Regional reproducibility of calibrated BOLD functional MRI:
Implications for the study of cognition and plasticity

Steffen N. Krieger, Claudine J. Gauthier, Dimo Ivanov, Laurentius Huber, Elisabeth Roggenhofer, Bernhard Sehm, Robert Turner, Gary F. Egan

Abstract
Calibrated BOLD fMRI is a promising alternative to the classic BOLD contrast due to its reduced venous sensitivity and greater physiological specificity. The delayed adoption of this technique for cognitive studies may stem partly from a lack of information on the reproducibility of these measures in the context of cognitive tasks. In this study we have explored the applicability and reproducibility of a state-of-the-art calibrated BOLD technique using a complex functional task at 7 tesla. Reproducibility measures of BOLD, CBF, CMRO2 flow-metabolism coupling n and the calibration parameter M were compared and interpreted for three ROIs. We found an averaged intra-subject variation of CMRO2 of 8% across runs and 33% across days. BOLD (46% across runs, 36% across days), and M (41% across days) showed significantly higher intra-subject variability. Inter-subject variability was found to be high for all quantities, though CMRO2 was the most consistent across brain regions. The results of this study provide evidence that calibrated BOLD may be a viable alternative for longitudinal and cognitive MRI studies.

© 2014 Elsevier Inc. All rights reserved.

Introduction
Functional Magnetic Resonance Imaging (fMRI) studies generally utilise blood oxygenation level dependent (BOLD) contrast to detect spatial and temporal patterns of brain activity (Ogawa et al., 1990). Local cerebral functional activation leads to changes in deoxyhaemoglobin concentration which drive the BOLD signal. BOLD signal changes are easily detectable with relatively high sensitivity. However, this technique has been the subject of debate due to its intrinsic physiological ambiguity. BOLD contrast arises from simultaneous changes in cerebral blood flow (CBF), cerebral blood volume (CBV) and cerebral metabolic rate of oxygen (CMRO2), but the quantitative relationships between these parameters are not fully understood, and are likely to differ across the lifespan and in disease (Buxton, 2010; Kim and Ogawa, 2012). Spatial specificity of the BOLD technique is diminished by the fact that downstream draining veins may show large BOLD signal changes during neuronal activation, even when not immediately proximal to the area of activation (Turner, 2002). Furthermore, the BOLD technique offers relatively poor stability characteristics. Recent studies demonstrated that temporal characteristics and amplitude of the BOLD signal can vary across brain regions, subjects, and even within subjects across days or runs (Aguirre et al., 1998; Duff et al., 2007; Fox et al., 2005). Differences in baseline blood oxygenation might be a possible origin of the high variability of BOLD (Lu et al., 2008). Finally, statistical mapping of significant functional BOLD activation can lead to inconsistent results even on short temporal scales (Wang et al., 2003).

Cognitive neuroimaging studies typically suffer from a low signal-to-noise ratio as compared to studies looking into primary cerebral areas, since they generally rely on the application of low-contrast or complex tasks that give rise to small signal changes. At the same time such studies rely on data with good stability characteristics and high spatial specificity to enable valid conclusions about the location of neuronal activity modulations during a task. The combined effects of physiological ambiguity and poor reproducibility of the BOLD signal may therefore severely limit the conclusiveness of group studies, when intrinsic technique-related variabiliy is combined with subtle group differences and high inter-subject variability. Within-subject longitudinal studies can be even more vulnerable to poor reproducibility, because this intrinsic BOLD variability cannot be averaged out as with group studies. These effects may be compounded in the case of learning studies, where inferences about effects on brain activity are often based on single functional scans at the individual level. The assessment of learning therefore...
relies on precise spatial and temporal detection of signal changes, while poor specificity and stability can mask learning-related changes or lead to artefactual effects, arising from variability in the technique. One solution to these problems is the use of more physiologically specific functional techniques that better reflect the underlying neuronal activity.

Arterial spin labelling (ASL) utilises magnetically labelled arterial blood water as an endogenous tracer in order to measure changes in CBF during neural activation (Calamante et al., 1999; Williams et al., 1993). ASL shows good stability characteristics even over long periods of time (Chen et al., 2011; Wang et al., 2003, 2011), providing a major advantage for longitudinal studies as compared with BOLD. When ASL and BOLD data are combined, CMRO$_2$ can be estimated (Davis et al., 1998). Calibrated fMRI techniques commonly rely on the simultaneous acquisition of CBF and BOLD signal during a gas-breathing manipulation, which can be hypercapnic (Hoge et al., 1999), hyperoxic (Chiarelli et al., 2007b) or a combination of both (Gauthier and Hoge, 2013). These breathing manipulations alter the measured BOLD signal, which is then extrapolated to its theoretical physiological maximum (M) via a biophysical model. Cerebral metabolic oxygen changes can then be calculated using this M calibration constant.

The calibrated BOLD technique inherently provides superior spatial specificity and offers physiologically more meaningful measures, compared with the classical BOLD approach. Importantly, this method has been shown to provide good stability characteristics at lower field strengths with simple functional stimuli (Leontiev and Buxton, 2007). Despite its theoretical advantages, few groups have adopted calibrated fMRI for the study of cognition and plasticity.

One reason may be the greater complexity of calibrated BOLD compared with classical BOLD, in terms of MRI setup, acquisition time and data analysis. Furthermore, most functional MRI software packages do not include tools to analyse calibrated BOLD data, and there is a lack of practical guidelines available for the adoption of this technique. Finally, the stability of calibrated BOLD for low-contrast and cognitive tasks has not been assessed, making it difficult to determine the benefit that a specific cognitive neuroimaging study could obtain by using calibrated BOLD compared with the effort required to implement the technique.

This study explores the applicability and reproducibility of a state-of-the-art calibrated BOLD technique when applying complex low-contrast functional tasks at 7 tesla. Ultra-high field fMRI provides increased spatial resolution, specificity and sensitivity compared to lower magnetic field approaches (Triantafyllou et al., 2005; Uludag et al., 2009). We show that calibrated BOLD can be reliably used for fMRI studies that apply complex low-contrast stimuli, such as cognitive tasks, also at high magnetic field strengths, with stability characteristics comparable to lower field strengths (Leontiev and Buxton, 2007). While this study was based on ultra-high field measurements, our conclusions extend beyond this field strength and help establish guidelines for the implementation of calibrated fMRI for both cross-sectional and longitudinal cognitive studies. We have assessed calibrated BOLD reproducibility across subjects and across days, and across runs within the same subjects. Furthermore, these estimates were compared with CBF and BOLD-based reproducibility estimates for the same group of subjects.

Methods

Thirteen healthy participants (7 women, 27 ± 3 years) were scanned on a Siemens Magneton 7 tesla whole body MRI scanner (Siemens Healthcare, Erlangen, Germany) according to the guidelines set by the local review board (Ethics Commission, Leipzig University). A 24-channel head coil (NOVA Medical Inc, Wilmington, MA, USA) was used for all acquisitions. All participants gave written informed consent for participation in the study.

Experimental design

Three calibrated BOLD scanning sessions were performed on each subject within a period of 7 days, with a maximum total daily scanning time of 60 minutes. Each session consisted of one ten-minute run with a combined gas breathing challenge and visuo-motor task (see Fig. 1) followed by three ten-minute runs with the visuo-motor task alone. A one-minute on/off block design was used for the task, starting with one minute rest. The first ten-minute run of each day included a combined visuo-motor and gas-inhalation design. Gas inhalation started after two minutes of acquisition and lasted five minutes, superimposed on the previously-described ten-minute task design (see Fig. 1). Simultaneous acquisition of BOLD and CBF-weighted images was performed via pulsed arterial spin labelling for all functional runs.

Functional stimulus

A visuo-motor task (Grafton et al., 2008) was used in this study in order to assess the technical applicability of calibrated fMRI in cognitive and longitudinal plasticity studies. The task was implemented in Presentation (v. 12.2.2, Neurobehavioral Systems Inc.). Participants controlled a short vertical bar displayed on a screen inside of the scanner bore using a frictionless MRI-compatible joystick. The joystick was used to precisely counteract the intrinsic movement of the vertical bar, and thus maintain the bar in the centre of the screen. A double continuous Gaussian function (one negative and one positive) with an overall period of 10 seconds was used as an intrinsic movement function. This function was repeated six times resulting in 60 seconds of functional activation. This provided sufficient time for participants to identify the pattern of the movement. A blank screen was shown during the 60 s rest block. Behavioural data of each scan was recorded. Performance error was defined as the distance between the moving bar and the centre of the screen.

Gas stimulus

During the first ten-minute acquisition, a combined hypercapnia-hyperoxia gas-breathing challenge (carbogen-7 – 7% CO$_2$/93% O$_2$) was delivered, simultaneously with the visuo-motor block design, starting after two minutes of acquisition and lasting for five minutes. A non-rebreathing mouthpiece connected to a three-way valve was used to deliver gases to the participants, while separating the inflow from the air exhaled by the subject. The exhaled air was guided out of the scanner bore by a separate pipe to avoid surrounding air susceptibility changes (Streicher et al., 2012). A medical doctor was present in the scanner room at all times during the breathing manipulation and was responsible for adjusting gas flow rates and monitoring pulse rate and arterial O$_2$ saturation measured via a pulse-oximeter on the subject’s index finger. Gas flows were adjusted manually on a flow-meter to a flow rate of 15 litres per minute. In the case of hyperventilation or collapse of the reservoir bag attached to the flow-meter during gas-
breathing, flow rates were temporarily increased to maintain a constant filling of the reservoir bag. An alarm bell was provided to all participants to allow them to interrupt the experiment in case of discomfort.

End-tidal \(O_2\) and \(CO_2\) values were recorded for all subjects throughout the whole scanning procedure using the \(CO_2\) and \(O_2\) modules of the MP150 BIOPAC physiological monitoring system (BIOPAC systems Inc., Goleta, CA, USA). The BIOPAC system was positioned outside of the scanner room. Inhaled and exhaled air samples were taken from the three-way valve attached to the mouthpiece via a small flexible tube connected directly to the gas sensor of the \(O_2\) and \(CO_2\) modules of the BIOPAC system. The BIOPAC system was calibrated before every use by adjusting the input resistances for both module sensors to the partial pressures of two gas mixtures not used during the actual experiment.

The first gas used to calibrate the monitoring system was pure oxygen, giving zero partial gas pressure of \(CO_2\) and 760 mmHg partial gas pressure of \(O_2\). The second gas used consisted of 7% \(CO_2\) (53.2 mmHg partial gas pressure) and 21% of \(O_2\) (159.6 mmHg partial gas pressure).

**Scanning parameters**

BOLD- and CBF-weighted functional MRI images were acquired simultaneously using a pulsed arterial spin labelling (PASL) flow attenuation inversion recovery (FAIR) (Kim, 1995) with quantitative imaging of perfusion using a single subtraction version (QUIPSII) (Wong et al., 1998) sequence. An adiabatic inversion pulse specifically designed for 7 tesla was used to obtain maximum inversion efficiency (Tr-FOCI) (Hurley et al., 2010). A short inversion pulse duration of 3.5 ms was chosen to minimise \(T_2^*\) relaxation effects and a nominal RF peak amplitude of 17.3 G was used to obtain a maximised inversion efficiency. 25 slices with no slice gap were acquired using a repetition time (TR) of 4 seconds and an echo time (TE) of 8.1 ms. Data was acquired via a standard echo-planar imaging (EPI) readout with a 90° flip angle. A GRAPPA factor of 3 and partial Fourier factor of 6/8 were used. The nominal isotropic resolution used was 3 mm with a field of view of 192 mm × 192 mm resulting in a matrix size of 64 × 64. The inversion times were \(T_1 = 700\) ms and \(T_2 = 1700\) ms and the tag-width was 90 mm.

**Data analysis**

FMRI datasets were corrected for motion, and all 12 runs of each subject were then co-registered using MCFLIRT (FSL, FMRIB, Oxford, UK). A general linear model (GLM) analysis across all functional runs from each subject was used to define two functional regions of interest (ROI) one including primary visual and the other including primary motor areas. The fMRI expert analysis tool (FEAT) v. 5.38 from the FSL software package was used to create activation maps using a 3 mm smoothing kernel. In order to preserve spatial information smoothing was only used for the ROI generation step. Cluster thresholding was used to define significant signal changes for BOLD and CBF-based ROI selection (\(p < 0.05\), corrected for multiple comparison). The intersection of significant BOLD and CBF signal changes was used to define the visual and motor ROIs. These ROIs were chosen, since primary areas were thought to be most likely to show detectable signal changes at the level of the individual, while being somewhat more immune to complex learning or habituation effects than areas exclusively involved in the more cognitive aspects of the task. The ROI selection is similar to the ones employed in longitudinal and learning studies. An additional grey-matter ROI was defined on the basis of thresholded absolute CBF maps generated via the Neuroloens software package (www.neuroloens.org). Voxels showing absolute flow values greater than 20 ml/min/100 g tissue were defined as grey matter and included in the mask (Pohmann, 2010).

All subsequent steps of the data analysis were performed using an in-house Mathworks Matlab (version 8.0.0.783 R2012b) code. Linear temporal interpolation was performed on tag and control datasets. Voxelwise subtraction of interpolated control and label images resulted in CBF-weighted images. BOLD time series were obtained from voxelwise addition. CBF and BOLD time courses were then ROI-averaged. Because there is no literature available describing the effects of hyperoxia on \(T_2\) of arterial blood at 7 tesla, a 10% underestimation of CBF was assumed, following the assumptions of earlier studies (Bulte et al., 2007; Gauthier and Hoge, 2013; Krieger et al., 2013). A linear BOLD signal drift correction was applied for runs without gas inhalation, although these corrections were only of minor amplitude. Furthermore, we tested for signal drifts within one-minute blocks during the calibration scans. No significant signal drifts could be detected.

The calibration constant \(M\) is needed for estimation of functionally-induced changes in cerebral oxygen metabolism. This study used the generalised calibration model (GCM) to obtain \(M\), since this model is compatible with combined hypercapnia and hyperoxia breathing manipulations such as carboxen-7 inhalation (Gauthier and Hoge, 2013). Based on the GCM the \(M\)-value can be calculated as follows:

\[
M = \frac{B_{OLD}}{1-CBF_t^{1-\beta}} = \frac{1}{\Sigma^{\text{CBF}}_t} \quad \text{with} \quad \Sigma^{\text{CBF}}_t = \frac{1}{\Sigma^{\text{CBF}}_t} \quad \text{where } \beta = \frac{1}{\alpha}
\]

where \(B_{OLD}\) is the fractional change in BOLD signal and \(C_{BF}\) is the relative CBF signal due to carboxen inhalation, \(\alpha\) represents the power law exponent describing the relationship between changes in cerebral blood flow and cerebral blood volume, assumed here to be 0.18 (Chen and Pike, 2010). The exponent \(\beta\) \((\text{Driver et al., 2010; Yablonisky and Haacke, 1994})\) is used to model the influence of deoxygenated haemoglobin on transverse relaxation (Boxerman et al., 1995) and is here assumed to have a value of 1.0 due to the high magnetic field used (Kida et al., 2000; Martindale et al., 2008). \(S_{O2}\) and \(S_{O20}\) represent venous \(O_2\) saturation during gas breathing periods and baseline (air breathing) periods respectively. Venous \(O_2\) saturation values were obtained from GCM modelling with an assumed value of 0.35 for baseline oxygen extraction fraction (OEF) (Bulte et al., 2012; Carpenter et al., 1991; Gauthier and Hoge, 2012; Lu and Ge, 2008). For more detailed information on the specific calibration model used in this study refer to (Gauthier and Hoge, 2013). \(M\) was estimated on each measurement day for each subject, under the assumption that the physiological conditions remain stable during one scanning session, giving a total of three \(M\)-values for each subject. Cerebral oxygen metabolism was calculated for each of the three task-only runs, for a total of nine ROI-based \(CMRO_2\) estimates per subject. Relative changes in \(CMRO_2\) were estimated using Eq. (2) with BOLD representing the fractional signal change of BOLD and CBF the relative CBF signal during task stimulation.

\[
CMRO_2 = \left(1 - \frac{B_{OLD}}{M}\right) \frac{1}{\beta} \quad \text{with} \quad \beta = \frac{1}{\alpha}
\]

The coupling between CBF and \(CMRO_2\) is described as the index \(n\), defined as the ratio between fractional changes in CBF and the fractional changes in \(CMRO_2\) in response to a functional challenge.

**Reproducibility**

Variability in BOLD, CBF, \(S_{O2}\), \(CMRO_2\) and \(n\) within and between MRI sessions was investigated. For that purpose three coefficients of variation (CV) (Tjandra et al., 2005) were computed. These characterised a) variability across runs on the same day for the same subject (\(C_{V_1}\)), b) across days for the same subject (\(C_{V_2}\)) and c) across all runs and sessions for all subjects (\(C_{V_3}\)). The CV is defined as the standard deviation divided by the mean signal change within the specific ROI, expressed as a percentage. A graphical representation of the different acquisitions included in each CV can be found in Fig. 2.
Variability estimates for each subject were derived on a daily basis. Coefficients of variation were calculated separately for each subject and each day resulting in 3 CVs per subject.

**CV1 - Same subject and day**
Variability estimates for each subject were derived on a daily basis. Coefficients of variation were calculated separately for each subject and each day resulting in 3 CVs per subject.

**CV2 - Same subject across days**
One coefficient of variation was calculated per subject using only the first run of each day.

**CV3 - Across subjects**
All runs and sessions from all subjects and all days were used to calculate one inter-subject CV. It should also be noted that task-evoked $\text{BOLD}_c$, $\text{CBF}_c$ and $\text{CMRO}_2$ were not calculated for the grey matter ROI as zero signal changes can be expected in response to the task within this mask since all task-relevant voxels were excluded from it. A statistical t-test was applied to the pooled data from each ROI in order to determine statistical significance of the results.

**Learning and habitation effects**
To assess the impact of habituation and learning on the stability estimates, we used motor performance from each subject as an additional regressor for statistical data analysis. Improvement in motor performance was expected to lead to decreased functional signals in primary motor areas (Floyer-Lea and Matthews, 2004; Ungerleider et al., 2002). The residual signal following regression on each individual acquisition was then used to perform the variability analyses previously described. These performance-corrected CVs may be useful to infer the variability of signals in longitudinal studies where no plasticity is expected to occur.

**Results**
Complete datasets from 13 subjects were analysed. These datasets contain three scanning sessions within 7 days and four scanning runs per day, with the first being the calibration scan and the last three being functional scans. Example CBF maps can be found in Fig. 3 for each subject. A good grey matter–white matter contrast can be found across all subjects. Frontal and occipital brain areas show slightly higher flow values than the rest of the grey matter. End-tidal oxygen values increased from an average baseline of 110.8 ± 6.0 mmHg up to 588.2 ± 104.8 mmHg during carbogen-7 inhalation. Carbon-dioxide end-tidal values increased from 39.1 ± 3.0 mmHg to 53.5 ± 3.6 mmHg.

Table 1 summarises the mean values for each quantity obtained by averaging across all days, runs and subjects. Mean $\text{BOLD}_c$ and $\text{CBF}_c$ signal changes during carbogen inhalation were larger in the motor ROI than in the visual ROI ($p = 0.03$). The average $\text{M}$-value was also larger in the motor ROI than in the visual ROI ($p = 0.04$). The largest mean $\text{BOLD}_c$ signal changes and $\text{M}$-values were, however, found in the grey matter ROI ($p = 0.01$). In the same region, $\text{CBF}_c$ changes due to carbogen inhalation were somewhat smaller than for the visual and motor ROI ($p = 0.01$). A GCM extrapolation therefore led to a higher $\text{M}$-value in grey matter than in other ROIs ($p = 0.01$).

In contrast with carbogen induced signal changes, similar $\text{BOLD}_t$, changes were found in both ROIs during functional activation. Task-evoked group average $\text{BOLD}_t$ changes, however, were different in both regions, with larger changes observed in the motor ROI ($p = 0.05$). Changes in oxygen metabolism were found to be comparable in both regions ($p = 0.23$), leading to a higher flow-metabolism coupling constant in the motor ROI than in the visual ROI ($p = 0.01$).

Figs. 4a to c summarise the reproducibility results for the visual (Fig. 4a), motor (Fig. 4b) and grey matter (Fig. 4c) ROI. CV1 (red), CV2 (blue) and CV3 (green) are group averaged values across all subjects. Within-day stability could not be assessed for $\text{M}$, $\text{SvO}_2$, $\text{CBF}_c$ and $\text{BOLD}_c$, as only one gas-breathing challenge was performed per day. The lowest coefficients of variation in all regions were found for venous oxygen saturation values showing a high consistency within and across subjects (CV < 5%). It should be noted, however, that $\text{SvO}_2$ is expected to be close to 100% in all cases, since an OEF of 35% is assumed within the GCM and very high flow and end-tidal oxygen values are consistently observed upon carbogen-7 inhalation. Gas induced $\text{CBF}_c$ signal changes showed lower variability across days and subjects (CV2 and CV3) than $\text{BOLD}_c$ and $\text{M}$ for all ROIs. $\text{BOLD}_c$ and $\text{M}$ variability was comparable within each ROI. A comparison across ROIs revealed that gas-related CVs across days and subjects for $\text{CBF}_c$, $\text{BOLD}_c$ and $\text{M}$ were largest within the motor ROI and lowest for the grey matter ROI. For all gas-related quantities, variability across subjects (CV2) was larger than variability within the same subject across days (CV3).

Amongst functionally-induced signal changes, $\text{CMRO}_2$ showed the best stability across runs (CV1) in the visual and motor ROI. The largest CV1 were found for $\text{BOLD}_c$ in both ROIs. $\text{BOLD}_t$, $\text{CBF}_c$ and $\text{CMRO}_2$ showed comparable CV1 in motor and visual ROI. The smallest CV2 across sessions was found for $\text{CMRO}_2$ in the visual ROI and the motor ROI. However, CV of $\text{BOLD}_c$ in the motor ROI was in a similar range as CV of $\text{CMRO}_2$. Averaging across subjects (CV3) revealed the CV of $\text{CBF}_c$ to be the lowest in both ROIs. Variability in $\text{BOLD}_c$, $\text{CMRO}_2$ and $\text{n}$ showed coefficients greater than 50% whereas all intra-subject CVs (CV1 and CV2) were below 50%.

Please cite this article as: Krieger, S.N., et al., Regional reproducibility of calibrated BOLD functional MRI: Implications for the study of cognition and plasticity. NeuroImage (2014), http://dx.doi.org/10.1016/j.neuroimage.2014.06.072
Inter-subjects test-retest plots across days for each quantity can be found in Figs. 5 and 6. Averaged signal changes for each quantity are plotted for the second and third day against the averaged data obtained on day 1. Data from different ROIs can be distinguished by colour, whereas days are encoded by shapes. Stability is determined by the distance between a given point and the unity line. Linear regressions have been applied to the data of each graph and the average coefficient of determination can be found in each plot. Fig. 5 shows the average signal changes related to gas inhalation and Fig. 6 functionally induced signal changes. The M-value (Fig. 5c) and blood flow plots (Fig. 5b) show less scattering than BOLD visualising better stability across subjects and within subjects of M and CBF than BOLD (Fig. 5a) due to gas inhalation and cognitive components. Oxygen metabolism estimates showed lower intra-subject variation than BOLD and CBF. Inter-subject variation was relatively high for all quantities with CMRO2 coefficients of variation being most consistent across ROIs. Taken together, these results indicate that calibrated fMRI may provide advantages over the BOLD approach.

**Discussion**

Despite the physiological ambiguity, poor stability characteristics and low spatial specificity of classical BOLD contrast, it is still the most commonly used method in longitudinal and group studies that assess cognitive brain function and task-induced functional activation. The utilisation of more direct physiological measures such as cerebral oxygen metabolism may have been limited so far by the lack of methodological validation and evidence for reasonable robustness and reproducibility. In this study, we have investigated the reproducibility characteristics of a state-of-the-art calibrated fMRI technique for cognitive studies, using a complex task with low-contrast primary sensory and cognitive components. Oxygen metabolism estimates showed lower intra-subject variation than BOLD and CBF. Inter-subject variation was relatively high for all quantities with CMRO2 coefficients of variation being most consistent across ROIs. Taken together, these results indicate that calibrated fMRI may provide advantages over the BOLD approach.
420 signal, both in terms of greater physiological specificity and in terms of reproducibility.

**Breathing manipulations**

Carbogen-induced changes in all quantities of this study showed good reproducibility characteristics with CV₂ comparable to those obtained at lower field strengths (Leontiev and Buxton, 2007). CV₂ and CV₃ of gas induced CBFc changes found in this study were, however, smaller than those obtained for BOLD, and M, in contrast to an earlier study at 3 tesla (Leontiev and Buxton, 2007). This discrepancy may be the result of implementation differences, both in terms of gas challenge and MRI acquisition. Different breathing manipulations may be used to perform a calibrated MRI experiment, and each has its own advantages in terms of balancing assumptions and the inclusion of low SNR measures which may reduce the reproducibility of the technique. The original calibrated BOLD approach was based on pure hypercapnic manipulations, under the assumption that this manipulation causes a

Fig. 4. Group averaged coefficients of variation for the visual (a), motor (b) and grey matter (c) BOLD CV₁ and CV₂, averaged across grey matter are smaller than those obtained in the visual and motor ROI. However, in all 3 ROIs only minor differences between across days intra-subject (CV₂) and inter-subject (CV₃) variability was observed. CMRO₂ shows the lowest CV₁ in both brain areas and CV₂ values comparable to BOLD.

Fig. 5. Average gas induced signal changes are plotted for the second and third day against the averaged data obtained on day 1. Data from different ROIs can be distinguished by colour, whereas days are encoded by shapes. Stability is determined by the distance between a given point and the unity line. M and CBF were more consistent across days within subjects and across subjects than the BOLD signal. However, all quantities show significant scattering from the unity line.
purely vascular change (Davis et al., 1998; Hoge et al., 1999; Kim et al., 1999). The assumption of zero change in CMRO2 during hypercapnia has, however, been challenged recently (Xu et al., 2011b; Zappe et al., 2008), and the hyperoxia calibrated BOLD method was introduced to circumvent this issue and improve participant comfort. It has also been postulated that the hyperoxia method may decrease variability compared to the hypercapnia approach since flow changes resulting from the breathing manipulation are not included in M estimation (Chiarelli et al., 2007b). The hyperoxia approach relies on the assumption of zero or negligible flow changes during oxygen inhalation, which has also been challenged (Bulte et al., 2007; Xu et al., 2011a). A generalised calibration technique allowing a combined hypercapnia-hyperoxia breathing challenge was utilised in this study, since this combination has been shown to lead to more robust estimations of the M.

Fig. 6. Average functionally induced signal changes are plotted for the second and third day against the averaged data obtained on day 1. While the BOLD data points show widespread scattering, CMRO2 data are close to the unity line reflecting matching data from each subject across days. Pooled data points in the CMRO2 plot compared to BOLD also suggests more predictable inter-subject reproducibility.

Fig. 7. Comparison of intra-subject coefficients of variation (CV1 and CV2) obtained with corrected (semi-transparent red and blue bars) and uncorrected (saturated red and blue bars) data for BOLD, CBF, CMRO2 and n. The correction scheme used in this study only had minor effects on the intra-subject coefficients of variation. However, intra-subject variation across days CV2 of BOLD reduced by approximately 15% after applying the correction algorithm.

Please cite this article as: Krieger, S.N., et al., Regional reproducibility of calibrated BOLD functional MRI: Implications for the study of cognition and plasticity, NeuroImage (2014), http://dx.doi.org/10.1016/j.neuroimage.2014.06.072
parameter (Gauthier and Hoge, 2013). Importantly, all these approaches suffer from the need to assume various physiological parameters such as flow-volume coupling ($\alpha$) and the relation between deoxyhaemoglobin concentration to the relaxation rate $R_2^*$. $R_2^*$ for hypercapnia or haematoctrit (Hct) and oxygen extraction fraction (OEF) for hyperoxia, which may show significant variability across subjects and brain regions (Blockley et al., 2012). Various techniques have recently been used to estimate brain OEF and most results are in the range of 0.3 to 0.4 (Bulte et al., 2012; Carpenter et al., 1991; Gauthier and Hoge, 2012; He and Yablonskiy, 2007; Ito et al., 2004; Jain et al., 2010; Lu and Ge, 2008; Qin et al., 2011; Raichle et al., 2001; Wise et al., 2013). Simulation of error using our current GCM-based results indicate that within this range the maximum error in $M$ due to underestimated OEF would be approximately $-30\%$ while overestimation could lead to a maximum error of $+20\%$. The potential effects of pathologies that impact OEF are yet unclear. For example, decreases in OEF have been found after brain injuries (Diringer et al., 2000) but the effect on $M$ and the accuracy of calibrated fMRI cannot be determined from the current data, as pathology would likely influence all measured components of the model. Recently, a hyperoxia calibrated fMRI approach has been introduced which is believed to overcome potential uncertainties in estimating $CMRO_2$ caused by assumption about flow-volume coupling and OEF (Driver et al., 2012). Furthermore, interleaved and combined hypercapnia and hyperoxia has been used to study absolute oxygen metabolism (Bulte et al., 2012; Gauthier and Hoge, 2012) as well as relative changes during functional activation while simultaneously estimating OEF, $\alpha$ and $\beta$ through a Bayesian modelling approach (Wise et al., 2013). These promising techniques could lead to improvements in calibrated fMRI implementation, though their applicability and reproducibility characteristics have yet to be assessed.

Another potentially confounding factor of hyperoxia based calibrated BOLD techniques might be the underestimated of gas breathing induced changes in CBF due to a shortening effect in longitudinal relaxation time $T_1$ of blood (Bulte et al., 2007; Gauthier and Hoge, 2013; Krieger et al., 2013). Commonly a $10\%$ underestimation of CBF is assumed based on increased oxygen saturation in arterial blood plasma. This value, however, is based on theoretical modelling. Venous and arterial blood $T_1$ have been measured ex-vivo at 1.5 tesla (Lu et al., 2003), 3 tesla (Lu et al., 2004; Noeske et al., 2000; Stanisz et al., 2005), 4.7 tesla (Dobre et al., 2007; Silvennoinen et al., 2003), 7 tesla (Dobre et al., 2007; Francis et al., 2008; Rane and Gore, 2013) and 9.4 tesla (Dobre et al., 2007). While only few studies have investigated venous blood $T_1$ in-vivo at these field strengths (Qin et al., 2011; Rane and Gore, 2013; Rooney et al., 2007; Wu et al., 2010; Zhang et al., 2013) and there is no in-vivo data available for arterial blood $T_1$. This lack of information is compounded by the fact that studies started investigating the effects of hyperoxia and hypercapnia on the longitudinal relaxation time of blood only very recently, and the results obtained so far do not provide reliable quantitative estimates (Dethrage et al., 2014; Krieger et al., 2014; Ma et al., 2014). Future studies will be necessary to quantitatively assess the effect of breathing gases to blood $T_1$, especially with regards to the consistency of these effects across subjects.

The oxygen and carbon dioxide concentrations of delivered gas mixtures vary between studies. Higher concentrations of oxygen and CO₂ result in larger BOLD signal changes, bringing the measured BOLD signal closer to M on the iso-CMRO₂ contour (Hoge et al., 1999). SNR-limited techniques such as ASL could benefit from increased signal changes induced by higher CO₂ concentrations and lead to improved reproducibility characteristics of the calibrated BOLD technique. On the other hand, high carbon-dioxide concentrations can induce significant levels of air hunger, discomfort or breathlessness, which can affect physiology and metabolism of cortical and sub-cortical brain regions (Brannan et al., 2001; Liotti et al., 2001). It is therefore important to use accurate control and recording systems for end-tidal gas values, as well as physiological correlates such as heart and breathing rate. Future studies could compare effects of various gas-mixtures to calibrated BOLD reproducibility.

Gas delivery also plays an important role in technical reliability. Whereas face masks and nasal cannulas are the most common modes of gas delivery, we have used a mouthpiece combined with a nose clip. This setup guaranteed that only the carbogen-7 gas mixture was inhaled, and no outside air can be entrained into the breathing circuit. Larger changes in end-tidal partial gas pressures and CBF have typically been observed for the same gas concentration when mouth pieces were used instead of face masks (Gauthier and Hoge, 2013; Krieger et al., 2013). Most face masks are not sealed, which leads to contamination of inhaled gas by room air. Cannulas commonly lead to even greater contamination from room air since there is still space in the nose around the cannula. Therefore end-tidal values are both lower and more variable than with face masks and cannulas than with mouthpieces. This variability in gas concentration inhaled by the subjects introduces greater uncertainty in calibrated BOLD acquisitions, thereby influencing their reliability. Furthermore, there is evidence that breathing through face masks induces larger air hunger than breathing through mouth pieces (Liotti et al., 2001). Air hunger in turn can result in significant neuronal cerebral activation which may diminish the effectiveness and accuracy of the calibration procedure (Brannan et al., 2001; Liotti et al., 2001). While the gas delivery method used here has shown good reproducibility, further improvements could be obtained using an end-tidal gas control system (Mark et al., 2010, 2011; Slessarev et al., 2007). This technique uses a feed-forward physiological model to precisely target a predefined end-tidal partial gas pressure. Cerebro-vascular reactivity (CVR) has been found to be higher when using an end-tidal control system compared to fixed gas concentrations (Tancredi and Hoge, 2013). This technique, while a possible improvement in terms of reproducibility, comes at the cost of further scanner time and represents an additional equipment investment. An alternative approach uses computer controlled feedback on a breath-by-breath basis to alter the inspired gas partial pressures to meet pre-determined end-tidal gas pressures (Wise et al., 2007). Future studies could determine whether the use of these gas-inhalation techniques provide an improvement in reliability which outweighs their additional cost.

In addition to the previously-described gas inhalation based fMRI approaches, recent studies have demonstrated the possibility of conducting calibrated fMRI experiments without any breathing manipulation. Combined gradient and spin echo data has been used to measure the relaxation rates $R_1$ and $R_2^*$. A biophysical model is then used to calculate the calibration constant based on the transverse relaxation rates (Blockley et al., 2012; Herman et al., 2013; Hyder et al., 2001; Kida et al., 2000). An alternative technique uses $T_2$:Relaxation-Under-Spin-Tagging (TRUST) MRI to quantify oxygen metabolism as a whole-brain average (Xu et al., 2009). TRUST MRI-based estimation of global cerebral oxygen metabolism has been found to be a very reliable method showing low intra-subject and inter-subject coefficients of variation (Liu et al., 2013b). The major advantage of these techniques compared with gas inhalation based calibrated fMRI is that no additional gas breathing equipment is needed, which makes the setup more cost-effective and reproducible across sites. The disadvantages and restrictions of the techniques described can, however, be severe, depending on the purpose of the specific study and the MRI setup available. Where- as TRUST MRI might be a powerful tool for measuring a global average in oxygen metabolism, this technique does not provide information about regional oxygen usage. Functional experiments can therefore not be conducted using this technique. The relatively large amount of RF pulses applied in this sequence restricts its application to lower field MRI due to specific absorption rate (SAR) limits. Similar power deposition restrictions are valid for the R2* calibrated fMRI techniques (Blockley et al., 2012; Herman et al., 2013). Additionally, the R2* approach has been found to be less reliable than hypercapnia or hyperoxia fMRI, potentially due to its greater sensitivity to magnetic field inhomogeneity (Blockley et al., 2012). Due to these restrictions, this study was conducted using a gas inhalation based calibrated fMRI technique rather than any of the promising non-gas inhalation based techniques mentioned.

---

Please cite this article as: Krieger, S.N., et al., Regional reproducibility of calibrated BOLD functional MRI: Implications for the study of cognition and plasticity, NeuroImage (2014), http://dx.doi.org/10.1016/j.neuroimage.2014.06.072
A number of gas-induced reproducibility features must be considered. All gas inhalation–related repeatability measures across days (CV2) were slightly lower within the grey matter ROI than within the visual and motor ROIs, confirming trends observed in ASL and calibrated fMRI at lower field strengths (Stone et al., 2013; Wang et al., 2011). This may be attributable to the relative sizes of each ROI. The visual and motor ROI represent only a relatively small section of the total grey matter, thereby leading to more variability than when signal is averaged across the whole non-task-specific cortical voxels of the brain.

The calibration parameter M values obtained in this study are somewhat smaller than those obtained by earlier 7 tesla studies (Driver et al., 2012; Gauthier et al., 2013a; Hall et al., 2013; Huber et al., 2013). This is an expected result due to the relatively short echo time used in this study. However, absolute comparisons of M-values across studies are generally difficult since different calibration approaches, MRI techniques, acquisition parameters and breathing manipulations can affect M (Krieger et al., 2013). Along these lines, the higher inter-session variability of M compared to CBF found in this study demonstrates the importance of measuring M for each subject and each day separately. Wrongly estimated M-values can lead to deviations in CMRO2, which may diminish the conclusiveness of a specific study (Krieger et al., 2013). While variability in M originating from technical uncertainty could be reduced by using an end-tidal gas control technique (Mark et al., 2010; Wise et al., 2007), physiological variability can only be accounted for by estimating M during each scanning session. Furthermore, our results indicate that regional differences in flow–metabolism coupling (n) and maximum BOLD signal changes (M) across ROIs may lead to inaccurate results when the calibration procedure is based on a global average rather than separate averaging for each brain region (Chiarelli et al., 2007a). In this context, future calibrated BOLD studies could seek to increase the spatial resolution of this technique to investigate the potential impact of partial volume effects on the estimation of M and CMRO2. Furthermore, higher resolution data would allow comparison with results obtained in recent rat studies which established the uniformity of M across cortical laminae (Herman et al., 2013).

Finally, venous oxygen saturation was associated with lower CVs than any other measure. This is not surprising as this measurement is expected to always reside within a very limited range of values. SvO2 is estimated via the GCM under the assumption of an OEF of 35%, and combined with the end-tidal O2 partial pressure values at baseline and during gas inhalation. Since baseline O2 partial pressure is similar both within and across subjects in healthy adults, and since gas-induced end-tidal values are reliably high within and across subjects, SvO2 variability is expected to be low.

Functional reproducibility

Several studies have assessed the reproducibility characteristics of BOLD and ASL for high-contrast motor and visual tasks (e.g. finger tapping or flickering checkerboard) at lower field strengths. There is evidence that using CBF as a marker for cerebral neuronal activity provides better intra-subject and inter-subject reproducibility than BOLD, especially over long periods of time (Chen et al., 2011; Tjandra et al., 2005; Wang et al., 2003, 2011). The ASL signal originates from the capillary and arteriolar vessel system, thereby providing better spatial specificity than BOLD, which is biased towards draining venules and veins that can be further away from the actual area of functional activation (Detre and Wang, 2002; Turner, 2002). Furthermore, the BOLD signal is vulnerable to susceptibility artefacts and baseline drifts, which have a less important impact on ASL (Detre and Wang, 2002). Task-evoked BOLD signal changes have also been found to be modulated by baseline blood oxygenation and cerebro-vascular reactivity, which contributes to the reduced reproducibility of BOLD (Liu et al., 2013a; Lu et al., 2008). The results obtained in this study indicate how these trends translate to ultra-high field and more complex cognitive tasks that lead to lower signal changes. Functional CBF changes showed better within-day intra-subject reproducibility (CV1) than BOLD, confirming the results of studies at lower field strengths (Tjandra et al., 2005; Wang et al., 2003, 2011). BOLD variability across days (CV2) was similar to variability across runs from the same day (CV3) which confirms results obtained by earlier studies of BOLD variability (Smith et al., 2005), leading to comparable reproducibility of BOLD and CBF over longer time periods. To our knowledge, there is no published literature comparing BOLD and CBF intra-subject reproducibility on a time scale of seven days. Our results suggest that the superior sensitivity of BOLD may balance out against the lower signal to noise ratio of PASL, despite the technical advantages of PASL in terms of long-term reproducibility, when using more complex cognitive tasks that only lead to modest functional signal changes.

A viable alternative to BOLD and CBF as surrogate neuronal markers of functional cerebral activation could be CMRO2. Earlier studies at 3 tesla observed comparable intra-subject repeatability for all three quantities (Leontiev and Buxton, 2007; Stone et al., 2013). The results of this study provide evidence that CMRO2 is superior to BOLD and CBF in terms of intra-subject stability when using more complex cognitive tasks (Figs. 4a, b and 6a to d). This effect may be enhanced by the use of ultra-high magnetic field. CMRO2 variability estimates across runs within the same day were found to be well below those of BOLD and CBF. These results confirm recent findings at 3 tesla with visual stimulation showing that calculated responses such as CMRO2 and n are relatively insensitive to baseline CBF changes across runs. Due to flow-metabolism coupling, changes in CBF can be assumed to affect the BOLD signal by a similar relative amount, leading to lower variation in CMRO2 and n than in CBF and BOLD (Leontiev and Buxton, 2007). Only recently, cerebral oxygen metabolism has started to be used to study cognitive processes (Ances et al., 2009; Gauthier et al., 2013b; Mohtasib et al., 2012). The intra-subject repeatability characteristics found in this study indicate that calibrated BOLD is a promising alternative to classical functional techniques such as BOLD and ASL for this type of study.

This study features some interesting results with regards to the coupling of cerebral blood flow and brain oxygen metabolism. The flow-metabolism coupling parameters n obtained in this study were comparable to those reported in earlier studies (Ances et al., 2008; Marrett and Gjedde, 1997; Roland et al., 1987; Seitz and Roland, 1992). The variability of n across brain areas and subjects found in this study could help understand and interpret the relatively large variability of the BOLD signal (Aguirre et al., 1998; Duff et al., 2007; Fox et al., 2001). The variability of n across sessions found in this study was smaller than that of CBF, and CMRO2, though larger than the variability of CMRO2. However, larger variability in n could be found within the same subject across days. Interestingly, M shows better intra-subject reproducibility across sessions than n in motor and visual ROIs which could be important for applications of calibrated fMRI to pathophysiology.

Inter-subject variability CV4 was found, as expected from previous studies, to be larger than intra-subject variability for BOLD, CBF, CMRO2 and n (Leontiev and Buxton, 2007; Stone et al., 2013; Tjandra et al., 2005). However, we anticipated that the relatively complex task used in this study would lead to smaller signal changes and higher inter-subject variability values compared to those obtained with high-contrast tasks such as flickering checkerboard. Without explicit measures such as B1 shimming or transmit SENSE, ultra-high field MRI scanners inevitably suffer from more severe transmit radiofrequency magnetic field inhomogeneities. Due to the higher field strength, these inhomogeneities also are larger across subjects than for lower field strengths. This may partly explain the higher inter-subject variability observed in this study. Calculated quantities such as CMRO2 and n showed consistent CV4 for both the visual and motor ROI, while both BOLD, and CBF, showed greater differences in variability coefficients across regions (Stone et al., 2013). More consistent stability across brain regions may improve the reliability of studies assessing cognitive...
processes in several cerebral areas. Our results therefore suggest that calibrated BOLD could be a powerful alternative to BOLD and CBF scans alone.

Habituation effects on reproducibility

The nature of the functional task chosen for this study may have had an impact on the reproducibility estimates obtained. Cerebral functional changes related to motor learning can be expected, as well as habituation effects in task performance and visual stimulation. Linear regression of habituation and learning effects from task performance measures did not, however, have a major impact on our stability assessments. Only minor differences in BOLDt, CBFt, CMRO2 and nCVs across runs or days in both visual and motor ROI were found following regression of performance changes (see Fig. 7). CVs for most measures remained almost identical before and after applying the correction. It should however be noted, that the correction scheme we have applied is based on the assumption that motor performance is linearly related to all quantities and signal changes. It could be that this scheme underestimates the learning and habituation effects, if a more complex relationship is at play. However, most recent studies assume a linear relationship between performance and MRI signal amplitude changes and therefore this study followed this assumption (Fernández-Seara et al., 2009; Grafton et al., 2002; Steele and Penhune, 2010). Future studies could seek to implement other mathematical models to investigate non-linear correlation between MRI signal changes and motor performance or test for different weightings between physiological measures.

Day to day differences in motor performance and within session habituation effects leading to differences in cerebral functional activation may furthermore be responsible for counter-intuitive trends in BOLD reproducibility. Coefficients of variation across runs (CVfl) were found to be larger than across days (CVfd) within the motor ROI for BOLDt (compare Fig. 4b). Increased motor activation was found at the beginning of each day. Participants require a certain period at the beginning of each session to recover the low performance error observed in the previous session. This could be due to the fact that the initial performance of the task requires greater attention than subsequent runs, leading to a higher BOLDt signal at the beginning of each scanning day (Seidler et al., 2004; Steele and Penhune, 2010). This within-session habituation effect may lead to the observed lower BOLDt reproducibility across runs than across days. It is furthermore possible that this habituation effect is associated with changes in flow-metabolic coupling in different learning phases, thereby leading to different impacts on BOLDt, CBF, and CMRO2 stability. This interpretation is supported by the fact that the visual area does not show this counter-intuitive CV effect and CV across runs and days were found to have comparable values.

Technical considerations for calibrated fMRI

Compared to the classical BOLD contrast, calibrated BOLD typically relies on the simultaneous acquisition of CBF and BOLD signals. Commonly used labelling schemes include pulsed arterial spin labelling (PASL) schemes such as FAIR (Kim, 1995) and PICORe (Wong et al., 1997) or pseudo-continuous arterial spin labelling (pCASL) (Dai et al., 2008; Wu et al., 2007) techniques. Recent studies have found that the pCASL approach provides superior SNR and stability (Chen et al., 2011; Wu et al., 2011, 2013), so pCASL may be the best ASL technique to use when conducting a calibrated BOLD study at lower field strengths (Gauthier and Hoge, 2013). However, due to limits in specific absorption rate (SAR) related to radio-frequency pulse energy deposition, the use of pCASL and Q2TIPS at ultra-high field strengths is complicated (Charif et al., 2012). For this reason, an optimised FAIR QUIPSII labelling scheme, found to provide superior reproducibility compared to PICORe QUIPSII and PICORe Q2TIPS (Cavuoso et al., 2009), was used in the current study.

The echo time (TE) is a crucial parameter when conducting PASL based calibrated fMRI studies, both in terms of BOLD signal scaling and compartmental origin, and with regards to potential BOLD contamination of the CBF-weighted MRI signal. Additional BOLD contribution to the PASL signal could lead to an overestimation of CBF, although extravascular BOLD signal contribution can be minimised by using short echo times, such as the one used in this study. However, the intravascular BOLD signal peaks at short echo times and may therefore contaminate the CBF-weighted PASL signal, especially in regions containing large veins (Ivanov et al., 2013). As the intravascular BOLD signal is believed to be smaller than its extravascular counterpart, minor BOLD contamination of the CBF-weighted PASL signal can be expected, especially at ultra-high magnetic field (Duong et al., 2003; Kim and Ogawa, 2012; Uludağ et al., 2009). Another confounding factor especially in ASL at ultra-high field is physiological noise, which scales with field strength. Larger voxel sizes translate into better SNR but at the same time increase the amplitude of physiological noise (Krüger and Glover, 2001; Triantafyllou et al., 2005). In this study we have used a voxel size that showed optimal SNR while minimising the impact of physiological noise on our data. Background suppression techniques commonly rely on multiple inversion pulses to null the background signal at acquisition time (Garcia et al., 2005; Ishimori et al., 2011). As this approach relies on a well-defined nulling time of all background signals, its application is limited to 3D readout or small number of slices. Furthermore, there is concern that background suppression will impact extravascular BOLD quantification and thereby change the relationship between blood flow and BOLD, which may lead to biases in calibrated fMRI quantification. Therefore background suppression has not been applied in this study. A short TE will also result in lower BOLD signal changes. This would in turn lead to a smaller M estimate. This is not expected to lead to biases in CMRO2 quantification, however, if the task-based BOLD response is measured using the same TE as both measures will then scale together. The BOLD signal in this case will however represent a larger intravascular contribution. A dual-echo sequence with a second TE optimized for BOLD would reduce this bias and such sequences should be used whenever possible.

Another restriction at ultra-high magnetic fields arises from the radiofrequency transmitting hardware (Duyn, 2012; Yacoub et al., 2001). Whereas it is standard at lower field strengths to use a body coil to label arterial spins globally in FAIR sequences, or selectively spins in the carotid arteries in pCASL sequences, such equipment is not yet available at 7 tesla. Therefore ASL at ultra-high field may suffer from reduction in functional SNR due to inflow of non-labelled spins. Fresh blood inflow may be especially problematic for hypercapnia-based calibrated BOLD. Carbon dioxide inhalation leads to global increases in blood flow and therefore rises in blood flow velocity, which in turn reduces the arterial arrival time. A shortening in arterial arrival time could diminish functional contrast of ASL at 7 tesla. Therefore it is indispensable (at each field strength) to carefully adjust ASL inversion times (TI) to achieve maximum functional SNR at rest and during gas inhalation. In this study the combination of TI1/TI2 = 700 ms/1700 ms provided the optimal contrast-to-noise ratio while guaranteeing no fresh blood inflow into the imaging slab before the EPI readout. In summary, while some improvements in stability would be possible using other ASL implementations, SAR and hardware limitations at 7 T have lead to the choice of the particular ASL implementation used in this study. Future studies could seek to use dedicated labelling coils to enable the use of better ASL techniques, thereby improving the reliability of calibrated fMRI experiments.

Conclusions

Calibrated BOLD fMRI is a promising technique for investigating functional brain activity on the basis of physiologically meaningful quantities. This study investigated important technical considerations for the implementation of calibrated fMRI and assessed the applicability...
of the calibrated BOLD technique in the context of cognitive and learning studies. CMRO₂ was found to be more suitable for longitudinal studies than BOLD and CBF alone. CMRO₂ was furthermore associated with a more consistent variability across brain regions than BOLD and CBF, which may provide better reliability for studies looking into several brain regions. While calibrated BOLD studies require increased scanning times and additional equipment, these disadvantages may be outweighed by the increased stability, higher spatial specificity and greater physiological interpretability of CMRO₂ estimates.

Acknowledgments

We thank Elisabeth Wladimirow and Domenica Wilf for help with data acquisition, subject handling and scanning organisation. We would also like to thank Robert Trampel for help with ethics application and Robert Trampel and Gabrielle Loehmann for helpful discussions.

References


