Chapter 3. Measuring Brain Function with MRI: The Interaction of Physics and Physiology

3.1. Introduction

The ways in which the NMR signal can be influenced by physiological changes in the brain are the subjects of this chapter. The fluctuations in the time course of the NMR signal in each voxel are the basic data available to the investigator. Understanding how these fluctuations arise from the interaction of the physiological environment with the magnetization physics is important. Both the desired signal—fluctuations consequent to neural activity changes—and the major sources of the noise interfering with detection of the signal are caused by similar processes. The relative and absolute magnitudes of signal and noise are influenced by the imaging method. Measurements and maps of brain function made with fMRI are subject to numerous confounds and caution must be used in interpreting the results.

Capsule Summary

For functional brain imaging, the desirable signal changes are those caused by neuronal activity; however, the electrical activity itself is not detectable using MRI. Instead, changes in local blood flow, volume, and oxygenation affect the population of water molecules and the microscopic distribution of the local magnetic field, both of which can affect the MRI signal intensity (see Table 3.1). These *hemodynamic* effects occur within a few seconds of changes in neural activity. There are also a number of physiologically based MRI signal changes that are *not* correlated to neural activity. These include blood flow changes with the cardiac cycle, magnetic field changes due to respiration and other complex movements, and gross motion of the subject's head.

3.2. Sources of Contrast in fMRI

Contrast was defined in Chapter 2 as being the difference in NMR signal intensity between two different tissue types. Traditional contrast mechanisms emphasize anatomical differences between spatially separated measurements. For example, images that are sensitive to differences in T1 will provide strong contrast between gray matter and white matter in the brain. With the advent of very fast methods for acquiring MR images, it has also become possible to measure physiological differences. The contrast is the signal difference between temporally separated measurements made in the same location, while the subject is in two (or more) different physiological states.

In functional brain imaging, it is changes in synaptic activity that are of direct interest. At present, this type of electrochemical activity is not measurable with MRI methods—the physical changes that might affect the NMR signal are just too small. However, local increases in synaptic activity are followed by local changes in the flow of blood. These larger-scale physical effects *do* produce measurable changes in the NMR signal, and are the basis of fMRI as it is currently practiced.

It is important to realize that fMRI is a doubly indirect measurement of changes in neural activity. These *hemodynamic* changes are not the physiological changes of interest; in addition, the actual NMR signal depends in a complex fashion on the details of the blood flow changes. Both levels of indirection (neural activity $\rightarrow$ blood flow, and blood flow $\rightarrow$ NMR intensity) can be interfered with and influenced by physiological and physical changes other than neural activity. Interpretation of fMRI brain activity maps must be made cautiously, and with knowledge of the possible interfering effects.

It is also important to realize that the hemodynamic changes are much more sluggish than neuronal events—it takes 4–6 seconds for the maximum MRI signal change to occur after the neuronal activity level increases, and it takes 4–6 more seconds for the MRI signal change to return to baseline when the neuronal activity ceases. These response times and the response magnitude vary from place to

place in the brain, and among subjects, further complicating interpretation of the results of an fMRI experiment.

The rest of this section will discuss four specific hemodynamic changes that influence the MRI signal intensity. These effects are summarized in Table 3.1. The precise physiological causes of these changes are still matters for research (e.g., the signaling mechanism that causes increased blood flow after neural activity increases is still unknown). Efforts to quantitate these changes with MRI are still being developed, as are methods to understand what factors influence the relative and absolute magnitudes of the different hemodynamic responses (e.g., does the decrease in blood vessel pliability with age affect the fMRI signal?).

Neural Activity Induces	Effect on NMR Images
Increased Blood Flow	⇒ New protons flow into slice ⇒ More protons are aligned with B_0 ⇒ Equivalent to a shorter T1 (protons are realigned faster) ⇒ NMR signal goes up [mostly in arteries]
Increased Blood Volume	⇒ Total deoxyhemoglobin increases ⇒ Magnetic field randomness increases ⇒ NMR signal goes down [near veins and capillaries]
"Oversupply" of oxyhemoglobin	⇒ Total deoxyhemoglobin decreases ⇒ Magnetic field randomness decreases ⇒ NMR signal goes up [near veins and capillaries]
Increased capillary perfusion	⇒ Inflowing spins exchange to parenchyma ⇒ NMR signal goes up [near capillaries only]

Table 3.1. The four principal physiological changes that accompany increased neural activity and which have significant effects on the NMR signal.

Blood Flow Increases with Neural Activity

Local changes in the rate of arterial blood flow into a neurally active region will change the NMR signal because the amount of water flowing into the imaged slice will change. Water molecules that enter the slice may not have previously been exposed to the excitation RF. If the image repetition time TR is less than about $3 \cdot T1$ ($T1 \approx 0.5$ s at $B_0 = 1.5$ Tesla), then protons in the slice will not recover their longitudinal magnetization completely between RF shots. Inflowing protons that have not "seen" RF before will have fully relaxed magnetization, which will be larger than the partially recovered magnetization of "stay-at-home" protons—those that have been in the slice for at least one TR interval. When the fresh protons see their first RF pulse, they will give a larger NMR signal than the protons they displaced. (This inflow effect also explains why major arteries in the brain often appear very bright in images.)

Most of the blood flow increase occurs in the arteries and arterioles near the site of increased neural activity. The result is that the increased MRI intensity consequent to neural activation occurs "upstream" (proximal) from the true site of activation. The amount of increase is also highly dependent on the details of the blood supply to the particular brain region (e.g., the extent to which the arteries are flowing across the slice or in the slice, since increased blood flow within the slice does not provide fresh protons, but just moves the stay-at-home protons around).

Blood Volume Increases with Neural Activity

Increased blood inflow in the brain is not matched immediately by correspondingly increased outflow; that is, blood volume tends to build up locally for a time. Most of this blood volume increase is in the venules and veins "downstream" (distal) from the site of neural activation. Under resting conditions, a typical gray matter voxel in the brain will be about 3–4% blood: 1% arterial, 1% in capillaries, and 1–2% in veins. During neural activation, the arterial and capillary components change only slightly, but the venous component may rise to 3% of the total voxel volume. This seems to be caused by the relative lack of smooth muscle tissue in venous tissue (compared to arterial vessel walls), making the veins downstream of an activated region behave somewhat like balloons that are inflated by the increased blood inflow.

Solid (parenchymal) brain tissue and oxygenated hemoglobin (oxyHb) have very similar magnetic properties. When immersed in the large B_0 field, they produce a small negative magnetic field—about 10^{-7} times the strength of B_0. Contrarily, deoxygenated hemoglobin (deoxyHb) produces a small positive magnetic field—also about 10^{-7} times the strength of B_0, but in the opposite direction to the extra field produced by parenchymal tissue and by blood containing oxyHb.

In healthy subjects, arterial hemoglobin is 95–98% in the oxyHb form (i.e., almost completely oxygen saturated) when breathing room air. After passage through the capillaries in the brain, the fraction of oxyHb falls to about 60% in resting gray matter; that is, about 35% of the oxygen is extracted. If this oxygen extraction fraction remains constant, increased venous blood volume will result in an increase in the amount of deoxyHb in the venules downstream from the neural activation site. Even a "large" venule is only about 300 μm (0.3 mm) in diameter, much smaller than the resolution of MRI voxels (1–3 mm). When the amount of deoxyHb in the venous network increases, the effect is an increase in the randomness of the local magnetic field at the microscopic level. In and near venules with deoxyHB, the magnetic field will be slightly larger than B_0. In the parenchyma farther away from the venules, the magnetic field will be slightly smaller than B_0. Since each voxel is a complex intertwining of blood vessels—ranging in size from 6 μm capillaries on up—and parenchymal tissue, increasing the amount of deoxyHB makes the magnetic field more heterogeneous (see Fig. 3.1). The consequence is an increase in the rate at which the NMR signal dephases; that is, T2 and T2* both decrease. When this occurs, the NMR signal intensity will decrease.

BOLD: Blood Oxygenation Level Dependence of the NMR Signal

The above picture is complicated by the fact that the oxygenation extraction fraction does *not* remain constant during increased neural activity. In particular, the blood flow increase is so large that there is an oversupply of oxyHb to the capillaries, where oxygen exchange to the tissue occurs. Although metabolic consumption of oxygen increases with synaptic activity, the amount of oxyHb supplied when synaptic activity increases is more than enough to supply this need. The result is that the blood leaving the capillaries will actually have more oxyHb during periods of neural activation than during periods of rest. A few seconds after neural activity steps up, the amount of oxyHb in the veins has actually increased over the rest state, from about 60% in rest to about 80% in activation. The total amount of deoxyHb in the venules has declined, even though the venous blood volume has increased. The net effect is an increase in T2 and T2*—since the microscopic magnetic field randomness is smaller in the venule network—and an increase in the NMR signal intensity originating from the vicinity of the veins downstream from the activated tissue. This is called the ***BOLD*** effect, which is the basis for the vast majority of fMRI experiments to date. (Even within blood vessels the magnetic field is inhomogeneous, since the deoxyHb is not spread uniformly throughout the plasma but is concentrated in red blood cells.

This intra-vascular effect is a significant component of the BOLD effect at 1.5–2.0 Tesla, but becomes less important at higher field strength [Boxerman].)

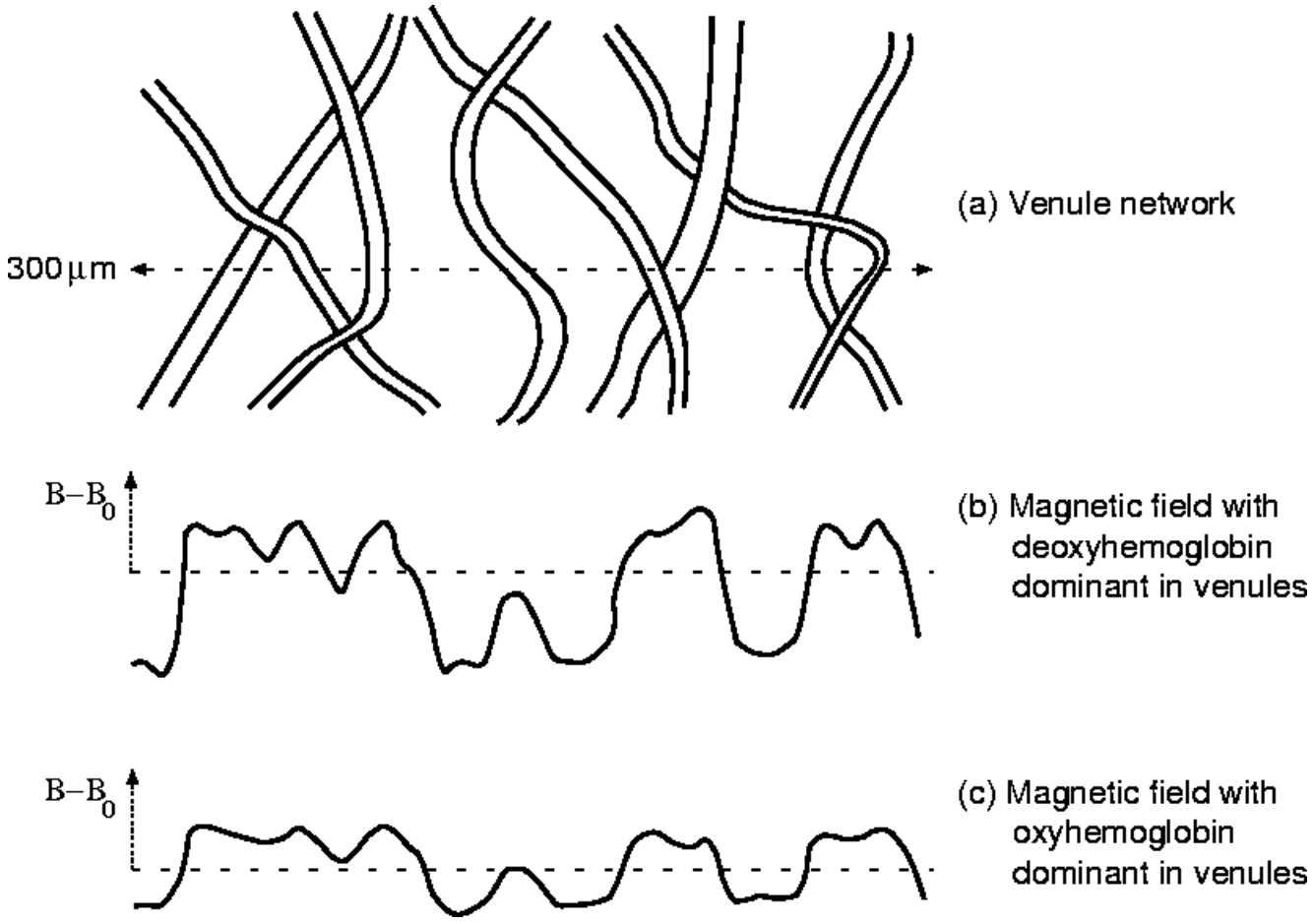

Figure 3.1. The heterogeneous structure of brain tissue at a microscopic scale produces a heterogeneous magnetic field, which in turn affects the rate of decay of the NMR signal during data readout. (a) A schematic of a network of venules, each 10–20 μm in diameter. (b) The magnetic field as a function of position along the dotted line when the hemoglobin in the venules is largely deoxygenated. (c) The magnetic field along the dotted line when the hemoglobin in the venules is largely oxygenated.

Spatial Localization of Neural Activation

Arterial flow increases and venous blood volume/oxyHb changes both occur some distance from the site of increased neural activity. Because the inflow effect depends strongly on the orientation of the arterial vessels supplying each slice, and also depends on the sequence in which the slice data are gathered, increased arterial blood flow is not generally considered a reliable method for mapping increases in neural activity. Changes in the NMR signal due to changes in venous blood volume and concentration of deoxyHb are less sensitive to the details of the geometry of the blood vessels and slice acquisition. To avoid having the image contrast include inflow effects, a small NMR flip angle can be used during the RF excitation, and/or a long TR can be used between image acquisitions.

The pooling of oxyHb in the venules downstream from the activation site can cause the activation detected by the BOLD technique to be spread out up to 10 mm away from the actual region of

increased neural activity. Collecting veins on the surface of the brain may show BOLD "activity". These veins are often in the sulci between brain gyri; when "activated", they may appear as long threads in the functional maps. Since collecting veins are larger than the venules immediately distal to the capillaries, they may actually contribute a larger BOLD effect than arises from the near vicinity of the true activation site. For this reason, the location of the peak activation (or "hot spot") may not be a precise characterization of the position of the neurally active region.

One way around this problem is to acquire images that are most sensitive to magnetic field changes in and around smaller blood vessels, and that are relatively insensitive to magnetic field changes from larger vessels. This can be done using a spin echo imaging method. As described in Chapter 2, this technique is most sensitive to magnetic field fluctuations over distances less than about 5–20 μm—about the size of capillaries and the immediate post-capillary venules. For comparison, a gradient echo imaging technique is sensitive to magnetic field changes over all spatial scales up to the size of a voxel (see Fig. 3.2). The hope is that the spin echo image changes detected between two different neural conditions will be confined to regions that are close to the neurally active parenchyma, since the small venules that will contribute most of the intensity changes will be physically close to the capillaries that supply the brain tissue.

The principal drawback to the use of spin echo imaging methods for fMRI is that the BOLD signal changes from smaller vessels, when measured with spin echoes, are 2–4 times smaller than those from larger vessels. At B_0=1.5 Tesla, a strong gradient echo NMR signal change with neural activation is 2% of the baseline (resting condition) signal level. The noise level in any measurement is 0.5–1%, so it takes many measurements (image acquisitions) to obtain reliable results—a minimum of 60 images of each slice, split evenly between two neural conditions. With spin echo acquisitions, the same level of validity will require at least 4 times as many images. The investigator must decide if the better spatial localization provided by spin echo imaging is worth the extra scanning time. Most fMRI work to date has involved relatively crude spatial questions (e.g., connections between the basal ganglia, the precentral gyrus, and the cerebellum), and so gradient echo methods have been satisfactory. However, gradient echo methods probably cannot reliably determine if an activation occurs on the anterior or posterior edge of the crown of a gyrus, for example.

It is unknown at what spatial scale the hemodynamic effects are controlled. The answer to this question will determine the physiological limit to spatial resolution. Capillaries themselves do not have smooth muscle or sphincters, so the blood flow must be regulated at an unascertained arteriolar scale, larger than the inter-capillary spacing of 20–30 μm (in cortical gray matter). There are some results that indicate such control is approximately at the cortical column scale of 0.3–0.5 mm, which raises the hope that blood-based functional imaging methods could provide a substantially complete mapping of the functional units of the cortex (a modest proposal). However, it is not clear that the arterial vessel network is conveniently mapped to the cortical column network; that is, does an individual hemodynamically controlled element in the brain correspond to at most one cortical column?

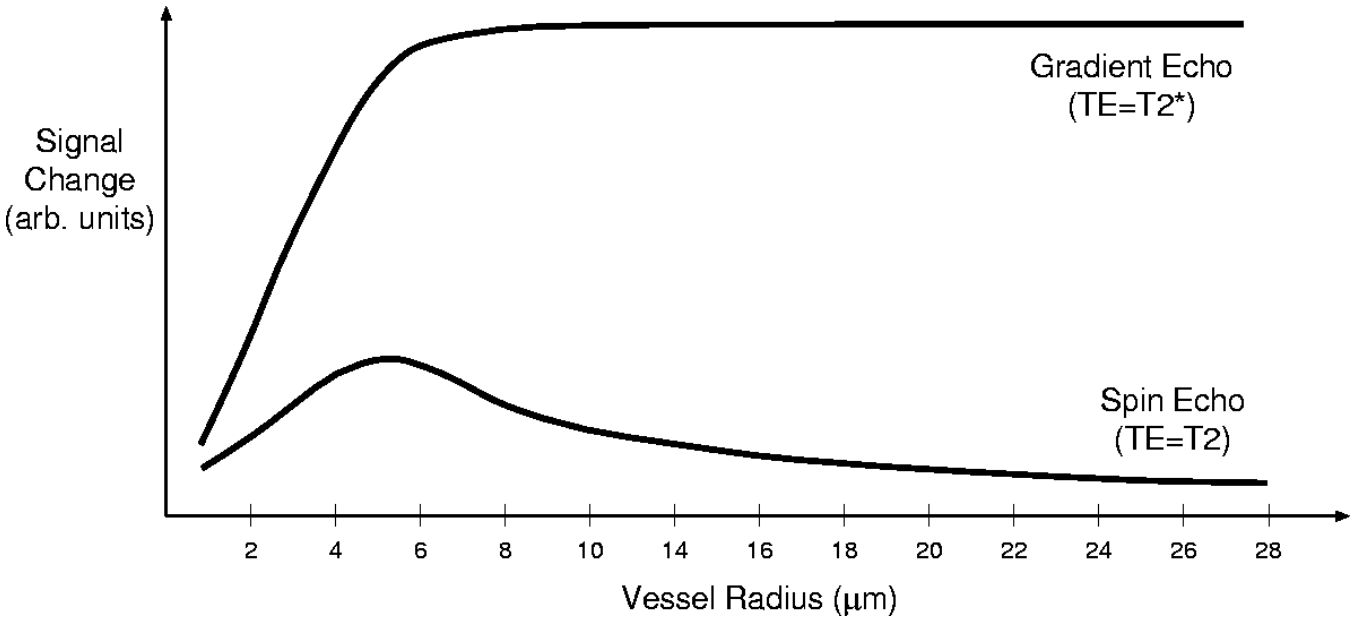

Figure 3.2. Sensitivity of gradient echo and spin echo imaging methods to the BOLD effect arising from blood vessels of different diameters [Boxerman]. Capillaries typically have a 3 mm radius. Gradient echo fMRI data is sensitive to all vessel sizes at or above small venules (5 mm radius). Spin echo fMRI data is sensitive mostly to magnetic field inhomogeneities near small venules. As a result, spin echo data may provide better functional localization that gradient echo data, but at the cost of a significant decrease in signal change.

Perfusion: Increased Flow in Capillaries

Detecting NMR signal changes that originate only at the capillary level of the circulation would presumably provide a more precise localization of the neural activation. In fact, it is possible to detect flow increases that occur only in capillaries due to an interesting physiological phenomenon: most of the water molecules that flow into a capillary do not flow out the other end. Instead, they go out into the parenchymal tissue through small pores in the capillary walls, and other water molecules enter to take their place. This water exchange does not occur in arteries or veins, which are sealed much tighter than capillaries (in the brain).

If it were possible to attach a label to arterial water molecules, wait for them to transit to the capillaries and exchange to the parenchyma, and then detect only labeled water molecules, this would provide a method for measuring the rate of flow into the capillary bed. This is possible using radioactive tracers with PET imaging (e.g., ^{15}O labeled H_2O). It is also possible using MRI, where proton magnetization is used to provide a short-lived label to the blood in the arteries at the base of the brain. These *arterial spin labeling* (ASL) imaging methods are not yet widespread, but hold the promise of providing more precise spatial localization than the BOLD technique. Their main drawback is their lack of sensitivity—the NMR signal change from a strong activation is even smaller with ASL methods than with spin echo imaging. It is likely that such techniques, even when widely available, will only be used in experiments that require highly accurate spatial discrimination.

3.3. Time Course of Hemodynamic Changes

It is quite practicable to acquire a 2D slice image every 100 ms at 3 mm resolution with single shot echo-planar imaging (EPI) methods. With 40 slices to cover most of the brain—the average left-right

dimension of an adult human brain is about 140 mm—the repetition time (TR) can be 2 s or less. Since the hemodynamic response time is about 4–6 s, the temporal details of the hemodynamic changes are observable with functional MRI. It is important to understand these details, since they form the basis for detection of neural activity.

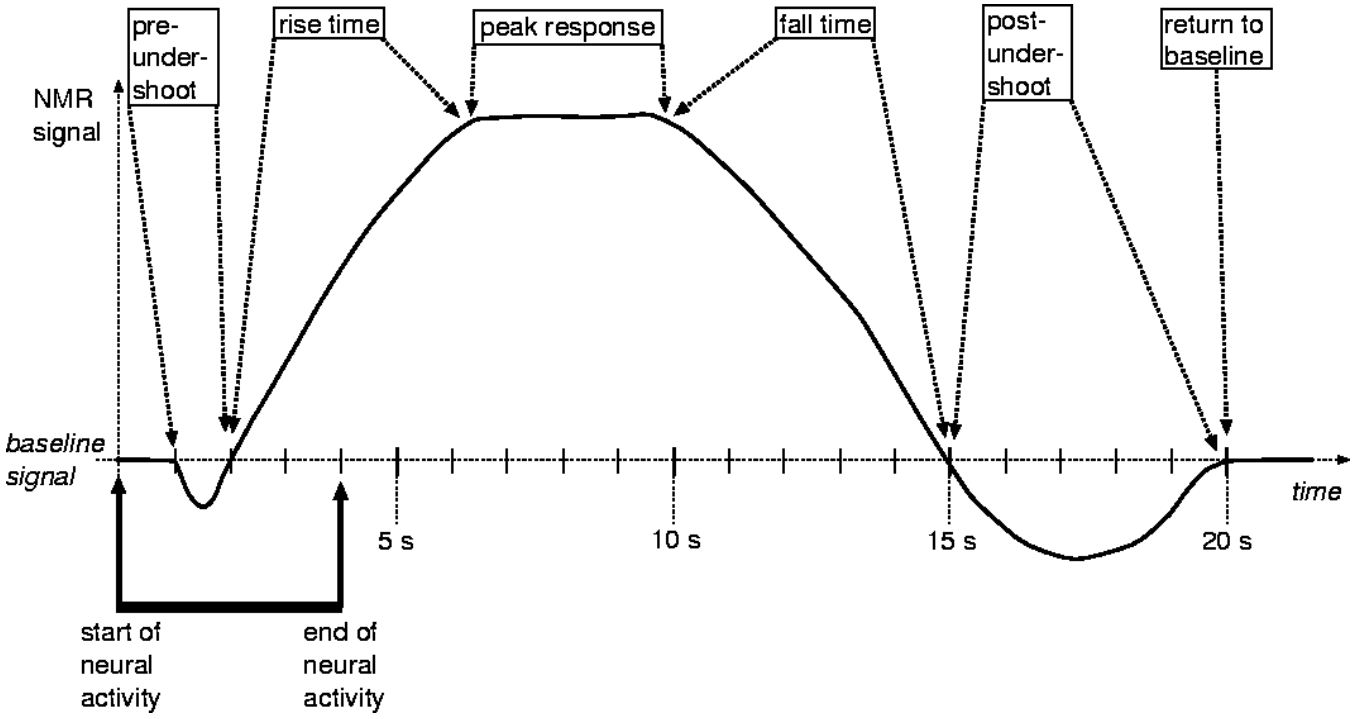

Figure 3.3. Schematic diagram of the principal features of the BOLD signal time course in a single voxel, resulting from a four second stimulation of neural activity. See the text for discussion and description. Not all features are present in all "active" voxels.

Time Course of Physiological Events and Resulting fMRI Signals

When neural activity increases in a brain region, the increased metabolic rate causes increased oxygen consumption. (There is also some evidence that lactic acid starts to builds, indicating an increase in non-oxidative glycolosis, but this is somewhat controversial.) Through an unknown signaling mechanism, within 1–2 seconds arteriolar sphincters proximal to the activated region relax and let in an increased amount of blood and red blood cells. If the neural activation remains constant, the inflow reaches a new plateau 4–5 s after the onset of the stimulus that caused the neural activity to increase. When the stimulus is removed, the blood inflow returns to its baseline level after another 4–5 s.

The time course of the change in the NMR signal from an activated region is similar to the rise-and-fall time course of the flow described above. The BOLD effect makes the NMR signal depend on the amount of and spatial distribution of deoxyHb in a voxel more than it depends directly on blood flow or volume. This complication adds some features to the fMRI time course that are not present in the arterial blood flow time course—see Fig. 3.3. The details of fMRI signal changes subsequent to changes in neural activity are subjects of ongoing research into whether they can provide additional physiological information about the brain and brain function. Significantly different results are often obtained by different investigators doing ostensibly similar experiments, and disparate explanations are proffered when similar results are obtained. This state of affairs is due to the great complexity of brain

physiology, to the many different imaging methods possible using NMR, to the extreme variety of the stimuli possible with human subjects, to the tremendous variability of the human brain between regions and between individuals, and to the vigor with which basic research into fMRI is being conducted.

The time from the stimulus onset to the peak response varies among brain voxels, between 4 and 6 seconds. This delay time is reproducible to within at least 0.3–0.5 s in any given voxel [Savoy?]. It has been suggested that voxels with longer delays correspond to those containing mostly veins more distal to the true activation site [Glover?]. If so, this would provide one way to refine the localization of fMRI activation maps without giving up the signal strength advantages of BOLD contrast. However, other results fail to support this hypothesis [Saad?]. In principle, if the delay time in any given voxel has little intrinsic variability, then with fast enough imaging it should be possible to use variations in BOLD response time as measures of neuronal signaling delays (on the order of 0.1 s). This desirable goal has yet to be realized. At present, there is no reliable method using fMRI to determine the temporal order in which different brain regions are activated. This has led to some interest in fusing the results of EEG and MEG techniques (high temporal/low spatial resolution) with fMRI (low temporal/high spatial resolution) to get the best of both worlds. [Belliveau].

A very controversial feature of the BOLD fMRI time course is the *pre-undershoot*. This is a small decrease in the NMR signal from a voxel starting about 1 s after stimulation and lasting about 1 s. This decrease is believed to come from the initial consumption of oxygen when neural metabolism increases, increasing the local amount of deoxyHb (as the oxyHb releases more of its oxygen). It takes about 2 s for the inflow of blood and oxyHb to arrive to cancel out this initial dip in the signal. Since the pre-undershoot is thought to be caused by a purely local increase in oxygen metabolism, and since the extra deoxyHb produced this way does not have much time to flow far into the venous vasculature, it is hoped that mapping the location of the pre-undershoot will provide a highly precise method for spatial localization. The pre-undershoot has only been seen with MRI at 4 Tesla, only in primary sensory and motor cortex, and only by a few investigators. The magnitude is very small and extensive signal processing is required to detect it. Much more work remains to be done to determine the utility of this effect for mapping and quantitating brain function. At best, it seems likely to be in the same category as spin echo and arterial spin labeled images: providing precise localization at the cost of extended data acquisitions.

After the cessation of the stimulus, inflow returns to baseline. The BOLD response follows the inflow curve down, but often does not stop at its pre-stimulus value; rather, it often goes below the pre-stimulus baseline and takes some extra time to rebound back up to its original value. This phenomenon is called the *post-undershoot*. It does not occur in all voxels. The duration of the post-undershoot is also a matter of controversy. The reported values seem to depend on the stimulus used to elicit brain activity, on the region of the brain being studied, on the imaging method, and on the signal processing techniques used to analyze the data. Most investigators find that the major portion of the post-undershoot is over within 10 s after the stimulus ends, but values up to 60 s have also been seen. The longer duration provides one explanation for the drift downwards in the fMRI signal baseline that is often observed.

One explanation for the post-undershoot is the hypothesis that the venous blood volume does not decrease back to baseline as rapidly as the flow decreases. In this *balloon model* [Buxton], the venous volume stays high for some time. Just after activation ceases, this extra blood volume will contain mostly oxyHb, and so the NMR signal will be higher than before the stimulus began. When a few seconds have passed, the blood flowing out of the capillaries is back at the normal lower oxygenation level and is replacing the oxyHb in the still distended veins with deoxyHb. This results in an increase in the total amount of deoxyHb in the voxel relative to the pre-stimulus condition, since the deoxyHb

concentration per unit of blood is the same as before the stimulus but the blood volume is larger. More deoxyHb distributed throughout a voxel results in a more random magnetic field, which causes the NMR signal to decay away faster, resulting in the post-undershoot. Only when the venous blood volume returns to normal (when the balloon deflates) does the total amount of deoxyHb in the voxel return to the pre-stimulus condition and only then will the NMR signal from the voxel increase back to its baseline level. The balloon model partially explains why the magnitude and duration of the post-undershoot can vary significantly between voxels. The exact distribution of veins of different sizes and different elasticities can be expected to depend strongly on location in the brain, resulting in a very heterogeneous distribution of the post-undershoot. In the balloon model, the presence of a strong post-undershoot in a voxel may be a marker for the voxel containing larger venules and veins, presumably more distal from the activation site. If this is true, the post-undershoot could be used to refine the spatial localization of BOLD imaging.

An alternative hypothesis explaining the post-undershoot involves the "coupling" and "uncoupling" of oxidative metabolism with oxygen supply in the brain [Fox, Frahm]. The oversupply of oxygen subsequent to neural activation is referred to as an uncoupling of these two phenomena. That is, there is little doubt that oxidative metabolism increases during activation, but it is also clear that the oxygen supplied delivered after neural activity increases (in the form of red blood cells containing oxyHb) is more than proportional to the oxygen required for metabolism. If neural or glial metabolism continues at a high level for some time after neural activation per se and after arterial blood flow both return to baseline, then the reverse uncoupling could occur. That is, the oxygen being delivered to the capillaries would be less than proportional to the oxygen required, so that the oxygen extraction fraction (conversion from oxyHb to deoxyHb) would have to increase. The greater concentration of deoxyHb in the veins would cause the NMR signal to drop, even if the venous blood volume did not change—it is the total amount of deoxyHb in a voxel and its spatial distribution that determines the changes in the NMR signal in T2- and T2*-weighted imaging methods.

It is important to know if the BOLD-measurable response to neural activity depends on the duration of the stimulus. This also is an area of ongoing research. The simplest possible case would be if the time course of data from extended activation could be explained as the superposition of time courses resulting from brief activations occuring in rapid succession. If this hypothesis is true, then knowing the response to a brief stimulation is all that is needed to synthesize the response to stimuli of arbitrary duration—see Fig. 3.4. If this hypothesis is not true, then knowing the history of previous stimuli and responses would be necessary to predict the incremental response to an additional stimulus.

To a first approximation, the superposition hypothesis appears to be true [Glover? Buckner?]. There are some experiments indicating that history effects need to be taken into account to accurately model the BOLD response [Friston, Vasquez]. There are some experiments indicating that the response to very brief stimuli (under 1 s) may not be additive [Bandettini]. Whether these deviations from the hypothesis are primarily neuronal or vascular has yet to be determined.

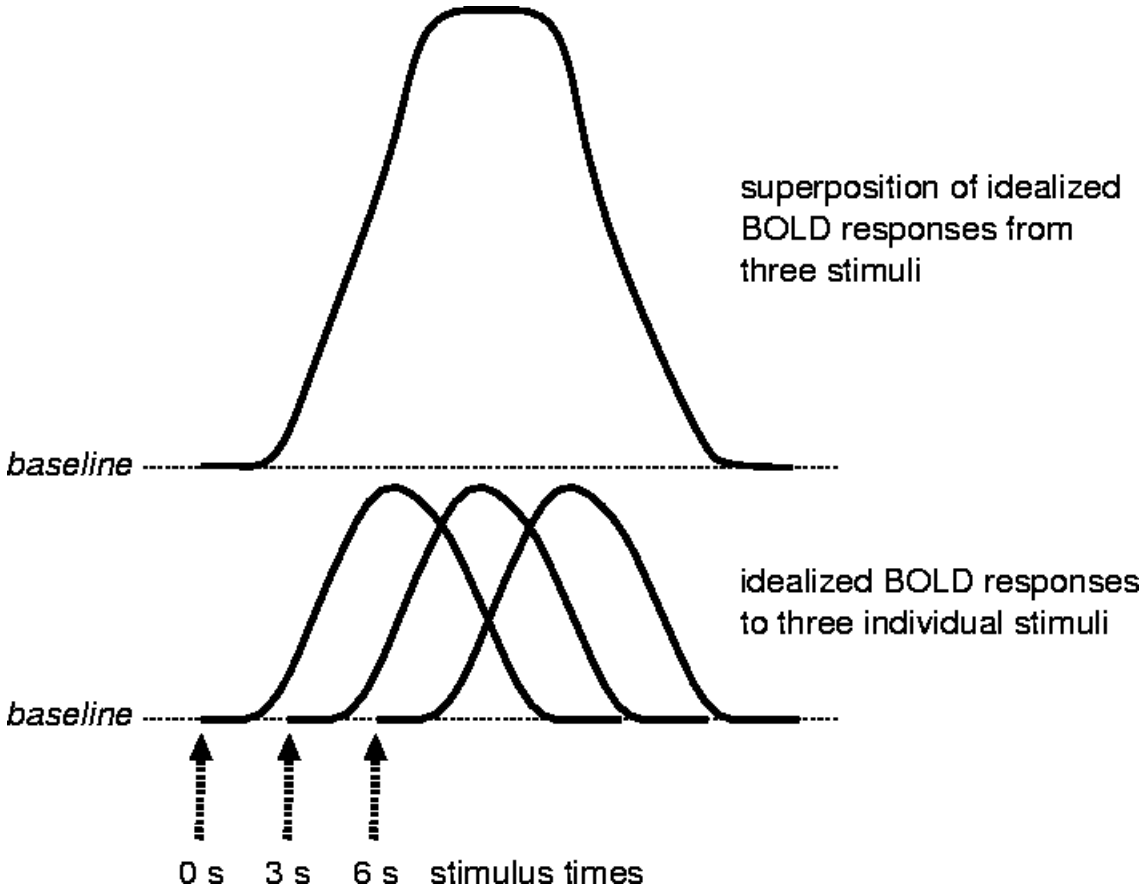

Figure 3.4. An example of superposition. The lower curves represent the BOLD responses to three brief individual neural activations. The upper curve is the sum of the three lower curves.

3.4. Artifacts in MRI and fMRI

The term *artifact* is MRI jargon for any effect in the imaging process that makes the results deviate from the ideal.

Susceptibility

The magnetic *susceptibility* of a material is a measure of its ability to generate a magnetic field when it is placed into an externally created magnetic field such as B_0. Most tissue is slightly *diamagnetic*, which means that it generates a small field opposing B_0. (Deoxyhemoglobin is *paramagnetic*—it produces a small field reinforcing B_0.) If a completely uniform diamagnetic object were placed into the scanner, it would simply reduce the magnetic field by about 1 part in 10^7, which would have no discernible effect on the images. However, human subjects are not completely uniform. The susceptibility of air is essentially zero, meaning that in the vicinity of air-tissue interfaces the magnetic field is varying rapidly in space. In the head, the nasal air passages are very close to the ventral frontal lobes; as a result, the magnetic field varies most strongly there.

Echo-planar imaging is quite sensitive to spatial inhomogeneities in the magnetic field, due to the long data readout times needed to acquire all the imaging data from one RF excitation. The form of the resulting image artifacts depends on the orientation of the slice relative to the direction in which the magnetic field varies most rapidly. Trying to correct for the effects of these field distortions is the main

reason for shimming the magnetic field (Chapter 2); however, shimming cannot correct distortions in B that occur over a region smaller than about 10–15 cm.

Field inhomogeneities through the slice will cause the signals from the top and bottom edges to be at different frequencies. Over a long readout time, these signals will then drift apart and cancel each other out. The result is a lower intensity signal, or even a complete dropout (a black hole in the image). The problem is visible at 1.5 Tesla but is much worse at 3 Tesla, where the effects of tissue susceptibility are twice as large.

Field inhomogeneities in the in-slice directions will cause echo-planar images to be distorted and possibly to have *ghosts*—faint duplicates of portions of the image displaced to incorrect locations. The cause of these effects is the distortion that the susceptibility-induced magnetic field inhomogeneities add to the scanner-controlled magnetic field inhomogeneities that are used to form the image (i.e., gradient fields). The distortions and ghosting occur in the phase-encoding direction, since that is scanned much more slowly than the frequency-encoding direction, so the frequency errors have longer to build up in that direction. These types of errors can be partially corrected with advanced EPI pulse sequence and image reconstruction software that respectively measure and compensate for the distortions in the data caused by the frequency changes. This situation is preferable to the unrecoverable loss of signal caused by through-slice field variations, since distorted data is better than no data. For this reason, coronal or sagittal EPI slices may be preferred if coverage of the ventral frontal lobe region is important, since axial slices will have the through-slice direction be in the same direction as the nasal-passage-to-brain change in tissue susceptibility.

Gross Motion

One of the most common difficulties in fMRI is motion of the subject's head during a scanning session. If some voxel is intrinsically 10% "brighter" (in the NMR sense) than its neighbors, then motion of 10% of a voxel dimension will cause the voxel's intensity to decline by about 1% since intrinsically "dimmer" tissue will have moved in to occupy some of the voxel space. The neighboring voxel(s) into which the intrinsically brigher tissue moves will correspondingly increase in intensity. A typical voxel dimension in fMRI is 3 mm, meaning that a 300 μm head movement can cause a signal change that is of the same order of magnitude as the BOLD effect.

The effects of head movement are most pronounced at the edge of the brain, where neighboring voxels may differ in intensity by 70%. (CSF is intrinscally very dark in rapid scan images. Its T1 is substantially longer than that of gray or white matter, so that when TR is small, the longitudinal magnetization in the CSF has little time to recover before the next RF excitation.)

Motion of tissue outside the image field-of-view can affect the image intensity time course as well. Notable examples are speaking and swallowing, either of which can noticeably change images near the base of the brain. This effect arises from susceptibility: the magnetic field generated by the tissue near the mouth propagates through space to the region of the brain. When the tissue moves, the magnetic field in the brain changes slightly, affecting the image data acquisition as described earlier.

An even grosser effect of motion occurs if the subject touches the RF coil. This will alter the its electrical properties, and the measurement of the RF signal that is the raw data for image reconstruction will be strongly affected. The result can be a 100% or more change in the image intensity—signal changes of this magnitude should be investigated to discover some experimental or equipment flaw, since such changes do not arise from the BOLD effect or other hemodynamic phenomena.

The consequences of movement strongly limit the kinds of experiments that can be performed with fMRI techniques. Only small movements of the subject's fingers and hands can be tolerated, for example, since any large motion will be communicated to the head. Spoken responses tend to add so

much artifact to the images that it is difficult to detect "true" (neurally related) activation. In a later chapter, methods for compensating for the effects of motion will be discussed; however, it is better to have uncorrupted data than to attempt to adjust for corruption. For this reason, it is best to design experiments that minimize subject motion, and to make sure that the subjects are comfortably fixed in position in the scanner.

Physiological Motions

Some motions are difficult to suppress entirely. Respiration also affect the images, partly due to the susceptibility effects of the moving tissue in the chest (particularly the diaphragm). This can produce changes in the signal intensity time course and can also produce time-dependent ghosts. There may also be an effect from the changes in the blood oxygen level through the respiratory cycle. If images are gathered fast enough, say with $TR < 2$ s, oscillations in the image time course at the respiratory frequency (0.2–0.25 Hz, or a period of 4–5 s) are clearly visible.

The more rapid cardiac cycle affects the image intensity time course as well. This effect is partly due to the constantly changing blood flow rate. As described in §3.2, blood flowing into a slice carries protons that have not had the same RF excitation history as protons that have been resident in the slice for a long time. The different history of these protons means that they will have a different net longitudinal magnetization, and so provide a different NMR signal. In addition to this effect, the actual movement of the blood during the image data readout interval will affect the image intensity slightly, as the protons in the blood move through the magnetic field gradients. Since blood flow velocity changes throughout the cardiac cycle, this type of movement artifact will also cause the image time course to fluctuate.

To resolve directly the signal fluctuations at the cardiac frequency—about 1.1 Hz, or a period of 0.9 s—requires an image repetition time TR of less than 0.45 s (at least twice as fast as the period of the fluctuations). This rapid scan rate is hard to achieve except in single-slice imaging. In whole brain imaging, where TR is much longer than the heartbeat period, cardiac-induced signal oscillations will look like extra noise added to the image intensity time course. This is usually the case for respiratory-induced signal changes as well, since it is common for whole brain imaging to have TR longer than the inter-breath period. Although these components of the signal may look like noise, and they definitely interfere with the detection of neurally-related signal changes, they are not technically "noise" since they are in fact a true NMR signal—just not the NMR signal of interest. In a later chapter, methods for compensating for these physiological fluctuations will be outlined.

Near the base of the brain, large quantities of blood are flowing in through the circle of Willis. The large pulsations here during the long echo-planar data readout time can also cause ghosting artifacts that will spread outward in the phase-encoding direction. (This can also occur in other imaging methods; artifacts are quite common in this region of the brain.) This is one reason that the phase-encoding direction is often chosen to be in the anterior-posterior or left-right directions, so that these artifacts will appear outside most of the brain. If the phase-encoding direction were in the inferior-superior direction, then artifacts originating at the base of the brain would appear up in the cortical regions, contaminating the fMRI signals. Spiral scan images are somewhat less sensitive to this type of flow artifact (which is caused by motion during the readout time) than are echo-planar images [Glover98].