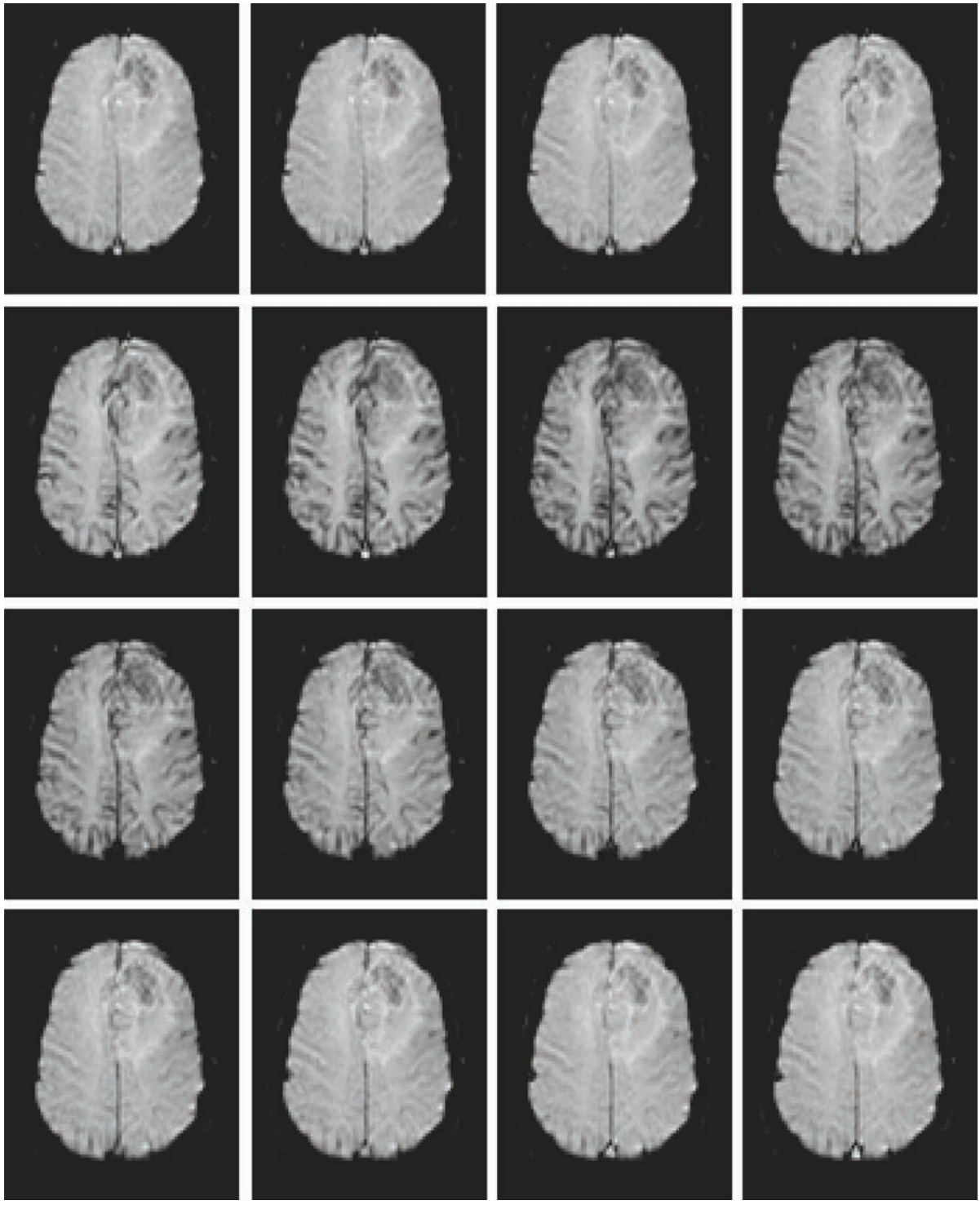

Figure 18.11 Sixteen images from a DSC-MRI examination of a low-grade glioma as the bolus of gadolinium passes through the vasculature. Time runs from left to right then down the rows, temporal resolution 1.2 s. As the bolus of gadolinium passes through the vasculature the signal intensities drop because of reductions in T_2^*.

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

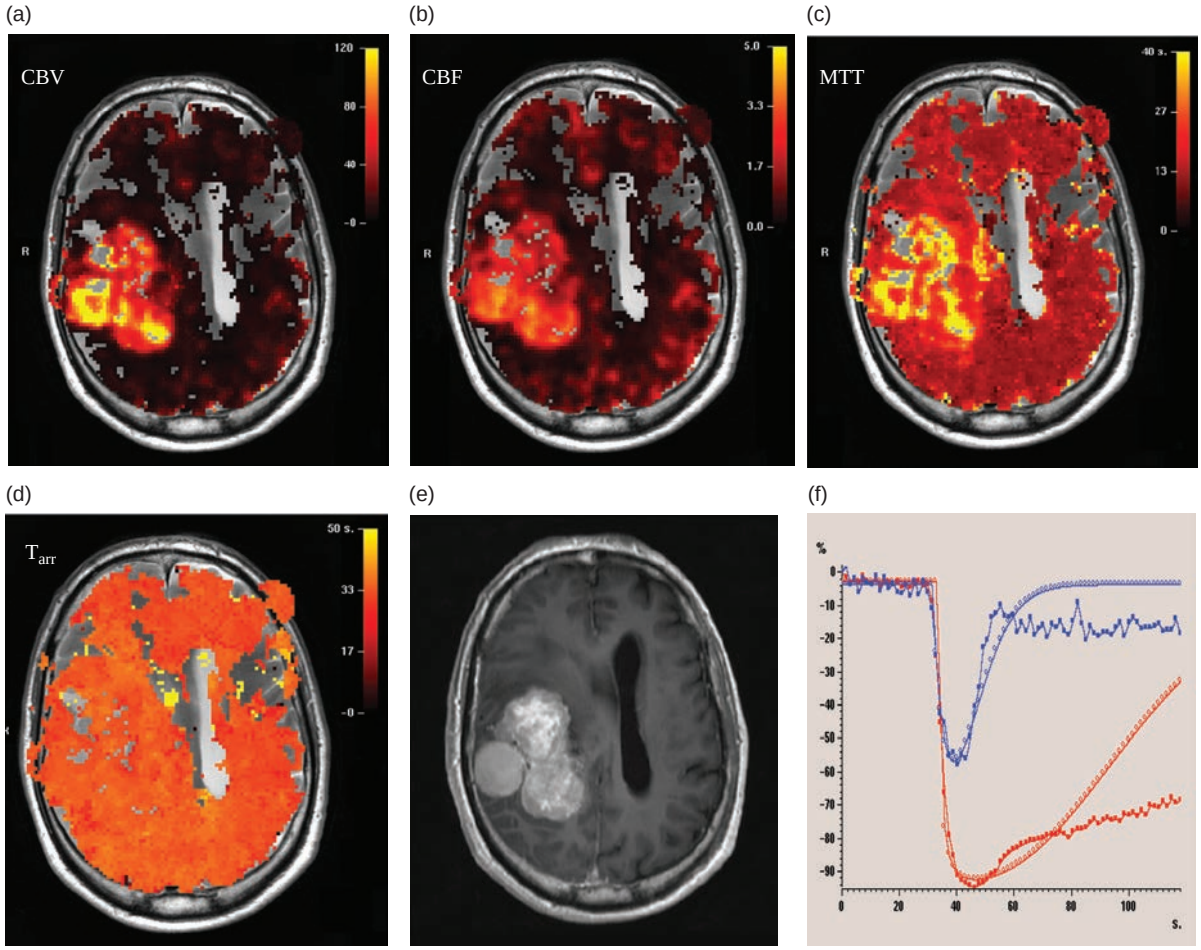

Figure 18.12 Perfusion parameter maps from DSC-MRI in a patient with a glioma. (a) CBV, (b) CBF, (c) MTT, (d) time of arrival (t_{arr}), (e) matching post-Gd T_1-weighted image, (f) time-intensity curves for normal brain (blue) and tumour (red). Note how the fitted gamma-variate curve deviates from the real data due to rapid leakage of Gd into the tumour.

and rCBF respectively, but have the disadvantage that they depend strongly on the shape of the bolus.

Full analysis for quantification uses deconvolution of the shape of the injected bolus, known as the arterial input function (AIF). This can be measured in the middle cerebral artery or any other easily identifiable vessel, but caution should be exercised in patients with stroke who may not have a normal arterial supply. Good SNR is essential for the deconvolution to work successfully and it is common to use spatial filtering (smoothing) to reduce the noise level in the images.

Advanced Processing and Quantification

The starting point for perfusion quantification, or for generating parametric maps, is to find the Gd concentration as a function of time for each voxel, $C_{tissue}(t)$. This can be done by assuming an inverse linear relationship between the T_2^* of the tissue and $C_{tissue}(t)$:

$$C_{tissue}(t) \propto \frac{1}{\Delta T_2^*} = -\frac{1}{TE} \cdot \ln\left[\frac{S_{tissue}(t)}{S_{tissue}(0)}\right]$$

where $S_{tissue}(t)$ is the tissue signal intensity at time t and $S_{tissue}(0)$ is the initial baseline signal intensity. From the concentration–time curve (Figure 18.13) we can measure some simple parameters such as the time of arrival (t_{arr}), time to peak (t_p) and maximum concentration C_p. Maps of these parameters are particularly useful in stroke studies, where a delayed t_{arr} is an indication of collateral blood supply. However, they are sensitive to differences in the bolus injection, which is a good reason for using a power injector in perfusion studies.

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

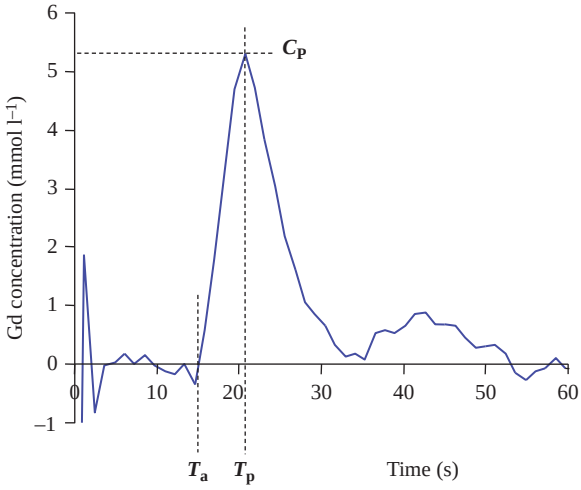

Figure 18.13 Concentration–time curve for white matter, showing the definition of various summary parameters. C_p denotes peak height; t bolus time arrival and t_p time to peak.

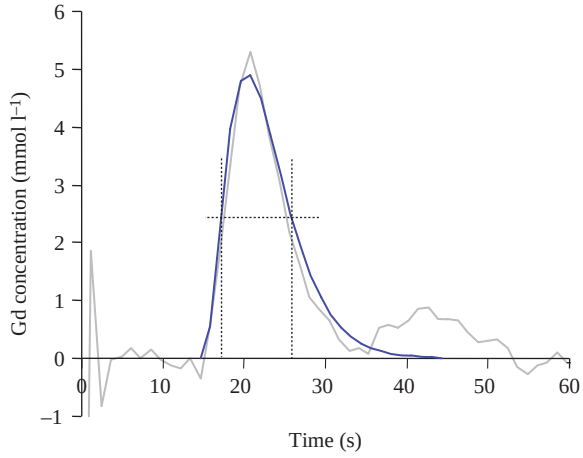

Figure 18.14 Gamma-variate fitting to the concentration–time curve. Starting points for the fitting parameters are found from the peak height and the times of the two half-heights.

For more accurate quantification, we use indicator dilution theory (see Further reading for some useful references). This is a well-known model which describes how an indicator (gadolinium) is distributed (diluted) in the blood supply following an instantaneous bolus injection. In the ideal case of an intact blood–brain barrier, the concentration–time curve has a gamma-variate shape:

$$C_{ideal}(t) = C_p \cdot \left(\frac{e}{rs}\right)^r \cdot (t - t_{arr})^r \cdot \exp\left(-\frac{(t - t_{arr})}{s}\right)$$

where e is exp(1), $e \approx 2.718$; r and s are related to the rate of increase and rate of decrease respectively. Of course the bolus is certainly not ideal, and the tissue concentration curve is modified by the shape of the input concentration curve, known as the arterial input function, AIF(t). Mathematically this is a convolution,

$$C_{tissue}(t) = C_{ideal}(t) \otimes AIF(t)$$

which means that if we measure the AIF, typically in a cerebral artery, we can find the ideal concentration–time curve by deconvolution. To do this, we divide the Fourier transform of $C_{tissue}(t)$ by the Fourier transform of AIF(t) and take the inverse Fourier transform of the result:

$$C_{ideal}(t) = FT^{-1}\left\{\frac{FT(C_{tissue}(t))}{FT(AIF(t))}\right\}$$

It is usual to fit a gamma-variate curve to the measured $C_{ideal}(t)$, using $r = s$ = full-width-half-maximum as starting values for the fitting (Figure 18.14). The fitted curve $C_{fit}(t)$ can then be integrated to provide the cerebral blood volume (CBV), mean transit time (MTT) and cerebral blood flow (CBF):

$$CBV = \frac{\kappa}{\rho}\int C_{fit}(t) \cdot dt$$

$$MTT = \frac{\int t \cdot C_{fit}(t) \cdot dt}{\int C_{fit}(t) \cdot dt}$$

$$CBF = \frac{CBV}{MTT}$$

where ρ is the density of brain tissue and κ is a constant that accounts for the difference in haematocrit between large and small vessels. There are many problems associated with quantification of perfusion from Gd bolus studies, not least being the requirement for good SNR for both the deconvolution and the fitting process. However, with care it is possible to achieve reliable results.

18.3.2 Arterial Spin Labelling

Arterial **S**pin **L**abelling (ASL) imaging is based on magnetically labelling protons in the arterial blood supply, usually with an inversion pulse. When images are then acquired from slices in the brain, there will be a very small signal loss compared with unlabelled images. This is because there is an in-flow enhancement even in the capillaries. The label decays in only

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

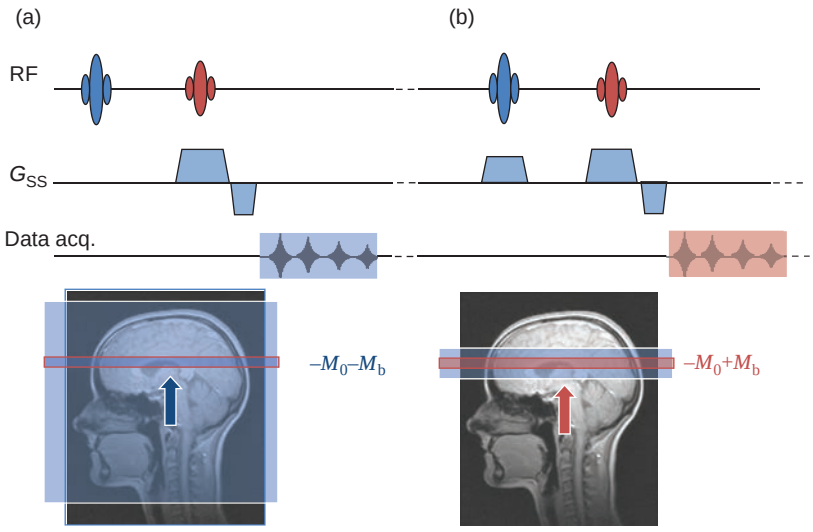

Figure 18.15 Schematic diagram for FAIR. (a) The labelled image is produced by applying a non-selective inversion pulse to the whole head while (b) the control image is produced with a slice-selective inversion pulse slightly wider than the image slice thickness.

4–5 s due to T_1 relaxation of the blood protons, so EPI is used for the image acquisition. There are several different ASL techniques, differing mainly in the way they apply the labelling and control pulses. We will only describe one, **F**low-sensitive **A**lternating **I**nversion **R**ecovery (FAIR), in order to illustrate the principles. You should check out the references at the end of this chapter for the other ASL techniques.

In the FAIR labelled images, an inversion pulse is applied to the whole brain before a single-slice EPI image is acquired (Figure 18.15). In the labelled image, inverted protons in the arterial supply are carried into the capillary bed of the imaged slice and exchange with protons in the tissues. Depending on the time delay between inversion and imaging, the signal intensity of protons in perfused tissue is reduced. For the control experiment the inversion pulse is applied only to the imaged slice. In the control experiment, arterial blood protons outside the slice, which are unlabelled, are carried into the image slice during the TI and the image will have a slightly higher SNR overall. Large bipolar crusher gradients (effectively diffusion weighting) are added to the pulse sequence to remove the signal in larger blood vessels, so one of the advantages of ASL over DSC-MRI is that it only measures capillary perfusion.

There are two main classes of ASL pre-pulse: pulsed and continuous. The pulsed ASL schemes include STAR, FAIR and PICORE, and can be easily understood; they simply apply a magnetization label to the up-stream arteries, which is then carried to the image volume. Pulsed ASL lends itself to time-resolved techniques, which can track the magnetized bolus as it travels through the vasculature and into the brain. Continuous ASL uses a much longer RF pulse, most commonly as a train of smaller RF pulses, known as pseudo-continuous ASL (PCASL). The train of pulses continues until a steady state is reached in the perfused tissues, typically 1800 ms, and is followed by a delay to allow time for the tagged protons to reach the imaging volume, another 1800 ms. PCASL provides much higher SNR than pulsed techniques, and especially when combined with a 3D readout sequence, gives diagnostic image quality in a very acceptable 3–4 min scan time. This is becoming the 'gold standard' for ASL, both in clinical practice and in research studies.

Clinical Applications of ASL

In the last few years, PCASL has become widely available on commercial scanners. With a 2D SE-EPI readout, or a 3D GRASE readout, this sequence is relatively robust and offers good SNR with a scan time of around 4 min. Although this is easily within the scope of many clinical centres, ASL is still not routine practice everywhere. One reason is that ASL is very sensitive to patient motion, and special techniques are needed to correct the images in post-processing. Also, most commercial packages offer only the perfusion map as output; it is not possible to calculate the arrival times of the tagged protons in ASL. This puts ASL at a disadvantage compared with DSC-MRI, or with CT-perfusion techniques, which can produce bolus arrival-time maps.

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

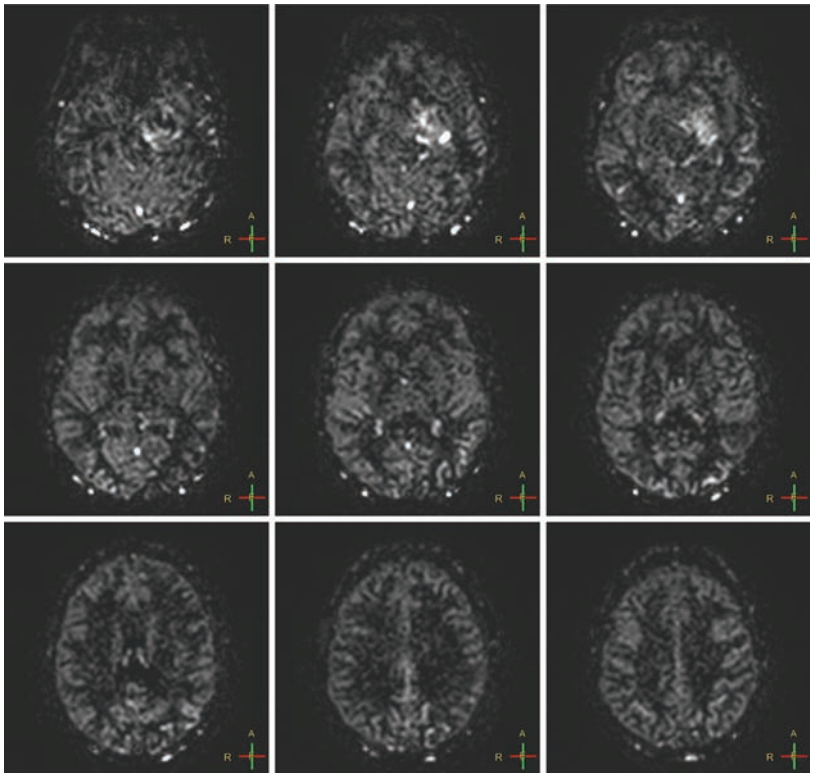

Figure 18.16 Arterial spin labelling perfusion images, showing higher perfusion as bright signal intensity. The focal lesion in the left basal ganglia has higher perfusion than normal grey matter and white matter.

In acute stroke, CT is still the preferred diagnostic method due to its speed and greater availability. In sub-acute stroke, and in other neurological diseases, ASL has good potential to add diagnostic value, particularly when the quantitative ASL methods are fully validated. For example, PET scanning shows a distinctive pattern of brain perfusion in dementia; in the future, this could be achieved using ASL at lower cost and without radioactive agents.

Outside of the brain, ASL is being used to image renal and liver perfusion. However, these methods are still in the research domain as they require sophisticated motion-compensation to correct for respiration. Non-neuro applications will continue to grow in the coming decade.

ASL Analysis

The difference in signal intensity can be shown to be related to M_0, T_1 and perfusion of the tissue. The theory for ASL is not derived from the indicator dilution theory, since the tracer (labelled water) is freely diffusible whereas gadolinium is merely transported by the blood. Instead the theory is derived from the Bloch equations using a two-compartment exchange model for water in the blood and tissues. For the FAIR technique

$$\Delta M = 2M_0 \frac{f}{\lambda}\left[\frac{\exp(-\mathrm{TI}\cdot R_{1a}) - \exp(-\mathrm{TI}\cdot R_1)}{R_1 - R_{1a}}\right]$$

where f is tissue perfusion, TI is the inversion time, R_1 is the relaxation rate ($1/T_1$) of the tissue and R_{1a} is the relaxation rate of arterial blood. λ is a constant called the blood–brain water partition coefficient, which is usually assumed to be 0.9 (i.e. 90% of water is in the brain tissues and only 10% in the intracranial blood vessels). Although literature values for M_0 and T_1 could be used in theory, in practice it is usual to acquire images at a series of TIs, and to calculate M_0 and T_1 maps from the control images. With the M_0 and T_1 maps the signal differences can be fitted to the above equation to produce a perfusion map (Figure 18.16).

18.4 Dynamic Contrast Enhancement: Permeability Imaging

DCE is generally used to describe the process of dynamically acquiring MRI images during the

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

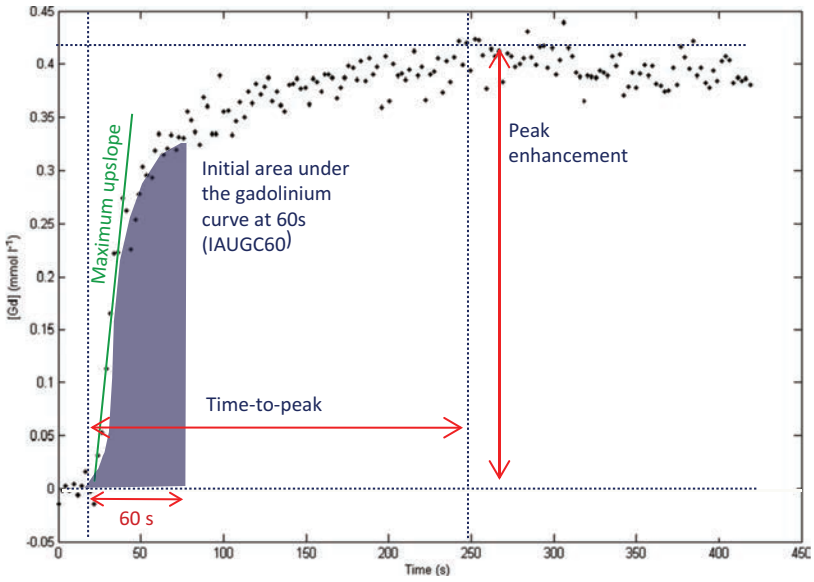

Figure 18.17 Four examples of non-model-based analysis metrics for DCE data (data points).

passage of an exogenous contrast agent, usually a standard chelate of gadolinium. In comparison to dynamic susceptibility contrast (DSC) techniques that exploit differences in T_2^* to create contrast in the images, DCE techniques rely on differences in T_1. Depending upon the anatomy of interest, a T_1-weighted acquisition may be performed either continuously during the wash-in and wash-out of the contrast agent or at specific time-points. In the liver, for example, it is common to acquire 3D T_1-weighted data pre-contrast, at the peak of arterial enhancement (arterial phase) and at the peak of venous enhancement (portal venous phase). A subsequent (delayed) phase may also be acquired. Since different lesions exhibit differential enhancement it is a requirement to have these different temporal phase acquisitions. If the data are acquired with a sufficiently high temporal resolution, then it is possible to perform pharmacokinetic analysis of the contrast agent uptake and wash-out in order to provide quantitative or semi-quantitative indices of the tissue, or more commonly, the tumour vascularity. The blood supply to a tumour is typically made up of very chaotic and immature blood vessels which have a very 'leaky' endothelium. If a contrast agent is administered as a tight bolus then it will leak out of the vessels into the tumour, causing a transient signal intensity increase in the T_1-weighted images. If images continue to be collected then the signal may start to decrease due to wash-out of the contrast agent.

There are various approaches to the quantification of contrast agent uptake. These can be divided into either model-based or non-model-based. The latter, non-model-based, methods use semi-quantitative indices that make no assumptions about the kinetics of the contrast agent uptake (Figure 18.17). For example, we can measure time-to-peak, maximum upslope, peak enhancement and the area under the signal-time curve for a fixed period after the start of contrast uptake. If the signal intensity is converted to [Gd(t)] then the latter index is referred to as the **I**nitial **A**rea **U**nder the **G**adolinium **C**urve (IAUGC). This often has a subscript referring to the number of seconds for which the area is calculated, e.g. $IAUGC_{60}$ or $IAUGC_{90}$.

Model-based methods to quantify contrast enhancement use a mathematical equation with a number of unknown parameters. The best-fit of the experimental data to the model provides values for these parameters. One of the most common models in use for DCE-MRI data analysis is the Tofts' model. This model is a simplification of the in vivo situation and assumes two 'compartments'; the vasculature (or more correctly the blood plasma) and the tissue into which the contract agent leaks, more correctly called the extravascular, extracellular space (EES), i.e. the interstitial space that contains neither blood vessels or cells. The Tofts' model is a fairly complex mathematical equation (see Box 'Mathematical Modelling: The MRI Catwalk') that yields a number of quantitative indices such as K^{trans}, k_{ep} and v_e (Figure 18.18).

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

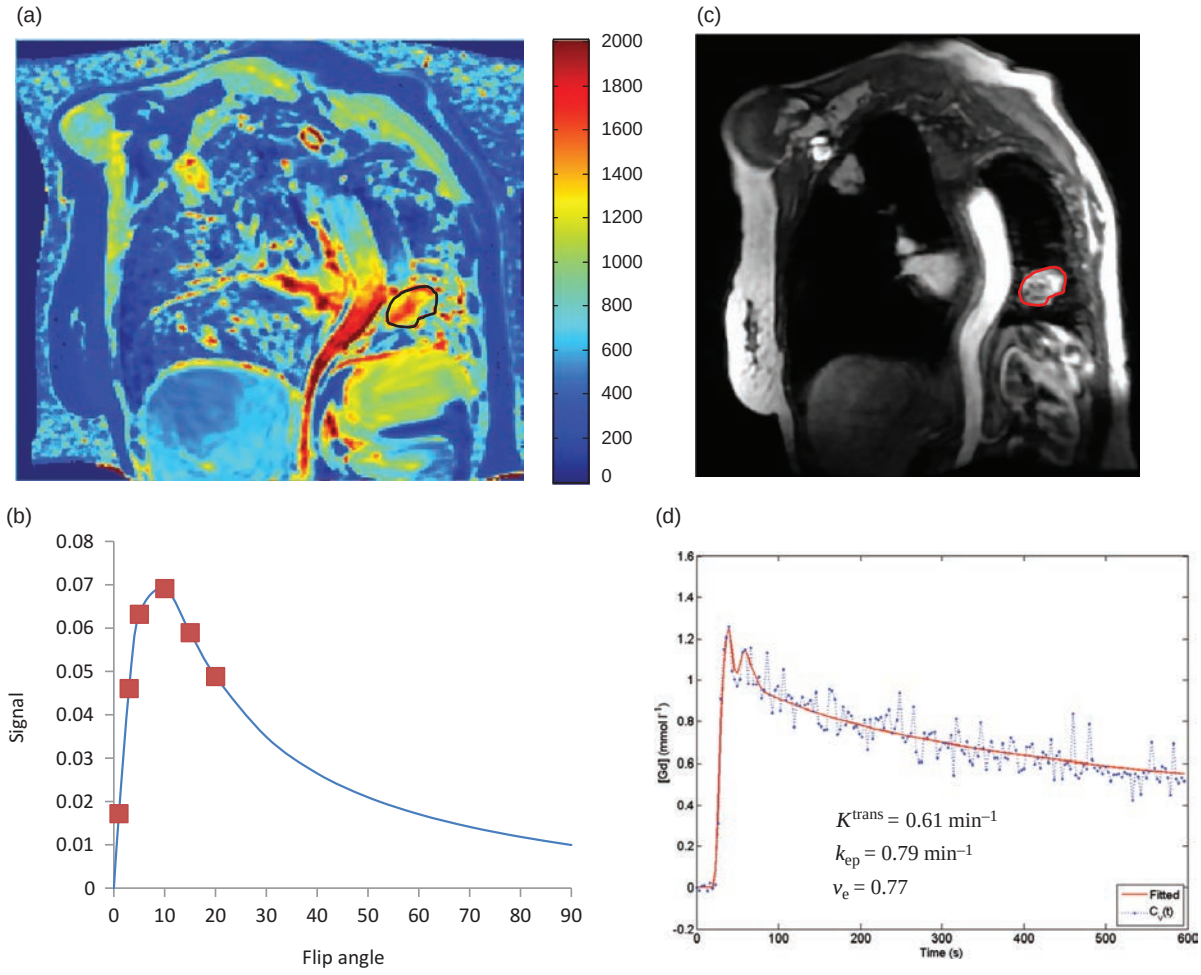

Figure 18.18 (a) shows a T_1 map calculated from six different flip angles using the signal equation for a T_1w spoiled gradient-echo sequence. The tumour is outlined. (b) Shows the signal/flip angle curve for the tumour ROI that was used to calculate the T_1. The six flip angles are marked. (c) Shows a single phase from the DCE acquisition at peak enhancement in the ROI. (d) Shows the Gd concentration time curve (blue data points) and the fitted Tofts' model (red line). Courtesy of Dr Andrew Gill, University of Cambridge, UK.

A large number of research studies have shown that these indices may be more sensitive biomarkers of tumour changes than standard measurements, e.g. the RECIST (**R**esponse **E**valuation **C**riteria **I**n **S**olid **T**umours) method which involves simply measuring the largest diameter of the tumour.

Mathematical Modelling: The MRI Catwalk

Using the standard Tofts' model the uptake of the contrast agent in a region of interest (ROI) can be approximated using the following equation

$$C_{tissue}(t) = \frac{K^{trans}}{1 - Hct} \cdot C_{art}(t) \otimes \exp\left(-k_{ep}(t - \tau)\right)$$

where $C_{tissue}(t)$ and $C_{art}(t)$ are the Gd concentration in tissue and arterial blood respectively, both measured as a function of time; Hct is the haematocrit; K^{trans} is the transfer constant from blood plasma to the EES; k_{ep} is the transfer constant from EES back to blood plasma; τ is the onset time of arterial Gd contrast uptake; and v_e is the total EES volume. $\otimes$ is a convolution operator. K^{trans}, k_{ep} and v_e are related by the equation

$$v_e = \frac{K^{trans}}{k_{ep}}$$

So, although you might expect the two transfer constants to have the same units, they don't. K^{trans} is measured in ml g^{-1} s^{-1}, while k_{ep} is measured in s^{-1}, and the EES volume v_e is measured in ml g^{-1}.

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

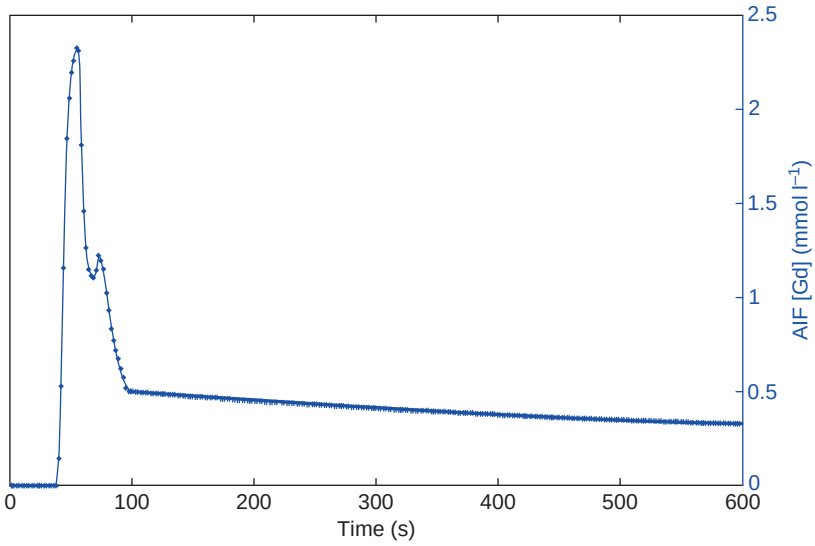

Figure 18.19 The model arterial input function (AIF) that was used in the Tofts' analysis for the data in Figure 18.18. Courtesy of Dr Andrew Gill, University of Cambridge, UK.

Since MRI only reports arbitrary signal intensities it is necessary to first determine the actual gadolinium concentration as a function of time [Gd(t)]. If we make the assumption that there is a linear relationship between signal intensity (S) and [Gd] then

$$[\mathrm{Gd}(t)] = \frac{S(t) - S(0)}{S(0) \cdot \mathrm{T}_{1,0} \cdot r_1}$$

where $S(t)$ is the signal at each time-point t, $S(0)$ is the signal pre-contrast injection, i.e. $t = 0$, $\mathrm{T}_{1,0}$ is the T_1 of the tissue prior to contrast agent injection and r_1 is the known T_1 relaxivity of the contrast agent. However, this linear relationship assumes that the TR of the sequence measuring $S(t)$ is approximately the same as $\mathrm{T}_{1,0}$ and that T_1 is independent of time. Therefore a non-linear model based on the signal equation of a T_1w spoiled gradient-echo sequence with known TR and flip angle $\alpha°$ can be used

$$S = S_0 \frac{\sin\alpha(1 - E_1)}{1 - \cos\alpha E_1}$$

where $\mathrm{E}_1 = \exp(-\mathrm{TR}/\mathrm{T}_1)$. This can be rearranged as follows

$$\frac{1}{\mathrm{T}_1(t)} = \frac{1}{\mathrm{TR}} \cdot \ln\left(\frac{S_0 \sin\alpha - S_0 \cos\alpha}{S_0 \sin\alpha - S(t)}\right)$$

Using this equation [Gd(t)] can be calculated from

$$[\mathrm{Gd}(t)] = \frac{1}{R_1}\left[\frac{1}{\mathrm{T}_1(t)} - \frac{1}{\mathrm{T}_{1,0}}\right]$$

It is therefore necessary to know the value of $\mathrm{T}_{1,0}$ for every voxel. There are various methods of determining $\mathrm{T}_{1,0}$ (see Section 19.2), with the multiple flip angle method quite popular for DCE techniques since the same 2D or 3D spoiled gradient-echo imaging sequence can be used.

While it is reasonably straightforward to determine $C_{\mathrm{tissue}}(t)$, it is rather more difficult to determine $C_{\mathrm{art}}(t)$, which is often termed the arterial input function (AIF) or vascular/vessel input function (VIF). First, it may be difficult to localize the main vessel feeding the tumour; second, since $C_{\mathrm{art}}(t)$ changes very rapidly, very high temporal resolution is required to accurately capture the AIF. For these reasons a number of 'model-based' AIFs have been described in the literature (Figure 18.19). These have typically been derived from either blood-sampling or by averaging a population of measured AIFs. The choice of AIF is still an area of contention and will depend upon the tissue being studied. Likewise, there are a number of other DCE models described in the literature and the Tofts' model may not be optimal in all applications.

18.5 Brain Activation Mapping Using the BOLD Effect

In recent years the term functional MRI or fMRI (note the small 'f') has become synonymous with brain activation imaging using the BOLD effect. This is in some ways related to ASL perfusion imaging, except that intrinsic T_2^* contrast rather than T_1 contrast is utilized. In BOLD fMRI we 'see the brain working'.

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

18.5.1 The BOLD Effect

The BOLD (**B**lood **O**xygenation **L**evel **D**ependent) effect was observed at the start of the 1990s in animal-based experiments. It was known that oxyhaemoglobin is diamagnetic (i.e. essentially non-magnetic) and that deoxyhaemoglobin is paramagnetic. This means that deoxygenated blood has a shorter T_2^* and hence lower MR signal than fully oxygenated blood. That the opposite was observed in visual cortex during sensory stimulation was initially surprising – the MR signal increasing or 'lighting up' at times when oxygen consumption was heightened due to neuronal activity as in Figure 18.20. The scientific consensus is now that the increased consumption of oxygen by neurons during activation is accompanied by a disproportionate increase in the supply of fully oxygenated blood, so that downstream from the site of activation the concentration of deoxyhaemoglobin decreases and so T_2^* is elevated and the MR signal increases. Gradient-echo EPI, with its strong T_2^* weighting, is the sequence of choice for many fMRI examinations.

Brain or Vein

In fMRI we detect blood oxygenation changes in the draining veins, i.e. downstream from the actual activation site (Figure 18.20). The difference in magnetic susceptibility between fully oxygenated and deoxygenated blood is about 9.5×10^{-7}. However, fully oxygenated blood has a magnetic susceptibility χ similar to that of the extravascular space in grey matter and so differences in oxygenation will affect the local field homogeneity and therefore change T_2^*.

The BOLD effect during activation is illustrated in Figure 18.20, where following neuronal activity, although there is a greater number of deoxygenated red blood cells, the increase in fresh oxygenated blood delivery results in a reduction of the concentration of deoxyhaemoglobin and therefore T_2^* increases, as does the MR signal.

MR perfusion can also detect neuronal activation by its sensitivity to the blood flow changes in the capillaries. It is thought that perfusion fMRI may pinpoint the actual activation site more directly than BOLD fMRI, although the sensitivity of BOLD is higher. The localized blood flow changes in activation can also be detected using non-EPI sequences such as spoiled gradient echo, although the high sensitivity to flow may result in vessels on the cortical surface being mistaken for activation.

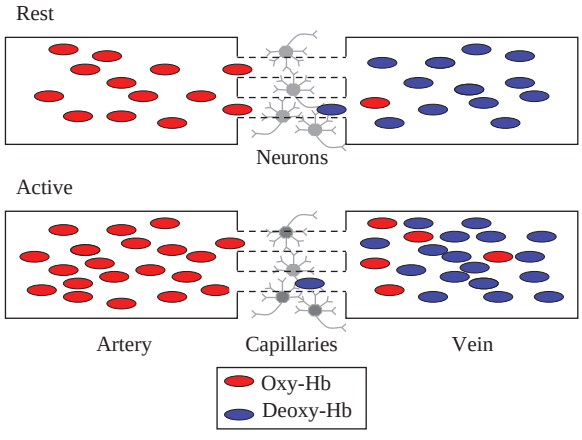

Figure 18.20 The origin of the BOLD effect. In activation (below) the over-provision of fully oxygenated blood leads to a reduction in deoxy-Hb and an increase in local T_2^* in the draining veins compared with the rest condition (above).

18.5.2 fMRI Acquisitions

To say that the 'brain lights up' during activation is a bit of an exaggeration as the actual signal intensity changes are no more than a few per cent. In fMRI we detect these by modulating the oxygenation level at the site of brain activity and look for correlated signal changes. Rapid scanning, usually with EPI, is carried out continuously while the subject performs various tasks (known as the paradigm). These are commonly arranged in a block design with periods of activity interspersed with periods of contrasting activity or rest. The periods of activity might involve motor tasks, stimulus presentation or cognitive activity (e.g. generating words, doing mental arithmetic, etc.). A block length or 'epoch' will typically be about 30 s, with perhaps three or four complete cycles of the two contrasting tasks. This is shown in Figure 18.21 for a simple on–off visual stimulus and in Box 'Clinical fMRI'.

Block design paradigms are robust and simple to arrange. However, for some mental tasks it may be difficult to generate extended periods of activation to form an epoch. Some brain events, e.g. hallucinations resulting from psychosis, may be of a transient and unpredictable nature. One solution to this is to use so-called **E**vent-**R**elated (ER) fMRI. In ER-fMRI scanning is carried out continuously at a higher image acquisition rate (e.g. once per second) and we simply

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

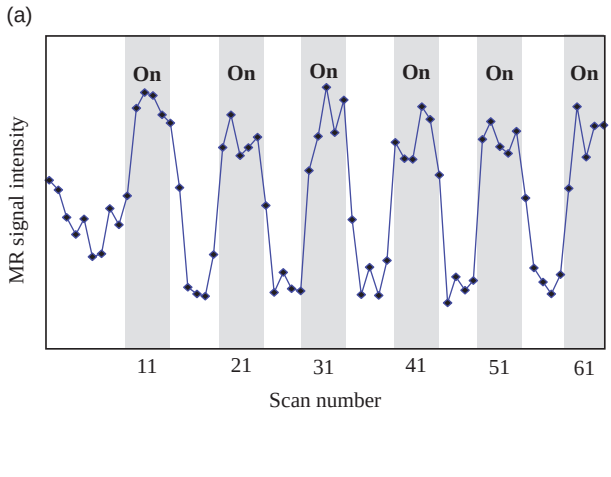

Figure 18.21 (a) fMRI BOLD signal changes and (b) area of activation in the visual cortex resulting from (c) a visual (reversing chequerboard) paradigm. The first scans are used to establish equilibrium and are discarded from the analysis. A haemodynamic delay is apparent between the paradigm changes and the signal response.

wait until the event occurs and look for corresponding changes in the MR signal that resemble the haemodynamic response of activation (see Box 'Haemodynamic Delay and Convolution').

Having acquired the time series of scans, there is much post-processing required. This is normally done offline on an independent computer. Finally, an image is produced that represents the statistical significance of signal changes correlated with the paradigm.

18.5.3 fMRI Processing

The common steps of fMRI processing are illustrated in Figure 18.22. First, the images are realigned or co-registered with themselves. This is necessary because even the smallest shift in the position of a voxel can generate significant signal changes. Stimulus-correlated motion would result in false-positive 'activations' unless removed in this way. Normally the first few volumes will be discarded to ensure that the magnetization is in a steady state, usually required for the realignment algorithm to be accurate. Post-acquisition realignment is required even if prospective slice positioning using navigator echoes has been used for the acquisition.

Data Meltdown

A typical fMRI scan may involve 24 slices of the brain, acquired with in-plane resolution of 128 × 128. We may collect 100 volumes in a single run, performing several runs on each subject. By the time we have done all the processing the data volume will have been at least trebled or quadrupled.

$$\text{Total raw data per run} = 24 \times 128 \times 128 \times 100 = 39\,321\,600 \text{ voxels} = 80 \text{ MB}$$

and this represents 5 min scanning only. Generating several gigabytes (1 GB = 1000 MB) in one day is not hard.

A second step is to spatially normalize the images to a standard brain space, often to the Talairach Brain Atlas. This aids the neurological interpretation of the resultant activation maps and it also allows data to be averaged over groups of subjects. For individual patients with structural brain abnormalities or tumours this step is omitted.

The third step is to smooth the data. This helps to boost the SNR but has to be applied carefully to avoid excessive loss of spatial resolution. A fourth step may be to de-trend or normalize the data according to an overall or global mean or to apply a high-pass temporal filter. This step is designed to remove any bias resulting from scanner drift, for example, over the acquisition.

Next, statistics are calculated. Various levels of sophistication are involved here, but the simplest is to subtract the rest images from the active images and look for significant increases in signal, characterized by the Z-score defined as

$$Z = \frac{\text{mean signal difference}}{\text{standard deviation}}$$

In practice Z needs to be greater than 3 for any degree of confidence in the results. Z-scores are related to p-values commonly quoted with regard to the normal

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

distribution (*t*-test). The resulting brain activation map is made by displaying only those voxels that have the appropriate statistical parameter (e.g. Z-score, *p*-value) greater than a given statistical threshold. Sometimes spatial extent thresholding is carried out to exclude isolated voxels or small groups and only show clusters of activation. Sophisticated analyses can be applied to obtain better inferences of the statistical significance of clusters of activated voxels. A number of statistical approaches other than *t*-tests are possible, including correlation, Fourier, wavelet and independent component analyses. Software packages are available through the academic domain as well as those supplied by MR manufacturers. It is common for an fMRI statistics program to model the haemodynamic effect either by introducing a fixed or variable time delay, or by a more complex convolution of the paradigm waveform over the time series with a notional haemodynamic response curve.

Finally the resultant statistical maps are combined with underlying anatomical information. It should be recognized that there may not be full geometric correspondence between the EPI BOLD images and the anatomical images (unless a similar EPI protocol is used for both). This is due to the inherent image quality limitations of EPI.

Haemodynamic Delay and Convolution

One of the complexities of BOLD imaging is that the effect we are detecting lags behind the actual firing of the neurons by as much as 6 s. This is known as the haemodynamic delay and can be allowed for in the data analysis. Similarly, the effect will outlast the neural activation to a similar degree. This makes the temporal resolution of fMRI rather limited. There is also a considerable undershoot post-activation. The effects of the delay and undershoot can be seen in the actual data of Figure 18.21 and schematically in Figure 18.22.

Mathematically the combination of haemodynamic delay with the stimulus waveform, a boxcar or series of impulses, can be described by a

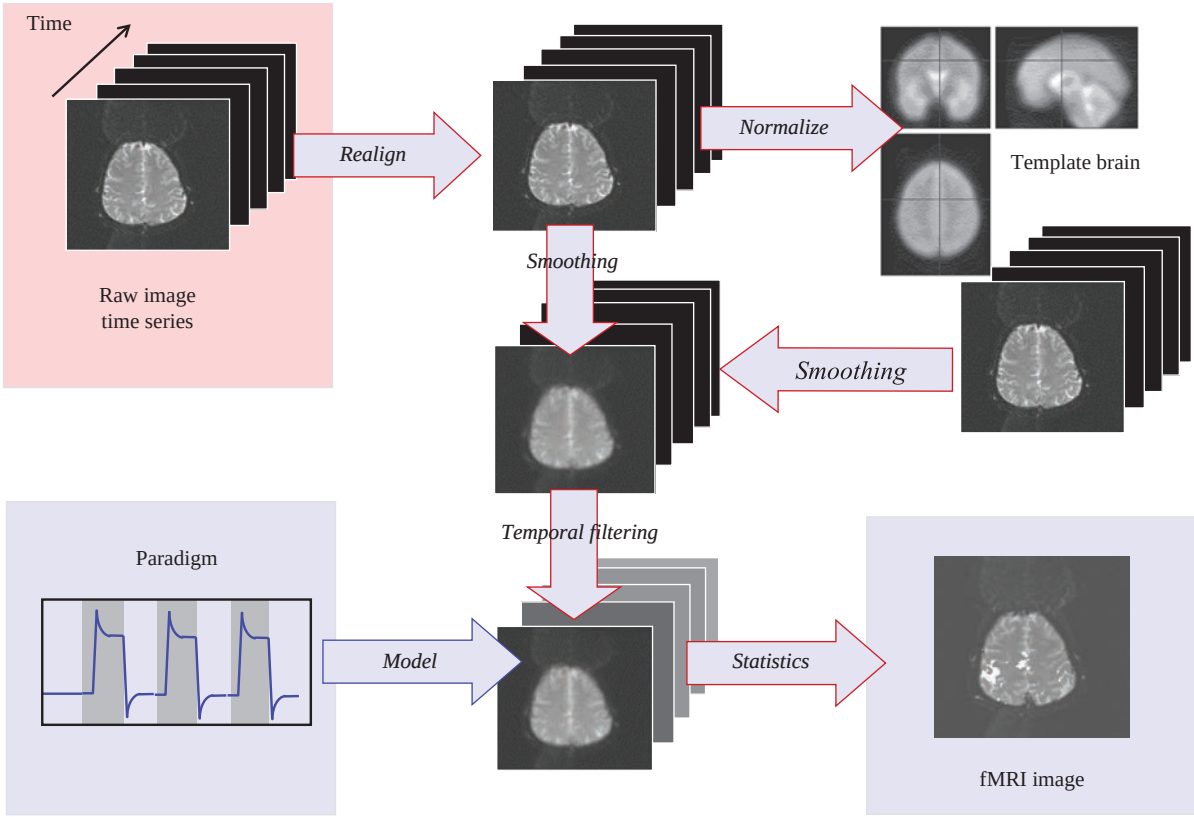

Figure 18.22 fMRI processing from the input of a time series of EPI data and a paradigm. The final image is a superposition of a statistical map on a raw image.

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

convolution process. By deconvolution of the actual response the haemodynamic response function (HRF) may be obtained. In the case of an impulse stimulus (because of the mathematical properties of the delta-function) the MR signal intensity changes can directly yield the HRF, although correction for the exact timing of the slices with respect to the stimulus timing is required.

One further feature of the HRF is an initial MR signal dip, thought to be due to the initial oxygen consumption increase, before the increase in blood delivery which causes the BOLD signal kicks in. This dip is usually too rapid to be observed with standard EPI sequences.

18.5.4 Interpreting fMRI: 'Blobology'

What are we looking at in a functional MR image? fMRI brain activation maps are normally presented as coloured 'blobs' superimposed on a greyscale anatomical background image, or as a colour overlay on a 3D surface-rendered image of cortical grey matter, as in Figure 18.23.

It is important to understand that the coloured blobs do not themselves signify brain activation, but represent areas of statistically different MR signals. The intensity of the colour represents the degree of statistical confidence that a voxel value or a group of voxels has changed according to the 'paradigm'. Actual signal changes may only be a small percentage or less. For this and other reasons the clinical interpretation of individual fMRI examinations is fraught with difficulty. BOLD signals can go down as well as up, and this so-called deactivation may be artefactual, due to inhibitory brain processes, or simply the result of the subject's mental processes, e.g. shifts in attention.

Clinical fMRI

Clinical fMRI can be used for patient selection and pre-surgical planning for epilepsy surgery or resection of brain tumours, or in the clinical evaluation of

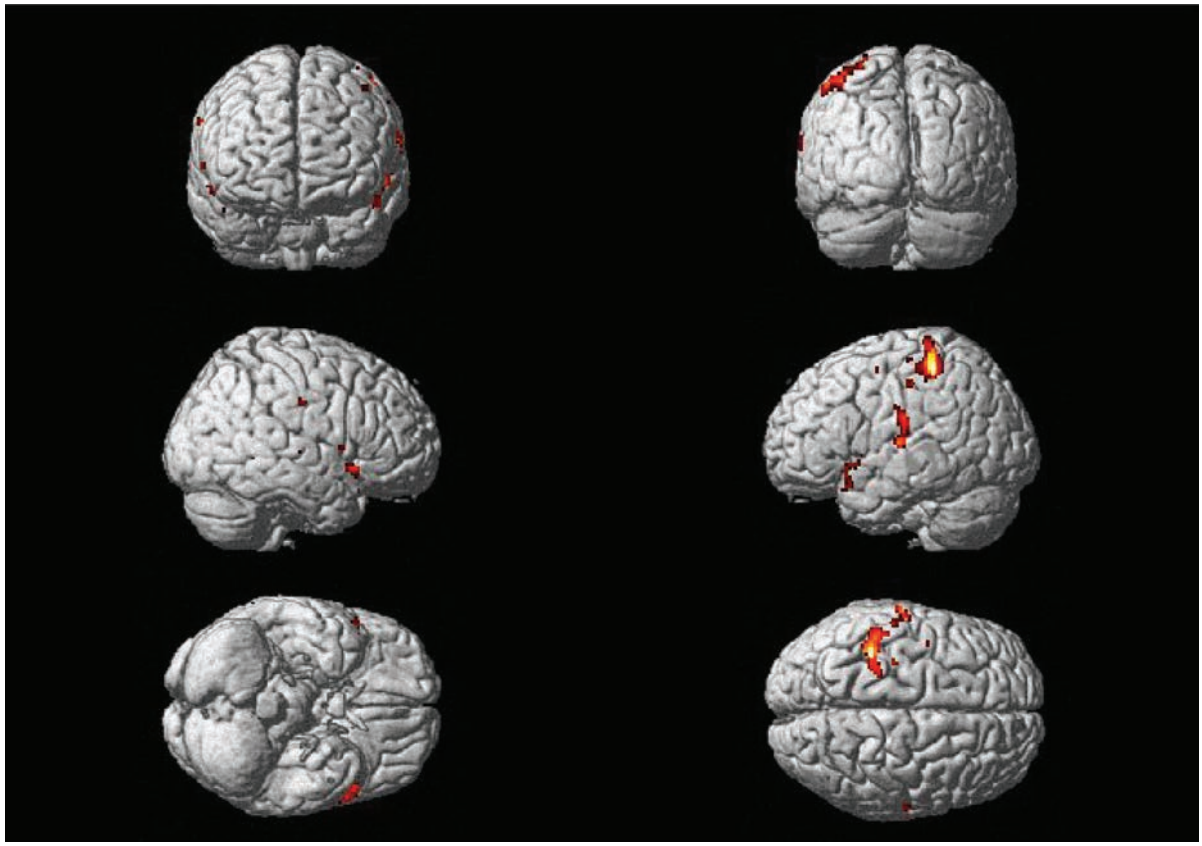

Figure 18.23 Colour fMRI rendered map. The statistical map is normalized to a standard brain atlas and rendered in 3D to show the activated areas.

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

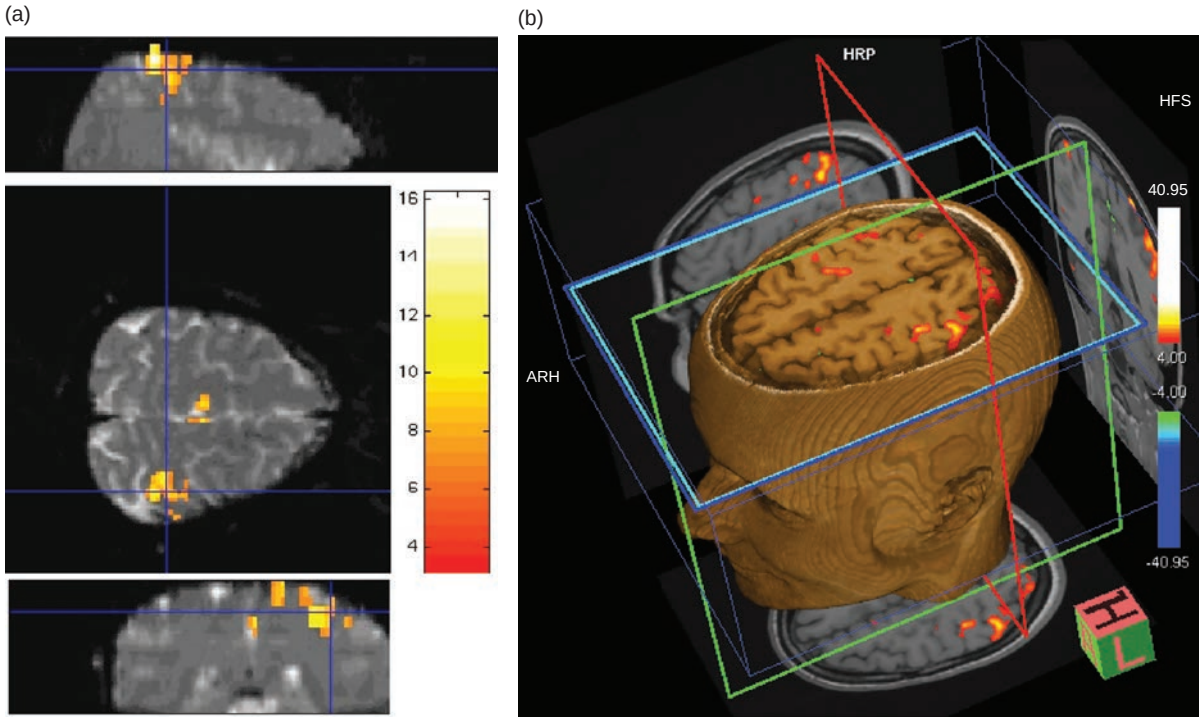

Figure 18.24 Clinical fMRI (a) orthogonal views with colour overlay showing the activation. (b) Real-time 3D display shown during the acquisition.

brain plasticity following injury. It can help with decision-making for patients with low-grade astrocytomas and arterio-venous malformations (AVMs) who have slight neurological impairments. It may be used in repeat studies for patients with slow-growing tumours or congenital lesions. Most clinical examinations are of the motor cortex for hand, foot and facial movement control (Figure 18.24a). One specialized but promising application is in the determination of the hemispheric dominance of language, as an alternative to the invasive Wada test.

Typically in clinical fMRI the question is, 'if we operate, how close will the surgeons have to go to particular key functional areas of the brain?' Getting a definitive answer to this is difficult for a number of reasons. Patients may have some brain abnormality, AVM, tumour, etc., which makes spatial normalization to a standard brain atlas impossible. They may have impaired ability to cooperate, e.g. patients with an impaired motor function, and may find it harder to keep still when exercising their affected side. There is much scope for error and misinterpretation arising from false-positive and false-negative responses. Pre-surgical fMRI findings can be confirmed at the time of operation by direct electrical stimulation in the open-skull situation. Further research will determine the role for fMRI in psychiatric illness.

Clinical fMRI doesn't have to be a long-winded examination. A typical clinical fMRI protocol might consist of a scout scan (three orthogonal views), followed by a 3D structural scan, e.g. MP-RAGE with matrix 256 × 256 × 120 and scan time 8 min. This would be followed by fMRI EPI scans for each functional area, each taking typically 5 min scan time. The EPI scans would have 24 transverse slices parallel to the anterior commissure–posterior commissure (AC–PC) line, 5 mm slice thickness, in-plane resolution 96 × 96 or 128 × 128, 220 mm field of view, 35–50 volume acquisitions, interscan interval ('TR') 3–6 s, block paradigm with a repeating pattern of 30 s activity followed by 30 s rest. To establish the position of the motor cortex requires two paradigms: first with left-hand finger motion versus rest, and second with right-hand finger motion versus rest. The total scan time for this examination is approximately 20 min plus patient handling time.

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019

18.5.5 Paradigm Shift

fMRI is extremely difficult to do well. Very good subject cooperation is required, even a small degree of patient head movement is intolerable and there are lots of data and much complicated analysis; but, more fundamentally, the paradigm itself may be flawed. Except for simple sensorimotor fMRI, it is quite hard to be certain that the subject is actually performing the task properly. Overt speech generally causes excessive head motion. Moreover, for subtle cognitive effects, the paradigm has to be psychologically effective, easily transferable to the inside of a scanner and capable of generating detectable BOLD responses. The scanner environment, with its excessive noise and claustrophobia-inducing space, presents a major limitation. Special stimulus delivery systems are required for the presentation of visual and audible stimuli. Despite these technical difficulties, fMRI opens up new avenues of research into cognitive neuroscience, psychiatric illness and neuropathology that are not possible with positron emission tomography (PET) due to its limited availability and the radiation dose involved.

T_2^* (Star) Wars: The Field Strength Question

As the BOLD effect is related to MR susceptibility, it should increase with field strength. Add to this the increase in SNR with B_0 and we would therefore expect the sensitivity of BOLD experiments to increase dramatically with the field strength. Reports in the literature suggest an increase in BOLD contrast of 30–40% for simple sensory-motor activations for 3 T compared with 1.5 T and greater again at 7 T. However, the overall picture is more complex as T_2^* is reduced at higher field strengths (T_2^* depends upon the field inhomogeneity in absolute units, i.e. microtesla). Moreover, in areas of particular neurological significance, e.g. the frontal or temporal lobes, the distortion and signal drop-out due to susceptibility differences may be excessive. Improving upon this is a particular focus of high-B_0 MRI research.

See also:

- Combining gradient and spin echoes: Section 12.5
- GE echo planar imaging: Section 13.4.4

Further Reading

Alsop DC, Detre JA, Golay X, *et al.* (2015) 'Recommended implementation of arterial spin-labelled perfusion MRI for clinical applications: a consensus of the ISMRM perfusion study group and the European consortium for ASL in dementia'. *Mag Res Med* 73:102–115.

Buxton RB (2009) *Introduction to Functional Magnetic Resonance Imaging: Principles and Technique*, 2nd edn. Cambridge: Cambridge University Press.

Calamante F, Thomas DL, Pell GS, Wiersma J and Turner R (1999) 'Measuring cerebral blood flow using magnetic resonance imaging techniques'. *J Cereb Blood Flow Metab* 19:701–735.

Jezzard P, Matthews PM, and Smith SM (eds) (2002) *Functional MRI: An Introduction to Methods*. Oxford: Oxford University Press.

Koh D-M and Collins DJ (2007) 'Diffusion-weighted MRI in the body: applications and challenges in oncology'. *Am J Roentgenol* 188:1622–1635.

Moonen CTW and Bandettini PA (eds) (2000) *Functional MRI*. Berlin: Springer-Verlag.

Schmitt R, Stehling MK and Turner R (2012) *Echo-Planar Imaging: Theory, Technique and Application*. Berlin: Springer-Verlag.

Takahara T, Imai Y, Yamashita T, Yasuda S, Nasu S and Van Cauteren M (2004) 'Diffusion weighted whole body imaging with background suppression (DWIBS): technical improvement using free breathing, STIR and high resolution 3D display'. *Radiat Med* 22:275–282.

Tofts PS (ed.) (2003) *Quantitative MRI of the brain: Measuring Changes Caused by Disease*. Chichester: John Wiley.

Tournier JD, Calamante F and Connelly A (2007) 'Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution'. *NeuroImage* 35:1459–1472.

Downloaded from https://www.cambridge.org/core. National Cheng Kung University, on 16 Sep 2019 at 02:12:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781107706958.019