NeuroImage **15,** 488–500 (2002)
doi:10.1006/nimg.2001.0990, available online at http://www.idealibrary.com on IDEAL®

Experimental Design and the Relative Sensitivity of BOLD and Perfusion fMRI

G. K. Aguirre,*[,1] J. A. Detre,† E. Zarahn,‡ and D. C. Alsop§

**Center for Cognitive Neuroscience and †Department of Neurology, University of Pennsylvania, Philadelphia, Pennsylvania 19104; ‡Department of Psychiatry, Columbia University, New York, New York 10032; and §Department of Radiology, Beth Israel Deaconess Medical Center, Boston, Massachusetts, and Harvard Medical School, Boston, Massachusetts 02215*

Received November 13, 2000; published online January 22, 2002

This paper compares the statistical power of BOLD and arterial spin labeling perfusion fMRI for a variety of experimental designs within and across subjects. Based on theory and simulations, we predict that perfusion data are composed of independent observations in time under the null hypothesis, in contrast to BOLD data, which possess marked autocorrelation. We also present a method (sinc subtraction) of generating perfusion data from its raw source signal that minimizes the presence of oxygen-sensitive signal changes and can be used with any experimental design. Empirically, we demonstrate the absence of autocorrelation in perfusion noise, examine the shape of the hemodynamic response function for BOLD and perfusion, and obtain a measure of signal to noise for each method. This information is then used to generate a model of relative sensitivity of the BOLD and perfusion methods for within-subject experimental designs of varying temporal frequency. It is determined that perfusion fMRI provides superior sensitivity for within-subject experimental designs that concentrate their power at or below ~0.009 Hz (corresponding to a "blocked" experimental design of 60-s epochs). Additionally, evidence is presented that across-subject hypothesis tests may be more sensitive when conducted using perfusion imaging, despite the better within-subject signal to noise obtained in some cases with BOLD.

INTRODUCTION

Arterial spin labeling (ASL) magnetic resonance (MR) imaging techniques permit the noninvasive quantification of regional brain tissue perfusion using labeled, inflowing arterial protons as an endogenous tracer (Detre and Alsop, 1999). An important feature of perfusion MR is that it can provide a physiologically relevant measure (cerebral blood flow; CBF) in absolute units (i.e., cc of blood/100 g of tissue/min) (Ye *et al.*, 2000). As a result, perfusion MR has found application in the same clinical settings in which ^{15}O PET scanning has been used. Perfusion MR has also been used to detect evoked changes in neural activity. This application is made possible by the same regional neurovascular coupling that is the basis of blood oxygen level-dependent (BOLD) imaging, although in the case of perfusion MR it is the change in blood flow itself, as opposed to the resultant changes in blood oxygenation, that provides imaging contrast.

The aim of this article is to compare, both theoretically and empirically, a particular perfusion fMRI approach (multislice, continuous ASL; Alsop and Detre, 1998) with BOLD fMRI for the purposes of detecting experimentally evoked changes in neural activity. This comparison might be conducted on any one of several levels, including the type of information that each technique provides, the relative susceptibility of each approach to particular artifacts or confounds, or differences in hardware requirements. The primary focus of this paper will be upon the relative signal and noise properties of the techniques. In particular, the statistical power of perfusion and BOLD approaches will be considered with reference to the temporal structure of different experimental designs and the testing of population (across-subject) hypotheses. We begin with a theoretical treatment of the signal and noise properties of perfusion data, and the sections below consider (i) different methods for deriving perfusion signals from source data, (ii) theoretical properties of the null-hypothesis power spectrum of perfusion data, and (iii) the presence of BOLD signal "contamination" in perfusion signals. Next, an empirical study of the relative power of BOLD and perfusion imaging techniques for different types of experimental designs is presented.

[1] To whom correspondence should be addressed at the Center for Cognitive Neuroscience, University of Pennsylvania, 3815 Walnut Street, Philadelphia, PA 19104-6196. Fax: (215) 898-1982. E-mail: aguirreg@mail.med.upenn.edu.

1053-8119/02 $35.00

THEORETICAL PROPERTIES OF PERFUSION DATA

While perfusion contrast sensitive to functional activation can be obtained using time series data in a manner exactly analogous to BOLD acquisitions, generation of image contrast that is quantifiable in perfusion units requires subtraction of image pairs with ("labeled") and without ("control") arterial spin labeling. The labeling (or control) is a spatially selective radiofrequency irradiation that precedes the image acquisition sequence. It may consist of a brief, spatially selective inversion (pulsed ASL) or continuous, velocity-driven adiabatic inversion (continuous ASL; CASL). Here we focus on the latter approach. The effect of the labeling is to reduce magnetization in the observation slice due to the presence of prelabeled spins perfusing that region. The control condition attempts to replicate the frequency-dependent and off-resonance effects of the labeling without producing significant arterial spin labeling. Subtraction of the labeled from the control images yields the signal due to perfusion.

The effects of ASL are independent of the pulse sequence used to sample the resulting changes in magnetization. This offers the potential of acquiring images with little susceptibility weighting, an advantage for examining brain regions with high susceptibility gradients such as the inferior frontal and temporal lobes. However, to maximize slice coverage and minimize acquisition times, many ASL implementations, including ours, still use gradient-echo echoplanar images for measuring magnetization. Thus, the raw image data also contain BOLD contrast, which is attenuated during subtraction of control and labeled pairs. This allows BOLD and perfusion effects to be compared within the same data set (Wong *et al.*, 1997).

Perfusion data therefore differ from BOLD data in that they are generated by subtraction of temporally adjacent observations. We discuss first a variety of techniques that might be used to realize this subtraction, then consider how selection of a particular subtraction method might impact the intrinsic noise properties of the derived signal, the presence of contaminating signal changes resulting from the BOLD effect, and the accuracy of the derived representation.

Different Methods for Subtraction

Consider a vector of n observations,

$$[C_1, L_1, C_2, L_2, \ldots, C_{(n/2)}, L_{(n/2)}],$$

corresponding to a series of temporally interleaved label (L) and control (C) images. Several methods might be conceived for extracting the differential perfusion (P) signal from this vector. The simplest is the adjacent subtraction method, in which

$$[P_1, P_2, \ldots, P_{(n/2)}] = [C_1 - L_1, C_2 - L_2, \ldots, C_{(n/2)} - L_{(n/2)}].$$

In the ideal case, label and control images would be acquired at precisely the same time, so that the only difference between the two images would be attributable to the effect of the label. Because of practical limitations of the method, and the basic requirements of the transit time of blood from the neck to the proximal cortex, at least 2 s elapse between the acquisition of a labeled image and the next control. During this time fluctuations in the signal value can create changes between the label and the control that are not attributable to the effect of label. While these signal fluctuations can be the random result of noise, they can also represent systematic changes in the magnetization as measured by the pulse sequence, i.e., the BOLD effect.

Other approaches, beyond simple subtraction, have been used in an attempt to minimize the effect of both random and systematic signal fluctuations in the source signal that are not the result of label. One method, proposed by Wong and colleagues in 1997, uses a surround average to dampen the effects of high-frequency changes in signal:

$$[P_1, P_2, P_3, \ldots, P_{(n/2)}] = \left[C_1 - L_1, C_2 - \frac{L_1 + L_2}{2}, C_3 - \frac{L_2 + L_3}{2}, \ldots, C_{(n/2)} - \frac{L_{(n/2)-1} + L_{(n/2)}}{2} \right].$$

We describe here yet another approach. The "sinc subtraction" method attempts to remove the effect of signal fluctuations by correcting for the temporal offset between the label and the control images. This is done by isolating the label image vector from the total vector, doubling the temporal resolution using sinc interpolation, and then resampling at the intermediate points in time. The result is an estimate of what the label image vector would be if it was obtained at the same time as the control images,

$$[L_1, L_2, \ldots, L_{(n/2)}] \Rightarrow [L_{1/2}, L_1, L_{3/2}, L_2, \ldots, L_{n-1/2}, L_{(n/2)}] \Rightarrow [L_{1/2}, L_{3/2}, \ldots, L_{n-1/2}],$$

where the fractional subscripts represent the projected time series at a point one TR earlier in time. Finally, the perfusion vector is generated as the difference between the control and the interpolated label images:

$$[P_1, P_2, \ldots, P_{(n/2)}] = [C_1 - L_{1/2}, C_2 - L_{3/2}, \ldots, C_{(n/2)} - L_{n-1/2}].$$

The validity (and therefore effectiveness) of the sinc interpolation approach depends upon the assumption that no meaningful perfusion signal-change exists at frequencies higher than that of the Nyquist frequency in the derivative data (the *P* vector). Naturally, this assumption is better satisfied as the TR of data acquisition is decreased.

A fourth option, termed "intertrial subtraction" by Yang and colleagues (2000), forgoes the attempt to place label and control images in absolute temporal register and instead pairs label and control images by their timing with relation to experimental events. Unlike the previously described subtraction techniques, this approach can be applied only to experimental designs that meet certain criteria: every experimental condition must have an even number of homogeneous stimulus events, and each of these events must have a duration that is an odd number of TRs in the raw data acquisition. A simple example of an experiment that meets these criteria is a sparse event-related fMRI design in which an identical stimulus is presented briefly every 18 s, while images are acquired at a TR of 2 s. The derivation of the perfusion signal then involves identifying pairs of label and control images that are acquired at equal points in time in relation to the onset of stimuli. By design, these label and control images will be drawn from separate trials. An advantage of the intertrial subtraction technique is that it most effectively removes any artifact in the measured perfusion signal introduced by BOLD effect. This is because systematic fluctuations in the source signal will have the same expectation at identical points in time following stimulus events. A disadvantage of this approach, as mentioned, is that it can be applied only to a subset of possible experimental designs of interest.

The Perfusion Null-Hypothesis Power Spectrum

BOLD fMRI data collected from human subjects in the absence of any experimental task or time varying stimuli demonstrate greater power at some frequencies compared to others. Specifically, there is increasing power at low frequencies, a distribution of power that is well characterized by a 1/frequency (1/*f*) function (Zarahn *et al.,* 1997), as well as other models (e.g., autoregressive model; Purdon and Weisskoff, 1998; Friston *et al.,* 2000). In addition to rendering ordinary parametric (Aguirre *et al.,* 1997; Zarahn *et al.,* 1997) and nonparametric (Aguirre *et al.,* 1998a) statistical tests invalid, this temporal autocorrelation causes relative reductions in sensitivity for some experimental designs (Aguirre and D'Esposito, 1999). Specifically, experiments with fundamental frequencies in the lower range (e.g., a boxcar design with 60-s epochs) will have reduced sensitivity, due to the presence of greater noise at these lower frequencies.

It is to be expected that the temporal autocorrelation of ASL perfusion data will differ from that of BOLD. Specifically, the subtraction methods that derive the perfusion time series from adjacent, temporally interleaved images will dampen long-time-scale autocorrelation present in the source noise. This is because the spectrum of the temporal derivative of a process with a $1/f$ spectrum is a flat line (i.e., the derivative of $1/x$ with respect to x is $-1/x^2$). Figure 1 presents the intrinsic temporal autocorrelation that would be expected in perfusion data derived using each of the subtraction methods described above, assuming $1/f$ noise in the original data and no additional sources of noise in the perfusion time series. Also shown for reference is the power spectrum of the original, unsubtracted data. As can be seen, all subtraction methods tend to flatten the power spectrum of the derivative time series, although gentle slopes in the spectra remain. Three of the derived time series also have reduced total power (i.e., noise) compared to the original time series. It should be noted that the curve presented for the intertrial subtraction method assumes a particular experimental design (that used in our empirical study, discussed below). Unlike the other subtraction methods, the form of the intrinsic temporal autocorrelations in perfusion data obtained with intertrial subtraction will vary depending upon the particular experimental design employed. This is because changes in experimental design alter the magnitude and distribution of temporal spacing between pairs of subtracted label and control images.

These simulations suggest that perfusion data, obtained with any of the subtraction methods, might be nearly independent under the null hypothesis, as opposed to BOLD fMRI data. The presence of only "white noise" in perfusion fMRI data would have some desirable consequences. First, the absence of serial correlation of the error terms would render unnecessary the use of the "modified" general linear model (Worsley and Friston, 1995) for the analysis of perfusion activation experiments. No form of intrinsic autocorrelation would need to be assumed, so virtually any traditional parametric or nonparametric statistical test might be used. For example, one might adopt data permutation methods to assess the significance of results (which are not valid for use with BOLD fMRI; Zarahn *et al.,* 1997)

A second consequence is that experimental designs with a concentration of power at low frequencies would become feasible. Consider an experiment that seeks to detect a slow, continuous change in neural activity over the course of several minutes. Such a design, perhaps used to study procedural learning, would be impracti-

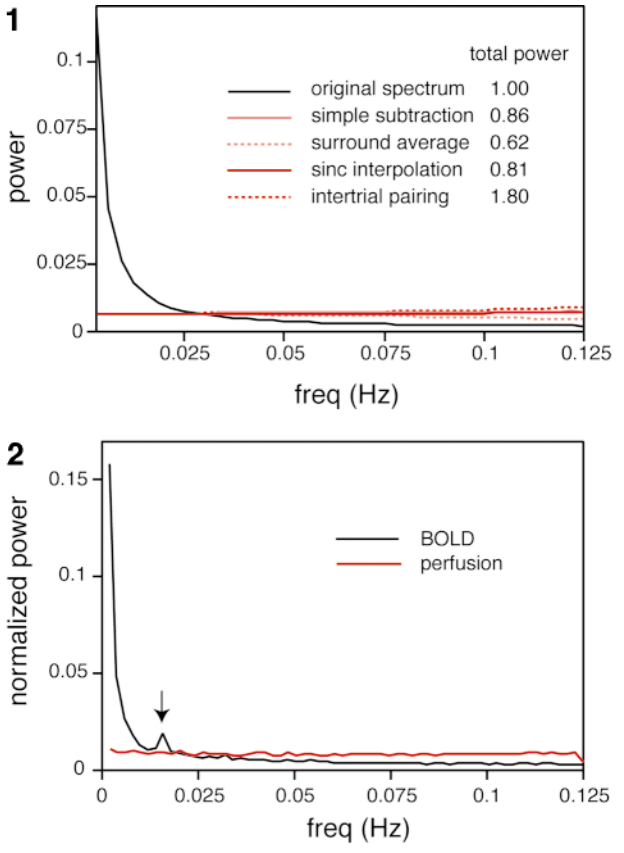


FIG. 1. Power spectra of time series derived from a simulated, null-hypothesis time series that contained noise with a 1/frequency power distribution. The power spectrum of the original vector is shown in black. Each of the derivative time series were generated using one of the three methods described in the text. The "total power" value reported is equivalent to the variance of the time series whose power spectrum is depicted. We assumed a time series of 248 images collected at a TR of 2. The data contained random, Gaussian noise, passed through a 1/*f* filter. No evoked signal changes were modeled for this simulation as this was intended to represent data under the null hypothesis. 100,000 simulations were performed for each of the subtraction methods, and the average power spectrum of the derived data was obtained.

FIG. 2. The average across-subject, voxel-averaged power spectrum observed for the BOLD and perfusion data sets, each normalized to unit total power. The fundamental frequency of the experimental design (visual stimulation) is indicated by an arrow.

cal for use with BOLD fMRI.[2] This is because the power of the experimental paradigm would lie within the greatly elevated noise range of the 1/*f* function. If perfusion data were temporally independent, however, such a design would be perfectly tractable using ASL methods. The signal to noise at a given frequency of blocked design would be dependent only on the perfusion transfer function, which is presumably low pass. Indeed, one could drive the notion of a "low-frequency design" to its logical limit and study a subject in one experimental condition during one 6-min scan and then another experimental condition during another 6-min scan, perhaps even on a different day. Such a design, used perhaps to place a subject in a particular behavioral "set," would have negligible power using BOLD fMRI but would be within the reach of perfusion in the absence of temporal autocorrelation.

[2] Slow time-course learning has been studied with BOLD fMRI, but to make such experiments feasible, the authors were required to "chop" the learning paradigm with another condition. In other words, the subject repeatedly alternated between the learning paradigm of interest and some control task (Shadmehr and Holcomb, 1997). While effective in this particular case, there may in general be undesirable behavioral consequences of task-switching during a learning paradigm.

Contamination of Perfusion Signal by Oxygenation Effects

As noted above, ASL perfusion data can be influenced by changes in tissue oxygenation in addition to reflecting changes in tissue perfusion. Both the label and the control images in the original time series are sensitive to the BOLD effect, as they are echoplanar, T2*-weighted images. As they are acquired at least 2 s apart, the label and control images will sample the BOLD response at different points along its slow, hemodynamic evolution. Consequently, the perfusion data that result from subtraction of adjacent pairs of BOLD sensitive images will be a combination of a true perfusion response plus the first derivative of the BOLD response times the TR. This is problematic if, for inferential reasons, one wishes to obtain a measure of blood-flow response that is uninfluenced by changes in blood oxygenation.

Table 1 presents the degree to which BOLD signal changes are expected to be present within perfusion data, under a variety of scanning and data analysis conditions, for a particular experimental design (that

TABLE 1

Ratio of BOLD Variance in the Derived Perfusion Signal to BOLD Variance Present in Original, Unsubtracted Signal

Subtraction method	Simulated TR of unsubtracted data (in ms)		
	4000	3000	2000
Simple subtraction	0.187	0.181	0.083
Surround average	0.057	0.034	0.009
Sinc subtraction	0.045	0.013	**0.002**

Note. These values were derived from a simulation that began with a time series containing an evoked BOLD response in the form of 31-s periods of neural activity every 62 s, convolved with a standard BOLD hemodynamic response function (Aguirre *et al.*, 1998b). No perfusion effect was simulated in the raw data as we wished to determine the magnitude of signal change in the derivative time series in the absence of a true perfusion effect. The ratio of the signal variance present in the derivative time series to the signal variance present in the unsubtracted, raw time series was assessed for each subtraction method for each of three different simulated TRs.

studied empirically below). The simulations began with a "raw" echoplanar signal that contained BOLD hemodynamic effects evoked over 31-s epochs, but did not contain any perfusion effect. The values presented in the table represent the proportion of BOLD signal variance that is present in a derived perfusion signal relative to the original signal variance. Notably, "contamination" values are not presented for the intertrial subtraction method, as this will always yield zero BOLD signal variance in the perfusion data for those experimental designs that meet its criteria.

Generally, the sinc interpolation method yields the least contamination, as does moving to shorter TRs. The degree of contamination will also vary depending upon the frequency structure of the experimental design. For the values presented in the table, an experimental design was assumed that concentrated task variance at a relatively low frequency (0.016 Hz), plus negligible power at higher harmonics. For the simple subtraction and surround average methods, power present at ever higher frequencies in the BOLD signal will produce ever greater contamination in the perfusion signal. Interestingly, this is in contrast to the sinc subtraction method, which will pass "contamination" variance from the BOLD signal to the perfusion signal only for power present at the perfusion data Nyquist frequency (e.g., 0.125 Hz for a TR of 4) or above. As there is negligible power present in BOLD hemodynamic response above approximately 0.15 Hz, the sinc subtraction approach should minimize BOLD contamination under all conditions.

The form in which BOLD contamination will appear in the perfusion data will also vary by subtraction method. For the sinc subtraction method, oxygen sensitivity appears as a high-frequency "ringing" artifact, while the adjacent subtraction method introduces large transients in the data at the onset and offset of neural activity. Simulations (not presented) indicate that the shape of the true, underlying hemodynamic response is best preserved in the perfusion data using the sinc subtraction approach.[3]

As was noted above, a vector sensitive to oxygenation effects (the BOLD signal) can be obtained by *adding* pairs of adjacent label and control images. The sinc-shifting method can be applied here as well, in that one can add the control images to the paired set of sinc-shifted label images. This method accurately recovers the BOLD response and does not introduce the artifacts that accompany simple addition of adjacent images.

In the empirical studies we present below, we have elected to use sinc subtraction to derive the perfusion data. Although the intertrial subtraction method could be used with our experimental paradigm, the design requirements of that approach would tend to limit the extension of our findings to other experimental designs. The sinc subtraction approach minimizes BOLD contamination, preserves the shape of the perfusion response, is predicted to provide a flat power spectrum under the null hypothesis, and can be applied to any functional perfusion imaging design.

[3] Although, as noted previously, the intertrial subtraction method will preserve the shape of the perfusion response perfectly for those experiments to which it is applicable.

EMPIRICAL PROPERTIES OF PERFUSION DATA

Next we describe an empirical comparison of BOLD and perfusion fMRI motivated by the theoretical work. The principal goal of the empirical study is to characterize the relative statistical power available to examine different types of experimental designs using BOLD and perfusion fMRI methods. As described above, perfusion images are expected to constitute independent observations in time under the null hypothesis. Depending upon the specific, relative levels of signal and noise present at different frequencies, one possible consequence of independence is that experimental designs in which low frequencies predominate will have better statistical power when examined with perfusion fMRI compared to BOLD methods. We wished to examine this notion empirically by characterizing the relative statistical power for a particular experimental design under BOLD and perfusion and then using this information to extrapolate the relative sensitivities of the methods for other experimental designs.

A further goal was to examine the relative statistical power of the two methods for population hypotheses. Neuroimaging questions are sometimes asked about groups of subjects, as opposed to results from a particular subject. For example, one might ask if the *population* from which a set of subjects is drawn possesses a hypothesized effect or if two different populations differ in the evocation of some effect. These types of hypotheses are appropriately tested within the context of a random-effects model (Friston *et al.,* 1999), in which a single effect measurement is obtained from each subject. A "second level" statistical test is then performed upon this group of effect values. In practice, these effect measures are the scaling values calculated for parameters of interest that model each subject's data. Such designs appropriately account for variability in the magnitude of the effect across subjects.

In the ideal case, all variability present in across-subject data would be explained by variability in the magnitude of evoked neural activity. In reality, there are several other sources of variability in the BOLD fMRI effect size. For example, between-subject differences in physiology likely produce variability in neurovascular coupling which, even in the presence of identical magnitudes of neural activity, would lead to different BOLD fMRI signal changes. These sources of

between-subject variability in the BOLD effect act to reduce the statistical power of random-effects analyses.

It is possible that the between-subject variability in task effect sizes will be different when measured with perfusion fMRI compared to BOLD fMRI. For example, it may be that changes in blood flow are more consistent across subjects than changes in the BOLD effect (which reflects, at least, blood oxygenation, blood volume, and blood flow). If so, then population analyses may be more powerful when conducted with perfusion imaging compared to BOLD imaging, even if the magnitude of signal change (relative to within-subject noise) is inferior for perfusion data compared to BOLD.

To examine the issues raised above, BOLD and perfusion data were obtained from 10 different subjects during episodic visual stimulation. Several measures were made of the data. First, average power spectra were obtained for the BOLD and perfusion data sets and compared to see if they differed in the degree of temporal autocorrelation present. Next, voxels within regions of interest in primary visual cortex that demonstrated a response to the task were identified. From these voxels, an average representation of the BOLD and perfusion evoked response was obtained. The average signal to noise was then estimated for both the perfusion and the BOLD data. This information, in concert with the estimates of the evoked response and intrinsic temporal autocorrelation, was used to generate statistical power curves to compare the relative sensitivity of BOLD and perfusion techniques for experimental paradigms with different fundamental frequencies. Finally, the relative across-subject variability in effect size was examined.

EXPERIMENTAL METHODS

Subjects and Behavioral Task

A total of 14 subjects were studied. Ultimately, the data from 4 of these subjects were rejected prior to statistical processing because of movement that exceeded predefined parameters in 1 case and the presence of radiofrequency noise of an unknown source in the other 3 (this noise impacted both the perfusion and the BOLD data). The remaining 10 subjects had a mean age of 23 (standard deviation of 3 years) and included of 6 males and 4 females. After providing informed consent, subjects viewed a back-lit projection screen from within the magnet bore through a mirror mounted on the head coil. The black, featureless screen was intermittently illuminated by a centrally located, circular checkerboard stimulus that flashed at 10 Hz. The subject was instructed to passively view the checkerboard, which would appear for 31 s every 62 s. Each individual scan lasted for just over 8 min, and a total of five scans were performed for each subject.

MRI Technique and Initial Data Processing

Imaging was carried out on a 1.5-T GE Horizon Echospeed scanner using the standard head RF coil. Foam-padded "head bars" were ratcheted onto the zygomatic arches of each subject to comfortably restrict head motion. High-resolution sagittal and axial T1-weighted images were obtained in every subject, followed by a series of chemical-shift images used to subsequently correct for image distortion. Functional imaging data were then obtained in eight axial slices with 3.75×3.75 mm in-plane resolution and 8 mm between-plane resolution with a 2-mm skip. Slices were selected to encompass the primary visual cortex. "Dummy" gradient and RF pulses preceded each scan to allow tissue to reach steady-state magnetization.

ASL images were obtained using flow-driven adiabatic inversion (Dixon *et al.*, 1986) CASL and gradient-echo echoplanar imaging. The perfusion sequence obtained 248 images per slice per scan, using a TR of 2 s and a TE of 22 ms. Labeling was performed with a 0.25 G/cm gradient and 35-mG RF irradiation applied 8 cm beneath the center of the acquired slices, which placed the labeling plane 4.5 cm below the lowest slice (Alsop and Detre, 1996). Temporally interleaved images with (labeled) and without (control) labeling were taken while controlling for off-resonance effects by applying an amplitude-modulated version of the labeling pulse (Alsop and Detre, 1998). Acquisition of each slice required 60 ms and the images were acquired in a sequential fashion (inferior to superior), the last slice being acquired 420 ms after the first one. A postlabeling delay of 700 ms was introduced between the end of continuous labeling and the slice acquisitions. This delay was selected as a relatively optimal point in the trade-off between maximizing perfusion functional activation and accurate quantification of regional CBF (Gonzalez-At *et al.*, 2000).

Offline data processing was performed using the VoxBo software package (http://www.voxbo.org). After image reconstruction and prior to motion correction, the data were sinc interpolated (by shifting the phase of the Fourier components) in time to correct for the differential timing of fMRI slice acquisition in space (Aguirre *et al.*, 1998b). The data were then motion corrected using a six-parameter, rigid-body, least squares realignment routine (Friston *et al.*, 1995). The odd and the even images of each time series were separately realigned to the first and second images (respectively) of the first scan. This was done as the label and control images have different image intensities, which would be interpreted erroneously as movement by the realignment routine if all images were referenced to the same target.

The BOLD and perfusion data were then derived. First, the series of label images was projected back in time by one TR using sinc interpolation. BOLD data

were obtained by taking the average of the label and the time-matched control images, while the perfusion data were generated by subtraction. Conversion to CBF values was effected using a modification of a previously described approach (Alsop and Detre, 1996) which assumes a labeling efficiency of 71% (Alsop and Detre, 1998). Briefly, the intensity of cerebral spinal fluid within the ventricles was used to calibrate the ASL induced signal change (Chalela *et al.*, 2000), which is equivalent to assuming that the label actually remains in the microvasculature rather than exchanging with tissue water. Failure of this assumption would lead to an underestimate of flow in short T1 tissues such as white matter and an overestimate of flow in long T1 tissues. As gray matter T1 is close to blood T1, quantification errors in normal gray matter should be minimal. The resulting BOLD and perfusion data sets contained images with an effective TR of 4 s. Both data sets were smoothed in space with a three-dimensional, 10-mm FWHM Gaussian kernel.

Characterization of Temporal Autocorrelation

Power spectra were determined (by multiplying the discreet Fourier transform of each time series by its complex conjugate) for each brain voxel for each scan for each subject. These voxel power spectra were averaged across the brain voxels to determine a "voxel-averaged power spectrum," or spatially averaged power spectrum, per scan. This approach ignores variability in the shape of the power spectrum that has been observed from voxel to voxel in a single subject's data (Friston *et al.*, 2000). Notably, the existence of such variability is also ignored (or compensated for) in the vast majority of univariate analyses of fMRI data, including SPM. Consequently, the results presented below apply to the average behavior of BOLD and perfusion data and might vary across different tissues (e.g., gray and white matter). The five BOLD and five perfusion power spectra (one from each scan) from each subject were averaged separately, providing an average BOLD and perfusion power spectrum per subject. Finally, these power spectra were averaged across subjects, providing an across-subject, spatially averaged power spectrum for the BOLD and perfusion data.

Creation of Statistical Maps

Voxel-wise analysis of the functional imaging data was conducted to identify voxels with a significant response to the visual stimulation. Appropriate statistical models were created for the concatenated BOLD and perfusion data for each subject. This analysis employed the modified general linear model of Worsley and Friston (1995) and used a reduced Fourier basis set of the first six harmonics of the 62-s trial window (six sines and six cosines) (Josephs *et al.*, 1997). A reduced set, as opposed to inclusion of all higher frequency harmonics, was selected to provide greater statistical power with the expectation of negligible loss in modeling, as little power exists within the hemodynamic transfer function above 0.15 Hz. Partial-*F* tests were used to evaluate the significance of the variance in the data explained by these 12 covariates together. A specific advantage of this analysis approach is that sensitivity is not dependent on the shape of the response (within the constraints of the reduced basis).

As has been observed previously, the BOLD data obtained here were found to have substantial temporal autocorrelation (1/*f* noise) (see Characterization of Temporal Autocorrelation, above). To account for this, a 1/frequency function was fit to the (square root of the) average BOLD power spectrum from each subject, ignoring those frequencies at which power attributable to task might be expected. The time-domain representation of the 1/*f* curve was placed within the *K* matrix (Worsley and Friston, 1995; Zarahn *et al.*, 1997) along with a filter designed to remove low-frequency confounds (below 0.014 Hz) and high-frequency noise at the Nyquist frequency (0.0125 Hz) and a high-pass kernel (a standard HRF; Aguirre *et al.*, 1998b). The removal of low frequencies and application of exogenous smoothing is necessary, even in the presence of the 1/*f* model, because there is substantial variability in the actual magnitude of the low-frequency power from voxel to voxel (Friston *et al.*, 2000). Finally, the BOLD data within each voxel were also normalized by the mean signal value for that voxel prior to statistical analysis. This is a standard step in BOLD data analysis and is motivated by the finding of "gain" or "scaling" effects within the raw data resulting from changes in the setting of the scanner's amplifiers from one session to the next. Note that this measure is different from correction for "global signal effects" (which was not performed in these analyses) and is not impacted by the presence of task variance within the data.

The perfusion data were not found to have any substantial temporal autocorrelation. Therefore, the analysis of these data was able to assume independence of the errors and did not require modeling of intrinsic temporal autocorrelation, "notch" filtering, or temporal smoothing with a hemodynamic response function. Mean normalization of the perfusion data was also not necessary as the calculation of CBF values from the raw perfusion data involves scaling (although in this case it is scaling by the signal present within the cerebral spinal fluid).

The primary visual cortex was defined upon each subject's T1 images by one of the authors (G.K.A.). The V1 region of interest comprised approximately 200 voxels, including both left and right hemispheres. Statistical maps corresponding to a partial-*F* test of the explanatory power of the Fourier basis set were created and masked with the predefined regions of interest. The maps were then thresholded at an *F* value

corresponding to a Bonferroni corrected, region-wise $\alpha = 0.05$, and any voxels within the search region that surpassed this threshold were identified.

An average evoked response during the 62-s blocks was determined for each subject. First, the average BOLD and perfusion signal across significantly activated voxels within the V1 ROI for each subject was obtained. These time series were then subjected to the analysis models described above, and the Fourier coefficients were determined. These coefficients define the best fit of the covariates of interest to the data and were used to define the average within-subject evoked response. The evoked response from each subject for each imaging method was scaled to have unit variance, and then the average of these responses across subjects was obtained.

The impulse response function implied by the evoked signal changes observed for the perfusion and BOLD data was then calculated. Assuming a step function of neural input and linearity and time invariance of the hemodynamic transformation, an estimate of the impulse response function was obtained for the BOLD and perfusion data by taking the first derivative of the (sinc interpolated) evoked response (Oppenheim *et al.*, 1983).

Theoretical Background for Power Analyses

The power of a statistical test is its ability to reject the null hypothesis when it is false. The assessment of power (within the setting of the general linear model) requires an assumption of the magnitude of the effect to be determined relative to the noise present. An appropriate form for this measure of signal:noise (here termed δ) is the ratio of the parameter estimate for a contrast of interest to the standard deviation of its estimator. Given the expression for the general linear model,

$$Y = G\beta + \varepsilon,$$

where Y is a data vector, G is a (filtered) matrix of covariates, β are the parameters, and ε are the errors, δ is calculated as

$$\delta = \frac{c\beta}{\sqrt{\mathrm{Var}[c\hat{\beta}]}},$$

where c is a row vector of contrast weights. Expanding the denominator for the case of serially correlated error terms (Worsley and Friston, 1995) we obtain

$$\delta = \frac{c\hat{\beta}}{\sigma\sqrt{c\,(G^T - G)^{-1}\,G^T - VG\,(G^T - G)^{-1}\,c^T}},$$

where V is the (filtered) variance–covariance structure of the errors.

Given δ derived from a "reference" experimental design (δ_R), it is possible to calculate the δ for another experimental design of interest (δ_I). (Details regarding this approach can be found in Zarahn and Slifstein, 2001) Assuming that the ratio of the magnitude of the impulse response function to the error variance is constant across designs, then the signal:noise value for the experiment of interest is

$$\delta_I = \delta_R \cdot \left[\frac{c_I\beta_I}{\sqrt{c_I\,(G_I^T G_I)^{-1}\,G_I^T V G_I\,(G_I^T G_I)^{-1}\,c_I^T}} \Big/ \frac{c_R\beta_R}{\sqrt{c_R\,(G_R^T G_R)^{-1}\,G_R^T V G_R\,(G_R^T G_R)^{-1}\,c_R^T}} \right].$$

Given δ, the degrees of freedom of the study, and the desired false-positive rate, power (the probability of detecting the presence of signal) can be derived from standard tables.

Estimation of the Noncentrality Parameter of the *t* Distribution and Power Analysis

In the current study, we wished to evaluate the statistical power of perfusion and BOLD techniques in the setting of experimental designs with different temporal structures. Given the hypothesized differences in the presence of intrinsic temporal autocorrelation between perfusion and BOLD, we expected that the two techniques would differ in their relative power for detecting different paradigm frequencies. For this study, we wished to treat the δ associated with our experimental design (31-s checkerboard flash every 62 s) for the perfusion and BOLD techniques as the reference δ. To obtain an estimate of δ from the perfusion and BOLD data sets which were not biased by statistical thresholding, we estimated the noncentrality parameter of the t distribution for all of the voxels within our predefined regions of interest. For reasonably high degrees of freedom, the noncentral t distribution is well approximated by a standard normal with mean equal to δ.

As our initial statistical analysis yielded F values (see Creation of Statistical Maps, above), it was first necessary to generate relevant t statistics from the analyses of the BOLD and perfusion data. To do so, the average, across-subject evoked response described above was projected onto the partial Fourier basis to define a single t contrast for all subjects. This approach does have the undesirable property of using information from the data itself to inform subsequent tests of the data, but the implied bias is both minimal and proper. The influence of any one subject's evoked response upon the average evoked response was slight (1 of 10), so that any bias toward detection of evoked responses within any subject would be minimal. In any

event, the *t* values thereby obtained for the BOLD and perfusion data sets would be equally inflated, rendering any such effect unimportant in the comparison between the modalities.

The population of (unthresholded) *t* values present within the V1 regions of interest was then obtained separately for the perfusion and BOLD data for each subject. We assumed that there were two populations of voxels within these regions: those voxels with no experimental effect (the *t* values of which are centered about zero) and those voxels with a true experimental effect (centered about some positive, nonzero point). The distribution of the *t* values within these regions should therefore approximate a bimodal Gaussian distribution (see Fig. 6). The histogram of the *t* values from each subject for each scanning method was fit with a function which was the sum of two Gaussians, with the parameters of mean, height, and width free to vary. Initial estimates of the mean were provided by one of the authors (G.K.A.) following visual inspection of the data, and subsequent fitting was performed using a gradient-expansion algorithm to compute a nonlinear least-squares fit (provided as a component of the Interactive Data Language package; Research Systems, Boulder, CO). The mean of the second (noncentral) Gaussian was taken as an estimate of the noncentrality parameter of the *t* distribution for each imaging method within each subject.[4] The mean of these values across subjects was also obtained, yielding the reference noncentrality parameter for the perfusion and BOLD data.

These noncentrality parameters were then taken as a δ_R, the reference signal:noise level for the perfusion and BOLD techniques. We then created hypothetical experimental designs that assumed a temporal structure different from that of our reference design. Specifically, the hypothetical designs were pure sinusoids of different frequencies but of equal total variance. These models of neural activity were convolved with the appropriate impulse response function (perfusion or BOLD—derived as described above) and the corresponding δ_I was calculated.

Evaluation of Across-Subject Variability in Effect Size

We wished to obtain an estimate of the magnitude of the evoked hemodynamic response associated with the perfusion and BOLD scanning techniques for each subject. To do so, the average time series was obtained from the voxels within primary visual cortex identified by the Fourier basis set. As was described above for the estimation of the noncentrality parameter, each average time series for each subject was then evaluated using a target vector created with the across-subject average evoked response. The unscaled β value associated with this contrast was thereby obtained for each subject for the perfusion and BOLD techniques. The population of β values obtained with each method was evaluated with an unpaired *t* test, thereby producing a within-region, random-effects analysis.

[4] For the BOLD data for two subjects, the entirety of the V1 ROI was filled with "activated" voxels. As a consequence, the two-Gaussian model provided a poor fit, and a single-Gaussian model was used instead.

RESULTS AND DISCUSSION

Power Spectra

Figure 2 presents the average, across-subject power spectra obtained using the BOLD and perfusion methods. As has been observed before, ever increasing power is present at ever lower frequencies in the BOLD data. The perfusion power spectrum, on the other hand, is essentially flat. This confirms that ASL perfusion data constitute independent observations. While these data were collected under experimental conditions, and therefore cannot strictly be taken as representing the null-hypothesis power distribution, it is simple to identify those frequencies at which the ongoing task impacted measured power. The fundamental frequency of the visual stimulation is indicated by an arrow, and power attributable to higher harmonics of the task are negligible. Notably, alteration of power at the task frequency is not readily visible in the perfusion power spectrum. This may be the result of a smaller number of voxels that show a perfusion effect or a smaller magnitude of perfusion signal change relative to noise, or both.

Evoked Responses

Statistical models were created to test for changes in the BOLD and perfusion data sets in response to the visual stimulation task. While the models created for the perfusion data were able to assume the independence of the observations, the BOLD analysis included a subject-specific representation of the 1/*f* power spectrum shown in Fig. 2. The analysis of the BOLD data also applied exogenous temporal smoothing in the form of a hemodynamic response function and filtering to remove power at frequencies below that of the task. These are standard components in many analyses of BOLD fMRI data and are mandated by the presence and variability of correlation in the error terms (Friston *et al.*, 2000). One consequence of the difference in model construction between the BOLD and the perfusion data is that roughly twice as many effective degrees of freedom were available to test hypotheses with perfusion compared to BOLD (603 versus approximately 300).

Figure 3 provides examples of areas of activation identified using the Fourier analysis for the BOLD and

perfusion data. For all subjects, a greater number of voxels was identified using BOLD compared to perfusion (mean 45 voxels per subject with perfusion within the region of interest versus 134 voxels with BOLD). Additionally, BOLD activation frequently extended far beyond the confines of the anatomically defined area of primary visual cortex, compared to the more localized perfusion signal changes. In all cases, the suprathreshold voxels identified for perfusion were a proper subset of those identified for BOLD. While it is the case that neural activity in areas beyond V1 would be expected in response to visual stimulation, it is also possible that some extension of the BOLD activation is driven by vascular, as opposed to true neural, changes. While this notion was not tested here, other studies that have compared the spatial selectivity of BOLD and perfusion responses suggest that this may be the case (Luh *et al.,* 2000).

Figure 4 presents the across-voxel average evoked response for BOLD and perfusion, averaged again across subject, and Fig. 5 presents the impulse response function (IRF) implied by these evoked responses. In keeping with previous findings (Yang *et al.,* 2000), the BOLD and perfusion IRFs share roughly the same shape, although the perfusion response is narrower in time.

Estimation of Noncentral *t* and Relative Within-Subject Power

The histogram of *t* values for all voxels within the (unthresholded) V1 region of interest was obtained for each subject under each method, and fit with a bimodal Gaussian model to estimate the signal:noise of the experiment. Figure 6 shows an example histogram from one subject's perfusion data, along with the Gaussian fits.

The average signal to noise (δ_R) was calculated across subjects for the BOLD and perfusion methods and revealed a roughly 40% better signal:noise for BOLD compared to perfusion (BOLD average δ, 6.7; perfusion average δ, 4.9). Coupled with estimates of the intrinsic temporal autocorrelation of the methods (Fig. 2) and the shapes of the hemodynamic responses (Fig. 5), the signal:noise values were used to calculate expected δ values for a range of different experimental designs (Fig. 7). The experimental designs assumed here were pure sinusoids of different frequencies, the results for which are readily comparable to that expected for a boxcar design with the same fundamental frequency. These calculated δ values are related to statistical power in a fairly straightforward manner (given a value for α and the effective degrees of freedom). Technically, the perfusion method will provide slightly better statistical sensitivity than BOLD for the same estimated δ value. This is because of the larger number of effective degrees of freedom available in perfusion analyses (although in practice this is a fairly small effect).

As can be seen from Fig. 7, BOLD enjoys a superior signal:noise ratio for all frequencies of experimental design, save for those at and below a frequency of 0.0089 Hz (the point at which the black line crosses below the red line). This frequency corresponds approximately to a blocked design with 60-s on and off epochs. Therefore, for most within-subject experimental designs, BOLD provides greater statistical sensitivity than perfusion. However, because of the absence of additional noise at low frequencies, perfusion methods do provide the better option for experimental designs with power concentrated at low frequencies.

There are several points at which the assumptions underlying this model of statistical power could be invalid or less relevant in practice. First, the BOLD data analyzed here had an effective TR of 4 s. BOLD data are typically collected at a shorter TR, thus affording a greater number of degrees of freedom. This objection is offset, however, by the fact that the BOLD data analyzed here were also the average of two adjacent observations. Therefore, any loss of degrees of freedom relative to a 2-s TR experiment would be (roughly) balanced by an increase in signal to noise. Second, a TE of 22 ms was used here for collection of the raw data, which represents a compromise between optimizing BOLD and perfusion sensitivity. Slightly greater signal to noise might be obtained for either method through refinements of the pulse sequence parameters. Notably, use of a shorter postlabel delay in the perfusion activation would have increased the magnitude of the perfusion signal change, although at a cost in the accuracy of blood-flow quantification. A third assumption of this model is a linear transform between neural activity and imaging signal. This assumption was recently examined simultaneously for perfusion and BOLD methods (Yang *et al.,* 2000). The primary finding was that equivalent nonlinearities are present in the response for BOLD and perfusion for closely spaced visual stimuli, although as with all tests of linearity under these circumstances it is difficult to ascribe the nonlinearity to the transform of stimulus to neural activity, or the transform of neural activity to imaging signal. In any event, these nonlinearities would tend to be present in higher frequency experimental designs and do not impact the conclusions made regarding relative power at lower frequencies.

Particular implementations of BOLD and perfusion methods at different centers might also alter the vertical separation between the power curves presented here. Despite all of these caveats, however, there is little doubt that perfusion will provide superior signal:noise for experiments with a preponderance of power at low frequencies.

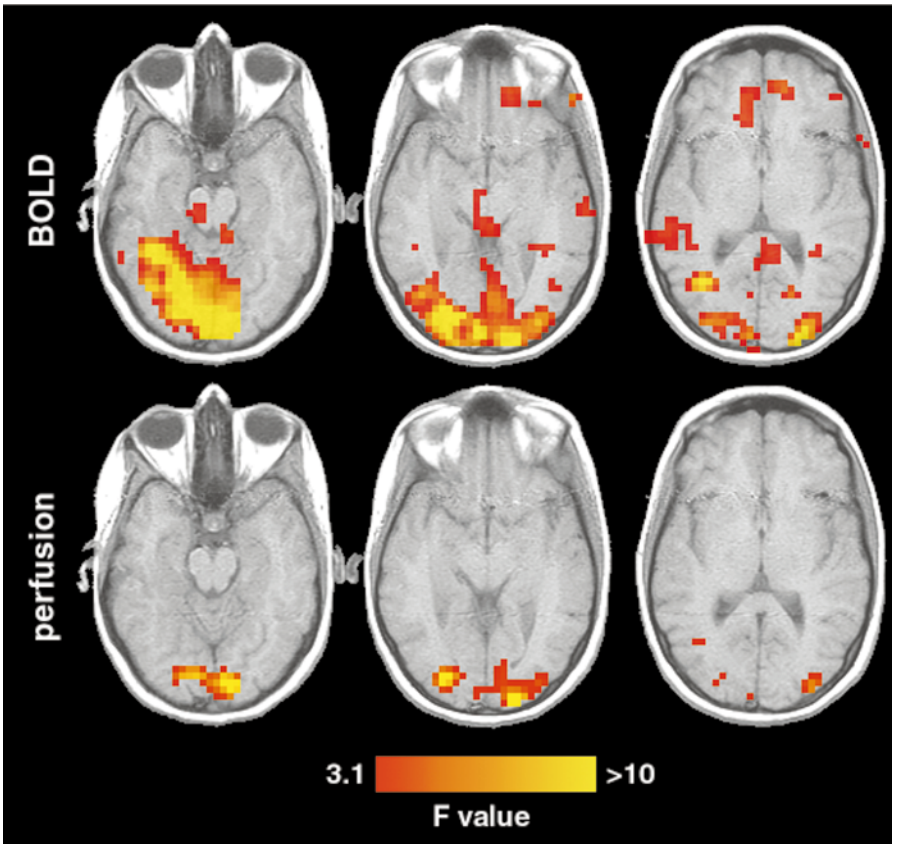


FIG. 3. Voxels with significant signal change in response to visual stimulation as detected by the Fourier basis set. The upper row shows the results for the BOLD data for one subject, while the lower row is the corresponding perfusion data. The maps are thresholded at a level corresponding to a region-wise α of 0.05, Bonferroni corrected for the number of voxels present within a primary visual cortex region of interest. Voxels beyond this region of interest are shown to allow comparison of the spatial extent of BOLD and perfusion activation.

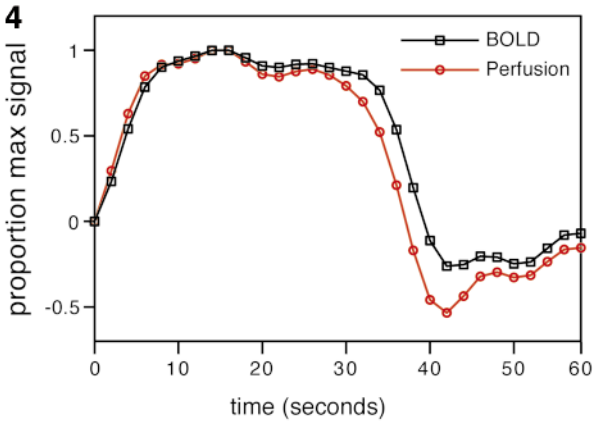
Relative Across-Subject Power

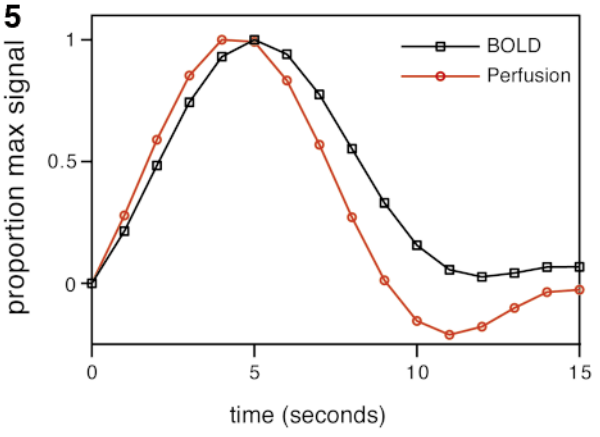
The β values associated with the evoked response within primary visual cortex were obtained for each subject for the perfusion and BOLD methods. A random-effects analysis of these data (corresponding to the hypothesis that there is an increase in imaging signal within the primary visual cortex in response to visual stimulation across subjects) was tested using an unpaired *t* test. The *t* value (9 *df*) associated with the set of perfusion β values was 10.7, while that associated with the BOLD data (9 *df*) was 8.9. Despite the finding that BOLD enjoys greater signal to noise *within* subjects, the perfusion method provided the more robust result *across* subjects. The explanation for this finding is that while the magnitude of the evoked signal change is smaller relative to within-subject noise for perfusion compared to BOLD, that evoked signal magnitude is more consistent across subjects.

To assess the reliability of this finding, we conducted a subgroup analysis. The 10 subjects were randomly assigned to two groups of 5 and random-effects *t* values calculated for the BOLD and perfusion data for each group. Group 1 was found to have a *t* value of 9.0 for the perfusion and 6.8 for the BOLD, while group 2 was found to have a *t* of 10.3 for the perfusion and 7.3 for the BOLD (4 *df* all cases). These findings suggest that hypotheses that address themselves to populations of subjects may be more profitably tested using perfusion imaging as opposed to BOLD fMRI. This recommendation is tempered by the appreciation that it is based on the limited observations reported here. To conclusively demonstrate that perfusion data have reduced across-subject variability in evoked responses compared to BOLD, several more sets of β values similar to that collected here would be required.

CONCLUSIONS

Previous studies have described well several desirable features of perfusion imaging. Primary among these is the provision of a physiologically relevant measure (i.e., cc of blood/100 g of tissue/min), in contrast to BOLD, which furnishes a signal that has no (simple) absolute interpretation. It is also generally accepted





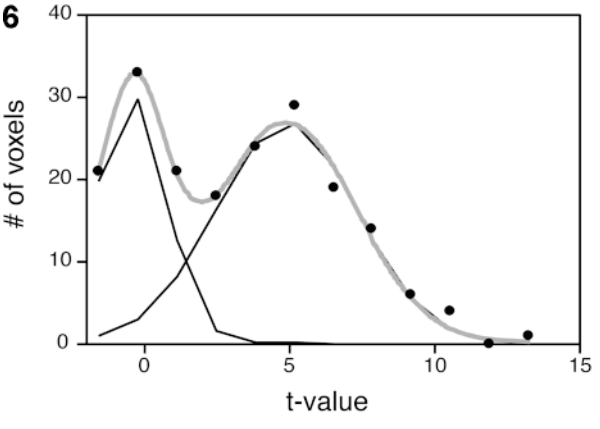


FIG. 4. The average across-subject, voxel-averaged evoked response from primary visual cortex for the BOLD and perfusion data sets. Visual stimulation began at time zero and continued for 31 s. The time series were set to start at zero and scaled to have a unit maximum.

FIG. 5. The impulse response functions implied by the evoked responses presented in Fig. 4, assuming a step function of neural activity and linearity of the transform of neural activity to imaging signal.

FIG. 6. An example fit for estimation of noncentral t (δ). Indicated by the filled circles is the histogram of all voxel t values from the (unthresholded) primary visual cortex region of interest for one subject's perfusion data. The black lines indicate the fit, obtained by least-squares estimation, of two Gaussians to the histogram data. The gray line is the fit formed by the sum of the two Gaussians. This analysis proceeded under the assumption that the region of interest contained two populations of voxels: those with no experimental effect (the t values of which would be centered about zero) and those with a positive experimental effect (distributed about some nonzero t value). For this example, the noncentral t value calculated was estimated to be 4.85.

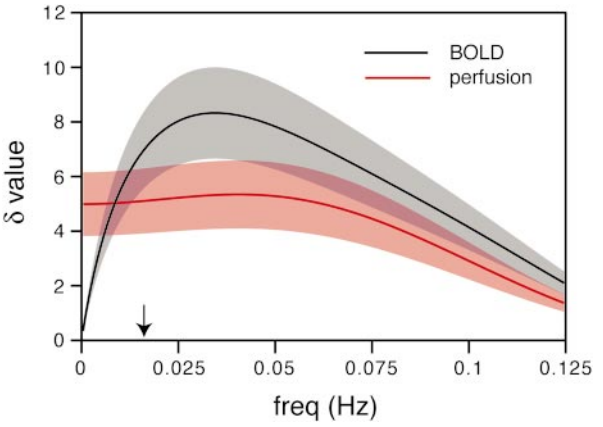


FIG. 7. Relative signal:noise (δ) values predicted for BOLD and perfusion as a function of frequency of an experimental design. These values were calculated as described in the text using the signal:noise values observed for the BOLD and perfusion data for our "reference" experimental design (31-s periods of flashing lights every 62 s), combined with estimations of the intrinsic temporal autocorrelation present in BOLD and perfusion data (Fig. 2) and the respective impulse response functions (Fig. 5). The shaded areas about each line indicate ±2 standard errors (calculated across subjects). The arrow indicates a frequency of 0.016 Hz, corresponding to the reference experimental design used here. The point of crossover occurs at 0.0089 Hz, corresponding to the fundamental frequency of a "blocked" experimental design with approximately 60-s epochs (i.e., 60 s of lights, 60 s of darkness, etc.).

that perfusion imaging provides results that more closely reflect the spatial distribution of changes in neural activity, as opposed to the "draining veins" believed to contribute to the spatial extent of BOLD activation (Duong *et al.*, 2000; Kim *et al.*, 2000; Luh *et al.*, 2000).

We have seen here that perfusion as applied to functional imaging provides advantages over BOLD in some circumstances. First, ASL perfusion methods generate data that are statistically independent over time under the null hypothesis. This implies that perfusion can be used to test hypotheses that address themselves to slow changes in neural activity, perhaps even changes that evolve over days, with reasonable sensitivity. Second, while BOLD provides superior statistical power for most experimental designs when conducted within subjects, there is evidence that perfusion methods are nonetheless more sensitive for hypotheses tested across populations of subjects.

ACKNOWLEDGMENTS

The authors thank Daniel Casasanto, Faez Siddiqi, and Irene Kan for their help with data acquisition. This work was supported in part by Grant HD39621.

REFERENCES

Aguirre, G. K., and D'Esposito, M. 1999. Experimental design for brain fMRI. In *Functional MRI* (P. A. Bandettini and C. Moonen, Eds.), pp. 369–380. Springer-Verlag, Berlin.

Aguirre, G. K., Zarahn, E., and D'Esposito, M. 1997. Empirical analyses of BOLD fMRI statistics. II. Spatially smoothed data collected under null-hypothesis and experimental conditions. *NeuroImage* **5:** 199–212.

Aguirre, G. K., Zarahn, E., and D'Esposito, M. 1998a. A critique of the use of the Kolmogorov–Smirnov (KS) statistic for the analysis of BOLD fMRI data. *Magn. Reson. Med.* **39:** 500–505.

Aguirre, G. K., Zarahn, E., and D'Esposito, M. 1998b. The variability of human BOLD hemodynamic responses. *NeuroImage* **8:** 360–369.

Alsop, D. C., and Detre, J. A. 1996. Reduced transit-time sensitivity in noninvasive magnetic resonance imaging of human cerebral blood flow. *J. Cereb. Blood Flow Metab.* **16:** 1236–1249.

Alsop, D. C., and Detre, J. A. 1998. Multisection cerebral blood flow MR imaging with continuous arterial spin labeling. *Radiology* **208:** 410–416.

Chalela, J. A., Alsop, D. C., Gonzalez-Atavales, J. B., Maldjian, J. A., Kasner, S. E., and Detre, J. A. 2000. Magnetic resonance perfusion imaging in acute ischemic stroke using continuous arterial spin labeling. *Stroke* **31:** 680–687.

Detre, J. A., and Alsop, D. C. 1999. Perfusion fMRI with arterial spin labeling. In *Functional MRI* (P. A. Bandettini and C. Moonen, Eds.), pp. 47–62. Springer-Verlag, Berlin.

Dixon, W. T., Du, L. N., Faul, D. D., Gado, M., and Rossnick, S. 1986. Projection angiograms of blood labeled by adiabatic fast passage. *Magn. Reson. Med.* **3:** 454–462.

Duong, T. Q., Kim, D. S., Ugurbil, K., and Kim, S. G. 2000. Spatiotemporal dynamics of the BOLD fMRI signals: Toward mapping submillimeter cortical columns using the early negative response. *Magn. Reson. Med.* **44:** 231–242.

Friston, K., Ashburner, J., Frith, C., Poline, J.-B., Heather, J., and Frackowiak, R. 1995. Spatial registration and normalization of images. *Hum. Brain Mapp.* **2:** 165–189.

Friston, K. J., Holmes, A. P., Price, C. J., Buchel, C., and Worsley, K. J. 1999. Multisubject fMRI studies and conjunction analyses. *NeuroImage* **10:** 385–396.

Friston, K. J., Josephs, O., Zarahn, E., Holmes, A. P., Rouquette, S., and Poline, J. 2000. To smooth or not to smooth? *NeuroImage* **12:** 196–208.

Gonzalez-At, J. B., Alsop, D. C., and Detre, J. A. 2000. Cerebral perfusion and arterial transit time changes during task activation determined with continuous arterial spin labeling. *Magn. Reson. Med.* **43:** 739–746.

Josephs, O., Rees, G., Turner, R., and Friston, K. J. 1997. Event-related fMRI. *NeuroImage* **5:** S483.

Kim, D. S., Duong, T. Q., and Kim, S. G. 2000. High-resolution mapping of iso-orientation columns by fMRI. *Nat. Neurosci.* **3:** 164–169.

Luh, W. M., Wong, E. C., Bandettini, P. A., Ward, B. D., and Hyde, J. S. 2000. Comparison of simultaneously measured perfusion and BOLD signal increases during brain activation with T-1-based tissue identification. *Magn. Reson. Med.* **44:** 137–143.

Oppenheim, A. V., Willsky, A. S., and Young, I. T. 1983. *Signals and Systems*. Prentice Hall, Englewood Cliffs, NJ.

Purdon, P. L., and Weisskoff, R. M. 1998. Effect of temporal autocorrelation due to physiological noise and stimulus paradigm on voxel-level false-positive rates in fMRI. *Hum. Brain Mapp.* **6:** 239–49.

Shadmehr, R., and Holcomb, H. H. 1997. Neural correlates of motor memory consolidation. *Science* **277:** 821–825.

Wong, E. C., Buxton, R. B., and Frank, L. R. 1997. Implementation of quantitative perfusion imaging techniques for functional brain mapping using pulsed arterial spin labeling. *NMR Biomed.* **10:** 237–49.

Worsley, K. J., and Friston, K. J. 1995. Analysis of fMRI time-series revisited—Again. *NeuroImage* **2:** 173–182.

Yang, Y., Engelien, W., Pan, H., Xu, S., Silbersweig, D. A., and Stern, E. 2000. A CBF-based event-related brain activation paradigm: Characterization of impulse-response function and comparison to BOLD. *NeuroImage* **12:** 287–297.

Ye, F. Q., Berman, K. F., Ellmore, T., Esposito, G., and van Horn, J. D. 2000. $H^{2}{}_{15}O$ PET validation of steady-state arterial spin tagging cerebral blood flow measurements in humans. *Magn. Reson. Med.* **44:** 450–456.

Zarahn, E., Aguirre, G. K., and D'Esposito, M. 1997. Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions. *NeuroImage* **5:** 179–197.

Zarahn, E., and Slifstein, M. 2001. A reference effect approach for power analysis in fMRI. *NeuroImage* **14:** 768–779.