Title

A Systematic Study of the Sensitivity of Partial Volume Correction Methods for the Quantification of Perfusion from Pseudo-continuous Arterial Spin Labeling MRI

Authors

Moss Y Zhao[1], Melvin Mezue[2], Andrew R Segerdahl[2], Thomas W Okell[2], Irene Tracey[2,3], Yingyi Xiao[4], Michael A Chappell[1]

Affiliations

1 Institute of Biomedical Engineering, University of Oxford, Oxford, United Kingdom

2 Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), Nuffield Department of Clinical Neuroscience, University of Oxford, Oxford, United Kingdom

3 Nuffield Division of Anaesthetics, Nuffield Department of Clinical Neuroscience, University of Oxford, Oxford, United Kingdom

4 St Hilda's College, University of Oxford, Oxford, United Kingdom

Corresponding Author

Moss Y. Zhao

Institute of Biomedical Engineering

University of Oxford

Oxford, OX3 7DQ, United Kingdom

yize.zhao@eng.ox.ac.uk

Abstract

Arterial spin labeling (ASL) MRI is a non-invasive technique for the quantification of cerebral perfusion, and pseudo-continuous arterial spin labeling (PCASL) has been recommended as the standard implementation by a recent consensus of the community. Due to the low spatial resolution of ASL images, perfusion quantification is biased by partial volume effects. Consequently, several partial volume correction (PVEc) methods have been developed to reduce the bias in gray matter (GM) perfusion quantification. The efficacy of these methods relies on both the quality of the ASL data and the accuracy of partial volume estimates. Here we systematically investigate the sensitivity of different PVEc methods to variability in both the ASL data and partial volume estimates using simulated PCASL data and *in vivo* PCASL data from a reproducibility study. We examined the PVEc methods in two ways: the ability to preserve spatial details and the accuracy of GM perfusion estimation. Judging by the root-mean-square error (RMSE) between simulated and estimated GM CBF, the spatially regularized method was superior in preserving spatial details compared to the linear regression method (RMSE of 1.2 vs 5.1 in simulation of GM CBF with short scale spatial variations). The linear regression method was generally less sensitive than the spatially regularized method to noise in data and errors in the partial volume estimates (RMSE 6.3 vs 23.4 for SNR=5 simulated data), but this could be attributed to the greater smoothing introduced by the method. Analysis of a healthy cohort dataset indicates that PVEc, using either method, improves the repeatability of perfusion quantification (within-subject coefficient of variation reduced by 5% after PVEc).

Keywords

Arterial Spin Labeling; Cerebral blood flow; Partial volume effects; Systematic analysis; Perfusion MRI; Repeatability; Reproducibility;

Abbreviations

Arterial spin labeling (ASL);

Continuous ASL (CASL);

Cerebral blood flow (CBF);

Cerebrospinal fluid (CSF);

Gray matter (GM);

Linear regression (LR);

Modified least trimmed squares (mLTS);

Pulsed ASL (PASL);

Pseudo-continuous ASL (PCASL);

Proton density (PD);

Post-labeling delay (PLD);

Point spread function (PSF);

Partial volume (PV);

Partial volume effects (PVE);

Partial volume correction (PVEc);

root-mean-square error (RMSE);

region of interest (ROI);

Signal-to-noise ratio (SNR);

Variational Bayesian (VB)

White matter (WM);

Within-subject coefficient of variation (wsCV);

Highlights

- PVEc is effective for ASL even when there are errors in the PV estimates.
- The spatially regularized method was superior at preserving spatial details.
- The linear regression method was less sensitive to errors in PV estimates.
- PVEc improves repeatability of CBF estimation in ASL MRI.

Introduction

Arterial Spin Labeling (ASL) MRI is a non-invasive technique to measure perfusion (cerebral blood flow, CBF) using blood-water as an endogenous tracer (Detre et al., 1992). A pair of images are obtained in an ASL experiment: label and control. The difference between the label and control image is proportional to the perfusion of the tissue in the imaging region, being a direct measure of the delivery of labeled blood-water created by RF inversion applied to the neck. The ASL Consensus Paper has recommended single-PLD PCASL (pseudo-continuous ASL labeling with a single post-labeling delay) as the standard implementation for ASL experiments due to the simplicity of data collection and CBF quantification (Alsop et al., 2014). Whilst obtaining one ASL image (single-PLD) is sufficient for perfusion quantification in principle, acquiring multiple ASL images (multi-PLD) enriches the information to estimate CBF more accurately (Van Osch et al., 2007). Hence multi-PLD remains popular in neuroimaging studies where greater acquisition time and analysis expertise are available.

The ASL signal obtained from the brain broadly arises from three sources: gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). The ASL signal from CSF is ideally zero on the basis that no labeled blood water should reach the CSF (Golay et al., 2004; Johnson et al., 2005). In the context of perfusion MRI, it is often the GM that is of primary interest due to its major role in the central nervous system and that subtle changes of GM CBF have been observed in studies of various diseases including dementia (Le Heron et al., 2014). However, ASL data has a low spatial resolution with a voxel size of 3 to 5 millimeters that is unlikely to encode a single type of tissue, particularly in cortical regions where the thickness of GM is between 2 to 4 millimeters (Henery and Mayhew, 1989), leading to partial volume effects (PVE). Whilst all MR imaging in the brain suffers from PVE, it is particularly noticeable in perfusion imaging because of the distinctly different perfusion properties of GM and WM: perfusion ratios of 3:1 is quite commonly assumed.

It is increasingly being recognized that it may be important to correct PVE to allow GM CBF quantified independently of any confounding effects of partial voluming with either WM or CSF. This may be especially relevant for studies involving patients with tissue atrophy (Mutsaerts et al., 2014) where changes in tissue content in voxels could mask or be misinterpreted as changes in perfusion. Several partial volume correction (PVEc) methods have been developed to correct PVE in perfusion estimation from ASL data including linear regression (LR) (Asllani et al., 2008) and a spatially regularized technique that uses spatial priors in a Bayesian inference framework for perfusion quantification (Chappell et al., 2011). In the LR approach, each voxel intensity is considered as the weighted sum of different tissue contributions. By assuming the perfusion of the central voxel to be equal to its neighbors in a small region, defined by a regression kernel, the perfusion from each type of tissue is estimated by solving a linear system of equations. The spatially regularized method, motivated by the use of spatial information in the LR approach, exploited adaptive spatial priors within an existing Bayesian approach to ASL perfusion quantification, adding a second WM component to the model used. Studies on multi-PLD PASL have demonstrated the superiority of the spatially regularized method for maintaining spatial details in the estimated GM CBF image over the LR approach (Chappell et al., 2011). Both PVEc methods require knowledge of the exact fraction of tissue in each voxel, known as the PV estimates, which are typically computed by tissue segmentation from a high-resolution T_1-weighted image. However, flaws in the segmentation procedures may introduce error into the PV estimates that could affect the accuracy of PVEc. In addition, PVEc will also be affected by limitations in ASL data acquisition such as background noise and mismatch in the point spread function (PSF) between the ASL acquisition and that of the structural image from which the PV estimates are obtained. The impact of these factors on PVEc methods has not been fully explored, and a comprehensive study is needed to investigate how PVEc methods respond to these factors.

In the present study, we investigate the sensitivity of existing PVEc methods to errors in the PV estimates when applied to PCASL. Two studies were performed: firstly, a simulation study was used to investigate the sensitivity of the methods to a variety of theoretical sources of variability including random error and bias in the PV estimates, along with the effect of noise on the ASL data and differences in the PSF between ASL and structural images. Simulated single-PLD PCASL data were generated using the tissue kinetic model and the parameters in the ASL Consensus Paper (Alsop et al., 2014). Simulated multi-PLD PCASL data were created using the parameters (PLD and bolus duration) of a previously published reproducibility study of multi-PLD PCASL in neuroimaging studies (Mezue et al., 2014). Secondly, we performed PVEc on a healthy cohort dataset obtained from this study (Mezue et al., 2014). The repeatability of CBF images from each PVEc method was assessed by computing the within-subject coefficient of variation (wsCV) under the assumption that PVEc should improve the overall repeatability of perfusion quantification by reducing random variability in measured CBF due to PVE in each subject.

Theory

A general description of ASL signal that incorporates both GM and WM contributions can be written as:

$$\Delta M(t) = P_{GM} \cdot \Delta M_{GM}(t) + P_{WM} \cdot \Delta M_{WM}(t) \quad [1]$$

where ΔM is the longitudinal magnetization difference between ASL label and control images, P is the PV estimates of each type of tissue, and t is the time since the start of RF inversion. It has been assumed that the contribution from CSF is zero since no ASL difference magnetization should be observed in the CSF. This formulation hold for both single and multi-PLD ASL data. Independent estimation of either gray or white matter magnetization, and hence their perfusion, is ill-posed in any given voxel since only a single measurement of magnetization (ΔM) is available. PVEc methods seek to address this by using neighboring voxels, differing in the way this information is incorporated into the estimation process.

Linear Regression

The linear regression approach to correct PVE in ASL assumes that the CBF in both GM and WM can be treated as constant within a 2D regression kernel with size $n \times n$ and centred at a voxel in which both GM and WM values are required. Assuming identical GM and WM CBF values in all the voxels, the ASL difference signal in the voxels within the kernel can be written in matrix form, based on Eq. [1], as:

$$\boldsymbol{\Delta M} = \boldsymbol{P}\, \overline{\boldsymbol{\Delta M}} \quad [2]$$

Where $\boldsymbol{\Delta M}$ is a vector of length n^2 containing all the ASL difference values, $\boldsymbol{P}$ is n^2 x 2 matrix of GM and WM PV estimates and $\overline{\boldsymbol{\Delta M}}$ is a vector with two entries for the gray and white matter perfusion in the kernel. A numerical solution can be obtained by the following equation that minimizes the squared error and is equivalent to linear regression:

$$\Delta\bar{M} = (\boldsymbol{P}^T \cdot \boldsymbol{P})^{-1} \cdot \boldsymbol{P}^T \cdot \Delta\boldsymbol{M} \qquad [3]$$

where$(P^T \cdot P)^{-1} \cdot P^T$ is the pseudo-inversion matrix of $\boldsymbol{P}$. This calculation is repeated at every voxel within the brain to arrive at a map of PV corrected GM and WM ASL difference signal. The method can also equally be applied to the final quantified perfusion image (Steketee et al., 2015; Zhao et al., 2016) and in principle can be used with different shapes of kernels, as long as they can be specified in terms of whole voxels.

Spatially Regularized

An alternative spatially regularized approach as proposed by Chappell et al (Chappell et al., 2011) adopts the formulation in equation [1], but subjects both the GM and WM to spatial priors within a Variational Bayesian (VB) inference scheme. This method was motivated by the same principle as the LR method, but it relaxes the strict kernel assumption of LR by allowing the spatial regularization to be adaptively determined by the data. In essence, the algorithm attempts to estimate both GM and WM contributions in every voxel, with each being subject to a prior distribution on the value derived from the estimated values in the neighbouring voxels, the so-called spatial prior (Groves et al., 2009; Penny et al., 2005). This is an iterative scheme whereby the voxel values are updated, followed by an update to the prior and so on. Through Bayes theory the estimated values in any given voxel reflect both the measured data in the voxel and the prior information, with the balance being determined automatically, hence making the spatial regularization adaptive to the data. Although it was originally demonstrated on multi-PLD ASL data, the approach is also applicable to single-PLD ASL.

Methods

Data

One hundred high-resolution PV estimates images (GM, WM, and CSF) were randomly extracted from the first release of the UK Biobank Study datasets to be used in the simulated experiments (Miller et al., 2016). These PV estimates were segmented from the T_1-weighted images of the same study. Scan parameters were: 3D MPRAGE; sagittal; in-plane acceleration factor = 2; TI/TR = 880/2000ms, voxel size = 1 × 1 × 1mm; matrix = 208 × 256 ×256; acquisition time = 4min54s.

ASL MRI data from eight subjects were retrospectively analysed from the reproducibility study on optimization and reliability of multi-PLD PCASL by Mezue et al (Mezue et al., 2014). Scan parameters were: PLD = 0.25, 0.5, 0.75, 1, 1.25, 1.5s; bolus duration = 1.4s; TR/TE = 4000/13ms; FOV = 220 × 220; matrix = 64 × 64; 24 sequential ascending slices (no slice gap); slice thickness = 4.95 mm; slice acquisition time = 45.2 ms; 12 consecutive label and control images; total scanning time: 6.4min. Two additional proton density (PD) weighted images with TR = 6000 ms (one using the head and the other the body coil) were acquired to calibrate estimated CBF into absolute units. The head calibration was collected to compute the equilibrium magnetization of the arterial blood, and the body calibration image of each subject was used to correct the ASL data for the uneven sensitivity profile of the head coil (Wu et al., 2011). This dataset was used in the healthy cohort experiments.

Partial Volume Estimation

PV estimates at the resolution of the ASL data were computed from each subject by affine registration and tissue segmentation. First, brain tissue was extracted from high-resolution T_1-weighted structural images to remove the non-brain components using the FSL tool BET (Jenkinson et al., 2002b). Then they were segmented to produce PV estimates for GM, WM, and CSF using the

FSL tool FAST (Zhang et al., 2001). Affine registration of T_1-weighted structural images and the brain mask of ASL resolution (computed by taking the mean of ASL control images across all PLDs) was performed using the FSL tool FLIRT with 6 degrees of freedom (Jenkinson et al., 2002a). For use in PVEc of ASL data, the high-resolution PV estimates were transformed to the same (low) resolution as the ASL image space by subsampling and interpolation as described by Chappell et al using the FSL tool *applywarp* (Chappell et al., 2011). PV estimates of CSF were not used in the analysis. Finally, the PV estimates of GM and WM with voxel intensity less than 10% were excluded from the analysis, by replacing values less than 0.1 by 0 in the relevant PV estimates. The resulting PV estimates were denoted as the reference PV estimates.

Simulated Data

Using the one hundred PV estimates from the individuals extracted from the Biobank data, simulated ASL difference datasets of both single and multi-PLD PCASL were created by applying the general kinetic model proposed by Buxton et al (Buxton et al., 1998). For each simulated dataset, four different patterns of GM CBF were applied using the parameters listed in TABLE 1, with variations in the GM CBF map as the following:

Simulated dataset 1 – Flat GM CBF: homogeneous GM CBF of 60 mL/100g/min throughout the whole brain.

Simulated dataset 2 – Hyper/hypo GM CBF: homogeneous GM CBF of 60 mL/100g/min with a hypoperfused (30 mL/100g/min, spherical region of radius 10 voxels) region to the right of the brain and a hyperperfused (90 mL/100g/min, spherical region eight voxels radius) region to the left of the brain.

Simulated dataset 3 – Slow Sinusoidal GM CBF Variation: GM CBF of 60 mL/100g/min with superimposed sinusoidal variations in all three spatial dimensions, with a period equal to 30

voxels and an amplitude of 10 mL/100g/min, resulting in a smoothly varying GM CBF over the approximate range 50–70 mL/100g/min.

Simulated dataset 4 – Fast Sinusoidal GM CBF Variation: GM CBF of 60 mL/100g/min with superimposed sinusoidal variations in all three spatial dimensions, with a period equal to 10 voxels and an amplitude of 10 mL/100g/min resulting in a smoothly varying GM CBF over the approximate range 50–70 mL/100g/min.

In all cases, the WM CBF was set to 20 mL/100g/min throughout the whole brain. The ASL difference signal was derived by summing the two kinetic curves after multiplication of each component by the reference PV estimates using the model developed by Buxton et al as in Equation [1]. For each simulated dataset of both single and multi-PLD, a TR of four seconds and a total acquisition time of four minutes were assumed based on the ASL Consensus Paper for single-PLD PCASL (Alsop et al., 2014) and the reproducibility study data used in healthy cohort experiments for multi-PLD PCASL (Mezue et al., 2014). Under such conditions, the single and multi-PLD simulated data include ten and five repeats respectively. White noise was introduced to simulate the MRI background noise using white noise according to $\mathcal{N}(0, SD)$, where SD is defined as the ratio of reference signal intensity to signal-to-noise ratio (SNR), i.e. $SD = ref(\Delta M_t(t))\ /SNR$. The reference signal intensity $ref(\Delta M_t(t))$ was chosen to be the maximum signal intensity of a kinetic curve at GM CBF = 60mL/100g/min, Δt = 0.7s, and τ = 1.8s. Figure 1 illustrates the process of creating the simulated data.

Table 1: List of simulation parameters

Parameter		Value
Grey Matter		
CBF	(ml/100g/min)	60 (flat), 30 (hypo), 90 (hyper)
Δt	(ms)	700
T_1	(ms)	1300
λ		0.98
White Matter		
CBF	(ml/100g/min)	20
Δt	(ms)	1000
T_1	(ms)	1100
λ		0.82
Arterial Blood		
T_1	(ms)	1600
Sequence		
τ	(ms)	1800
PLD	(ms)	1800 (Single-PLD)
		300, 600, … 1800 (Multi-PLD)
α		0.91
Repeat		10 (Single-PLD)
		5 (Multi-PLD)

CBF: cerebral blood flow, Δt : arrival time, τ: bolus duration, PLD: postlabeling delay, α: labeling efficiency, λ: blood tissue partition coefficient.

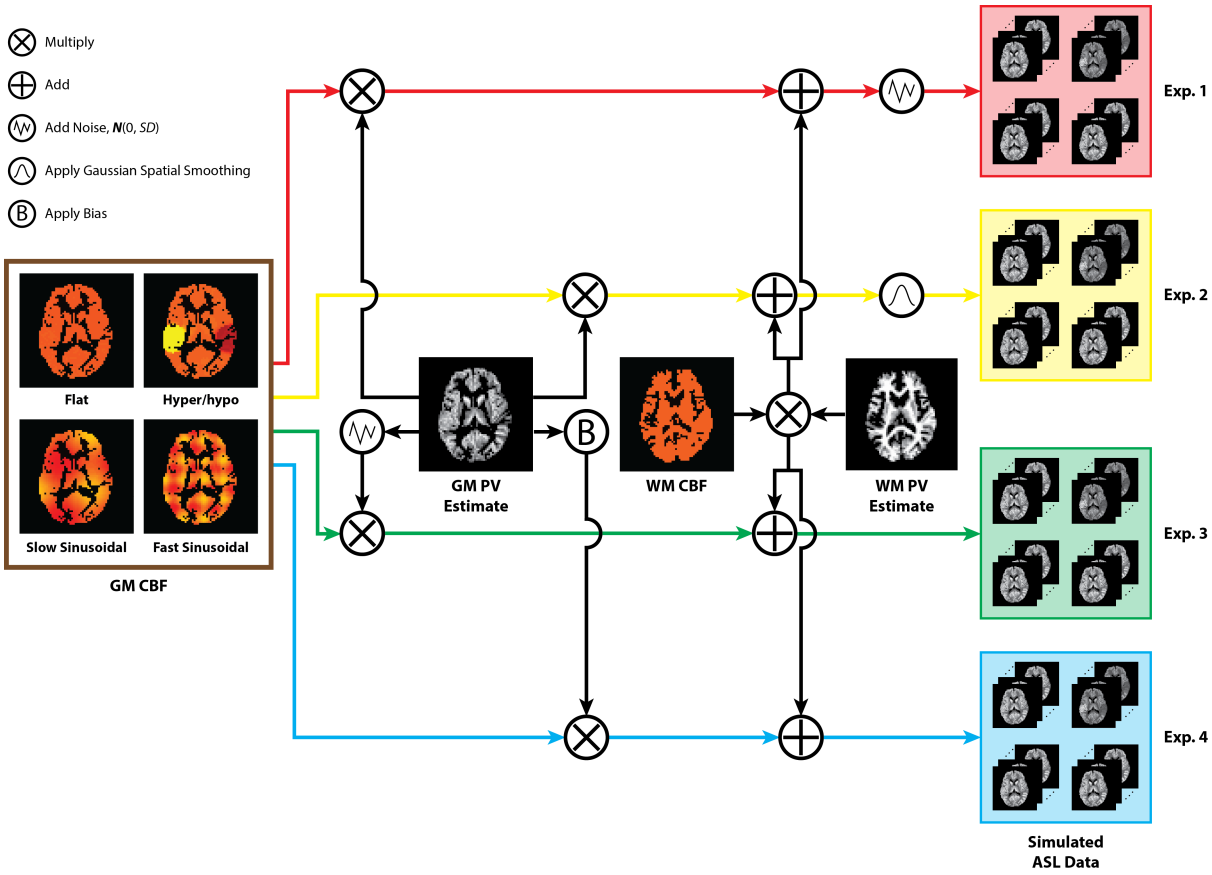


Figure 1: Schematic of the process of generating data for the simulation based experiments. Four datasets were created for both single and multi-PLD ASL data based on the CBF maps shown on the left. In each case the GM CBF map was used to generate simulated ASL difference data according to the kinetic model of (Buxton et al., 1998) and then multiplied by the GM PV estimate from an individual that had already been downsampled to the same space. To this was added a WM ASL difference signal generated using the same model, the WM PV estimates from the same individual and a presumed WM GBF of 20 ml/100g/min. Data were generated using PV estimates from 100 individuals. In experiment 1 (red), white noise was added to the simulated ASL difference data. In experiment 2 (yellow), Gaussian smoothing was applied to simulated ASL data (no noise was added). In experiment 3 (green), random errors were applied to PV estimates during the generation of the difference data. In experiment 4 (blue), bias was introduced to PV estimates during the generation of the difference data. For the data for both experiments 3 and 4, any adjustment made to the GM PV estimates was mirrored in the WM PV estimates, this process is not shown on the diagram.

Perfusion Estimation and Partial Volume Correction

In both single and multi-PLD ASL datasets, perfusion was estimated by fitting the kinetic curve to ASL difference data using spatial VB method in the FSL tool BASIL (Chappell et al., 2009). PVEc was performed by LR (Asllani et al., 2008) and spatially regularized approaches (Chappell et al., 2011). The LR method was implemented in two ways: (1) application of LR on ASL difference data before perfusion estimation; (2) application of LR on estimated CBF maps after perfusion estimation. For both implementations, the regression kernel size was set to 3×3 and 5×5. In the spatially regularized approach, 200 iterations were performed in each analysis to ensure convergence. Figure 2 shows the CBF estimation and PVEc process.

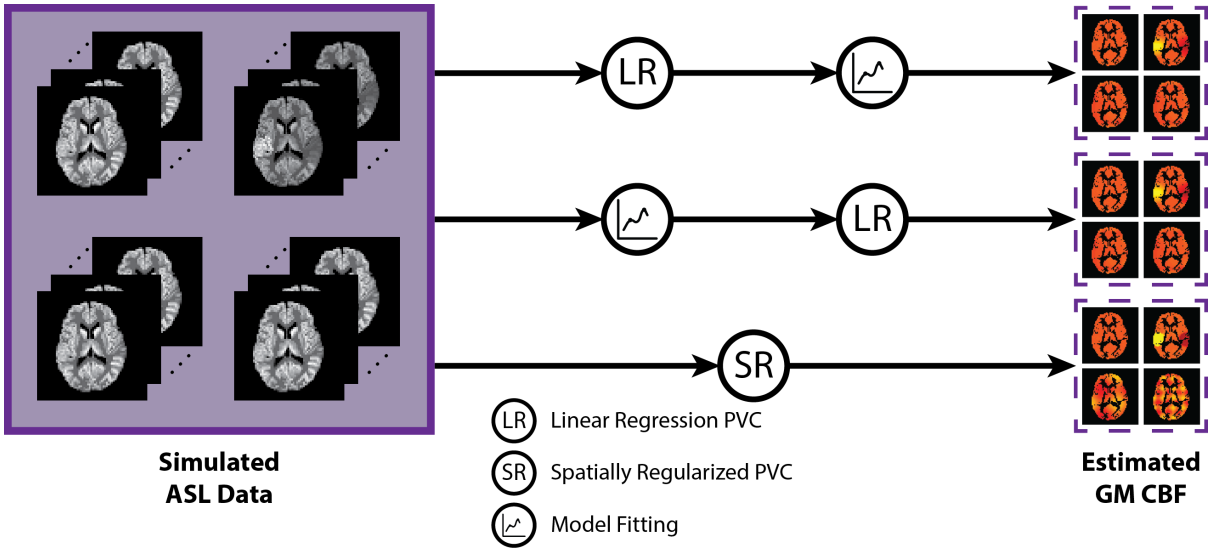


Figure 2: Perfusion estimation and partial volume correction. PVEc was performed using three methods – linear regression on ASL data, linear regression on CBF map, and spatially regularized method.

For the healthy cohort experiments, voxel-wise calibration was performed to compute GM CBF in the absolute unit (ml/100g/min) using the method proposed by Okell et al (Okell et al., 2013).

Specifically, a coil sensitivity map was derived by dividing the head calibration image by the body calibration image. The ASL difference image was corrected using the coil sensitivity map before model-fitting. A PD image was obtained by dividing the head calibration image by the coil sensitivity map. This PD image was then corrected by multiplying by the factor $\left(\frac{1}{1-e^{-TR/T_{1,tissue}}}\right)$, where $T_{1,tissue} = 1.3s$, to compute the M_{0t} image (Alsop et al., 2014). The M_{0a} image was computed by multiplying the M_{0t} image by the blood-water partition coefficient of 0.9 (Mezue et al., 2014). Then, a 2-D median filter of kernel size 3 × 3 was applied to the M_{0a} image to reduce noise. Due to the PVE on the edge of the M_{0a} image, the signal intensities of the voxels on the edge of the brain were substantially less than those inside the cortex, causing overestimation to the absolute CBF values on the edge of the brain. Such issue was resolved by erosion and extrapolation techniques. First the M_{0a} image was eroded by three voxels. The eroded voxels were extrapolated by taking the mean in a 5 × 5 neighborbood to create the corrected M_{0a} image. Finally, the estimated CBF images were calibrated by the corrected M_{0a} and an assumed inversion efficiency of 0.88 (Okell et al., 2013).

Simulation Study

Four numerical experiments were conducted to compare the effects on PVEc of variation in the degree of noise in the ASL data and PV estimates. They were performed using both single-PLD and multi-PLD simulated ASL data. For each experiment, the experimental parameters are listed in Table 2, and the results were assessed by root-mean-square error (RMSE) and region of interest (ROI) analysis.

Experiment 1: impact of background noise of ASL data on PVEc

This experiment investigated the influence of background noise in ASL data on the results of each PVEc method. The same reference PV estimates were used in each PVEc method as were used to

create the simulated data for each subject. This experiment provided a baseline for the subsequent experiments under the ideal case where the true PV estimates were known and thus can be used for PVEc.

Experiment 2: impact of resolution mismatch on PVEc.

As discussed, one of the limitations of PVEc is an unmatched PSF between structural and ASL images even when differences in matrix size (and thus voxel dimensions) are accounted for. This was modeled by blurring the simulated noise-free ASL data with a 3D isotropic Gaussian blurring kernel defined by a range of standard deviation values θ ($\theta = \sqrt{8ln2}/FWHM$ and $FWHM$ values are listed in Table 2). PVEc and CBF estimation were performed on the blurred ASL data using the reference PV estimates.

Experiment 3: impact of random errors in PV estimates on PVEc

This experiment assessed the sensitivity of each PVEc method to variabilities in the PV estimates. Random errors were added to each of the one hundred reference PV estimates assuming that the errors could be modeled as $\mathcal{N}(0, \sigma)$, and a range of σ were considered as indicated in Table 2. A total of one hundred distorted PV estimates were created using the Biobank datasets.

Experiment 4: impact of bias in PV estimates on PVEc

This experiment aimed to compare PVEc methods when there is a systematic bias in PV estimates. Biased PV estimates were generated in two steps: (1) altering the reference GM PV estimates using a non-linear transformation to create biased GM estimates; (2) computing the biased WM PV

estimates using the reference GM and WM PV estimates and the newly derived biased GM PV estimates in the following:

$$PV_{bias,GM} = PV_{ref,GM}^{\exp(b)}$$
$$PV_{bias,WM} = PV_{ref,GM} + PV_{ref,WM} - PV_{bias,GM} \quad [4]$$

where PV_{ref} and PV_{bias} are the reference and biased PV estimates respectively, b is the bias factor, which is applied to reference GM PV estimates using the range of values given in Table 2. Figure 3 illustrates the effect of non-linear transformation with different b values. This function exchanges PV in GM for WM, assuming that there is some inherent bias in the PV estimation process between these two tissue types. The biased PV estimates were used in PVEc and CBF estimation using the three methods for noise-free simulated ASL data.

TABLE 2: Experimental parameters

Parameter	Value
Experiment 1	
SNR	5, 10, 15, 20, ∞
Experiment 2	
FWHM	0, 1.5, 3, … , 7.5, 9 mm
SNR	∞
Experiment 3	
$\boldsymbol{\sigma}$	0, 0.025, 0.050, … , 0.175, 0.200
SNR	∞
Experiment 4	
$\boldsymbol{b}$	-0.5, -0.4, … , 0, 0.1, …, 0.5
SNR	∞

SNR: signal-to-noise ratio in simulated ASL data; FWHM: full width at half maximum for PSF; σ**:** standard deviation in random errors of PV estimates; b**:** bias factor in biased error of PV estimates

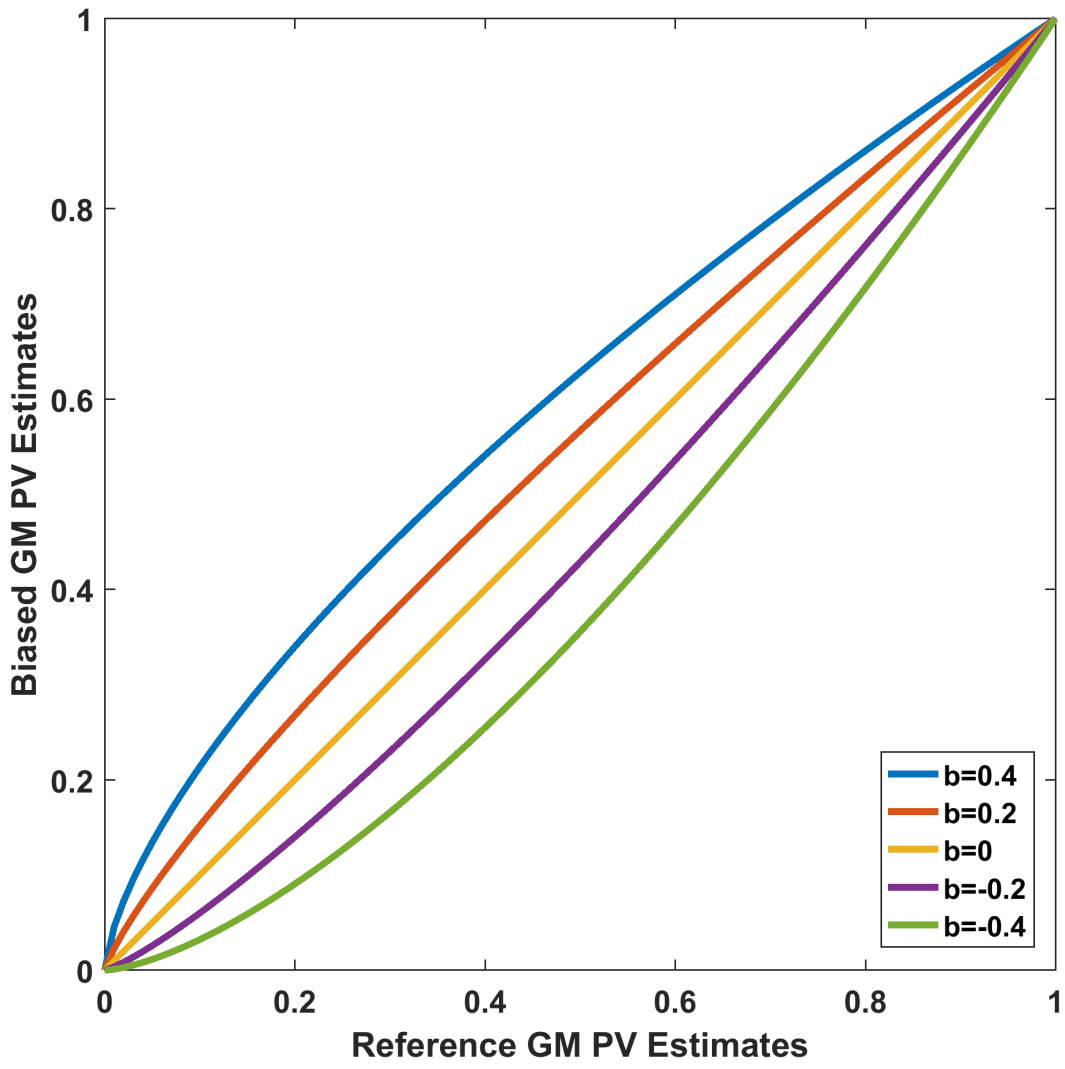


Figure 3: Biased PV Estimates. Biased GM PV Estimates were created by applying a non-linear transformation to the reference GM PV estimates.

Evaluation Metrics for Partial Volume Correction

RMSE Analysis

In evaluating the accuracy of PV corrected perfusion for the simulated data, RMSE was computed for each PVEc method in the following equation:

$$RMSE_{CBF} = \sqrt{\frac{\sum_{j=1}^{N}(\hat{f} - f)^2}{N}} \qquad [5]$$

where $\hat{f}$ and f are estimated and simulated CBF respectively, N is the total number of voxels within the brain mask in each model-fitting experiment.

ROI Analysis

The ROI based method introduced by Asllani et al (Asllani et al., 2008) was also used to examine the effects of different PVEc methods. Nine ROIs were defined based on the voxel intensity of estimated PV maps. Each ROI covered a 10% range of the PV estimates, and voxel intensity less than 10% were omitted. The resulting ranges were: 10% - 20%, 20% - 30%, … 90% - 100%. In each ROI, the mean GM CBF with and without PVEc was computed. The ROI methodology was originally proposed on the basis that pure GM CBF should be independent of the fraction of GM PV estimates because perfusion is an implicit measurement of the density of tissue CBF at each voxel. For the simulations, this will hold true for the homogenous GM CBF case. For the two cases of sinusoidal variation, it is still expected to be approximately true over the ensemble of brain voxels, albeit with a larger variability within the ROI. This is based on the assumption that the ROIs defined here are a pseudo-random sampling of the brain and the sinusoidal variations used are applied without any imposed correlation to the brain structure. The ROI analysis is no longer valid for the hyper/hypo perfusion case and thus was not computed.

The ROI method was extended to allow quantitative comparisons of the PVEc performance by considering a linear regression of GM CBF and GM PV estimates for ROI in the range 10% – 90% using a least-square fit. The slope of this regression was recorded, a zero slope would be expected to represent good PVEc as it would imply independence of estimated CBF with PV estimates.

Repeatability Study

We assessed the repeatability of the PVEc methods by analyzing the inter-session coefficient of variation of GM CBF calculated from the healthy cohort data collected by Mezue et al (Mezue et al., 2014). The hypothesis for the analysis on real data was that the inter-session variability should reduce after PVEc, on the assumption that PVE will introduce between session variability due to the different position of the brain in each scanning session and that PVEc removes this bias from the amount of GM tissue in each voxel. There will in practice be other sources of physiological variability apart from those related to PVE, including those identified by Clement et al (Clement et al., 2017). Thus, although we hypothesize that a reduction might be achieved by an effective PVEc strategy, it coult at best only partially improve the repeatability. Ideally, the GM CBF after PVEc would be identical between each session if there are no flaws in registration (and in the absence of noise) and if PVEc removes PVE completely.

A transformation matrix (Matrix 1) from ASL image to high-resolution space was computed by registering the CBF map without PVEc to the structural image of each subject. An initial transformation matrix was obtained using the FSL tool FLIRT between the calibration and structural images. This transformation was then refined with a boundary-based registration (BBR) cost function that uses the boundary of WM, which was segmented from the structural image. This implementation is available in the *asl_reg* function part of the BASIL toolbox in FSL. For each subject, a second transformation matrix (Matrix 2) as obtained by registering the high-resolution T_1-weighted image to the standard brain (MNI152_T1_2mm) using the FSL tool FLIRT (Jenkinson and Smith, 2001). Matrix 2 was used to initialize the estimation of the spline coefficients (warp) from structural to standard using the FSL tool FNIRT (Jenkinson et al., 2012). Finally, the GM CBF map was transformed to the standard space using the warp and Matrix 1 with the FSL tool *applywarp*.

We evaluated the repeatability of each PVEc method using within-subject coefficients of variation (wsCV) within three ranges of GM PV estimates ROIs: 10% - 40%, 40% - 70%, and 70% - 100%. Within-subject coefficients of variation were computed as the ratio between the standard deviation of voxel-wise absolute differences of repeated GM CBF estimates and the mean repeated GM CBF estimates in the following equation:

$$wsCV = \frac{SD(f_1 - f_2)}{mean(f_1, f_2)} \times 100\% \qquad [6]$$

where f is the GM CBF. The metric was computed for the three repeated experiments (session, week, and month repeat).

Results

Simulation study: single-PLD data (kernel size for 3×3 *the LR methods)*

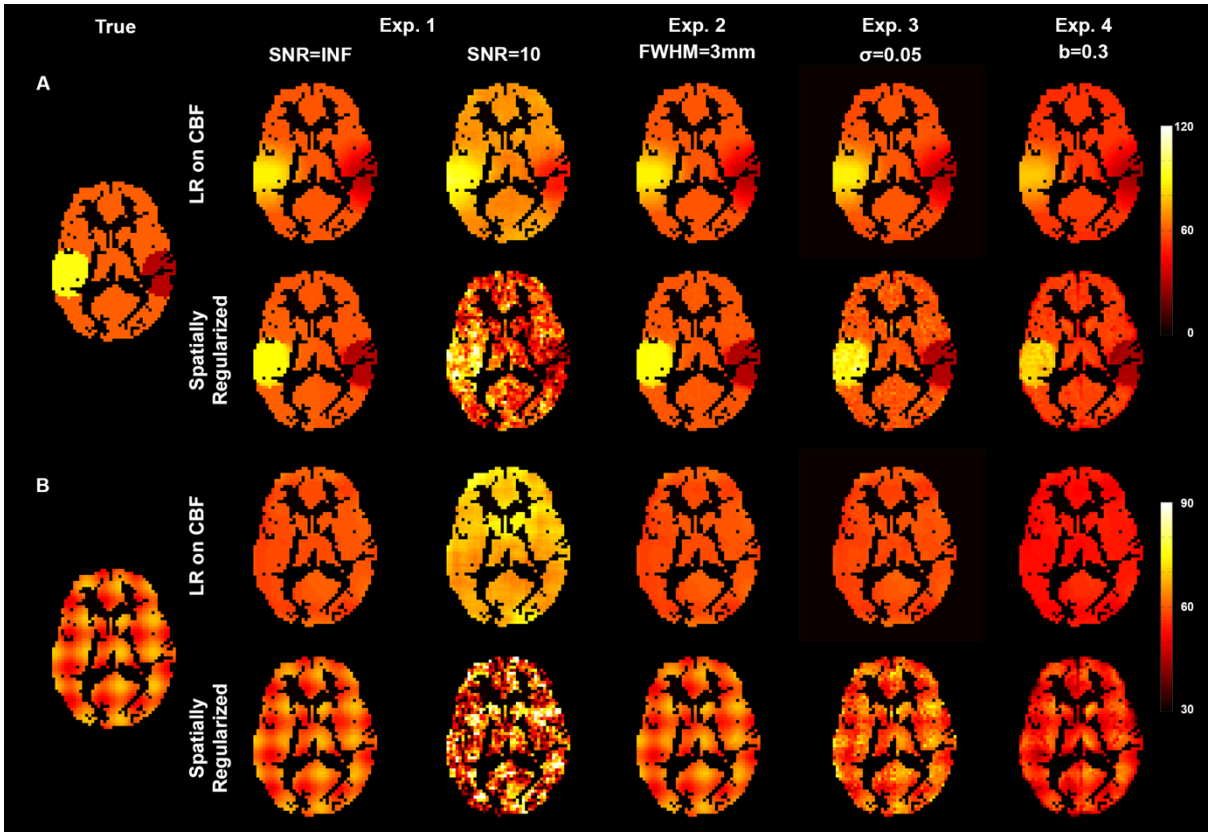


Figure 4: Estimated GM CBF map for single-PLD simulated data from a single subject. (A) Hyper/hypo GM CBF; (B) Fast sinusoidal GM CBF variation.

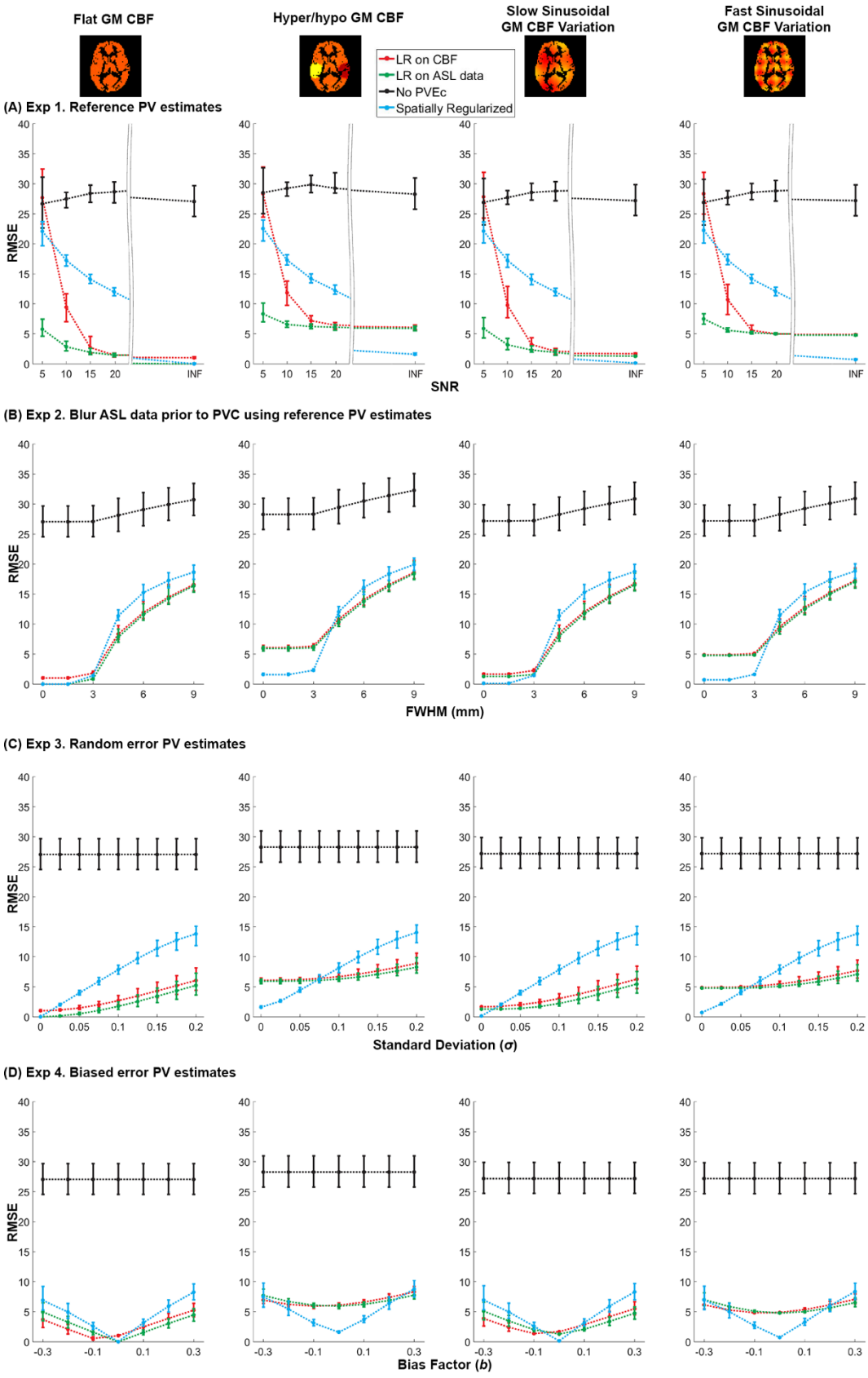


Figure 5: RMSE between estimated and simulated GM CBF using noise-free simulated single-PLD data.

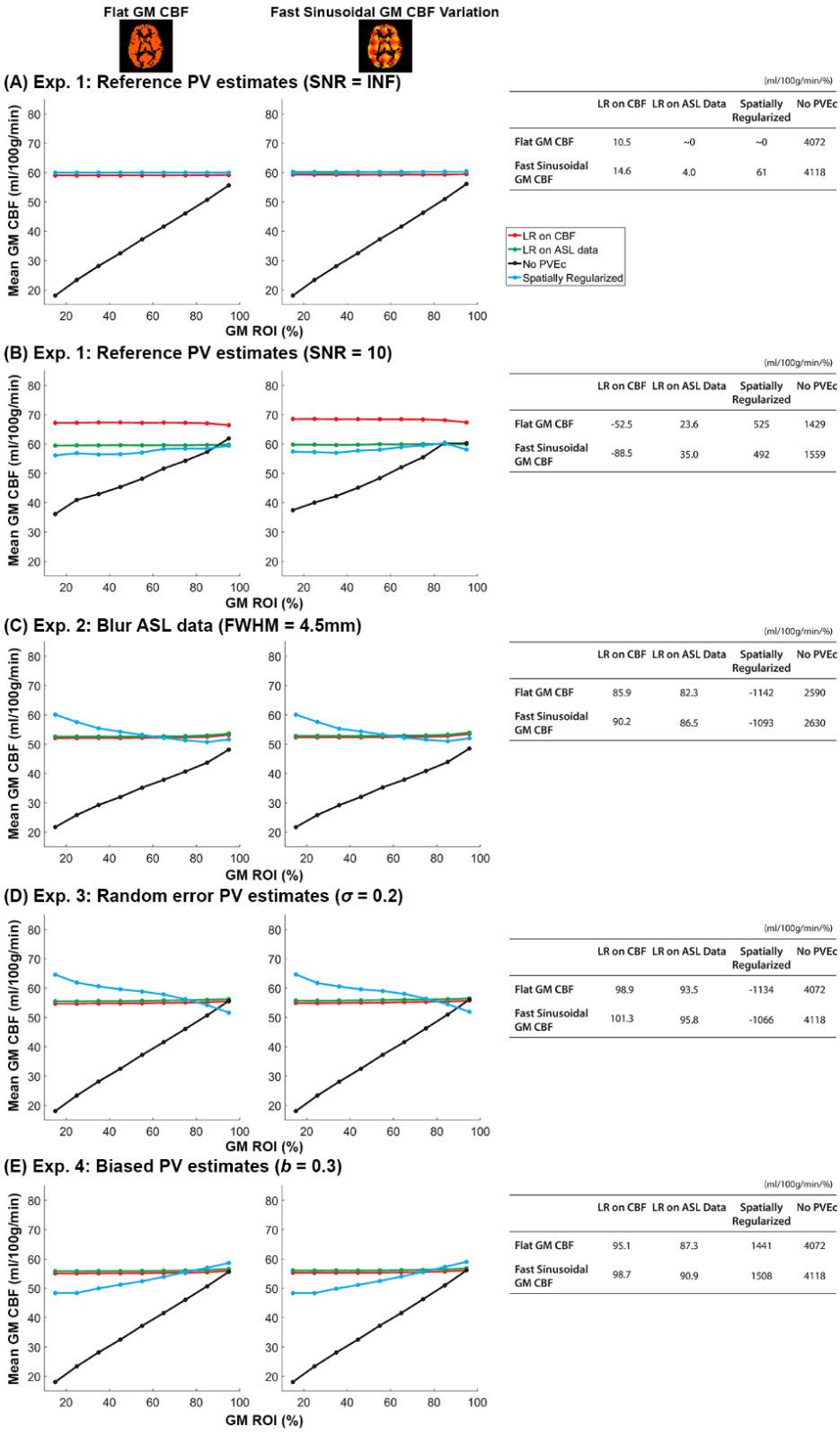


Figure 6: Example ROI curves for all experiments using single-PLD data. The first row of curves shows the ROI analysis in a perfect situation (SNR=INF and reference PV estimates). Subsequent plots

reveal the changes of mean GM CBF observed in each experiment for a specific choice of experimental parameter. Each adjacent table shows the estimated slope values for the linear regression fitting between estimated GM CBF and GM ROI.

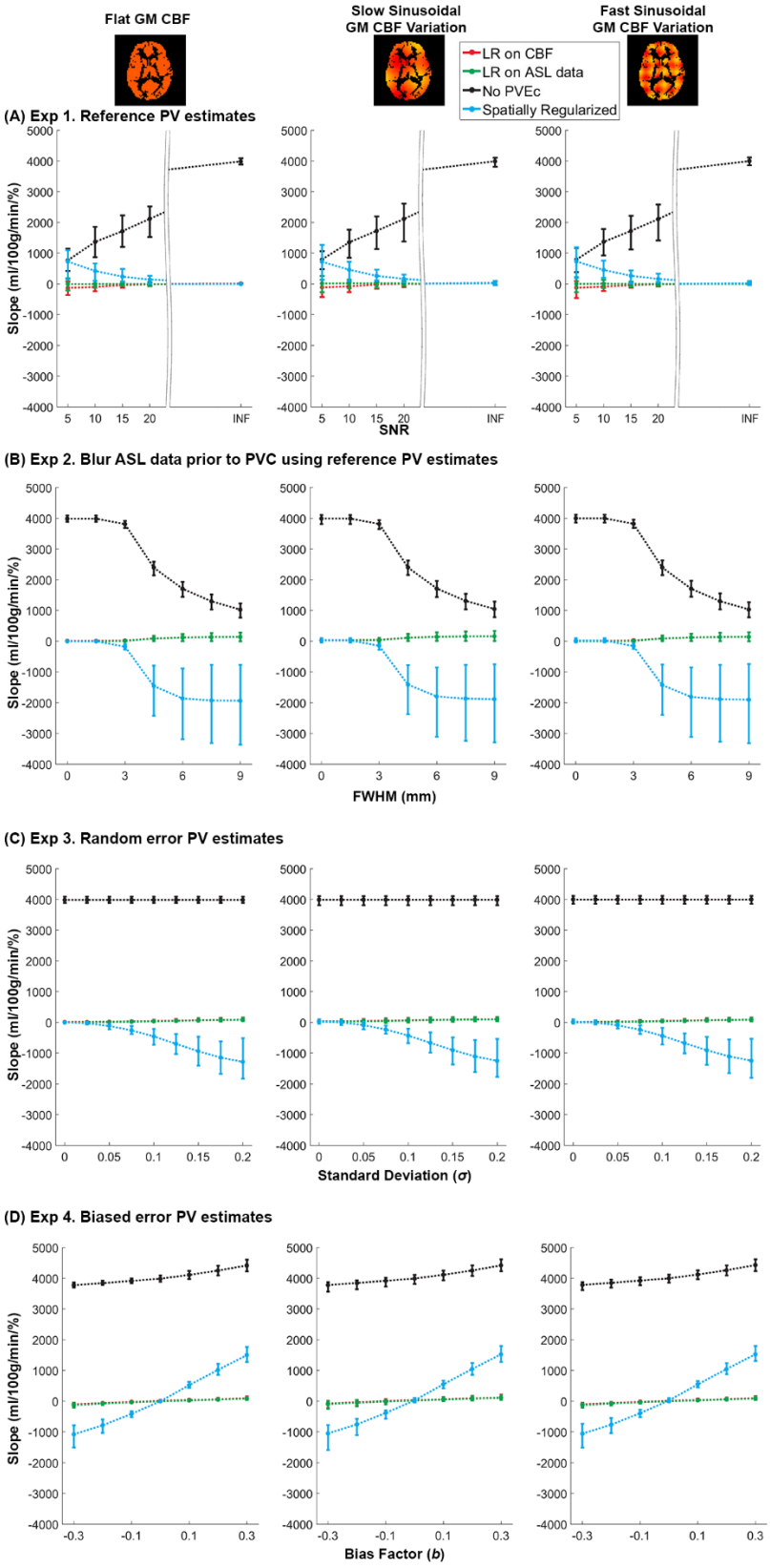


Figure 7: Estimated slope values of the extended ROI analysis using single-PLD data.

Figure 4 shows the estimated GM CBF maps of an example subject from the simulation of single-PLD PCASL data after applying different PVEc for selected cases within each of the four experiments. Figure 5 shows the RMSE between estimated and simulated ground truth GM CBF over all subjects for each experiment. Figure 6 shows example ROI curves and the corresponding slope values for each PVEc method for selected cases within each experiment. Figure 7 shows the calculated slope values across all the subjects for each experiment.

In experiment 1, the reference PV estimates were used within the PVEc to investigate the sensitivity of different PVEc methods to background noise on ASL data. For no noise (SNR=INF), all of the methods were able to correct for PVE, but the spatially regularized method appeared to be more accurate in capturing short scale variations in GM CBF, as can be seen most evidently in the results for hypo/hyper regions in Figure 4 and also from the lowest errors when SNR=INF in Figure 5A for the two cases with short scale CBF variations: hyper/hypo CBF and fast sinusoidal variation. As the SNR was degraded, all of the methods became less accurate at GM CBF estimation, with LR applied to the ASL data appearing to be least sensitive to noise, followed by LR applied to the CBF map and finally the spatially regularized method. Noticeably in Figure 4 as the noise increased the results from the spatially regularized method appeared more 'noisy', whereas the maps from LR remained smooth and errors between true and estimated CBF manifested as global CBF differences, as can be seen for example in Figure 5B where the GM CBF in all ROI was greater than the true (mean) value.

The high accuracy of all the PVEc methods for no noise translated to small slope values in the ROI analysis, Figures 6A and 7A: no correction resulted in a slope of the order of 4000 ml/100g/min/%, at least two orders of magnitude larger than the PVEc results. As the noise increased the slope value for no correction became less steep, for example, values around 1500 ml/100g/min/% at SNR=10 in

Figure 6B. For all of the PVEc methods as noise increased, the absolute value of the slope increased mirroring the reduction of accuracy seen in the RMSE results.

In experiment 2, the impact of mismatch between the PSF of the ASL data and the structural image was examined. Effects on both RMSE and slope were only obvious once the FWHM of the Gaussian kernel applied to the ASL data exceeded the voxel size of the ASL image (3.5mm). As the FWHM increased beyond this value the RMSE increased to approximately the same degree for all PVEc methods, and it was also associated with a small increase in RMSE for the non-corrected case. For the spatially regularized method, the errors introduced by PSF mismatch were manifested in overestimation at lower PV of GM and underestimation at higher values, as seen in Figure 6C (in an example with FWHM of 4.5 mm) and as a negative slope at larger FWHM in Figure 7B. However, Both LR methods only produced relatively small increases in slope as the PSF FWHM increase, the observed increase in RMSE for these examples being associated with global errors in CBF.

In experiment 3, the impact of random errors in the PV estimates was analysed for each PVEc method. Similarly to experiment 1, random errors on PV estimates led to an increase in RMSE for all the PVEc methods, as shown in Figure 5C, but the increase in error was more marked with the spatial method compared to either LR approach. The slope metric again showed overestimation at low PV GM by the spatial method, as shown in Figure 7C, whereas both LR methods estimated the same GM CBF in all ROIs on average even as the error on GM PV estimate increased, thus their slope remained comparatively close to zero. The increase in RMSE seen for the LR methods was due to consistently underestimating GM CMF in all ROIs, for example in Figure 6D both LR methods are underestimating the true (mean) simulated GM CBF (60 ml/100g/min).

In experiment 4, the impact of bias in the PV estimate on PVEc methods was examined. Bias affected all methods and led to an increase in RMSE with an increase in the degree of bias (bias factor: *b*), as shown in Figure 5D, and the spatial method appeared to be most sensitive to increases in bias. The ROI analysis showed that for increased positive bias the spatial method tended to underestimate in low and overestimate in high GM PV regions, resulting in positive slope values; vice versa for negative bias, as shown in Figure 7D. Both LR methods tended to globally underestimate GM CBF at higher bias, but retain slope values relatively close to zero, for example in Figure 6E.

Results of simulated experiments using multi-PLD data can be found in supplementary materials. Looking at the GM CBF map of multi-PLD ASL data in Figure 11, the impact of background noise was more substantial than on the single-PLD data for the spatially regularized method. In terms of perfusion quantification accuracy in noisy data, LR when applied to the data or the estimated perfusion map when using multi-PLD data were nearly identical in terms of RMSE (Figure 12A) as opposed to single-PLD data in Figure 5A. Apart from these, the results of simulated experiments using multi-PLD data were similar with the results of single-PLD data.

The results from using LR methods of kernel size of 3×3 are shown in supplementary materials (Figure 15 to 21). Although the overall trend of the results was similar with using the kernel size of 5×5, the LR methods became more accurate with the smaller kernel size due to less smoothing effects. However, it was still less accurate than the spatially regularized method when reference PV estimates were used in the noise-free simulated data, as shown in Figure 16A and Figure 19A.

Healthy Cohort Data

Healthy Cohort Experiments

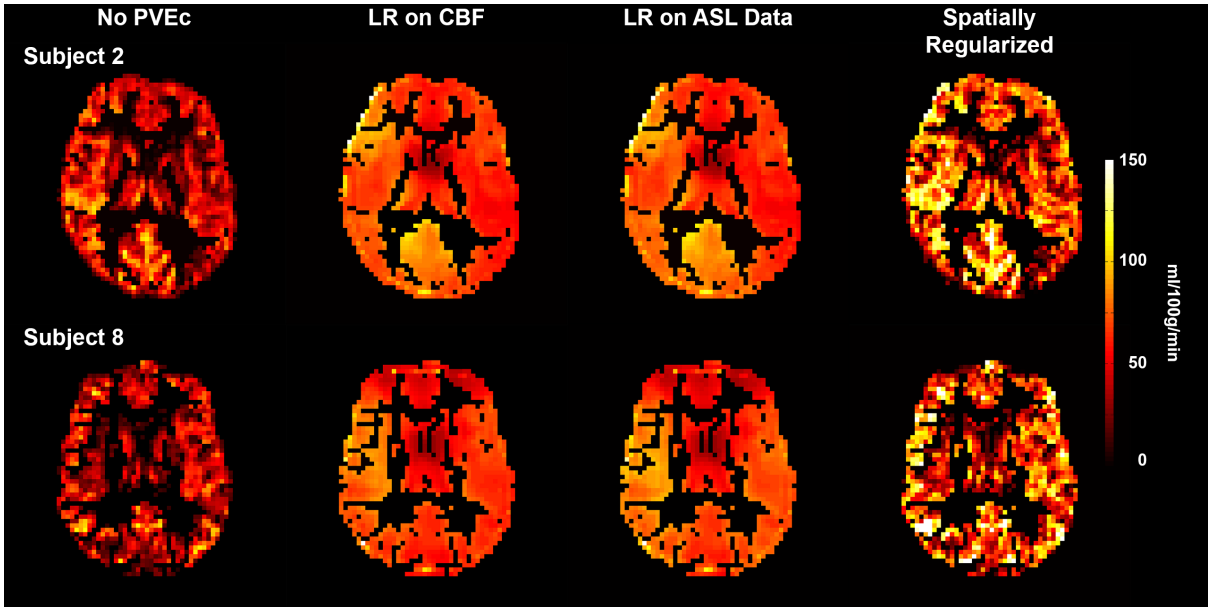


Figure 8: Estimated GM CBF maps of two example subjects using different PVEc methods.

Figure 8 shows the estimated GM CBF before and after PVEc from two subjects within the healthy cohort data. Similar to the simulated experiments, the results of the LR methods are nearly identical, with more spatial variability visible in the maps from the spatially regularized method.

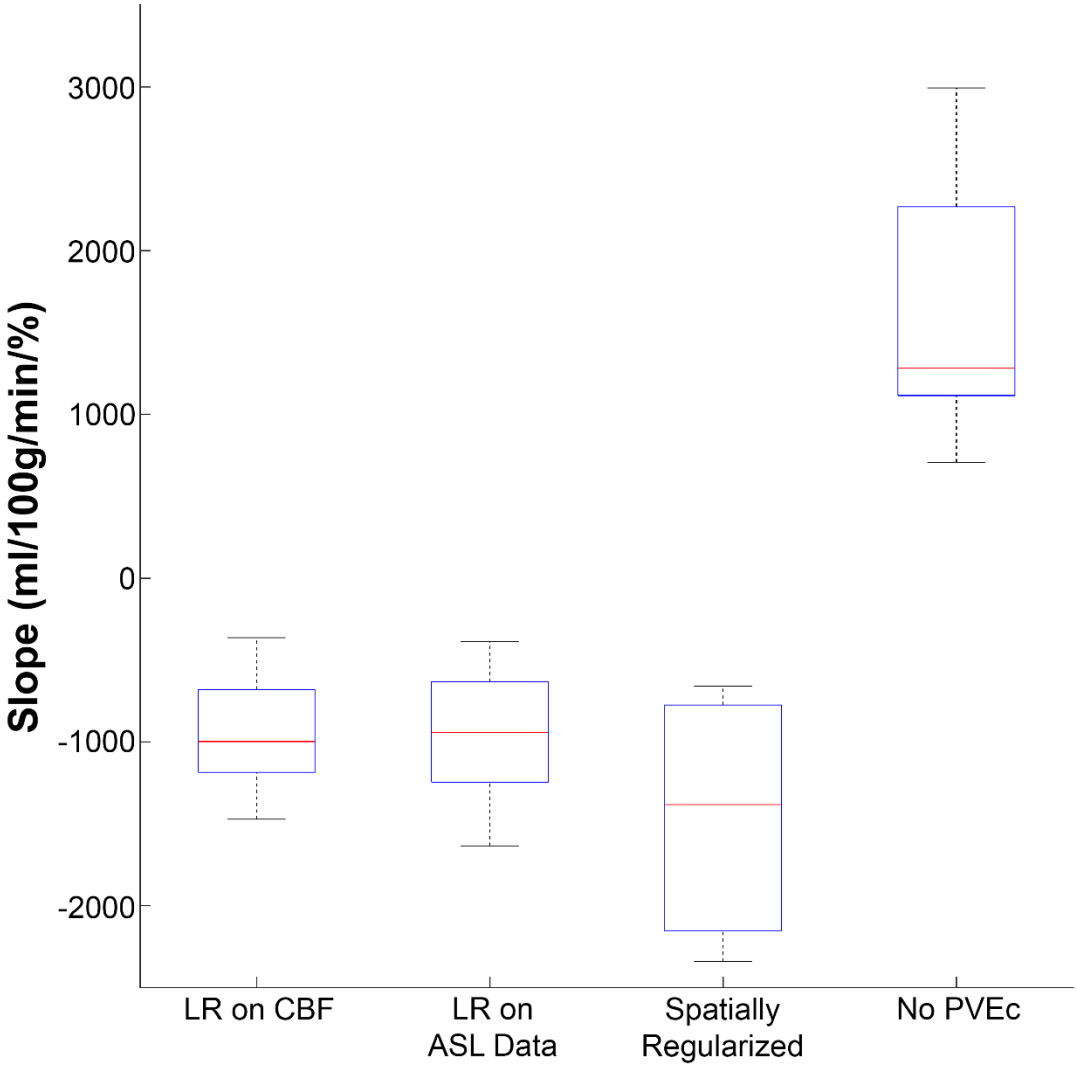


Figure 9: Estimated slope values of the extended ROI analysis on healthy cohort data. The slope values of all eight subjects are plotted for each PVEc method.

Figure 9 shows the estimated slope values of the extended ROI analysis on healthy cohort data. The analysis was performed on data from the first scan session of each subject. Without PVEc, the slope values are positive, in agreement with the simulations. After PVEc, the slope value becomes negative with a median value around -1000 ml/100g/min/% for the LR methods, but more negative for the

spatially regularized method. Similarly to the simulations, the two LR methods achieved nearly identical slope results.

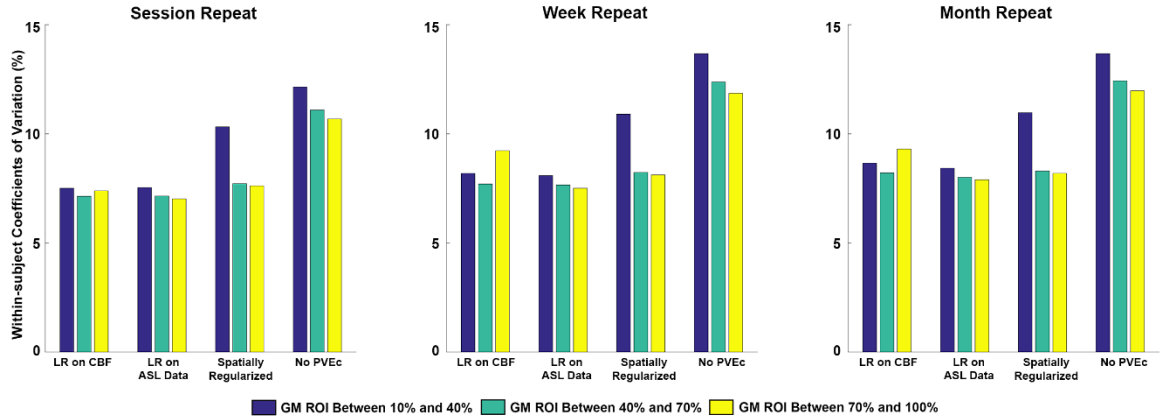


Figure 10 Within-subject coefficient of variation of the PVEc methods in three ROIs. All PVEc methods improve repeatability of perfusion quantification, and the repeatability is higher in greater GM ROIs.

Figure 10 shows the wsCV of each PVEc method (including no PVEc) for the three ROIs. Overall, the wsCV reduced after performing PVEc. Comparing between each PVEc method, the two LR methods show similar performance in ROIs containing voxels with less than 70% GM, whereas the LR on ASL data approach appears to be better in reducing variability for ROIs containing voxels with greater than 70% PV GM. The spatially regularized method has similar repeatability to the LR methods for ROIs containing voxels with greater than 40% GM, but not in those ROIs with voxels below this threshold, although it was still more repeatable than no correction.

Discussion

In this work, we have investigated the sensitivity of two different but similarly motivated PVEc methods for perfusion estimation from ASL on PCASL data. We have simulated different types of variabilities in ASL data acquisition as well as PV estimates including background noise in the ASL data itself, differences in PSF between ASL and structural data, and errors in the PV estimates. A range of artificial spatial distributions was considered in GM CBF, consistent with those that have been used previously by Chappell et al (Chappell et al., 2011), to assess how each PVEc method is able to restore spatial details. Additionally, the impact of PVEc on repeatability was examined using a cohort of healthy individuals previously used to examine the repeatability of PCASL. The findings of the study are threefold: (1) all PVEc methods reduce the bias of PVE in GM CBF quantification substantially regardless of the variabilities in data acquisition and PV estimates; (2) the spatially regularized method has shown superiority to the two LR methods in preserving spatial variations using data with high SNRs, but it is more sensitive to noise and errors in the PV estimates; (3) PVEc improves the repeatability of perfusion quantification using multi-PLD PCASL by reducing the bias from PVE.

Simulation study

It has already been shown that the spatially regularized method appears to have greater ability to preserve spatial details in the GM CBF image over the LR methods. This was confirmed in this study and was most obvious where GM CBF varies over shorter length scales and in less noisy datasets. This finding was demonstrated qualitatively in the form of estimated GM CBF and quantitatively in the RMSE analysis and the ROI analyses. Specifically, at low noise levels, the RMSE for the spatially regularized method was always lower than LR method for simulations in which the GM CBF map contained short scale variations. Due to the assumption of the LR method that tissue CBF is homogeneous within the kernel of a fixed size, a substantial degree of ‘smoothing’ arises after PVEc,

which appears to be similar to blurring the data with a Gaussian Filter (Asllani et al., 2008). Our current analysis not only confirms the previous findings but also give insights on the quality of the preserved spatial variability at different noise levels. Noticeably the LR methods appeared to be relatively insensitive to noise in data and errors and bias in the PV estimates; however, this appears to be related to the smoothing effect of the kernel based approach and results in a trade-off in the detail of the final GM CBF image. For example, the LR methods tended to return relatively small slope values (in the range 0 – 100 ml/100g/min/%) from the extended ROI analysis, consistent with good correction for PVE. However, this was often associated with a consistent error in GM CBF at all values of GM PVE, see for example Figure 6 where the GM CBF values returned by the LR methods were often consistently different from the true (mean) value of 60 ml/100g/min used in the simulation, and this error was largely invariant with the degree of noise or error in PV estimates used in PVEc. This effect is consistent with what can be seen in Figure 4, where the maps from LR are always noticeably smooth appearing implying that errors are being smoothed away, but at the same time so are spatial features.

In studying the impact due to PSF mismatch, we have observed an increase in error with FWHM for both PVEc and non-PVEc methods. The estimation accuracy reduces primarily only when the FWHM imposed became larger than the voxel size as shown in Figure 5B. The effect on the non-PVEc results is consistent with the corresponding increase in the influence of PVE that will occur when the true resolution of the data degrades. The more noticeable increase in RMSE for GM CBF estimation under PVEc will then be a combination of greater PVE in the data and the imperfect correction due to the use of a GM PV estimate map that does not match the true effect. Even for relatively modest FWHM beyond the voxel size gave rise to errors of up to 20 ml/100g/min in GM CBF. When the PSF of ASL data is less than the voxel size, the results here imply that this should not be a major concern for PVEc of ASL. This is true for certain readout schemes such as multi-TI 2D echo-planar imaging.

However, there are notable examples of ASL acquisitions where this could be problematic. For instance, acquisitions using 3D GRASE readouts which tend to have an extended PSF in the slice encoding direction. The original work of Chappell et al. (Chappell et al., 2011) did a correction for this effect prior to PVEc using a 'deblurring' method. The present study implies that this should be considered where PSF mismatches might arise. It is also noteworthy that the methodology adopted here to model mismatches in PSF is the same as the spatial smoothing that is sometimes applied to ASL data to improve visualisation. The results imply that care should be taken when spatially smoothing ASL data prior to quantification where PVEc is being used. In practice, the added spatial homogeneity introduced by the assumptions of PVEc methods means that smoothing of the data prior to quantification is unlikely to be necessary.

An extended ROI analysis has been developed to compute the slope value of the classic ROI curve proposed by Asllani et al (Asllani et al., 2008). The purpose was to quantify the observed approximately linear dependency between GM CBF and GM ROI. A positive slope value indicates overestimation in low ROI or underestimation in high ROI, whereas a negative slope value indicates the opposite. For the non-PVEc GM CBF results, the estimated slope value ranged between 2500 to 4000 ml/100g/min/% dependent upon the noise in the data. This matches the already observed positive correlation between estimated GM CBF and PV estimates. These slope values correspond to GM CBF being underestimated between 25 to 40 ml/100g/min in ROIs with the lowest proportion of GM. Notably, the slope value reduced for data without PVEc when the noise increased, implying that care needs to be taken with either the standard ROI analysis or the slope metric calculated here when this is applied for the evaluation of PVEc efficacy on real data.

For the spatially regularized method under the impact of errors and noise, the estimated slope moves away from zero, which indicates a reduced consistency of GM CBF with GM PV. The tendency

was for either over- or under-estimation in voxels with low GM PV and the opposite at high GM PV, with the best accuracy in the middle of the range. For the two LR methods, the slope values were comparatively near to zero in all experiments even as the parameters of the experiment were varied. Whilst this would be consistent with LR being an effective PVEc approach and thus relatively insensitive to the sources of error studies here, it would appear that this arises largely from the smoothing effect of the LR kernel. Even when the slope value was small, implying effective correction, there was global under or over estimation of CBF. Thus the resulting GM CBF images were very smooth and thereby had a fairly consistent value for all voxels irrespective of GM PV estimate. On the contrary, the genuine variations in GM CBF were being suppressed leading to spatially specific errors that are not captured by ROI analysis method. The ROI curve method has previously been used to judge the effectiveness of PVEc on real ASL data. This study implies that this should be interpreted with care, as methods that introduce greater smoothness in the estimated CBF images can appear to be very successful using this approach, despite introducing error into the CBF image.

Healthy Cohort Experiments

Results from the analysis using healthy cohort data have shown an increase of estimated GM CBF in voxels containing low GM PV after PVEc, which is consistent with existing PVEc studies (Ahlgren et al., 2014; Asllani et al., 2008; Chappell et al., 2011). The estimated slope values for the non-PVEc method were between 600 to 3000 ml/100g/min%, indicating a positive correlation with GM PV estimate as shown in Figure 9. Cross referencing with the simulation results in Figure 7A, such values would be consistent with data containing noise in the range we might expect to observe in ASL data of this duration of acquisition. However, the range would imply there is some variability within the group in either the noise or the degree to which the data from each subject exhibits PVE.

The slope values after LR PVEc were around -1000 ml/100g/min/%, which is more negative than were observed in the simulation study for LR correction. A broader range of slope values (between -2500 to -500 ml/100g/min/%) were observed in results of the spatially regularized method. This latter result could be explained from the simulations by a combination of the different errors simulated. However, the variability in slope value observed and inconsistency with the LR PVEc method means that care needs to be taken in interpreting these results using the simulations since the simulated scenarios might not be fully capturing variations in GM CBF that are observed in practice.

The repeatability of CBF quantification was improved after PVEc as evidenced by the reduction of wsCV for all repeated experiments. For the results without PVEc, the repeatability decreased with a longer gap between repeated scans, where session repeat was the most repeatable and month repeat the least. This agrees with the findings of the original reproducibility study using the same dataset although the current study used voxelwise calibration instead of CSF-based approach in the original reproducibility study (Mezue et al., 2014). Together with the results of the ROI analysis, the nearly identical results of the two LR methods imply that it makes no difference whether the LR method is applied before or after CBF quantification. Although wsCV of the spatially regualrized method was over 10% for GM ROI between 10% to 40%, it fell to the same level with the LR method for GM ROI greater than 40%. This confirmed the simulation results that spatially regularized method worked more consistently in medium and high ranges of PV estimates.

The metric of wsCV has been widely used to assess the repeatability of CBF quantification as well as PVEc studies in the past (Ahlgren et al., 2016; Mutsaerts et al., 2015; Petersen et al., 2010). Ahlgen et al have reported an improved repeatability after LR applied to single-PLD PCASL data (Ahlgren et al., 2016). In particular, the mean within-voxel coefficient of variation for GM reduced from 30% to 25%

after correcting PVE by the LR method. Although the present study separated GM ROI into three intervals, we have observed a similar improvement (wsCV reduced by 5%) after the same correction (LR) method were implemented. In addition, Ahlgen et al also investigated the effect of smoothing on CBF images in terms of repeatability, which, as might be expected, was also found to improve repeatability between scans. This implies that the smoothing observed in the GM CBF map after LR PVEc might make a substantial contribution to making the CBF measurements more repeatable apart from correcting for PVE. This is consistent with the observations of Ahlgren at al. that the modified least trimmed squares (mLTS) PVEc method, which introduces less smoothing, was also associated with lower repeatability. While Ahlgren et al found that the mLTS method for PVEc actually made the ASL measurements less repeatable than no correction, this study has shown the spatially regularized method does offer improvements in repeatability.

Limitations

A limitation of the study is the reliance on PCASL simulated data for the baseline experiments. Despite the four different types of GM CBF patterns created, the simulation does not necessarily capture physiological spatial variations in perfusion. This may explain some discrepancies seen under comparisons with the healthy cohort data. However, the simulations have provided a well-controlled environment in which to systemically study sensitivity to error for different PVEc methods. Unlike previous simulation studies of this kind, here one hundred individual structural images were employed to generate a simulated cohort that captured more natural variation in the population, seeking to avoid any bias in the results from choosing a single subject on which to base the PV estimates.

We have not sought to investigate all possible PVEc methods that have been proposed here and have concentrated on comparing the simplest LR approach to the other major class of methods to use spatial information. A potential extension would be to consider mLTS and a wider range of kernel sizes and shapes (such as a circular rather than square kernel) for LR. However, based on the extra exploration found in the supplementary material for a smaller kernel size of 3 × 3 and existing literature on PVEc (Asllani et al., 2008; Chappell et al., 2011), we believe that broadly the conclusions of this study would hold for other variants on LR. Although we have considered the impact of PVEc on multi-PLD PCASL in this study on CBF estimation, we have not sought to examine how this might affect estimates of arterial transit time, which is of interest in hemodynamic studies in its own right (MacIntosh et al., 2010). This is something that could be explored using the same framework in the future.

The current study specifically investigated the influence of variabilities on PVEc for ASL perfusion estimation in individuals and the corresponding repeatability in a cohort of healthy subjects. An open question is the degree to which PVEc might be useful when performing group studies to look for changes in perfusion associated with activation or disease, and whether the different PVEc methods are more of less favorable than each other for this application. This might partially be answered using simulations like those examined here, although further work using different PVEc methods within ASL groups studies would still be valuable to address these questions.

Conclusions

In this work, we have investigated the impact of different sources of error on PVEc methods when applied to PCASL perfusion images. In simulated experiments, the spatially regularized method demonstrated consistent greater ability to preserve spatial details in CBF maps but was also shown

to be sensitive to noise and errors in the PV estimates. The LR method was consistently less sensitive to noise and errors, but this appeared to be largely due to the smoothing effect of the kernel employed, thus the estimated GM CBF was less accurate than the spatially regularized method whenever there were features in the data represented over short length scales. The results from a repeatability study in healthy individuals have shown that both methods reduce intersession variability, the greater reduction coming from LR potentially being another effect of the spatial smoothing associated with the method.

Acknowledgments

This work was supported by the following grants and institutions: EP/P012361/1 (EPSRC, UK), Guarantors of Brain (UK), COST Action BM1103. The authors would like to thank Martin Craig for technical support for this study.

References

Ahlgren, A., Wirestam, R., Lind, E., Ståhlberg, F., Knutsson, L., 2016. A linear mixed perfusion model for tissue partial volume correction of perfusion estimates in dynamic susceptibility contrast MRI: Impact on absolute quantification, repeatability, and agreement with pseudo-continuous arterial spin labeling. Magn. Reson. Med. C, 1–12. doi:10.1002/mrm.26305

Ahlgren, A., Wirestam, R., Petersen, E.T., Ståhlberg, F., Knutsson, L., 2014. Partial volume correction of brain perfusion estimates using the inherent signal data of time-resolved arterial spin labeling. NMR Biomed. 27, 1112–1122. doi:10.1002/nbm.3164

Alsop, D.C., Detre, J.A., Golay, X., Günther, M., Hendrikse, J., Hernandez-Garcia, L., Lu, H., Macintosh, B.J., Parkes, L.M., Smits, M., van Osch, M.J.P., Wang, D.J.J., Wong, E.C., Zaharchuk, G., 2014. Recommended implementation of arterial spin-labeled perfusion MRI for clinical applications: A consensus of the ISMRM perfusion study group and the european consortium for ASL in dementia. Magn. Reson. Med. doi:10.1002/mrm.25197

Asllani, I., Borogovac, A., Brown, T.R., 2008. Regression algorithm correcting for partial volume effects in arterial spin labeling MRI. Magn. Reson. Med. 60, 1362–1371. doi:10.1002/mrm.21670

Buxton, R.B., Frank, L.R., Wong, E.C., Siewert, B., Warach, S., Edelman, R.R., 1998. A general kinetic model for quantitative perfusion imaging with arterial spin labeling. Magn. Reson. Med. 40, 383–396. doi:10.1002/mrm.1910400308

Chappell, M.A., Groves, A.R., MacIntosh, B.J., Donahue, M.J., Jezzard, P., Woolrich, M.W., 2011. Partial volume correction of multiple inversion time arterial spin labeling MRI data. Magn. Reson. Med. 65, 1173–1183. doi:10.1002/mrm.22641

Chappell, M.A., Groves, A.R., Whitcher, B., Woolrich, M.W., 2009. Variational Bayesian Inference for a Nonlinear Forward Model. IEEE Trans. Signal Process. 57, 223–236. doi:10.1109/TSP.2008.2005752

Clement, P., Mutsaerts, H.-J., Václavů, L., Ghariq, E., Pizzini, F.B., Smits, M., Acou, M., Jovicich, J., Vanninen, R., Kononen, M., Wiest, R., Rostrup, E., Bastos-Leite, A.J., Larsson, E.-M., Achten, E., 2017. Variability of physiological brain perfusion in healthy subjects – A systematic review of modifiers. Considerations for multi-center ASL studies. J. Cereb. Blood Flow Metab. 0271678X17702156. doi:10.1177/0271678X17702156

Detre, J.A., Leigh, J.S., Williams, D.S., Koretsky, A.P., 1992. Perfusion imaging. Magn. Reson. Med. 23, 37–45. doi:10.1002/mrm.1910230106

Golay, X., Hendrikse, J., Lim, T.C.C., 2004. Perfusion imaging using arterial spin labeling. Top. Magn. Reson. Imaging 15, 10–27. doi:10.1097/00002142-200402000-00003

Groves, A.R., Chappell, M.A., Woolrich, M.W., 2009. Combined spatial and non-spatial prior for inference on MRI time-series. Neuroimage 45, 795–809. doi:10.1016/j.neuroimage.2008.12.027

Henery, C.C., Mayhew, T.M., 1989. The cerebrum and cerebellum of the fixed human brain: efficient and unbiased estimates of volumes and cortical surface areas. J. Anat. 167, 167–180.

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002a. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17, 825–841. doi:10.1016/S1053-8119(02)91132-8

Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL. Neuroimage.

doi:10.1016/j.neuroimage.2011.09.015

Jenkinson, M., Pechaud, M., Smith, S., 2002b. BET2 - MR-Based Estimation of Brain , Skull and Scalp Surfaces. Elev. Annu. Meet. Organ. Hum. Brain Mapp. 17, 2002. doi:citeulike-article-id:1179617

Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Med. Image Anal. 5, 143–156. doi:10.1016/S1361-8415(01)00036-6

Johnson, N.A., Jahng, G.-H., Weiner, M.W., Miller, B.L., Chui, H.C., Jagust, W.J., Gorno-Tempini, M.L., Schuff, N., 2005. Pattern of Cerebral Hypoperfusion in Alzheimer Disease and Mild Cognitive Impairment Measured with Arterial Spin-labeling MR Imaging: Initial Experience1. Radiology 234, 851–859. doi:10.1148/radiol.2343040197

Le Heron, C.J., Wright, S.L., Melzer, T.R., Myall, D.J., Macaskill, M.R., Livingston, L., Keenan, R.J., Watts, R., Dalrymple-Alford, J.C., Anderson, T.J., 2014. Comparing cerebral perfusion in Alzheimer's disease and Parkinson's disease dementia: an ASL-MRI study. J. Cereb. Blood Flow Metab. 34, 1–7. doi:10.1038/jcbfm.2014.40

MacIntosh, B.J., Filippini, N., Chappell, M.A., Woolrich, M.W., Mackay, C.E., Jezzard, P., 2010. Assessment of arterial arrival times derived from multiple inversion time pulsed arterial spin labeling MRI. Magn. Reson. Med. 63, 641–647. doi:10.1002/mrm.22256

Mezue, M., Segerdahl, A.R., Okell, T.W., Chappell, M.A., Kelly, M.E., Tracey, I., 2014. Optimization and reliability of multiple postlabeling delay pseudo-continuous arterial spin labeling during rest and stimulus-induced functional task activation. J. Cereb. Blood Flow Metab. 34, 1919–27. doi:10.1038/jcbfm.2014.163

Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L.R., Griffanti, L., Douaud, G., Okell, T.W., Weale, P., Dragonu, I., Garratt, S., Hudson, S., Collins, R., Jenkinson, M., Matthews, P.M., Smith, S.M., 2016. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat. Neurosci. 19. doi:10.1038/nn.4393

Mutsaerts, H.J.M.M., Richard, E., Heijtel, D.F.R., Van Osch, M.J.P., Majoie, C.B.L.M., Nederveen, A.J., 2014. Gray matter contamination in arterial spin labeling white matter perfusion measurements in patients with dementia. NeuroImage Clin. 4, 139–144. doi:10.1016/j.nicl.2013.11.003 Review Article

Mutsaerts, H.J.M.M., van Osch, M.J.P., Zelaya, F.O., Wang, D.J.J., Nordhøy, W., Wang, Y., Wastling, S., Fernandez-Seara, M.A., Petersen, E.T., Pizzini, F.B., Fallatah, S., Hendrikse, J., Geier, O., Günther, M., Golay, X., Nederveen, A.J., Bjørnerud, A., Groote, I.R., 2015. Multi-vendor reliability of arterial spin labeling perfusion MRI using a near-identical sequence: Implications for multi-center studies. Neuroimage 113, 143–152. doi:10.1016/j.neuroimage.2015.03.043

Okell, T.W., Chappell, M.A., Kelly, M.E., Jezzard, P., 2013. Cerebral blood flow quantification using vessel-encoded arterial spin labeling. J. Cereb. Blood Flow Metab. 33, 1716–24. doi:10.1038/jcbfm.2013.129

Penny, W.D., Trujillo-Barreto, N.J., Friston, K.J., 2005. Bayesian fMRI time series analysis with spatial priors. Neuroimage 24, 350–362. doi:10.1016/j.neuroimage.2004.08.034

Petersen, E.T., Mouridsen, K., Golay, X., 2010. The QUASAR reproducibility study, Part II: Results from a multi-center Arterial Spin Labeling test-retest study. Neuroimage 49, 104–113.

Steketee, R.M.E., Bron, E.E., Meijboom, R., Houston, G.C., Klein, S., Mutsaerts, H.J.M.M., Mendez Orellana, C.P., de Jong, F.J., van Swieten, J.C., van der Lugt, A., Smits, M., 2015. Early-stage differentiation between presenile Alzheimer's disease and frontotemporal dementia using

arterial spin labeling MRI. Eur. Radiol. doi:10.1007/s00330-015-3789-x

Van Osch, M.J.P., Hendrikse, J., Van Der Grond, J., 2007. Sensitivity comparison of multiple vs. single inversion time pulsed arterial spin labeling fMRI. J. Magn. Reson. Imaging 25, 215–221. doi:10.1002/jmri.20823

Wu, W.C., Jiang, S.F., Yang, S.C., Lien, S.H., 2011. Pseudocontinuous arterial spin labeling perfusion magnetic resonance imaging-A normative study of reproducibility in the human brain. Neuroimage 56, 1244–1250. doi:10.1016/j.neuroimage.2011.02.080

Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging 20, 45–57. doi:10.1109/42.906424

Zhao, M.Y., Petersen, E.T., Chappell, M.A., 2016. Cerebral Perfusion Quantification Using Turbo-QUASAR Arterial Spin Labelling MRI : A Model-based Approach. Proc. MEIBioeng 2–4.

Supplementary Material

Simulation study: multi-PLD data (kernel size 5×5 *for the LR methods)*

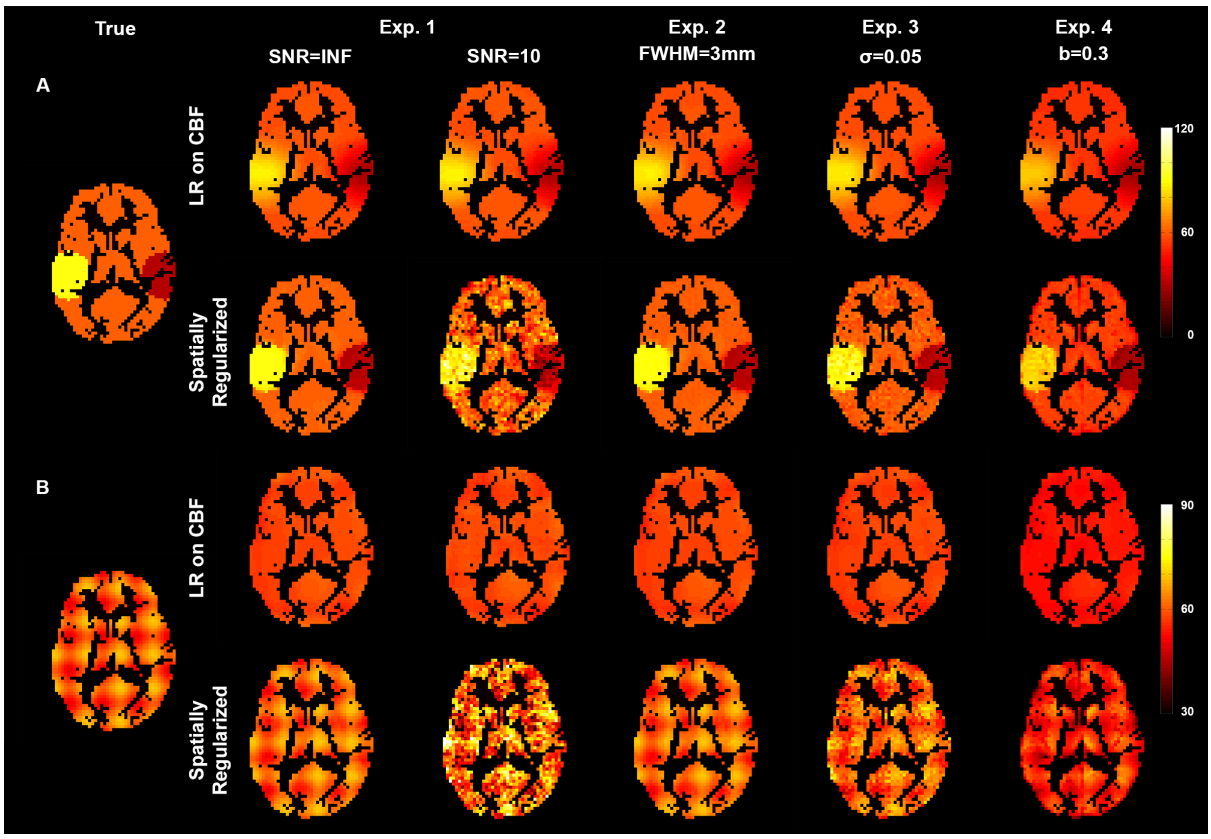


Figure 11: Estimated GM CBF map for multi-PLD simulated data using PV estimates of subject 1 as the reference PV estimates (kernel size 5×5 for the LR methods). (A) Hyper/hypo GM CBF; (B) Fast sinusoidal GM CBF variation.

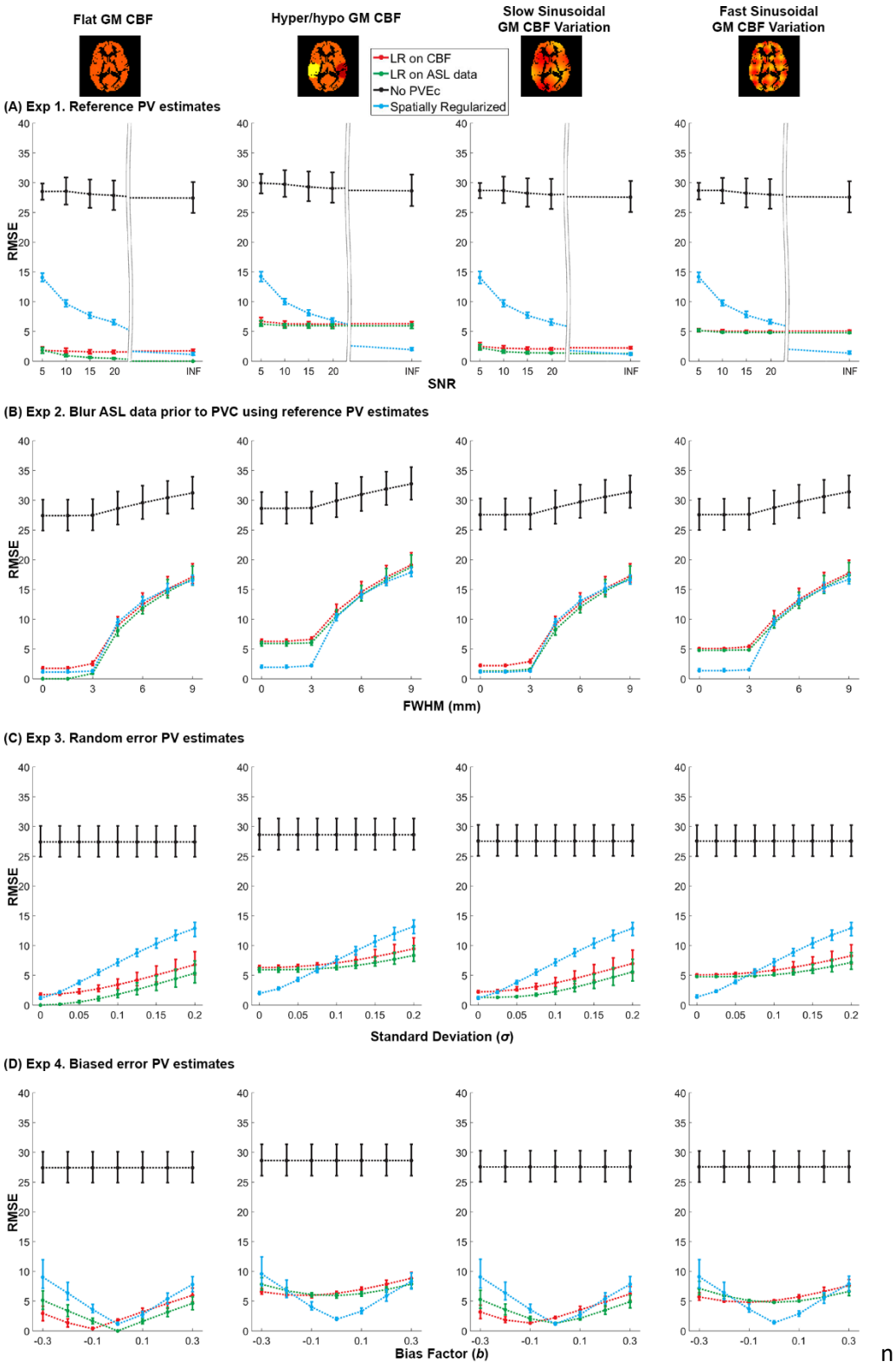


Figure 12: RMSE between estimated and simulated GM CBF using noise-free simulated multi-PLD data (kernel size 5×5 for the LR methods).

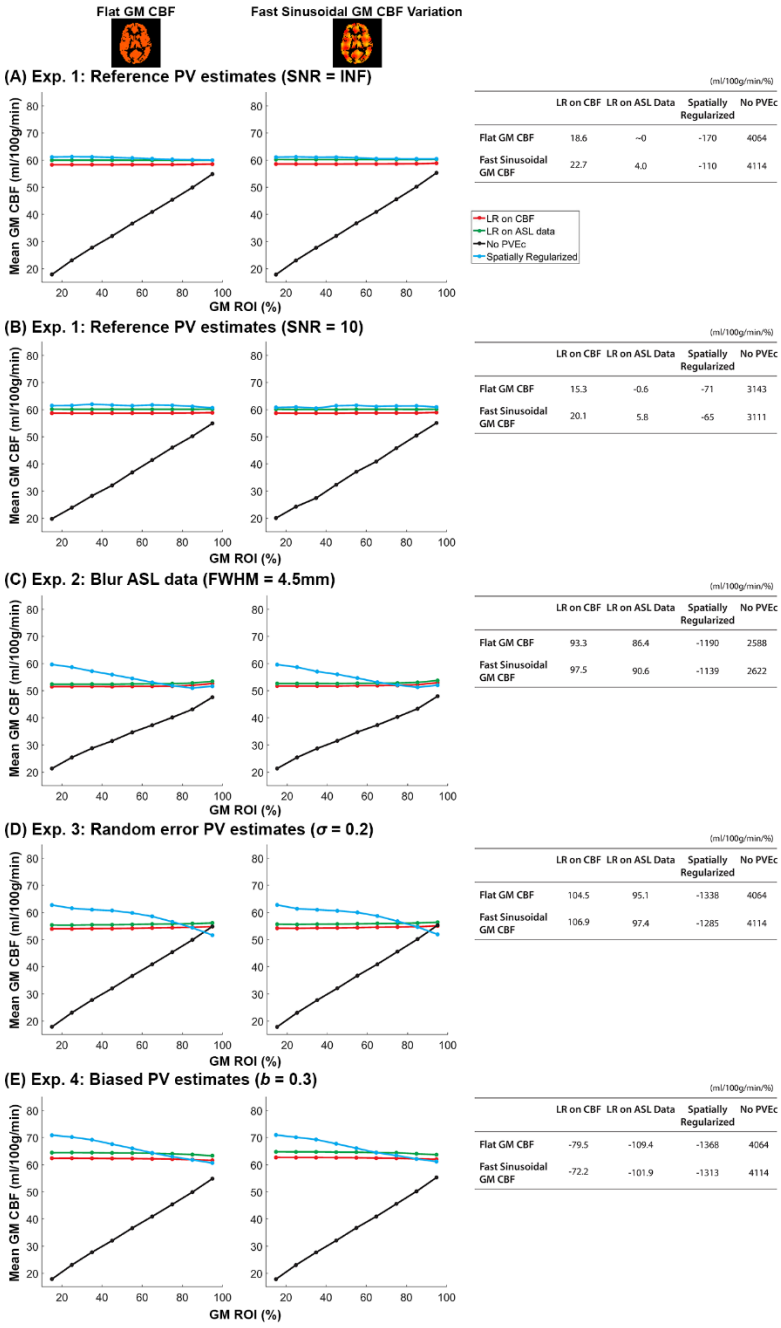


Figure 13: ROI curve for all experiments using multi-PLD data (kernel size 5×5 for the LR methods). The first row of curves shows the ROI analysis in a perfect situation (SNR=INF and reference PV estimates). Subsequent plots reveal the changes of mean GM CBF in each experiment. Each adjacent table shows the estimated slope values for the linear regression fitting between estimated GM CBF and GM ROI.

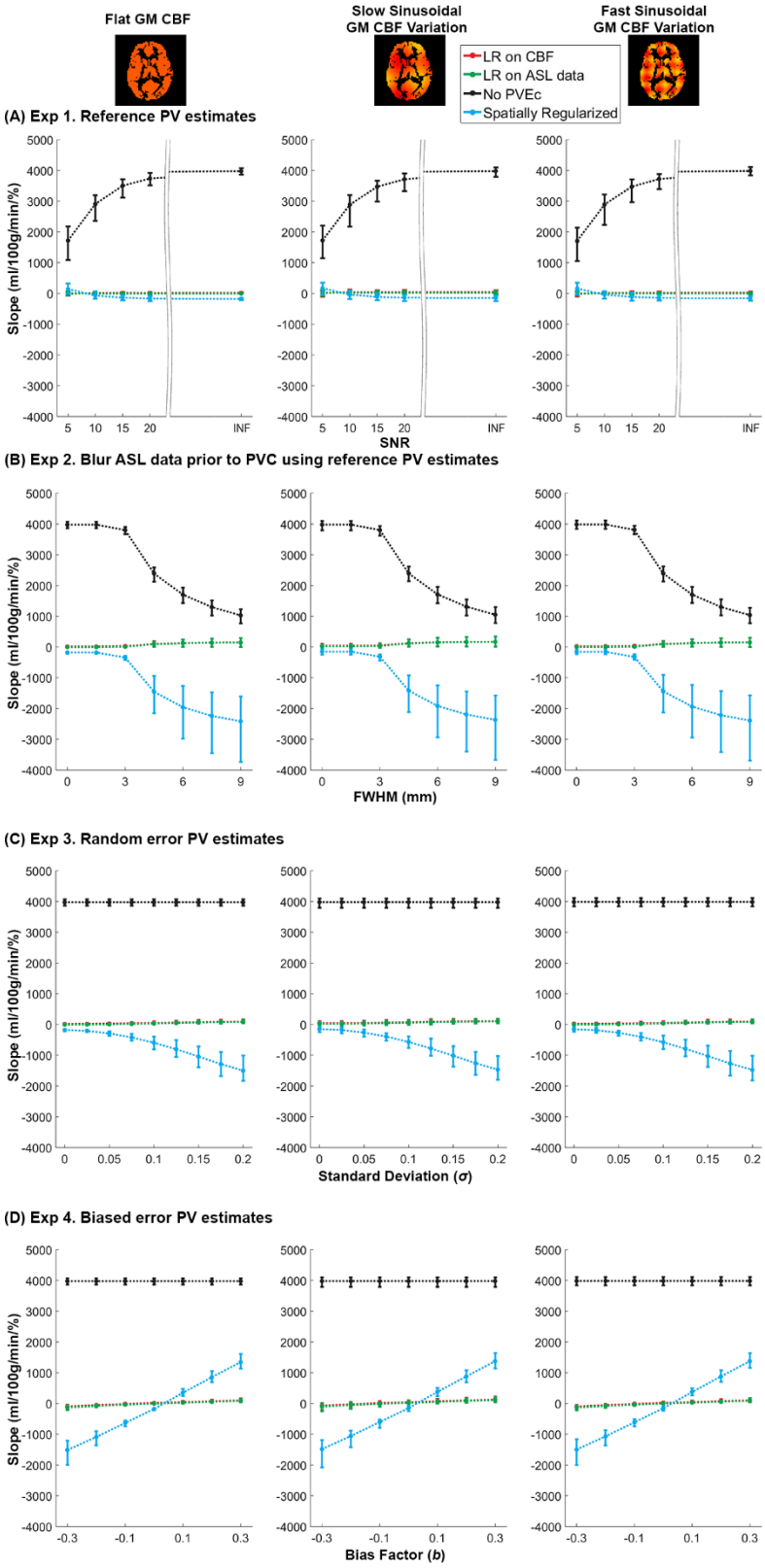


Figure 14: Estimated slope values of the extended ROI analysis using multi-PLD data (kernel size 5×5 for the LR methods).

Simulation study: single-PLD data (kernel size 3×3 for the LR methods)

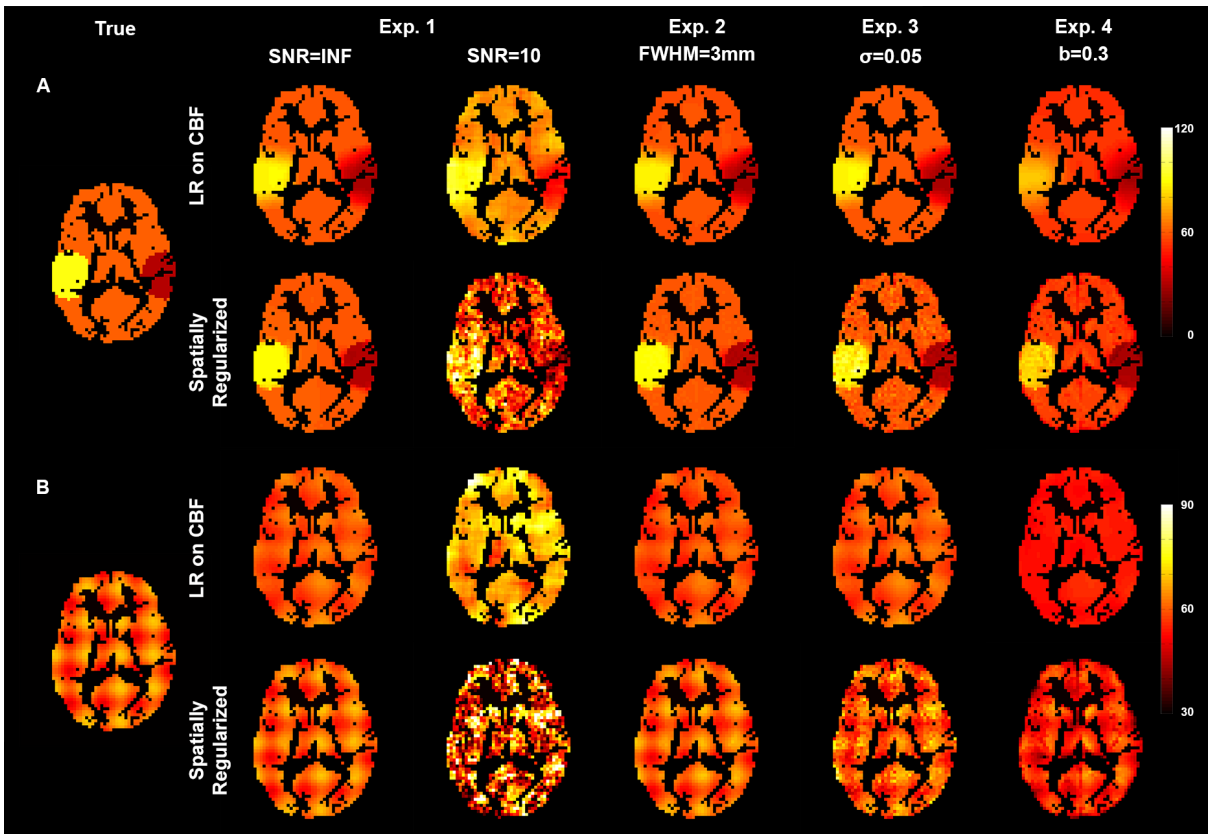


Figure 15: Estimated GM CBF map for single-PLD simulated data using PV estimates of subject 1 as the reference PV estimates (kernel size 3×3 for the LR methods). (A) Hyper/hypo GM CBF; (B) Fast sinusoidal GM CBF variation.

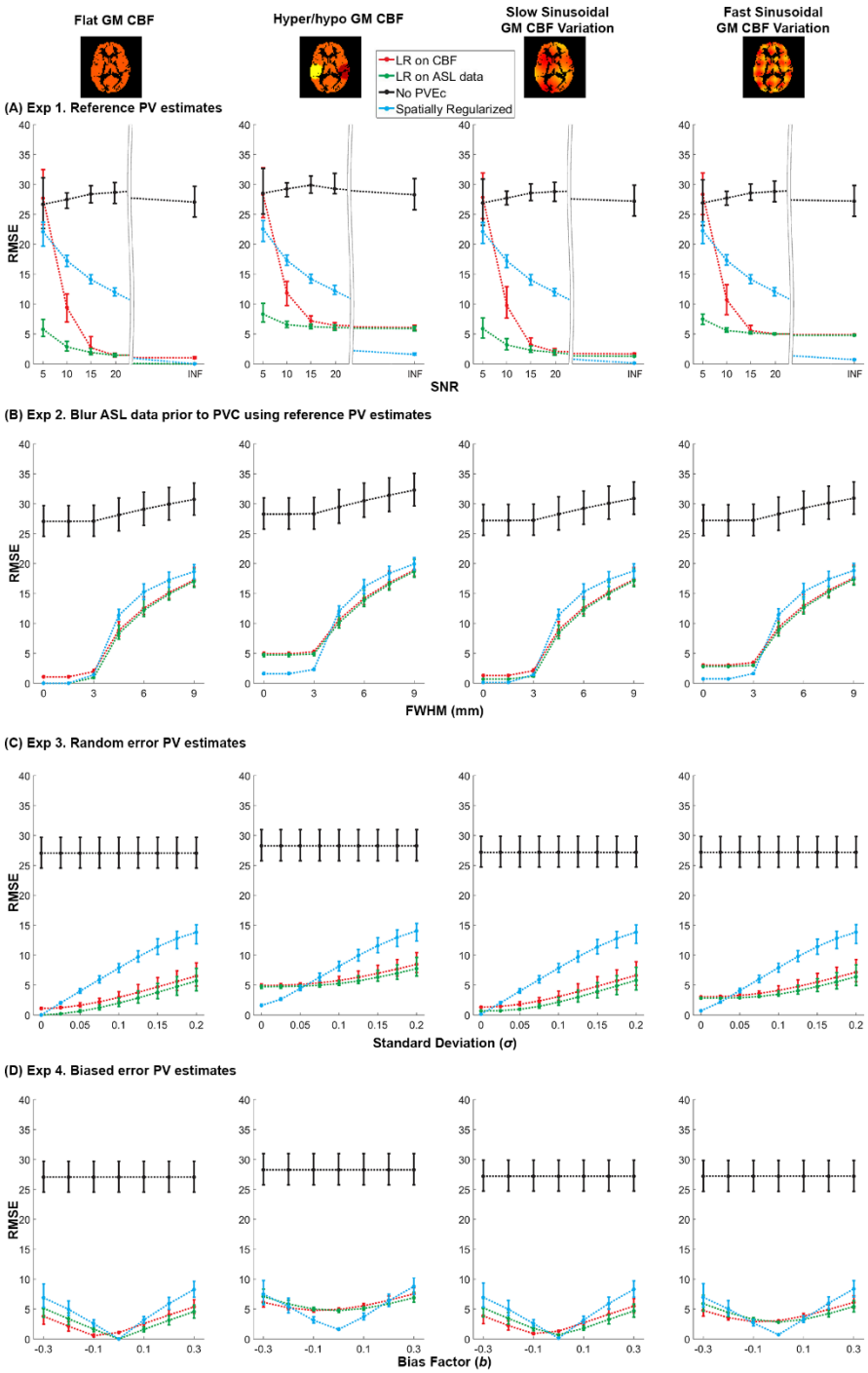


Figure 16: RMSE between estimated and simulated GM CBF using noise-free simulated single-PLD data (kernel size 3×3 for the LR methods).

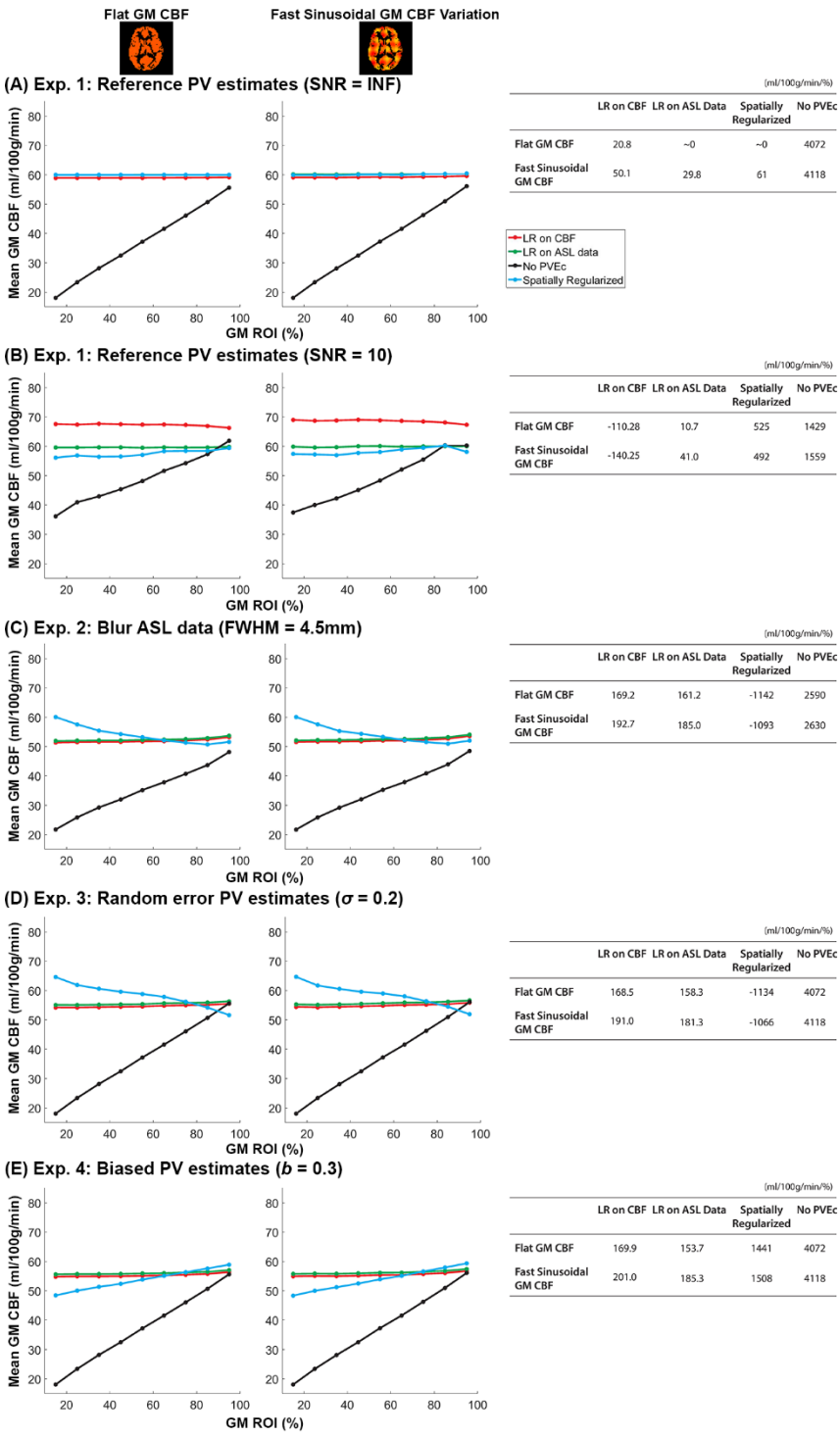


Figure 17: ROI curve for all experiments using single-PLD data (kernel size 3×3 for the LR methods). The first row of curves shows the ROI analysis in a perfect situation (SNR=INF and reference PV estimates). Subsequent plots reveal the changes of mean GM CBF in each experiment. Each adjacent table shows the estimated slope values for the linear regression fitting between estimated GM CBF and GM ROI.

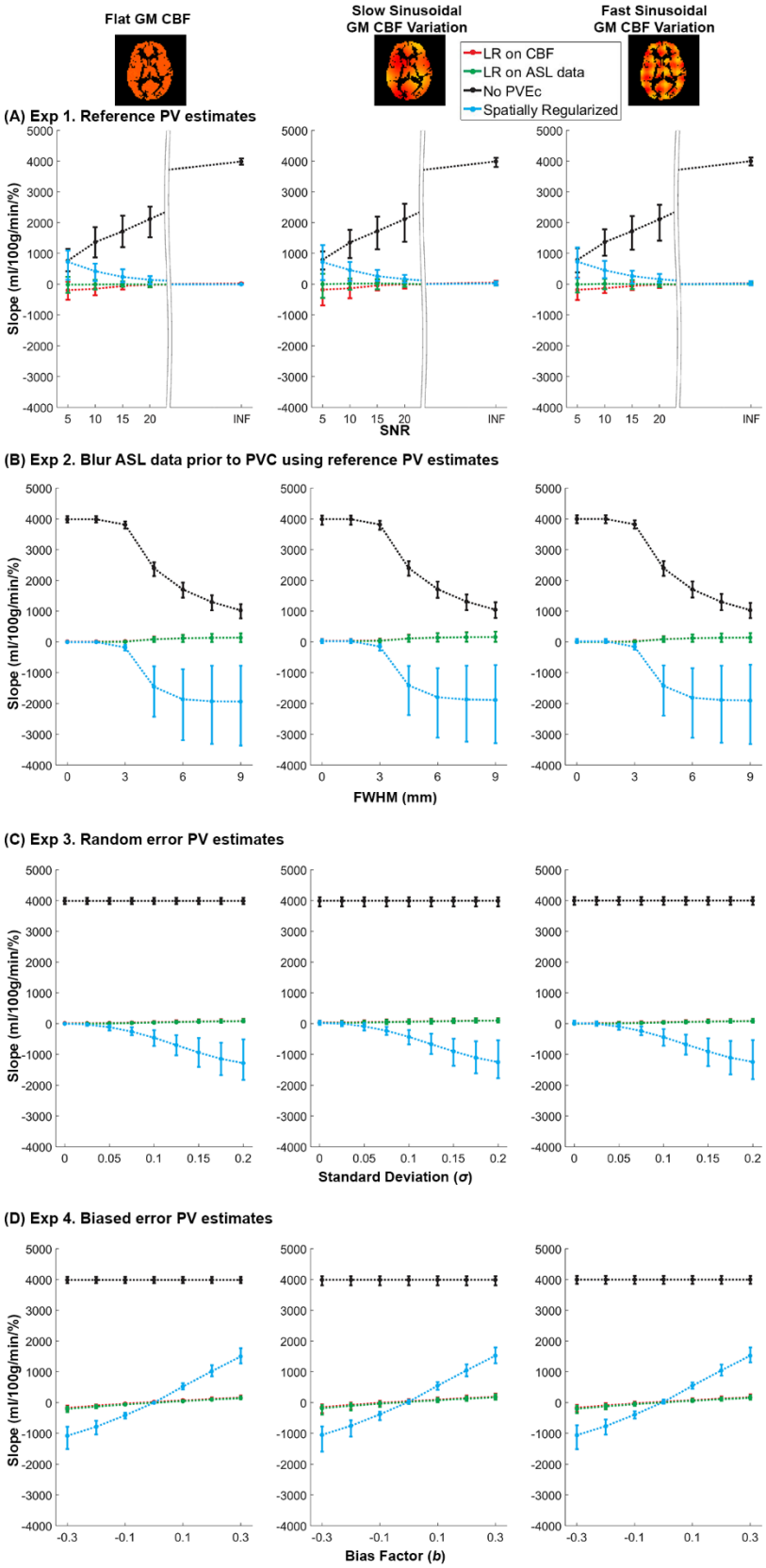


Figure 18: Estimated slope values of the extended ROI analysis using single-PLD data (kernel size 3×3 for the LR methods).

Simulation study: multi-PLD data (kernel size 3 × 3 for the LR methods)

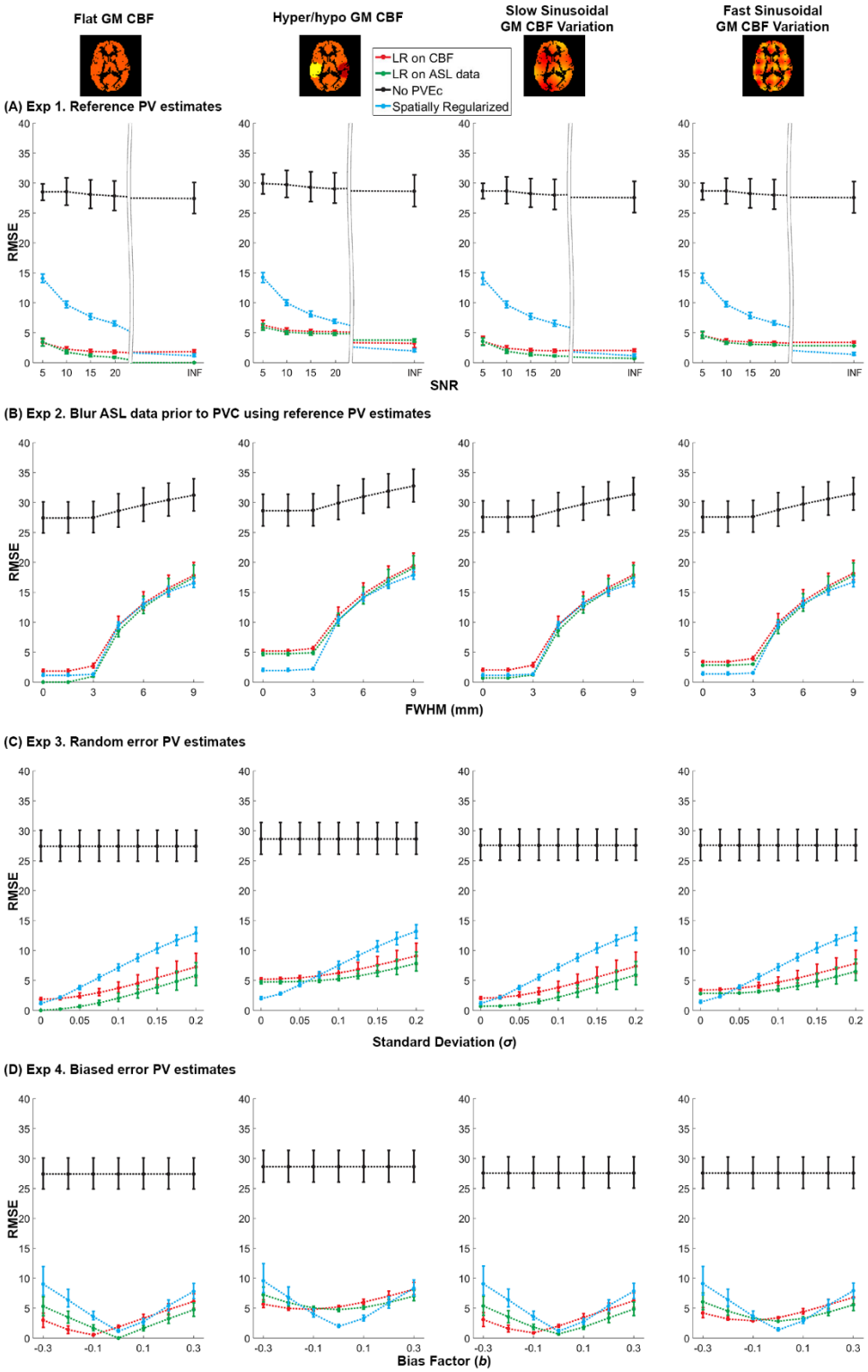
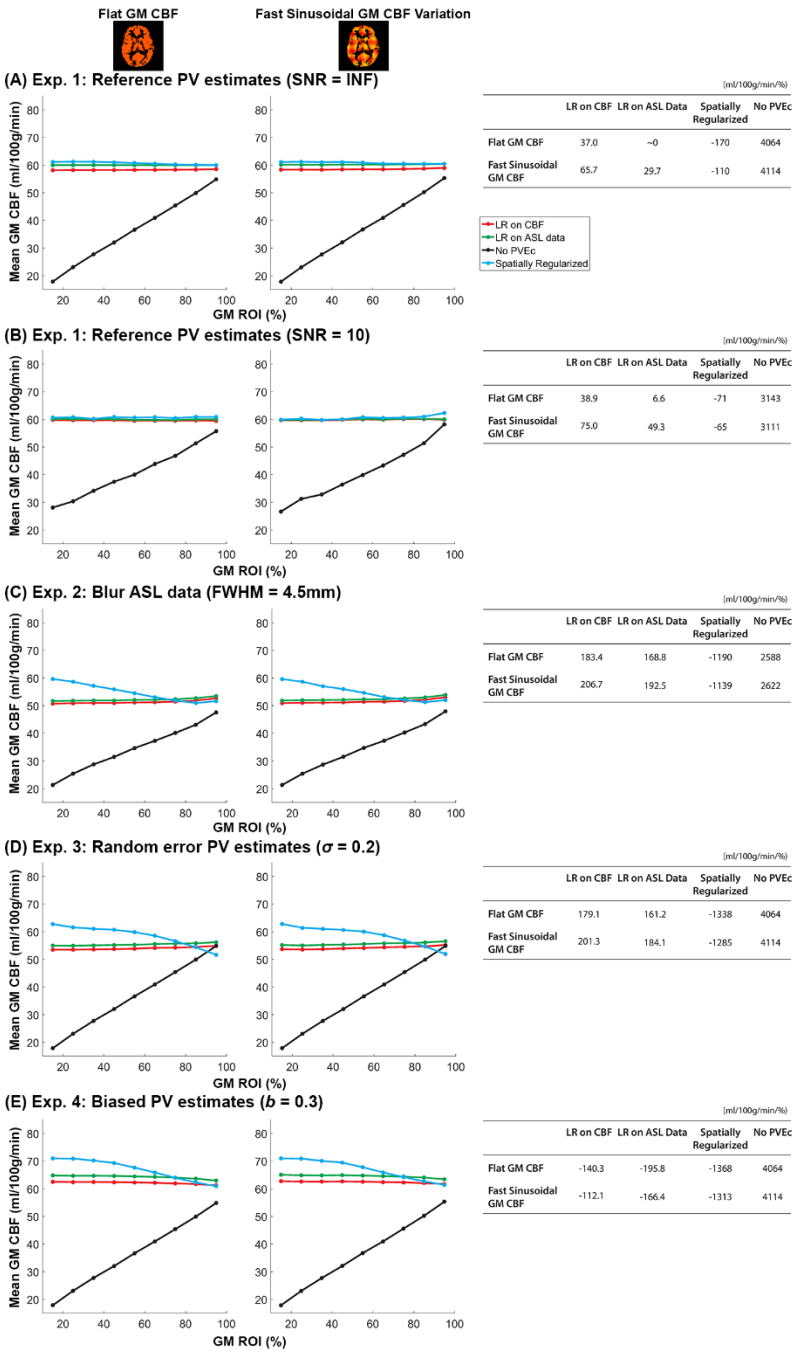


Figure 19: RMSE between estimated and simulated GM CBF using noise-free simulated multi-PLD data (kernel size 3 × 3 for the LR methods).



(ml/100g/min/%)

	LR on CBF	LR on ASL Data	Spatially Regularized	No PVEc
Flat GM CBF	37.0	~0	-170	4064
Fast Sinusoidal GM CBF	65.7	29.7	-110	4114

(ml/100g/min/%)

	LR on CBF	LR on ASL Data	Spatially Regularized	No PVEc
Flat GM CBF	38.9	6.6	-71	3143
Fast Sinusoidal GM CBF	75.0	49.3	-65	3111

(ml/100g/min/%)

	LR on CBF	LR on ASL Data	Spatially Regularized	No PVEc
Flat GM CBF	183.4	168.8	-1190	2588
Fast Sinusoidal GM CBF	206.7	192.5	-1139	2622

(ml/100g/min/%)

	LR on CBF	LR on ASL Data	Spatially Regularized	No PVEc
Flat GM CBF	179.1	161.2	-1338	4064
Fast Sinusoidal GM CBF	201.3	184.1	-1285	4114

(ml/100g/min/%)

	LR on CBF	LR on ASL Data	Spatially Regularized	No PVEc
Flat GM CBF	-140.3	-195.8	-1368	4064
Fast Sinusoidal GM CBF	-112.1	-166.4	-1313	4114

Figure 20: ROI curve for all experiments using multi-PLD data (kernel size 3×3 for the LR methods). The first row of curves shows the ROI analysis in a perfect situation (SNR=INF and reference PV estimates). Subsequent plots reveal the changes of mean GM CBF in each experiment. Each adjacent table shows the estimated slope values for the linear regression fitting between estimated GM CBF and GM ROI.

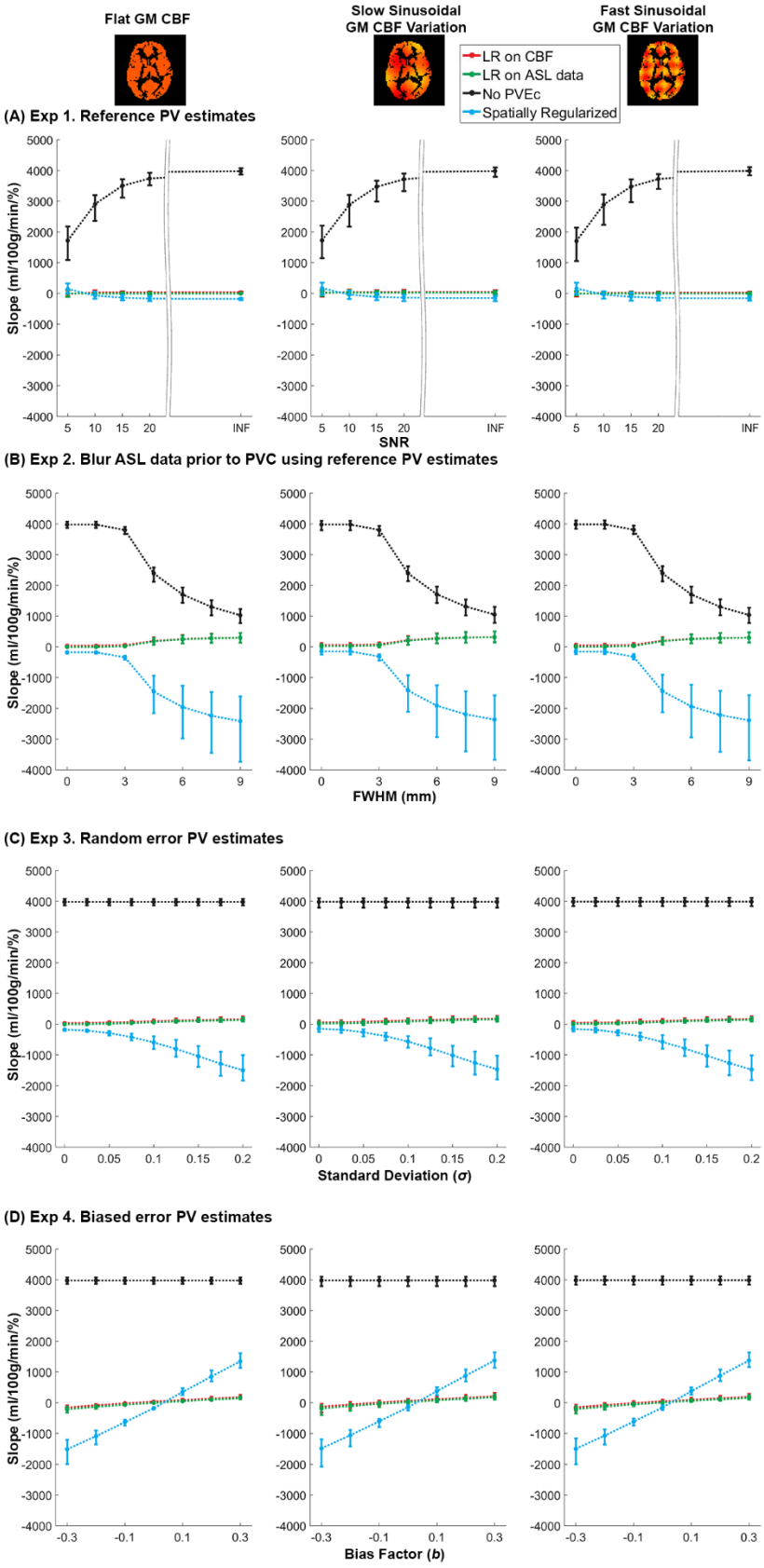


Figure 21: Estimated slope values of the extended ROI analysis using multi -PLD data (kernel size 3×3 for the LR methods).