SCIENTIFIC REPORTS

OPEN

Improving the Assessment of Breath-Holding Induced Cerebral Vascular Reactivity Using a Multiband Multi-echo ASL/BOLD Sequence

Received: 27 June 2018
Accepted: 28 February 2019
Published online: 25 March 2019

Alexander D. Cohen & Yang Wang

Breath holding (BH) is a viable vasodilatory stimulus for calculating functional MRI-derived cerebral vascular reactivity (CVR). The BH technique suffers from reduced repeatability compared with gas inhalation techniques; however, extra equipment is needed to perform gas inhalation techniques, and this equipment is not available at all institutions. This study aimed to determine the sensitivity and repeatability of BH activation and CVR using a multiband multi-echo simultaneous arterial spin labelling/blood oxygenation level dependent (ASL/BOLD) sequence. Whole-brain images were acquired in 14 volunteers. Ten subjects returned for repeat imaging. Each subject performed four cycles of 16 s BH on expiration interleaved with paced breathing. Following standard preprocessing, the echoes were combined using a T2*-weighted approach. BOLD and ASL BH activation was computed, and CVR was then determined as the percent signal change related to the activation. The "M" parameter from the Davis Model was also computed by incorporating the ASL signal. Our results showed higher BH activation strength, volume, and repeatability for the combined multi-echo (MEC) data compared with the single-echo data. MEC CVR also had higher repeatability, sensitivity, specificity, and reliability compared with the single-echo BOLD data. These data support the usefulness of an MBME ASL/BOLD acquisition for BH CVR and M measurements.

Cerebral vascular reactivity (CVR) is a measure of a blood vessel's response to a vasoactive stimulus, such as the manipulation of arterial levels of CO_2. Recently, functional magnetic resonance imaging (fMRI) has been applied to measure CVR and responses to vasodilatory stimuli. These include blood oxygenation level dependent (BOLD) fMRI, where changes in blood oxygenation in response to a vasodilatory stimulus are measured[1–6], and arterial spin labelling (ASL) MRI, where blood flow is measured directly by magnetically tagging blood flowing into the brain[7–9].

Several techniques are used to manipulate arterial CO_2 levels. One method uses gas inhalation with varying concentrations of CO_2 to create elevated CO_2 levels in the blood[7,8,10–13]. These methods, however, require extra equipment that may not be available at all institutions. Breath-holding (BH) fMRI is a viable alternative approach to provide a vasodilatory stimulus to measure of CVR[1,4,14]. For example, Kastrup *et al.* found BH CVR measurements were comparable to CVR computed using gas inhalation techniques[5]. In another study, healthy volunteers simulated poor BH performance, and repeatable results were found when end-tidal CO_2 measures were convolved with a double gamma variate hemodynamic response function (HRF) and used as regressors in a BH activation model[4]. However, end tidal CO_2 monitoring equipment is not available at all institutions.

ASL fMRI is an attractive complement to BOLD fMRI. While the BOLD response results from the combination of several factors including cerebral blood volume (CBV), cerebral blood flow (CBF), and blood oxygenation, ASL measures CBF directly by magnetically tagging blood flowing into the brain. Studies have compared BOLD and ASL CVR measurements and found regional and global similarities between the two techniques[7,15]. In addition, sequences have been developed to acquire BOLD and ASL images simultaneously and have been

Medical College of Wisconsin, Department of Radiology, Milwaukee, WI, USA. Correspondence and requests for materials should be addressed to A.D.C. (email: acohen@mcw.edu) or Y.W. (email: yangwang@mcw.edu)

used for CVR measurements[15,16]. These sequences also allow for quantitative CBF to be computed using ASL data collected at the same time as the BOLD data.

These techniques can also be used to calibrate the fMRI signal and evaluate neurovascular coupling[17,18]. These experiments require the collection of both CBF and BOLD images, preferably simultaneously. As mentioned, the BOLD response is complicated and is related to CBV, CBF, and the cerebral metabolic rate of oxygen consumption ($CMRO_2$). Davis *et al*. modelled the relationship between these parameters[17], Equation (1). Here, the subscript zero denotes the baseline condition. The constant α describes the relationship between changes in CBV to changes in CBF. The constant β is related to susceptibility changes. The parameter M represents the maximum BOLD percent signal change and can be determined using a hypercapnic challenge where the $CMRO_2$ ratio is assumed to be 1. Once M is known, $CMRO_2$ can be estimated during a separate task or during the resting state. It is important to note that $\Delta BOLD/BOLD$ and, as a result, M and CVR, changes with echo time; however, since neither is an absolute physiological quantity, this does not matter in the scope of this study.

$$\frac{\Delta BOLD}{BOLD_0} = M\left(1 - \frac{CBF}{CBF_0}^{\alpha-\beta}\frac{CMRO_2}{CMRO_{2,0}}^{\beta}\right) \quad (1)$$

BH can be completed following either inspiration or expiration. Although BH following inspiration is easier to perform, studies have shown BH following expiration is more repeatable[19]. Furthermore, BH following inspiration has been shown to be biphasic, consisting of an initial signal dip followed by a signal increase[3,20]. The BH response depends on the length of the BH, with larger, more robust responses occurring with longer BH durations[4,21]. Of course, a trade-off exists between BH length and subject tolerance.

Despite the advantages of BH fMRI, several issues exist. First, BH imaging relies on subject compliance, although respiratory bellows can be used to monitor subject compliance. Second, motion artefacts tend to be increased with BH protocols[22]. The BH stimulus is also non-quantitative and can vary across subjects and time points. As such, repeatability of BH fMRI is limited and inter- and intra-subject variability are relatively high, especially in the absence of end-tidal CO_2 measures[4].

Multiband (MB), or simultaneous multi-slice (SMS), imaging has been incorporated into fMRI studies[23–26]. These sequences excite multiple slices simultaneously and can be used to increase spatial and/or temporal resolution. One recent study showed higher activation sensitivity for SMS data compared with conventional echo-planar imaging (EPI) in response to a gas inhalation challenge[27]. Additionally multi-echo (ME) EPI techniques have shown higher sensitivity in BOLD acquisitions[28–33]. BOLD contrast is maximized when the echo time (TE) is equal to T2*. Thus, echoes can be combined by weighting each echo by the voxelwise T2*[28,30,34].

Recently, an MB, ME simultaneous ASL/BOLD (MBME ASL/BOLD) sequence was developed to acquire whole-brain ASL and BOLD images using a total of four echoes, allowing echo combination and denoising to increase BOLD sensitivity[35]. One study used this sequence to evaluate resting-state functional connectivity and found increased BOLD network size and strength following echo combination and denoising[35]. Another study acquired finger-tapping task fMRI data using this sequence[36]. This study found significantly higher temporal signal-to-noise ratio (tSNR) and task activation following echo combination and denoising[36]. The main advantage of MB imaging for this sequence is reduced effects of T1-decay of the tagged blood and reduced interslice labelling delay times and total readout times. Thus, whole-brain simultaneous ASL/BOLD data can be acquired with whole-brain PW data. This can also help mitigate the effects of longer readout times associated with ME imaging.

In this study, we used the MBME ASL/BOLD sequence to acquire BH fMRI data with four echoes. ASL and BOLD BH-CVR was analysed pre- and post-echo combination in the absence of end-tidal CO_2 measurements. The simultaneously acquired ASL signal was incorporated to compute the "M" parameter in the Davis Model[17]. In addition, we analysed the repeatability of the task activation, CVR, and M in the subjects that were scanned twice within a two-week period. Although all of the limitations of BH fMRI cannot be addressed, it was hypothesized that echo combination would lead to higher activation and repeatability of activation, CVR, and M, and would provide a means for the robust calculation of CVR and M without additional equipment.

Materials and Methods

This study received approval from the Medical College of Wisconsin's Institutional Review Board, and was conducted according to the ethical standards outlined in the Declaration of Helsinki. All subjects provided written informed consent before participating. In total, 14 right-handed, healthy adult volunteers (six male, eight female; mean age 29.8 +/− 8.3 years; age range 20–50 years) were recruited for this study. Ten subjects were able and willing to return within two weeks of their initial imaging session for a repeat scan. Subjects were asked to refrain from intake of caffeine for six hours before the MRI exam.

Imaging. Imaging was performed on a GE Healthcare (Waukesha, WI) 3T MR750 system with a body transmit coil and a 32-channel NOVA (Wilmington, MA) receive head coil. High-resolution anatomical images were acquired to provide accurate coregistration with the functional images. A T1-weighted magnetization-prepared rapid acquisition with gradient echo (MPRAGE) was collected with the following parameters: TR/TE = 7.3/3.0 ms; flip angle (FA) = 8°; field of view (FOV) = 256 mm, 1 × 1 × 1 mm^3 resolution; bandwidth (BW) = 62.5 kHz; and TI = 900 ms.

Each subject also underwent an MBME ASL/BOLD scan using the pulse sequence described in ref.[35]. Briefly, following a pseudo-continuous ASL (pCASL) labelling block and post-labelling delay (PLD) a single shot MBME EPI readout was executed. This sequence had the following parameters: pCASL labelling time = 1.5 s, PLD = 1.5 s, number of echoes = 4; TE = 9.1,25,39.6,54.3 ms; TR = 4.0 s; MB-factor = 4; number of excitations = 11 (total slices = 11 × 4 = 44); FOV = 240 mm; resolution = 3 × 3 × 3 mm^3; 80 × 80 matrix; FA = 90°; radio frequency pulse width = 6400 ms. Echoes were acquired consecutively as part of one shot. A partial k-space acquisition

was also employed with a partial Fourier factor of 0.75, and in-plane acceleration was utilized with R = 2. Blipped-controlled aliasing in parallel imaging (blipped-CAIPI)[26] was employed with FOV-shift = 1/3 to reduce the g-factor noise amplification caused by the slice-unaliasing in MB imaging.

During these scans, a BH task was employed. Scans began with 44 s of paced breathing, followed by four cycles of a 16 s of BH on expiration, 16 s of self-paced recovery breathing, and then 24 s of paced breathing. Scans ended with an additional 24 s of paced breathing. The paced breathing portions consisted of alternating 3 s inspiration and expiration blocks. Scans lasted 356 s, which included 64 s of calibration repetitions collected at the beginning of the acquisition for reconstruction of the functional images.

Image Reconstruction. Specific details regarding MBME ASL/BOLD reconstruction and the sequence itself can be found in ref.[35]. Briefly, slices were unaliased using a slice-GRAPPA (generalized autocalibrating partial parallel acquisition) algorithm[26] applied separately for each echo. In-plane aliasing was performed following slice-unaliasing using a traditional 1D-GRAPPA algorithm[37]. Coils were combined using a sum-of-squares technique, and partial k-space was reconstructed using the homodyne method[38].

Preprocessing. The anatomical MPRAGE image was skull-stripped and transformed to Montreal Neurological Institute (MNI) space using Advanced Normalization Tools (ANTS, http://stnava.github.io/ANTs). First, the MPRAGE image was affine-registered to MNI space with 12 degrees of freedom using a mutual information metric. Next, the registration was refined using a nonlinear symmetric normalization algorithm with a cross-correlation metric. In addition, individual WM and GM probability maps were extracted using the FAST segmentation function in FSL[39].

Data preprocessing was performed on each of the four echoes separately, using a combination of AFNI[40] (https://afni.nimh.nih.gov/afni) and FSL[41] (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki). First, each echo was skull-stripped and despiked using *3dSkullStrip* and *3dDespike* respectively in AFNI. Next, the first echo was volume registered and coregistered to the anatomical MPRAGE image using an affine transform with 12 degrees of freedom and *epi_reg* in AFNI. The transformation matrices from the volume and anatomical registrations were then applied to the remaining three echoes. The four echoes were transformed to MNI space using the transformation matrix output from the MPRAGE-MNI registration. A perfusion-weighted (PW) time series was generated by from the first-echo data by high-pass filtering the signal with a cut-off frequency of 0.09 Hz and then demodulating the result by multiplying by $\cos(\pi n)$[42].

Echo Combination. Following preprocessing, the four acquired echoes were combined using the T_2^*-weighted technique[34,43]. First, the voxelwise mean across time of each individual echo dataset was fit to an exponential function using log linear regression to estimate T_2^* (Equation (2)). Here, $\overline{S_0}$ is the signal immediately after excitation, and TE_n represents the nth echo time. The voxelwise T_2^* was then used to determine the weights, $w(T_2^*)$ (Equation (3)), which were used in a weighted summation of the echoes.

$$S(TE_n) = \overline{S_0} \cdot e^{-TE_n/T_2^*} \tag{2}$$

$$w(T_2^*) = \frac{TE_n \cdot e^{-TE_n/T_2^*}}{\sum_n TE_n \cdot e^{-TE_n/T_2^*}} \tag{3}$$

fMRI Processing and BH Response Analysis. The above procedures resulted in three datasets that underwent further processing for fMRI analyses: second-echo (E2, TE = 25 ms), ME combined, (MEC), and PW. The second echo was chosen to mimic a typical BOLD fMRI acquisition. All data were blurred with a 4.5 mm FWHM (full width at half maximum) Gaussian kernel. The E2 and MEC data were detrended with a third-order polynomial, and label-control oscillations were regressed out of the data by including a column of alternating −1 s and 1 s in the design matrix.

For both the BOLD and PW data, the BH response was determined with a general linear model using *3dDeconvolve* in AFNI. For the BOLD data, following *3dDeconvolve*, a restricted maximum likelihood model (*3dREMLfit)* was used to model temporal autocorrelations in the data. This program uses an ARMA(1,1) to model the time series noise correlation in each voxel. The PW time series is the result of a filtering and demodulation process. As a result, it is not as susceptible to temporal autocorrelation compared to BOLD[44]. Perfusion data have been found to have minimal temporal autocorrelation, and the perfusion time series has been shown to be temporally statistically independent[44]. BH regressors were generated by convolving a square wave, with ones during BH periods and zeros otherwise, with a double gamma-variate hemodynamic response function (HRF). The BH hemodynamic response is slow, with the peak occurring after the BH period. Thus, most studies time shift the BH regressor by several seconds to better model the response[21,45,46]. Moreover, the BH response delay varies across the brain by as much as +/−8 seconds[2,4,46,47]. To account for this, the BH regressor was shifted from −2*TR to 8*TR in steps of TR, and for each voxel, the regressor that resulted in the highest positive t-score was chosen.

CVR Calculation. CVR was calculated for the E2 (CVR_{E2}), MEC (CVR_{MEC}), and PW (CVR_{PW}) data as the percent signal change of the BH task response. This was computed by dividing the beta coefficient of the BH response by the mean signal.

M Computation. The Davis model describes the BOLD signal change in terms of CBF and $CMRO_2$, (Equation (1)). The model contains two constants, α and β, which must be assumed. A range of values has been used previously, with α typically equal to 0.2[48] and β ranging from 1–1.5[49–52]. Recent research has suggested β

is closer to 1 at field strengths >3T[18,51]. For this analysis, $\alpha = 0.2$ and $\beta = 1$. Under the hypercapnic condition (i.e., BH), $CMRO_2/CMRO_{2,0}$ is assumed to be 1. Thus, M can be calculated using Equation (4). The BOLD signal change and CBF ratio were calculated using the BOLD and PW BH activation beta values and the baseline BOLD and PW signals, respectively. M was computed using both E2 (M_{E2}) and MEC (M_{MEC}) datasets.

$$M = \frac{\frac{\Delta BOLD}{BOLD_0}}{1 - \left(\frac{CBF}{CBF_0}\right)^{\alpha - \beta}} \tag{4}$$

Statistical Analysis. For each subject and dataset, mean GM tSNR was computed on a voxelwise basis and defined as the mean signal divided by the standard deviation of the noise. Noise was defined as the residual between each voxel's best fit to the model and the signal itself.

All individual activation maps were thresholded at an uncorrected $p < 0.001$. For the BOLD data, the t-score of the BH activation was extracted both from GM and using an overlap mask that was created for each subject from voxels active in both the E2 and MEC datasets. For the PW data, the t-score was extracted from GM and active voxels. The fraction of active voxels in GM was also computed. Finally, the amount of variance explained by the regressor was computed. These metrics were compared across the E2, MEC, and PW datasets using Bonferroni-corrected paired t-tests to compare individual means.

The mean time series was extracted from all GM voxels, voxels with $0.01 < p < 0.05$, and voxels with $p < 0.001$ for the BOLD data for all subjects. To examine potential improvements with the combined-echo data in active voxels, a mask was created using the active voxels in the E2 data. That same mask was applied to the MEC data. Thus, the same voxels were analysed for the E2 and MEC datasets.

The mean CVR and M values were extracted from GM for each subject and dataset and compared between the E2 and MEC data. CVR and M were also averaged across subjects to create group mean CVR and M maps for both datasets. Statistical comparisons between E2 and MEC data were made using a paired t-test for CVR and M. To account for some subjects having multiple scans, values were averaged for those subjects and the averaged values were used for the subsequent t-test.

Repeatability Analysis. Repeatability of the BH activation, CVR, and M was also analysed. In total, ten subjects had two scans. Repeatability of the BH activation was evaluated using the Dice coefficient, computed using Equation (5), which provides a measure of the degree of overlap of active voxels between scans collected at different times. Here, A_1 and A_2 represent thresholded activation maps ($p < 0.001$, uncorrected) from time point 1 (TP1) and time point 2 (TP2), respectively. The repeatability of the CVR and M measurements was analysed across voxels using Equation (6), where N is the number of voxels and x is either CVR or M. The Dice coefficient and the mean repeatability of CVR and M were compared between the E2, MEC, and PW data.

$$D = \frac{2|A_1 \cap A_2|}{|A_1| + |A_2|} \tag{5}$$

$$Repeatability = 1 - \left(\frac{1}{N}\right) \sum_{i \in brain}^{N} abs\left(\frac{x_i^{TP1} - x_i^{TP2}}{x_i^{TP1} + x_i^{TP2}}\right) \tag{6}$$

A spatial correlation analysis was performed by correlating CVR and M from TP1 with TP2. Correlation values were determined using Pearson correlation on an individual subject basis, and then compared between E2 and MEC datasets using a paired t-test following transformation to Fisher's z-score.

Finally, test-retest reliability was analyzed using the intraclass correlation coefficient (ICC). Specifically, ICC(3,1) was used. ICC(3,1) ranges from 0 to 1 with a value of 1 indicating perfect reliability. ICC(3,1) was computed on a voxelwise basis using *3dLME* in AFNI. Mean values of repeatability and ICC(3,1) were extracted from GM. There remains some debate as to what constitutes a "reliable" ICC value. One often quoted guideline classifies $ICC < 0.4$ as poor, $0.4 < ICC < 0.6$ as fair, $0.6 < ICC < 0.75$ as good, and $ICC > 0.75$ as excellent[53]. Thus, ICC maps were thresholded at 0.4 and 0.6, and the percentage of GM voxels meeting those thresholds was computed.

GM/WM Contrast. CVR and M tend to be higher in GM compared to WM. GM/WM contrast was estimated by dividing the mean CVR and M values in GM by the mean CVR and M values in WM respectively. Thus, higher values represent higher GM/WM contrast. GM/WM contrast was compared between the E2 and MEC datasets using Bonferroni-corrected paired t-tests to compare individual means.

Results

tSNR. The tSNR was computed in GM following preprocessing, and was significantly higher for the MEC vs. E2 data (103.4 ± 20.0 vs. 75.9 ± 15.5, $p < 1e\text{-}10$). tSNR for the PW data was 3.17 ± 0.83. These values were computed to verify that echo combination improved tSNR. The values themselves are not necessarily meaningful.

BH Activation. In general, BH activation was widespread and tended to be higher in GM. Qualitatively, activation volume and strength were higher for the MEC data compared with the E2 and PW data. Example BH activation maps from one representative subject are shown in Fig. 1. These trends were also seen in the quantitative data (Table 1). In general, the BH activation t-score was higher for MEC vs. E2 and PW data. The fraction of GM

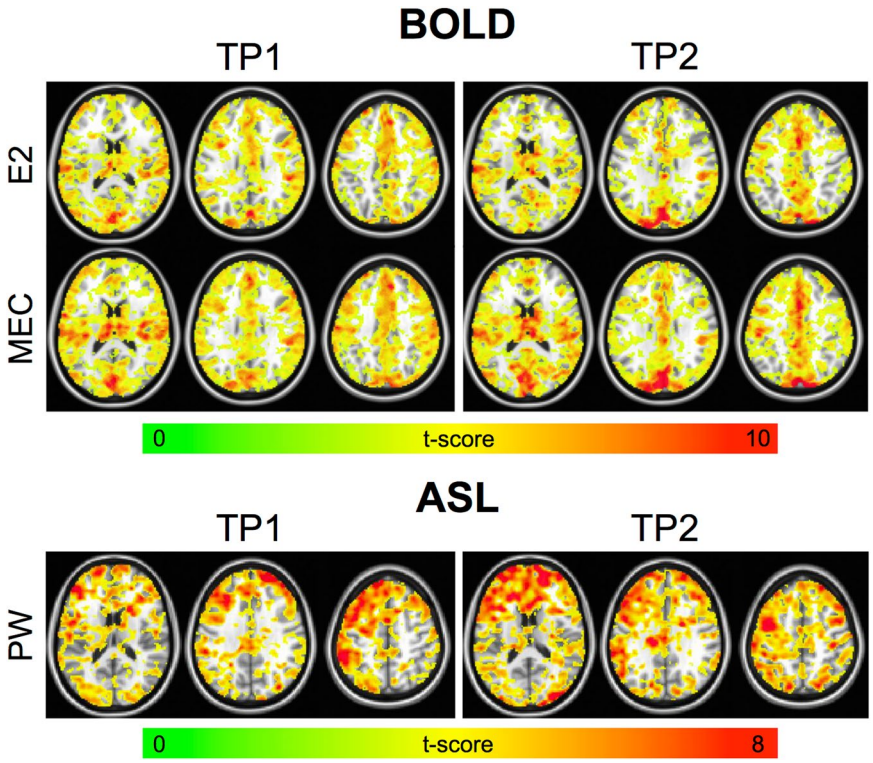

Figure 1. BH activation maps for one representative subject. Results are shown for the E2, MEC, and PW datasets. The MEC data showed higher activation strength and volume compared with the E2 and PW data. More activation was seen in GM compared with WM in all cases; however, more WM activation was seen for the MEC data.

	t-score, GM	t-score, Overlap	Fraction of Active Voxels	Variance Explained
E2	3.75 (0.73)	5.22 (0.57)	0.440 (0.133)	0.213 (0.054)
MEC	4.20 (0.78)	5.68 (0.75)	0.526 (0.136)	0.243 (0.063)
PW	3.70 (0.77)	5.07 (0.42)	0.476 (0.133)	0.189 (0.050)
P-value	**,***P < 0.001	**,***P < 0.001	**P < 1e-10	**,***P < 0.001

Table 1. Comparison of quantitative results for E2 and MECDN data. Parentheses indicate standard deviation. Abbreviations: GM = gray matter; E2 = single echo (Echo 2, TE = 25ms); MEC = Multiecho combined; P-values are the results of a Bonferroni corrected paired t-test. *E2 > PW, **MEC > E2, ***MEC > PW.

voxels that were active was significantly higher for the MEC vs. E2 and PW data. Finally, the amount of variance explained by the regressor was significantly higher for the MEC vs. E2 and PW data.

Mean BOLD time series extracted from GM and active voxels are shown in Fig. 2a. In general, combined-echo time series were cleaner with less variance across subjects compared with the single-echo data. In particular, this can be seen in the "barely" active voxels ($0.01 < p < 0.05$). Signal fits from a representative voxel are shown in Fig. 2b for BOLD and PW data.

CVR. Units for CVR measurements are percent. Mean GM CVR_{MEC} was significantly less than CVR_{E2} (1.35 ± 0.21 vs. 1.69 ± 0.33, $p = 6.0e\text{-}7$). Example individual-subject and group-averaged BOLD CVR maps are shown in Fig. 3. BOLD CVR maps were relatively robust across subjects and time points, showing similar patterns. Higher GM/WM contrast was seen for CVR_{MEC} compared with CVR_{E2}, which can be observed on both the individual subject and group maps.

Mean GM $CVR_{PW} = 78.3 \pm 27.4$. CVR_{PW} maps were much noisier and less robust across time and subjects than the BOLD maps (Fig. 3). Furthermore, low CBF in WM coupled with the inherently low SNR of ASL led to spuriously high CVR in the WM. Thus, CVR_{PW} is only displayed in GM. The group maps show similar patterns across time points. For example, heightened CVR is observed in the posterior cingulate cortex.

M. Units for M measurements are also percent. The M results mirrored the CVR results. As with CVR, mean M extracted from GM was significantly lower for M_{MEC} vs. M_{E2} (4.4 ± 0.9 vs. 5.1 ± 1.0, $p = 7.0e\text{-}7$). Example group-averaged M maps are shown in Fig. 4. Higher GM/WM contrast was seen for the M_{MEC} data versus the M_{E2}

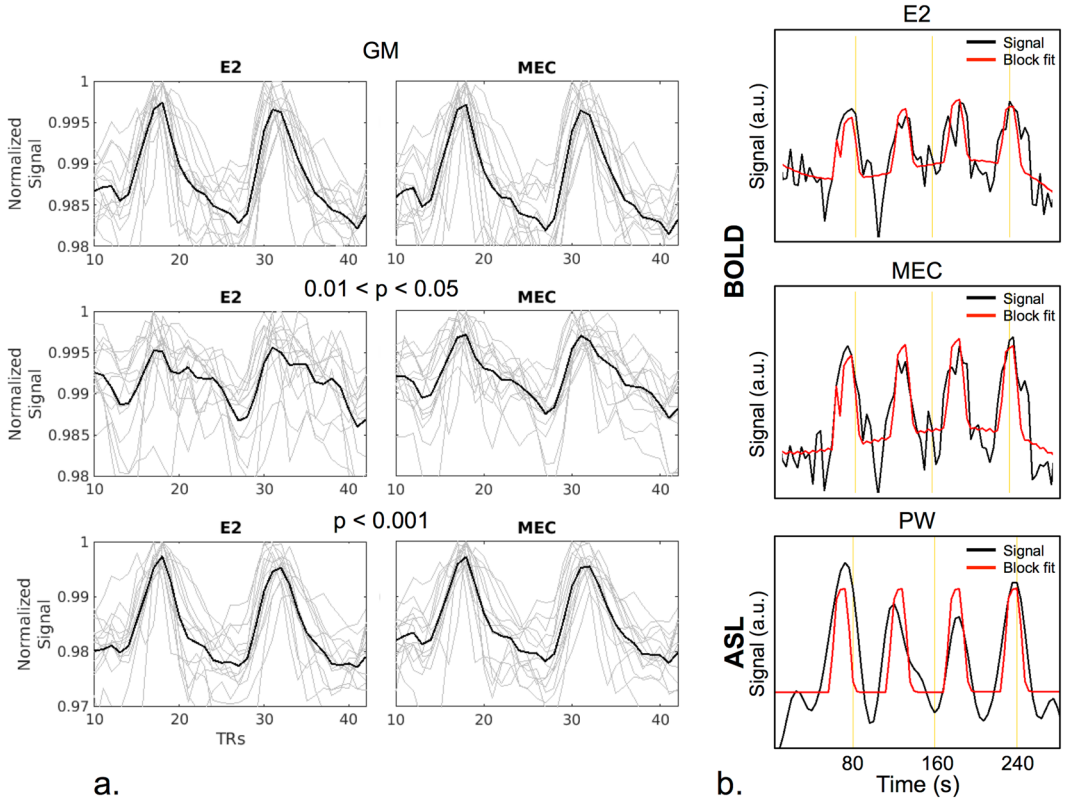

Figure 2. BOLD and PW time series. (**a**) Mean BOLD time series extracted from the GM (top), voxels with $0.01 < p < 0.05$ (middle), and voxels with $p < 0.001$ (bottom). Light grey curves show individual subject results and black curves show the group mean. Signals are shown following regression of label/control oscillations and detrending. Slight qualitative improvements are seen for MEC vs. E2 datasets for the GM and $p < 0.001$ cases. A significant improvement is seen for the MEC data compared with the E2 data for the $0.01 < p < 0.05$ case. The signal is cleaner and there is less variance across subjects. Active voxels were defined in the E2 dataset, and those same voxels were used to extract the MEC time series. (**b**) Example BOLD and PW time series and fits from a representative voxel. The MEC data are cleaner compared with the E2 data. The PW fit was less accurate compared with the BOLD fits.

data. This can be seen on both the individual subject and group maps. M also varied substantially across the brain, with higher values in the visual cortex, the default mode network, and major blood vessels.

Repeatability and GM/WM contrast. Repeatability and GM/WM contrast results are shown in Table 2. The Dice coefficient was significantly higher for the MEC vs. E2 data and for the BOLD vs. PW data. CVR and M repeatability was significantly higher for the MEC vs. E2 data and for the BOLD vs. PW data. CVR and M GM/WM contrast were also significantly higher for the MEC data compared with the E2 data.

The spatial correlation of CVR between time points is shown for the BOLD data in Fig. 5 for one representative subject and averaged across subjects. CVR_{MEC} had significantly stronger spatial correlation compared with CVR_{E2} (Fisher's z-score, 1.03 ± 0.13 vs. 0.75 ± 0.13, $p = 2.2e\text{-}4$). The same trend was observed for M with M_{MEC} having a significantly stronger spatial correlation compared to M_{E2} (Fisher's z-score, 0.51 ± 0.14 vs. 0.36 ± 0.08, $p = 8.2e\text{-}5$). Spatial correlation for CVR_{PW} (0.29 ± 0.07) was lower compared to BOLD CVR.

The mean GM ICC for CVR and M for E2 and MEC datasets fell in the "fair" reliability category. ICC(3,1) for the CVR data was higher for the MEC vs. E2 data (0.50 vs. 0.44), and ICC(3,1) for the M data was also higher for the MEC vs. E2 data (0.40 vs. 0.36), but lower compared to CVR. For CVR_{E2}, 45.3% and 24.0% of GM voxels had ICCs > 0.4 and 0.6, respectively compared with 55.6% and 32.8%, respectively, for CVR_{MEC}. For M_{E2}, 30.4% and 14.2% of GM voxels had ICCs > 0.4 and 0.6, respectively, compared with 37.7% and 19.0%, respectively, for M_{MEC}. ICC(3,1) for the PW data was "fair" (0.41).

Discussion

In this study, an MBME ASL/BOLD sequence was used to acquire BH fMRI data from 14 volunteers. Ten of the volunteers were able to return and had usable repeat scans acquired within two weeks of their initial scans. BOLD and PW BH activation statistics and CVR were computed. The multi-echo BOLD data was combined using the T2*-weighted technique. In addition, the PW data was used to compute the "M" parameter in the Davis model. Repeatability and GM/WM contrast of CVR and M were also evaluated. BH activation strength, volume, and repeatability increased for the combined-echo data compared to the single echo data. Repeatability and GM/WM contrast of CVR and M were higher for the combined-echo data compared to the single echo data. The PW data

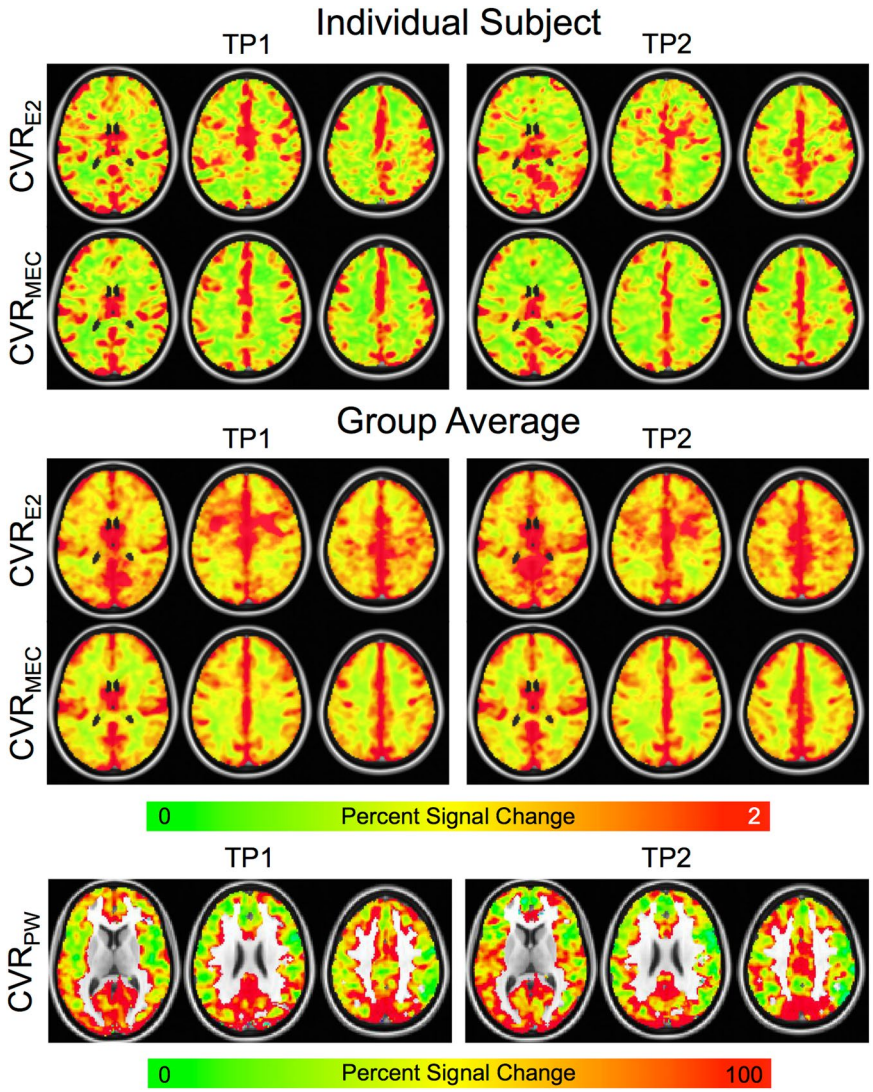

Figure 3. BOLD and PW CVR maps from the two time points. Individual subject (top) and group (middle) BOLD CVR maps are shown. The CVR_{MEC} maps appear cleaner with higher GM/WM contrast compared with the CVR_{E2} maps. The CVR_{MEC} maps are also more similar across time points compared with the CVR_{E2} maps. This can be seen on both an individual subject and group basis. Group CVR_{PW} maps (bottom) are also shown from the two time points. CVR_{PW} maps were much noisier and less robust across time and subjects than the BOLD CVR maps; however, the group maps do show similar patterns across time points. For example, heightened CVR was seen in the posterior cingulate cortex and visual cortex.

was noisier and less robust compared to the BOLD data. Overall, this study showed improved CVR and M maps could be acquired using a multi-echo approach.

It is important to note that the data in this study were collected using an MBME ASL/BOLD sequence that included a pCASL tagging module at the start. For the BOLD analysis, to remove the ASL effects from the echoes, the label/control signal oscillations were regressed from the data prior to running the general linear model. This was accomplished by including a column of alternating −1 s and 1 s in the design matrix and has been used in previous dual-echo ASL/BOLD studies[35,54,55]. One downside to the pCASL labelling module is a lengthening of the TR. In this study, TR = 4.0 s, which is relatively long for a BOLD fMRI study. Although only 73 functional time points were acquired, the collection of four echoes, while slightly increasing the TR, led to a significant increase in the tSNR. This can help compensate for the long TR, because fewer time points are needed to detect activation with a higher tSNR[56]. Previous work using the MBME ASL/BOLD sequence also has demonstrated this relationship[36]. Despite the long TR, robust BH activation was seen in both the E2 and MEC datasets.

In this study, a subset of subjects was imaged twice. Repeatability of BH activation was analysed using the Dice coefficient, and the repeatability of CVR and M was computed using a modified percent difference (Equation (6)) and the ICC. The Dice coefficient is dependent on the activation threshold (i.e., p-value), and the reliability of a similar spatial overlap method[57] has been shown to decrease with an increasing threshold[58]. The Dice coefficient was higher for the MEC vs. E2 data. In fact, on average, 68% of active voxels overlapped between the two time

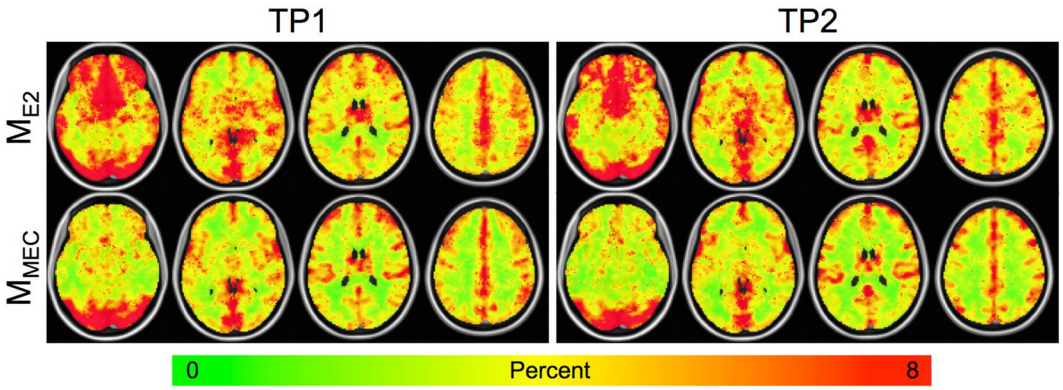

Figure 4. Group-averaged M maps. As with the CVR maps, the M_{MEC} maps appear cleaner with higher GM/WM contrast compared to the M_{E2} maps, and are more similar across time points compared with the M_{E2} maps. M varies substantially across the brain with higher values in the visual cortex and blood vessels.

	Dice Coefficient	Repeatability	Reliability (ICC)	GM/WM Contrast
CVR_{E2}	0.562 (0.134)	0.748 (0.022)	0.442 (0.242)	1.382 (0.159)
CVR_{MEC}	0.678 (0.115)	0.786 (0.030)	0.497 (0.245)	1.660 (0.215)
CVR_{PW}	0.516 (0.133)	0.702 (0.033)	0.413 (0.248)	N/A
P-Value	**,***P < 0.001	*,**,***P < 0.01	N/A	**P < 1e-10
M_{E2}	N/A	0.711 (0.023)	0.361 (0.239)	1.484 (0.487)
M_{MEC}		0.742 (0.026)	0.399 (0.247)	1.864 (0.649)
P-Value		P = 1.5e-5	N/A	P < 1e-10

Table 2. Group averages for reproducibility and sensitivity/specificity results. The Dice coefficient was computed using active voxels (P < 0.001, uncorrected). Mean reproducibility and reliability was extracted from gray matter. Abbreviations: E2 = single echo (Echo 2, TE = 25ms); MEC = Multiecho combined; ICC = intraclass correlation coefficient; PW = Perfusion Weighted; PWDN = Perfusion Weighted Denoised; N.S. = Not Significant, $^{*}CVR_{E2} > CVR_{PW}$, $^{**}CVR_{MEC} > CVR_{E2}$, $^{***}CVR_{MEC} > CVR_{PW}$.

points for the MEC data and 56% of active voxels overlapped for the E2 data at a stringent threshold of P = 0.001. These values are similar to or greater than those for the Dice coefficients for most fMRI studies[54,59,60]. This may be mostly due to the large volume of active voxels (>67% on average); nonetheless, it indicates BOLD BH activation is repeatable even without end-tidal CO_2 measurements. Similar results were observed for the repeatability of CVR and M, with repeatability of 78.6% and 73.5%, respectively for M_{MEC}. Overall, repeatability metrics were lower for the PW compared with BOLD data. One study calculated the reproducibility of functional connectivity density using the same metric in Equation (4), and found reproducibility ranging from 59–88%[58].

CVR was calculated as the percent signal change resulting from the BH response activation. In addition to qualitative improvements in the quality of CVR_{MEC} maps at the group level, noticeable improvements were also seen at the individual level (Fig. 3). Higher GM/WM contrast was observed for the CVR_{MEC} maps compared with the CVR_{E2} maps. This can also be seen in Fig. 4a, where MEC signal traces for the individual subjects (light grey lines) show a reduced spread compared with the single echo signal traces, especially for the low activation case. Mean CVR_{MEC} and M_{MEC} were reduced compared with CVR_{E2} and M_{E2}, respectively. This was caused by the averaging of the echoes. Signal from shorter echo times, with a lower percent signal change, was averaged with the signal from longer echo times, with a higher percent signal change. Since, in general, the signal is higher at shorter echo times, this led to an overall reduced percent signal change.

The mean GM/WM contrast was computed. This metric assumes higher CVR in grey matter than white matter. Therefore increased values were deemed desirable. This metric also provides a measure that can be compared between data types (i.e., E2 vs. MEC), and, in general, the assumption of low CVR in WM has been verified[4,11,61]. Here, we found echo combination significantly increased both CVR and M GM/WM contrast. For the PW data, spuriously high CVR was seen in WM. Therefore, CVR_{PW} GM/WM contrast was not computed.

The BOLD signal response to BH has been shown to increase with BH duration[4,21]. For example, Magon *et al.* collected data with BH durations of 9, 15, and 21s[21]. They found higher signal amplitude and reproducibility using 21 s BH durations but noted that BH durations of 15 s resulted in acceptable reproducibility across sessions and seemed "to be the best paradigm to catch the variability of the response of the population"[21]. We chose a BH duration of 16 s with the goal of producing robust signal changes while still being feasible for the vast majority of subjects to complete.

Murphy *et al.* performed a comprehensive analysis of nine different regressors used to model the BH response[1]. They found the sine/cosine regressor explained as much variance as the end-tidal CO_2 regressor. For

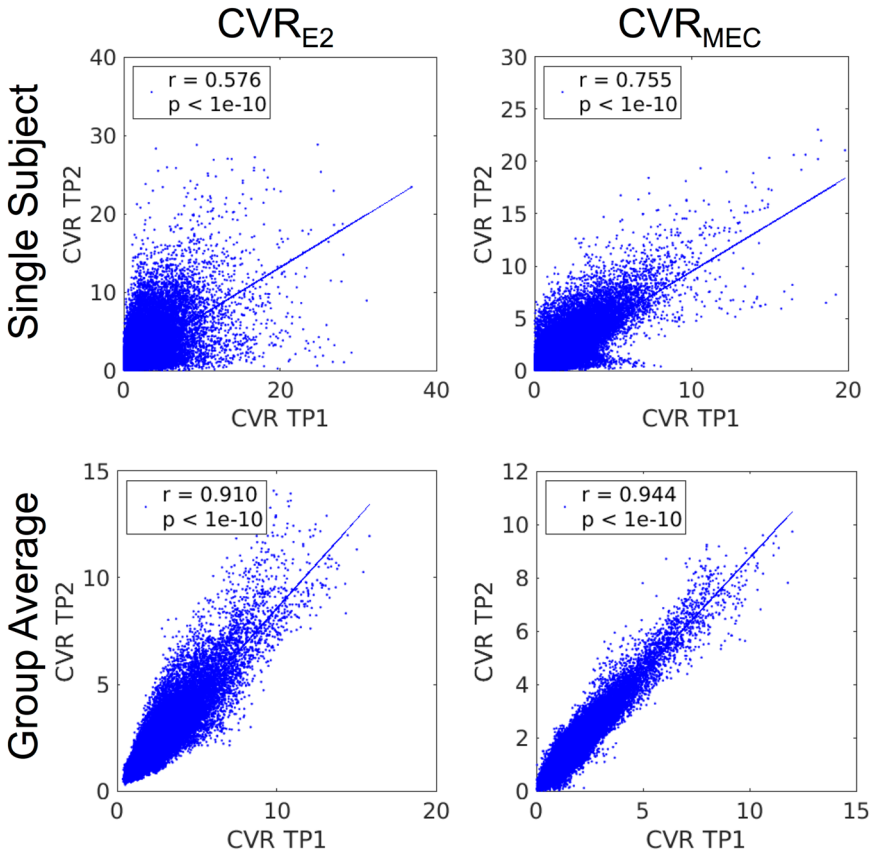

Figure 5. Spatial correlation of BOLD CVR between time points for one representative subject (top) and for the group CVR maps (bottom) for the E2 (left) and MEC (right) datasets. A strong correlation between time points is seen in both datasets at the individual and group levels; however, an increased correlation is seen for the MEC data for the single subject and group cases. Less spread in the CVR values is seen for the MEC data.

our study, the BH duration was much shorter than the post-BH breathing duration (16 s vs. 44 s, respectively) compared with 20 s vs. 30 s, respectively, reported in Murphy *et al.*[1]. Thus, the sine wave model was not ideal. As a result, a voxelwise phase-shifted square wave convolved with a double-gamma variate HRF was used in this study.

The simultaneous collection of ASL and BOLD data allowed CVR to be calculated with two different imaging contrasts from a single acquisition. ASL offers several advantages over BOLD imaging. First, ASL provides a direct and potentially quantitative measure of blood flow. Baseline CBF can also be measured. However, ASL suffers from low SNR, severely reducing the quality of the ASL CVR maps. In general, ASL CVR maps were of lower quality than the BOLD CVR maps, especially at the individual level. This was especially true in the white matter, where low CBF and SNR resulted in spuriously high CVR. Activation strength, volume, and repeatability were all lower compared with the respective BOLD measures. Future studies of ASL CVR in disease may benefit from a regional analysis to boost SNR. Furthermore, a longer period of rest at the beginning of the scan may provide a more accurate measurement of baseline CBF. Background suppression, in which the background signal is reduced using saturation and inversion pulses, could also be employed to increase ASL SNR. One recent study recommended background suppression for 2D dual-echo ASL acquisitions[62] after finding that the large CBF signal gains offset the slight BOLD sensitivity losses.

The simultaneous collection of ASL and BOLD data also allowed "M" to be calculated. This calculation involves two parameters that must be assumed in the model: α and β. The values for these parameters vary in the literature, with α typically ranging from 0.2–0.38[48] and β ranging from 1–1.5[49–52]. Recent research has supported lower values for these parameters. For example, Griffeth and Buxton found the accuracy of the Davis model was better when using the optimized values of $\alpha = 0.14$ and $\beta = 0.91$[50]. The computation of M only relies on the difference between these parameters, and changing these parameters should only change the quantitative value of M. The M analysis was repeated with $\beta = 1.3$ (not shown), and the findings of improved repeatability, sensitivity, and specificity with echo combination remained valid.

The value of M also varies widely in the literature, ranging from ~4–15% at 3T[63], and is dependent on a number of factors other than α and β including field strength and TE[64–66]. Our calculated values of M were toward the low end of that range. This is likely due to the shorter than average TE value (compared to the literature) for the second echo and T2* echo combination as M is linearly dependent on TE[66]. We also saw a heterogeneous distribution of M values throughout the brain, with increased values in GM and visual cortex, indicating a voxelwise measurement of M is necessary for accurate $CMRO_2$ calculation. Very few publications show maps of M. One

study reported M maps and had a similarly heterogeneous appearance[67]. Other studies also found heightened M values in GM and the visual cortex using a CO_2 gas inhalation challenge[17,68].

A robust, repeatable measurement of M is important as errors in M can propagate to errors in $CMRO_2$, which can affect calibrated fMRI and neurovascular coupling measurements. Studies have shown that CBF, $CMRO_2$, and their coupling are dependent on the baseline physiology of the brain and can affect the measured BOLD response. In fact, neurovascular coupling has shown to be altered with normal aging during childhood[69], Alzheimer's disease[70], and caffeine intake[71], and in tumours[14,72]. The repeatability of M was lower compared to CVR. This was likely because it relies on ASL, which suffers from low SNR thus increasing the noise of the M measurements.

This study was not without limitations. First, the TR was relatively long (4.0 s) due to inclusion of the pCASL module. This reduced the number of time points acquired. Because the ability to detect changes of a certain effect size increases with the number of time points, our statistical power may have been limited. However, robust activation statistics were still obtained, and collecting and combining four echoes increased tSNR and greatly compensated for this effect. Birn *et al*. found the Dice coefficient and ICC of resting-state networks increased with scan time[59]. Future studies should examine the effects of multiple echoes using a pure BOLD fMRI acquisition with a shorter TR and possibly longer imaging time. Also, we did not have the capability to measure end-tidal CO_2. Thus, all measures of CVR are only semiquantitative measures of percent signal change, and repeatability and reliability may be somewhat limited. Further, it was not possible to untangle motion effects from true arterial CO_2 changes. We tried to maximize data reliability in the absence of CO_2 measures by using end-expiration breath holds and paced breathing between breath holds, which have been shown to increase BH repeatability[19]. Ideally, a precisely controlled prospective end-tidal gas delivery system with end-tidal CO_2 output monitoring should be used[73]; however, these systems are not available at all institutions. Additional studies should examine the effects of multiple echoes using gas inhalation techniques with end-tidal CO_2 measurements. Regardless, the benefits of collecting multiple echoes for BH activation, CVR, and M measurements should still be valid. Finally, there are several concerns regarding translating this technique to the clinic. Individual subject CVR and M maps showed greater spatial variability compared to group maps. This could impact clinical translation where individual repeatability is critical. Furthermore, clinical translation to patients with cerebrovascular disease may be problematic as patients become hypoxic at different rates during BH, which lowers the BOLD signal[74]. Absolute arterial CO_2 values also may be necessary. Additional studies are needed in patients with cerebrovascular disease.

In conclusion, we evaluated BH activation and CVR and M repeatability using an MBME ASL/BOLD sequence. We found that echo combination led to higher BOLD activation strength, volume, and repeatability, and higher CVR and M repeatability and reliability. We also compared two models for computing BOLD BH activation. These results suggest ME approaches are advantageous for computing BOLD activation, CVR, and M using BH fMRI.

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Murphy, K., Harris, A. D. & Wise, R. G. Robustly measuring vascular reactivity differences with breath-hold: normalising stimulus-evoked and resting state BOLD fMRI data. *NeuroImage* **54**, 369–379, https://doi.org/10.1016/j.neuroimage.2010.07.059 (2011).
2. van Niftrik, C. H. *et al*. Fine tuning breath-hold-based cerebrovascular reactivity analysis models. *Brain and behavior* **6**, e00426, https://doi.org/10.1002/brb3.426 (2016).
3. Thomason, M. E. & Glover, G. H. Controlled inspiration depth reduces variance in breath-holding-induced BOLD signal. *Neuroimage* **39**, 206–214, https://doi.org/10.1016/j.neuroimage.2007.08.014 (2008).
4. Bright, M. G. & Murphy, K. Reliable quantification of BOLD fMRI cerebrovascular reactivity despite poor breath-hold performance. *NeuroImage* **83**, 559–568, https://doi.org/10.1016/j.neuroimage.2013.07.007 (2013).
5. Kastrup, A., Kruger, G., Neumann-Haefelin, T. & Moseley, M. E. Assessment of cerebrovascular reactivity with functional magnetic resonance imaging: comparison of CO(2) and breath holding. *Magn Reson Imaging* **19**, 13–20 (2001).
6. Lythgoe, D. J., Williams, S. C., Cullinane, M. & Markus, H. S. Mapping of cerebrovascular reactivity using BOLD magnetic resonance imaging. *Magn Reson Imaging* **17**, 495–502 (1999).
7. Zhou, Y., Rodgers, Z. B. & Kuo, A. H. Cerebrovascular reactivity measured with arterial spin labeling and blood oxygen level dependent techniques. *Magn Reson Imaging* **33**, 566–576, https://doi.org/10.1016/j.mri.2015.02.018 (2015).
8. Tancredi, F. B. & Hoge, R. D. Comparison of cerebral vascular reactivity measures obtained using breath-holding and CO_2 inhalation. *Journal of cerebral blood flow and metabolism: official journal of the International Society of Cerebral Blood Flow and Metabolism* **33**, 1066–1074, https://doi.org/10.1038/jcbfm.2013.48 (2013).
9. Mark, C. I. *et al*. Precise control of end-tidal carbon dioxide and oxygen improves BOLD and ASL cerebrovascular reactivity measures. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine* **64**, 749–756, https://doi.org/10.1002/mrm.22405 (2010).
10. Poublanc, J. *et al*. Measuring cerebrovascular reactivity: the dynamic response to a step hypercapnic stimulus. *J Cereb Blood Flow Metab* **35**, 1746–1756, https://doi.org/10.1038/jcbfm.2015.114 (2015).
11. Fisher, J. A. *et al*. Assessing cerebrovascular reactivity by the pattern of response to progressive hypercapnia. *Hum Brain Mapp*, https://doi.org/10.1002/hbm.23598 (2017).
12. Donahue, M. J. *et al*. Routine clinical evaluation of cerebrovascular reserve capacity using carbogen in patients with intracranial stenosis. *Stroke* **45**, 2335–2341, https://doi.org/10.1161/STROKEAHA.114.005975 (2014).
13. Blockley, N. P., Harkin, J. W. & Bulte, D. P. Rapid cerebrovascular reactivity mapping: Enabling vascular reactivity information to be routinely acquired. *NeuroImage* **159**, 214–223, https://doi.org/10.1016/j.neuroimage.2017.07.048 (2017).
14. Pillai, J. J. & Mikulis, D. J. Cerebrovascular reactivity mapping: an evolving standard for clinical functional imaging. *AJNR. American journal of neuroradiology* **36**, 7–13, https://doi.org/10.3174/ajnr.A3941 (2015).
15. Mandell, D. M. *et al*. Mapping cerebrovascular reactivity using blood oxygen level-dependent MRI in Patients with arterial steno-occlusive disease: comparison with arterial spin labeling MRI. *Stroke* **39**, 2021–2028, https://doi.org/10.1161/STROKEAHA.107.506709 (2008).
16. Faraco, C. C. *et al*. Dual echo vessel-encoded ASL for simultaneous BOLD and CBF reactivity assessment in patients with ischemic cerebrovascular disease. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine* **73**, 1579–1592, https://doi.org/10.1002/mrm.25268 (2015).
17. Davis, T. L., Kwong, K. K., Weisskoff, R. M. & Rosen, B. R. Calibrated functional MRI: mapping the dynamics of oxidative metabolism. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 1834–1839 (1998).

18. Blockley, N. P., Griffeth, V. E. & Buxton, R. B. A general analysis of calibrated BOLD methodology for measuring CMRO2 responses: comparison of a new approach with existing methods. *NeuroImage* **60**, 279–289, https://doi.org/10.1016/j.neuroimage.2011.11.081 (2012).
19. Scouten, A. & Schwarzbauer, C. Paced respiration with end-expiration technique offers superior BOLD signal repeatability for breath-hold studies. *NeuroImage* **43**, 250–257, https://doi.org/10.1016/j.neuroimage.2008.03.052 (2008).
20. Li, T. Q., Kastrup, A., Takahashi, A. M. & Moseley, M. E. Functional MRI of human brain during breath holding by BOLD and FAIR techniques. *NeuroImage* **9**, 243–249, https://doi.org/10.1006/nimg.1998.0399 (1999).
21. Magon, S. *et al*. Reproducibility of BOLD signal change induced by breath holding. *Neuroimage* **45**, 702–712, https://doi.org/10.1016/j.neuroimage.2008.12.059 (2009).
22. Moreton, F. C., Dani, K. A., Goutcher, C., O'Hare, K. & Muir, K. W. Respiratory challenge MRI: Practical aspects. *Neuroimage Clin* **11**, 667–677, https://doi.org/10.1016/j.nicl.2016.05.003 (2016).
23. Chen, L. *et al*. Evaluation of highly accelerated simultaneous multi-slice EPI for fMRI. *NeuroImage* **104**, 452–459, https://doi.org/10.1016/j.neuroimage.2014.10.027 (2015).
24. Feinberg, D. A., Setsompop, K. & Ultra-fast, M. R. I. of the human brain with simultaneous multi-slice imaging. *J. Magn. Reson.* **229**, 90–100, https://doi.org/10.1016/j.jmr.2013.02.002 (2013).
25. Moeller, S. *et al*. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine* **63**, 1144–1153, https://doi.org/10.1002/mrm.22361 (2010).
26. Setsompop, K. *et al*. Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magn. Reson. Med.* **67**, 1210–1224, https://doi.org/10.1002/mrm.23097 (2012).
27. Ravi, H., Liu, P., Peng, S. L., Liu, H. & Lu, H. Simultaneous multi-slice (SMS) acquisition enhances the sensitivity of hemodynamic mapping using gas challenges. *NMR Biomed* **29**, 1511–1518, https://doi.org/10.1002/nbm.3600 (2016).
28. Poser, B. A. & Norris, D. G. Investigating the benefits of multi-echo EPI for fMRI at 7 T. *Neuroimage* **45**, 1162–1172, https://doi.org/10.1016/j.neuroimage.2009.01.007 (2009).
29. Kundu, P., Inati, S. J., Evans, J. W., Luh, W. M. & Bandettini, P. A. Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *NeuroImage* **60**, 1759–1770, https://doi.org/10.1016/j.neuroimage.2011.12.028 (2012).
30. Posse, S. Multi-echo acquisition. *NeuroImage* **62**, 665–671, https://doi.org/10.1016/j.neuroimage.2011.10.057 (2012).
31. Kundu, P. *et al*. Integrated strategy for improving functional connectivity mapping using multiecho fMRI. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 16187–16192, https://doi.org/10.1073/pnas.1301725110 (2013).
32. Kundu, P., Santin, M. D., Bandettini, P. A., Bullmore, E. T. & Petiet, A. Differentiating BOLD and non-BOLD signals in fMRI time series from anesthetized rats using multi-echo EPI at 11.7 T. *NeuroImage* **102**, 861–874 (2014).
33. Kundu, P. *et al*. Robust resting state fMRI processing for studies on typical brain development based on multi-echo EPI acquisition. *Brain Imaging Behav* **9**, 56–73, https://doi.org/10.1007/s11682-014-9346-4 (2015).
34. Poser, B. A., Versluis, M. J., Hoogduin, J. M. & Norris, D. G. BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: parallel-acquired inhomogeneity-desensitized fMRI. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine* **55**, 1227–1235, https://doi.org/10.1002/mrm.20900 (2006).
35. Cohen, A. D., Nencka, A. S., Lebel, R. M. & Wang, Y. Multiband multi-echo imaging of simultaneous oxygenation and flow timeseries for resting state connectivity. *PloS one* **12**, e0169253, https://doi.org/10.1371/journal.pone.0169253 (2017).
36. Cohen, A. D., Nencka, A. S. & Wang, Y. Multiband multi-echo simultaneous ASL/BOLD for task-induced functional MRI. *PloS one* **13**, e0190427, https://doi.org/10.1371/journal.pone.0190427 (2018).
37. Griswold, M. A. *et al*. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn. Reson. Med.* **47**, 1202–1210, https://doi.org/10.1002/mrm.10171 (2002).
38. Noll, D. C., Nishimura, D. G. & Macovski, A. Homodyne detection in magnetic resonance imaging. *IEEE Trans. Med. Imaging* **10**, 154–163, https://doi.org/10.1109/42.79473 (1991).
39. Zhang, Y., Brady, M. & Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* **20**, 45–57, https://doi.org/10.1109/42.906424 (2001).
40. Cox, R. W. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and biomedical research, an international journal* **29**, 162–173 (1996).
41. Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W. & Smith, S. M. Fsl. *Neuroimage* **62**, 782–790, https://doi.org/10.1016/j.neuroimage.2011.09.015 (2012).
42. Chuang, K. H. *et al*. Mapping resting-state functional connectivity using perfusion MRI. *NeuroImage* **40**, 1595–1605, https://doi.org/10.1016/j.neuroimage.2008.01.006 (2008).
43. Posse, S. *et al*. Enhancement of BOLD-contrast sensitivity by single-shot multi-echo functional MR imaging. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine* **42**, 87–97 (1999).
44. Aguirre, G. K., Detre, J. A., Zarahn, E. & Alsop, D. C. Experimental design and the relative sensitivity of BOLD and perfusion fMRI. *NeuroImage* **15**, 488–500, https://doi.org/10.1006/nimg.2001.0990 (2002).
45. Kastrup, A., Li, T. Q., Glover, G. H. & Moseley, M. E. Cerebral blood flow-related signal changes during breath-holding. *AJNR Am J Neuroradiol* **20**, 1233–1238 (1999).
46. Birn, R. M., Smith, M. A., Jones, T. B. & Bandettini, P. A. The respiration response function: the temporal dynamics of fMRI signal fluctuations related to changes in respiration. *NeuroImage* **40**, 644–654, https://doi.org/10.1016/j.neuroimage.2007.11.059 (2008).
47. Bright, M. G., Bulte, D. P., Jezzard, P. & Duyn, J. H. Characterization of regional heterogeneity in cerebrovascular reactivity dynamics using novel hypocapnia task and BOLD fMRI. *NeuroImage* **48**, 166–175, https://doi.org/10.1016/j.neuroimage.2009.05.026 (2009).
48. Chen, J. J. & Pike, G. B. BOLD-specific cerebral blood volume and blood flow changes during neuronal activation in humans. *NMR in biomedicine* **22**, 1054–1062, https://doi.org/10.1002/nbm.1411 (2009).
49. Mark, C. I., Fisher, J. A. & Pike, G. B. Improved fMRI calibration: precisely controlled hyperoxic versus hypercapnic stimuli. *NeuroImage* **54**, 1102–1111, https://doi.org/10.1016/j.neuroimage.2010.08.070 (2011).
50. Griffeth, V. E. & Buxton, R. B. 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, https://doi.org/10.1016/j.neuroimage.2011.05.077 (2011).
51. Shu, C. Y. *et al*. Quantitative beta mapping for calibrated fMRI. *NeuroImage* **126**, 219–228, https://doi.org/10.1016/j.neuroimage.2015.11.042 (2016).
52. Boxerman, J. L., Hamberg, L. M., Rosen, B. R. & Weisskoff, R. M. MR contrast due to intravascular magnetic susceptibility perturbations. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine* **34**, 555–566 (1995).
53. Cicchetti, D. V. *Guidelines, criteria and rules of thumb for evaluating normed and standardized assessment instruments in psychology*. (American Psychological Association, 1994).
54. Zhu, S., Fang, Z., Hu, S., Wang, Z. & Rao, H. Resting state brain function analysis using concurrent BOLD in ASL perfusion fMRI. *Plos One* **8**, e65884, https://doi.org/10.1371/journal.pone.0065884 (2013).

55. Wang, Z. Improving cerebral blood flow quantification for arterial spin labeled perfusion MRI by removing residual motion artifacts and global signal fluctuations. *Magn Reson Imaging* **30**, 1409–1415, https://doi.org/10.1016/j.mri.2012.05.004 (2012).
56. Murphy, K., Bodurka, J. & Bandettini, P. A. How long to scan? The relationship between fMRI temporal signal to noise ratio and necessary scan duration. *NeuroImage* **34**, 565–574, https://doi.org/10.1016/j.neuroimage.2006.09.032 (2007).
57. Duncan, K. J., Pattamadilok, C., Knierim, I. & Devlin, J. T. Consistency and variability in functional localisers. *Neuroimage* **46**, 1018–1026, https://doi.org/10.1016/j.neuroimage.2009.03.014 (2009).
58. Tomasi, D., Shokri-Kojori, E. & Volkow, N. D. High-Resolution Functional Connectivity Density: Hub Locations, Sensitivity, Specificity, Reproducibility, and Reliability. *Cereb Cortex* **26**, 3249–3259, https://doi.org/10.1093/cercor/bhv171 (2016).
59. Birn, R. M. *et al*. The effect of scan length on the reliability of resting-state fMRI connectivity estimates. *Neuroimage* **83**, 550–558, https://doi.org/10.1016/j.neuroimage.2013.05.099 (2013).
60. Bennett, C. M. & Miller, M. B. How reliable are the results from functional magnetic resonance imaging? *Annals of the New York Academy of Sciences* **1191**, 133–155, https://doi.org/10.1111/j.1749-6632.2010.05446.x (2010).
61. Thomas, B. P., Liu, P., Park, D. C., van Osch, M. J. & Lu, H. Cerebrovascular reactivity in the brain white matter: magnitude, temporal characteristics, and age effects. *Journal of cerebral blood flow and metabolism: official journal of the International Society of Cerebral Blood Flow and Metabolism* **34**, 242–247, https://doi.org/10.1038/jcbfm.2013.194 (2014).
62. Ghariq, E., Chappell, M. A., Schmid, S., Teeuwisse, W. M. & van Osch, M. J. P. Effects of background suppression on the sensitivity of dual-echo arterial spin labeling MRI for BOLD and CBF signal changes. *NeuroImage* **103**, 316–322, https://doi.org/10.1016/j.neuroimage.2014.09.051 (2014).
63. Shu, C. Y., Sanganahalli, B. G., Coman, D., Herman, P. & Hyder, F. New horizons in neurometabolic and neurovascular coupling from calibrated fMRI. *Progress in brain research* **225**, 99–122, https://doi.org/10.1016/bs.pbr.2016.02.003 (2016).
64. Shu, C. Y. *et al*. Brain region and activity-dependent properties of M for calibrated fMRI. *NeuroImage* **125**, 848–856, https://doi.org/10.1016/j.neuroimage.2015.10.083 (2016).
65. Hare, H. V., Blockley, N. P., Gardener, A. G., Clare, S. & Bulte, D. P. Investigating the field-dependence of the Davis model: Calibrated fMRI at 1.5, 3 and 7T. *NeuroImage* **112**, 189–196, https://doi.org/10.1016/j.neuroimage.2015.02.068 (2015).
66. Hare, H. V. & Bulte, D. P. Investigating the dependence of the calibration parameter M on echo time. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine* **75**, 556–561, https://doi.org/10.1002/mrm.25603 (2016).
67. Lu, H., Hutchison, J., Xu, F. & Rypma, B. The Relationship Between M in "Calibrated fMRI" and the Physiologic Modulators of fMRI. *The open neuroimaging journal* **5**, 112–119, https://doi.org/10.2174/1874440001105010112 (2011).
68. Kazan, S. M. *et al*. Physiological basis of vascular autocalibration (VasA): Comparison to hypercapnia calibration methods. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine*, https://doi.org/10.1002/mrm.26494 (2016).
69. Schmithorst, V. J. *et al*. Evidence that neurovascular coupling underlying the BOLD effect increases with age during childhood. *Human brain mapping* **36**, 1–15, https://doi.org/10.1002/hbm.22608 (2015).
70. Cantin, S. *et al*. Impaired cerebral vasoreactivity to CO2 in Alzheimer's disease using BOLD fMRI. *NeuroImage* **58**, 579–587, https://doi.org/10.1016/j.neuroimage.2011.06.070 (2011).
71. Perthen, J. E., Lansing, A. E., Liau, J., Liu, T. T. & Buxton, R. B. Caffeine-induced uncoupling of cerebral blood flow and oxygen metabolism: a calibrated BOLD fMRI study. *NeuroImage* **40**, 237–247, https://doi.org/10.1016/j.neuroimage.2007.10.049 (2008).
72. Agarwal, S., Sair, H. I., Yahyavi-Firouz-Abadi, N., Airan, R. & Pillai, J. J. Neurovascular uncoupling in resting state fMRI demonstrated in patients with primary brain gliomas. *Journal of magnetic resonance imaging: JMRI* **43**, 620–626, https://doi.org/10.1002/jmri.25012 (2016).
73. Fierstra, J. *et al*. Measuring cerebrovascular reactivity: what stimulus to use? *J Physiol* **591**, 5809–5821, https://doi.org/10.1113/jphysiol.2013.259150 (2013).
74. Fisher, J. A., Venkatraghavan, L. & Mikulis, D. J. Magnetic Resonance Imaging-Based Cerebrovascular Reactivity and Hemodynamic Reserve. *Stroke* **49**, 2011–2018, https://doi.org/10.1161/STROKEAHA.118.021012 (2018).

Acknowledgements

We thank R. Marc Lebel, PhD, for assistance with the ASL portion of the pulse sequence, and Ajit Shankaranarayanan, PhD, and Matthew J. Middione, PhD, from GE Healthcare for providing the source code of the GE multiband sequence. We also thank Lydia Washechek for editorial assistance. This work was partially supported by the Medical College of Wisconsin and a grant from the Daniel M. Soref Charitable Trust (to Yang Wang).

Author Contributions

A.D.C. contributed to the study design, processed all MRI data, conducted statistical analyses, wrote the main manuscript text and prepared all figures. Y.W. contributed to the study design, interpretation of results, supervised data analysis, and reviewed and made modifications on all drafts of the manuscript.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.