Chapter

# 15 Go with the Flow: MR Angiography

## 15.1 Introduction

**M**agnetic **R**esonance **A**ngiography (MRA) uses a variety of methods to visualize the blood flow within arteries and veins. There are two main classes of MRA technique that can create angiographic-type images: those that exploit the intrinsic contrast created by moving spins and those that rely on the use of an injected contrast agent. The former endogenous contrast methods are now typically classified by the term '**N**on-**C**ontrast' **MRA** (NC-MRA) to separate them from techniques that require administration of an external (or exogenous) contrast agent, typically referred to as '**C**ontrast-**E**nhanced' **MRA** (CE-MRA). NC-MRA techniques use a variety of pulse sequences to create the desired contrast between blood moving within vessels, and the stationary background tissue.

In this chapter we will see that:

- flowing blood has different appearances on spin-echo and gradient-echo pulse sequences, compared to static tissues;
- NC-MRA sequences utilize in-flow, phase effects, or the inherent vessel pulsatility to produce angiographic images;
- CE-MRA sequences can be acquired as 3D and 4D, meaning that we can observe the temporal changes in blood flow;
- intracranial venography is possible using susceptibility-weighted imaging methods.

## 15.2 Effect of Flow in Conventional Imaging

The effect of moving spins on the magnetic resonance signal has been known since the earliest days of NMR and was rapidly exploited in the development of the first commercial imaging systems in the mid to late 1980s to produce angiographic images. The main effect of moving spins can be classified according to their influence on the longitudinal (time-of-flight) or transverse (phase shift) magnetization. In this chapter we will refer to these effects on blood, as shorthand for the effect on the actual moving spins.

### 15.2.1 Time-of-Flight Effects

The **T**ime-**O**f-**F**light (TOF) effect in MRI arises due to the blood flow between the RF pulses. In a spin-echo sequence all blood within the imaging slice of thickness $\Delta z$ will experience the 90° excitation pulse. During the time between the 90° excitation pulse and the 180° refocusing pulse (TE/2), the excited blood will have flowed, either partially or completely, out of the slice and fresh, i.e. unexcited, blood will flow into the slice. We know that only blood that experiences both a 90° and a 180° pulse will create a spin echo at the echo time. The spin-echo signal within the vessel will therefore depend upon the fraction of spins that experience the excitation pulse and are still present within the slice at the time of the 180° refocusing pulse. If the blood completely washes out of the slice between the two pulses then the lumen appears dark or hypointense. However, in the case of very slow flow, the wash-in of unsaturated spins with a large longitudinal magnetization between TRs, together with the lack of wash-out, may lead to a bright signal. Figure 15.1 shows the TOF effect in spin-echo imaging.

TOF effects in gradient-echo imaging generally appear as a signal hyperintensity due to the in-flow of fresh blood, that has not experienced any prior RF pulses, into the imaging slice between TRs (Figure 15.2). The degree of enhancement will depend upon the velocity ($v$), slice thickness ($\Delta z$) and TR. When $v$ is equal to or greater than $\Delta z$/TR, the blood is completely replaced with fresh, i.e. unsaturated, blood during TR, resulting in a maximum signal. When $v$ is less than $\Delta z$/TR there is only partial replacement of spins resulting in reduced signal

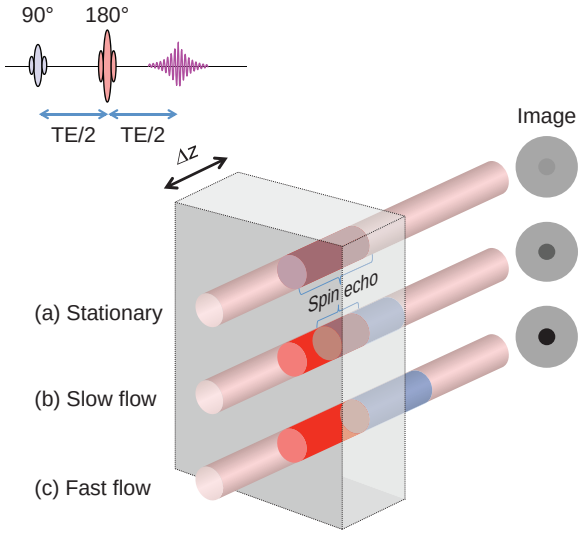


**Figure 15.1** Time-of-flight effect in spin-echo imaging. The 90° pulse (blue) excites all the spins in the slice of thickness Δz. (**a**) Stationary blood will experience both a 90° and a 180° (red) pulse yielding a spin-echo signal (purple). (**b**) Slow-flowing blood may not completely leave the slice between the 90° and 180° pulses, resulting in a spin-echo signal only from the blood experiencing both (purple). (**c**) Fast-flowing blood will only experience the 90° pulse and no spin echo will occur resulting in zero signal. This is known as wash-out.

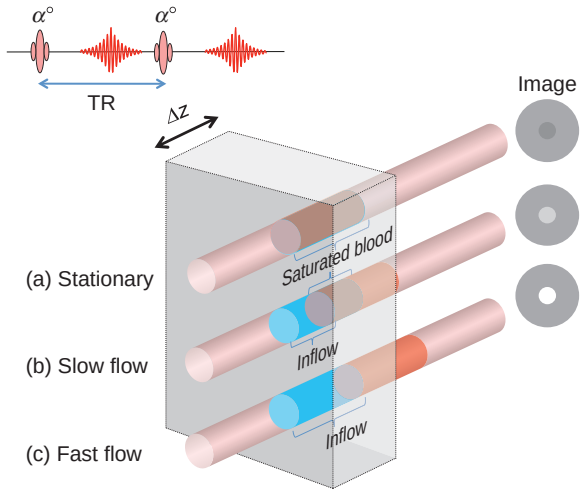


**Figure 15.2** Time-of-flight effect in gradient-echo imaging. The $\alpha°$ pulse (red) excites all the spins in the slice of thickness Δz. (**a**) Stationary blood will experience both $\alpha°$ pulses yielding a partially saturated signal (purple). (**b**) Slow-flowing blood may not completely leave the slice during TR and only a proportion of the blood will experience both pulses (purple). However, unsaturated blood (blue) that has not experienced any prior RF pulses will enter the slice during TR. The signal will be a combination of the unsaturated (blue) and saturated (purple) blood. (**c**) Fast-flowing blood will be completely replaced during the TR period, resulting in maximal signal in the second echo.

intensity. Figure 15.2 shows the TOF effect in gradient-echo imaging.

## 15.2.2 Flow Artefacts

As you should hopefully know by now, spins in the presence of a magnetic field gradient will accumulate phase. Whereas stationary spins accumulate phase in proportion to time ($T$), moving spins accumulate phase in proportion to time-squared ($T^2$) (see Box 'Velocity-Induced Phase Shift'). Since different phase-encoding steps will occur at different points in the cardiac cycle, unless the sequence is ECG gated, the temporal changes in flow velocity will result in temporal changes in phase. Following Fourier transformation these phase modulations will result in discrete 'sidebands' or ghosts in the images along the phase-encoding direction.

One way to reduce these artefacts for bright blood imaging is to eliminate the phase shift for moving spins. We are already used to the idea of eliminating stationary tissue phase shifts in a pulse sequence by using an equal area, but opposite sign (polarity), gradient pulse. For example, the slice-selection gradient is followed by a negative rephasing gradient to compensate for the phase shift induced across the slice thickness by the slice-select gradient. For stationary spins the phase shift is given by the product of the gradient amplitude $G$ and its duration $T$, known as the zeroth-order moment $M_0 = G{\cdot}T$. However, for spins moving with a constant velocity the phase shift is proportional to the first moment $M_1 = G{\cdot}T^2$. In this case a three-lobed gradient can be used to eliminate the phase shift for both stationary spins and those moving with a constant velocity. The three-lobed gradient can be considered as the sum of two back-to-back bipolar gradients with opposite polarities where the second bipolar pair compensates for the velocity-induced phase shift caused by the first pair.

Since these gradient pulses compensate for both the zeroth (stationary) and first-order (velocity) induced phase shifts their application in a pulse sequence is often referred to as **Gradient Moment Nulling** (GMN). Alternative names include velocity compensation or flow compensation (although it is strictly velocity, not flow, that is compensated). Figure 15.3 shows zeroth and first-order gradient

moment nulled waveforms and the effect of first-order nulling on pulsatile flow ghosting. Phase shifts induced by higher orders of motion, e.g. acceleration (a second-order effect), would require a four-lobed gradient waveform to compensate for zeroth, first- and second-order motion. The first three orders of GMN can be represented by binomial gradient amplitudes, e.g., 1:–1, 1:–2:1, and 1:–3:3:–1. The time required to play out these GMN waveforms invariably increases the minimum TE. Often, as a compromise, only first-order motion is compensated. Alternatively the use of pulse sequences with very short TEs reduces the time available for dephasing and may be preferable to the use of a longer TE associated with GMN. Figure 15.3 also shows a 2D gradient-echo pulse sequence with and without velocity compensation, and the effect of velocity compensation on reducing flow ghosting.

If all the blood within a vessel is moving at the same velocity (known as 'plug flow') then the phase shift across the vessel will be constant. However, usually there is a range of blood velocities across a vessel. In long, straight vessels the flow profile is parabolic, where the velocity is almost zero at the vessel walls and maximum in the centre. This is often called 'laminar flow' since adjacent layers of fluid flow past each other without mixing. Since a blood vessel is typically only a few voxels across, each voxel will contain a range of velocities, particularly those near the vessel wall where the range of velocities can be

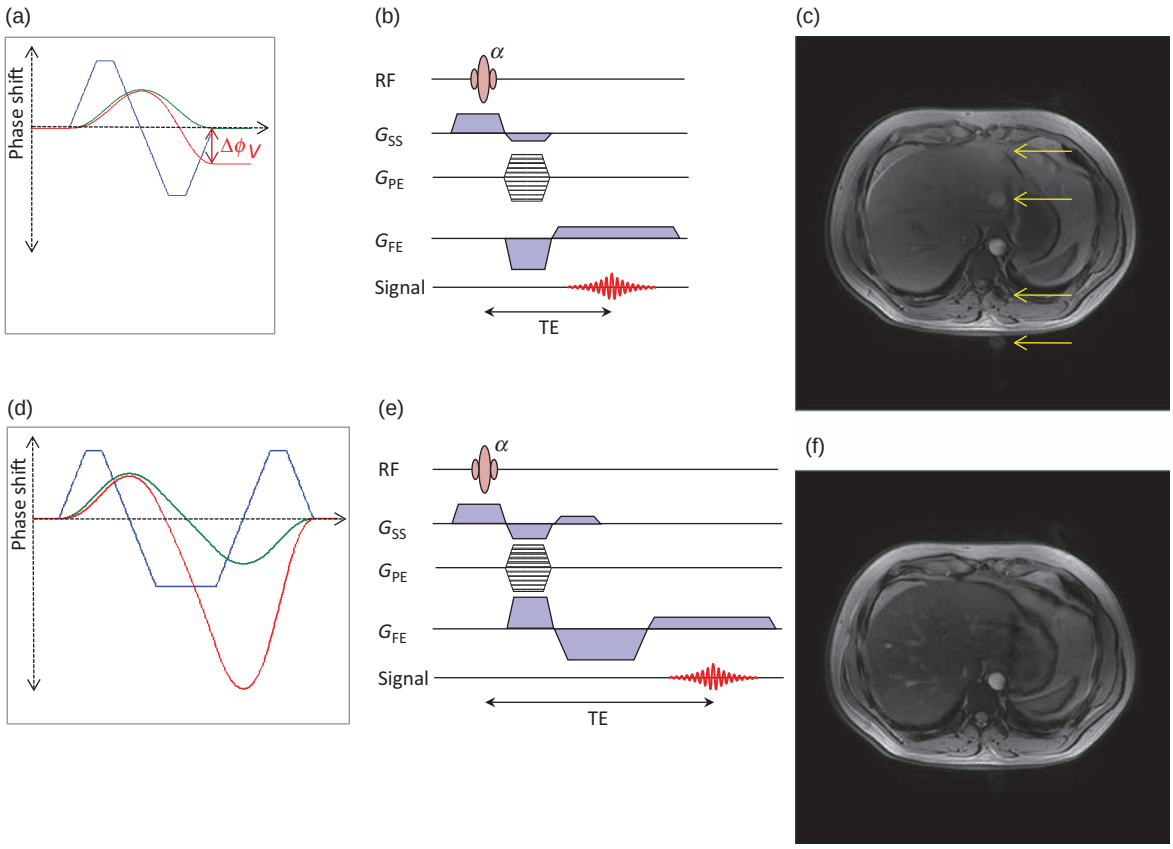


**Figure 15.3** Gradient moment nulling. (a) Zeroth-order ($M_0$) nulling uses a bipolar gradient (blue) with equal areas but different signs (1:–1). There is zero phase shift for stationary spins (green) but a net phase shift $\Delta\phi$ for spins moving with a constant velocity (red). (b) A standard 2D gradient-echo pulse sequence without velocity compensation results in (c) ghosts of the aorta (arrows) due to the phase shifts induced by the pulsatile blood flow. (d) First-order ($M_1$) nulling uses a three-lobed gradient with areas in the ratio 1:–2:1. There is now zero phase shift for both stationary spins and those moving with constant velocity. (e) A 2D gradient-echo sequence with velocity compensated $G_{SS}$ and $G_{FE}$ results in (f) an image where the ghosting is effectively eliminated by the velocity compensation. Note that $G_{PE}$ is not usually velocity-compensated since the amplitude and hence the phase shift is very small when sampling the centre of k-space.

quite large. Individual voxels will therefore contain a range of velocities and hence phase shifts. If the range of intravoxel phase shifts is large enough then signal cancellation can occur and the edges of vessels may appear with reduced signal and the apparent width of the vessel may appear smaller. This effect may also be seen distal to pathologies such as a vascular stenosis, where turbulent flow can potentially overestimate the degree of stenosis.

GMN can also be used in combination with $T_2$w TSE imaging. Flow in the cerebrospinal fluid can often cause signal dephasing in the spinal cord. The use of velocity compensation gradients can reduce the appearance of these artefacts.

GMN is an effective method for reducing ghosting artefacts and ensuring that the vascular signal is bright. If bright flow is not a requirement then ghosting artefacts can be most effectively dealt with by eliminating the signal from flowing blood before it enters the imaging slice, using adjacent spatial pre-saturation bands (see Section 7.2.5). Spins within the pre-saturation slabs are flipped into the transverse plane and have their magnetization dephased by spoiler gradients. These saturated spins then flow into the imaging slice, where they have zero signal. A consistently low intra-lumenal signal means that artefacts are not produced. The disadvantage is that the TR must be slightly increased because of the additional RF and gradient pulses required for the pre-saturation bands.

**Velocity-Induced Phase Shift**

Nuclei precess at the Larmor frequency

$$\omega = \gamma B_0$$

Phase ($\phi$) is the time integral of frequency ($\omega$), i.e.

$$\phi = \int \omega \, dt$$

In a magnetic field gradient along the $x$ axis, for example, $G_x$, this phase shift becomes

$$\phi = \gamma \int (B_0 + G_x \cdot x) \, dt$$

For spins moving at a constant velocity $v$ along the $x$ direction, the phase shift is

$$\phi = \gamma \int (B_0 + G_x \cdot (x + vt)) \, dt$$

If we consider a single gradient pulse of amplitude $G$ and duration $T$ then the phase shift due to velocity alone will be

$$\phi = \gamma \int_0^T vGt \, dt = \left[\frac{1}{2}\gamma vGt^2\right]_0^T = \frac{1}{2}\gamma vGT^2$$

The product $GT^2$ is usually called the 'first moment' of the gradient, $M_1$, so we can write

$$\phi = \frac{1}{2}\gamma vM_1$$

If a second gradient of the same duration ($T$) but opposite amplitude ($-G$) immediately follows the first then we have a bipolar gradient pulse with a total phase shift of

$$\phi = -\gamma vM_1$$

If we repeat the acquisition with the polarity of the bipolar pulse reversed and subtract the phases, then we have the phase difference $\Delta\phi$, which can be expressed as the difference in the two moments of the bipolar pulses $\Delta M_1$

$$\Delta\phi = \gamma v \Delta M_1$$

The velocity encoding or '*venc*' parameter is defined as the velocity that produces a phase shift of $\pi$ radians or 180°

$$\pi = \gamma \cdot venc \cdot \Delta M_1$$

therefore,

$$venc = \frac{\pi}{\gamma \cdot \Delta M_1}$$

## 15.2.3 The MR Angio Pulse Sequence Jungle

Since angiography is a very popular application for MR, there is a whole new set of acronyms available from the different manufacturers. MR angiography sequences can be divided into three main types: non-contrast (NC) MRA, contrast-enhanced (CE) MRA and susceptibility MRA. There are four types of NC-MRA: the very common time-of-flight effect, phase-contrast scans, sequences based on 3D TSE and finally those based on fully rewound gradient echo. CE-MRA techniques use a bolus injection of gadolinium contrast agent, and can be classified as either time-resolved (also known as 4D) or fluoro-triggered.

Table 15.1 is a guide to the different names used by the main manufacturers for MRA sequences.

**Table 15.1** Acronyms for MR angio sequences; see Glossary for more details

| Generic name | GE Healthcare | Hitachi | Philips | Siemens | Toshiba |
|---|---|---|---|---|---|
| NC-MRA using time-of-flight effect | TOF, MOTSA | TOF | TOF, Multi-chunk MRA | TOF | TOF |
| NC-MRA using phase-contrast effect | PC | PC | PC, Q-Flow | PC | PS |
| NC-MRA using 3D TSE | Inhance 3D Deltaflow | VASC FSE | TRANCE | NATIVE-SPACE | FBI, CIA |
| NC-MRA using fully rewound GE | IFIR | VASC ASL | b-TRANCE | NATIVE-TrueFISP | Time-SLIP |
| Time-resolved (4D) CE MRA | TRICKS | TRAQ | 4D-TRAK | TWIST | LDRKS |
| Fluoro-triggered CE MRA | Smart-Prep, Fluoro Triggered | FLUTE | BolusTRAK | CARE Bolus | Visual Prep |
| Susceptibility-weighted imaging | SWAN | BSI | SWIp | SWI | FSBB |

## 15.3 Non-Contrast MR Angiography

NC-MRA refers to techniques used to create angiograms without the use of an exogenous contrast agent, e.g. gadolinium. The original NC-MRA techniques are based on TOF and PC. However, in recent years, other techniques have been developed, partly to address concerns associated with **N**ephrogenic **S**ystemic **F**ibrosis (NSF) (see Section 20.7) and partly to address imaging in areas for which TOF and PC are less successful, such as the abdomen and peripheral vasculature.

### 15.3.1 Time-of-Flight Angiography

Time-of-flight MRA was developed in the mid to late 1980s and utilizes the fact that flowing blood appears bright on gradient-echo imaging due to the TOF effect discussed above, particularly if steps are taken to reduce the signal from the stationary background tissue. Typically, first-order GMN is used to reduce signal dephasing. Images are acquired either as multiple 2D slices or as 3D volume acquisitions, with 3D angiographic images produced by stacking the slices and performing a **M**aximum **I**ntensity **P**rojection (MIP) algorithm in which parallel rays are cast through the 3D volume and a 2D projection image is created using the maximum pixel intensity traversed by the ray, as shown in Figure 15.4.

In a 2D TOF technique multiple sequential thin (approximately 1.5 mm) slice gradient-echo images are acquired perpendicular to the direction of flow. Usually a relatively large flip angle, e.g. 60°, is used to saturate the stationary tissue, resulting in good vascular contrast. Figure 15.5 shows example images from a 2D TOF study of the extracranial circulation.

**Applications of 2D TOF**

2D TOF is often used for fast vascular localizers, particularly in the carotid arteries where the slices are perpendicular to the vessel. Superior saturation bands are applied to provide directional flow selectivity. Since the slice thickness is relatively thick compared to the in-plane pixel size, the reformatted images can show a 'stair-casing' artefact following MIP. The relatively long TE associated with exciting a thin slice and the use of GMN also means that complex flow dephasing secondary to a stenosis can result in an apparent overestimation of the length and diameter of the stenosis, as shown in Figure 15.6.

In 3D TOF technique a relatively thick slab of tissue is excited and then sub-divided into thin slices or partitions by a second phase-encoding process along the slice-selection direction, followed by a 3D

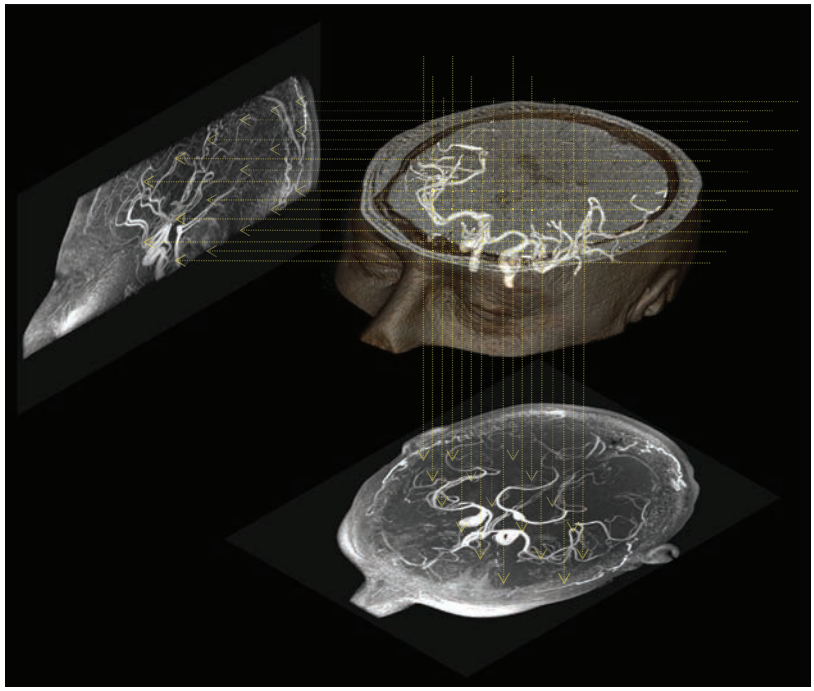

**Figure 15.4** Maximum intensity projection (MIP) algorithm. Parallel rays (dotted lines) are projected through the data volume. The maximum pixel brightness is then displayed on the projection images. This has the effect of collapsing all the data along the ray into a single image. The data volume may be rotated a few degrees and the MIP repeated. If the projection images are then displayed in a movie loop, they provide all the necessary visual cues to allow '3D' perception of the data.

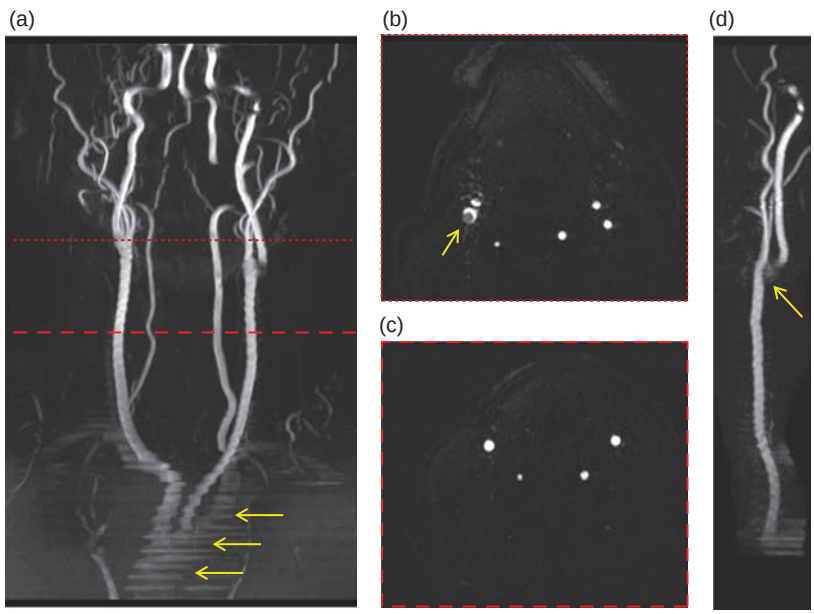


**Figure 15.5** 2D time-of-flight (TOF) MRA. (**a**) Coronal MIP from a stack of 240 slices through the extracranial circulation. The alternating signal intensity around the aortic arch (arrows) is due to pulsatile flow. (**b**) An axial slice at the location of the dashed red line, through the right carotid bifurcation. Note that complex flow around the carotid bulb causes signal loss (arrow). (**c**) A single 2 mm slice at the location of the dashed red line. The use of a superior sat band means that only the blood flowing inferior to superior, i.e. in the common carotid and vertebral arteries, is visible. (**d**) A MIP around just the left carotid bifurcation. Note the loss of signal (arrow) where the blood flows in-plane.

FT to reconstruct the images. The advantage of 3D encoding is that much thinner slices can be reconstructed (typically as little as 0.7 mm), which improves the resolution and hence vessel conspicuity by reducing the partial volume effects. Since $\Delta z$ is now the slab thickness, from our equation $v = \Delta z/\mathrm{TR}$, the threshold velocity is quite high. Therefore 3D TOF methods are best applied intracranially where there is very fast-flowing blood. Since venous blood in the intracranial circulation flows much slower, 3D

TOF methods cause an intrinsic saturation of slow flow without the need for saturation bands. Figure 15.7 shows example images from a 3D TOF study of the intracranial circulation.

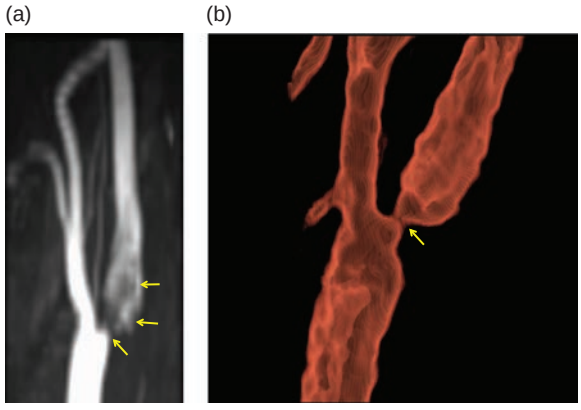


**Figure 15.6** 2D TOF overestimation of stenosis. (a) MIPed 2D TOF in a patient with a short 70% stenosis of the internal carotid artery. Note the signal loss and hence overestimation of the stenosis (arrows). (b) The true extent of the stenosis is shown in this rendering from a 3D contrast-enhanced MRA with a very short TE that minimizes the signal dephasing.

### Applications of 3D TOF

In 3D TOF we try to avoid saturating the signal from spins as they flow deeper into the volume, hence we use a lower flip angle compared to 2D TOF, e.g. 30°. 3D acquisitions also means that the SNR in individual images is also significantly improved compared to 2D methods. However, this improvement in SNR also applies to the background tissue which limits the contrast and visibility of some small distal vessels. Magnetization transfer techniques can also be used to reduce background tissue signal since blood does not demonstrate a significant MT effect, but brain parenchymal signal is reduced. A low flip angle may not be enough to avoid spin saturation with depth so special RF pulses known as ramped excitation or **T**ilted **O**ptimized **N**on-saturating **E**xcitation (TONE) can be used. These pulses effectively increase the flip angle with distance into the volume. Alternatively multiple 3D slabs can be used; these should be slightly overlapped to eliminate the artefacts that can occur at the edges of 3D slabs, as shown in Figure 15.8.

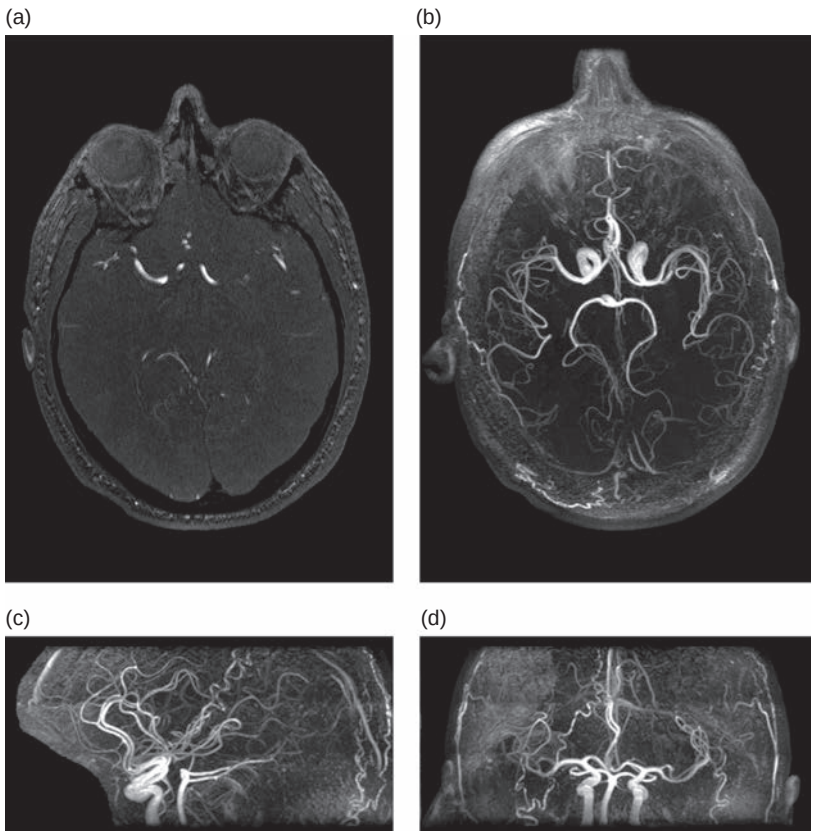


**Figure 15.7** 3D TOF MRA. (a) A single 1 mm slice from a 3D TOF acquisition through the main intracranial vessels. (b) The top-down MIP from all 172 slices, (c) sagittal projection and (d) coronal projection. Note the relatively high signal in the stationary background tissue compared to 2D TOF.

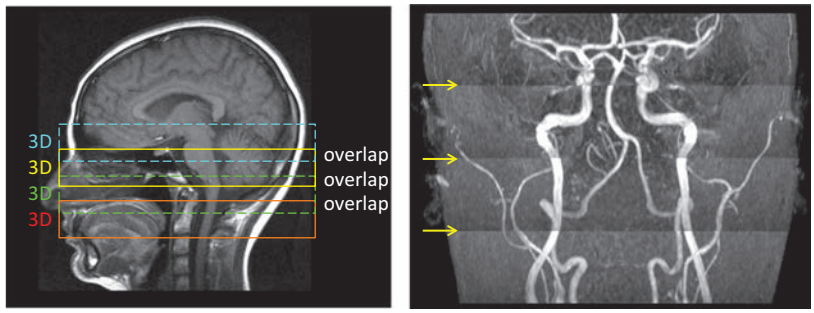


**Figure 15.8** Multiple Overlapping Thin Slab Acquisition (MOTSA). Multiple thin 3D slabs are acquired to try to reduce the progressive signal saturation associated with thick 3D slabs. The slabs are overlapped to avoid artefacts at the slab boundary. Even so there is often some residual artefact at the slab boundaries (arrows), known as Venetian blind artefact.

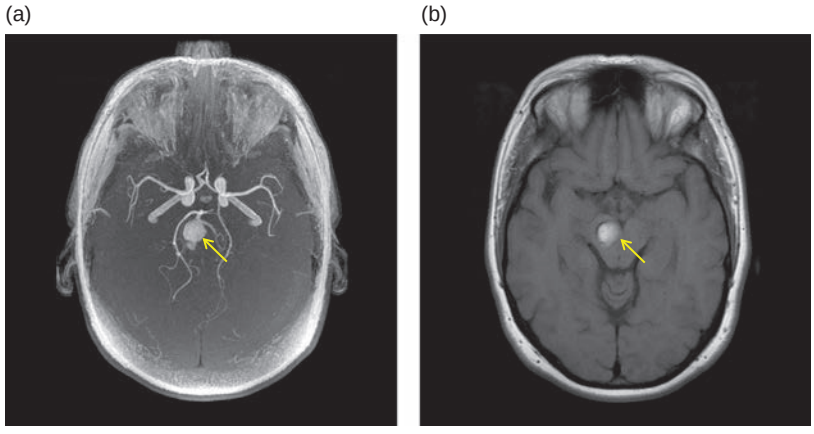


**Figure 15.9** Cavernous angioma. (a) MIP from a 3D TOF study showing a cavernous angioma (arrow) that appears hyperintense. This could be mistaken for flowing blood within an aneurysm. (b) $T_1$w spin echo where the angioma also appears hyperintense, indicating the presence of short $T_1$ blood breakdown products.

Since TOF MRA exploits the changes in the longitudinal magnetization of flowing blood to create contrast, any tissue with a short $T_1$ can also appear bright in a TOF angiogram. This can be particularly problematic in certain pathologies that result in tissues with very short $T_1$s, such as methaemoglobin in haemorrhage. These tissues appear hyperintense in TOF and may be mistaken for flowing blood. Figure 15.9 shows a flow-mimicking artefact in a patient with a cavernous angioma.

## 15.3.2 Phase Contrast Angiography

**P**hase **C**ontrast (PC) MRA exploits the changes in the phase of blood's transverse magnetization as it moves along a magnetic field gradient. We have already seen how a bipolar imaging gradient gives zero phase shift for stationary spins, but a non-zero phase shift for moving spins. In a PC pulse sequence additional bipolar velocity-encoding gradients are applied along each of the three gradient directions to create a linear relationship between the velocity of the blood and the phase of the MR signal. This relationship is scaled by setting a user-controlled velocity encoding value (*venc*). Since we have 360° of unique phase available, flow in one direction, relative to the gradient, is allocated 0° to +180°, while flow in the opposite direction is allocated 0° to −180°. The *venc* is the maximum velocity, along each direction, that will result in a 180° phase shift. Since the MRI signal is acquired in quadrature, it is possible to create images of this phase shift and hence velocity.

Gradient-echo sequences, due to their short TR, are most commonly used for PC-MRA acquisitions. However, variations in the static magnetic field uniformity produce phase shifts in the background stationary tissue. To eliminate these phase shifts, two acquisitions are performed along each gradient direction, with the bipolar velocity-encoding gradients reversed in polarity for the second acquisition. The phase images for both acquisitions are then calculated and subtracted, cancelling the stationary background phase and leaving only positive and negative phase shifts depending upon the direction of blood flow. Finally, to suppress background pixels, e.g. air, where the phase is random, the phase subtraction image is

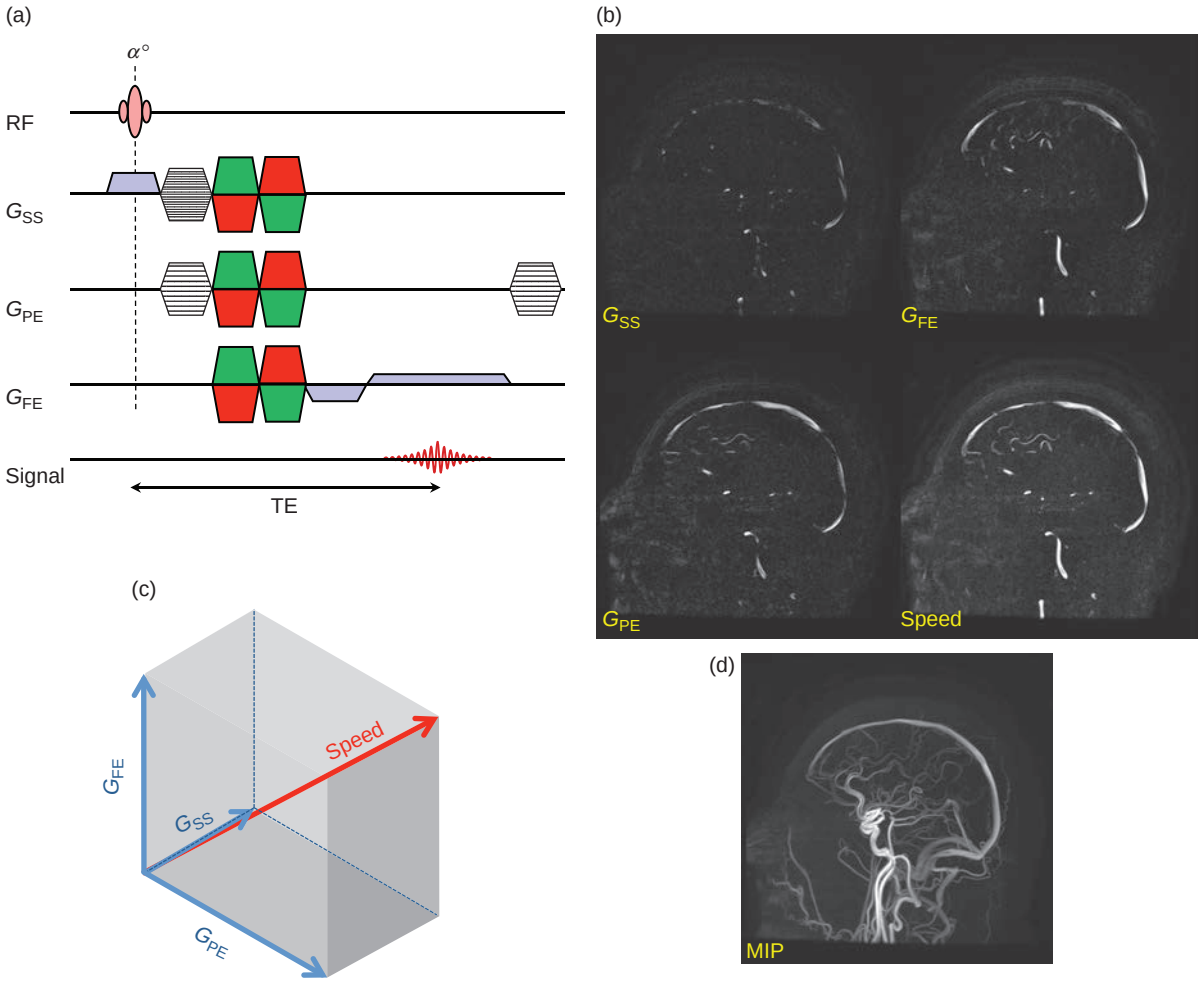


**Figure 15.10** Principles of 3D phase contrast (PC) MRA. (a) Bipolar flow encoding gradients (green and red) are applied separately in subsequent TRs and then along each direction. The phase images for each encoding along a particular direction are then subtracted. (b) Images from a mid-line sagittal slice from a 3D PC acquisition. The three velocity components are shown as well as the combination as a speed image. (c) The three velocity vectors are combined into a single speed value. (d) MIP of the full 3D speed images. Since the phase shift is proportional to the velocity the faster-flowing blood appears brighter.

multiplied, pixel-by-pixel, with the conventional magnitude image.

PC MRA is directionally sensitive, since only blood moving in the same direction as the bipolar flow-encoding gradient results in a phase shift. Blood vessels follow a fairly tortuous path throughout the body so it is generally necessary to encode along all three gradient directions. This means that a pair of alternate polarity velocity-sensitizing gradients must be applied along each gradient axis in turn, as shown in Figure 15.10, i.e. six acquisitions are required. However, in practice this can be reduced to four: one velocity sensitive acquisition along each direction followed by one acquisition with zero velocity sensitization (first-order GMN) along all axes. This requires subtracting each of the velocity-sensitive phase images from the same reference image, which is suboptimal from an SNR perspective. In practice more efficient methods, such as Hadamard encoding schemes that sensitize in two directions simultaneously, address this issue. Even so, a PC MRA study with velocity sensitization in all three directions will take at least four times as long as an equivalent TOF study.

The individual phase images are usually combined to produce a single 3D angiogram by calculating the

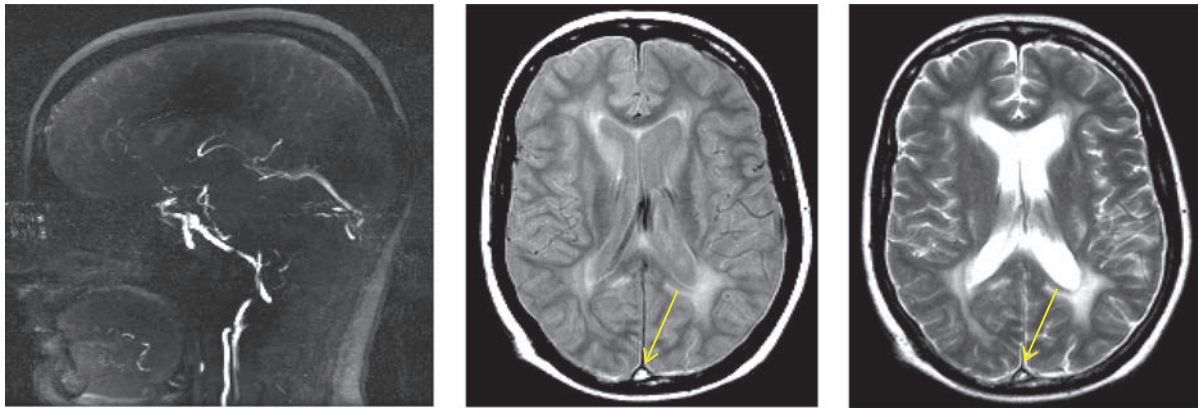


**Figure 15.11** Superior sagittal sinus thrombosis. MIP from a 3D PC MRA acquisition. Note the absence of signal due to sinus thrombosis. High signal, i.e. no flow wash-out, is also seen in the sagittal sinus (arrow) on PDw and $T_2$w images.

resultant flow magnitude image $|v|$ from the $x$, $y$ and $z$, i.e. slice-select, phase-encoding and frequency-encoding directions for each pixel using

$$|v| = \sqrt{v_x^2 + v_y^2 + v_z^2}$$

The resultant magnitude image has no directional information and is often termed a 'speed' image. Also, since each velocity value is squared in the calculation, all the positive and negative velocity information is eliminated and so any directional information is lost, although the greyscale is still proportional to the speed.

**Applications of 3D PC**

In 3D PC MRA you excite a slab of tissue, with each slice encoding having velocity sensitization applied along each of the required directions, usually all three. This makes 3D PC studies quite time-consuming and you may have to sacrifice some resolution in the phase-encoding direction or employ parallel imaging techniques to achieve an acceptable acquisition time. Like TOF MRA the speed images from each 3D slice are processed by the MIP algorithm to produce the standard angiographic display.

With PC MRA you have to set a *venc*. If you know the velocities in the vessels of interest then you should set your *venc* to be about half the peak velocity. It is possible to use a 2D PC MRA method whereby you prescribe just a single slice which is thick enough to cover the vessels of interest. This produces a projection angiogram of the blood flow within that thick slice. Since the method is quick, i.e. only a single slice, you can acquire images with different *venc*s to find the most suitable value for your time-consuming 3D PC acquisition. These 2D projection images are sometimes called *venc* localizers. Finally, since only moving spins give signals on PC MRA images, they do not have the same problem with short-$T_1$ blood-breakdown products as TOF MRA images. PC MRA is therefore the technique of choice in conditions such as sinus thrombosis, as shown in Figure 15.11

## 15.3.3 3D ECG-Triggered TSE-Based NC-MRA

Both TOF and PC MRA methods do not work particularly well outside the head and neck. The use of Gd-based contrast-enhanced MRA has been the method of choice for body and peripheral MRA for a number of years. However, concerns over the safety of Gd-based contrast agents, particularly related to NSF, have sparked an interest in the development of NC-MRA methods in the rest of the body. ECG-triggered 3D TSE methods, originally called **F**resh **B**lood **I**maging (FBI) by Toshiba, have been used to obtain large FOV angiograms in a variety of regions, although most commonly in the periphery. The method uses two interleaved 3D TSE readouts. The first TSE echo train is gated to systole and the second to diastole. In the systolic images, fast arterial flow dephases the MRI signal and blood appears dark. In the diastolic image the arterial flow is slower, hence the dephasing is less and so blood appears with a high signal. Venous flow is

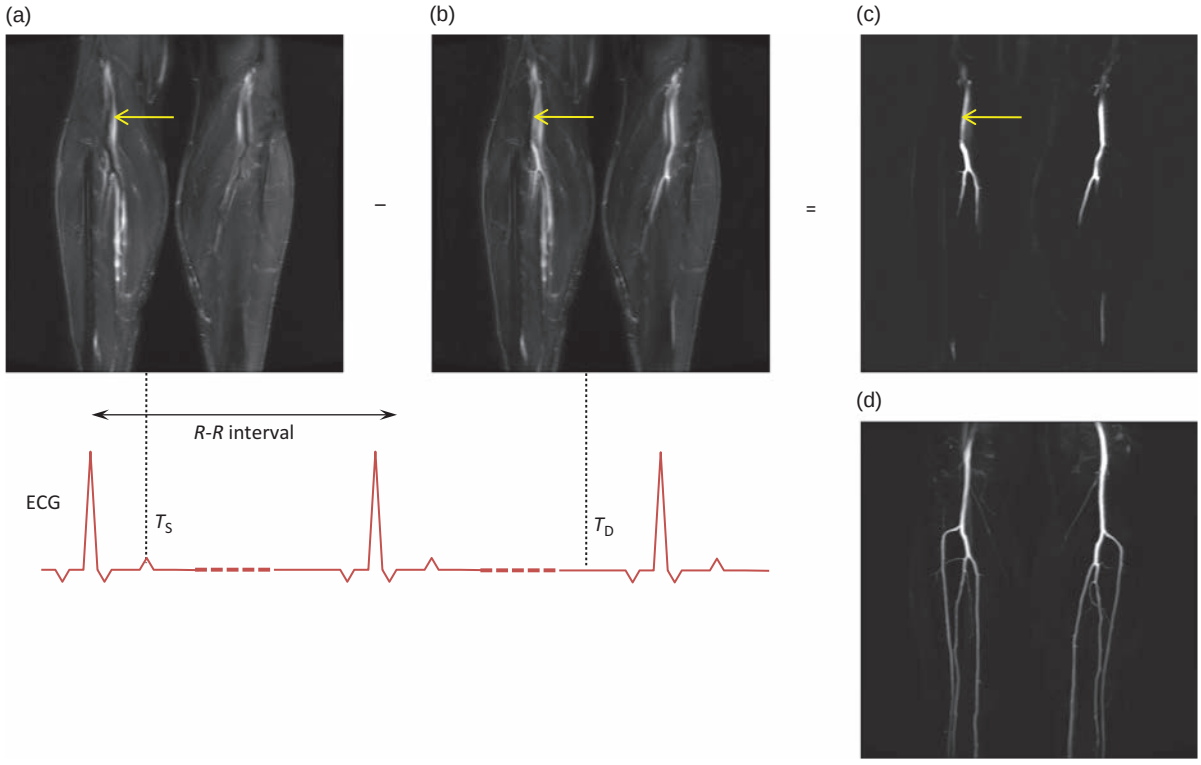


**Figure 15.12** Fresh blood imaging. This method uses two 3D multi-shot TSE acquisitions. (**a**) The first echo train is timed to systole, where fast-moving systolic blood is dephased (arrow). (**b**) The second echo train is acquired 1 to 3 *R–R* intervals later and is timed to diastole, where blood velocities are reduced and there is less dephasing. (**c**) Since venous blood and stationary tissue are identical in both acquisitions, subtraction of the images results in only signal from the arteries. (**d**) Coronal MIP of subtracted data.

relatively slow and constant throughout the cardiac cycle and appears bright in both acquisitions. Likewise, stationary background tissue appears similar in both acquisitions. A simple magnitude subtraction of the systolic images from the diastolic images results in suppression of the background and venous signals, leaving only the arterial signal as shown in Figure 15.12. An additional STIR preparation helps to reduce the background signal from fat. **N**on-contrast MR of **Ar**T**er**I**e**s and **VE**ins (NATIVE) by Siemens, **TR**iggered **A**ngiography **N**on-**C**ontrast **E**nhanced (TRANCE) by Philips and Inhance 3D Deltaflow by GE Healthcare are similar techniques.

Limitations of these techniques include spatial misregistration between the two acquisitions, as only part of the 3D acquisition can be performed in each heartbeat, and image blurring due to $T_2$ relaxation during the TSE readout. This can be reduced by the use of parallel imaging to shorten the echo train. It is also important to identify the optimal systolic and diastolic trigger delays required for each body area. An 'ECG-Prep' scan such as a single-slice, multiphase, single-shot TSE (SS-TSE) with progressive ECG trigger delays can be used to visually identify the optimal systolic (darkest intravascular signal) and diastolic (brightest intravascular signal) trigger delays.

The scan time for these methods is approximately four minutes for a single 3D location. Multiple vascular territories can be covered by moving the patient and repeating the acquisition. Given the differences in flow profiles, it is often necessary to repeat the ECG-Prep in order to optimize the timings for each location. Since this technique primarily relies on having a good differential flow between systole and diastole, its primary application is in peripheral angiography.

## 15.3.4 Balanced Steady-State-Free-Precession

**B**alanced **S**teady-**S**tate-**F**ree-**P**recession (bSSFP) is a term widely used in the MR angiography literature as a generic term for fully rewound gradient-echo sequences used in MRA. In this chapter, instead of

using our normal term (fully rewound gradient echo) we will use the abbreviation bSSFP. This sequence has an inherently high signal from blood which is relatively independent of flow. ECG and/or respiratory-triggered 3D bSSFP has been used in a number of body areas including imaging the coronaries and thoracic aorta. In bSSFP both arteries and veins appear bright as well as background tissue. An inversion pulse applied prior to the bSSFP readout can null the background signal. Commercial implementations include NATIVE-TrueFISP by Siemens, b-TRANCE by Philips, **I**nhance in**F**low **I**nversion **R**ecovery (IFIR) by GE Healthcare and **time**-**S**patial **L**abelling **I**nversion **P**ulse (time-SLIP) by Toshiba.

### Applications of NC-MRA: Renal Arteries

Balanced SSFP methods are used for renal MRA, particularly in patients with impaired renal function in whom Gd-based agents are contraindicated. In this application, the 180° inversion pulse is used to invert the signal from a large region encompassing the kidneys, inferior vena cava and other tissues. Following the inversion pulse the longitudinal magnetization of the tissue and the blood in the inferior vena cava will recover due to $T_1$ relaxation and, after a TI of around 1 s, will become zero. In addition, during this TI period arterial blood that originates outside of the inverted region will flow into the imaging region. A spectral inversion pulse is applied

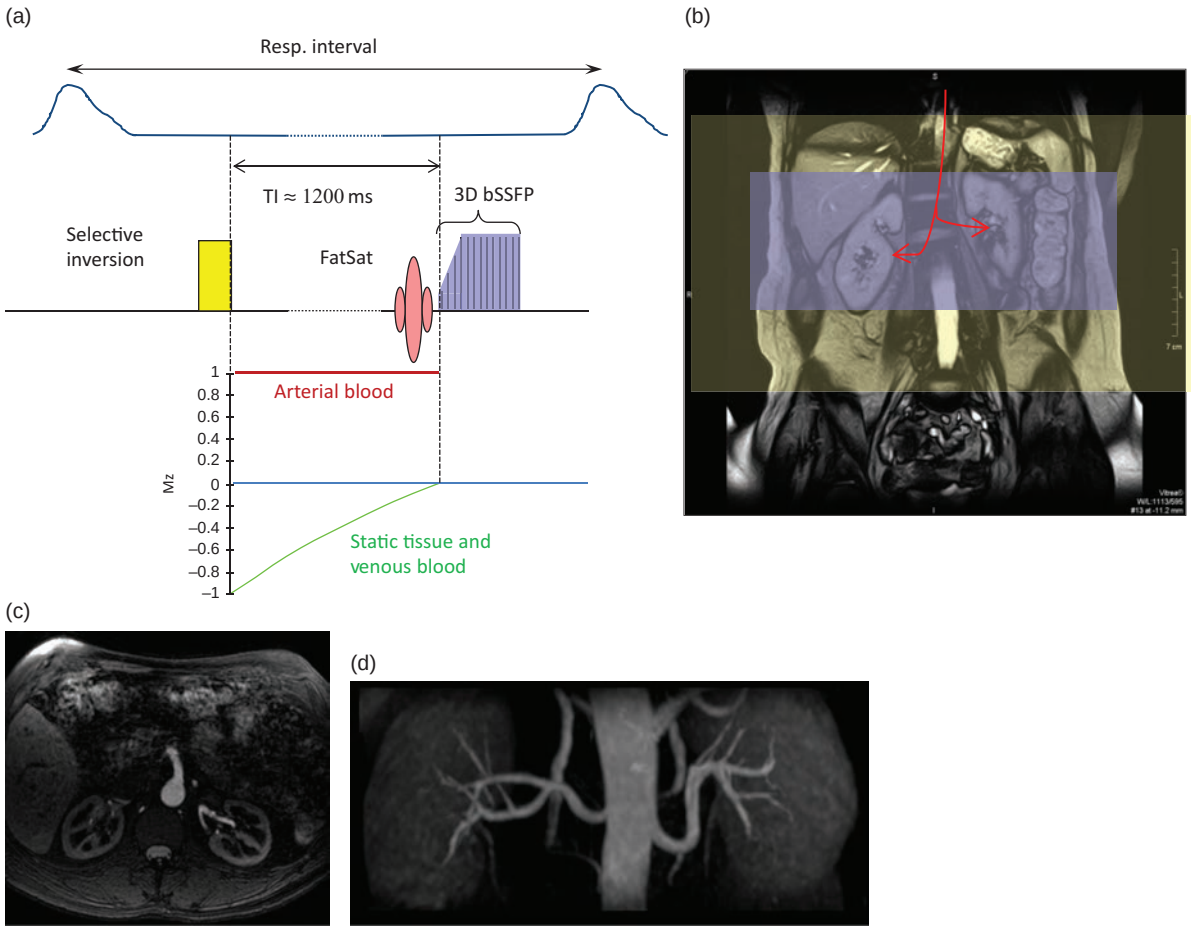


**Figure 15.13** NC-MRA of the renal arteries. (**a**) The pulse sequence has a thick slab selective inversion pulse (yellow) which inverts a region covering both the kidneys and the inferior vena cava. (**b**) During the inversion time (TI) non-inverted blood from the heart (red arrows) flows into the kidneys. At the same time the inverted magnetization recovers due to $T_1$. After the TI period the static tissue and venous blood should be approximately zero. A segmented 3D bSSFP readout, which yields a high signal from blood, is then performed. The 3D bSSFP slab (blue) is positioned directly over the kidneys and renal arteries. A spectral inversion recovery pulse is applied just prior to the bSSFP readout to further suppress the signal from fat. (**c**) A single axial slice from the 3D bSSFP slab. Note the high signal from the blood. (**d**) Coronal MIP demonstrating the renal arteries.

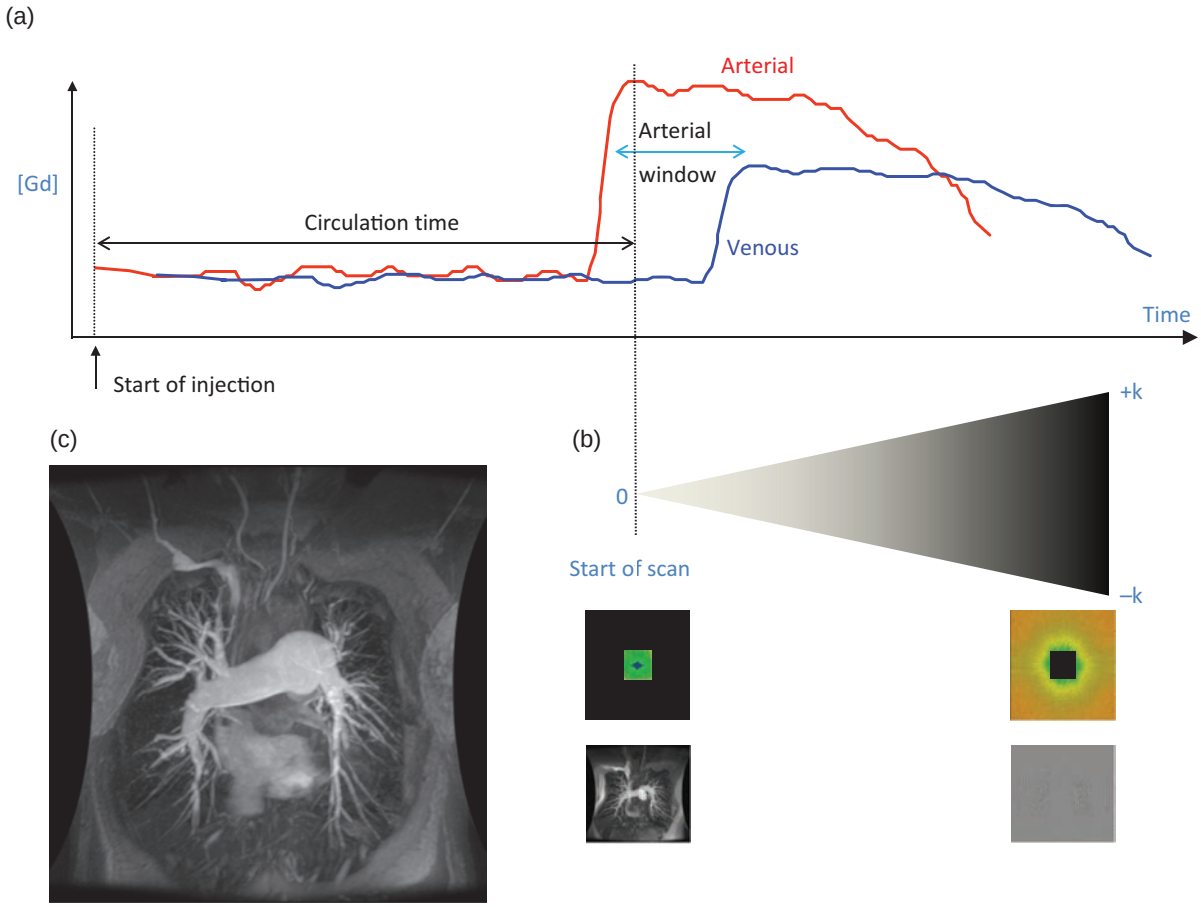


**Figure 15.14** Basic principles of contrast-enhanced MRA. (a) Following injection of the contrast agent there will be a delay, known as the circulation time, before the agent reaches the area of interest. There will then be a short period of time, the arterial window, when the contrast is mainly in the arterial circulation and has not reached the venous circulation. (b) Since this arterial window is generally shorter than the full image acquisition, it is usual to start data acquisition from the centre of k-space outwards. This ensures that most of the image contrast is acquired during the arterial window and the periphery of k-space, i.e., the edge information, is acquired later when venous contamination is less of an issue. (c) With the correct timing good-quality arterial-phase angiograms can be obtained, often in a breath-hold.

just prior to the bSSFP readout to reduce the high signal from fat. The bSSFP readout slab is positioned just covering the kidneys and renal arteries. The whole sequence is triggered by the respiratory waveform, with part of the 3D bSSFP acquisition acquired in each respiratory cycle, to minimize misregistration artefact. The overall scan time is around two to four minutes depending upon patient cooperation and the desired spatial resolution. Since the venous and background signals are naturally suppressed, there is no requirement for subtraction of data. The inversion and imaging regions are usually independently positioned to allow targeting of the desired vessels. As with other techniques, the final projection angiogram is produced using a MIP. Figure 15.13 shows the principle of the method for renal NC-MRA.

## 15.4 Contrast-Enhanced MRA

In CE-MRA a Gd-based contrast agent, injected as a rapid bolus, is used to produce an angiogram. We can think of this as a variant of the 3D TOF techniques since we are exploiting differences in the longitudinal magnetization to yield vascular contrast. But, unlike TOF methods, we are using Gd to shorten blood's $T_1$ relaxation time, rather than relying on the TOF effect, to cause enhancement of the blood signal. CE-MRA is usually used to acquire large field-of-view 3D angiograms in the coronal or sagittal plane without suffering from spin saturation. A fast 3D $T_1$w gradient-echo sequence, with very short TE and TR, is used to capture the first-pass transit of the contrast agent bolus through the anatomy of interest. The TE is kept as

short as possible by not using flow-compensation and by employing fractional echo data collection.

## 15.4.1 Timing Matters

The basic premise of CE-MRA is that a bolus administration of a Gd-based contrast agent will transiently reduce the $T_1$ of blood during its first pass through the anatomy of interest (see Box 'Reducing the $T_1$ of Blood'). This will have the effect of substantially increasing the vascular signal. Image acquisition has to be timed to coincide with when the contrast agent concentration is maximal in the area of interest and before the contrast agent reaches the venous circulation and the images become contaminated by the venous return. Because the acquisition time will typically be much longer than this arterial-only time window, we arrange the acquisition so that the centre of k-space, which primarily contributes to the contrast in an image, is acquired when the arterial contrast agent concentration is at its peak. This typically involves acquiring the centre of k-space first. Figure 15.14 shows the relationship between circulation time, arterial contrast agent concentration and data acquisition for a centrically encoded 3D acquisition. Acquiring data too early with respect to the injection can result in poor arterial enhancement and edge enhancement artefacts, while acquiring data too late not only gives poor arterial enhancement but also increases unwanted venous signal. Various methods have been proposed over the years to optimize the timing, ranging from the administration of a small separate test bolus, to administration of the full bolus with triggering performed either automatically using a 1D 'tracker' region placed inside a vessel or using a 'real-time' fluoroscopic triggering system. Some manufacturers offer a 'real-time' fluoroscopic triggering system in which the operator sees the contrast agent arrival in the desired area and then manually activates the 3D acquisition.

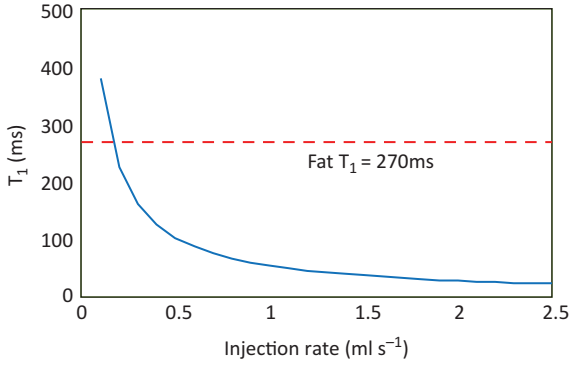


**Figure 15.15** Contrast agent $T_1$ shortening. The curve shows the reduction in blood $T_1$ as a function of 0.5 M Gd-DTPA injection rate (in ml $s^{-1}$) for a cardiac output of 5 l $min^{-1}$. Since fat is usually the brightest tissue in $T_1$w imaging, the $T_1$ of fat at 1.5 T (approx. 270 ms) is also shown.

### Reducing the $T_1$ of Blood

The reduction in $T_1$ achievable with a paramagnetic contrast agent is given by

$$\frac{1}{T_{1,post}} = \frac{1}{T_{1,pre}} + r_1 \cdot [C_A]$$

where $T_{1,pre}$ is the $T_1$ of blood prior to administration of the agent (typically 1200 ms), $T_{1,post}$ is the $T_1$ of blood following administration, $r_1$ is the longitudinal relaxivity of the contrast agent and $[C_A]$ is the concentration of the agent in the blood. Standard Gd-DTPA

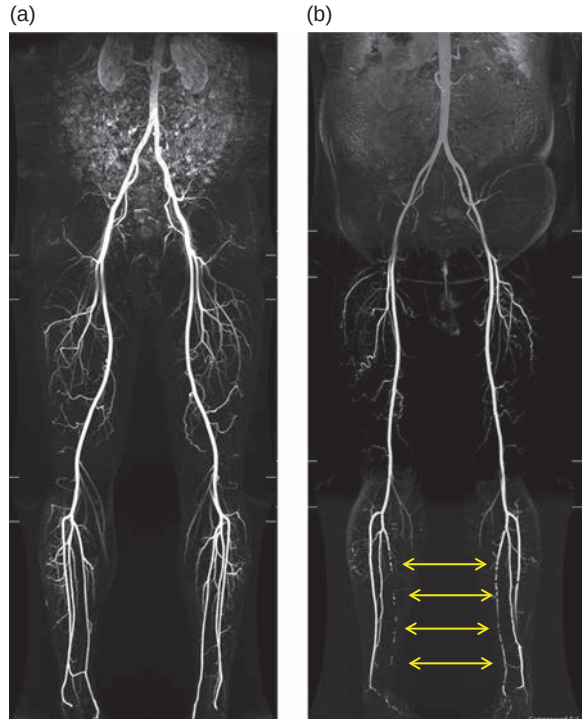


**Figure 15.16** Multiple-station peripheral CE-MRA. (a) Three-station moving table subtraction peripheral MRA using a single bolus of contrast at 3 T (patient with no vascular abnormalities). (b) A similar acquisition in a patient with run-off disease in the calf (arrows). Images courtesy of Dr Paul Malcolm, Norwich, UK.

(see Section 3.10) has a concentration of 0.5 mol $l^{-1}$ and an $r_1$ of approximately 4.5 $s^{-1}$ $mM^{-1}$ at 1.5 T.

For first-pass studies the dynamic contrast agent concentration $[C_A]$ is given by

$$[C_A] = \frac{\text{injection rate (ml s}^{-1})}{\text{cardiac output (l s}^{-1})}$$

Figure 15.15 shows the reduction in the $T_1$ of blood for standard 0.5 M Gd-based contrast agent.

### Applications of Moving Table MRA

Contrast-enhanced MRA is ideally suited to large field of view (FOV) coverage since data can be acquired in the coronal or sagittal planes without spin saturation. However, some regions of the body, e.g. the peripheral vasculature, are much larger than the maximum FOV of most MRI scanners. To overcome this FOV limitation we can use so-called 'moving table' or 'bolus chase' techniques, where the patient is automatically moved through the scanner by a fixed distance and a further MRA acquisition performed. The process is then repeated until the total desired coverage is obtained. The entire peripheral vasculature can be covered in typically three or four slightly overlapping 'stations', as shown in Figure 15.16.

CNR is maximized by first performing a mask run without contrast to use for background subtraction. SNR is high on moving table peripheral MRA because a single, large contrast bolus is imaged multiple times as it travels down the torso and legs. However, timing can be complicated because it is important for the rate at which the table is moving to roughly match the rate at which contrast flows down the legs. In practice, it is virtually impossible for MRA data acquisition and table movement to keep up with the contrast agent flow. The time for table movement between stations should be as short as possible, typically less than 3 s.

## 15.4.2 4D (Dynamic) CE-MRA

Four-dimensional (4D) MRA are methods to acquire dynamic 3D CE-MRA data. This eliminates the

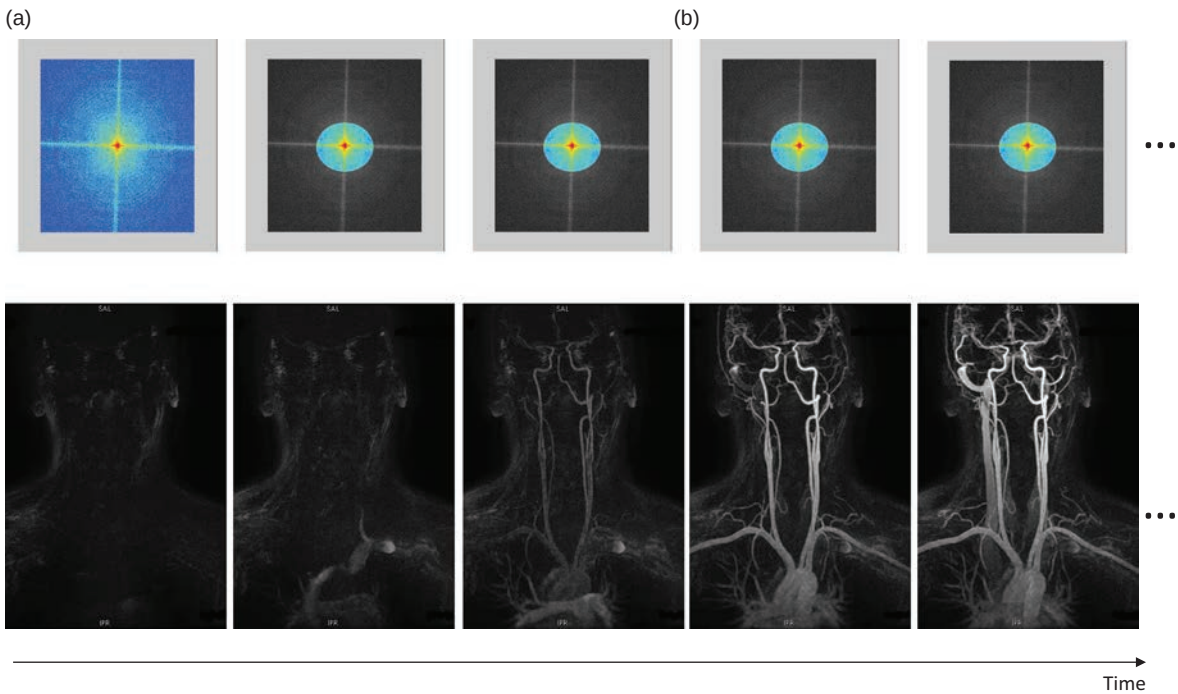


**Figure 15.17** Principle of view-sharing or 'keyhole' imaging. (a) Full k-space (colour image) is acquired at the start of the acquisition before the contrast is injected. (b) The contrast is injected but only the centre of k-space (colour region) is acquired to keep the acquisition time short. Once the acquisition is completed the peripheral k-space data are copied to produce a full k-space acquisition. Each temporal phase is then reconstructed. Each of the images is a MIP of the full 3D acquisition for that temporal phase.

necessity of accurate timing since data can be acquired continuously before, during and after the contrast agent injection in order to demonstrate temporal changes in the circulation of the contrast agent through the vasculature. Since, as discussed above, the time to acquire complete 3D data is generally much longer than the contrast agent circulation time it is necessary to use methods of k-space data sharing in order to reduce the effective temporal sampling time. The simplest method of k-space data sharing is to acquire fully sampled 3D k-space at either the start or end of the dynamic acquisition, but to only acquire data from the centre of k-space during the dynamic acquisition. This is sometimes referred to as 'key-hole imaging'. Full 3D k-space data for each temporal phase are then created by using the same peripheral data from the full k-space acquisition, as shown in Figure 15.17. Clearly in such a situation only the centre of k-space, i.e. the low spatial frequency information, is updated in each temporal phase.

Slightly more sophisticated data-sharing algorithms also update the periphery of k-space during the dynamic acquisition but at a slower rate than the centre. For example, in the **T**emporally **R**esolved **I**maging of **C**ontrast **K**inetic**S** (TRICKS) technique

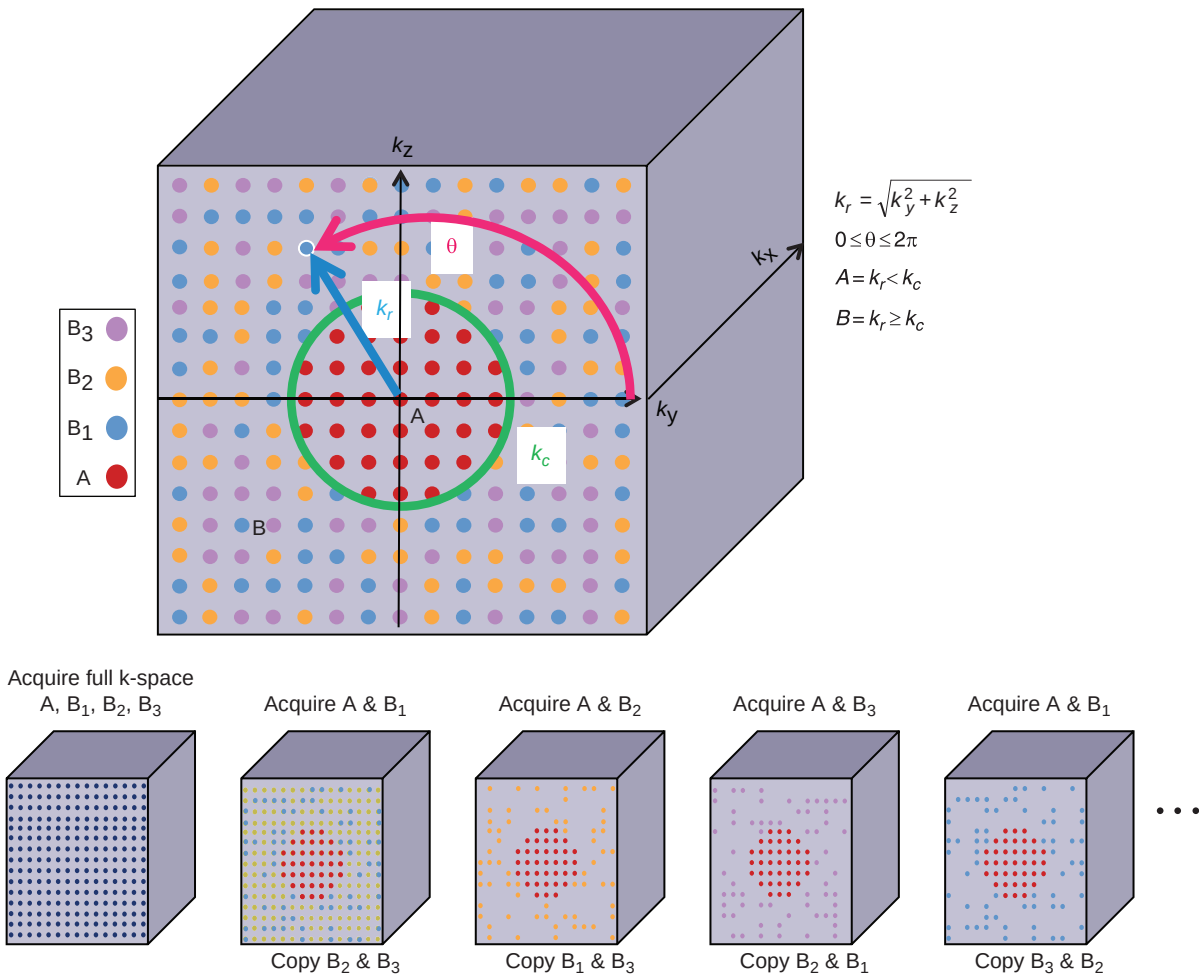


**Figure 15.18** TWIST k-space scheme. The $k_y$–$k_z$ space is divided into two regions A and B, by a circle with radius $k_c$. The points in the B region are randomly allocated to one of three equally sized groups: $B_1$, $B_2$ and $B_3$. The acquisition starts by fully acquiring k-space, then acquires the k-space regions in the pattern A–$B_1$–A–$B_2$–A–$B_3$. Regions A and $B_1$ are combined with $B_2$ and $B_3$ from the initial acquisition to create a full k-space; this represents the first phase. Phase 2 is created from the newly acquired regions A and $B_2$, $B_1$ is copied from the previous k-space and $B_3$ from the initial k-space. The third phase uses newly acquired regions A and $B_3$, $B_2$ copied from phase 2 and $B_1$ from phase 1.

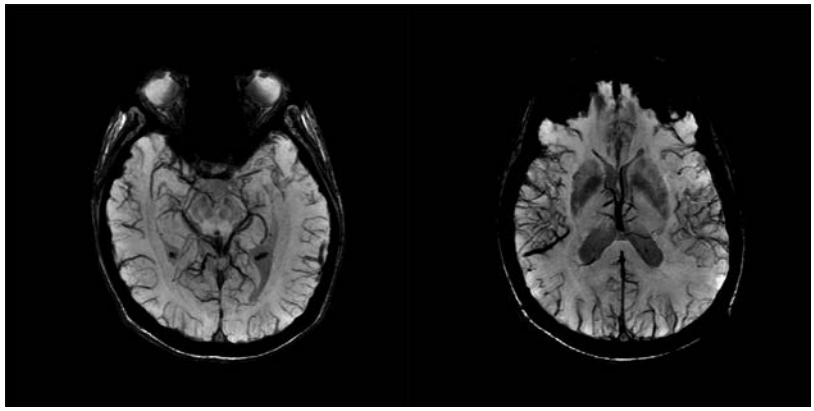

**Figure 15.19** Susceptibility-weighted images emphasizing the venous circulation.

available on GE Healthcare systems, k-space, i.e. each $k_z$–$k_y$ view, is divided into a central region (A) and three concentric outer rings (B, C and D). Initially all three regions are acquired $A_1$, $B_1$, $C_1$ and $D_1$ and then the central region is acquired every time one of the peripheral rings is acquired, i.e. $A_2$, $B_2$, $A_3$, $C_2$, $A_4$, $D_2$, $A_5$, $B_3$, … etc. Full k-space data for each temporal phase ($\phi_n$) is then reconstructed by sharing the data, i.e. $\phi 1 = [A_2\ B_1\ C_1\ D_1]$, $\phi 2 = [A_3\ B_2\ C_1\ D_1]$, $\phi 3 = [A_4\ B_2\ C_2\ D_1]$, … etc. Similarly, the **T**ime-resolved angiography **W**ith **I**nterleaved **S**tochastic **T**rajectories (TWIST) method (Siemens systems) also has a central region A but then the whole of the peripheral k-space is randomly, hence the use of the term stochastic, assigned to one of three groups $B_1$, $B_2$ and $B_3$. The A region is then sampled more frequently than the B region and the data shared to create full k-space, as shown in Figure 15.18. Note that the order of sampling in any particular B group is smoothly varying based on the angle $\theta$ and radial distance $k_r$ of an individual sampling point, with every odd point acquired on the way out towards the maximum $k_r$ and every even point on the way back in towards $k_c$. The full range of parallel imaging techniques can also be applied with these methods to further reduce the apparent temporal resolution.

### 15.4.3 Contrast Agents

The majority of commercially available contrast agents are based upon particular chelates of gadolinium. Different ligands result in different contrast agents having different relaxivities ($r_1$). The majority of ligands have $r_1$ values in the range 4.3 to 6.7 $mM^{-1}\ s^{-1}$. Gadobenate dimeglumine is an example of an agent that weakly and reversibly binds to human serum albumin (HSA), increasing the $r_1$ value to around 11 $mM^{-1}\ s^{-1}$, while gadofosveset binds strongly to HSA and has an $r_1$ value of around 30 $mM^{-1}\ s^{-1}$ and acts as an intravascular/blood-pool contrast agent. It should be noted that $r_1$ values are dependent upon the media in which they are placed – blood is different to saline – as well as the Larmor frequency at which $r_1$ is measured. Ultra-small paramagnetic iron oxide (USPIO) nanoparticles were evaluated as blood-pool agents for MRA but these have not become commercially available.

## 15.5 Susceptibility-Weighted Imaging

**S**usceptibility-**W**eighted **I**maging (SWI) is a method of enhancing susceptibility differences in tissues. The basic SWI acquisition is a long-TE 3D gradient-echo acquisition, flow-compensated in all three directions, in which both the magnitude and phase data are collected. The phase data is high-pass-filtered to remove the low-spatial frequency phase variations that occur across the field-of-view due to poor magnetic field shimming or air–tissue interfaces. The resulting phase mask is then used in combination with the magnitude image to emphasize either positive or negative, or both, phase shifts. Susceptibility-weighted imaging, like the BOLD techniques, is sensitive to deoxyhaemoglobin in the venous circulation. In SWI the negative phase mask is used to enhance the contrast for veins and is hence sometimes referred to as **S**usceptibility-**W**eighted **A**ngiography (SWA). The processed SWA images can then be displayed by the use of a **m**inimum **I**ntensity **P**rojection (mIP)

to demonstrate the venous vasculature. Alternative methods of acquiring SWA include the acquisition of multi-echo gradient-echo images followed by a weighted sum of the images obtained at different echo times without the use of the phase data. Figure 15.19 shows an example of susceptibility-weighted images.

## Further Reading


Carr JC and Carroll TJ (eds) (2012) *Magnetic Resonance Angiography: Principles and Applications*. New York: Springer.

Haacke EM, Mittal S, Wu Z, Neelavalli J and Chenga Y-CN (2009) 'Susceptibility-weighted imaging: technical aspects and clinical applications, part 1'. *Am J Neuroradiol* 30:19–30

Hartung MP, Grist TM and Francois CJ (2011) 'Magnetic resonance angiography: current status and future directions'. *J Cardiov Magn Reson* 13:19. Available online at www.jcmr-online.com/content/13/1/19 [accessed 8 May 2015].

Hashemi RH and Bradley WG Jr (2010) *MRI The Basics*, 3rd edition. Baltimore, MD: Lippincott, Williams & Wilkins.

Wheaton AJ and Miyazaki M (2012) 'Non-contrast enhanced MR angiography: physical principles'. *J Magn Reson Imaging* 36:286–304.