

# Quantitative assessment of the reproducibility of functional activation measured with BOLD and MR perfusion imaging: Implications for clinical trial design

Teddy Tjandra,<sup>a,b</sup> Jonathan C.W. Brooks,<sup>a,b</sup> Patricia Figueiredo,<sup>c</sup> Richard Wise,<sup>a,b</sup>  
Paul M. Matthews,<sup>a</sup> and Irene Tracey<sup>a,b,\*</sup>

<sup>a</sup>Oxford Centre for Functional Magnetic Resonance Imaging of the Brain, Department of Clinical Neurology, University of Oxford, England, UK

<sup>b</sup>Department of Human Anatomy and Genetics, University of Oxford, England, UK

<sup>c</sup>IBILI at Faculty of Medicine, University of Coimbra, Portugal

Received 20 October 2004; revised 4 March 2005; accepted 5 April 2005

Available online 25 May 2005

**BOLD contrast is the most commonly used functional MRI method for studies of brain activity. However, the underlying physiological processes giving rise to measured BOLD signal changes (which include contribution from changes in cerebral blood flow (CBF), cerebral blood volume (CBV) and cerebral metabolic rate of oxygen consumption (CMRO<sub>2</sub>)) vary substantially between sessions and subjects. To determine whether direct CBF measurement is a more reliable technique, we compared the localisation of activation and reproducibility of relative signal change measured by optimised BOLD versus CBF measured using the arterial spin labelling (ASL) technique. Data were collected within the primary sensorimotor cortex in normal healthy controls performing a simple finger-tapping task over three imaging sessions (two on same day and one on a different day). The displacement between the foci of BOLD and CBF activation was less than the linear dimension of one voxel (2.4 mm), however, BOLD activation was significantly closer to the nearest draining vein compared to CBF activation ( $P = 0.030$ ). For the relative signal change measurement, we found that CBF has a lower inter-subject variation than BOLD ( $P < 0.05$ ), enabling a smaller sample size for any given effect size, although the intra-subject variation across sessions for CBF was not significantly different from BOLD. BOLD imaging provides the optimal contrast for exploratory brain activation mapping, however, for a single time-point group study, CBF has reduced variance. In addition, the reduction of variance over time using CBF measurements (non-significant) suggests it could potentially provide a more useful approach when assessing longitudinal activation changes.**

© 2005 Elsevier Inc. All rights reserved.

*Keywords:* FMRI; BOLD; Perfusion; Clinical trials; Brain

---

\* Corresponding author. Department of Human Anatomy and Genetics, Pain Imaging Neuroscience Group, South Parks Road, Oxford University, Oxford OX1 3QX, England, UK. Fax: +44 1865 282675.

*E-mail address:* irene.tracey@anat.ox.ac.uk (I. Tracey).

Available online on ScienceDirect ([www.sciencedirect.com](http://www.sciencedirect.com)).

## Introduction

Functional magnetic resonance imaging (fMRI) has been used to detect human brain activity changes associated with motor, sensory or cognitive processes. Since its first application over 10 years ago to detect human brain areas processing vision, it has been applied widely to study normal human brain function (Bandettini et al., 1992; Belliveau et al., 1991; Kwong et al., 1992; Ogawa et al., 1993). There has been a growing interest in using fMRI for clinical studies of the brain (see Matthews and Jezzard, 2004 for a review) and, in particular, for longitudinal studies of patients to monitor either disease progression or therapeutic intervention (Johansen-Berg et al., 2002; Kumari et al., 2002). The combination of fMRI with administration of central nervous system (CNS) acting drugs provides a powerful opportunity to determine site and efficacy of drug action in the human brain (Rogers et al., 2004; Tracey, 2001; Wise et al., 2002, 2004).

Despite its popularity, there are potential problems with the commonly applied BOLD functional imaging in these applications. First, the signal change can vary substantially between subjects and across sessions (Aguirre et al., 1998; McGonigle et al., 2000). This variability reduces the statistical power in longitudinal design. Secondly, in many studies, it has been noted that the signal within a brain active region varies widely between individuals for reasons that are poorly understood (Miezin et al., 2000; Rajapakse et al., 1998), which again limits the applicability of this method in patient and drug studies. A third problem with fMRI is that the exact location of signal change may not define exactly where the presynaptic neuronal activity is found (Logothetis et al., 2001). This is because a joint measure of changes associated with the cerebral blood flow (CBF), cerebral blood volume (CBV) and the cerebral metabolic rate of oxygen consumption (CMRO<sub>2</sub>) is being assumed to produce contrast in BOLD fMRI. The complex interactions of these underlying mechanisms make changes in the BOLD signal difficult to interpret.

Direct CBF measurements using MRI provide an attractive alternative. Recent studies have compared BOLD with CBF measurement using the arterial spin labelling perfusion imaging technique (Aguirre et al., 2002; Luh et al., 2000; Wang et al., 2003a,b; Yang et al., 1998). Such reports suggest that CBF measurement provides improved sensitivity relative to BOLD particularly for low-frequency tasks and is associated with lower inter-subject variability (Aguirre et al., 2002; Wang et al., 2003a). However, such studies (with the exception of Wang et al., 2003a) were performed using non-optimised BOLD acquired simultaneously with the ASL technique.

Here, we investigate further both the intra- and inter-subject group variability of the two modalities for a simple motor hand-tapping task using three separate analyses. A first analysis explored the location of activation for BOLD and CBF within the primary sensorimotor (PSM) area and compared the Euclidean distance relative to the small draining vein defined using an MR venogram (Reichenbach and Haacke, 2001). Next, we assessed the reproducibility of the intra-subject variation of the relative signal change arising from motor activity measured by BOLD and CBF across three sessions (two same day, one different day) within three separate pre-defined regions of interest: the entire PSM and separate anatomically defined primary motor (M1) and primary sensory (S1) cortices. Within these regions, we also measured the inter-subject variation of the relative signal change for the group between BOLD and CBF. Finally, we investigated the reliability of each modality based on a retrospective power calculation for the sample size required to produce a given statistical significance. This work extends the current understanding of the variability of the signal change measured by BOLD and CBF and allows more accurate estimation of the sample sizes required (Fig. 1).

## Materials and methods

### Subjects

6 normal right-handed subjects (age,  $26.2 \pm 4.2$ , three male) were recruited. They performed a visually cued four-finger-tapping experiment using the right hand in a block-design paradigm consisting of 8 cycles of 20 s OFF (resting) and 32 s ON (active) periods at a frequency of 2 Hz. Each subject had their right four fingers taped together for one synchronous hand-tapping movement whilst maintaining the thumb and wrist stationary with a hand brace. An in-house built wooden brace was designed with a horizontal wooden bar pegged across the edges at the back of the device, 3 cm high from the base. In use, the palm of the hand rested on the base, and the fingers touched the wooden bar for each tap. This kept the amplitude and angle of the tapping motion from the base constant for all subjects in each scan. Each subject was told to minimise any limb movement. Head motion was minimised through the use of restraints such as padding on both sides of the temples with tape across the forehead. Each subject repeated the experiment three times: two sessions performed within the same day separated by a 30-min break outside the scanner and the third on a separate day 1 week later. All subjects were requested to refrain from caffeinated drinks 3 h before each study, as caffeine could cause an overall increase in the BOLD contrast during activation (Mulderink et al., 2002).

### MR scanning

Imaging was performed on a 3 T Varian/Siemens MRI scanner. BOLD fMRI was acquired first using a single-shot gradient-echo (GRE) echo-planar imaging (EPI) sequence (208 volumes plus 4

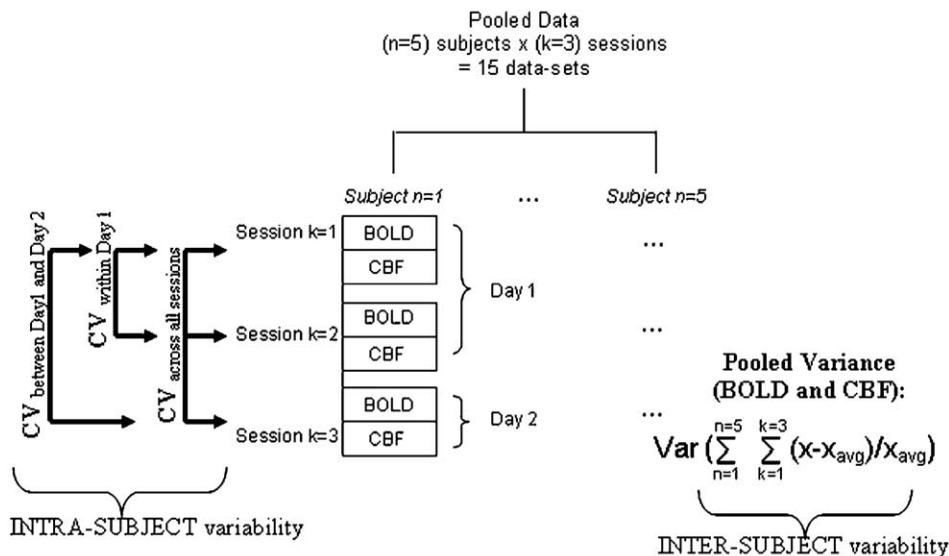


Fig. 1. A schematic diagram for the assessment of the reproducibility of the relative signal change measured by BOLD and CBF modalities in a motor hand-tapping experiment for  $n = 5$  in  $k = 3$  sessions (two within same day, one on a different day). The intra-subject variability for 'same-day,' 'between-day' and 'across all sessions' experimental comparisons was each measured as the coefficient of variation (CV) of the relative signal change: (i) between sessions #1 and #2 on Day 1, (ii) between sessions #1 and #3 and (iii) all three sessions, respectively. Note that in our experiment, Day 1 and Day 2 is separated by a one-week interval. For the inter-subject variability analysis, the pooled data was collated (5 subjects  $\times$  3 sessions = 15 data-sets). The relative signal change for each session was normalised to the subject's average value of the relative signal change across all three sessions. The variance of these normalised relative signal changes for the pooled data was calculated for both BOLD and CBF and their results compared for significant difference by using an F-test.

dummy scans, TR/TE = 2000/30 ms to optimise signal-to-noise ratio (SNR), flip angle = 75°, 21 contiguous axial slices of 6 mm thickness parallel to the anterior–posterior commissure (AC–PC), covering the whole brain from the base of the cerebellum to the vertex, FOV = 256 × 256 mm<sup>2</sup>, matrix = 64 × 64). Subsequently and in the same imaging session, perfusion data were acquired using Pulsed Arterial Spin Labelling (PASL) technique with QUIPSS-II (Quantitative Imaging of Perfusion using a Single Subtraction) method (Wong et al., 1997, 1998a,b). An adiabatic inversion pulse was employed for labelling followed by optimised inversion time delays  $TI_1 = 700$  ms and  $TI_2 = 1500$  ms, chosen so as to minimise the flow-through effect whilst allowing sufficient SNR (Luh et al., 2000; Wong et al., 1998a,b). Interleaved tag and control images were acquired using a GRE-EPI sequence, with acquisition parameters: TR/TE = 2000/20 ms and flip angle = 75°. The imaging region consisted of 5 contiguous axial slices of 6 mm thickness acquired parallel to the AC–PC line from the vertex of the brain. This encompassed the primary motor area based on an anatomical observation from the ‘omega’ landmark on the axial plane and the ‘hook’ shape identified on the sagittal view (Yousry et al., 1997). The imaging slab with FOV = 256 × 256 mm<sup>2</sup> and matrix = 64 × 64 was separated by a 15-mm spacing from the 100-mm-thick inversion slab. A 5-slice volume of the equilibrium brain tissue magnetisation,  $M_0$ , was acquired to normalise the difference perfusion maps using the same sequence parameters as mentioned above for QUIPSS-II. A 21-slice volume of the equilibrium magnetisation of the whole brain was acquired as a reference image for subsequent registration from 5 slices to 21 slices.

During one of the three separate imaging sessions, a high-resolution whole brain structural scan was acquired for each subject using a T1-weighted sequence inversion recovery (IR) 3D Turbo FLASH (TR/TE = 30/5 ms, TI = 500 ms, flip angle = 20°, 64 × 3 mm axial slices parallel to the AC–PC plane, FOV = 256 × 256 mm<sup>2</sup>, matrix = 256 × 256). In addition, an MR venogram was acquired to define the location of small veins in the sub-millimetre range within the region of interest for each subject using a T2\*-weighted Turbo Echo sequence. This technique utilises the differences in magnetic susceptibility between deoxyhaemoglobin in venous blood (as an intrinsic contrast agent) and the surrounding tissue to generate contrast (Reichenbach and Haacke, 2001). The acquisition parameters were: TR/TE = 40/30 ms, flip angle = 30°, 64 slices of 2 mm thickness covering the whole brain from the base to the vertex parallel to the AC–PC line, FOV = 256 × 160 mm<sup>2</sup>, matrix = 512 × 384 for a 0.5 × 0.4 × 2 mm<sup>3</sup> resolution.

#### Image analysis

Analysis was carried out using FEAT (FMRI Expert Analysis Tool) Version 5.1, part of FSL (FMRIB’s Software Library, <http://www.fmrib.ox.ac.uk/fsl>). The following pre-statistics processing was applied for BOLD FMRI data; motion correction using MCFLIRT (Jenkinson and Smith, 2002); spatial smoothing using an “optimum” Gaussian kernel of full width half maximum (FWHM) 5 mm (Ye et al., 1997); mean-based intensity normalisation of all volumes by the same factor; high-pass temporal filtering (Gaussian-weighted LSF straight line fitting, with  $\sigma = 39.0$  s). Time-series statistical analysis was carried out using FILM (FMRIB’s Improved Linear Model) with local autocorrelation correction (Woolrich et al., 2001). For intra-subject analysis,

the  $Z$  (Gaussianised T/F) statistic images for each subject were thresholded using clusters determined by  $Z > 2.3$  and a (corrected) cluster significance threshold of  $P < 0.01$  (Friston et al., 1994; Forman et al., 1995; Worsley et al., 1992) with a requirement of at least four contiguous voxels. Random effects group analyses for BOLD and CBF for the group  $Z$  statistic images were thresholded using  $Z > 1.8$  and a cluster significance threshold of  $P < 0.05$ , corrected for multiple comparisons (Forman et al., 1995), to take into account the low SNR in CBF and a small sample size. Registration to individual high-resolution structural images and the MNI (Montreal Neurological Institute) standard brain (Collins et al., 1994) was carried out using FLIRT (FMRIB’s Linear Image Registration Tool) (Jenkinson and Smith, 2002) using 9 and 12 degrees of freedom (DOF) respectively.

Perfusion maps were calculated by pair-wise subtraction of the time-matched *tag* and *control* images, normalised by the equilibrium brain tissue magnetisation, as described by Wong et al. (1997). The pre-statistics processing and the post-processing statistical significance based on clustering applied to perfusion images were the same as BOLD for both intra-subject and within-group analyses (except for no prewhitening correction of CBF data since there is no significant temporal autocorrelation compared to BOLD data (Wang et al., 2003b)). The registration from 5 slices to 21 slices (i.e. CBF to BOLD) was performed with FLIRT using 3 DOF before further transformation to the high-resolution structural and MNI standard brain using 12 DOF.

#### Data analysis

Regions of interest (ROIs) were defined within the corresponding (i.e. contralateral) hemisphere of each individual’s high-resolution T1-weighted structural image using the anatomical landmark (omega shape) in axial plane (Yousry et al., 1997). Anatomical masks for three regions of interest (ROI) within each subject were defined as: (a) the primary motor cortex (M1; the posterior half of the precentral gyrus, corresponding to Brodmann area 4), (b) the primary sensory cortex (S1; the postcentral gyrus, corresponding to Brodmann areas 3, 1 and 2) and (c) the primary sensorimotor cortex (PSM), which is the sum of both M1 and S1. For each subject, the ROI extended on the lateral surface of the cortex to the midline and from the vertex of the brain to the level of the superior aspect of the lateral ventricles caudally. The PSM mask excludes the gap of the central sulcus separating the anterior bank of S1 and the posterior bank of M1.

Each anatomical mask was used to calculate the mean signal change for the hand-tapping activation. For calculation of the location of activation, only the PSM mask was used. For this, we proposed using the centre of gravity (COG) measurement technique which is the weighted (by signal change) location of the centre of activation within a cluster volume of neighbouring voxels that are significantly activated.

Group analysis was conducted for all subjects in each session for each modality to find the degree of overlap between sessions and to determine the extent of the activation within primary sensorimotor cortex (PSM). The MNI coordinates were noted for the PSM area and were compared as a reference for intra-subject analyses for the location of activation. In addition, a frequency map was computed: this was defined as the overlay of group activation in which the thresholded  $Z$  score map from each session was binarised and superimposed in order to illustrate the common area

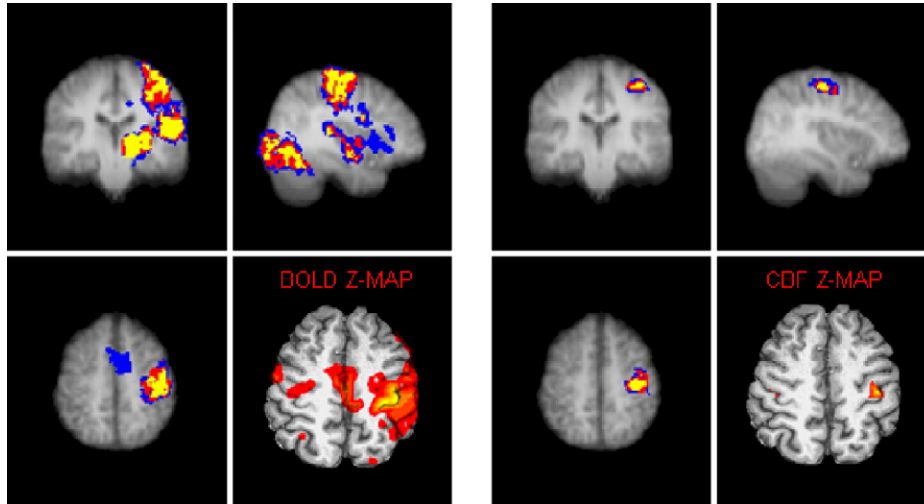


Fig. 2. The frequency map of the group thresholded Z-score image for all three sessions were superimposed on one another for BOLD (left column) and CBF (right column) in axial, coronal and sagittal views. Each frequency map is colour coded using blue, yellow and red to represent the overlap of one, two and three sessions to show the difference in the spatial extent of the functional activation for BOLD and CBF for the primary sensorimotor cortex with respect to different sessions. A Z-score image of the BOLD and CBF activation in one subject is also shown on the bottom right corner respectively.

of overlap and to qualitatively compare the extent of activation between BOLD and CBF (see Fig. 2).

#### Testing for the variability of the location of activation

To test if BOLD and CBF co-localise, the displacement of the location of activation from CBF to BOLD for each session in each subject was calculated in terms of the Cartesian  $X$ - $Y$ - $Z$  coordinates from the COG measurement within each anatomical mask. The mean displacement across the group was calculated. A draining vein was used as a second reference point to measure the relative displacement and distance for each modality. The rationale was to determine whether BOLD activation is closer to large vessels compared to CBF activation. Each draining vein was defined for each subject as the estimate of the nearest vein to the motor hand area (see Fig. 3). The displacement with respect to the draining vein was calculated for each subject in all three sessions for each coordinate. The group mean Euclidean distance for BOLD and CBF was compared using a two-tailed paired  $t$  test.

#### The reproducibility of the signal change within each ROI

Fig. 2 shows a schematic diagram for the assessment of the reproducibility of the signal change measurement for BOLD and CBF. The mean signal changes over all sessions were determined for both BOLD and CBF in each ROI. The variance from the pooled data ( $5 \text{ subjects} \times 3 \text{ sessions} = 15 \text{ data-sets}$ ) for each modality was calculated ( $\text{Var} = \sum(x_i - x_{\text{avg}})^2 / (N - 1)$ ) in each ROI. This variance for BOLD and CBF measurements was compared using an  $F$ -to- $p$  ratio of variance test ( $F = \text{larger Var} / \text{smaller Var}$ ) to test for any significant difference in the inter-subject variability of the measured relative signal change. For the purpose of normalisation in this pooled variance calculation, the signal change measured from each session for each subject was subtracted followed by division from the respective individual's average signal change across sessions (see Fig. 1). This allows adjustment for the difference in magnitude between BOLD (in the order of 1–2%) and CBF (in the order of 30–40%) signal change measurements.

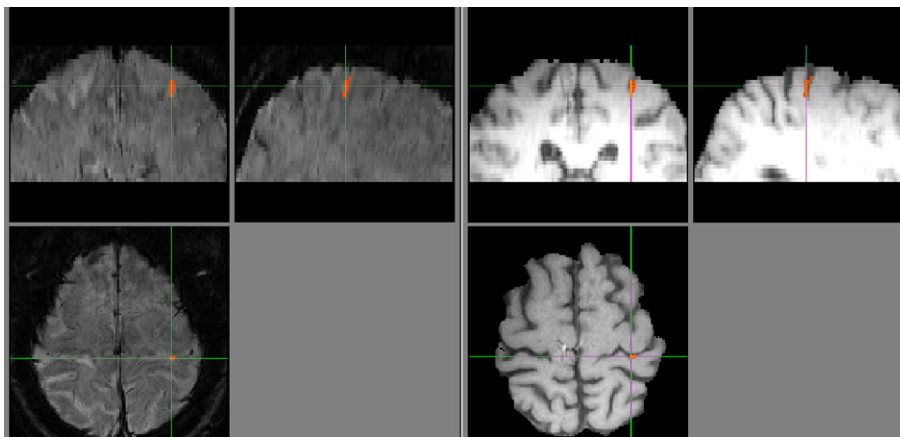


Fig. 3. The small draining vein (red) defined for one subject and overlaid onto the subject's venogram map (left) compared to the overlay on the anatomical structural (right). The course of the small vessel was observed from the venogram map over 5 slices axially, coronally and sagittally. The point of draining vein closest to the motor hand area is chosen as the reference point (green cross-lines).

The coefficient of variation (CV) was used to estimate the intra-subject reproducibility for each ROI in three experimental comparisons: (1) all three sessions, (2) two sessions from the same day and (3) two sessions between different days. For the latter between-day comparison, we used the first session on each day. Each subject’s CV was defined as the standard deviation of the task-related signal change across sessions divided by the mean of the task-related signal change across sessions for each comparison in each ROI.

$$CV = (\sigma_i/x_{avg}) * (100\%), \text{ where } \hat{\sigma}_i = \sqrt{\sum \frac{(x_i - \bar{x})^2}{n - 1}}$$

and  $n$  = number of sessions compared

The group means for this intra-subject CV measurement were calculated for each experimental comparison, and the results for BOLD and CBF were compared with two-tailed paired  $t$  tests for each ROI to test for any significant difference in the intra-subject CV across the group. The standard deviation of this CV measurement for BOLD and CBF was calculated. These values were compared between the two modalities for each of the above-mentioned experimental comparisons using an  $F$  test. This tested for any significant difference in the variability of the intra-subject variation within the study group.

In addition, to further interrogate the source of variation and assess whether common components contribute to the variance in BOLD and CBF, the pooled data for the signal change measurement for BOLD were tested for any significant correlation with that for CBF within each ROI using a Pearson correlation.

*Testing for the required sample size for different effect sizes*

To assess the utility of both modalities, a retrospective power calculation was done to compare the number of subjects needed to detect statistically significant 10%, 20% and 30% signal difference or experimental “effect size” (null hypothesis was zero activation) using the inter-subject variance across all sessions within the PSM. The sample size was estimated using MacANOVA software

for a single-sample two-tailed  $t$  test (<http://www.stat.umn.edu/macanova>).

For all tests described above, significance was accepted if  $P < 0.05$ .

**Results**

Full data-sets obtained from 5 subjects are reported. The sixth subject was considered an outlier since the signal change measured for BOLD was at least 2 standard deviations smaller than the group average for two out of three sessions.

*Variability in the localisation of activation changes*

The group mean vector averaged across sessions showed only a small and non-significant displacement between CBF and BOLD measurements:  $X = -2.2 \pm 3.2$  mm;  $Y = 0.4 \pm 1.8$  mm;  $Z = -0.1 \pm 4.5$  mm (Table 1). However, the group mean Euclidean distance to the nearest draining vein for BOLD ( $7.8 \pm 2.3$  mm) was significantly smaller compared to CBF ( $10.2 \pm 3.5$  mm) ( $t = -2.409$ ,  $P = 0.030$ , two-tailed) (Fig. 4).

*Reproducibility of signal change*

Table 2 summarises the results of the group average of the pooled signal change measurement for all subjects and all sessions, the pooled variance of this measurement and the group mean of intra-subject CV across sessions and day difference for BOLD and CBF in three regions of interest (ROI). Fig. 5 illustrates the mean signal change across sessions for each subject in each ROI, whilst Fig. 6 illustrates the group mean CV and the standard deviation for comparisons across all-sessions, within the same day and between days in each ROI.

The group mean signal change in the primary sensorimotor cortex (PSM) for BOLD was  $1.23 \pm 0.34\%$  whilst that for CBF was  $40.55 \pm 5.24\%$ . There was a significantly lower pooled variance for CBF measurement (0.012) than for BOLD measurement (0.032)

Table 1  
Results for the analysis of the location of activation for BOLD and CBF measured by the centre of gravity (COG)

Subject	Draining vein coordinates (X, Y, Z)			Average BOLD location across sessions* (X, Y, Z)			Average CBF location across sessions* (X, Y, Z)			With respect to the draining vein							
										BOLD displacement* (X, Y, Z)		CBF displacement* (X, Y, Z)		BOLD distance		CBF distance	
1	-42.7	-27.9	58.3	-43.0	-23.1	55.1	-40.9	-24.7	54.3	-0.3	4.8	-3.2	1.8	3.2	-4.0	6.0	6.2
2	-29.2	-14.4	61.0	-38.1	-11.9	57.3	-40.3	-13.5	50.4	-8.9	2.5	-3.7	-11.1	0.9	-10.6	10.0	15.6
3	-31.1	-36.2	64.3	-40.1	-32.7	60.5	-37.8	-33.4	60.7	-9.0	3.5	-3.8	-6.7	2.8	-3.6	10.5	8.3
4	-43.8	-23.9	54.6	-42.9	-28.7	56.4	-36.1	-26.0	61.7	0.9	-4.8	1.8	7.7	-2.1	7.1	5.2	11.2
5	-38.9	-20.6	49.7	-40.2	-18.9	56.8	-38.3	-19.4	59.2	-1.3	1.7	7.1	0.6	1.2	9.5	7.4	10.0
Group mean (SD)	-37.1 (6.7)	-24.6 (8.2)	57.6 (5.7)	-40.9 (2.1)	-23.0 (8.2)	57.2 (2.0)	-38.7 (2.0)	-23.4 (7.4)	57.3 (4.8)	-3.7 (4.8)	1.6 (3.7)	-0.4 (4.8)	-1.5 (7.4)	1.2 (2.1)	-0.3 (8.4)	7.8 (2.3)	10.2 (3.5)

Data for the location of the draining vein nearest to the primary sensorimotor cortex (PSM) and the respective BOLD and CBF location of activation was shown for each subject in terms of the X,Y, Z coordinates (the first 4 columns). Each draining vein was defined using an MR venogram. The average location of activation for BOLD and CBF shown for each subject is the average across all three sessions and was calculated from the centre of gravity (COG) of a significant cluster within PSM with threshold of  $Z > 2.3$ ,  $P < 0.01$ . The next 4 columns show the displacement of BOLD and CBF location of activation and the respective Euclidean distance, with respect to the draining vein. The group mean  $\pm$  standard deviation (SD) for each column is also shown. All units are in millimeter.

\* Note for simplicity, the standard deviation for each subject across sessions for the COG measurement (the location and the distance) is not shown. For the testing of significant difference in the displacement and distance between BOLD and CBF, the pooled data (i.e. each individual data for all sessions) for the group (not shown here) was used.

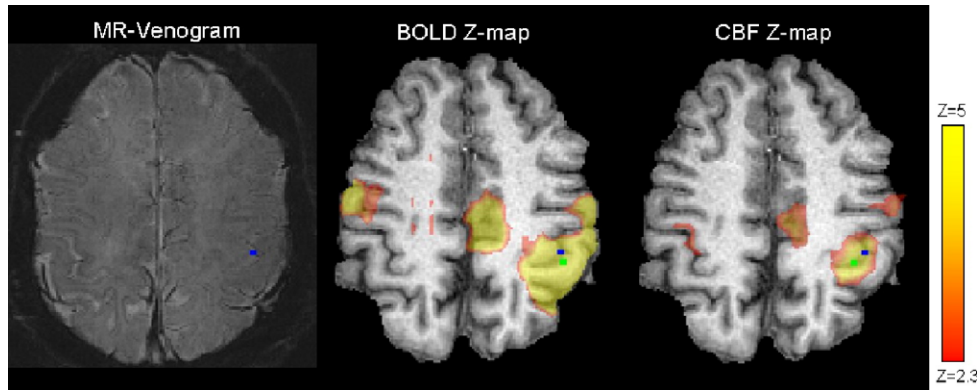


Fig. 4. A plot of the motor-hand tapping activation pattern (Z-score activation map;  $2.3 \leq Z \leq 5$ ) seen in BOLD and CBF modalities for Subject #4 in session #2, overlaid onto the subject's structural image. The respective localisation of BOLD and CBF activation measured by the weighted centre of gravity (COG) metric (see green square) was shown within the primary sensorimotor cortex (PSM) region-of-interest relative to the draining vein (see blue square), as defined by the MR Venogram as a small dark circle closest to the 'omega' shape. This example illustrates the fact that BOLD COG is closer to the draining vein than that for CBF.

( $F_{1,14} = 2.67$  at  $P = 0.05$ ; see Table 2 and Fig. 5). The group mean of the intra-subject CV measurement across all sessions for BOLD was  $17 \pm 14\%$  (range: 2–39%), whilst that for CBF was  $13 \pm 3\%$  (range: 9% to 17%) (see Table 2 and Fig. 6). The standard deviation of this CV measurement for BOLD was significantly greater than for CBF for all experimental comparisons ( $F$  tests,  $P < 0.05$ ) (Fig. 6). The relative signal changes of the pooled data within PSM for BOLD ( $1.23 \pm 0.34\%$ ) showed a trend towards a positive correlation with those for CBF ( $40.55 \pm 5.24\%$ ) ( $r = 0.498$ ,  $P = 0.059$ ).

Considering the sub-regions of PSM separately, the group mean signal change in the primary motor cortex (M1) across all sessions for BOLD was  $1.39 \pm 0.91\%$ , whilst for CBF it was  $42.89 \pm 9.65\%$ . There was a significantly lower pooled variance for CBF measurement (0.015) than for BOLD measurement (0.063) ( $F_{1,14} = 4.20$  at  $P = 0.01$ ; see Table 2 and Fig. 5). The group mean of the intra-subject CV measurement across all sessions for BOLD was  $24 \pm 19\%$  (range: 5–55%), whilst that for CBF was  $14 \pm 6\%$  (range: 4–20%) (see Table 2 and Fig. 6). The standard deviation of this CV measurement for BOLD was significantly greater than CBF for 'across all sessions' experimental comparison ( $F$  test,  $P < 0.05$ ) (see Fig. 6). The relative signal changes of the pooled data within

M1 for BOLD showed a significant positive correlation with those for CBF ( $r = 0.566$ ,  $P = 0.028$ ).

The group mean signal change in the primary somatosensory cortex (S1) across all sessions for BOLD was  $1.19 \pm 0.55\%$ , whilst that for CBF was  $39.52 \pm 4.10\%$ . There was a trend towards a significant difference for the pooled variance between BOLD and CBF measurements; BOLD having greater variability (0.064) than CBF (0.029) ( $F$  test  $P = 0.07$ ; see Table 2 and Fig. 5). The group mean of the intra-subject CV measurement across all sessions for BOLD was  $24 \pm 20\%$  (range: 2–44%), whilst that for CBF was  $16 \pm 5\%$  (range: 10–21%) (see Table 2 and Fig. 6). Within S1, the standard deviation of this CV measurement for BOLD was significantly greater than CBF for all experimental comparisons ( $F$  test  $s$ ,  $P < 0.05$ ) (see Fig. 6). The relative signal changes of the pooled data within S1 for BOLD did not show any significant correlation with that for CBF.

#### Implications for sample size calculation

Table 3 shows the required sample size using the retrospective power calculation to detect different effect sizes and was obtained using the information from the pooled variances of the signal

Table 2  
Results for the analysis of the reproducibility of signal changes for BOLD and CBF

Region of interest	PSM ROI		M1 ROI		S1 ROI	
	BOLD	CBF	BOLD	CBF	BOLD	CBF
A. Group average of the mean signal change $\pm$ SD (%)	$1.23 \pm 0.39$	$40.55 \pm 6.72$	$1.39 \pm 0.91$	$42.89 \pm 10.51$	$1.19 \pm 0.56$	$39.52 \pm 6.88$
B. Pooled variance	$0.032 \pm 0.025$	$0.012 \pm 0.006$	$0.063 \pm 0.035$	$0.015 \pm 0.005$	$0.064 \pm 0.050$	$0.029 \pm 0.022$
C. F ratio: BOLD vs. CBF	2.67*		4.20**		2.21	
D. Mean CV $\pm$ SD (%)						
1). All sessions	$17 \pm 14$	$13 \pm 3$	$24 \pm 19$	$14 \pm 6$	$24 \pm 20$	$16 \pm 5$
2). Same day	$15 \pm 18$	$14 \pm 3$	$18 \pm 13$	$13 \pm 8$	$18 \pm 18$	$14 \pm 4$
3). Different days	$21 \pm 15$	$4 \pm 2$	$20 \pm 20$	$11 \pm 10$	$30 \pm 25$	$10 \pm 5$

Results of the reproducibility analysis for BOLD and CBF measurement in each region of interest (ROI): PSM, M1 and S separately: (A) The group average of the mean signal change (%)  $\pm$  standard deviation (SD) across subjects for three sessions, (B) The pooled variance ( $\pm$  error), (C) The F ratio test of the pooled variance between BOLD and CBF, (D) The mean Coefficient of Variation, CV for the group for each comparison: (1) all three sessions, (2) two sessions within the same day, and (3) two sessions in different days are also shown. Note that \* denotes significance at  $P = 0.05$  while \*\* denotes significance at  $P = 0.01$ .

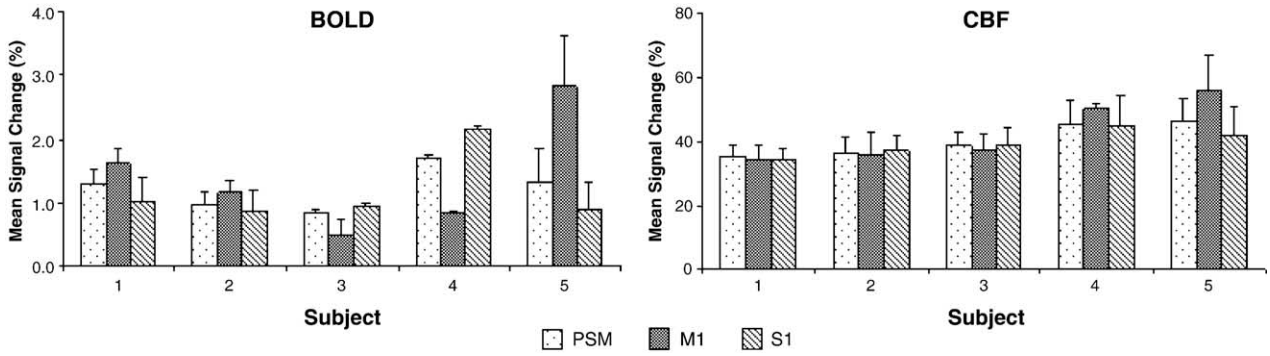


Fig. 5. A plot of the mean signal change (%) average across three session for each subject for BOLD (left) and CBF (right) measurement within three regions of interest: primary sensorimotor (PSM), primary motor (M1) and primary sensory (S1) cortices. Error bars denote one standard deviation.

change measurement in PSM. We estimate that CBF-based fMRI should require fewer subjects than BOLD to detect a similar change over time.

**Discussion**

We have investigated the reproducibility and reliability of the BOLD signal change and CBF measurement during a simple finger movement task. Motor activation was found within contralateral primary sensorimotor cortex (PSM) for all subjects in each session. The group activation map showed a larger area of BOLD activation for all sessions compared to CBF (see Fig. 2), reflecting the superior sensitivity of BOLD compared to ASL. The average common area of overlap for CBF group activation across sessions (~60%) is larger than for BOLD (~40%). As the statistical threshold changes for determining BOLD activation, it is expected that the relative percentage of overlap between CBF and BOLD will change. However, as this is not the main objective of our study, we have not investigated this relationship further.

By focussing on activation in the primary sensorimotor cortex, we have demonstrated that small spatial differences in the activations centres of gravity (COG) may be detected between CBF and BOLD. The first part of our analyses shows only minor spatial differences in the activation measured from the centre of gravity (COG) between CBF and BOLD with activation in the primary sensorimotor cortex. However, we found that BOLD

activation was significantly closer to the draining vein compared to CBF. This supports the notion that CBF may more accurately localise brain activation due to its greater sensitivity to changes in the capillary bed local to the activated neuronal population (Lee et al., 1999; Talagala and Noll, 1998; Wong et al., 1997) than the venule side—as is the case for BOLD contrast.

The inter-subject variability for CBF changes was significantly lower than for BOLD changes, which was reflected in the smaller pooled variance (see Table 2 and Fig. 5). The group mean CV for BOLD intra-subject variability was greater than that for CBF (see Table 2), although this could not be demonstrated significantly for each specific serial quantitative assessments (‘across all three sessions’, ‘two sessions on the same day’ or ‘two sessions on different days’) presumably due to the small sample size. There was a significant degree of variability in this intra-subject variation within the group (see Fig. 5). CBF had significantly lower variability than BOLD for all experimental comparisons in PSM and S1, as well as for ‘across all three sessions’ experimental comparison in the M1 region. This is because of the lower standard deviation for the group mean CV measurements in CBF compared to BOLD (see Table 2 and Fig. 6). BOLD shows particularly greater variability of this intra-subject variation than CBF for ‘sessions on different days’ experimental comparison since the 95% confidence interval for the paired *t* tests has a wider range in each region of interest (data not shown) compared to the sessions ‘on the same day’ and ‘across all three sessions’. These observations have important implications for the design of clinical

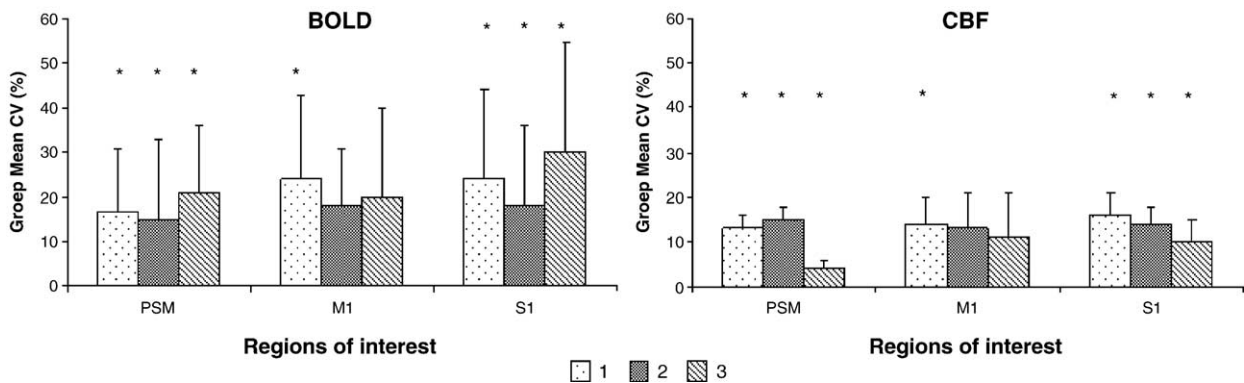


Fig. 6. A plot of the group mean Coefficient of Variation (CV) for each experimental comparison: 1) across all sessions, 2) within the same day and 3) between different days; within three regions of interest; primary sensorimotor (PSM), primary motor (M1) and primary sensory (S1) cortices for BOLD (left) and CBF (right). Error bars denote one standard deviation of the group mean CV. \* denotes a significant difference between the standard deviation of the group mean CV for a given comparison between BOLD and CBF in a given region of interest (for clarity, significant differences are shown on both BOLD and CBF data).

Table 3

Results for the calculation of the sample size ( $\pm$  error) for BOLD and CBF measurements for each given effect size

Modality	Pooled variance	Sample size for an effect size		
		10%	20%	30%
BOLD	0.032	28 $\pm$ 22	9 $\pm$ 8	6 $\pm$ 5
CBF	0.012	12 $\pm$ 7	5 $\pm$ 4	4 $\pm$ 3

A summary of the results for the sample size calculation ( $\pm$  error) for BOLD and CBF with effect sizes 10%, 20% and 30%, standardised using the pooled variance of the signal change measurement across all three sessions and five subjects within the primary sensorimotor cortex. The test for each modality is for a single-sample two-tailed test with  $\alpha = 0.05$  and power = 0.80.

research in which serial observations are to be compared. It is apparent from the power calculation in Table 3 that CBF allows more efficient trial design than BOLD.

Our results extend the observation (Aguirre et al., 2002; Wang et al., 2003b) showing that CBF gives significantly lower inter-subject variability (see Fig. 6). In addition, we demonstrated a trend for CBF to have lower intra-subject variability for sessions performed on the same day, different days and across all sessions. CBF measurement of the relative signal changes is also more consistent (reproducible) across different subcortical areas of motor hand area; the primary sensorimotor cortex and the primary sensory cortex, with a trend towards significance for the primary motor cortex. The greater inter-subject variability of BOLD measurements across sessions could be a consequence of multiple factors (e.g. CBF, CBV and CMRO<sub>2</sub>) contributing to its changes. This also includes differences in biological causes such as hormonal changes, arterial concentration of CO<sub>2</sub> and haematocrit. The reduced intra-subject variability in CBF compared to BOLD could also be due to the inherent subtraction of label and control images in the ASL measurement which gives a more stable baseline than BOLD.

Our results for the correlation analyses show that common components of BOLD and CBF measurement contribute to the variances determined. However, differences in the BOLD and CBF correlation measurements were found between M1 and S1. This could be due to the differences in the underlying neuronal activation, and interestingly a recent study examining differences between the haemodynamic response for BOLD and CBF across M1 and supplementary motor area (SMA) has also shown correlation differences between these regions (Obata et al., 2004). Obata and colleagues found that the average BOLD response from the beginning to the end of the motor task was different between M1 and SMA, whilst CBF responses in the two regions were similar.

Despite this, the evidence for greater accuracy of localisation with CBF is less conclusive. Firstly, the group standard deviation for the CBF Euclidean distance relative to the draining vein is greater than BOLD (Table 1). This is likely due to the relatively low SNR of our CBF measurements. Secondly, the difference in the group mean Euclidean distance ( $t = 2.4$  mm) between the two modalities is within the inherent resolution of collected data ( $4 \times 4 \times 6$  mm<sup>3</sup>) and the high-resolution structural space to which it was registered ( $1 \times 1 \times 3$  mm<sup>3</sup>). The difference could be due at least in part to registration problems during transformation from functional to structural space. However, this difference in the distance measured for BOLD and CBF relative to the draining vein is within 5 voxels of the inherent resolution of the venogram ( $0.5 \times 0.4 \times 2$

mm<sup>3</sup>) from which the draining vein for each subject was depicted. This finding could partially be explained by the fact that we used a gradient-echo (GRE) EPI sequence which is known to suffer from an increased sensitivity to venous contributions (Duong et al., 2003; Zhao et al., 2004). Our reason for choosing this sequence, which potentially biases our result, is because it has a higher SNR as compared to a spin-echo (SE) EPI sequence.

Nevertheless, it still cannot be inferred from our data that CBF is potentially more co-localised to neuronal activation, as activation may be shifted towards the feeding artery rather than the capillary site, although there is no literature to support that this occurs. It is known, however, that venous drainage can extend as far as a few centimetres away from the capillary, therefore a potentially larger concern for BOLD than CBF. A potential direction to explore this disparity in displacement would be to use a more precise electrophysiological method as an independent measure of neuronal activity for comparison between BOLD and CBF localisation of activation. A study performed by Disbrow et al. (2001) had found that BOLD activation was weighted towards the draining vein within the central sulcus of the motor area in monkeys, compared to the recorded electrical potentials, which presumably reflect the site of neuronal activity. In addition, other work has shown that CBF is perhaps a better prescriptor of the neuronal activity (Lee et al., 1999). Studies on spatial specificity at sub-millimetre resolution have shown that perfusion-based fMRI can be used to map individual functional columns provided that large-vessel contaminations are minimised or eliminated (Duong et al., 2001). Further investigations are clearly needed to prove unequivocally that CBF provides a more reliable marker for the location of brain activation such that one could use it for determining healthy tissue margins during pre-surgical planning for brain tumour resection.

In conclusion, although BOLD imaging gives a high signal-to-noise ratio (SNR) for studies of activation, perfusion (CBF) measurements in PSM show significantly lower inter-subject variability. This enables a smaller sample size for CBF-based studies compared to BOLD studies to detect any given effect size. CBF also has a trend for lower intra-subject variability when comparing sessions within the same day or between different days. This finding suggests that CBF may provide a better outcome measure than BOLD for studies demanding serial observations of magnitude of change, for instance, in assessment of a drug intervention in clinical trials.

## Acknowledgments

The authors are grateful to Peter Hobden, Stuart Clare and Peter Jezzard for their help with the experiments. This work was supported by the McDonnell-Pew Scholarship (TT), Wellcome Trust (RGW), Dr. Hadwen Trust for Humane Research (JCWB), Medical Research Council (PMM) and Higher Education Funding Council for England (IT).

## References

- Aguirre, G.K., Zarahn, E., D'Esposito, M., 1998. The variability of human BOLD hemodynamic responses. *NeuroImage* 8, 360–369.
- Aguirre, G.K., Detre, J.A., Zarahn, E., Alsop, D.C., 2002. Experimental design and the relative sensitivity of BOLD and perfusion fMRI. *NeuroImage* 15, 488–500.



- Bandettini, P.A., Wong, E.C., Hinks, R.S., Tikofsky, R.S., Hyde, J.S., 1992. Time course EPI of human brain function during task activation. *Magn. Reson. Med.* 25, 390–397.
- Belliveau, J.W., Kennedy Jr., D.N., McKinstry, R.C., Buchbinder, B.R., Weisskoff, R.M., Cohen, M.S., Vevea, J.M., Brady, T.J., Rosen, B.R., 1991. Functional mapping of the human visual cortex by magnetic resonance imaging. *Science* 254, 716–719.
- Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J. Comput. Assist. Tomogr.* 18, 192–205.
- Disbrow, E.A., Slutsky, D.A., Roberts, T.P., Krubitzer, L.A., 2001. Functional MRI at 1.5 tesla: a comparison of the blood oxygenation level-dependent signal and electrophysiology. *Proc. Natl. Acad. Sci. U.S.A.* 97, 9718–9723.
- Duong, T.Q., Kim, D.S., Ugurbil, K., Kim, S.G., 2001. Localized cerebral blood flow response at submillimeter columnar resolution. *Proc. Natl. Acad. Sci. U. S. A.* 98, 10904–10909.
- Duong, T.Q., Yacoub, E., Adriany, G., Hu, X., Ugurbil, K., Kim, S.G., 2003. Microvascular BOLD contribution at 4 and 7 T in the human brain: gradient-echo and spin-echo fMRI with suppression of blood effects. *Magn. Reson. Med.* 49, 1019–1027.
- Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C., 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn. Reson. Med.* 33, 636–647.
- Friston, K.J., Worsley, K.J., Frackowiak, R.S., Mazziotta, J.C., Evans, A.C., 1994. Assessing the significance of focal activations using their spatial extent. *Hum. Brain Mapp.* 1, 214–220.
- Jenkinson, M., Smith, S., 2002. A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5, 143–156.
- Johansen-Berg, H., Dawes, H., Guy, C., Smith, S.M., Wade, D.T., Matthews, P.M., 2002. Correlation between motor improvements and altered fMRI activity after rehabilitative therapy. *Brain* 125, 2731–2742.
- Kumari, V., Gray, J.A., Honey, G.D., Soni, W., Bullmore, E.T., Williams, S.C., Ng, V.W., Vythelingum, G.N., Simmons, A., Suckling, J., Corr, P.J., Sharma, T., 2002. Procedural learning in schizophrenia: a functional magnetic resonance imaging investigation. *Schizophr. Res.* 57, 97–107.
- Kwong, K.K., Belliveau, J.W., Chesler, D.A., Goldberg, I.E., Weisskoff, R.M., Poncelet, B.P., Kennedy, D.N., Hoppel, B.E., Cohen, M.S., Turner, R., 1992. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc. Natl. Acad. Sci. U. S. A.* 89, 5675–5679.
- Lee, S.P., Silva, A.C., Ugurbil, K., Kim, S.G., 1999. Diffusion-weighted spin-echo fMRI at 9.4 T: microvascular/tissue contribution to BOLD signal changes. *Magn. Reson. Med.* 42, 919–928.
- Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., Oeltermann, A., 2001. Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150–157.
- Luh, W.M., Wong, E.C., Bandettini, P.A., Ward, B.D., 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.
- Matthews, P.M., Jezzard, P., 2004. Functional magnetic resonance imaging. *J. Neurol. Neurosurg. Psychiatry* 75, 6–12.
- McGonigle, D.J., Howseman, A.M., Athwal, B.S., Friston, K.J., Frackowiak, R.S., Holmes, A.P., 2000. Variability in fMRI: an examination of intersession differences. *NeuroImage* 11, 708–734.
- Miezin, F.M., Maccotta, L., Ollinger, J.M., Petersen, S.E., Buckner, R.L., 2000. Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *NeuroImage* 11, 735–759.
- Mulderink, T.A., Gitelman, D.R., Mesulam, M.M., Parrish, T.B., 2002. On the use of caffeine as a contrast booster for BOLD fMRI studies. *NeuroImage* 15, 37–44.
- Obata, T., Liu, T.T., Miller, K.L., Luh, W.M., Wong, E.C., Frank, L.R., Buxton, R.B., 2004. Discrepancies between BOLD and flow dynamics in primary and supplementary motor areas: application of the balloon model to the interpretation of BOLD transients. *NeuroImage* 21, 144–153.
- Ogawa, S., Menon, R.S., Tank, D.W., Kim, S.-G., Merkle, H., Ellerman, J.M., Ugurbil, K., 1993. Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging: a comparison of signal characteristics with a biophysical model. *Biophys. J.* 64, 803–812.
- Rajapakse, J.C., Kruggel, F., Maisog, J.M., von Cramon, D.Y., 1998. Modeling hemodynamic response for analysis of functional MRI time-series. *Hum. Brain Mapp.* 6, 283–300.
- Reichenbach, J.R., Haacke, E.M., 2001. High-resolution BOLD venographic imaging: a window into brain function. *NMR Biomed.* 14, 453–467.
- Rogers, R., Wise, R.G., Painter, D.J., Longe, S.E., Tracey, I., 2004. An investigation to dissociate the analgesic and anesthetic properties of ketamine using functional magnetic resonance imaging. *Anesthesiology* 100, 292–301.
- Talagala, S.L., Noll, D.C., 1998. Functional MRI using steady-state arterial water labeling. *Magn. Reson. Med.* 39, 179–183.
- Tracey, I., 2001. Prospects for human pharmacological functional magnetic resonance imaging (phMRI). *J. Clin. Pharmacol. Suppl.*, 21S–28S.
- Wang, J., Aguirre, G.K., Kimberg, D.Y., Detre, J.A., 2003a. Empirical analyses of null-hypothesis perfusion FMRI data at 1.5 and 4 T. *NeuroImage* 19, 1449–1462.
- Wang, J., Aguirre, G.K., Kimberg, D.Y., Roc, A.C., Li, L., Detre, J.A., 2003b. Arterial spin labeling perfusion fMRI with very low task frequency. *Magn. Reson. Med.* 49, 796–802.
- Wise, R.G., Rogers, R., Painter, D., Bantick, S., Ploghaus, A., Williams, P., Rapeport, G., Tracey, I., 2002. Combining fMRI with a pharmacokinetic model to determine which brain areas activated by painful stimulation are specifically modulated by remifentanyl. *NeuroImage* 16, 999–1014.
- Wise, R.G., Williams, P., Tracey, I., 2004. Using fMRI to quantify the time dependence of remifentanyl analgesia in the human brain. *Neuropsychopharmacology* 29, 626–635.
- Wong, E.C., Buxton, R.B., Frank, L.R., 1997. Implementation of quantitative perfusion imaging techniques for functional brain mapping using pulsed arterial spin labeling. *NMR Biomed.* 10, 237–249.
- Wong, E.C., Buxton, R.B., Frank, L.R., 1998a. Quantitative imaging of perfusion using a single subtraction (QUIPSS and QUIPSS II). *Magn. Reson. Med.* 39, 702–708.
- Wong, E.C., Buxton, R.B., Frank, L.R., 1998b. A theoretical and experimental comparison of continuous and pulsed arterial spin labeling techniques for quantitative perfusion imaging. *Magn. Reson. Med.* 40, 348–355.
- Woolrich, M.W., Ripley, B.D., Brady, M., Smith, S.M., 2001. Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage* 14, 1370–1386.
- Worsley, K.J., Evans, A.C., Marrett, S., Neelin, P., 1992. A three-dimensional statistical analysis for CBF activation studies in human brain. *J. Cereb. Blood Flow Metab.* 12, 900–918.
- Yang, Y., Frank, J.A., Hou, L., Ye, F.Q., McLaughlin, A.C., Duyn, J.H., 1998. Multislice imaging of quantitative cerebral perfusion with pulsed arterial spin labeling. *Magn. Reson. Med.* 39, 825–832.
- Ye, F.Q., Smith, A.M., Yang, Y., Duyn, J., Mattay, V.S., Ruttimann, U.E., Frank, J.A., Weinberger, D.R., McLaughlin, A.C., 1997. Quantitation of regional cerebral blood flow increases during motor activation: a steady-state arterial spin tagging study. *NeuroImage* 6, 104–112.
- Yousry, T.A., Schmid, U.D., Alkadhi, H., Schmidt, D., Peraud, A., Buettner, A., Winkler, P., 1997. Localization of the motor hand area to a knob on the precentral gyrus. A new landmark. *Brain* 120, 141–157.
- Zhao, F., Wang, P., Kim, S.G., 2004. Cortical depth-dependent gradient-echo and spin-echo BOLD fMRI at 9.4 T. *Magn. Reson. Med.* 51, 518–524.