

# **A novel Bayesian approach to accounting for uncertainty in fMRI-derived estimates of cerebral oxygen metabolism fluctuations**

**Aaron B. Simon<sup>1,5</sup>, David J. Dubowitz<sup>2</sup>, Nicholas P. Blockley<sup>3</sup>, and Richard B. Buxton<sup>\*2,4</sup>**

<sup>1</sup>Department of Bioengineering, University of California San Diego, La Jolla, CA, USA

<sup>2</sup>Keck Center for Functional Magnetic Resonance Imaging, Department of Radiology, University of California San Diego, La Jolla, CA, USA

<sup>3</sup>FMRIB Centre, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK<sup>[1]</sup><sub>SEP</sub>

<sup>4</sup>Kavli Institute for Brain and Mind, University of California San Diego, La Jolla, CA, USA

<sup>5</sup>Medical Scientist Training Program, University of California San Diego, La Jolla, CA, USA

## **\*Correspondence:**

Dr. Richard B. Buxton,

University of California San Diego,

9500 Gilman Drive, MC 0677,

La Jolla, CA 92093-0677, USA

Email: rbuxton@ucsd.edu

## Abstract

Calibrated blood oxygenation level dependent (BOLD) imaging is a multimodal functional MRI technique designed to estimate changes in cerebral oxygen metabolism from measured changes in cerebral blood flow and the BOLD signal. This technique addresses fundamental ambiguities associated with quantitative BOLD signal analysis; however, its dependence on biophysical modeling creates uncertainty in the resulting oxygen metabolism estimates. In this work, we developed a Bayesian approach to estimating the oxygen metabolism response to a neural stimulus and used it to examine the uncertainty that arises in calibrated BOLD estimation due to the presence of unmeasured model parameters. We applied our approach to estimate the CMRO<sub>2</sub> response to a visual task using the traditional hypercapnia calibration experiment as well as to estimate the metabolic response to both a visual task and hypercapnia using the measurement of baseline apparent  $R_2'$  as a calibration technique. Further, in order to examine the effects of cerebral spinal fluid (CSF) signal contamination on the measurement of apparent  $R_2'$ , we examined the effects of measuring this parameter with and without CSF-nulling. We found that the two calibration techniques provided consistent estimates of the metabolic response on average, with a median  $R_2'$ -based estimate of the metabolic response to CO<sub>2</sub> of 1.4%, and  $R_2'$ - and hypercapnia-calibrated estimates of the visual response of 27% and 24%, respectively. However, these estimates were sensitive to different sources of estimation uncertainty. The  $R_2'$ -calibrated estimate was highly sensitive to CSF contamination and to uncertainty in unmeasured model parameters describing flow-volume coupling, capillary bed characteristics, and the iso-susceptibility saturation of blood. The hypercapnia-calibrated estimate was relatively insensitive to these parameters but highly sensitive to the assumed metabolic response to CO<sub>2</sub>.

**Keywords:** calibrated BOLD, cerebral oxygen metabolism, fMRI, Bayesian estimation

## 1 Introduction

Functional magnetic resonance imaging (fMRI) based on Blood Oxygenation Level Dependent (BOLD) contrast is an important tool for the study of human cognition because of its impressive ability to localize sources of evoked neural activity and its safe and noninvasive nature. However, while BOLD imaging has been highly useful for answering the question of *where* cognitive processes take place, it has had much less success in answering *how much* activity is associated with a particular process. This is largely because BOLD imaging is not directly sensitive to the electrical events that mediate neural signaling, nor to any single activity-related physiological process (e.g. blood flow, oxygen metabolism), but rather to fluctuations in the rate of MR signal decay attributable to changes in the quantity of deoxyhemoglobin in the cerebral vasculature – a quantity that depends on the relative values of several physiological variables (Buxton, 2013).

The fundamental gap between what BOLD measures—namely the change in  $R_2^*$  between two neurophysiological states—and what is typically of interest, a metric of state change-associated neural activity, has inspired considerable effort on understanding the biophysical processes that transform neural activity changes into detectable changes in  $R_2^*$ , with the goal of making it possible to quantitatively interpret BOLD signal changes in terms of fundamental physiological processes. From the inception of BOLD imaging, the link between deoxyhemoglobin and MR contrast was understood (Bandettini et al., 1992; Ogawa et al., 1992), and seminal biophysical modeling work in the early 1990s established the first quantitative links between physiological variables such as hemoglobin oxygen saturation, blood volume, and hematocrit and MR signal evolution (Ogawa et al., 1993; Yablonskiy and Haacke, 1994). In 1998 Davis et al. made a critical step in linking BOLD contrast changes to changes in neural activity by positing that by combining BOLD imaging with arterial spin labeling (ASL), an MR contrast directly sensitive to cerebral blood flow (CBF) (Detre et

al., 1992; Wong et al., 1997), and a simple biophysical model of  $R_2^*$  decay, one could quantitatively estimate a stimulus-evoked change in cerebral oxygen metabolism ( $CMRO_2$ ), a physiological parameter thought to capture the integrated energy costs of the underlying neural activity (Davis et al., 1998).

The Davis model was a breakthrough for the quantitative experimental study of the BOLD response because it presented a straightforward approach to converting BOLD and ASL measurements into  $CMRO_2$  measurements. The model divided the physiological basis of the BOLD response into three general components: (1) Changes in the Oxygen Extraction Fraction (OEF), captured by the ratio of  $CMRO_2$ -CBF changes, (2) changes in cerebral venous blood volume (CBV), captured by literature-derived relationships between CBF and CBV, and (3) the quantity of deoxyhemoglobin in the baseline state, a scaling parameter captured through a separate “calibration” experiment. Originally this calibration experiment entailed the experimentally demanding measurement the BOLD and CBF responses to hypercapnia (Davis et al., 1998); however, recently, theoretical and experimental studies have suggested that the measurement of  $R_2'$ , the rate of signal decay that may be refocused by a spin echo, in the baseline state provides similar information without the need for inhaled gases (Blockley et al., 2012; 2015).

The results of numerous calibrated BOLD studies emphasize the new information beyond what is available from the BOLD response that measurements of  $CMRO_2$  can provide as well as the feasibility of obtaining this information experimentally (Buxton et al., 2014). However, estimates of the  $CMRO_2$  response can depend significantly on assumptions about unmeasured and often unknown physiological parameters, parameters that are implicit to, but not explicitly defined in the Davis model. These physiological parameters could be factors relevant for the BOLD signal model, such as vascular volume changes, or factors relevant for the calibration experiment, such as the assumption of an iso-metabolic response to hypercapnia, or assumptions about the method used to measure  $R_2'$ .

These unmeasured parameters constitute sources of uncertainty in experimental estimates of the CMRO<sub>2</sub> response that are not easily accounted for using the Davis model or other similarly heuristic models (Griffeth et al., 2013).

In this work we describe a novel approach to account for this uncertainty in an experimental context. In this framework a detailed biophysical model that incorporates a range of important physiological variables is used as a forward model, allowing one to predict experimental measurements for a given set of physiological parameters. In addition to predicting the BOLD response to activation, the model also predicts the experimental results of the calibration experiment, either the BOLD change in the hypercapnia experiment or the signal decay curve during a spin echo experiment for the R<sub>2</sub>' approach.

In order to use this model to estimate the CMRO<sub>2</sub> response consistent with a particular set of measurements, it must be inverted. To do this we adopt a Bayesian approach, assigning a prior probability distribution to each parameter in the model, then sampling the parameter space according to the likelihood that a given set of parameters could generate our measured data given an explicitly stated statistical model, building up a collection of samples that can be used to determine a posterior probability distribution for the CMRO<sub>2</sub> response that reflects the uncertainty of that estimate given our *a priori* uncertainty about the values of the unmeasured model parameters.

We applied this approach to an experiment in which we measured BOLD and CBF responses in the human visual cortex to a visual stimulus (flickering checkerboard) and to a hypercapnia stimulus. In addition, we measured local baseline R<sub>2</sub>' using a modified technique that was designed to control for uncertainty due to partial volume effects of tissues with different T<sub>2</sub> values, and to test for the contribution of cerebrospinal fluid (CSF) effects to the apparent value of R<sub>2</sub>'. We then sought to answer the following questions: What is the uncertainty associated with an estimate of the CMRO<sub>2</sub> response if only the CBF and R<sub>2</sub>\* (BOLD) responses to a stimulus are measured? How much does

the uncertainty decrease if the baseline apparent  $R_2'$  is measured in addition to the CBF and  $R_2^*$  responses? How much does the uncertainty decrease if the  $\text{CO}_2$  stimulus is assumed to be iso-metabolic and is used to inform the estimate of the response to the visual task?

Through this analysis we found that, while considerable uncertainty in the  $\text{CMRO}_2$  response may exist even in the absence of population variability or measurement noise, this uncertainty can be reduced by improving our prior knowledge of just a few of the many unmeasured parameters in the model. Through focused investigation of these parameters and better understanding of how they vary across populations of interest, much more precise estimates of the  $\text{CMRO}_2$  response may be obtainable through functional MRI techniques.

## 2 Theory

BOLD signal decay has its origin in the paramagnetic nature of deoxygenated hemoglobin. In the presence of a strong magnetic field, deoxyhemoglobin, which is found in increasingly high concentrations as one moves down the vascular tree from arteries to veins, produces a magnetic dipole that perturbs the local magnetic field around it, leading to a loss of coherence in the precession of proton magnetic moments across an imaging voxel and ultimately signal decay. A complete model of this process would require precise knowledge of the architecture of the tissue vasculature as well as the hemoglobin oxygen saturation at each point in the vascular tree (Gagnon et al., 2015).

Similarly, precisely capturing the shape of this signal decay curve in an experiment would require many closely spaced and high signal-to-noise samples. At this time, neither of these is available for estimating the  $\text{CMRO}_2$  response to a neural stimulus in the human brain. As a result, we must necessarily make many simplifications to our model of signal decay, through which we hope to capture most of the salient features. Similarly, in our experimental measurements, time constraints limit both the signal-to-noise of our measurements and the number of samples we can obtain from a

decay curve, requiring that we use simple metrics to capture the salient features of the signal decay (here apparent  $R_2'$  in the baseline state and the apparent change in  $R_2^*$  in response to a stimulus).

In the discussion that follows we will first describe the forward signal model that we have adopted for this work, identifying the key physiological parameters that arise from it and motivating the modeling decisions that went into its development. We will then describe how we invert this model using a Bayesian probabilistic framework to compare simulated decay curves with experimental data.

## **2.1 Tissue composition**

### **2.1.1 Key assumptions of the model**

Brain tissue within an imaging voxel of typical volume ( $\sim 30\text{-}100\text{mm}^3$ ) or region of interest (ROI) is highly heterogeneous, consisting of neural parenchyma, a diverging and converging vascular tree that continuously changes in vessel diameter and oxygen saturation as one moves from artery to vein, and, potentially, non-cellular compartments such as cerebral spinal fluid (CSF). The best way to discretize this structure into a finite and tractable number of compartments is unclear, and investigators have made different choices in how to do so. The original Davis model considers only signal coming from an extravascular parenchymal compartment, which is subject to field inhomogeneity produced by a single, uniformly oxygenated venous compartment (Davis et al., 1998). Xiang He et al. expanded this model to include the signal from a single, uniform intravascular compartment as well as a CSF compartment (He and Yablonskiy, 2006). Following He et al., Dickson et al. adopted a model with intravascular, extravascular, and interstitial/CSF compartments, although they attempted to relax the assumption of uniform vessel diameter in the He model by modeling the vascular compartment as containing vessels coming from a distribution of sizes (Dickson et al., 2010), though with uniform saturation. Uludag et al., and later Griffeth et al., excluded the CSF compartment but subdivided the vascular compartment into arteries, capillaries,

and veins, assigning unique but uniform vessel sizes and oxygen saturations to each sub-compartment (Griffeth and Buxton, 2011; Uludağ et al., 2009). For this work we have attempted to balance these approaches, adopting a five compartment model containing parenchyma, arteries, capillaries, veins, and CSF. Each of these compartments is assigned a fractional volume of the total imaging volume, denoted  $V_p$ ,  $V_a$ ,  $V_c$ ,  $V_v$ , and  $V_e$ , respectively. The magnitude of the measured signal  $S(t)$  at time  $t$  after excitation is considered to be a weighted sum of the signal contributions of each compartment.

$$S(t) = \sum_{x=p,a,c,v,e} \rho_x V_x W_x S_x(t) \quad (1)$$

where the parameters  $\rho_x$ ,  $W_x$ , and  $S_x(t)$  represent the spin density, T<sub>1</sub>-weighting, and transverse signal for each compartment, respectively. By subdividing the vascular compartment and allowing the volume of each sub-compartment to be uncertain, we hope to be able to capture the important features of a continuously changing vascular tree without needing detailed knowledge of its architecture.

### 2.1.2 Signal evolution in the brain parenchyma compartment

Transverse signal decay in the parenchymal compartment ( $S_p(t)$ ) is caused by magnetic field inhomogeneity that arises from deoxyhemoglobin molecules in the intravascular space. Multiple factors affect the rate of parenchymal signal decay, as well as its behavior in response to a spin echo pulse. Important factors include the saturation of the blood, the blood volume, the orientation of the blood vessels to the main magnetic field, and the blood vessel diameters. As discussed above, none of these parameters has a single value in a physical volume of brain tissue; however, with some simplifying assumptions, an approximate model of the tissue signal behavior becomes mathematically tractable. The key assumptions made in this model are as follows: (1) within each vascular sub-compartment, oxygen saturation and vessel radius are uniform; (2) the vessels that comprise each sub-compartment may be thought of as randomly oriented and distributed, infinitely



long cylinders of uniform magnetic susceptibility; (3) vessels in the arterial and venous compartments are large enough (diameter  $\gg 10\mu\text{m}$ ) that the signal decay they produce is not affected by the diffusion of protons through the extravascular space; while those in the capillary compartment are small enough (diameter  $< 10\mu\text{m}$ ) to be significantly effected by diffusion; (4) the effects of each vascular compartment on the tissue compartment are multiplicative. In the limit where the decay produced by each vascular compartment is truly mono-exponential  $R_2^*$  decay, this is equivalent to assuming that the  $R_2^*$  contributions of each vascular compartment are additive. For example, if the presence of the venous, capillary, and arterial compartments were each to cause the signal in the tissue compartment to decay by 10% alone, their combined effect would be to cause the tissue signal to decay by approximately 27% rather than 30%. This approach prevents the total signal loss from being greater than the original signal and is an intuitive extension of previous mono-exponential models and single vascular compartment models (Yablonskiy and Haacke, 1994).

Under these conditions we can describe the parenchymal signal evolution with the equation

$$S_p(t) = e^{-R_{2,p}t} \cdot S_{p,a}(t) \cdot S_{p,c}(t) \cdot S_{p,v}(t) \quad (2)$$

where  $R_{2,p}$  is the deoxyhemoglobin independent rate of  $R_2$  decay in the parenchyma, and  $S_{p,a}$ ,  $S_{p,c}$ , and  $S_{p,v}$  are the contributions of each of the vascular compartments to the extravascular signal decay.

Because of the relatively large diameters of the arterioles and veins, the effect of the arterial and venous compartments on the parenchymal signal can be described by analytic expressions for signal decay developed by Yablonskiy and Haacke (Yablonskiy and Haacke, 1994).

$$S_{p,x}(t) = \begin{cases} e^{-V_x \cdot f_{c,x}\left(\frac{t}{\tau_{c,x}}\right)} & t < \frac{SE}{2} \\ e^{-V_x \cdot f_{c,x}\left(\frac{SE-t}{\tau_{c,x}}\right)} & \frac{SE}{2} \leq t < SE \\ e^{-V_x \cdot f_{c,x}\left(\frac{t-SE}{\tau_{c,x}}\right)} & SE \leq t \end{cases} \quad (3)$$

$$f_{c,x}\left(\frac{t}{\tau_{c,x}}\right) = \frac{1}{3} \int_0^1 du \frac{(2+u)\sqrt{1-u}}{u^2} \left(1 - J_0\left(1.5\left(\frac{t}{\tau_{c,x}}\right)u\right)\right) \quad (4)$$

$$\tau_{c,x}^{-1} = \frac{4}{3} \gamma \cdot \pi \cdot \Delta\chi_0 \cdot Hct(|Y_{off} - Y_x|) \cdot B_0 \quad (5)$$

In Equations 3-5,  $\tau_{c,x}$  describes the fundamental rate constant for signal decay. The parameters  $\gamma$ ,  $\Delta\chi_0$ , and  $B_0$  represent the gyromagnetic ratio, susceptibility difference between fully oxygenated and deoxygenated hemoglobin, and the main magnetic field strength, respectively. The values of these parameters are well understood and may be derived from the literature (Haynes, 2014; Spees et al., 2001). They are listed in Table 1. The parameter  $Hct$  represents the hematocrit of the blood and typically is not measured for a functional MRI experiment, though it may vary considerably across the healthy population (Nicoll et al., 2012). The parameter  $Y_x$  represents the fractional hemoglobin saturation of the blood in compartment  $x$ . For arterial blood this parameter may be measured directly with a pulse oximeter. For venous blood this value is typically unknown, and has been found to range from 50-75% in healthy subjects (Lu and Ge, 2008). The parameter  $Y_{off}$  is a poorly understood but important biophysical parameter representing the hemoglobin saturation that would produce no magnetic susceptibility difference between blood and tissue.

#### 2.1.2.1 The significance of $Y_{off}$

Water, proteins, and lipids, the chief constituents of blood and neural tissues, are all diamagnetic, with magnetic susceptibilities of -0.719, -0.774, and -0.670ppm, respectively (He and Yablonskiy, 2009). Because red blood cells (RBCs) contain more protein than the surrounding plasma, when fully oxygen saturated they are actually more diamagnetic than the plasma and produce field inhomogeneity. Based on the relative concentrations of protein and water in RBCs and plasma, Spees and others determined that RBCs have the same susceptibility as the surrounding plasma when hemoglobin is 95% saturated (Spees et al., 2001). Reasoning that plasma should have a similar biochemical composition to brain tissue, Uludag et al. and also Griffeth et al. assumed that at an oxygen saturation of 95%, blood ceases to produce the field inhomogeneity that leads to signal decay (Griffeth and Buxton, 2011; Uludağ et al., 2009). Others have not considered this term in their

analysis, simply assuming  $Y_{off} = 1$  (Dickson et al., 2010; He and Yablonskiy, 2006). Based on literature values for the biochemical composition and magnetic susceptibility of grey matter (He and Yablonskiy, 2009), and using the formula from Spees et al. for calculating blood susceptibility we calculate that  $Y_{off}$  could be as low as 0.89, depending upon the hematocrit and concentrations of non-heme iron in the grey matter and blood. The parameter  $Y_{off}$  is important because it essentially scales the rate of signal decay produced for vessels with a given oxygen saturation. The closer this parameter is to  $Y_x$ , the less signal decay deoxyhemoglobin in compartment  $x$  produces.

### 2.1.2.2 Diffusion and the effect of the capillary compartment

The effect of the capillary compartment on parenchymal signal decay is more challenging to model. In the capillary compartment, the vessels are considered to be small enough that the analytic equations describing signal evolution around large vessels cannot be applied. At this scale, water proton diffusion through the inhomogeneous magnetic field leads to signal dephasing that cannot be completely recovered by a spin echo pulse. Several investigators have developed Monte Carlo models to describe extravascular signal evolution due to blood vessels of this scale and have summarized their results in phenomenological models that describe the apparent rates of  $R_2$  and  $R_2^*$  for vessels of a few selected radii (Ogawa et al., 1993; Uludağ et al., 2009). In order to have more freedom to vary the physiological parameters describing the capillary compartment, we chose to reproduce the models rather than using their summary phenomenological descriptions.

Briefly, the tissue is treated as a collection of rectangular prisms each containing a single vessel. The vessels are treated as very long, straight cylinders of a specified radius and the volume of the enclosing prisms was determined such that the vessels would occupy a specified fractional volume of their prisms. The affect of a vessel on the surrounding magnetic field is determined by its oxygen saturation ( $Y_c$ ), fractional volume ( $V_c$ ), radius ( $a$ ), orientation with respect to the main magnetic field, and hematocrit ( $Hct_c$ ), taken here to be 76% of the large vessel hematocrit (Sakai et

al., 1989). To simulate signal evolution in this environment, 16 square frequency fields were generated, corresponding to vessels with orientation angles that were evenly distributed between 0 and  $\pi$ . Into each frequency field, 2000 simulated protons were randomly placed. The protons were then allowed to diffuse through the fields and accumulate magnetic moment phase offsets. At each time point in the simulation the complex sum of all proton magnetic moments was calculated. The output of this simulation was a signal evolution curve, discretely sampled at times corresponding to experimental measurements. Inline Supplemental Figure 1 shows the effect of varying  $Y_c$ ,  $Hct$ ,  $V_c$ , and  $a$  on capillary-induced parenchymal signal decay ( $S_{p,c}(t)$ ) for a spin-echo experiment in which the spin-echo time is 48ms or 98ms.

To save computational time, the capillary model was not computed for every combination of  $Y$ ,  $Hct$ ,  $V_c$ , and  $a$  that was encountered in the Bayesian sampling scheme. Instead, Lookup tables corresponding to  $Y_c = [0.3, 0.4, \dots, 1.0]$ ,  $Hct = [0.25, 0.3, \dots, 0.5]$ ,  $V_c = [0.001, 0.006, \dots, 0.061]$ , and  $a = [1, 1.5, \dots, 4\mu\text{m}]$  were generated and stored for each measured time point. For sets of parameters in between these sampled values, linear interpolation of these lookup tables was used to approximate the capillary contribution to the extravascular signal. The value of  $D$ , the diffusion constant, was fixed at  $1\mu\text{m}^2/\text{ms}$  (Ogawa et al., 1993). While, there is some uncertainty in the value of this literature-derived parameter, the important parameter, in terms of determining the effect of diffusion on signal decay is not  $D$  alone, but the characteristic diffusion time  $\tau_D = a^2/D$ , which is on the order of the time required for a water molecule to travel the distance of a field generating vessel's radius (Yablonskiy and Haacke, 1994). Thus by allowing  $a$  to range from 1-4 $\mu\text{m}$ , we are effectively allowing  $\tau_D$  to vary by a factor of 16.

### 2.1.3 Signal evolution in the intravascular compartment

The magnetic environment within a blood vessel is highly complex and heterogeneous due to the presence of red blood cells. Without knowledge of the distribution, shape, and movement of these

cells, as well as the plasma around them, it is very challenging to develop a theoretical framework for describing the transverse signal evolution in the intravascular compartments. For this reason, we have adopted an empirically derived phenomenological model of intravascular signal decay based on the work of Blockley et al. and Zhao et al. (Blockley et al., 2012; Zhao et al., 2007). In this model, intravascular signal decay at time  $t$  about a spin echo at time  $SE$  is described by the piecewise continuous mono-exponential decay function

$$S_x(t) = \begin{cases} e^{-R_{2,x}^* t} & t < \frac{SE}{2} \\ e^{-R_{2,x}(2t-SE)-R_{2,x}^*(SE-t)} & \frac{SE}{2} \leq t < SE \\ e^{-R_{2,x}SE-R_{2,x}^*(t-SE)} & SE \leq t \end{cases} \quad (6)$$

where  $R_{2,x}$  and  $R_{2,x}^*$  are functions of the Hematocrit ( $Hct_x$ ) and hemoglobin oxygen saturation ( $Y_x$ ) of the intravascular compartment:

$$R_{2,x}^* = (14.9Hct_x + 14.7) + (302.1Hct_x + 41.8)(1 - Y_x)^2 \quad (7)$$

$$R_{2,x} = (16.4Hct_x + 4.5) + (165.2Hct_x + 55.7)(1 - Y_x)^2 \quad (8)$$

#### 2.1.4 Signal evolution in the CSF compartment

In their model for transverse signal decay, He et al. suggested that an imaging volume nominally containing brain tissue contains a finite volume of CSF or extracellular fluid (He and Yablonskiy, 2006). Due to the differing biochemical makeups of grey matter and CSF, they reasoned that the magnetic moment of the CSF compartment could precess about the transverse plane at a slightly different frequency than that of the brain tissue. Because the measured signal is a complex sum of its constituents, they proposed that dephasing between a CSF compartment and a parenchymal compartment could contribute to signal loss. Fitting experimental data to their model, they estimated that this compartment had an off resonance frequency of approximately 5Hz at 3T and comprised ~5% of an average nominally gray matter voxel. Using this same model, Dickson et al. estimated the off resonance frequency to be approximately 7Hz and the fractional volume to be ~4%

of a gray matter volume (Dickson et al., 2010). Since this dephasing could be refocused by a spin-echo pulse, we reasoned that such an effect could contribute significantly to the apparent rate of  $R_2'$ , biasing the baseline measurement used to calibrate an estimate of CMRO<sub>2</sub>. As such we included such a compartment in our model, with a signal contribution described by the equation

$$S_e(t) = \begin{cases} e^{-R_{2,e}t - 2\pi i \Delta\nu \cdot t} & t < \frac{SE}{2} \\ e^{-R_{2,e}t - 2\pi i \Delta\nu \cdot (SE - t)} & \frac{SE}{2} \leq t < SE \\ e^{-R_{2,e}t - 2\pi i \Delta\nu \cdot (t - SE)} & SE \leq t \end{cases} \quad (9)$$

where  $R_{2,e}$  is the rate of  $R_2$  decay for CSF and  $\Delta\nu$  is the off resonance frequency between gray matter parenchymal tissue and CSF.

## 2.2 T<sub>1</sub>-weighting in each compartment

The signal model described above is used to simulate the signal decay that would be measured for experiments with three different types of pulse sequences. The first sequence is a PICORE QUIPSS-II ASL sequence with a dual echo readout designed for the simultaneous measurement of  $R_2^*$  and CBF-weighted signals (Wong et al., 1998). The second sequence is a Gradient Echo Sampling of Spin Echo (GESSE) sequence (Yablonskiy and Haacke, 1997) used to measure the baseline apparent  $R_2'$ . The third sequence is a variation of the GESSE sequence in which a fluid attenuated inversion recovery (FLAIR) module is added to the standard GESSE sequence in an effort to remove the effects of CSF contamination on the measurement of  $R_2'$ . Because of the different  $T_1$  preparations used in each of these experiments, the T<sub>1</sub>-weighting term  $W_x$  was modified to reflect each experiment. The equations used to describe this parameter were

$$W_x = \begin{cases} 1 - e^{-\frac{TI_2}{T_{1,x}}} & \text{ASL} \\ 1 - e^{-\frac{TR}{T_{1,x}}} & \text{GESSE} \\ 1 - \left(2 - e^{-\frac{TR - TI}{T_{1,x}}}\right) e^{-\frac{TI}{T_{1,x}}} & \text{FLAIR GESSE} \end{cases} \quad (10)$$

where TI refers to the time between the inversion and excitation pulses in the FLAIR GESSE sequence, and TI2 refers to the time between inversion and excitation pulses in the ASL sequence. This inversion time (TI2) is the relevant time for  $T_1$  relaxation because a pre-saturation pulse is applied to the imaging volume immediately before the inversion pulse. For the CSF compartment a  $T_1$  of 4000ms was assumed based on literature findings (Lin et al., 2001; Lu et al., 2005) and similarly the  $T_1$  of the parenchyma was taken to be 1200ms (Lu et al., 2005; Wansapura et al., 1999), values appropriate for a field strength of 3T. The  $T_1$  of blood was calculated for each vascular compartment as a function of hematocrit and oxygen saturation using linear interpolation of the functions described by Lu et al. for 92% and 69% oxygen-saturated blood (Lu et al., 2004). Inline Supplemental Figure 2 displays calculated blood  $T_1$  values across the hematocrit and oxygen saturation ranges relevant to this study.

### 2.3 The metabolic response to neural stimulation

The signal model described above approximates the MR signal behavior for a given set of arterial, capillary, and venous oxygen saturations and compartment volumes, as well as additional model parameters such as hematocrit. As the model is designed to describe changes in metabolism in response to a stimulus,  $CMRO_2$  must be related to oxygen saturation in the baseline and stimulus states. The fundamental equation used to relate oxygen saturation to  $CMRO_2$  is the Fick equation,  $CMRO_2 = \varepsilon \cdot CBF \cdot (Y_a - Y_v) = \varepsilon \cdot CBF \cdot Y_a \cdot OEF$ , where  $\varepsilon$  represents the oxygen carrying capacity of a milliliter of blood and OEF is the oxygen extraction fraction. The change in oxygen metabolism associated with a stimulus is thus

$$r = \frac{CMRO_{2,stim}}{CMRO_{2,0}} = \frac{CBF_{stim}Y_{a,stim}OEF_{stim}}{CBF_0Y_{a,0}OEF_0} \quad (11)$$

where the subscript *stim* indicates the stimulus state and the subscript 0 indicates the baseline state.

The change in CBF is measured, and  $Y_a$  is easily measured through pulse oximetry. Thus determining

the change in  $CMRO_2$  is reduced to determining which values of  $Y_v$  (or equivalently, which values of OEF) are consistent with the measured data in the baseline and stimulus states. Because the capillary compartment is intermediate to the arterial and venous compartments, its oxygen saturation is modeled as a weighted averaged of the arterial and venous saturations, with the Greek character  $\kappa$  representing the venous weighting.

$$Y_c = (1 - \kappa) \cdot Y_a + \kappa \cdot Y_v. \quad (12)$$

## 2.4 Relationships between baseline and stimulus blood volumes

Blood volume changes are modeled through exponential flow-volume coupling constants as described by Griffeth et al. (Griffeth and Buxton, 2011). Because some experimental work has been done to ground these models for venous volume (Chen and Pike, 2009), capillary volume (Stefanovic et al., 2008), and total CBV (Grubb et al., 1974), the models used here are described by the equations

$$\frac{V_{v,stim}}{V_{v,0}} = \left( \frac{CBF_{stim}}{CBF_0} \right)^{\phi_v}, \frac{V_{c,stim}}{V_{c,0}} = \left( \frac{CBF_{stim}}{CBF_0} \right)^{\phi_c}, \text{ and } \frac{CBV_{stim}}{CBV_0} = \left( \frac{CBF_{stim}}{CBF_0} \right)^{\phi},$$

where the subscript *stim* indicates the stimulus state and the subscript 0 indicates the baseline state, and the flow-volume coupling parameters  $\phi$ ,  $\phi_v$ , and  $\phi_c$  are typically assumed to be independent of the nature or magnitude of the hemodynamic stimulus (Griffeth and Buxton, 2011). From these equations, the change in arterial volume may be calculated. The volume of the parenchymal compartment is modeled as shrinking to accommodate to the growth of the intravascular compartments as in (Griffeth and Buxton, 2011).

## 2.5 Inverting the model to estimate stimulus induced changes in $CMRO_2$

The discussion above describes our approach to simulating the signal decay curve associated with a particular experiment and a particular set of model parameters. From these decay curves, we can simulate the characteristic measurements that we would make during a particular experiment. For the dual echo ASL experiment, these measurements are the fractional change in CBF and the



absolute change in apparent value of  $R_2^*$  between rest and stimulation. For the GESSE and FLAIR GESSE experiments the measurement is the apparent baseline  $R_2'$ . The following describes how we invert this model to determine the probability distribution of CMRO<sub>2</sub> changes that are consistent with our model, our explicitly defined assumptions about the possible values of unmeasured parameters, and our experimental measurements.

Note that we refer to both  $R_2^*$  and  $R_2'$  as ‘apparent’ in this work. We do this because, strictly speaking, characterizing signal decay by the parameters  $R_2^*$  and  $R_2'$  only describes systems undergoing mono-exponential decay. Under such ideal conditions, signal decay at time  $t$  after excitation may be described simply by the equation

$$S(t) = S_0 e^{-R_2^* t} = S_0 e^{-(R_2 + R_2') t} \quad (13)$$

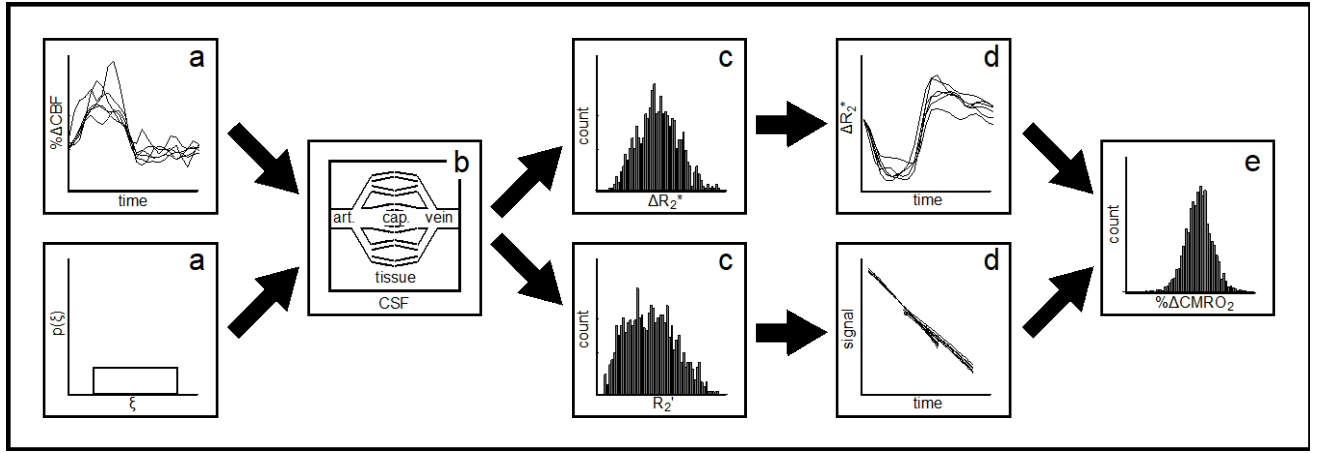
where  $R_2'$  represents the rate of signal loss than can be recovered by a spin-echo pulse and  $R_2$  represents the rate of decay that is irrecoverable. In the system described here, certain components are modeled as undergoing strict mono-exponential decay (e.g. the intravascular compartments); however, the aggregate system is not. As such, if  $R_2^*$  and  $R_2'$  are measured, their apparent values will depend on precisely when (i.e. at what time points on the decay curve) signal measurements are made. Thus the value of  $R_2'$  measured by one experimental protocol (e.g. GESSE) is not the same as the value of  $R_2'$  measured in the same system by a different protocol (e.g. asymmetrical spin-echo). To account for this, we calculate  $R_2^*$  and  $R_2'$  from our simulated decay curves in the same way that we do for the experiments themselves, which is described in detail in Section 3.6 below.

To estimate the uncertainty in an estimate of the CMRO<sub>2</sub> response attributable to uncertainty in our measurements, as well as to unmeasured model parameters, we adopted a Bayesian probabilistic model. In the context of this model, the probability that the underlying, unmeasured model parameters, denoted here by the array  $\xi$ , and the measurement parameters,  $\psi$ , take on a particular set of values given the measured data,  $y$ , can be expressed by Bayes’ Rule

$$p(\xi, \psi | y) \propto p(\xi)p(\psi)p(y|\xi, \psi) \quad (14)$$

In this equation  $p(\xi, \psi | y)$  represents the probability that  $\xi$  and  $\psi$  take on a particular set of values given the measured data. The expressions  $p(\xi)$  and  $p(\psi)$  are termed prior probability densities and represent our beliefs about the possible values of these parameters before data is collected. The expression  $p(y | \xi, \psi)$  is termed the likelihood function and represents the probability that we would obtain the observed measurements,  $y$ , given that  $\xi$  and  $\psi$  take on a particular set of values. In this study, the array  $\xi$  contains the elements  $OE_{stim}$ ,  $OE_{0}$ ,  $V_a$ ,  $V_v$ ,  $V_a$ ,  $V_e$ ,  $\phi$ ,  $\phi_v$ ,  $\phi_c$ ,  $R_{2t0}$ ,  $a$ ,  $Hct$ ,  $Y_{off}$ ,  $\Delta v$ , and  $\kappa$ . If in the analysis we were to use assumptions about the CMRO<sub>2</sub> response to CO<sub>2</sub> to constrain the estimate of the response to the visual task, then  $r_{CO_2}$  (the ratio of CMRO<sub>2</sub> after and before CO<sub>2</sub> inhalation) would be included in  $\xi$ . The array  $\psi$  contains the population means  $\mu$  and variances  $\sigma^2$  associated with each of the measurements used in the estimation of the CMRO<sub>2</sub> response.

Because Equation 14 cannot be solved analytically, we estimate  $p(\xi, \psi | y)$  by drawing random samples from the posterior probability density. By drawing sufficient samples, we slowly build up a picture of the entire probability distribution (Figure 1). In this study we are most interested in the posterior probability distribution of  $r$ , the CMRO<sub>2</sub> response to a stimulus defined in Equation 11. Although this is not one of the parameters in  $\xi$ , it may be easily calculated for each sample from the parameters in  $\xi$  and  $\psi$  using Equation 11. The algorithm we used to sample this posterior probability distribution is described in Section 3.7 below.



**Figure 1: Schematic of Bayesian estimation approach:** (a) Measured CBF responses to a stimulus of interest and prior beliefs about unmeasured model parameters are input into a detailed model as probability distributions. (b) The model describes signal evolution for a volume of brain containing five compartments (tissue, CSF, arteries, capillaries, veins). (c) The model is used to simulate measurements of apparent  $\Delta R_2^*$  and  $R_2'$  for random samples from the probability distributions described by column a. (d) Simulated  $\Delta R_2^*$  and  $R_2'$  are sampled based on their probability of representing the population means of the measured apparent  $\Delta R_2^*$  and  $R_2'$  values. (e) Over repeated sampling the posterior probability distributions of the unmeasured model parameters are estimated, including the parameter of interest, the percentage change in CMRO<sub>2</sub> associated with the stimulus.

### 3 Materials and methods

#### 3.1 Subjects

Six healthy adult subjects (three female) participated in this study (ages 22-29 years). The study was approved by the institutional review board at the University of California San Diego, and written informed consent was obtained from all participants.

#### 3.2 Imaging

Simultaneous BOLD and CBF-weighted images were acquired on a GE Discovery 750 3T scanner with a dual-echo arterial spin labeling (ASL) PICORE QUIPSS II sequence (Wong et al., 1998) with a spiral readout. Eight slices (5mm thick/1mm gap) were acquired, with slices 2-5 (from inferior to superior) aligned by visual inspection with the calcarine sulcus. Pulse sequence parameters were as follows: TR 3.0s, TI1/TI2 700/1800ms, TE1 3.3ms, TE2 30ms, 90° flip angle, FOV 256mm,

matrix 64x64. In addition, a cerebral spinal fluid (CSF) reference scan and a minimum contrast scan were acquired for use in quantifying CBF (Chalela et al., 2000; Wang et al., 2005). The CSF and minimum contrast scans were single-shot spiral EPI images with TE 3.3ms, TR 4s and TE 11ms TR 2s, respectively, and the same in-plane parameters as the ASL scan. A field map was also acquired for use in correcting distortions in the spiral images due to the inhomogeneity of the magnetic field.

Measurements used to estimate  $R_2'$  were made using a gradient echo sampling of spin echo (GESSE) imaging sequence (Yablonskiy and Haacke, 1997). The GESSE image slices were aligned with the centers of the ASL images and had the same in plane resolution and field of view. However, in order to reduce the effects of through-plane gradients on the estimate of  $R_2'$ , these images were acquired with a slice thickness of 2mm, with the gap between slices increased to 4mm. Two pairs of GESSE image series were collected for this study. Each pair of images consisted of an early spin echo series and a late spin echo series, which were acquired as separate scans. The early spin echo occurred 48ms after excitation and the late spin echo, 98ms after excitation. The samples of each spin echo decay curve were collected asymmetrically. Around the early spin echo curve, 64 samples were collected from 42.77-82.59ms after excitation at intervals of 0.63ms. Around the late spin echo curve, samples were collected from 62.78-102.59ms with the same interval spacing. For the first pair of image series, no T1 preparation pulses were used, and the TR was 2s. For the second pair of image series, an inversion pulse was added before the excitation pulse with the intention of minimizing the cerebral spinal fluid (CSF) signal in the 2<sup>nd</sup>-5<sup>th</sup> slices from the bottom of the image stack. A TR of 3.5s and inversion time (TI) of 1.38s was chosen to null the CSF signal, based on an assumed CSF  $T_1$  of approximately 4000ms (Lin et al., 2001; Lu et al., 2005). The third slice from the bottom of the stack was chosen to occur at the CSF null time. Each image slice was acquired sequentially from inferior to superior with a spacing of 110ms. Based on equation 10, we estimated that the CSF signal in the FLAIR GESSE images would be less than 8% of its completely relaxed value in the 5th slice,

which was least optimally nulled, or less than 20% of the CSF signal in the standard GESSE images. Inline Supplemental Figure 3 displays FLAIR GESSE signal intensity as a fraction of GESSE signal intensity across slices 2-5 for a single subject, with signal suppression in the ventricles consistent with the theoretical prediction.

### **3.3 Visual stimulus paradigm**

Visual Stimuli were produced using MATLAB® (2009a, The MathWorks, Natick, MA) with the Psychophysics Toolbox extensions (Pelli, 1997). The visual stimulus consisted of an 8Hz black and white flickering radial checkerboard with a central region (visual angle  $\sim 1.5^\circ$ ) that was maintained an iso-luminescent gray. The visual stimulus was projected onto a screen, which the subject could view through a head coil mounted mirror.

Each study contained two visual task runs during which simultaneous BOLD and CBF weighted images were acquired. During each scan cardiac and respiratory activity were recorded using a pulse oximeter and respiratory bellows that were built into the MRI scanner. Throughout each of the runs, subjects were asked to fixate on the center of the screen. In order to maintain the subjects' attention, random numbers (0-9) were displayed in the gray central region of the screen at 1s intervals. The subjects were instructed to press a button on a response box each time a number was displayed twice in a row. The first functional run was used to locate a region of interest (ROI) in the visual cortex. The stimulus paradigm began with 24s of rest followed by 6 cycles of 24s-stimulus, 24s-rest. The second functional run was used to quantify the CBF and BOLD responses to a visual stimulus. The visual stimulus used for this run was the same as that used for ROI localization, however the timing of the stimuli was altered to improve estimation of the baseline signal and to allow for full recovery from the post-stimulus undershoot between stimulus cycles. This functional run began with 72s rest, followed by 6 cycles of 24s-stimulus, 48s-rest and ended with an additional 60s period of rest.

### 3.4 CO<sub>2</sub> stimulus paradigm

Throughout the imaging session, each subject wore a non-rebreathing facemask (Hans Rudolph, KS, USA). The inspiratory port was connected to a tube that was open to the air in the scanner room but could be connected to a large gas-tight bag filled with a premixed gas (5% CO<sub>2</sub>, 21% O<sub>2</sub>, balance N<sub>2</sub>) by turning a valve.

Each study contained a CO<sub>2</sub> stimulus run during which BOLD and CBF weighted images were again acquired. The duration of the run was 9min. For the first 3.5 minutes the inspiratory port was open to room air while baseline measurements were acquired. A valve was then turned, switching the source of air to the 5% CO<sub>2</sub> mixture for a period of three minutes. For the final 2.5 minutes, the inspiratory port was again open to room air. Throughout the run, the subject was asked to keep his or her eyes open and focus on the center of the projection screen. To maintain subject attention, the number repetition task was again employed throughout the run.

### 3.5 Preprocessing

Raw ASL images were first corrected for distortions due to the inhomogeneity of the magnetic field (Noll et al., 2005). The first four images of each scan were discarded to allow the MRI signal to reach steady state. All functional runs were motion corrected and registered to the first visual task run using AFNI software (Cox, 1996). In order to minimize BOLD contamination of the CBF measurements, CBF-weighted image series were produced from the raw first-echo ASL images by surround subtraction (Liu and Wong, 2005).  $R_2^*$ -weighted (i.e. BOLD-weighted) images were produced from both the first echo and second echo ASL images by surround averaging (Liu and Wong, 2005) and used to calculate quantitative  $R_2^*$  image series as described below.

All quantitative analysis was performed in a region of interest (ROI) within the visual cortex, which was defined by the BOLD and CBF response of each subject to the first visual task. Statistical analysis for ROI selection was performed with a general linear model approach for the analysis of ASL data as described by Perthen (Perthen et al., 2008). Briefly, a stimulus regressor was produced by convolving the stimulus pattern with a gamma density function (Boynton et al., 1996). Cardiac and respiratory signals were used as regressors to account for the non-stimulus related signal variance produced by physiological processes (Glover et al., 2000; Restom et al., 2006). A constant and a linear term were also included as regressors. An anatomical mask that included only gray matter voxels in the posterior half of the brain and within slices 2-5 from inferior to superior was produced for each subject, and ROI selection was restricted to this region. Voxels exhibiting CBF or BOLD activation were detected after correcting for multiple comparisons using AFNI AlphaSim (Cox, 1996), using an overall significance threshold of  $p = 0.05$  given a minimum cluster size of four voxels. For each subject, an active visual cortex region of interest (ROI) was defined as those voxels exhibiting both CBF and BOLD activation independently.

Estimates of the deoxyhemoglobin-related  $R_2'$  in the brain may be biased by the presence of air-tissue interfaces, which produce magnetic field inhomogeneity across the brain that, in turn, produces  $R_2'$ -type signal decay that is unrelated to the quantity and distribution of deoxyhemoglobin in an imaging voxel (Dickson et al., 2010; He and Yablonskiy, 2006). We used a method described by Dickson and others to correct our GESSE images for the effects of through-plane gradients before using them for quantitative analysis, using phase images derived from the GESSE image series themselves to generate field maps and making the assumptions that the slice profile was approximately rectangular and that the field inhomogeneity could be approximated as linear gradients through each voxel. Details of the field correction procedure may be found in (Dickson et al., 2010).

After correction, all GESSE images series were registered to the first visual task ASL image series using AFNI software (Cox, 1996).

### 3.6 Quantitative data analysis

Before performing quantitative analysis on each subject, the CBF-weighted and  $R_2^*$ -weighted image series from the second visual task run and CO<sub>2</sub> run, as well as the four GESSE image series, were spatially averaged across the ROI defined by the subject's response to the first visual task, reducing each four dimensional data set (3 spatial dimensions plus time) to a one-dimensional time series. From each pair of  $R_2^*$ -weighted time series, a single time series representing the apparent  $R_2^*$  at every sample  $t$  was then made using the equation,  $R_2^* = \ln \left( S \left( \frac{TE1}{TE2} \right) \right) / (TE2 - TE1)$ , where  $S(TE1)$  and  $S(TE2)$  represent the remaining signal at the first and second echo times, respectively, at time point  $t$ , assuming mono-exponential decay. The BOLD or  $\Delta R_2^*$  response to a stimulus was defined as the difference between the mean value of  $R_2^*$  during the baseline (rest) period and the stimulus period. The CBF response to a stimulus,  $f$ , was defined as the ratio of the mean CBF in the stimulus period to the mean CBF in the rest period. For the visual stimulus, the stimulus period was defined as the last 12sec of each visual stimulus cycle and the baseline period the initial and final 60s periods of rest (120s total). For the CO<sub>2</sub> stimulus, the stimulus period was defined as the final two minutes of the stimulus and the rest period as the first 3min before the stimulus was applied.

To estimate the apparent  $R_2'$  in the baseline state, the GESSE time series were assumed to represent mono-exponential signal decay around a spin echo, which can be described by the equation

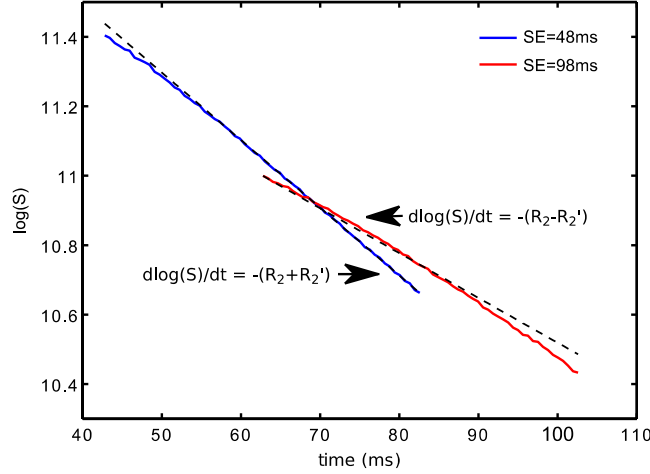
$$S(t) = \begin{cases} S_0 e^{-(R_2 + R_2')t} & t < \frac{SE}{2} \\ S_0 e^{-(R_2 - R_2')t - R_2' SE} & \frac{SE}{2} \leq t < SE \\ S_0 e^{-(R_2 + R_2')t + R_2' SE} & SE \leq t \end{cases} \quad (15)$$

During the period when the signals are both sampled, the early (48ms) spin-echo decay curve is in the third time regime and the late (98ms) spin-echo decay curve is in the second time regime. The



apparent  $R_2'$  could thus be calculated as one half the difference in slopes between the logarithms of the late and early spin echo decay curves (Figure 2)

$$R_2' = \frac{1}{2} \left( \frac{d\ln(S(t)_{SE=98ms})}{dt} - \frac{d\ln(S(t)_{SE=48ms})}{dt} \right) = \frac{(R_2' - R_2) + (R_2' + R_2)}{2} \quad (16)$$



**Figure 2: Dual GESSE imaging approach to the estimation of  $R_2'$ .** Representative FLAIR GESSE signal decay curves from a single subject (subject 3) shown. The GESSE sequences measure signal decay asymmetrically about SE times of 48 ms (blue curve) and 98 ms (red curve). Far from the spin echo, the log decay curves become approximately linear, with slopes of  $-(R_2+R_2')$  and  $-(R_2-R_2')$ , respectively.  $R_2'$  can be calculated from the difference in these slopes.

### 3.7 Sampling from the posterior probability distribution of the CMRO<sub>2</sub> change

To sample the posterior probability density of  $r$ , the activation CMRO<sub>2</sub> value normalized to the baseline CMRO<sub>2</sub> value, we used a simple algorithm we designed to efficiently gather samples while fully exploring the parameter space. For this work, we wished to isolate uncertainty in the CMRO<sub>2</sub> estimate stemming from two sources: 1) uncertainty in the measured values due to measurement error and response variability across the population sample; and 2) uncertainty due to unknown values of underlying physiological variables in the model. For this reason, we considered two cases in the analysis: “absolute” uncertainty, which includes contributions from both measurement uncertainty and physiological uncertainty (case 1); and “intrinsic” uncertainty due just to the physiological uncertainty (case 2). Our approach was as follows:

1. We first chose a statistical model for each measurement  $y_i$  (e.g.  $f, \Delta R_2^*, R_2', Y_a$ ). In order to assess the uncertainty attributable to variance in the measurements across the sample population, we assumed a simple model, that each per subject measurement  $y_i^j$  comes from a Gaussian distribution of unknown mean  $\mu_i$  and variance  $\sigma_i^2$ , and that each measurement is independent and identically distributed. To isolate the intrinsic uncertainty due to uncertainty in the unmeasured model parameters alone, we assumed that the sample mean for each measurement is identical to the population mean.
2. Because the population variance of the measurements,  $\sigma^2$ , was not a parameter of interest, we next marginalized the joint posterior probability distribution by integrating over  $\sigma^2$ .

$$p(\xi, \mu | y) \propto \int_{\sigma_1^2} \dots \int_{\sigma_i^2} p(\xi) p(\mu) p(\sigma^2) p(y | \xi, \mu, \sigma^2) d\sigma_1^2 \dots d\sigma_i^2 \quad (17)$$

Because the likelihood function  $p(y_i | \xi, \mu, \sigma^2)$  depends only on  $\mu_i$  and  $\sigma_i^2$ , Equation 17 may be rewritten as

$$p(\xi, \mu | y) \propto p(\xi) \prod_i \int_{\sigma_i^2} p(\mu_i) p(\sigma_i^2) p(y_i | \mu_i, \sigma_i^2) d\sigma_i^2 \propto p(\xi) \prod_i p(\mu_i | y_i) \quad (18)$$

3. For the Gaussian statistical model, assuming the non-informative prior  $p(\mu_i, \sigma_i^2) \propto \frac{1}{\sigma_i^2}$ , the marginal posterior probability function  $p(\mu_i | y_i)$  has the form of a  $t$ -distribution (Gelman et al., 2014)

$$\left. \frac{\mu_i - \bar{y}_i}{s_i / \sqrt{n}} \right| y_i \sim t_{n-1} \quad (19)$$

where  $\bar{y}_i$  and  $s_i$  denote the sample mean and standard deviation of measurement  $i$  and  $n$  is the number of samples. In the case in which measurement variance is ignored, the parameters in  $\sigma^2$  are irrelevant, and marginal posterior probability function for  $\mu_i$  can be written

$$p(\mu_i | y_i) \propto \begin{cases} 1 & \mu_i = \bar{y}_i \\ 0 & \mu_i \neq \bar{y}_i \end{cases} \quad (20)$$

4. To construct samples of the posterior distribution,  $p(\xi, \mu | y)$ , we drew random samples from the prior distributions  $p(\xi)$  and marginal posterior distributions  $p(\mu | y)$  such that for each set of samples, the forward model would generate the sampled population means of  $R_2'$  and  $\Delta R_2^*$ . To do this we first drew random samples from the prior distributions,  $p(\xi_i)$ , of each of the unmeasured physiological variables with the exceptions of  $OEf_0$  and  $OEf_{stim}$ . The prior distribution used for each variable in  $\xi$  was uniform (i.e. equal probability density) within a finite range of values (see Table 1), which we denote as  $p(\xi_i) \sim U(\xi_{i,a}, \xi_{i,b})$  where  $\xi_{i,a}$  is the lowest non-zero probability value of parameter  $\xi_i$  and  $\xi_{i,b}$  is the highest non-zero probability value.
5. We then drew random samples of  $\mu_i$  from  $p(\mu_i | y_i)$  for each measurement used in the estimation process.
6. If no calibration data was used in the estimation process (either from measuring  $R_2'$  or from measuring the CBF and  $R_2^*$  responses to  $CO_2$ ), then  $OEf_0$  was simply drawn from the prior distribution  $p(OEf_0)$ . If calibration data was used, then the value of  $OEf_0$  was fit such that the forward model for  $R_2'$  or the forward model for the  $\Delta R_2^*$  response to  $CO_2$  would produce the sampled value of  $\mu_{R_2'}$  or  $\mu_{\Delta R_2^*, CO_2}$ , respectively, depending upon the chosen calibration experiment.
7. After sampling or fitting for  $OEf_0$ , the parameter  $OEf_{stim}$  was fit such that the forward model of the  $\Delta R_2^*$  response to the stimulus of interest would produce the sampled value of  $\mu_{\Delta R_2^*, stim}$ .
8. Because  $p(OEf_0)$  and  $p(OEf_{stim})$  were uniform, if the fit values of  $OEf_0$  and  $OEf_{stim}$  were within the supported domains of their prior distributions, the sample was accepted, and the value of  $r$  was calculated for that sample using Equation 11. Otherwise the sample was rejected.

9. Steps 4-7 were repeated until sufficient samples were obtained. Samples were processed in batches of 1000. After each batch was processed, half of the complete sample set was randomly and repeatedly resampled. The 95% central interval was calculated for each partial sample, and sufficient samples were assumed to have been collected when the upper and lower bounds of the 95% central intervals from the partial data sets achieved variances of less than 1%.

## 4 Results

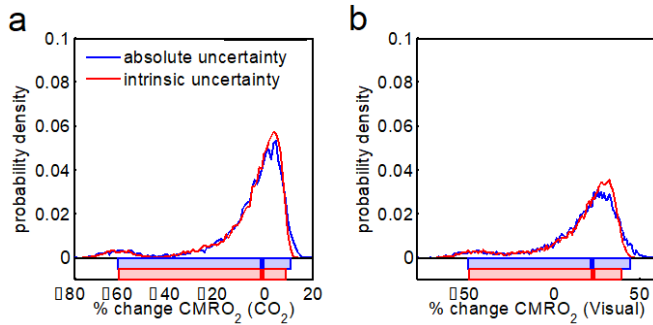
### 4.1 Summary of baseline and stimulus response measurements

Across subjects average measured baseline apparent  $R_2'$  was  $3.94 \pm 0.64 \text{ s}^{-1}$  using the standard GESSE protocol and  $3.05 \pm 0.41 \text{ s}^{-1}$  using the FLAIR GESSE protocol (mean  $\pm$  std). Average CBF responses to  $\text{CO}_2$  and to the visual stimulus, with respect to baseline, were  $24 \pm 5\%$  and  $69 \pm 16\%$ , respectively. The average  $\Delta R_2^*$  responses to  $\text{CO}_2$  and the visual stimulus were  $-0.63 \pm 0.16 \text{ s}^{-1}$  and  $-0.74 \pm 0.17 \text{ s}^{-1}$ , respectively. The average arterial oxygen saturation ( $Y_a$ ) was  $0.99 \pm 0.01$ . Responses for individual subjects may be seen in Table 2.

### 4.2 Estimating the $\text{CMRO}_2$ response from CBF and $\Delta R_2^*$ alone

Quantitative estimation of the  $\text{CMRO}_2$  response typically involves a calibration experiment, which is designed to capture characteristics of the baseline state that affect the magnitude of the BOLD response associated with a given change in CBF and  $\text{CMRO}_2$ . Our first question was what could be determined about the magnitude of the  $\text{CMRO}_2$  response, given our prior uncertainty about the unmeasured variables in our model, without such a calibration experiment. Figure 3 shows the posterior probability distributions of the  $\text{CMRO}_2$  responses to (a)  $\text{CO}_2$  and (b) the visual stimulus with and without consideration of measurement noise. The shaded bars represent 95% central

intervals for each estimate. Even without the effect of measurement uncertainty, little may be concluded about the magnitude of the CMRO<sub>2</sub> response without additional information. To the CO<sub>2</sub> stimulus, the estimated responses were -1.5% [-62 – 10%] (median [95% CI]) and -1.2% [-62 – 8.6%], accounting for and ignoring measurement uncertainty, respectively. To the visual stimulus, the estimated responses were 22% [-51 – 44%] and 22% [-50 – 39%].

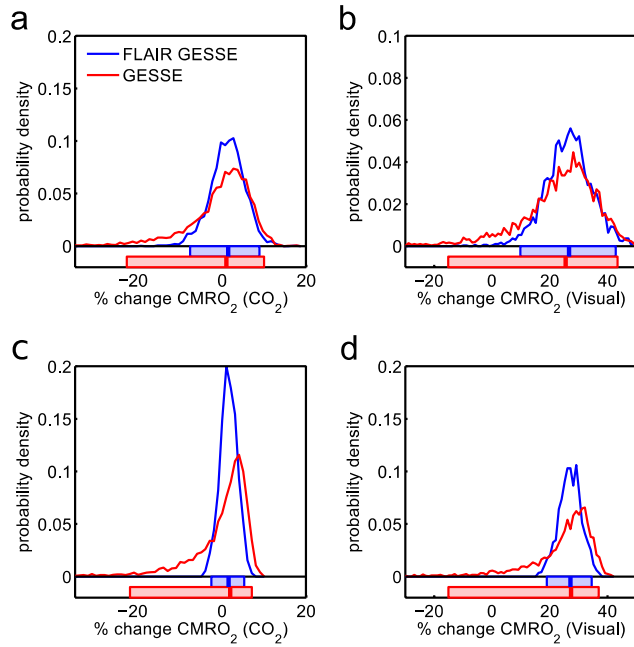


**Figure 3: Uncertainty in the estimate of the CMRO<sub>2</sub> response based on measurements of R<sub>2</sub><sup>\*</sup> and CBF responses alone.** Uncertainty due to both measurement variance and intrinsic uncertainty in unmeasured model parameters (blue line) and due to intrinsic uncertainty alone (red line) in the estimate of the CMRO<sub>2</sub> response to (a) CO<sub>2</sub> and (b) a visual stimulus. Shaded bars indicate 95% central intervals. Without additional information from a calibration experiment, little may be concluded about the response to either stimulus.

### 4.3 Calibrating the CMRO<sub>2</sub> estimate by measuring baseline apparent R<sub>2</sub>'

We next asked how much our estimate of the CMRO<sub>2</sub> response would be improved by using a baseline measurement of the apparent R<sub>2</sub>' to calibrate the BOLD response to each stimulus. Figure 4 shows the uncertainty associated with estimating the CMRO<sub>2</sub> response based on R<sub>2</sub>' calibration. Compared with no calibration, R<sub>2</sub>' calibration greatly decreased uncertainty in the estimate of the CMRO<sub>2</sub> response. Using the FLAIR preparation to suppress CSF contamination of the signal further reduced the uncertainty in the estimate. The estimated response to the CO<sub>2</sub> stimulus, accounting for measurement noise, was 1.4% [-7.6 – 9.1%] (median [95% CI]) for the FLAIR GESSE estimate and 1.0% [-23 – 10%] for the standard GESSE estimate. The estimated response to the visual stimulus

was 27% [9.9 – 43%] for the FLAIR GESSE estimate and 25% [-15 – 43%] for the standard GESSE estimate. Accounting only for intrinsic uncertainty, the estimated response to the CO<sub>2</sub> stimulus was 1.5% [-2.5 – 5.3%] for the FLAIR GESSE estimate and 1.9% [-22 – 7.0%] for the standard GESSE estimate. The estimated response to the visual stimulus was 27% [19 – 34%] for the FLAIR GESSE estimate and 27% [-15 – 37%] for the standard GESSE estimate.



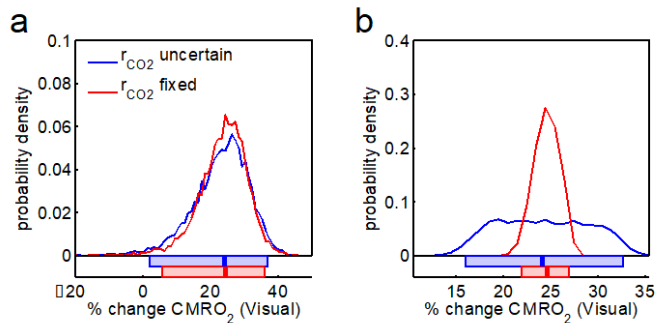
**Figure 4: Uncertainty in estimate of the CMRO<sub>2</sub> response when calibrated by measurement of baseline apparent R<sub>2</sub>' using either a standard GESSE or a CSF-nulling (FLAIR) GESSE protocol.** Blue curves indicate FLAIR GESSE-based estimates, while red curves indicate standard GESSE-based estimates. Top row shows probability density curves accounting for both measurement variance and intrinsic uncertainty in estimates of response to (a) the CO<sub>2</sub> stimulus and (b) the visual stimulus. Bottom row (c and d) show density curves accounting only for intrinsic uncertainty. Shaded bars indicate 95% central intervals.

#### 4.4 Calibrating the CMRO<sub>2</sub> response estimate to the visual stimulus by assuming the response to CO<sub>2</sub>

Often, the CMRO<sub>2</sub> response to a stimulus of interest is calibrated by measuring the CBF and R<sub>2</sub>\* changes that result from breathing air containing elevated levels of CO<sub>2</sub>, under the assumption that it does not produce a change in CMRO<sub>2</sub> (Barzilay et al., 1985; Chen and Pike, 2010; Davis et al.,

1998; Mark et al., 2011; Sicard and Duong, 2005). However, this assumption remains controversial and some studies have shown either increases (Horvath et al., 1994; Yang and Krasney, 1995) or decreases (Xu et al., 2011; Zappe et al., 2008). As we reported above, our estimates of the CO<sub>2</sub> response based on apparent R<sub>2</sub>' are consistent with CO<sub>2</sub> having a negligible effect on CMRO<sub>2</sub>, albeit with some uncertainty. As such we asked what the uncertainty in our estimate of the CMRO<sub>2</sub> response to a visual stimulus would be if we either assumed that the CMRO<sub>2</sub> response to CO<sub>2</sub> was negligible or assigned to it a modest prior uncertainty ( $p(r_{CO_2} \sim U(0.95, 1.05))$ ) and used it to calibrate the response to the visual stimulus instead of using the baseline R<sub>2</sub>'. All other parameters of the model maintained their respective prior uncertainties.

Figure 5 displays the posterior uncertainty of the CMRO<sub>2</sub> response to the visual stimulus under these conditions. When it was assumed that the CMRO<sub>2</sub> response to CO<sub>2</sub> was negligible, CO<sub>2</sub> calibration significantly reduced the uncertainty in the estimate of the response to the visual stimulus. Accounting for measurement uncertainty, the response was estimated to be 24% [5.6 – 36%] and ignoring measurement uncertainty, 25% [22 – 27%]. However, even a small amount of uncertainty in the CO<sub>2</sub> response produced considerable uncertainty in the estimate of the response to the visual stimulus. Accounting for measurement uncertainty, the response was estimated to be 24% [1.8 – 37%], and ignoring measurement uncertainty, 24% [16 – 33%].



**Figure 5: Posterior uncertainty associated with estimate of the CMRO<sub>2</sub> response to visual stimulus when calibrated by assumed response to CO<sub>2</sub>.** Blue curves indicate conditions under which the CMRO<sub>2</sub> response to a separately measured CO<sub>2</sub> stimulus is assigned a finitely wide but otherwise uninformative prior probability distribution. Red curves indicate conditions under which

the response to CO<sub>2</sub> is assumed to be iso-metabolic (a) Posterior probability densities for estimates of CMRO<sub>2</sub> response to visual stimulus, accounting for both measurement and intrinsic uncertainty. (b) Posterior probability distributions to the same stimuli, accounting only for intrinsic uncertainty. Shaded bars indicate 95% central intervals.

## 5 Discussion

In this work we developed a novel Bayesian approach to estimating the CMRO<sub>2</sub> response to a neural stimulus based on calibrated BOLD data and used it to examine the uncertainty that arises in estimates of the CMRO<sub>2</sub> response when the values of unmeasured physiological model parameters are not precisely known. We examined estimates calibrated by a traditional hypercapnia experiment, in which the measured changes in CBF and  $R_2^*$  in response to breathing elevated levels of CO<sub>2</sub> are used to obtain information about baseline deoxyhemoglobin, as well as estimates calibrated by the more novel technique of measuring baseline apparent  $R_2'$ . In order to minimize the effect of multi-exponential  $T_2$  decay on the estimate of  $R_2'$ , we employed a novel measurement approach using two GESSE imaging sequences with different spin-echo times. Further, in order to examine the effects of CSF-contamination on the measurement of apparent  $R_2'$ , we measured this parameter with and without a CSF-nulling preparation pulse. We found that the two calibration approaches provided highly comparable CMRO<sub>2</sub> response estimates on average. However, we found that while uncertainty due to measurement variance was a significant source of uncertainty in our estimates, likely due to the small sample size of the study, we also found that prior uncertainty in the unmeasured model parameters produced considerable intrinsic uncertainty in our estimates, and that the magnitude of this uncertainty depended upon the choice of calibration experiment.

### 5.1 Sources of intrinsic uncertainty in $R_2'$ -calibrated CMRO<sub>2</sub> estimates

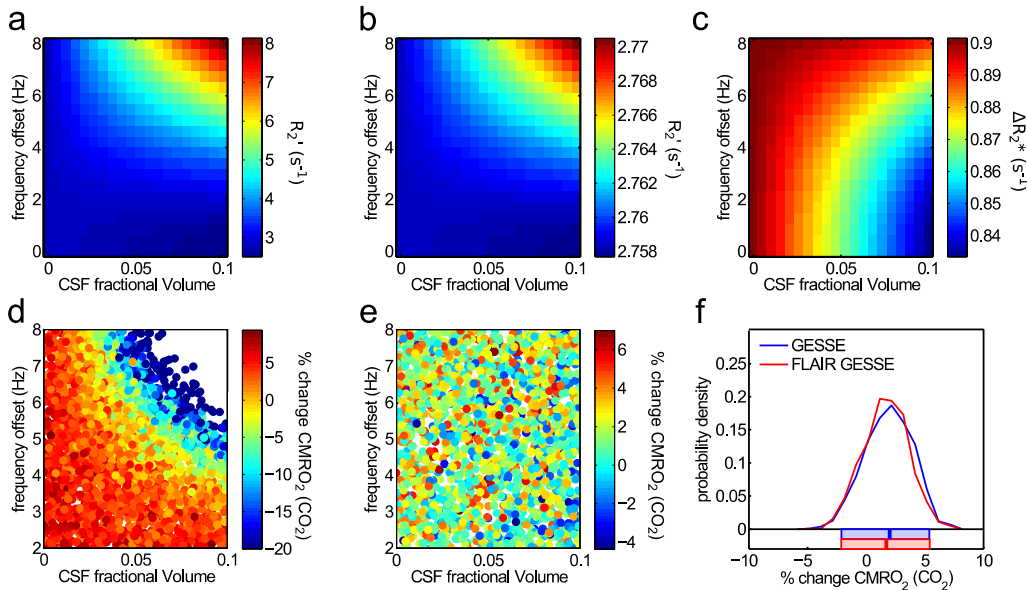
As originally envisioned, the calibration experiment was designed to capture information about the quantity of deoxyhemoglobin (approximately the product of hematocrit, venous blood



volume, and oxygen extraction fraction) in a voxel in the baseline state, a quantity determined in theoretical work to strongly influence the magnitude of the  $R_2^*$  (BOLD) response to a given change in CBF and  $CMRO_2$  (Davis et al., 1998). In both theoretical and experimental work, the apparent spin echo-recoverable transverse relaxation rate,  $R_2'$ , has been found to reflect this baseline deoxyhemoglobin state, making it a potentially valuable calibration metric (Blockley et al., 2012; Fujita et al., 2006; Kida et al., 2000; Yablonskiy, 1998). We found that measuring baseline apparent  $R_2'$  did significantly reduce the uncertainty associated with the  $CMRO_2$  estimate, compared with that of a calibration-free estimate, but that several unmeasured model parameters contributed significantly to the intrinsic uncertainty associated with the measurement.

The largest source of this uncertainty in our analysis was associated with the potential for CSF contamination of a nominally gray matter imaging volume in the  $R_2'$  experiment, which in this model could produce spin-echo recoverable signal loss. Figure 6 illustrates the how CSF partial volume produces estimation uncertainty. Figures 6a and 6b display simulations of apparent  $R_2'$  as a function of CSF fractional volume and resonance frequency offset for an otherwise fixed set of model parameters. Figure 6a displays simulations associated with a standard GESSE protocol, while Figure 6b displays simulations associated with a protocol optimized to minimize the CSF signal. In this simulation, even a very small ( $<0.05$ ) CSF fraction could bias the measurement of  $R_2'$  by as much as 50% using the standard approach. Suppression of the CSF signal minimized this bias. Figure 6c displays simulations of apparent  $\Delta R_2^*$  for an iso-metabolic 50% increase in CBF. CSF volume and off-resonance also affected the value of  $\Delta R_2^*$  in this simulation; however, the effect size was on the order of a few percent, orders of magnitude smaller than the  $R_2'$  effect. Because the measurement bias produced by CSF contamination is greater for  $R_2'$  than for  $\Delta R_2^*$ , it may cause one to overestimate the  $CMRO_2$  response if it is not accounted for. If the size and off resonance frequency of this compartment are unknown, then attempting to account for them through a prior distribution

introduces uncertainty into the CMRO<sub>2</sub> estimate. Figures 6d and 6e show how uncertainty in CSF volume and off-resonance create uncertainty in the estimated CMRO<sub>2</sub> response to the CO<sub>2</sub> stimulus when R<sub>2</sub>' is measured without or with CSF suppression, respectively. While these parameters strongly influence the estimate based on the standard GESSE measurement, they have negligible influence on the FLAIR GESSE estimate. Figure 6f shows the probability distributions associated with standard or FLAIR GESSE derived estimates of the CMRO<sub>2</sub> response to CO<sub>2</sub> if, instead of defining a finite prior uncertainty on  $V_e$  and  $\Delta\nu$ , the values of these parameters are fixed at 0.035 and 5Hz, respectively, comparable to the values reported in (He and Yablonskiy, 2006) and (Dickson et al., 2010). Under these conditions, the estimates derived from standard and FLAIR GESSE become highly consistent.

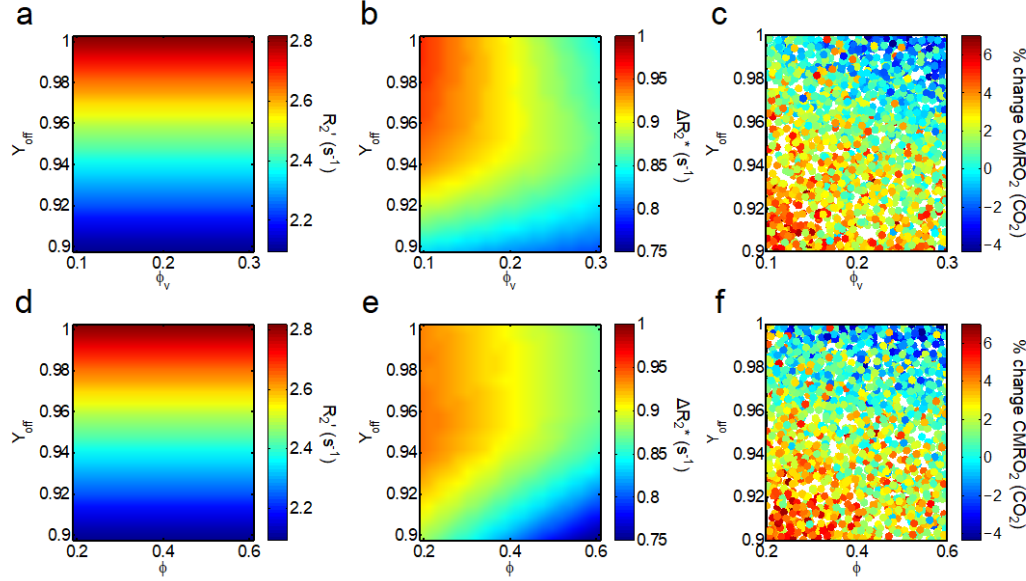


**Figure 6: Effect of an off-resonance CSF compartment on the measurement of  $R_2'$ ,  $\Delta R_2^*$  and the estimate of the CMRO<sub>2</sub> response.** (a) Simulated standard GESSE measurement of apparent  $R_2'$  for otherwise fixed model parameters as a function of CSF compartment fractional volume and off-resonance frequency. (b) Simulated FLAIR GESSE measurement of apparent  $R_2'$  for same model parameters. (c) Simulated measurement of  $\Delta R_2^*$  in response to a 50% increase in CBF and 0% increase in CMRO<sub>2</sub> as a function of CSF fractional volume and off-resonance frequency. (d) Scatter plot showing samples from the posterior distribution of the estimated CMRO<sub>2</sub> response to CO<sub>2</sub> given the sampled CSF fractional volume and off-resonance frequency based on calibration with the standard GESSE protocol. Color indicates the sample estimate. (e) Scatter plot showing samples

from the posterior distribution of the estimated CMRO<sub>2</sub> response to CO<sub>2</sub> based on calibration with the FLAIR GESSE protocol. (f) Posterior probability distributions associated with estimate of CMRO<sub>2</sub> if fractional volume and off-resonance frequency are assumed to be fixed rather than uncertain. For a 3.5% fractional volume and 5Hz off-resonance frequency, estimates based on the standard and FLAIR GESSE R<sub>2</sub>' measurements become highly comparable. Shaded bars indicate 95% central intervals. For clarity, only intrinsic uncertainty is reflected in the probability distributions.

Even after suppression of CSF contamination, the intrinsic uncertainty associated with R<sub>2</sub>' calibration is not negligible. This is because while R<sub>2</sub>' and  $\Delta R_2^*$  are similarly sensitive to variation in the baseline hematocrit, oxygenation and volume of the venous compartment (the drivers of the first-order BOLD effect) they are disparately sensitive to variation in several other unmeasured model parameters. The parameters for which this is particularly problematic are the blood-tissue iso-susceptibility parameter  $Y_{off}$ , the flow-volume coupling parameters (e.g.  $\phi$  and  $\phi_v$ ), and the parameters that describe the fundamental characteristics of the capillary bed,  $\kappa$  and  $a$ . That  $Y_{off}$ ,  $\phi$ , and  $\phi_v$  contribute significantly to posterior uncertainty in the CMRO<sub>2</sub> response is not surprising. The  $\phi$  variables describe changes in volume that accompany changes in blood flow. The measurement of R<sub>2</sub>' used to calibrate the CMRO<sub>2</sub> estimate is made in the baseline state and thus contributes no information about blood volume changes (Figures 7a and 7d).  $Y_{off}$  is a parameter that essentially increases the effective oxygen saturation in each compartment by a fixed amount. This has a large effect on the measurement of R<sub>2</sub>' (Figures 7a and 7d), because this measurement reflects the absolute quantity of “visible” deoxyhemoglobin in the volume (i.e., deoxyhemoglobin that contributes to a BOLD signal change). It has a much smaller effect on  $\Delta R_2^*$  (Figures 7b and 7e) because this measurement describes a change in deoxyhemoglobin, with both the baseline and stimulus states affected similarly by the value of  $Y_{off}$ . As a result, if  $Y_{off}$  is assumed to be close to one, then for a given baseline physiological state, more “visible” deoxyhemoglobin is expected, meaning that R<sub>2</sub>' is expected to be large. For a given measured R<sub>2</sub>',  $\Delta R_2^*$ , and  $f$ , this means that the estimate of the

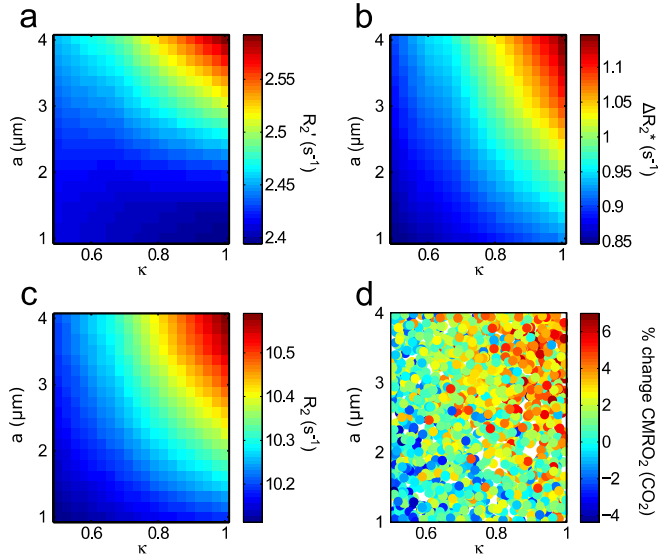
CMRO<sub>2</sub> response will be smaller. If  $Y_{off}$  is assumed to be lower, the estimate will be larger. Similarly, if  $\phi$  or  $\phi_v$  are assumed to be larger, the CMRO<sub>2</sub> response estimate must be smaller for a given set of measurements (figures 7c and 7f).



**Figure 7: Effects of  $Y_{off}$ ,  $\phi_v$ , and  $\phi$  on the measurement of  $R_2'$  and  $\Delta R_2^*$  and the estimate of the CMRO<sub>2</sub> response.** Left column, Simulated measurements of apparent  $R_2'$  for otherwise fixed model parameters as a functions of (a)  $Y_{off}$  and  $\phi_v$  or (d)  $Y_{off}$  and  $\phi$ . Middle column, Simulated measurements of  $\Delta R_2^*$  in response to a 50% increase in CBF and 0% increase in CMRO<sub>2</sub> as functions of (b)  $Y_{off}$  and  $\phi_v$  or (e)  $Y_{off}$  and  $\phi$ . Right column, scatter plots show samples from the posterior distributions of the estimated CMRO<sub>2</sub> response to CO<sub>2</sub> as functions of (c)  $Y_{off}$  and  $\phi_v$  or (f)  $Y_{off}$  and  $\phi$ . Color indicates the sample estimate.

The effects of  $\kappa$  and  $a$  on the estimate of CMRO<sub>2</sub> are more subtle. Both parameters scale the magnitude of the effect the capillary compartment has on extravascular signal decay, the first by determining the oxygen saturation of the capillary vessels, the second by defining the characteristic size of the capillary vessels. As capillary oxygen saturation drops and size increases, the rate of extravascular signal decay produced by those vessels is increased. Some of this decay is refocused by a spin echo, and is thus captured by the measurement of  $R_2'$ . However, some of this decay is not refocused, and may be thought of as more  $R_2$ -like decay. As a result of the increase in  $R_2$  decay, the  $\Delta R_2^*$  produced by a given set of model parameters increases more with increases in  $\kappa$  and  $a$  than

does  $R_2'$ . The result is that if  $\kappa$  and  $a$  are assumed to be larger, the estimated CMRO<sub>2</sub> response also will be larger (Figure 8).



**Figure 8: Effects of capillary diameter and oxygen saturation on the measurement of  $R_2'$ ,  $R_2$ , and  $\Delta R_2^*$  and the estimate of the CMRO<sub>2</sub> response.** (a) Simulated measurement of apparent  $R_2'$  for otherwise fixed model parameters as a function of capillary radius ( $a$ ) and arterial-venous oxygen saturation weighting ( $\kappa$ ). (b) Simulated measurement of  $\Delta R_2^*$  in response to a 50% increase in CBF and 0% increase in CMRO<sub>2</sub> as a function of  $a$  and  $\kappa$ . Simulated measurement of apparent  $R_2$  as a function of  $a$  and  $\kappa$ . (d) Scatter plot showing samples from the posterior distribution of the estimated CMRO<sub>2</sub> response to CO<sub>2</sub> as a function of  $a$  and  $\kappa$ . Color indicates the sample estimate.

## 5.2 Reducing intrinsic uncertainty in $R_2'$ - calibrated CMRO<sub>2</sub> estimates through CSF Suppression

In this work, we suggested that we could reduce the uncertainty associated with CSF contamination by adding a FLAIR preparation to the GESSE imaging sequence, and it is important to consider whether this FLAIR preparation achieved its intended purpose. While it is difficult to directly measure the effect of CSF contamination on signal decay, the measurements we made of  $R_2'$  with and without CSF-nulling were consistent with our imaging volume containing a CSF compartment of similar size and off-resonance frequency as measured previously (Dickson et al., 2011; He and Yablonskiy, 2006). Further, while a direct measurement of the effect size of CSF contamination on  $R_2'$  estimation has not been previously reported, a first-order comparison of other

recent  $R_2'$  estimates with our own suggests some consistency in effect size across studies and imaging methods. For example, in a 2006 study at 1.5T, Fujita et al. measured baseline apparent  $R_2'$  in the visual cortex using a gradient echo sampling of free induction decay and echo (GESFIDE) sequence without CSF-nulling (Fujita et al., 2006), calculating an average apparent  $R_2'$  of  $2.48s^{-1}$  across subjects. Because to a first order approximation, apparent  $R_2'$  is expected to be proportional to magnetic field strength (Yablonskiy, 1998), this would equate to an  $R_2'$  of  $\sim 5s^{-1}$  at 3T, the field strength in this study. In this study we measured an average apparent  $R_2'$  of  $\sim 4s^{-1}$  without and  $\sim 3s^{-1}$  with CSF-nulling. The relatively high numbers reported by Fujita et al. compared with our own are consistent with the hypothesis that CSF contributes to  $R_2'$ -like decay, with remaining quantitative discrepancies possibly explained by differences in the pulse sequences used to measure apparent  $R_2'$ . Similarly, in a recent study using asymmetrical spin echo (ASE) to measure  $R_2'$  without CSF suppression our group found an average  $R_2'$  of  $\sim 3s^{-1}$  (Blockley et al., 2015). Again, accounting for differences in the methodologies used to estimate apparent  $R_2'$  this estimate is consistent with the CSF-contaminated  $R_2'$  estimate in this work (in our simulations ASE estimates of  $R_2'$  are approximately  $\sim 1s^{-1}$  lower than those made using our current dual-GESSE technique due to the effect of approximately quadratic exponential signal decay near the spin echo (Yablonskiy, 1998) – data not shown).

It is possible that some of the difference in  $R_2'$  that was observed between the GESSE and FLAIR GESSE-based measurements is attributable not to CSF contamination, but to white matter contamination. White matter has a  $T_1$  of  $\sim 800ms$ , and thus has increased  $T_1$  weighting in the FLAIR GESSE images vs. the GESSE images (1.56 vs. 1.13 relative to gray matter based on Equation 10) (Lu et al., 2005; Wansapura et al., 1999). Because blood volume is significantly lower in white matter than gray matter, significant white matter contamination of a nominally gray matter voxel will decrease the apparent  $R_2'$ , and this effect should be magnified in the FLAIR GESSE images relative

to the GESSE images. We considered the effect size that white matter contamination could have on the measurement of  $R_2'$  and  $\Delta R_2^*$ , as well as the uncertainty associated with estimating the CMRO<sub>2</sub> response, in order to determine whether a white matter compartment of uncertain volume needed to be included in our model. Inline Supplemental Figure 4 displays the difference between simulated  $R_2'$  measurements for GESSE AND FLAIR GESSE-based measurements as a function of increasing white matter contamination. In these simulations, significant white matter contamination explained a small portion of the observed difference in apparent  $R_2'$  between GESSE and FLAIR GESSE measurements, but concomitant CSF contamination was required to fully explain the observed difference. Further, because white matter contamination affects  $\Delta R_2^*$  measurements similarly to  $R_2'$  measurements, it is less likely to contribute strongly to uncertainty in the CMRO<sub>2</sub> response (Inline Supplemental Figure 4).

These findings lend support to the idea that CSF partial-volume effects have an effect on signal decay that is reasonably well-approximated by the model proposed in (He and Yablonskiy, 2006), which was adopted here, and suggest that CSF-suppression should be included in protocols designed to measure  $R_2'$ . However, it will be important in the future to consider whether the highly simplified CSF compartment model adopted here and in previous work is sufficient to describe the range of effects such a compartment could have on the MR signal.

### **5.3 Sources of intrinsic uncertainty in CO<sub>2</sub> – calibrated CMRO<sub>2</sub> estimates**

Breathing air containing elevated concentrations of CO<sub>2</sub> was the original technique proposed for calibrating BOLD studies and is still the most commonly used today. Though performing this calibration experiment is challenging and may be contraindicated in some patient populations, we found here that it has the advantage of being much less susceptible to the sources of uncertainty that affect  $R_2'$  calibration, likely because variation in the unmeasured parameters of the model affects the

measurement of apparent  $\Delta R_2^*$  similarly in both the calibration and activation experiment. We did find, however, that this technique is highly sensitive to the assumption that the response to  $\text{CO}_2$  is iso-metabolic, such that even slight prior uncertainty in this response eliminates its advantage over  $R_2'$ . This could be important to an investigator interested in measuring the  $\text{CMRO}_2$  response to a neural stimulus in a population that could be metabolically sensitive to  $\text{CO}_2$  or who is interested in the response to  $\text{CO}_2$  itself.

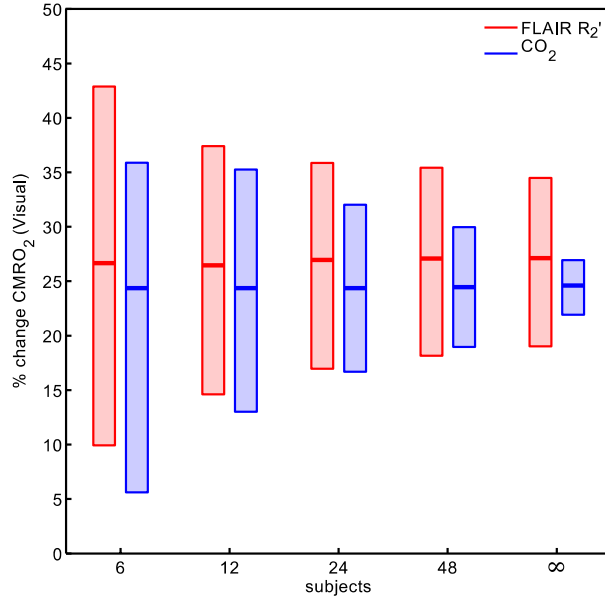
#### **5.4 Comparability of $R_2'$ and $\text{CO}_2$ calibration**

In this study we found the median estimates of the  $\text{CMRO}_2$  response to the visual stimulus based on  $R_2'$  and  $\text{CO}_2$  calibration to be highly comparable (26% and 24%, respectively). Similarly we found the median  $R_2'$ -calibrated estimate of the  $\text{CMRO}_2$  response to  $\text{CO}_2$  (~1.5%) to be consistent with the assumption that the  $\text{CO}_2$  response is approximately iso-metabolic. This is encouraging because it suggests that these two calibration techniques, at least on average, provide comparable calibration and thus have the potential to be used interchangeably based on the experimental needs of the investigator.

Interestingly, while the intrinsic uncertainty associated with hypercapnia calibration (taking the iso-metabolic assumption to be valid) was found to be significantly lower than the intrinsic uncertainty associated with  $R_2'$  calibration, the absolute uncertainty of the two estimates was highly comparable in this study. To get an idea of how many subjects would be required before hypercapnia calibration would demonstrate a significant advantage over  $R_2'$  calibration given the measurement variance and noise model employed in this study, we repeated the estimates of the  $\text{CMRO}_2$  response for each modality, assuming the same measurement variances that were found here but an increased sample size. Figure 9 shows simulated estimates of the median and 95% central interval for each approach as a function of increasing sample size. For  $R_2'$  calibration, absolute uncertainty became



dominated by intrinsic uncertainty for samples greater than approximately 24 subjects, while for hypercapnia, uncertainty continued to decrease for larger samples. This suggests that hypercapnia calibration may demonstrate more significant advantages over  $R_2'$  calibration in studies with large cohorts, provided that the iso-metabolic assumption is valid given the experimental conditions.



**Figure 9: Simulated effect of sample size on uncertainty for  $R_2'$  - and  $CO_2$  - calibrated  $CRMO_2$  response estimates.** Simulated medians and 95% central intervals are plotted for FLAIR  $R_2'$  (red) and  $CO_2$  (blue) based calibration as a function of sample size assuming constant measurement variance across sample sizes.

## 5.5 Study limitations

There are several limitations to this study. First, as we have acknowledged above, while we have made an effort here to incorporate into our model of blood oxygenation sensitive signal decay parameters that capture the most salient features of a realistic imaging volume, including multiple intravascular compartments, an extravascular parenchymal compartment, and a contaminating CSF compartment, we have by necessity greatly reduced the complexity of a true vascular network. The decisions we made in this process have introduced an additional source of uncertainty into our

estimates, uncertainty that is difficult to quantify without a gold standard for comparison. As efforts continue, for now in animal models, to better understand and describe the vascular structure and characteristics of brain tissue (Gagnon et al., 2015), resulting model improvements should be readily incorporable into the framework we have presented here.

Second, in choosing particular prior probability distributions for each of the parameters in this model, we have undoubtedly influenced the estimated posterior probability distribution of the CMRO<sub>2</sub> response. In choosing prior distributions, we often had few reports from the literature with which to constrain our decisions, and so typically simply tried to make each distribution broad enough that it would encompass reported estimates by a wide margin. As such, we recognize that some might consider the prior constraints on some of the variables too tight and others too broad. In addition, some might consider it inappropriate to make the priors on each parameter independent, when it could be reasonable to assume that some, such as intravascular compartment volumes, are correlated. In this work, we are not advocating the use of these particular priors in future studies and would certainly expect that if another investigator had additional information to constrain one or more of them further, that she would do so. Rather we are describing an analysis framework within which an investigator may explicitly state her assumptions about each model parameter in terms of prior probability distributions and account for the uncertainty in her findings given those stated assumptions. In addition, we have identified several parameters of which prior uncertainty strongly influences posterior uncertainty and which therefore deserve focused study if suspected to vary systematically across different populations of interest. For example  $Y_{off}$  is dependent upon the concentration of non-heme iron both in blood and in brain tissue and thus could change significantly with aging (Zecca et al., 2004), while flow-volume coupling constants and capillary characteristics could be affected by vascular disease.

Third, as is the case with any statistical analysis, the model we have chosen to represent the uncertainty in our measurements influences the level of uncertainty in our estimates of the CMRO<sub>2</sub>

response. Because we did not make sufficient independent measurements at the individual subject level to confidently estimate measurement variance, we chose to assume that the variance in measurements across subjects was attributable to measurement error with a distribution that was Gaussian and independent across measurements. This may prove to be a somewhat conservative approach to modeling this variability, as it is likely that some of the variance across measurements is correlated (e.g. subjects with a large CBF response also have a large BOLD response) and that accounting for this correlation would reduce the uncertainty in our final estimates. Given enough repeated and independent measurements at the individual subject level to estimate measurement error, a more optimal approach might be to apply the approach described here to the analysis of each individual subject and then to combine the resulting posterior probability distributions within a hierarchical model to determine the group posterior probability distribution. Such an approach would, of course, require additional consideration of how to treat the prior uncertainty of the unmeasured parameters, for example whether a given parameter should be assumed to have an uncertain but constant value across the experimental population, or whether it should be assumed to vary across the group. We are investigating the effects of such assumptions in an upcoming study, but felt that such an investigation was beyond the scope of this present work, which was primarily to investigate the intrinsic sources of uncertainty in estimates of the CMRO<sub>2</sub> response.

## **5.6 Use of detailed biophysical models in experimental analysis**

In order to examine the sources of uncertainty in our estimate of the CMRO<sub>2</sub> response, we developed a detailed model that explicitly takes into account what we believe to be the salient physiological parameters of the BOLD response. The model itself is not particularly novel, as it is built from the collected work of many previous investigators. However, our application of such a detailed model to the analysis of experimental data, instead of a more heuristic model such as the Davis model or Griffeth model (Davis et al., 1998; Griffeth et al., 2013), is, to our knowledge, unique

to the calibrated BOLD literature. There are both advantages and disadvantages to this approach. We found the chief disadvantage to be the computational time required to invert the model in order to estimate the CMRO<sub>2</sub> response from the BOLD and CBF measurements, which is trivial with a model such as the Davis model. Sampling the posterior distribution of the CMRO<sub>2</sub> response, which we implemented in MATLAB, took between 15 minutes and an hour depending on the calibration experiment used and the number of samples needed to stabilize the estimate of the 95% central interval. Improvements in computational time are very likely obtainable with more a efficient sampling algorithm; however, for investigators interested in making voxel-wise estimates, time costs could prove to be an important consideration.

However, this approach also offers several important advantages. First, as we have discussed at length here, it allows an investigator to explicitly state her assumptions about the parameters of the model, incorporate any experimental information she deems appropriate to constrain these parameters, and state her uncertainty about her conclusions explicitly in terms of the assumptions she has made and the measurements she has acquired. Second, the more flexible framework of the detailed model makes it possible to analyze experimental data that cannot be easily handled by a more heuristic model. For example, the Davis model considers only the effects of a single venous-like compartment on signal decay, implicitly assuming that more saturated arterial blood does not contribute significantly to the BOLD response. This is not an unreasonable assumption in a population of young, healthy subjects whose arterial blood is near the iso-susceptibility oxygen saturation ( $Y_{off}$ ). However, it is unclear how this model would accommodate a population of hypoxemic subjects, whose arterial blood could contribute significantly to the BOLD response. Investigators interested in studying subjects with cardiovascular or pulmonary disease, or who are interested in the effects of altitude on CMRO<sub>2</sub>, may find a detailed model more suitable for analyzing their experimental data.

Third, this approach allows the relatively simple incorporation of many sources of data that are challenging to integrate into existing heuristic models. For example, an investigator could measure each subject's hematocrit and integrate this information directly into the model if it were thought to vary systematically across populations of interest. Alternatively, an investigator could employ one of several recently developed techniques (e.g. TRUST, VSEAN, QUIXOTIC, PROM) to directly measure baseline venous oxygen saturation in order to further constrain the estimate of baseline OEF (Bolar et al., 2011; Fan et al., 2012; Guo and Wong, 2012; Lu and Ge, 2008).

## **5.7 Conclusions**

The need to use biophysical models with unmeasured parameters in order to estimate the CMRO<sub>2</sub> response to a neural stimulus from a calibrated BOLD experiment introduces uncertainty into the estimated response. We have described here an approach to accounting for this uncertainty and examined the principal sources of uncertainty associated with two promising methods of calibration, measurement of baseline apparent  $R_2'$  and measurement of the CBF and  $R_2^*$  responses to CO<sub>2</sub> inhalation. We found that these two calibration techniques provided consistent estimates of the CMRO<sub>2</sub> response on average. In addition we found that while  $R_2'$  calibration was sensitive to several of the unmeasured parameters in our model, CO<sub>2</sub> calibration was sensitive primarily to the assumed CMRO<sub>2</sub> response to the calibrating CO<sub>2</sub> stimulus. Understanding the sources of uncertainty in CMRO<sub>2</sub> response estimates should help guide experimental design in future calibrated BOLD experiments. In addition, the greater flexibility of a detailed model should facilitate the application of the calibrated BOLD approach to the study of populations in which cerebral vascular physiology may differ from that of normal controls.

## **6 Acknowledgements**

We would like to thank Zachary Smith for his assistance with performing the hypercapnia experiments for this project. This work was supported by NIH grants NS036722, NS085478, NS081405, NS053934, NS075812 and EB000790.

## 7 References

- 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.
- Barzilay, Z., Britten, A.G., Koehler, R.C., Dean, J.M., Traystman, R.J., 1985. Interaction of CO<sub>2</sub> and ammonia on cerebral blood flow and O<sub>2</sub> consumption in dogs. *Am. J. Physiol.* 248, H500–7.
- Blockley, N.P., Griffeth, V.E.M., Buxton, R.B., 2012. A general analysis of calibrated BOLD methodology for measuring CMRO<sub>2</sub> responses: Comparison of a new approach with existing methods. *NeuroImage* 60, 279–289. doi:10.1016/j.neuroimage.2011.11.081
- Blockley, N.P., Griffeth, V.E.M., Simon, A.B., Dubowitz, D.J., Buxton, R.B., 2015. Calibrating the BOLD response without administering gases: comparison of hypercapnia calibration with calibration using an asymmetric spin echo. *NeuroImage* 104, 423–429. doi:10.1016/j.neuroimage.2014.09.061
- Bolar, D.S., Rosen, B.R., Sorensen, A.G., Adalsteinsson, E., 2011. QUantitative Imaging of eXtraction of oxygen and TIssue consumption (QUIXOTIC) using venular-targeted velocity-selective spin labeling. *Magn. Reson. Med.* 66, 1550–1562. doi:10.1002/mrm.22946
- Boynton, G.M., Engel, S.A., Glover, G.H., Heeger, D.J., 1996. Linear systems analysis of functional magnetic resonance imaging in human V1. *J. Neurosci.* 16, 4207–4221.
- Buxton, R.B., 2013. The physics of functional magnetic resonance imaging (fMRI). *Rep. Prog. Phys.* 76, 096601. doi:10.1088/0034-4885/76/9/096601
- Buxton, R.B., Griffeth, V.E.M., Simon, A.B., Moradi, F., 2014. Variability of the coupling of blood flow and oxygen metabolism responses in the brain: a problem for interpreting BOLD studies but potentially a new window on the underlying neural activity. *Frontiers in Neuroscience* 8. doi:10.3389/fnins.2014.00139
- Chalela, J.A., Alsop, D.C., Gonzalez-Atavales, J.B., Maldjian, J.A., Kasner, S.E., Detre, J.A., 2000. Magnetic resonance perfusion imaging in acute ischemic stroke using continuous arterial spin labeling. *Stroke* 31, 680–687.
- Chen, J.J., Pike, G.B., 2009. BOLD-specific cerebral blood volume and blood flow changes during neuronal activation in humans. *NMR Biomed* n/a–n/a. doi:10.1002/nbm.1411
- Chen, J.J., Pike, G.B., 2010. Global cerebral oxidative metabolism during hypercapnia and hypocapnia in humans: implications for BOLD fMRI. *Journal of Cerebral Blood Flow & Metabolism* 30, 1094–1099. doi:10.1038/jcbfm.2010.42
- Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173.
- Davis, T.L., Kwong, K.K., Weisskoff, R.M., Rosen, B.R., 1998. Calibrated functional MRI: mapping the dynamics of oxidative metabolism. *Proc. Natl. Acad. Sci. U.S.A.* 95, 1834–1839.
- Detre, J.A., Leigh, J.S., Williams, D.S., Koretsky, A.P., 1992. Perfusion imaging. *Magn. Reson. Med.* 23, 37–45.
- Dickson, J.D., Ash, T.W.J., Williams, G.B., Harding, S.G., Carpenter, T.A., Menon, D.K., Ansorge, R.E., 2010. Quantitative BOLD: The effect of diffusion. *J. Magn. Reson. Imaging* 32, 953–961. doi:10.1002/jmri.22151
- Dickson, J.D., Ash, T.W.J., Williams, G.B., Sukstanskii, A.L., Ansorge, R.E., Yablonskiy, D.A.,

2011. Quantitative phenomenological model of the BOLD contrast mechanism. *Journal of Magnetic Resonance* 212, 17–25. doi:10.1016/j.jmr.2011.06.003
- Fan, A.P., Benner, T., Bolar, D.S., Rosen, B.R., Adalsteinsson, E., 2012. Phase-based regional oxygen metabolism (PROM) using MRI. *Magn. Reson. Med.* 67, 669–678. doi:10.1002/mrm.23050
- Fujita, N., Matsumoto, K., Tanaka, H., Watanabe, Y., Murase, K., 2006. Quantitative study of changes in oxidative metabolism during visual stimulation using absolute relaxation rates. *NMR Biomed* 19, 60–68. doi:10.1002/nbm.1001
- Gagnon, L., Sakadzic, S., Lesage, F., 2015. Quantifying the microvascular origin of BOLD-fMRI from first principles with two-photon microscopy and an oxygen-sensitive nanoprobe. ... of Neuroscience.
- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2014. *Bayesian Data Analysis*, Third Edition. CRC Press, Boca Raton.
- Glover, G.H., Li, T.Q., Ress, D., 2000. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn. Reson. Med.* 44, 162–167.
- Griffeth, V.E.M., Blockley, N.P., Simon, A.B., Buxton, R.B., 2013. A New Functional MRI Approach for Investigating Modulations of Brain Oxygen Metabolism. *PLoS ONE* 8, e68122. doi:10.1371/journal.pone.0068122
- Griffeth, V.E.M., Buxton, R.B., 2011. A theoretical framework for estimating cerebral oxygen metabolism changes using the calibrated-BOLD method: Modeling the effects of blood volume distribution, hematocrit, oxygen extraction fraction, and tissue signal properties on the BOLD signal. *NeuroImage* 58, 198–212. doi:10.1016/j.neuroimage.2011.05.077
- Grubb, R.L., Raichle, M.E., Eichling, J.O., Ter-Pogossian, M.M., 1974. The effects of changes in PaCO<sub>2</sub> on cerebral blood volume, blood flow, and vascular mean transit time. *Stroke* 5, 630–639.
- Guo, J., Wong, E.C., 2012. Venous oxygenation mapping using velocity-selective excitation and arterial nulling. *Magn. Reson. Med.* 68, 1458–1471. doi:10.1002/mrm.24145
- Haynes, W.M. (Ed.), 2014. *CRC Handbook of Chemistry and Physics*, 94th Edition (Internet Version 2014). CRC Press/Taylor and Francis, Boca Raton, FL.
- He, X., Yablonskiy, D.A., 2006. Quantitative BOLD: Mapping of human cerebral deoxygenated blood volume and oxygen extraction fraction: Default state. *Magn. Reson. Med.* 57, 115–126. doi:10.1002/mrm.21108
- He, X., Yablonskiy, D.A., 2009. Biophysical mechanisms of phase contrast in gradient echo MRI. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13558–13563. doi:10.1073/pnas.0904899106
- Horvath, I., Sandor, N.T., Ruttner, Z., McLaughlin, A.C., 1994. Role of nitric oxide in regulating cerebrocortical oxygen consumption and blood flow during hypercapnia. *Journal of Cerebral Blood Flow & Metabolism* 14, 503–509. doi:10.1038/jcbfm.1994.62
- Kida, I., Kennan, R.P., Rothman, D.L., Behar, K.L., Hyder, F., 2000. High-Resolution CMRO<sub>2</sub> Mapping in Rat Cortex: A Multiparametric Approach to Calibration of BOLD Image Contrast at 7 Tesla. *Journal of Cerebral Blood Flow & Metabolism* 20, 847–860. doi:10.1097/00004647-200005000-00012
- Lin, C., Bernstein, M., Huston, J., Fain, S., 2001. Measurements of T1 Relaxation times at 3.0T: Implications for clinical MRA. *Proceedings of the 9th Annual Meeting of ISMRM*, Glasgow, 2001 1391.
- Liu, T.T., Wong, E.C., 2005. A signal processing model for arterial spin labeling functional MRI. *NeuroImage* 24, 207–215. doi:10.1016/j.neuroimage.2004.09.047
- Lu, H., Clingman, C., Golay, X., van Zijl, P.C.M., 2004. Determining the longitudinal relaxation time (T1) of blood at 3.0 Tesla. *Magn. Reson. Med.* 52, 679–682. doi:10.1002/mrm.20178
- Lu, H., Ge, Y., 2008. Quantitative evaluation of oxygenation in venous vessels using T2-Relaxation-

- Under-Spin-Tagging MRI. *Magn. Reson. Med.* 60, 357–363. doi:10.1002/mrm.21627
- Lu, H., Nagae-Poetscher, L.M., Golay, X., Lin, D., Pomper, M., van Zijl, P.C.M., 2005. Routine clinical brain MRI sequences for use at 3.0 Tesla. *J. Magn. Reson. Imaging* 22, 13–22. doi:10.1002/jmri.20356
- Mark, C.I., Fisher, J.A., Pike, G.B., 2011. Improved fMRI calibration: Precisely controlled hyperoxic versus hypercapnic stimuli. *NeuroImage* 54, 1102–1111. doi:10.1016/j.neuroimage.2010.08.070
- Nicoll, D., Lu, C., Pignone, M., McPhee, S.J. (Eds.), 2012. *Pocket Guide to Diagnostic Tests*, 6 ed. McGraw-Hill, New York.
- Noll, D.C., Fessler, J.A., Sutton, B.P., 2005. Conjugate phase MRI reconstruction with spatially variant sample density correction. *IEEE Trans. Med. Imaging* 24, 325–336. doi:10.1109/TMI.2004.842452
- Ogawa, S., Menon, R.S., Tank, D.W., Kim, S.G., Merkle, H., Ellermann, 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. doi:10.1016/S0006-3495(93)81441-3
- Ogawa, S., Tank, D.W., Menon, R., Ellermann, J.M., Kim, S.G., Merkle, H., Ugurbil, K., 1992. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proc. Natl. Acad. Sci. U.S.A.* 89, 5951–5955.
- Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10, 437–442.
- Perthen, J.E., Lansing, A.E., Liao, J., Liu, T.T., Buxton, R.B., 2008. Caffeine-induced uncoupling of cerebral blood flow and oxygen metabolism: A calibrated BOLD fMRI study. *NeuroImage* 40, 237–247. doi:10.1016/j.neuroimage.2007.10.049
- Restom, K., Behzadi, Y., Liu, T.T., 2006. Physiological noise reduction for arterial spin labeling functional MRI. *NeuroImage* 31, 1104–1115. doi:10.1016/j.neuroimage.2006.01.026
- Sakai, F., Igarashi, H., Suzuki, S., Tazaki, Y., 1989. Cerebral blood flow and cerebral hematocrit in patients with cerebral ischemia measured by single-photon emission computed tomography. *Acta Neurol. Scand., Suppl.* 127, 9–13.
- Sicard, K.M., Duong, T.Q., 2005. Effects of hypoxia, hyperoxia, and hypercapnia on baseline and stimulus-evoked BOLD, CBF, and CMRO<sub>2</sub> in spontaneously breathing animals. *NeuroImage* 25, 850–858. doi:10.1016/j.neuroimage.2004.12.010
- Spees, W.M., Yablonskiy, D.A., Oswood, M.C., Ackerman, J.J., 2001. Water proton MR properties of human blood at 1.5 Tesla: Magnetic susceptibility, T<sub>1</sub>, T<sub>2</sub>, T<sub>2</sub><sup>\*</sup>, and non-Lorentzian signal behavior. *Magn. Reson. Med.* 45, 533–542.
- Stefanovic, B., Hutchinson, E., Yakovleva, V., Schram, V., Russell, J.T., Belluscio, L., Koretsky, A.P., Silva, A.C., 2008. Functional reactivity of cerebral capillaries. *Journal of Cerebral Blood Flow & Metabolism* 28, 961–972. doi:10.1038/sj.jcbfm.9600590
- Uludağ, K., Müller-Bierl, B., Ugurbil, K., 2009. An integrative model for neuronal activity-induced signal changes for gradient and spin echo functional imaging. *NeuroImage* 48, 150–165. doi:10.1016/j.neuroimage.2009.05.051
- Wang, J., Qiu, M., Constable, R.T., 2005. In vivo method for correcting transmit/receive nonuniformities with phased array coils. *Magn. Reson. Med.* 53, 666–674. doi:10.1002/mrm.20377
- Wansapura, J.P., Holland, S.K., Dunn, R.S., Ball, W.S., 1999. NMR relaxation times in the human brain at 3.0 tesla. *J. Magn. Reson. Imaging* 9, 531–538.
- 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., 1998. Quantitative imaging of perfusion using a single subtraction (QUIPSS and QUIPSS II). *Magn. Reson. Med.* 39, 702–708.
- Xu, F., Uh, J., Brier, M.R., Hart, J., Yezhuvath, U.S., Gu, H., Yang, Y., Lu, H., 2011. The influence of carbon dioxide on brain activity and metabolism in conscious humans. *Journal of Cerebral Blood Flow & Metabolism* 31, 58–67. doi:10.1038/jcbfm.2010.153
- Yablonskiy, D.A., 1998. Quantitation of intrinsic magnetic susceptibility-related effects in a tissue matrix. Phantom study. *Magn. Reson. Med.* 39, 417–428.
- Yablonskiy, D.A., Haacke, E.M., 1994. Theory of NMR signal behavior in magnetically inhomogeneous tissues: the static dephasing regime. *Magn. Reson. Med.* 32, 749–763.
- Yablonskiy, D.A., Haacke, E.M., 1997. An MRI method for measuring T2 in the presence of static and RF magnetic field inhomogeneities. *Magn. Reson. Med.* 37, 872–876.
- Yang, S.P., Krasney, J.A., 1995. Cerebral blood flow and metabolic responses to sustained hypercapnia in awake sheep. *Journal of Cerebral Blood Flow & Metabolism* 15, 115–123. doi:10.1038/jcbfm.1995.13
- Zappe, A.C., Uludağ, K., Oeltermann, A., Ugurbil, K., Logothetis, N.K., 2008. The Influence of Moderate Hypercapnia on Neural Activity in the Anesthetized Nonhuman Primate. *Cerebral Cortex* 18, 2666–2673. doi:10.1093/cercor/bhn023
- Zecca, L., Youdim, M.B.H., Riederer, P., Connor, J.R., Crichton, R.R., 2004. Iron, brain ageing and neurodegenerative disorders. *Nat Rev Neurosci* 5, 863–873. doi:10.1038/nrn1537
- Zhao, J.M., Clingman, C.S., Närväinen, M.J., Kauppinen, R.A., van Zijl, P.C.M., 2007. Oxygenation and hematocrit dependence of transverse relaxation rates of blood at 3T. *Magn. Reson. Med.* 58, 592–597. doi:10.1002/mrm.21342

## 8 Tables

Variable				
Parameters				
Parameter	Parameter	Prior	Literature	Reference(s)
Symbol	Significance	Distribution	Value(s)	
$V_{a0}$	Fractional baseline	U(0.005,0.015)	0.01	Roland et al., 1987;
	arterial blood volume			Weber et al., 2008
$V_{c0}$	Fractional baseline	U(0.01,0.03)	0.02	Roland et al., 1987;
	capillary blood volume			Weber et al., 2008
$V_{v0}$	Fractional baseline	U(0.01,0.03)	0.02	Roland et al., 1987;
	venous blood volume			Weber et al., 2008

$V_{e0}$	Baseline CSF volume	U(0,0.1)	0.045, 0.049	He and Yablonskiy, 2006; Dickson et al., 2010
$\phi$	exponential CBF-CBV coupling constant	U(0.2,0.6)	0.4	Grubb et al., 1974
$\phi_v$	exponential CBF- $V_v$ coupling constant	U(0.1,0.3)	0.2	Chen and Pike, 2009
$\phi_c$	exponential CBF- $V_c$ coupling constant	U(0.1,0.3)	0.2	Stefanovic et al., 2008
$Y_{off}$	blood-tissue isosusceptibility saturation	U(0.9,1)	0.89, 0.95, 1	He et al., 2009; Uludağ et al., 2009; He and Yablonskiy, 2006
$\kappa$	weighting constant for capillary $O_2$ sat.	U(0.5,1)	0.6	Tsai et al., 2003
$a$	characteristic radius of capillary vessels	U(1 $\mu$ m,4 $\mu$ m)	2.5	Ogawa et al., 1993
<b>Hct</b>	Hematocrit	U(0.35,0.5)	Male: 0.39-0.49, Female: 0.35-0.45	Nicoll et al., 2012
$\Delta\nu$	CSF-tissue off-resonance frequency	U(2Hz,8Hz)	5Hz, 7Hz	He and Yablonskiy, 2006; Dickson et al.,

				2010
$R_{2t0}$	Non-hemoglobin tissue $T_2$ relaxation rate constant	$U(7.5s^{-1}, 12.5s^{-1})$	$7.5s^{-1}, 12.5s^{-1}$	Lu et al., 2005; Wansapura et al., 1999
$OE_{F_0}$	Oxygen Extraction Fraction in Baseline State	$U(0.2, 0.5)$	0.25-0.5	Lu and Ge, 2008
$OE_{F_{stim}}$	Oxygen Extraction Fraction in stimulus State	$U(0.05, 0.95)$	NA	NA

## Fixed

### Parameters

Parameter Symbol	Parameter Significance	Value Used	Literature Value(s)	Reference(s)
$\rho_t$	brain parenchyma water density	0.84	0.84	Herscovitch and Raichle, 1985
$\rho_e$	CSF water density	1	1	NA
$\rho_{a,c,v}$	blood water density	0.87	0.87	Herscovitch and Raichle, 1985
$T_{1p}$	brain parenchyma $T_1$	1200ms	1122ms, 1283ms	Lu et al., 2005; Wansapura et al. 1999
$T_{1e}$	CSF $T_1$	4000ms	3817ms,	Lu et al., 2005; Lin

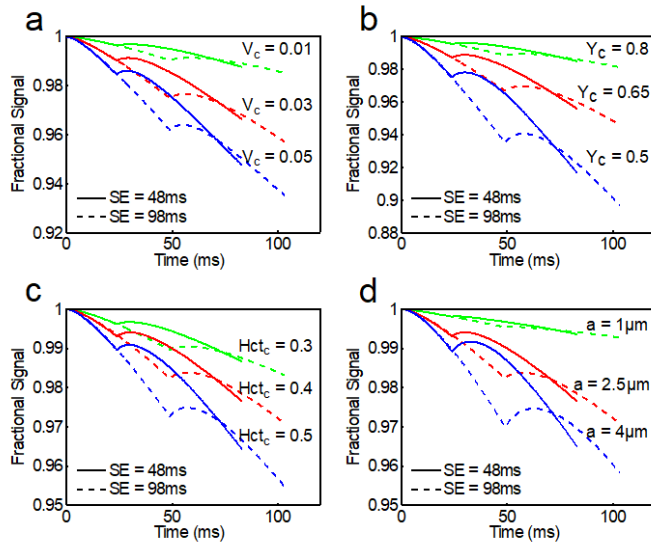
			4163ms	et al, 2001
$\Delta\chi_0$	susceptibility difference: de/oxygenated hemoglobin	0.264ppm	0.264ppm	Spees et al., 2001
$\gamma$	gyromagnetic ratio	$2.675 \times 10^8 \text{ rad ms}^{-1} \text{ T}^{-1}$	$2.675 \times 10^8 \text{ rad ms}^{-1} \text{ T}^{-1}$	Haynes, 2014
$B_0$	Magnetic field strength	3T	NA	NA
$D$	Diffusion constant for water in tissue	$1 \mu\text{m}^2 \text{ ms}^{-1}$	$1 \mu\text{m}^2 \text{ ms}^{-1}$	Ogawa et al., 1993

**Table 1:** Parameter Names and Values

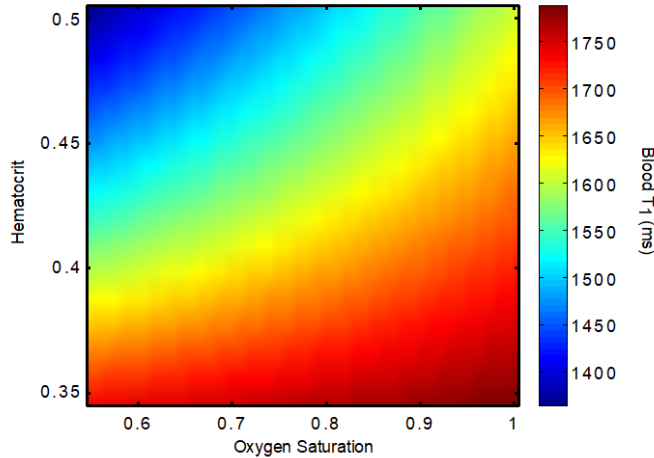
Measurement	Sub 1	Sub 2	Sub 3	Sub 4	Sub 5	Sub 6	mean	std
$R_2' \text{ (s}^{-1}\text{)}$	3.31	3.58	3.50	4.21	5.22	3.84	3.94	0.64
FLAIR $R_2' \text{ (s}^{-1}\text{)}$	2.56	3.34	3.29	2.67	3.68	2.74	3.05	0.41
$\Delta R_2^* \text{-CO}_2 \text{ (s}^{-1}\text{)}$	-0.89	-0.71	-0.57	-0.59	-0.36	-0.67	-0.63	0.16
$\Delta \text{CBF-CO}_2 \text{ (\%)}$	17.82	24.78	23.10	30.10	29.57	17.40	23.79	5.02
$\Delta R_2^* \text{-Visual (s}^{-1}\text{)}$	-0.68	-0.89	-0.64	-0.46	-0.97	-0.79	-0.74	0.17
$\Delta \text{CBF-Visual (}\% \text{)}$	43.30	72.85	60.46	87.17	90.57	60.11	69.08	16.44
$Y_a$	0.98	0.98	0.97	0.99	1.00	1.00	0.99	0.01
age	24	28	23	29	22	23	24.83	2.67

**Table 2:** Summary of Experimental Measurements

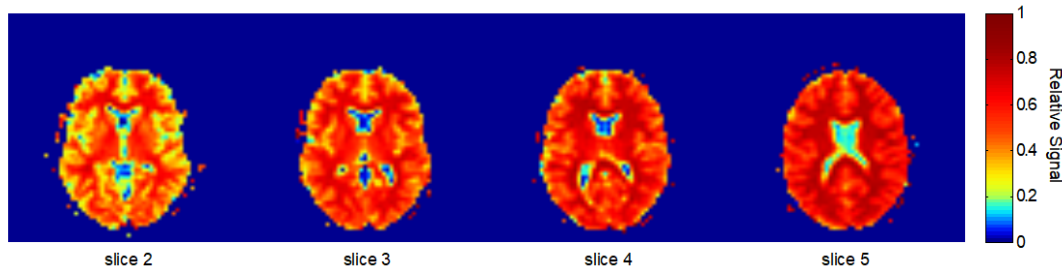
## 9 Inline Supplemental Figure Captions



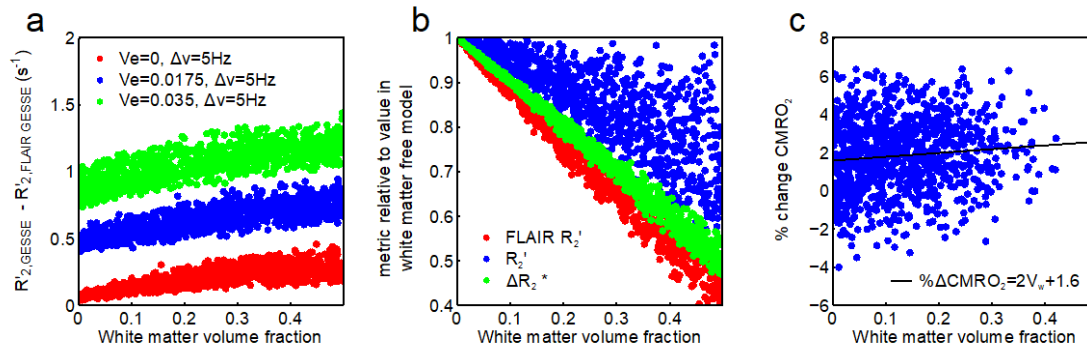
**Inline Supplemental Figure 1: Simulated tissue signal decay due to capillary compartment.** Decay curves simulate capillary contribution to tissue signal decay for spin echo experiments with spin echo times of 48ms (solid lines) and 98ms (dashed lines) as a function of (a) capillary volume ( $V_c$ ), (b) oxygen saturation ( $Y_c$ ), (c) hematocrit ( $Hct_c$ ), and (d) characteristic capillary radius ( $a$ ). Unless otherwise specified,  $V_c = 0.01$ ,  $Y_c = 0.75$ ,  $Hct_c = 0.4$ ,  $a = 2.5\mu m$ .



**Inline Supplemental Figure 2: Assumed Blood  $T_1$  values based on oxygen saturation and hematocrit.** The  $T_1$  of blood was calculated for each vascular compartment as a function of hematocrit and oxygen saturation using linear interpolation of the functions described by (Lu et al., 2004).



**Inline Supplemental Figure 3: CSF suppression in FLAIR GESSE images.** Images from representative subject show signal intensity in FLAIR GESSE images at spin echo time of 48ms as a fraction of signal intensity in standard GESSE images. Consistent with prediction based on Equation 10, FLAIR GESSE signal within the lateral ventricles is <20% of the standard GESSE signal in slice 5, which was least optimally nulled.



**Inline Supplemental Figure 4: Simulated effect of white matter partial volume on measurement of apparent  $R_2'$  and  $CMRO_2$  response.** As a first approximation, the white matter volume was considered to be an essentially inert, bloodless volume of tissue that relaxed with a baseline  $T_2$  decay time of  $\sim 80$ ms but otherwise did not contribute to  $R_2'$  or  $R_2^*$  (Lu et al., 2005; Wansapura et al., 1999). (a) Simulated difference in apparent  $R_2'$  based on standard or FLAIR GESSE imaging due to white matter partial volume effects. In this simulation, all model parameters were allowed to vary as described by Table 1 except for  $V_e$  and  $\Delta v$ , which were fixed as described in the figure legend. While increasing white matter contamination did increase the difference in apparent  $R_2'$  measured by the two modalities, the effect was small compared to that produced by CSF off-resonance in this model. (b) Simulated effect of white matter partial volume on the measurement of  $R_2'$  and  $\Delta R_2^*$ . In this simulation all model parameters were allowed to vary as described by Table 1. The y-axis displays the simulated apparent value of the metric indicated in the legend for a given set of model parameters and white matter partial volume, relative to the simulated apparent value for the same set of model parameters but without a white matter partial volume. Increasing white matter partial volume decreased the apparent  $R_2'$  obtained with FLAIR GESSE and  $\Delta R_2^*$  similarly, though there was a small systematic difference in the rate of decrease due to the increased  $T_1$ -weighting of white matter in the FLAIR GESE images. The highly variable effect of white matter contamination on the  $R_2'$  obtained with GESSE was due to the effects of CSF contamination. (c) Simulated effect of white matter contamination on the estimate of the  $CMRO_2$  response to  $CO_2$  when FLAIR GESSE is used to measure  $R_2'$ . Scatter plot shows simulated samples from the posterior distribution of the estimated  $CMRO_2$  response (ignoring measurement noise) for a model including an uncertain white matter

partial volume. Increasing white matter partial volume ( $V_w$ ) produced a systematic effect on the distribution of the estimated CMRO<sub>2</sub> response ( $\% \Delta \text{CMRO}_2$ ); however, it was small even as  $V_w$  approached half of the voxel volume.