



Partial volume correction in arterial spin labeling perfusion MRI: A method to disentangle anatomy from physiology or an analysis step too far?

Michael A. Chappell^{a,b,c,d,*}, Flora A. Kennedy McConnell^{a,b,c}, Xavier Golay^e,
Matthias Günther^{f,g,h}, Juan A. Hernandez-Tamamesⁱ, Matthias J. van Osch^j, Iris Asllani^{k,l}

^a Radiological Sciences, Mental Health and Clinical Neurosciences, School of Medicine, University of Nottingham, Nottingham, UK

^b Sir Peter Mansfield Imaging Center, School of Medicine, University of Nottingham, Nottingham, UK

^c Nottingham Biomedical Research Centre, Queens Medical Centre, University of Nottingham, Nottingham, UK

^d Wellcome Centre for Integrative Neuroimaging, Nuffield Department of Clinical Neurosciences, University of Oxford, UK

^e Department of Brain Repair and Rehabilitation, UCL Institute of Neurology, London, UK

^f Fraunhofer MEVIS, Bremen, Germany

^g University Bremen, Bremen, Germany

^h mediri GmbH, Heidelberg, Germany

ⁱ Radiology and Nuclear Medicine Department, Erasmus MC, Rotterdam, the Netherlands

^j C.J. Gorter Center for High Field MRI, Radiology Department, Leiden University Medical Center, Leiden, the Netherlands

^k Clinical Imaging Sciences Centre, Department of Neuroscience, University of Sussex, UK

^l Department of Biomedical Engineering, Rochester Institute of Technology, Rochester, NY, United States

ARTICLE INFO

Keywords:

Arterial spin labeling
Partial volume effect
Perfusion
Cerebral Blood Flow
MRI
fMRI
Atrophy
Aging and dementia

ABSTRACT

The mismatch in the spatial resolution of Arterial Spin Labeling (ASL) MRI perfusion images and the anatomy of functionally distinct tissues in the brain leads to a partial volume effect (PVE), which in turn confounds the estimation of perfusion into a specific tissue of interest such as gray or white matter. This confound occurs because the image voxels contain a mixture of tissues with disparate perfusion properties, leading to estimated perfusion values that reflect primarily the volume proportions of tissues in the voxel rather than the perfusion of any particular tissue of interest within that volume. It is already recognized that PVE influences studies of brain perfusion, and that its effect might be even more evident in studies where changes in perfusion are co-incident with alterations in brain structure, such as studies involving a comparison between an atrophic patient population vs control subjects, or studies comparing subjects over a wide range of ages. However, the application of PVE correction (PVEc) is currently limited and the employed methodologies remain inconsistent.

In this article, we outline the influence of PVE in ASL measurements of perfusion, explain the main principles of PVEc, and provide a critique of the current state of the art for the use of such methods. Furthermore, we examine the current use of PVEc in perfusion studies and whether there is evidence to support its wider adoption.

We conclude that there is sound theoretical motivation for the use of PVEc alongside conventional, ‘uncorrected’, images, and encourage such combined reporting. Methods for PVEc are now available within standard neuroimaging toolboxes, which makes our recommendation straightforward to implement. However, there is still more work to be done to establish the value of PVEc as well as the efficacy and robustness of existing PVEc methods.

1. Introduction

The Partial Volume Effect (PVE) is a widely recognized phenomenon in medical imaging, reflecting the fact that the spatial resolutions of the images are lower than the variations in the anatomy of the tissues they seek to capture. For example, even ‘high’ resolution anatomical MRI, with voxels sizes smaller than 1 mm isotropic, will contain many voxels which lie across a boundary between tissues identified as gray matter

(GM), white matter (WM), and cerebral spinal fluid (CSF). Whilst PVE is widely acknowledged, its influence on MRI techniques that seek to derive quantitative measurements of physiological processes in the brain, such as perfusion, is often neglected. This article considers the effect of partial voluming on quantitative perfusion measurements using Arterial Spin Labeling (ASL) MRI and whether correction for this effect is warranted given the algorithms currently available. Many of these arguments are shared with other perfusion modalities, such as dynamic

* Corresponding author at: University of Nottingham, Beacon Hub Offices, A floor, Medical School, QMC, Nottingham, NG7 2UH, UK.

E-mail address: michael.chappell@nottingham.ac.uk (M.A. Chappell).

<https://doi.org/10.1016/j.neuroimage.2021.118236>.

Received 19 February 2021; Received in revised form 25 May 2021; Accepted 2 June 2021

Available online 4 June 2021.

1053-8119/© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

susceptibility contrast perfusion MRI, O^{15} - H_2O PET, CT perfusion, and closely related methods like FDG-PET. There are modality specific influences, however, which means that the impact and appropriateness of correction for PVE may be different for each modality. For example, the difference in the width of the point spread function between PET and MRI introduces an additional source of partial voluming in PET applications (Greve et al., 2016).

The aim of this article is firstly to outline the influence of PVE in ASL measurements of tissue-specific perfusion and the implications this has for analysis using ASL perfusion measurements. We then seek to make recommendations for how to control for the effects of PVE, including an overview of the main principles of partial volume correction (PVEc), and a critique of the current state of the art for the use of such methods. Furthermore, we examine the current use of PVEc in perfusion studies and whether there is evidence to support its wider adoption.

2. The sources of partial volume effects in perfusion measurements

Interest in hemodynamic and perfusion imaging has grown significantly in the last decade since changes to perfusion and microvasculature have been found to occur earlier than structural changes in several disease processes (Cada et al., 2000; Iturria-Medina et al., 2016; Jong et al., 1999). As such, perfusion imaging is playing an increasingly important role in understanding and management of diseases such as stroke, cancer, Alzheimer's, Parkinson's, dementia with Lewy bodies, fronto-temporal dementia, and other vascular and neurodegenerative brain diseases (Grade et al., 2015; Telischak et al., 2014). ASL is particularly well suited for clinical research because it is non-invasive and consensus on robust and reliable acquisition is now available (Alsop et al., 2015). This is evidenced by the steady increase in the number of publications employing ASL, including a growing number where it is used to image perfusion in a wide range of clinical applications.¹

The typical definition of perfusion in physiology is the process of delivering blood to the capillary bed in a tissue. Perfusion is important because it relates directly to the function of the vascular system to deliver nutrients to the tissue and remove waste. The seminal work by Kety and Schmidt (1948) introduced the concept of measuring the perfusion of a whole organ and the units of perfusion that are still widely used to this day: ml blood/ 100 g tissue/ min. Subsequent developments have aimed to 'zoom in' and measure not only the total blood flow to an organ, but also the distribution of blood delivery within an organ. Hence, the goal of current perfusion imaging techniques is to provide spatially resolved maps of capillary blood delivery within organs which are subsequently used to quantify perfusion within individually defined regions of interest and/or voxels. However, a voxel is not an anatomically or physiologically meaningful volume and there is no guarantee that perfusion measured in one voxel in one region of an organ in one subject can be meaningfully compared to a different voxel in a different region and/or a different individual. For brain perfusion imaging, it has become conventional to separate cortical gray from white matter perfusion in ASL measurements, reflecting the gross differences in perfusion (and other hemodynamic properties) between these two anatomically different tissues as well as the different roles they play in normal brain function and disease.

ASL is typically imaged with a voxel volume of $\sim 3.5 \times 3.5 \times 5 \text{ mm}^3$, which is more than an order of magnitude lower than that of anatomical MRI. This volume is 'large' when compared to spatial variations of tissue content in the brain. For example, the average thickness of the cortex in young healthy adults has been found to be $\sim 2.5 \text{ mm}$ (Fischl and Dale, 2000). It is, therefore, inevitable that most ASL voxels will contain a mixture of the different constituent tissues: primarily GM, WM and CSF. It is generally accepted that GM has somewhere between 2 and 5

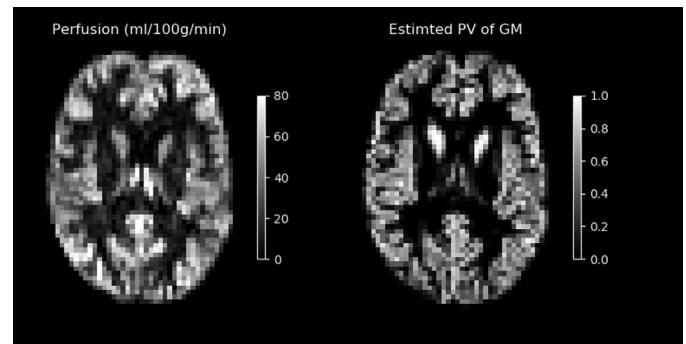


Fig. 1. Comparing an ASL perfusion image (left) with the estimated partial volumes (PV) of gray matter in the same individual on a matched imaging grid (right). Estimated PV of gray matter is obtained by down-sampling a GM segmentation from a 3D T1 anatomical scan to the ASL resolution.

times the perfusion of WM, according to PET studies (Donahue et al., 2006; Huang et al., 1983), whilst CSF is not perfused. The perfusion values measured by ASL in any given voxel will inherently reflect a weighted average of the perfusion of the mixture of tissues present in that voxel. A comparatively modest partial volume mixture of 80% GM and 20% CSF, as might survive most thresholding criteria in a group analysis and thus count as a 'GM voxel', would give rise to an underestimate of perfusion by $\sim 24\%$ in theory (Asllani et al., 2008). Even for healthy young brains, most imaged voxels present a mixture of signals from at least two different tissues. This effect will be more pronounced in aging and diseased brains due to atrophy and tissue damage.

Fig. 1 illustrates the influence of PVE on ASL perfusion images, comparing a perfusion-weighted image to a down-sampled map of GM partial volumes segmented from a T1-weighted anatomical image in the same individual. The measured ASL perfusion image reflects many of the spatial characteristics of the GM PV image, indicating the degree to which structural information directly influences the appearance of the perfusion image. The ASL perfusion image might thus be described as a structural image with perfusion information modulated on top, or that the brain anatomy is 'shining through' the perfusion.

Arguably, PVE in ASL images may not be obviously problematic when images are being visually read for the presence of substantial regions of hyper- or hypperfusion, such as in stroke and brain tumors. In such situation it is sufficient to examine the overall perfusion within voxels. However, PVE could be substantially misleading for quantification of perfusion when examining differences of perfusion between individuals, e.g., in a study population, or when seeking to detect subtle changes in perfusion, such as in neurodegenerative diseases (for example, hippocampal flow in aging and Alzheimer's disease) or diaschisis effects in stroke. In such cases, genuine perfusion differences or changes may be confused for, or masked by, differences or changes in PVE between corresponding voxels in the same anatomical location. Changes in PVE are expected to be exacerbated in aging and dementia populations due to atrophy and other disease-related changes in the brain anatomy.

Whilst various methods have been developed for the correction of PVE (PVEc) in ASL, (Ahlgren et al., 2014; Asllani et al., 2008; Chappell et al., 2011; Johnson et al., 2005; Kandel et al., 2015; Liang et al., 2013; Petr et al., 2013; Wiersma et al., 2006) with varying levels of sophistication and assumptions on perfusion in cerebral tissue (these are reviewed in more depth in 'Correcting for Partial Volume Effects' in the Supplementary Materials), no consensus has emerged on the best method to use, despite comparative studies examining their theoretical and practical performance (Petr et al., 2018; Zhao et al., 2017). And, there remain issues around the estimation of tissue partial volumes used in all correction algorithms mentioned above (for more detail see 'Confounds for PVEc' in the Supplementary Materials).

¹ A pubmed search with 'Arterial Spin Labeling MRI' returns over 300 publications per year since 2015.

Like many functional MRI techniques, ASL resolution is chosen to achieve sufficient SNR by sampling from larger voxel volumes than used in anatomical imaging. PVE will have less of an impact at higher spatial resolution, hence (Donahue et al., 2006) found an increase in GM perfusion of 10–20% in high ($2.5 \times 2.5 \times 3.0 \text{ mm}^3$) resolution ASL that they attributed to a reduction in PVE. Higher resolution ASL data have been made possible by advances in technology including 3D or Simultaneous Multi-Slice (SMS) readouts. For example, the Human Connectome Project Ageing and Development cohorts acquired ASL with voxel dimensions 2.5 mm isotropic in a 5:29 min acquisition using SMS (Harms et al., 2018; Li et al., 2015). Despite this, PVE in such high resolution ASL may still be substantial, especially in older cohorts (Álvarez et al., 2019). ASL performed at higher magnetic field ($>3\text{T}$), exploiting the inherent enhancement in SNR and the further SNR benefits for ASL imaging of longer T1 times potentially, offers higher resolution, e.g., (Álvarez et al., 2019; Bause et al., 2016; Luh et al., 2013; Zuo et al., 2013), and might enable sub-millimeter voxel dimensions albeit with limited brain coverage (Huber et al., 2019). However, the advanced technology needed to achieve high resolution are not widely available and not accessible to many clinical studies.

PET imaging is the other major research field in which PVE has been examined, having similar motivation to ASL investigations due to the low imaging resolution compared to brain anatomy (Erlandsson et al., 2012; Greve et al., 2016; Thomas et al., 2016). PET researchers have likewise investigated correction approaches, some of which have been the basis for or were based upon those used in ASL. However, at this point there has not been widespread adoption of PVEc for PET, partially because availability of high-resolution structural brain images is not as routine for PET as for ASL. Furthermore, the relationship between tissue type and PET tracer signal is potentially more complex than for ASL perfusion, since the influence of the differences in point spread function between the anatomical and PET imaging makes a more substantial contribution to PVE on top of effects arising from mixing different tissue types within a voxel (commonly termed in the PET literature as the Tissue Fraction Effect). The influence of the PSF in ASL PVE is typically ignored, though there are a few studies that have examined or attempted to correct for PSF related effects in ASL (Chappell et al., 2011; Zhao et al., 2017). The combination of tissue mixing and PSF are found in other comparatively low resolution non-cartesian MRI methods, such as ^{23}Na imaging, where PVEc methods have also been used (Kim et al., 2021; Niesporek et al., 2015).

There is a sound theoretical foundation to suggest that accounting for PVE when using ASL will provide greater insight into the underlying physiology beyond what is possible with existing perfusion images. The evidence that it makes a practical difference is, however, currently limited, not least because of the lack of studies employing methods to correct for PVE and a lack of consistency between those that do. We have identified three areas in which PVE in ASL has an impact, and in which some form of correction may be warranted:

- Quantification of perfusion in regions of interest,
- Detecting voxelwise perfusion differences between groups,
- Detecting genuine perfusion changes when accompanied by atrophy.

In the rest of this article, we highlight the impact of PVE in these cases and suggest ways forward for the community when considering the role of managing or correcting for PVE. In doing so we seek to guide the community to a more consistent approach, as well as motivate further awareness and focused research in this area.

3. The impact of PVE in ASL studies

In this section we examine the influence of PVE when ASL is used to examine perfusion within or between groups of individuals in a study, seeking to highlight the specific issues that arise from PVE in common applications of ASL.

3.1. Influence of PVE on quantitative perfusion in regions of interest (ROIs)

In ASL studies, as with many fMRI studies, it is common to calculate perfusion in ROIs, including the calculation within a whole-brain ROI as well as anatomical ROIs defined from an atlas. Quite apart from the usual issues around differences in resolution between the data and the space in which ROIs might be defined, for ASL, PVE has important consequences for *quantitative perfusion measurement from an ROI*, since the tissue content of the ROI will directly affect the apparent perfusion value returned.

The consequences of PVE in ROI-based quantification are:

- systematic under-estimation of cortical GM perfusion due to the mixture of tissues present within the ROI;
- increased variability between individuals due to inherent differences in the proportions of different tissues within the ROI due to individual anatomy and position of the brain relative to the acquisition grid. This will reduce the statistical power to detect differences in perfusion, and will increase the reported between-subject variability beyond that due to genuine physiological variation;
- increased variability between studies due to methodological differences arising from both variation in acquisition, e.g. the resolution used, leading to variation in the magnitude of the intrinsic PVE; and from variation in analysis, e.g. how ROI are defined and whether they are applied in the native space of the ASL data or in template space.

It is common for ASL studies to report whole-brain mean perfusion as a single representative metric from each individual. Since GM is the most highly perfused tissue in the brain, and hence is the easiest within which to detect perfusion using ASL, as well as being the tissue in which perfusion is likely to vary in most functional studies and many pathologies, it is generally accepted that this ROI should be restricted only to the GM and thus the whole-brain GM mean perfusion is reported. We believe it is likely that a measure of whole-brain GM perfusion will become the de facto imaging derived phenotype (Elliott et al., 2018; Miller et al., 2016) from ASL.

With typical ASL resolution it is almost impossible to define a ‘pure’ GM ROI: the number of voxels that do not have some contamination with WM or CSF is very limited. The influence of PVE on whole-brain GM mean perfusion using ASL can be seen in reports from the literature, where typical values are in the range 30–40 ml/100 g/min, e.g. (Petersen et al., 2010), lower than the expected GM perfusion of 50–70 ml/100 g/min. Whilst this consequence of PVE might be consistent within a study population, since an average value is being taken across many voxels each affected by PVE, in practice there will be systematic differences between populations due to anatomy (e.g. due to increasing atrophy with age and pathology). There will also be systematic differences between studies since no one consistent approach is taken to define GM voxels, e.g., differences in methods for segmentation and the source of partial volume estimates. In practice, it is common for the details related to the definition of the GM ROI to be incompletely reported, meaning that it isn’t even possible to establish whether there might be systemic differences between studies.

3.2. The influence of PVE on voxelwise perfusion statistics

Partial volume effects are particularly problematic when voxelwise statistical tests are applied across groups of individuals. Group analysis normally includes the additional step of spatial normalization from subject image space to a template image space, e.g., MNI152. Whilst spatial normalization might be able, under ideal conditions, to align sulci and gyri between individuals and thus match voxels across the group, it cannot overcome PVE present within the original ASL data. The consequences of PVE on detecting voxelwise perfusion differences are:

- in each individual, a different mixture of tissues will have been captured in spatially equivalent voxels, leading to differences in perfu-

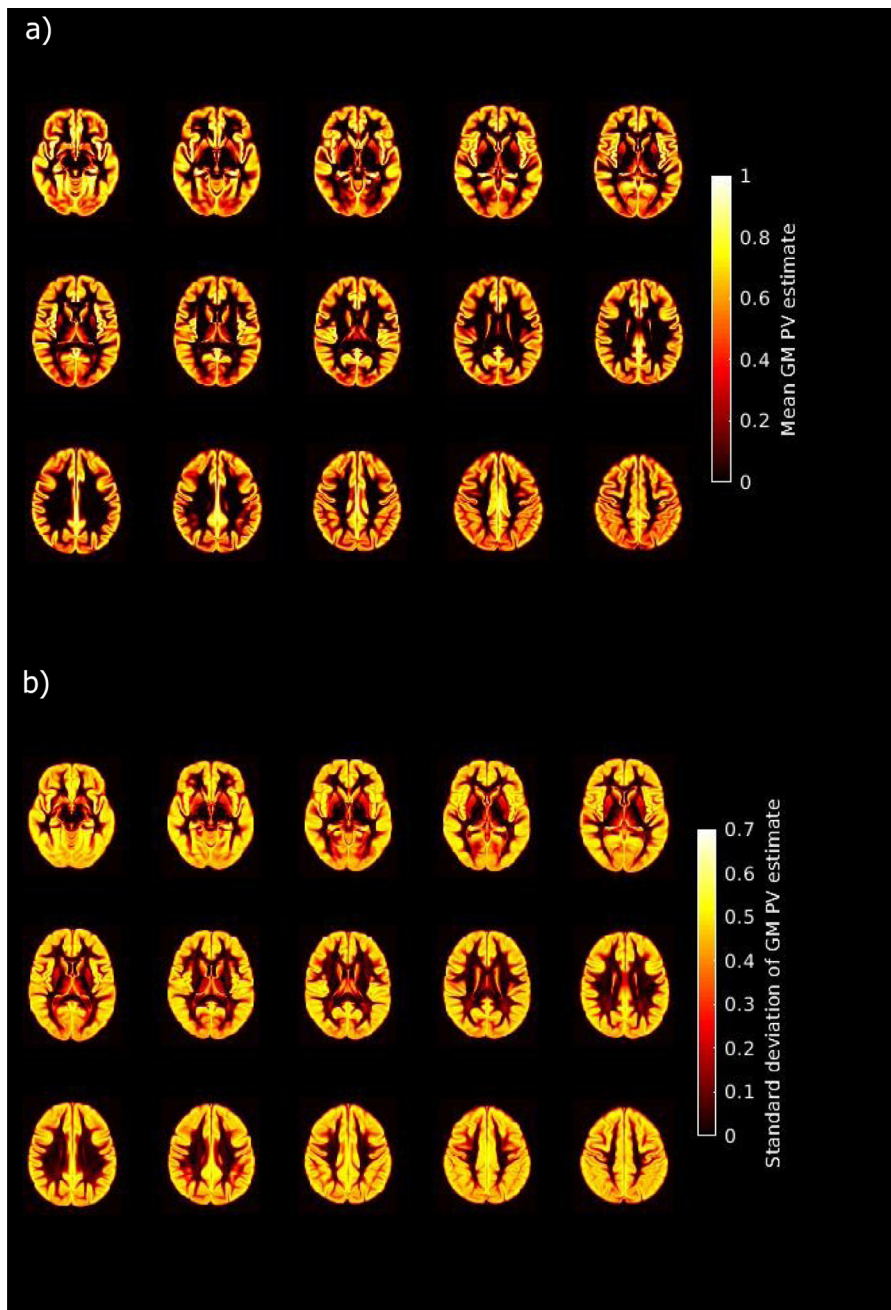


Fig. 2. Variability in brain structure across a group as a source of potential variability in apparent perfusion. The images show a) the mean PV of GM and b) the variability in the PV of GM (expressed as the standard deviation) across a group of 14,503 individuals from the UK Biobank Imaging study. For methodological details see Supplementary Materials.

sion values on an individual level that directly reflect the mixture of tissues present within the measurement voxel, for example in a voxel identified within the GM based on segmentation there will be an underestimation of perfusion related to the contribution of CSF and WM tissue;

- a reduction in statistical power (false negatives) to detect differences due to added variability in apparent perfusion that arise purely from PVE.
- false positive findings due to consistent differences in PVE between groups (e.g. due to sex) thereby indicating statistically significant differences, whereas the perfusion within individual tissues is not different between the groups.

This effect is illustrated in Fig. 2, which shows the variability in the proportion of GM tissue within voxels across 14,503 individuals from the UK Biobank Imaging Study (Miller et al., 2016) when transformed into MNI152 (2 mm) template image space. The standard deviation map in Fig. 2 reflects the variability due to PVE alone, and largely arises

from a mismatch in PVE in the original ASL data due to differences in underlying brain structure (although additional variability will also be introduced by imperfect spatial normalization algorithms (Petr et al., 2013)). This variability, which is over 30% in regions of cortical GM, will translate directly into variability in the apparent perfusion within the voxel and hence additional variability in any statistical analysis and the statistical power available to detect genuine perfusion differences. Separately, differences in PVE between groups occur through variation in anatomy and will give rise to a systemic difference in apparent perfusion in affected brain regions. These have been shown to make a measurable contribution to age-related perfusion changes, for example, accounting for an apparent 10–12% increase in the age-related perfusion difference between men and women (Asllani et al., 2009).

3.3. The influence of PVE when coincident with atrophy

With the use of ASL in clinical applications, it became evident early on that partial volume effect could be a major confound in studies of

aging and dementia (Johnson et al., 2005). This was, in part, due to the parallel discussion in PET literature of the influence of PVE on FDG- and O^{15} -PET studies (Meltzer et al., 2000; Samuraki et al., 2007). Atrophy is present in aging and has been observed in many neurological and psychiatric disorders (Pini et al., 2016). Given the influence of PVE on perfusion images, it is possible to observe apparent perfusion changes between groups that arise from atrophy alone. Moreover, in atrophic regions loss of GM might mask true physiological changes in cortical perfusion.

The consequences of PVE on detecting genuine perfusion changes when accompanied by atrophy are:

- Systematic differences between groups due to underlying differences in GM volumes that are a direct consequence of pathology leading to misinterpretation of hemodynamic alterations from ASL data where structural changes are the true underlying cause of the signal change.
- Masking of potential compensatory increases in perfusion in regions of GM atrophy (Alsop et al., 2008).
- Potential overestimation of hypoperfusion in regions where both perfusion and GM volume are reduced. While this could be helpful for identification of pathological changes, the combined loss of GM and decreased perfusion is typically associated with more severe pathology such as later stages of dementias.

Reductions in cortical volume of $\sim 0.5\%$ per year have been observed in healthy aging, with more pronounced changes in Alzheimer's disease (Fjell et al., 2009), which would directly translate into a similar order of magnitude reduction in apparent perfusion. In practice, there is evidence for a complex relationship between changes in atrophy in perfusion with age and in disease disease (Chen et al., 2011; Chen et al., 2012) potentially reflecting both apparent and genuine changes in perfusion. In a clinical diagnostic context, the apparent enhancement of pathological regions due to atrophy combined with hypoperfusion might be helpful to increase the sensitivity for identification of affected regions as a perfusion-based biomarker for neurodegeneration. From both a clinical and research perspective, information on both functional (hypo- and hyper-perfusion) and structural (atrophy) disease processes is still potentially relevant for diagnostic and disease progression purposes. Thus, although areas of atrophy but normal perfusion are very common in aging, areas of perfusion change without atrophy might indicate the early stages of disease progression in dementias. Being able to separate true GM hypoperfusion from apparent hypoperfusion due to atrophy might help distinguish, at an early stage, pathology from normal age-induced changes.

4. Managing PVE in ASL studies

In this section we examine strategies that can be used to manage the influence of PVE in studies that employ ASL, particularly the value of correction methods. We seek to make recommendations based on usage in the literature and our experience using these techniques.

4.1. Restricting analysis to only gray matter voxels

A simple solution to PVE that has been adopted in the literature is to only include voxels identified as being within the GM, normally based on the segmentation of an anatomical image. For ROI-based analysis this involves further restricting the ROI, for a voxelwise analysis this involves a restriction on which voxels contribute to the group analysis which might be implemented on a subject-by-subject basis or based on an average (e.g. atlas based) definition of where GM is present. The most common example of restricting analysis is the definition of whole-brain GM perfusion as discussed in 2.1, but this approach has also been used with other atlas defined anatomical regions.

Substantial variation is seen in the literature as to how this restriction is implemented and often details needed to reproduce the process

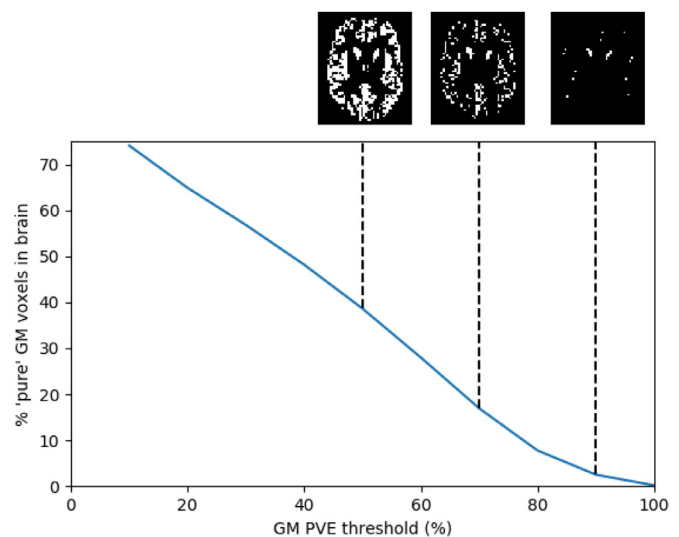


Fig. 3. Variation in the percentage of brain voxels that count as 'pure' GM for different thresholds on the downsampled PV estimate for the data in Fig. 1. Shown above the plot are corresponding images for the ROI created by that threshold for example cases (50, 70 and 90% GM PV).

are missing. Segmentation algorithms, as applied to an anatomical image, may produce any combination of a GM segmentation, indicating voxels where GM is present or the dominant tissue type; a map of the probability of GM being present in each voxel; and a map of partial volume estimates of GM within each voxel. The use of either of the latter two maps requiring the choice of a threshold to define which voxels will be counted as GM. Additionally, this segmentation will have been performed at the resolution of the anatomical image and thus either the perfusion image will need to be up-sampled to the grid of the anatomical image, or the GM information down-sampled onto the grid of the ASL data. Each of these choices, and the algorithm used, potentially leading to subtly different definitions of the final restricted set of GM voxels.

As already noted, it is practically impossible to define 'pure' GM voxels for ASL given the typical data resolution. Fig. 3 illustrates this issue when using a threshold on the GM PV estimate to define a whole-brain GM ROI for the data shown in Fig. 1. Choosing a high threshold to exclude WM and CSF partial volumes leads to very few remaining voxels that they are unlikely to be truly representative of the whole brain; using a lower threshold leads to a more plausible mask but there will be residual contamination with other tissues. In practice, therefore, it is necessary to choose a pragmatic threshold on GM PV estimate and thus there will remain some degree of PVE in the resulting analysis. The process of restricting to GM voxels could also lead to systematic differences based on variation in the cortical thickness of different brain regions or in different populations. Regions or individuals with thicker cortex will have a higher proportion of GM and thus contribute more voxels to the resulting GM mask.

Overall, we would not advocate this approach to managing PVE in ASL since at best it only reduces and cannot eliminate PVE. However, given that there might exist situations where correction for PVE (to be discussed in the next section) might not be feasible, we make some recommendations on how this approach could be used in a way that minimizes systematic differences and offers transparency and reproducibility.

- GM ROIs should be defined in terms of a partial volume estimate, where available, or failing that, the probability of GM tissue in the voxel. A GM segmentation image (a binary mask produced by the segmentation algorithm which will indicate voxels in which GM is the dominant tissue) should not be used to control for PVE.

- A threshold of 70% GM partial volume should be used to define ‘pure’ GM. The >70% threshold recommended here is a pragmatic and conservative choice reflecting the comparatively low resolution of typical ASL data compared to cortical thickness. As can be seen from Fig. 3 this leads to a mask that samples GM perfusion from across the brain, even if it is sparse compared to the voxels that contain some GM tissue. This pragmatic threshold could still lead to a bias, across regions of the brain, or between particular individuals, with those with thicker cortex, contributing more voxels for the calculation of the ROI perfusion.
- Studies should specifically report on the process via which the GM is identified, including: the algorithm used to arrive at PV estimates; the method used to down-sample these estimates or up-sample the perfusion image (including how registration was performed to arrive at any transformations between image spaces, and the interpolation method used); the PV threshold used. Where data sharing permits, the perfusion images and associated partial volume maps should be made available alongside the report of a study.

4.2. Partial volume effect correction

Various PVE correction (PVEc) methods have been specifically developed for ASL to overcome the mixed tissue problem and directly estimate pure GM (and in some cases WM) perfusion at each voxel within the brain. Early methods assumed constant WM perfusion or constant (both spatially as well as between subjects) ratios of WM/GM perfusion values (Johnson et al., 2005; Villain et al., 2008). More recently ‘spatially regularized’ approaches have been developed based on the assumption that, in a small neighbourhood, perfusion in ‘pure’ gray matter is constant (Ahlgren et al., 2014; Asllani et al., 2008; Chappell et al., 2011; Kandel et al., 2015; Liang et al., 2013; Petr et al., 2013; Wiersma et al., 2006) (these are discussed in more depth in ‘Correcting for Partial Volume Effects’ in the Supplementary Materials). By analyzing neighbouring voxels and assuming constant perfusion of each tissue type, the perfusion of the different tissue types can be estimated by exploiting the variation in partial volume fractions from one voxel to the next. Borogovac et al. (Borogovac et al., 2010) examined the sensitivity of ASL to map brain function over an extended time period and found increased sensitivity to CBF changes using PVEc. However, there is little other systematic literature that examines whether this potential advantage translates into improved statistical power for detection of subtle differences in perfusion more generally. There are also potential drawbacks of these PVE correction methods. These methods oversimplify physiological variability within the subject and also may lead to spatial smoothing of the data, potentially reducing spatial specificity. Moreover, these methods lead to increases in noise, an inherent effect of estimating more than one perfusion value from a single set of data at each voxel. PVEc methods are also dependent upon the accuracy of the partial volume estimates and are thus subject to errors associated with the source images for these (normally anatomical imaging) and the algorithms used to make these estimates (as discussed further in ‘Confounds for PVEc’ in the Supplementary Materials).

While PVEc methods are increasingly available in ASL analysis tools, there is limited experience with their use in the literature. Based on this, we recommend that:

- Studies should attempt partial volume correction in addition to performing their analysis without the correction (or vice versa).
- Correction should use a spatially regularized method to avoid imposing strict assumptions on perfusion values in the data (for more details on PVEc methods see ‘Correction for Partial Volume Effects’ found in the Supplementary Materials).
- Where partial volume estimates are defined at a higher resolution than the ASL data, e.g. from the segmentation of a structural image, they should be transformed into the lower resolution of the ASL data by a method that sums the tissue contributions within the voxels, not

by simple linear interpolation (for more details see ‘Transforming higher resolution partial volume estimates to ASL resolution’ found in the Supplementary Materials)

- Care should be taken when comparing results with and without PVEc, but differences should motivate further investigation into how PVE or PVEc affect the results of group analysis.

It might be unusual to advocate performing two analyses of data, and it is important that account is taken of the resulting multiple comparisons and that both analyses are reported. But we believe it is justified given the potential influence of PVE on perfusion, as well as the interpretation of the findings and hence on the conclusions drawn from an ASL study. Moreover, by also reporting non-PVEc values comparison with current body of literature remains possible. Alongside a wider use of PVEc, further work is required to validate PVEc results using model systems in which it is possible to make tissue specific, potentially invasive, perfusion measurements.

4.3. Additional considerations when atrophy is present

Early studies using FDG-PET have demonstrated, using a ratio-based partial volume correction method, apparent hypometabolism independent of atrophy in Alzheimer’s Disease (Villain et al., 2008). While some studies have found a substantial difference in the results obtained with and without correction for PVE in ASL (Borogovac et al., 2010; Steketee et al., 2015), others have detected only subtle differences (Abad et al., 2016; Binnewijzend et al., 2013). Some studies have used anatomical parameters (tissue volume extracted from partial volume estimates or cortical thickness) as regressors in the analysis of the uncorrected CBF images. Using cortical thickness as a regressor, (Chen et al., 2011), found only “subtle” difference in the results, suggesting a minor dissociation between CBF and anatomical changes in the brain due to aging. However, regressor based approaches to account for PVE are not the same as PVEc, and it is thus more difficult to draw conclusions from them about the utility of partial volume correction in ASL perfusion imaging of disease.

Whilst there is more literature addressing PVE in the study of perfusion changes in disease where changes in GM volume are expected, there is no clear consensus on how to manage PVE and thus whether PVEc is warranted. Given the potential bias associated with restricting analysis to GM voxels as discussed in Section 3.1, this approach wouldn’t be appropriate in the context of anatomical difference or change. Hence, our recommendations would be to use PVEc, following the approach outlined in Section 3.2 and undertaking analysis with and without correction for PVE. In this case, it is critical to examine differences in the results between these two analyses and determine whether they can be attributed to differences in GM partial volume.

An emerging alternative to PVEc for the detection of abnormal perfusion in an individual is to attempt to predict a personalized ‘normal’ perfusion map based on anatomical data. For example, Kandel et al. (2015) attempted to model the relationship between anatomy and perfusion using a general linear model that included as regressors the partial volume estimates along with other image-derived features. It is too early to make any recommendations for the use of such an approach and there are ongoing projects evaluating similar methodologies. However, the approach might in the future represent an alternative to attempting to correct the perfusion image for partial volume effects, offering more sensitivity to detect perfusion changes in an individual, for example pathological changes in a patient.

4. Conclusion

The partial volume effect results in systematic underreporting of gray matter perfusion in ASL studies due to the presence of other, less perfused, tissues within ROIs and individual voxels. PVE is a substantial, and often overlooked issue, particularly when ASL is used in group studies and/or clinical populations. Use of GM masking is at best only a

partial solution and must be accompanied by consistent, comprehensive reporting. Experimental results to date have not shown a consistent benefit of correction for PVE, but a variety of methods have been employed in only a small number of studies. The strong conceptual argument for PVE correction, and the ongoing improvement in methods motivate continued evaluation, especially in studies of older or diseased populations where atrophy is expected. Greater use of correction, alongside conventional analyses, and more systematic studies of the influence of PVE and merits of correction are required for the field to move forward in this area and form the basis of a future consensus. Hence, we recommend that studies should attempt and report on the result of partial volume correction in addition to performing their analysis without the correction (or vice versa).

Declaration of Competing Interest

MAC receives royalties for commercial licencing of the FMRIB Software Library and sales of the Oxford Neuroimaging Primers (Oxford University Press) book series. XG is CEO of Gold Standard Phantoms. MG is CEO of mediri GmbH. MJPvO receives Research support from Philips.

Credit authorship contribution statement

Michael A. Chappell: Conceptualization, Writing – original draft. **Flora A. Kennedy McConnell:** Investigation, Writing – review & editing. **Xavier Golay:** Writing – review & editing. **Matthias Günther:** Writing – review & editing. **Juan A. Hernandez-Tamames:** Conceptualization, Writing – original draft. **Matthias J. van Osch:** Conceptualization, Writing – original draft. **Iris Asllani:** Conceptualization, Writing – original draft.

Acknowledgments

We thank the European Union COST agency (through COST Action BM1103) for sponsorship of the workshops that have stimulated the discussion from which this paper has emerged. MC and FKM have received support from the Engineering and Physical Sciences Research Council UK (EP/P012361/1). FKM is supported by the Beacon of Excellence in Precision Imaging, University of Nottingham. MJPvO is supported by Horizon2020 (project: CDS-QUAMRI, Project number 634541) and the Netherlands organization for Scientific Research Grant (Award Number 016.160.351). XG is supported by the National Institute for Health Research University College London Hospitals Biomedical Research center.

The data used in this work for Fig. 2 were obtained from UK Biobank Resource under Application Number 43172. We are grateful to UK Biobank for making the data available, and to all UK Biobank study participants, who generously donated their time to make this resource possible.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.118236.

References

Abad, V.M., Garcia-Polo, P., O'Daly, O., Hernandez-Tamames, J.A., Zelaya, F., 2016. ASAP (Automatic Software for ASL Processing): a toolbox for processing arterial spin labeling images. *Magn. Reson. Imaging* 34, 334–344. doi:10.1016/j.mri.2015.11.002.

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., Casement, M., Bazelaire, C.de, Fong, T., Press, D.Z., 2008. Hippocampal hyperperfusion in Alzheimer's disease. *Neuroimage* 42, 1267–1274. doi:10.1016/j.neuroimage.2008.06.006.

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., Osch, M.J.P., Wang, D.J.J., Wong, E.C.,

Zaharchuk, G., 2015. 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, 102–116. doi:10.1002/mrm.25197.

Álvarez, M.G.M., Stobbe, R.W., Beaulieu, C., 2019. High resolution continuous arterial spin labeling of human cerebral perfusion using a separate neck tagging RF coil. *PLoS ONE* 14, e0215998. doi:10.1371/journal.pone.0215998.

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.

Asllani, I., Habeck, C., Borogovac, A., Brown, T.R., Brickman, A.M., Stern, Y., 2009. Separating function from structure in perfusion imaging of the aging brain. *Hum. Brain Mapp.* 30, 2927–2935. doi:10.1002/hbm.20719.

Bause, J., Ehnes, P., Mirkes, C., Shajan, G., Scheffler, K., Pohmann, R., 2016. Quantitative and functional pulsed arterial spin labeling in the human brain at 9.4 t. *Magn. Reson. Med.* 75, 1054–1063. doi:10.1002/mrm.25671.

Binnewijzend, M.A.A., Kuijter, J.P.A., Benedictus, M.R., Flier, W.M., van derWink, A.M., Wattjes, M.P., Berckel, B.N.M.van, Scheltens, P., Barkhof, F., 2013. Cerebral blood flow measured with 3D Pseudocontinuous arterial spin-labeling MR imaging in Alzheimer disease and mild cognitive impairment: a marker for disease severity. *Radiology* 267, 221–230. doi:10.1148/radiol.12120928.

Borogovac, A., Habeck, C., Small, S.A., Asllani, I., 2010. Mapping brain function using a 30-day interval between baseline and activation: a novel arterial spin labeling fMRI approach. *J. Cereb. Blood Flow Metab.* 30, 1721–1733. doi:10.1038/jcbfm.2010.89.

Cada, A., Torre, J.C., Lima, F.G., 2000. Chronic cerebrovascular ischemia in aged rats: effects on brain metabolic capacity and behavior. *Neurobiol. Aging* 21, 225–233. doi:10.1016/s0197-4580(00)00116-0.

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. *Magnetic resonance in medicine* 65, 1173–1183. doi:10.1002/mrm.22641.

Chen, J.J., Rosas, H.D., Salat, D.H., 2011. Age-associated reductions in cerebral blood flow are independent from regional atrophy. *Neuroimage* 55, 468–478.

Chen, J.J., Salat, D.H., Rosas, H.D., 2012. Complex relationships between cerebral blood flow and brain atrophy in early Huntington's disease. *Neuroimage* 59, 1043–1051. doi:10.1016/j.neuroimage.2011.08.112.

Donahue, M.J., Lu, H., Jones, C.K., Pekar, J.J., Zijl, P.C.M.van, 2006. An account of the discrepancy between MRI and PET cerebral blood flow measures. A high-field MRI investigation. *NMR Biomed.* 19, 1043–1054. doi:10.1002/nbm.1075.

Elliott, L.T., Sharp, K., Alfaro-Almagro, F., Shi, S., Miller, K.L., Douaud, G., Marchini, J., Smith, S.M., 2018. Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature* 562, 210–216. doi:10.1038/s41586-018-0571-7.

Erlundsson, K., Buvat, I., Pretorius, P.H., Thomas, B.A., Hutton, B.F., 2012. A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. *Phys. Med. Biol.* 57, R119–R159. doi:10.1088/0031-9155/57/21/r119.

Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc. Natl. Acad. Sci.* 97, 11050–11055. doi:10.1073/pnas.200033797.

Fjell, A.M., Walhovd, K.B., Fennema-Notestine, C., McEvoy, L.K., Hagler, D.J., Holland, D., Brewer, J.B., Dale, A.M., 2009. One-year brain atrophy evident in healthy aging. *J. Neurosci.* 29, 15223–15231. doi:10.1523/jneurosci.3252-09.2009.

Grade, M., Tamames, J.A.H., Pizzini, F.B., Achten, E., Golay, X., Smits, M., 2015. A neuro-radiologist's guide to arterial spin labeling MRI in clinical practice. *Neuroradiology* 57, 1181–1202. doi:10.1007/s00234-015-1571-z.

Greve, D.N., Salat, D.H., Bowen, S.L., Izquierdo-Garcia, D., Schultz, A.P., Catana, C., Becker, J.A., Svarer, C., Knudsen, G.M., Sperling, R.A., Johnson, K.A., 2016. Different partial volume correction methods lead to different conclusions: an 18F-FDG-PET study of aging. *Neuroimage* 132, 334–343. doi:10.1016/j.neuroimage.2016.02.042.

Harms, M.P., Somerville, L.H., Ances, B.M., Andersson, J., Barch, D.M., Bastiani, M., Bookheimer, S.Y., Brown, T.B., Buckner, R.L., Burgess, G.C., Coalson, T.S., Chappell, M.A., Dapretto, M., Douaud, G., Fischl, B., Glasser, M.F., Greve, D.N., Hodge, C., Jamison, K.W., Jbabdi, S., Kandala, S., Li, X., Mair, R.W., Mangia, S., Marcus, D., Mascali, D., Moeller, S., Nichols, T.E., Robinson, E.C., Salat, D.H., Smith, S.M., Sotiropoulos, S.N., Terpstra, M., Thomas, K.M., Tisdall, M.D., Ugurbil, K., Kouwe, A.van der, Woods, R.P., Zöllei, L., Essen, D.C.V., Yacoub, E., 2018. Extending the human Connectome project across ages: imaging protocols for the lifespan development and aging projects. *Neuroimage* 183, 972–984. doi:10.1016/j.neuroimage.2018.09.060.

Huang, S.C., Carson, R.E., Hoffman, E.J., Carson, J., MacDonald, N., Barrio, J.R., Phelps, M.E., 1983. Quantitative measurement of local cerebral blood flow in humans by positron computed tomography and 15O-Water. *J. Cereb. Blood Flow Metab.* 3, 141–153. doi:10.1038/jcbfm.1983.21.

Huber, L., Uludağ, K., Möller, H.E., 2019. Non-BOLD contrast for laminar fMRI in humans: CBF, CBV, and CMRO2. *Neuroimage* 197, 742–760. doi:10.1016/j.neuroimage.2017.07.041.

Iturria-Medina, Y., Sotero, R.C., Toussaint, P.J., Mateos-Pérez, J.M., Evans, A.C., 2016. Early role of vascular dysregulation on late-onset Alzheimer's disease based on multi-factorial data-driven analysis. *Nat. Commun.* 7, 11934. doi:10.1038/ncomms11934.

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 experience. *Radiology* 234, 851–859. doi:10.1148/radiol.2343040197.

Jong, G.L.D., Farkas, E., Stienstra, C.M., Plass, J.R.M., Keijsers, J.N., Torre, J.C., Luiten, P.G.M., 1999. Cerebral hypoperfusion yields capillary damage in the hippocampal CA1 area that correlates with spatial memory impairment. *Neuroscience* 91, 203–210. doi:10.1016/s0306-4522(98)00659-9.

- Kandel, B.M., Wang, D.J.J., Detre, J.A., Gee, J.C., Avants, B.B., 2015. Decomposing cerebral blood flow MRI into functional and structural components: a non-local approach based on prediction. *Neuroimage* 105, 156–170. doi:10.1016/j.neuroimage.2014.10.052.
- Kety, S.S., Schmidt, C.F., 1948. The nitrous oxide method for the quantitative determination of cerebral blood flow in man: theory, procedure and normal values. *J. Clin. Invest.* 27, 476–483. doi:10.1172/jci101994.
- Kim, S., Song, J., Yoon, J., Kim, K., Chung, J., Noh, Y., 2021. Voxel-wise partial volume correction method for accurate estimation of tissue sodium concentration in ²³Na-MRI at 7 T. *NMR Biomed.* 34, e4448. doi:10.1002/nbm.4448.
- Li, X., Wang, D., Auerbach, E.J., Moeller, S., Ugurbil, K., Metzger, G.J., 2015. Theoretical and experimental evaluation of multi-band EPI for high-resolution whole brain pCASL Imaging. *Neuroimage* 106, 170–181. doi:10.1016/j.neuroimage.2014.10.029.
- Liang, X., Connelly, A., Calamante, F., 2013. Improved partial volume correction for single inversion time arterial spin labeling data. *Magn. Reson. Med.* 69, 531–537. doi:10.1002/mrm.24279.
- Luh, W., Talagala, S.L., Li, T., Bandettini, P.A., 2013. Pseudo-continuous arterial spin labeling at 7 T for human brain: estimation and correction for off-resonance effects using a Prescan. *Magn. Reson. Med.* 69, 402–410. doi:10.1002/mrm.24266.
- Meltzer, C.C., Cantwell, M.N., Greer, P.J., Ben-Eliezer, D., Smith, G., Frank, G., Kaye, W.H., Houck, P.R., Price, J.C., 2000. Does cerebral blood flow decline in healthy aging? A PET study with partial-volume correction. *J. Nucl. Med.* 41, 1842–1848. doi:10.1001/archpsyc.1987.01800190037006.
- 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, 1523–1536. doi:10.1038/nn.4393.
- Niesporek, S.C., Hoffmann, S.H., Berger, M.C., Benkhedah, N., Kujawa, A., Bachert, P., Nagel, A.M., 2015. Partial volume correction for in vivo ²³Na-MRI data of the human brain. *Neuroimage* 112, 353–363. doi:10.1016/j.neuroimage.2015.03.025.
- Petr, J., Mutsaerts, H.J.M.M., Vita, E.D., Steketee, R.M.E., Smits, M., Nederveen, A.J., Hofheinz, F., Hoff, J., van den, Asllani, I., 2018. Effects of systematic partial volume errors on the estimation of gray matter cerebral blood flow with arterial spin labeling MRI. *Magn. Reson. Mater. Phys. Biol. Med.* 1–10. doi:10.1007/s10334-018-0691-y.
- Petr, J., Schramm, G., Hofheinz, F., Langner, J., Hoff, J., van den, 2013. Partial volume correction in arterial spin labeling using a Look-Locker sequence. *Magn. Reson. Med.* 70, 1535–1543. doi:10.1002/mrm.24601.
- Pini, L., Pievani, M., Bocchetta, M., Altomare, D., Bosco, P., Cavedo, E., Galluzzi, S., Mariz-zoni, M., Frisoni, G.B., 2016. Brain atrophy in Alzheimer's disease and aging. *Ageing Res. Rev.* 30, 25–48. doi:10.1016/j.arr.2016.01.002.
- Samuraki, M., Matsunari, I., Chen, W.P., Yajima, K., Yanase, D., Fujikawa, A., Takeda, N., Nishimura, S., Matsuda, H., Yamada, M., 2007. Partial volume effect-corrected FDG PET and grey matter volume loss in patients with mild Alzheimer's disease. *Eur. J. Nucl. Med. Mol. Imaging* 34, 1658–1669. doi:10.1007/s00259-007-0454-x.
- Steketee, R.M.E., Bron, E.E., Meijboom, R., Houston, G.C., Klein, S., Mutsaerts, H.J.M.M., Orellana, C.P.M., Jong, F.J.de, Swieten, J.C.van, Lugt, A.van der, Smits, M., 2015. Early-stage differentiation between presenile Alzheimer's disease and frontotemporal dementia using arterial spin labeling MRI. *PubMed - NCBI Eur. Radiol.* 26, 244–253. doi:10.1007/s00330-015-3789-x.
- Telischak, N.A., Detre, J.A., Zaharchuk, G., 2014. Arterial spin labeling MRI: clinical applications in the brain. - PubMed - NCBI. *J. Magn. Reson. Imaging JMRI* 41. doi:10.1002/jmri.24751, n/a-n/a.
- Thomas, B.A., Cuplov, V., Bousse, A., Mendes, A., Thielemans, K., Hutton, B.F., Erlandsson, K., 2016. PETPVC: a toolbox for performing partial volume correction techniques in positron emission tomography. *Phys. Med. Biol.* 61, 7975–7993. doi:10.1088/0031-9155/61/22/7975.
- Villain, N., Desgranges, B., Viader, F., Sayette, V.de la, Mézenge, F., Landeau, B., Baron, J.C., Eustache, F., Chételat, G., 2008. Relationships between hippocampal atrophy, white matter disruption, and gray matter hypometabolism in Alzheimer's disease. *J. Neurosci. Off. J. Soc. Neurosci.* 28, 6174–6181. doi:10.1523/jneurosci.1392-08.2008.
- Wiersma, J., Deichmann, R., Ordidge, R., Turner, R., 2006. Removing the effects of CSF partial voluming on fitted CBF and arterial transit times using FAIR, a pulsed arterial spin labeling technique. *Magn. Reson. Mater. Phys. Biol. Med.* 19, 115–123. doi:10.1007/s10334-006-0034-2.
- Zhao, M.Y., Mezue, M., Segerdahl, A.R., Okell, T.W., Tracey, I., Xiao, Y., Chappell, M.A., 2017. 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, 384–397. doi:10.1016/j.neuroimage.2017.08.072.
- Zuo, Z., Wang, R., Zhuo, Y., Xue, R., Lawrence, K.S.St., Wang, D.J.J., 2013. Turbo-FLASH based arterial spin labeled perfusion MRI at 7 T. *PLoS ONE* 8, e66612. doi:10.1371/journal.pone.0066612.