Effects of systematic partial volume errors on the computation of mean gray matter cerebral blood flow with Arterial Spin Labeling MRI

Petr, J.; Mutsaerts, H. J. M. M.; de Vita, E.; Steketee, R. M. E.; Smits, M.; Nederveen, A. J.; Hofheinz, F.; van den Hoff, J.; Asllani, I.;

Originally published:

June 2018

Magnetic Resonance Materials in Physics, Biology and Medicine 31(2018)6, 725-734

DOI: https://doi.org/10.1007/s10334-018-0691-y

Noname manuscript No.
(will be inserted by the editor)

Effects of systematic partial volume errors on the estimation of gray matter cerebral blood flow with Arterial Spin Labeling MRI

Jan Petr · Henri J.M.M. Mutsaerts · Enrico De Vita · Rebecca M.E. Steketee · Marion Smits · Aart J. Nederveen · Frank Hofheinz · Jörg van den Hoff · Iris Asllani

Received: date / Accepted: date

Short title: Effects of partial volume errors on ASL GM CBF

Jan Petr
Bautzner Landstrae 400 01328 Dresden Tel.: +49 351 260 2709 E-mail: j.petr@hzdr.de

JP, FH, JVDH
Helmholtz-Zentrum Dresden-Rossendorf, Institute of Radiopharmaceutical Cancer Research, Dresden, Germany

JP, HM, IA
Department of Biomedical Engineering, Rochester Institute of Technology, Rochester, NY, USA

HM
Sunnybrook Research Institute, Toronto, Canada

HM, AN
Department of Radiology, Academic Medical Center Amsterdam, the Netherlands

HM
Department of Radiology, University Medical Center Utrecht, the Netherlands

EDV
Lysholm Department of Neuroradiology, National Hospital for Neurology and Neurosurgery, London, United Kingdom

EDV
Department of Biomedical Engineering, School of Biomedical Engineering & Imaging Sciences, Kings College London, Kings Health Partners, St Thomas Hospital, London, SE1 7EH, United Kingdom

RS, MS
Department of Radiology and Nuclear Medicine, Erasmus MC – University medical center Rotterdam, the Netherlands

JVDH
Department of Nuclear Medicine, University Hospital Carl Gustav Carus, Technical University Dresden, Dresden, Germany

Abstract Objectives: Partial volume correction (PVC) is an important step in arterial spin labeling (ASL) MRI that is used to separate perfusion from structural effects when computing the mean gray-matter (GM) perfusion. There are three main methods to perform this correction: (1) GM-threshold, which include only voxels with GM volume above a preset threshold; (2) GM-Weighted, which uses voxelwise GM contribution combined with thresholding; or (3) PVC, which applies a spatial linear regression algorithm to estimate the flow contribution of each tissue at a given voxel. In all cases, GM volume is obtained using PV maps extracted from the segmentation of the T1w image. As such, PV maps contain errors due to the difference in readout-type and spatial resolution between ASL and T1w images. Here, we estimated these errors and evaluated their effect on the performance of each PV-correction method in computing GM CBF.

Materials and Methods: Twenty-two volunteers were scanned using 2D EPI and 3D spiral ASL. For each PV-correction method, GM CBF was computed using PV maps simulated to contain estimated errors due to spatial resolution mismatch and geometric distortions which are caused by the mismatch in read-out between ASL and T1w images. Results were analyzed to assess the effect of each error on estimation of GM CBF from ASL data.

Results: Geometric distortion had the largest effect on the 2D EPI data whereas the 3D spiral was most affected by the resolution mismatch. The PVC method outperformed the GM-Threshold even in the presence of combined errors from resolution mismatch and geometric distortions. The quantitative advantage of PVC was 16% without and 10% with the combined errors for both 2D and 3D ASL. Consistent with theoretical expectations, for error-free PV maps, the PVC method extracted the true GM CBF. In contrast, GM-Weighted overestimated GM CBF by 5% whereas GM-Threshold under-

estimated it by 16%. The presence of PV-map errors decreased the calculated GM CBF for all methods.

Conclusion: The quality of PV maps presents no argument for preferring the GM-Threshold method to PVC in clinical applications of ASL.

Keywords arterial spin labeling · perfusion magnetic resonance imaging · cerebral blood flow · partial volume

INTRODUCTION

The quality of arterial spin labeling (ASL) perfusion MRI has improved significantly with the increased availability of multi-channel coils [1], 3T scanners [2], more advanced 3D readouts [3,4], background-suppression schemes [5,6], and improvements in labeling efficiency [7]. As a result, the application of ASL in clinical research has been on the rise [8,9].

Despite these developments, the spatial resolution of ASL has remained relatively low. Therefore, the majority of voxels in ASL images contain a mixture of perfusion signals from gray matter (GM), white matter (WM), and CSF, a phenomenon referred to as partial-volume effects. Since GM perfusion is reported to be 2–4.5 times higher than WM perfusion [10,11], and CSF is not perfused, the measured ASL signal in a given voxel is dependent on the fractional contributions of GM and WM, i.e., its tissue composition. Consequently, differences in measured perfusion across regions within a subject or for the same region across subjects could, to a varying degree, be attributable to differences in tissue composition rather than actual changes in perfusion. This hypothesis was borne out in an ASL study of healthy aging which showed that around 50 % of the CBF difference between healthy elderly and young participants could be accounted for by atrophy alone [12]. It is therefore of principal relevance to separate changes in CBF from structural variability in studies when the latter cannot be ruled out [13,14].

In this context, several approaches for partial volume (PV) correction have been proposed. Originally, a fixed ratio between GM and WM perfusion was assumed [15]. Later, a linear regression was used to model the relation of the observed CBF and GM and WM PV-maps, to solve for the unknown PV-corrected GM and WM CBF within a spatially defined kernel [10]. The main focus of the subsequent work on PV correction in ASL has been to reduce

the smoothing effects inherent in the original algorithm [16,17]. Less attention, however, has been given to the quality of the GM and WM PV maps from which the algorithms receive the structural information, and which can influence the results of the PV correction [18]. Typically, the PV maps (obtained from the segmentation of high-resolution T1-weighted ($T1_w$) images) are co-registered and resampled to match the low-resolution ASL image space. Therefore, differences in geometric distortion and effective resolution between the ASL and $T1_w$ readouts are propagated to the PV maps and subsequently in the PV correction process and CBF computation [19–22]. Furthermore, differences in ASL readout schemes [23] can also influence the efficacy of the PV correction and thus may lead to bias and additional variance in estimating mean GM CBF in data from multi-center studies. However, the magnitude of the errors stemming from differences in processing strategies and readout schemes has not yet been systematically quantified in either PV corrected or PV uncorrected data.

In this study, we have addressed these issues by estimating the errors in PV maps due to geometric distortion and resolution mismatch between ASL and $T1_w$ images. To test for the effect of ASL readout on the magnitude of these errors, two of the commonly used ASL readout schemes – 2D EPI and 3D spiral – were analyzed. The goal was to explore the effect of these errors on the computation of mean GM CBF using different methods for PV correction. Simulated CBF images, constructed based on real ASL data acquired on twenty-two subjects with the two different readout schemes were analyzed using the PVC method and two different GM thresholding methods. The results of this study provide further insight into utility of PV correction and sources of error in computation of mean GM CBF.

MATERIALS AND METHODS

Imaging

Twenty-two healthy young volunteers (mean age 22.6±2.1 years, 9 men) were scanned twice at two imaging centers operating two different 3T MRI systems: Discovery MR750 (GE Healthcare, USA) and Intera (Philips Healthcare, Best, Netherlands), both equipped with the same 8-channel head coil (InVivo, Gainesville, USA). The study was approved by the local medical research ethics committees of both sites: the Erasmus MC – University Medical Center Rotterdam and the Academic Medical Center, Amsterdam and was conducted according to the Declaration of Helsinki. All participants gave written informed consent and received financial compensation for participation.

Each imaging session included a pseudo-continuous ASL (pCASL) scan and a 1 mm isotropic 3D $T1_w$ structural scan. On the Philips scanner, pCASL was acquired using a 2D gradient-echo single-shot echo planar imaging (EPI) readout with voxel size $3\times3\times7$ mm^3, TE/TR 17 ms/4000 ms, labeling duration 1650 ms, post-labeling delay 1525 ms, number of averages 33, resulting in the total acquisition time 4:33 min. On the GE scanner, a 3D fast spin-echo interleaved stack-of-spirals (3D spiral) readout with 512 points on 8 spirals was used with acquisition resolution $3.8\times3.8\times4$ mm^3, reconstructed by default to $1.9\times1.9\times4$ mm^3 voxel size, TE/TR 10.5 ms/4600 ms, labeling duration 1450 ms, post-labeling delay 1525 ms, number of averages 3, and total acquisition time 4:29 min. A detailed description of the recruitment criteria and imaging protocols can be found in [24].

Image processing

Images were processed using in-house Matlab (MathWorks, Nattick, USA) routines using SPM12 (Wellcome Trust Centre for Neuroimaging, London, UK). Image processing was geared towards estimation of CBF-quantification errors associated with: (1) Differences in geometric deformation between the ASL and T1w images due to the difference in contrast between the two modalities; (2) Differences between the acquisition and effective resolution of the ASL images. In both cases, the quantification errors were estimated independently for the 2D EPI and 3D spiral ASL readout sequences.

Below, we detail the main processing steps, also outlined in Figure 1, and provide the rationale behind each step. Briefly: following the processing of the ASL and $T1_w$ data (step A), the geometric distortions (step B) and effective resolution (step C) were estimated and subsequently applied to estimate their effect on GM CBF calculation (step D).

A. Pre-processing

Each subject's $T1_w$ image was segmented into GM and WM PV maps (denoted pGM and pWM) using CAT12 (C. Gaser, Structural Brain Mapping group, Jena University Hospital, Jena, Germany).

For the 2D EPI data, all control and label pairs were motion corrected and pair-wise subtracted. For the 3D spiral data, the perfusion-weighted images were provided, by default, directly from the scanner. CBF images, referred to as CBF_{real}, were quantified using the single-compartment model for both types of ASL readouts [25].

In addition to the CBF_{real} image, a pseudo CBF (CBF_{pseudo}) image was also computed based on the pGM and pWM maps assuming a GM-WM CBF ratio of 3.4 [10] (Figure 1A). In contrast to the CBF_{real} image, the CBF_{pseudo}

image perfectly matches the $T1_w$ image in terms of geometric deformations and spatial resolution. The CBF_{real} image was upsampled using B-spline interpolation to match the voxel size of the $T1_w$ space so that a direct comparison with CBF_{pseudo} could be made to estimate the deformations and resolution as detailed in sections (B) and (C).

B. Estimation of geometric distortions between the $T1_w$ and ASL images

Three commonly used co-registration algorithms were applied to co-register the ASL and T1w images represented by the CBF_{real} and CBF_{pseudo} images, respectively: Rigid (6-parameters rigid-body transformation), Nonlinear (affine transformation followed by low-dimensional nonlinear transformation using discrete cosine basis functions [26]), or DARTEL (high-dimensional warping with Diffeomorphic Registration Algorithm [27]). For Rigid, mutual information criterion was used. For Nonlinear and DARTEL, sum-of-squares criterion was used while first adjusting the GM and WM CBF in CBF_{pseudo} to minimize the voxel-wise difference. For each subject, six deformation fields were computed, one from each co-registration algorithm – Rigid, Nonlinear, DARTEL – repeated for each ASL readout scheme – 2D EPI and 3D spiral (Figure 1B). We assume that the DARTEL algorithm has sufficient degrees of freedom to fully reflect the geometric distortions between the ASL and T1w images [28], and was thus used as a comparative reference.

C. Estimation of the difference between the effective spatial resolution of the ASL and T1w images

The difference in spatial resolution between the ASL and T1w images needs to be accounted to avoid errors in GM CBF analysis. For that purpose, the true resolution of the ASL images needs to be used instead of the ASL ac-

quisition resolution or the reconstructed voxel size. The reconstructed voxel size does not need to match the resolution, and often in the case of 3D acquisitions, the acquisition resolution and the true resolution differ because of substantial smoothing effects of the acquisition and may lead to GM CBF underestimation.

Here, we estimated the effective spatial resolution of each ASL readout by assuming a Gaussian point-spread function (PSF) and by quantifying the resolution difference between T1w and ASL images, represented by CBF_{pseudo} and CBF_{real} images, respectively, that were co-registered using the Nonlinear method, which reduces the influence of mis-registration (step B). Resolution of the $T1_w$ images was assumed to be 1 mm^3 isotropic with Gaussian PSF. First, the spatially varying GM and WM CBF was estimated using PV correction through linear regression to account for perfusion non-uniformity. Then, the estimated values were used to generate a CBF_{pseudo} image. Lastly, the effective resolution of ASL images was estimated by iteratively smoothing of the CBF_{pseudo} image with a varying anisotropic Gaussian kernel. Downhill simplex method was used to find the minimum root mean square deviation between the CBF_{real} and CBF_{pseudo}. Three iterations of PV correction and iterative smoothing were used.

D. Simulation of CBF images and PV maps

The goal of this step was to simulate a reference CBF image and PV maps that contained different amount of geometric deformations (step B, Figure 1) and with and without estimating the effective resolution (step C, Figure 1) in order to evaluate the errors associated with each. To this end, we applied the three estimated deformation fields obtained from Rigid/Nonlinear/DARTEL transformation (Figure 1B), to the original pGM and pWM maps (Figure 1A). The resulting pGM and pWM maps were then downsampled to the original

low-resolution ASL space. The image matrix of the ASL image was preserved. The downsampled pGM and pWM values were obtained by integrating in the high-resolution space and assuming a Gaussian PSF with either the acquisition resolution of the ASL images (referred to as *acq*) or the estimated effective resolution obtained in step C (referred to as *eff*).

Six sets of pGM and pWM images were thus produced for each subject and each ASL readout: Rigid-acq, Rigid-eff, Nonlinear-acq, Nonlinear-eff, DARTEL-acq, and DARTEL-eff. CBF image was simulated for each subject using the PV-maps from DARTEL-eff ($\mathrm{CBF}_{\mathrm{DARTEL-eff}}$). Each type of investigated error was produced by the respective difference between these six PV-maps when used for calculating GM CBF from the reference CBF image ($\mathrm{CBF}_{\mathrm{DARTEL-eff}}$). Specifically: the DARTEL-eff set was used as the reference assuming that DARTEL-eff and DARTEL-acq PV maps reflect purely the difference in acquired and effective resolution; the Rigid-eff and Nonlinear-eff contained residual geometric distortions due to suboptimal $\mathrm{T1_w}$ to ASL co-registration not fully encompassing the differences between the two sequences; the Rigid-acq and Nonlinear-acq contained both geometric distortions and resolution mismatches.

For the quantitative analyses of error propagation, Gaussian noise with a standard deviation of 4 mL/100g/min was added to the $\mathrm{CBF}_{\mathrm{DARTEL-eff}}$ images [29]. For all simulations, a uniform GM CBF of 80 mL/100g/min and GM CBF/WM CBF ratio of 3.4 was assumed [10].

Quantitative evaluation of errors

The mean GM CBF of the $\mathrm{CBF}_{\mathrm{DARTEL-eff}}$ images was obtained by thresholding the low-resolution pGM maps at threshold levels of 10, 20, 30, ..., 90, and 95 %. Three methods for calculating the GM CBF were compared:

- **GM-Threshold** – obtained as mean CBF within the GM region of interest (GM-ROI);
- **GM-Weighted** – the mean CBF within the GM-ROI divided by the mean pGM within the same GM-ROI, i.e. weighted mean;
- **partial volume correction (PVC)** – the mean of PV-corrected GM CBF (within the GM-ROI) computed as per the Asllani method using a $5 \times 5 \times 1$ kernel [10].

To ensure that the results were not dependent on the absolute quantification of CBF, the obtained GM CBF value was divided by 80 mL/100g/min, the preset GM CBF value of the simulated data. A mean relative error in GM CBF calculation was calculated as the average error across all participants.

RESULTS

The mismatch between effective and acquisition resolutions for both the 2D EPI and 3D spiral ASL images is qualitatively shown in Figure 2. For the 2D EPI readout, the mismatch was mainly in the L-R and A-P directions with effective resolution FWHM, mean (s.d.) 3.0 (0.15) $\times$ 3.2 (0.16) $\times$ 7.4 (0.35) mm^3, on average around 10% larger as compared to the acquisition resolution of $3 \times 3 \times 7$ mm^3. The difference was more pronounced for the 3D spiral readout for which the the effective resolution across the whole population was 4.6 (0.20) $\times$ 4.3 (0.37) $\times$ 11.8 (0.39) mm^3 versus the acquisition resolution of $3.8 \times 3.8 \times 4$ mm^3, representing on average a 200% mismatch in the slice-encoding direction.

The mean relative errors in GM CBF calculated using the three different methods – GM-Threshold, GM-weighted, and PVC – are shown in Figure 3 for all pGM thresholds. Results for the commonly-used pGM threshold of 70% are also summarized in Table 1. We picked a relatively high pGM threshold to ensure a better validation with the PVC method even in elderly populations.

For the calculation based on PV maps without errors (DARTEL-eff), as expected, the PVC method fully extracted the GM CBF signal, whereas the standard methods, GM-Threshold and GM-Weighted, scaled with the partial volume effects (Figure 3) for both the 2D EPI and 3D spiral readouts. For GM-ROI with 70% threshold (GM-ROI$_{70\%}$), the GM-Threshold method underestimated GM CBF by 16.5% whereas the GM-Weighted method overestimated it by more than 5.3%.

For the 2D EPI readout, the largest sources of errors were due to geometric deformations typical of the 2D EPI sequence. The error in PVC results increased by 3% when the Nonlinear registration was used to model the deformations, and by 8% with Rigid registration.

For the 3D spiral acquisition, the resolution mismatch was the main source of additional error: 6.0% and 12.5% in CBF images computed with the GM-Threshold and PVC methods, respectively. For the CBF computed with the GM-Weighted, a 6.1% overestimation turned to a 6.8% underestimation of GM CBF.

Independently, errors due to the geometric distortions and resolution mismatch were 16.5% and 18.0% lower for the PVC method than the GM-Threshold method for the 2D EPI and 3D spiral readout, respectively. For the combined geometric distortion and resolution mismatch effects, the PVC method yielded 13.2% and 10.3% lower error than GM-Threshold, respectively. The reason for these effects adding less error for GM-Threshold than PVC is the already high absolute error of GM-Threshold for optimal PV maps. While PVC had 5.3% and 6.0% lower error that the GM-Weighted method for the 2D EPI and the 3D spiral readout, respectively, without assuming the combined effects of distortion and resolution. With these effects, PVC had 5.3% and 6.2% higher error, respectively, than the GM-Weighted method.

The GM values in each of the simulated GM maps were compared with the reference DARTEL-eff GM map, see the joint histograms in Figure 4. For the 2D EPI, a large variance around the line of identity is visible for the Rigid and Nonlinear registration. For the 3D spiral, a large deviation from the line of identity is visible for the acquisition resolution. These results indicate that the largest source of error for 3D spiral readout are related to the resolution mismatch whereas, for the 2D EPI, the geometric deformations are the main source of errors.

DISCUSSION

We measured the effects of two sources of errors – geometric distortion and resolution mismatch – in computing GM CBF using ASL images. These errors are inherent to the current ASL methods, which are based on the extraction of PV maps from the segmentation of the high-resolution T1w images. Geometric distortions result mainly from the difference in readout between ASL and T1w whereas resolution mismatch is due to the intrinsic difference in spatial resolution between the two modalities. The effects of these errors were estimated for three different PV-correction methods: GM-Threshold, GM-Weighted and PVC, and two different ASL readout sequences: 2D EPI and 3D spiral. While the presence of these effects is generally acknowledged [18,30,20,19,22], we have now provided empirical data on how they affect PV maps and consequently the GM CBF calculation.

The key result of this study was that errors associated with resolution mismatch and geometric deformation had similar effects on the mean GM CBF independent of whether PV correction was used. This result, combined with the fact that PVC outperformed the other two PV-correction methods, should pave the way for a broader application of PVC in clinical settings. However,

PV maps can contain additional errors, such as those resulting from inadequate segmentation of the T1w image. In this case, the use of linear regression is not optimal as it assumes that the PV maps (the independent variables) are error free. Using a method such as total least squares, which accounts for uncertainties in the independent variables, might thus be an effective alternative. Another drawback of the PVC method is the use of a spatial kernel which inherently leads to blurring of the PV-corrected CBF images. The blurring may be especially problematic in multi-center studies as equal level of smoothing might be difficult to achieve for sequences with different resolution and matrix size. Performing the PV correction on a global basis and controlling for the smoothness of the resulting corrected CBF maps, for example by modeling the results using B-spline functions should be a subject of further research.

A novel approach for estimating the PV maps directly in the ASL native space using fractional signal modeling was recently introduced by Petr et al. and Ahlgren et al. [19,20]. Better agreement between the PV and CBF maps was demonstrated in both studies as compared to the PV maps estimated from the high-resolution T1w images. This improvement is most likely due to the use of the same readout for PV estimation as for the ASL acquisition. Such PV maps can potentially lead to improved PV-correction as they do not suffer from the detrimental effects of the resolution and deformation issues presented in this study. This statement, however, needs to be further validated.

We have shown that while the PVC method for computation of GM CBF was the one most affected by the geometric deformations (in 2D EPI ASL) and resolution discrepancies (in 3D spiral ASL), it nevertheless outperformed the commonly used GM-Threshold method irrespective of the PV-map quality. With error-free PV maps, PVC outperforms the GM-Weighted method. However, the GM CBF overestimation inherent to GM-Weighted method cancels out in the presence of PV-map errors due to deformation and resolution issues.

The GM-Weighted method thus can have a smaller error than PVC method on certain conditions. Nevertheless, the GM-weighted method is still more dependent on the tissue volume than PVC, and a general advantage of the GM-weighted method over PVC is not warranted for regional and voxel-wise studies.

Although the resolution mismatch was relatively small for the 2D EPI acquisition, our findings show that geometric deformations due to EPI susceptibility artifacts can cause around 8% underestimation in GM CBF when the rigid-body transformation is used to co-register across the modalities. Using SPMs nonlinear co-registration [26] reduced this effect by 5%. In contrast, geometric distortion had negligible effects on the multi-segment 3D spiral readout, which was, however, significantly affected by the resolution mismatch between the effective and acquisition ASL resolution. The effective resolution we estimated was about 200% higher in the slice-encoding direction. This can be explained by a convolution effect of acquisition point-spread-function, motion, and smoothing performed during the scanner's reconstruction phase. Using the expected acquisition resolution instead of the estimated effective resolution can underestimate the mean CBF significantly regardless of PV-correction. Similar effect was shown previously by Zhao et al. and Petr et al., however, the results are not directly comparable as a simulated smoothing with an isotropic Gaussian kernel was used [18,22]. Here, we showed that downsampling the PV maps to the estimated effective resolution before the mean GM CBF calculation can solve this issue. However, to achieve this , the effective resolution needs to be estimated either as performed in the current study or through using literature values, provided that a similar acquisition method was used. The mean GM CBF increased on average 8 % when CBF images were "deblurred". Compensating for the resolution mismatch by deconvolution of CBF data rather than by correctly downsampling the PV maps is likely to be more

prone to noise and computationally demanding. On the other hand, the results of the PVC method are not affected by the mismatch and will be similar when a regional mean CBF is evaluated. However, deconvolution might be still preferable to increase the visual quality of the images and to improve the localization of the perfusion signal in voxel-wise analysis. Deblurring by deconvolution was used by Chappell et al. and by Boscolo Galazzo et al. for 3D GRASE datasets [16,31]. However, these datasets all contained time-series (i.e. multiple post-labeling delay or several control/label repetitions) and therefore the point-spread-function could be estimated from noise autocorrelation.

Another limitation of our study is that here we focused on only two ASL readout schemes. The GM CBF errors might add up differently for other ASL readout types or even different acquisition parameters. We regret the lack of 3D GRASE data, which is another frequently used ASL acquisition. However, we anticipate that the geometric distortion and effective resolution of 3D GRASE in most cases will lie between those for 2D EPI and 3D spiral [32].

Also, while only global GM CBF was evaluated, it can be hypothesized that similar magnitude of error can be expected also in voxelwise analysis. This, however, remains to be verified in the future.

There were several differences between the two acquisitions used in this study that may have added to the observed discrepancy in the GM CBF calculation between the acquisitions. Because no M0 image was obtained for the 2D EPI sequence, B1 inhomogeneities may be present in the 2D EPI CBF images that were not present in the 3D spiral data. We expect that this difference does not influence the results since simulated CBF maps were used. Motion correction could not be performed for 3D spiral data, however, only very little motion was observed in the 2D EPI data. Hence, considering that both 2D EPI and 3D spiral data were acquired on the same participants, we expect the motion effect to be negligible.

Inter-vendor differences in $T1_w$ segmentation can also affect the PV correction, however the assessment of these effects was outside the scope of the current study. Here, we used $T1_w$ images from a single vendor to exclude this potential source of variability.

Another drawback of this study was the choice of the DARTEL transformation as reference for ASL-T1w registration. We did not compare this with acquiring ancillary parameters, such as a B0 field or repeating the acquisition with reversed phase-encoding directions [33]. While DARTEL is typically not used for ASL-T1w registration, it has been previously shown that non-linear registration performs well for EPI images [34], and can even outperform geometric distortion correction using the B0 field [28]. This was the basis of our assumption for using DARTEL in lieu of ground truth in this study where participants were young healthy volunteers. This assumption might not hold in diseased populations, due to poorer SNR and the presence of ASL image artifacts, and precludes the use of DARTEL in clinical studies to actually obtain the deformation field.

For the resolution estimation, an anisotropic Gaussian point-spread function was assumed. For the 3D spiral acquisition, the estimated effective resolution is slightly lower (11.8 mm vs 8 mm FWHM in slice-encoding direction) than in a previous study by Oliver at al. [30] and in good agreement with a study by Vidorreta et al. ($4.6 \times 4.3 \times 11.8$ mm^3 vs $4.64 \times 4.64 \times 9.04$ mm^3 FWHM) [4]. For the 2D EPI, the AFNI toolbox estimated the resolution from noise autocorrelation to be $3.3 \times 3.6 \times 6.2$ mm^3 vs our $3.0 \times 3.2 \times 7.4$ mm^3 resolution matching method (under 2% difference in voxel volume). Both these differences could be explained by the additive effect of several factors including motion, vessel artifacts, suboptimal co-registration, and segmentation errors. Specifically, the values estimated here for 2D EPI probably include (in comparison to AFNI results) segmentation errors. And the 3D spiral values

contain, in comparison to values by Oliver et al. [30] and Vidorreta et al. [4] obtained for an optimal acquisition, other factors as motion [35], or smoothing during the reconstruction. This discrepancy between the anticipated and actual estimated effective resolution emphasized the need for the effective resolution to be estimated for each study independently in order to ensure that PVC is performed on the correct resolution.

CONCLUSION

Careful minimization of geometric distortion and effective resolution differences between the ASL and T1w images significantly reduces errors in GM CBF calculation, independent of PV-correction method used, including no correcting for PV effects at all. Even without the minimization of these differences, PVC remains the most accurate way to calculate GM CBF. Therefore, errors in PV maps caused by geometric deformations and resolution issues present no argument for excluding PVC from the analysis of ASL data. Furthermore, the quantification errors induced by the ASL-T1w differences have a similar magnitude as the CBF effects in many clinical studies. Therefore, minimizing the errors resulting from the ASL-T1w differences are key to increasing the statistical power of ASL imaging and in being able to separate perfusion from structural effects in clinical studies.

AUTHORS' CONTRIBUTION

Study conception and design: JP, HJMMM, IA, JVDH. Acquisition of data: RMES, MS, AJN, Analysis and interpretation of data: JP, HJMMM, FH, IA. Drafting of manuscript: JP, HJMMM, IA. Critical revision JP, HJMMM, IA, EDV, RMES, MS, AJN, FH, JVDH.

Compliance with Ethical Standards

Conflicts of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Research involving Human Participants and/or Animals

Informed consent Informed consent was obtained from all individual participants included in the study.

ACKNOWLEDGMENTS

We thank Matthias J.P. van Osch for his insightful thoughts and comments. The authors wish to thank the COST-AID(European Cooperation in Science and Technology–Arterial spin labeling Initiative in Dementia) Action BM1103 for facilitating meetings for researchers to discuss various application of ASL. Part of this work was supported by The Gleason Foundation (IA).

REFERENCES

References

1. J.C. Ferré, J. Petr, E. Bannier, C. Barillot, J.Y. Gauvrit, Improving quality of arterial spin labeling MR imaging at 3 Tesla with a 32-channel coil and parallel imaging., Journal of magnetic resonance imaging **35**(5), 1233 (2012). DOI 10.1002/jmri.23586
2. J.A. Detre, J.S. Leigh, D.S. Williams, A.P. Koretsky, Perfusion imaging, Magn Reson Med **23**, 37 (1992)

3. M. Günther, K. Oshio, D.A. Feinberg, Single-shot 3D imaging techniques improve arterial spin labeling perfusion measurements., Magn Reson Med **54**(2), 491 (2005). DOI 10.1002/mrm.20580
4. M. Vidorreta, E. Balteau, Z. Wang, E. De Vita, M.A. Pastor, D.L. Thomas, J.A. Detre, M.A. Fernández-Seara, Evaluation of segmented 3D acquisition schemes for whole-brain high-resolution arterial spin labeling at 3 T., Nuclear Magn Reson in Biomed **27**(11), 1387 (2014). DOI 10.1002/nbm.3201
5. F.Q. Ye, J.a. Frank, D.R. Weinberger, a.C. McLaughlin, Noise reduction in 3D perfusion imaging by attenuating the static signal in arterial spin tagging (ASSIST)., Magn Reson Med **44**(1), 92 (2000)
6. D.M. Garcia, G. Duhamel, D.C. Alsop, Efficiency of inversion pulses for background suppressed arterial spin labeling., Magn Reson Med **54**(2), 366 (2005). DOI 10.1002/mrm.20556
7. W. Dai, D. Garcia, C. de Bazelaire, D.C. Alsop, Continuous flow-driven inversion for arterial spin labeling using pulsed radio frequency and gradient fields., Magn Reson Med **60**(6), 1488 (2008)
8. N.A. Telischak, J.A. Detre, G. Zaharchuk, Arterial spin labeling MRI: Clinical applications in the brain, J Magn Reson Imaging **41**(5), 1165 (2015). DOI 10.1002/jmri.24751
9. S. Haller, G. Zaharchuk, D.L. Thomas, K.O. Lovblad, F. Barkhof, X. Golay, Arterial Spin Labeling Perfusion of the Brain: Emerging Clinical Applications, Radiology **281**(2), 337 (2016). DOI 10.1148/radiol.2016150789
10. I. Asllani, A. Borogovac, T.R. Brown, Regression algorithm correcting for partial volume effects in arterial spin labeling MRI., Magn Reson Med **60**(6), 1362 (2008). DOI 10.1002/mrm.21670
11. K. Zhang, H. Herzog, J. Mauler, C. Filss, T.W. Okell, E.R. Kops, L. Tellmann, T. Fischer, B. Brocke, W. Sturm, H.H. Coenen, N.J. Shah, Comparison of cerebral blood flow acquired by simultaneous [15O]water positron emission tomography and arterial spin labeling magnetic resonance imaging., J Cerebr Blood F Met **34**(8), 1373 (2014). DOI 10.1038/jcbfm.2014.92
12. I. Asllani, C. Habeck, A. Borogovac, T.R. Brown, A.M. Brickman, Y. Stern, Separating function from structure in perfusion imaging of the aging brain., Human Brain Mapping **30**(9), 2927 (2009). DOI 10.1002/hbm.20719
13. E.E. Bron, R.M.E. Steketee, G.C. Houston, R.a. Oliver, H.C. Achterberg, M. Loog, J.C. van Swieten, A. Hammers, W.J. Niessen, M. Smits, S. Klein, Diagnostic classification

of arterial spin labeling and structural MRI in presenile early stage dementia., Human Brain Mapping **35**(9), 4916 (2014). DOI 10.1002/hbm.22522

14. R.M.E. Steketee, E.E. Bron, R. Meijboom, G.C. Houston, S. Klein, H.J.M.M. Mutsaerts, C.P. Mendez Orellana, F.J. de Jong, J.C. van Swieten, A. van der Lugt, M. Smits, Early-stage differentiation between presenile Alzheimer's disease and frontotemporal dementia using arterial spin labeling MRI, European Radiology **26**(1), 244 (2016). DOI 10.1007/s00330-015-3789-x
15. J.A. Aston, V.J. Cunningham, M.C. Asselin, A. Hammers, A.C. Evans, R.N. Gunn, Positron emission tomography partial volume correction: estimation and algorithms, Journal of Cerebral Blood Flow & Metabolism **22**(8), 1019 (2002)
16. M.A. Chappell, A.R. Groves, B.J. MacIntosh, M.J. Donahue, P. Jezzard, M.W. Woolrich, Partial volume correction of multiple inversion time arterial spin labeling MRI data., Magn Reson Med **65**(4), 1173 (2011). DOI 10.1002/mrm.22641
17. X. Liang, A. Connelly, F. Calamante, Improved partial volume correction for single inversion time arterial spin labeling data., Magn Reson Med **69**(2), 531 (2013). DOI 10.1002/mrm.24279
18. M.Y. Zhao, M. Mezue, A.R. Segerdahl, T.W. Okell, I. Tracey, Y. Xiao, M.A. Chappell, A systematic study of the sensitivity of partial volume correction methods for the quantification of perfusion from pseudo-continuous arterial spin labeling MRI, NeuroImage **162**, 384 (2017)
19. J. Petr, G. Schramm, F. Hofheinz, J. Langner, J. van den Hoff, Partial volume correction in arterial spin labeling using a Look-Locker sequence, Magn Reson Med **70**(6), 1535 (2013). DOI 10.1002/mrm.24601
20. A. Ahlgren, R. Wirestam, E.T. Petersen, F. Ståhlberg, L. Knutsson, Partial volume correction of brain perfusion estimates using the inherent signal data of time-resolved arterial spin labeling., NMR in biomedicine **27**(9), 1112 (2014). DOI 10.1002/nbm.3164
21. R. Oliver, E. De Vita, X. Golay, D. Thomas, in *Proceedings of the 22nd Joint Annual Meeting of ISMRM-ESMRMB, Milan* (2014), p. 418
22. J. Petr, H.J.M.M. Mutsaerts, E. De Vita, J. Maus, J. van den Hoff, I. Asllani, in *ISMRM '16: Proceedings of the 24th Scientific Meeting and Exhibition of International Society for Magnetic Resonance in Medicine*, vol. 24 (ISMRM, 2016), vol. 24, p. 1496
23. H.J. Mutsaerts, M.J. van Osch, F.O. Zelaya, D.J. Wang, W. Nordhøy, Y. Wang, S. Wastling, M.A. Fernandez-Seara, E. Petersen, F.B. Pizzini, S. Fallatah, J. Hendrikse, O. Geier, M. Günther, X. Golay, A.J. Nederveen, A. Bjørnerud, I.R. Groote, Multi-vendor reliability of arterial spin labeling perfusion MRI using a near-identical

sequence: Implications for multi-center studies, NeuroImage **113**, 143 (2015). DOI 10.1016/j.neuroimage.2015.03.043

24. H.J.M.M. Mutsaerts, R.M.E. Steketee, D.F.R. Heijtel, J.P.A. Kuijer, M.J.P. van Osch, C.B.L.M. Majoie, M. Smits, A.J. Nederveen, Inter-vendor reproducibility of pseudo-continuous arterial spin labeling at 3 tesla., PloS one **9**(8), e104108 (2014). DOI 10.1371/journal.pone.0104108
25. D.C. Alsop, J.A. Detre, X. Golay, M. Günther, J. Hendrikse, L. Hernandez-Garcia, H. Lu, B.J. MacIntosh, L.M. Parkes, M. Smits, M.J.P. van Osch, D.J.J. Wang, E.C. Wong, G. Zaharchuk, 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 **73**(1), 102 (2015)
26. J. Ashburner, K.J. Friston, et al., Nonlinear spatial normalization using basis functions, Human brain mapping **7**(4), 254 (1999)
27. J. Ashburner, A fast diffeomorphic image registration algorithm, Neuroimage **38**(1), 95 (2007)
28. H. Huang, C. Ceritoglu, X. Li, A. Qiu, M.I. Miller, P.C. van Zijl, S. Mori, Correction of b0 susceptibility induced distortion in diffusion-weighted images using large-deformation diffeomorphic metric mapping, Magnetic resonance imaging **26**(9), 1294 (2008)
29. Y. Qiu, A. Borogovac, A. Laine, J. Hirsch, I. Asllani, in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (2014), pp. 6687–6690
30. R. Oliver, L. Ly, C. Wang, H. Beadnall, I. Boscolo Galazzo, M. Chappell, X. Golay, E. De Vita, D. Thomas, M. Barnett, in *Proceedings of the 24th Annual Meeting of ISMRM, Singapore* (2016), p. 1899
31. I. Boscolo Galazzo, M.A. Chappell, D.L. Thomas, X. Golay, P. Manganotti, E. De Vita, in *Proceedings of the 22th Joint Annual Meeting of ISMRM-ESMRMB, Milan* (2014), p. 2704
32. H.J. Mutsaerts, J. Petr, D.L. Thomas, E.d. Vita, D.M. Cash, M.J.v. Osch, X. Golay, P.F. Groot, S. Ourselin, J.v. Swieten, R. Laforce, F. Tagliavini, B. Borroni, D. Galimberti, J.B. Rowe, C. Graff, F.B. Pizzini, E. Finger, S. Sorbi, M. Castelo Branco, J.D. Rohrer, M. Masellis, B.J. MacIntosh, on behalf of the GENFI investigators, Comparison of arterial spin labeling registration strategies in the multi-center genetic frontotemporal dementia initiative (genfi), Journal of Magnetic Resonance Imaging DOI 10.1002/jmri.25751

33. J.L. Andersson, S. Skare, J. Ashburner, How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging, Neuroimage **20**(2), 870 (2003)
34. J. Kybic, P. Thévenaz, A. Nirkko, M. Unser, Unwarping of unidirectionally distorted epi images, IEEE transactions on medical imaging **19**(2), 80 (2000)
35. J. Petr, H. Mutsaerts, E. De Vita, Z. Shirzadi, S. Cohen, C. Blokhuis, D. Pajkrt, F. Hofheinz, J. van den Hoff, I. Asllani, in *Proceedings of the 33rd Annual Meeting of ESMRMB, Vienna, Austria* (2016), p. 1

FIGURE LEGENDS

Fig. 1: Diagram of image processing pipeline For each column, the input images (top row), the processing steps (middle row), and the output images (bottom row) are displayed. (A) **Preprocessing:** A CBF_{real} was computed and upsampled to the image matrix of the $T1_w$ image; A CBF_{pseudo} was also created using pGM and pWM tissue maps from segmentation of the $T1_w$ image. (B) **Estimation of geometric distortions** between CBF_{real} and CBF_{pseudo} was done using Rigid, Nonlinear or DARTEL. (C) **Estimation of the effective resolution** of ASL images was done via iterative resolution matching of CBF_{real} and CBF_{pseudo} images. (D) **Simulation of CBF images** containing deformations introduced in steps (B) and (C) was done by applying the deformations to the original pGM, and pWM maps (A) and recomputing the CBF image based on these maps.

Fig. 2: A CBF_{real} image and pGM maps co-registered to the ASL image using the three different transformations (Rigid-eff/Nonlinear-eff/DARTEL-eff) and two different resolutions (DARTEL-acq/DARTEL-eff) are shown for the 2D EPI (top row) and 3D spiral (bottom row) ASL readouts. DARTEL-eff should be the closest match of the pGM maps to the CBF_{real} image in terms of both resolution and deformations. The difference between PV maps containing deformations (Rigid-eff, Nonlinear-eff) were more pronounced in the 2D EPI images. White lines help to notice prolongation artifact in the phase-encoding direction typical for EPI acquisition. The 3D spiral sequence had only minor deformations and therefore the PV maps (Rigid-eff/Nonlinear-eff/DARTEL-eff) were similar for all co-registrations. Visible difference in resolution can be seen between the 3D spiral images DATEL-eff and DARTEL-acq and also between the DARTEL and the CBF_{real} image.

Fig. 3: Mean relative error in GM CBF calculation is shown for the 2D EPI (top row) and 3D spiral (bottom row) ASL sequences. The columns represent cases when the PV maps used for calculation contained (left to right) no errors (DARTEL-eff PV maps), deformation errors with non-linear coregistration (Nonlinear-eff), deformation errors (Rigid-eff), resolution errors (DARTEL-acq), resolution and deformation errors with non-linear coregistration (Nonlinear-acq), and both deformation and resolution errors (Rigid-acq). Note that without errors in the PV maps, the GM-Threshold method underestimates the CBF and the GM-Weighted overestimates the CBF. Resolution and deformation errors both cause underestimation of the GM CBF.

Fig. 4: Joint histograms of 2D EPI and 3D Spiral readout schemes for a randomly selected subject. The line of identity, corresponding to a perfect match of the PV volumes, is indicated in white.

TABLES

Type	PV type	DARTEL eff	Nonlinear eff	Rigid eff	DARTEL acq	Nonlinear acq	Rigid acq
2D EPI	GM-Threshold	-16.5	-19.3	-21.6	-16.9	-19.6	-21.9
	GM-Weighted	5.3	1.2	-2.6	4.4	0.4	-3.4
	PVC	< 0.01	-3.0	-8.0	-0.9	-3.9	-8.7
3D spiral	GM-Threshold	-18.1	-18.7	-20.0	-24.1	-24.6	-25.7
	GM-Weighted	6.1	5.5	3.5	-6.8	-7.4	-9.2
	PVC	< 0.01	-0.3	-4.5	-12.6	-12.9	-15.4

Table 1: Mean relative error in GM CBF calculation for the 2D EPI and 3D spiral readouts is shown in [%] relative to the preset reference GM CBF value of 80 mL/min/100 g. Results are shown for GM-ROI$_{70\%}$. Results for all other threshold levels are shown in Figure 3.