ExploreASL: An image processing pipeline for multi-center ASL perfusion MRI studies

Henk J.M.M. Mutsaerts a,b,c,d,e,*,1, Jan Petr d,f,1, Paul Groot b, Pieter Vandemaele e, Silvia Ingala a, Andrew D. Robertson s, Lena Václaví b, Inge Groote i, Hugo Kuijf f, Fernando Zelaya k, Owen O’Daly k, Saima Hilal l,m,n, Alle Meije Wink b, Ilse Kant c,d, Matthew W.A. Caan p, Catherine Morgan q, Jeroen de Bresser i, Elisabeth Lysvik j, Anouk Schrantee b, Astrid Bjørnebekk s, Patricia Clement e, Zahra Shirzadi t, Joost P.A. Kuijer a, Viktor Wottschel a, Udnuna C. Anazodo s,t, Dasja Pajkrt w, Edo Richard d, Reinoud P.H. Bokkers z, Liesbeth Reneman b, Mario Maselli l, Matthias Günther as,ab,ae, Bradley J. MacIntosh f, Eric Achten e, Michael A. Chappell ad, Matthias J.P. van Osch h, Xavier Golay ae, David L. Thomas ae, Enrico De Vita af, Atle Bjørnerud i,ag, Aart Nederveenb, Jeroen Hendrikse c, Iris Asllani d,ah, Frederik Barkholt a,ae,ai

a Department of Radiology and Nuclear Medicine, Amsterdam Neuroscience, Amsterdam University Medical Center, Location VUmc, Amsterdam, the Netherlands
b Radiology and Nuclear Medicine, Amsterdam Neuroscience, Amsterdam University Medical Centers, Location Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
c Radiology, University Medical Center Utrecht, Utrecht, the Netherlands
d Kate Gleason College of Engineering, Rochester Institute of Technology, NY, USA
e Ghent Institute for Functional and Metabolic Imaging (GI4MI), Ghent University, Ghent, Belgium
f Helmholtz-Zentrum Dresden-Rossendorf, Institute of Radiopharmaceutical Cancer Research, Dresden, Germany
g Schlegel-UW Research Institute for Aging, University of Waterloo, Waterloo, Ontario, Canada
h C.J. Gorter Center for High Field MRI, Department of Radiology, Leiden University Medical Center, Leiden, the Netherlands
i Department of Diagnostic Physics, Oslo University Hospital, Oslo, Norway
j Image Sciences Institute, University Medical Center Utrecht, Utrecht, the Netherlands
k Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, UK
l Department of Pharmacology, National University of Singapore, Singapore
m Memory Aging and Cognition Center, National University Health System, Singapore
n Saw Swee Hock School of Public Health, National University of Singapore, Singapore
o Department of Intensive Care, University Medical Centre, Utrecht, the Netherlands
p Department of Biomedical Engineering and Physics, Amsterdam University Medical Center, Location Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
q School of Psychology and Centre for Brain Research, University of Auckland, Auckland, New Zealand
r Department of Radiology, Leiden University Medical Center, Leiden, the Netherlands
s The Anabolic Androgenic Steroid Research Group, National Advisory Unit on Substance Use Disorder Treatment, Oslo University Hospital, Oslo, Norway
t Sunnybrook Research Institute, University of Toronto, Toronto, Canada
u Department of Medical Biophysics, University of Western Ontario, London, Canada
v Imaging Division, Lawson Health Research Institute, London, Canada
w Department of Pediatric Infectious Diseases, Emma Children’s Hospital, Amsterdam University Medical Center, Location Academic Medical Center, Amsterdam, the Netherlands
x Department of Neurology, Donders Institute for Brain, Behavior and Cognition, Radboud University Medical Center, Nijmegen, the Netherlands
y Neuroscience, Amsterdam University Medical Center, Location Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
z Department of Radiology, Medical Imaging Center, University of Groningen, Groningen, the Netherlands
aa Fraunhofer MEVIS, Bremen, Germany
ab University of Bremen, Bremen, Germany
ac Mediri GmbH, Heidelberg, Germany
ad Institute of Biomedical Engineering, Department of Engineering Science & Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK
ae UCL Queen Square Institute of Neurology, University College London, London, UK

* Corresponding author. Dep. of Radiology and Nuclear Medicine, PK -1, De Boelelaan 1117, 1081, HV, Amsterdam, the Netherlands.

https://doi.org/10.1016/j.neuroimage.2020.117031

Available online 8 June 2020
1053-8119/© 2020 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Arterial spin labeling (ASL) has undergone significant development since its inception, with a focus on improving robust and reproducible clinical ASL image processing, we developed the software package ExploreASL, allowing standardized analyses across centers and scanners.

The procedures used in ExploreASL capitalize on published image processing advancements and address the challenges of multi-center datasets with scanner-specific processing and artifact reduction to limit patient exclusion. ExploreASL is self-contained, written in MATLAB and based on Statistical Parameter Mapping (SPM) and runs on multiple operating systems. To facilitate collaboration and data-exchange, the toolbox follows several standards and recommendations for data structure, provenance, and best analysis practice.

ExploreASL was iteratively refined and tested in the analysis of >10,000 ASL scans using different pulse-sequences in a variety of clinical populations, resulting in four processing modules: Import, Structural, ASL, and Population that perform tasks, respectively, for data curation, structural and ASL image processing and quality control, and finally preparing the results for statistical analyses on both single-subject and group level. We illustrate ExploreASL processing results from three cohorts: perinatally HIV-infected children, healthy adults, and elderly at risk for neurodegenerative disease. We show the reproducibility for each cohort when processed at different centers with different operating systems and MATLAB versions, and its effects on the quantification of gray matter cerebral blood flow.

ExploreASL facilitates the standardization of image processing and quality control, allowing the pooling of cohorts which may increase statistical power and discover between-group perfusion differences. Ultimately, this workflow may advance ASL for wider adoption in clinical studies, trials, and practice.
brain tumors (Warmuth et al., 2003), as well as pharmacological applications (Wang et al., 2011; MacIntosh et al., 2008; Handley et al., 2013). Following the consensus recommendations for the acquisition and quantification of ASL images (Alsop et al., 2015), ASL became ready for large multi-center observational studies and clinical trials (Jack et al., 2010; Almeida et al., 2018; Blokhuis et al., 2017).

However, despite the consensus in clinical implementation and image acquisition (Alsop et al., 2015), ASL image processing (Wang et al., 2008; Shin et al., 2016; Melbourne et al., 2016; Chappell et al., 2010; Li et al., 2019; Mato Abad et al., 2016; Bron et al., 2014) remains disparate among research laboratories. In previous ASL studies, detailed description of all processing steps is often lacking. Clinical studies are often performed without proper quality control (QC) or with arbitrary QC metrics. This hampers both the interpretation and reproducibility of individual studies as well as meta-analyses of multiple studies. A consensus on the best practices to robustly process ASL data would facilitate comparison of as well as meta-analyses of multiple studies. A consensus on the best practices to robustly process ASL data would facilitate comparison of results across centers and studies, avoid duplicate development, and speed up the translation into clinical practice, as is advocated by the practices to robustly process ASL data would facilitate comparison of as well as meta-analyses of multiple studies. A consensus on the best practices to robustly process ASL data would facilitate comparison of results across centers and studies, avoid duplicate development, and speed up the translation into clinical practice, as is advocated by the following consensus recommendations for the acquisition and quantification of ASL images (Alsop et al., 2015), ASL became ready for large multi-center observational studies and clinical trials (Jack et al., 2010; Almeida et al., 2018; Blokhuis et al., 2017).

For these reasons, the software package ExploreASL was initiated through the COST-action BM1103 “OSIPI (www.osipi.org). With the aim of developing a comprehensive pipeline for reproducible multi-center ASL image processing. To date, ExploreASL has been used in more than 30 studies consisting of more than 10,000 ASL scans from three MRI vendors - GE, Philips, Siemens, with pulsed ASL and pseudo-continuous ASL (PCASL) sequences, see the list of studies in the Supplementary material. The primary aims of ExploreASL are to increase the comparability and enable pooling of multi-center ASL datasets, as well as to encourage and facilitate cross-pollination between clinical investigators and image processing method developers.

2. Theory: Software overview

ExploreASL is developed in MATLAB (MathWorks, MA, USA, tested with versions 2011–2019) and uses Statistical Parametric Mapping 12 routines (SPM12, version 7219) (Ashburner et al., 2012; Flandin et al., 2008). Here, we describe the implementation of ExploreASL version 1.0.0, which is available for free for non-commercial use on www.ExploreASL.org or at https://github.com/ExploreASL/ExploreASL. ExploreASL provides a fully automated pipeline that comprises all the necessary steps from data import and structural image processing to cerebral blood flow (CBF) quantification and statistical analyses. Unique features of ExploreASL include:

- Self-contained software suite: all third-party toolboxes are included in the installation, compatible with Linux, macOS, Windows, and Windows Subsystem for Linux and supporting multi-threading; ExploreASL requires a Matlab installation to run interactively but a compiled version of ExploreASL is also available. The system requirements are a single-core CPU, 4 Gb RAM, minimum 1 Gb free disk space. When running multiple instances of ExploreASL in parallel, 2.5 Gb RAM and 1 Gb disk space is recommended per instance;
- Flexible data import from different formats including (enhanced) DICOM, Siemens MOSAIC variant, Philips PAR/REC, NIfTI and Brain Imaging Data Structure (BIDS) (Gorgolewski et al., 2016), with automatic detection of control-label or label-control order;
- Data management: anonymization, compression of image files, and optional defacing;
- Modular design: automatically iterates over all available subjects and scans, allows investigators to change/replace each sub-module, allows to suspend and resume processing at any point - pipeline steps are tracked using a system of lock and status files. Before executing a submodule for a given subject and session, a lock file is created to avoid parallel access to the data. When the submodule is finished, the lock file is removed and a status file is created. When the pipeline is restarted - e.g. after a computer or program crash - the already processed steps are skipped;
- Image processing optimized for: multiple centers, different ASL implementations from GE/Philips/Siemens (Mutsaerts et al., 2018), both native/standard-space analysis, advanced ASL markers – e.g. spatial coefficient-of-variation (CoV) (Mutsaerts et al., 2017), asymmetry index (Kurth et al., 2015), and partial volume correction (PVC) (Aslani et al., 2008);
- ‘Low quality’ option allowing for quick pipeline testing by running all image processing with fewer iterations and lower spatial resolution;
- Extensive QC and data provenance: visual QC for all intermediate and final images, comparison with perfusion templates from different ASL implementations, progress report with processing history (provenance).

ExploreASL requires the following input data: ASL and T1-weighted (T1w) images, and optionally FLAIR and M0 scans. Other options are binary images with lesion masks and additional ROIs either in the T1w or FLAIR space. The ASL acquisition parameters that cannot be extracted from the DICOM data (e.g. PLD, labeling duration) need to be provided in either a study-configuration file or a JSON sidecar. As output, ExploreASL provides the processed structural and quantified CBF NIfTI files in the T1w and ASL native space, as well as in the standard space, see Supplementary Fig. 1. Detailed information for users, including a manual, a step-by-step walkthrough, and video tutorials, is provided on the ExploreASL website www.ExploreASL.org.

In the following sections, we review each processing step of the four ExploreASL modules as outlined in Fig. 1 and summarized in Table 1. Each section starts with a brief methodological review including the rationale within the context of ASL processing, followed by a detailed description of the ExploreASL implementation, and ending with a discussion of emerging developments and potential future improvements.

3. Theory: Implementation

3.1. Import module

To avoid manual restructuring of arbitrary data structures from the scanner or other sources (Nichols et al., 2017), ExploreASL uses a flexible input data/directory description scheme based on regular expressions and converts the data to a BIDS-compatible data structure (Gorgolewski et al., 2016); the full BIDS ASL extension is currently in development (bi ds.neuroimaging.io). The input images can be NIfTI format, conventional or enhanced DICOM, Philips PAR/REC, or Siemens mosaic format, which are then converted to NIfTI using dcm2niix, taking into account vendor-specific scale slopes in private tags (Li et al., 2016). ASL images can be provided as control-label time-series, a single perfusion-weighted image, or an already quantified CBF image, from any 2D or 3D readout schemes, and from any MRI vendor. Before an image is processed, ExploreASL first computes and aligns the center-of-mass of each image to the origin of the world coordinates to deal with potentially incorrectly stored orientations. A tolerance of 50 mm is used for the center-of-mass offset to avoid resetting correct initial alignments. Additionally, ExploreASL provides an overview of missing and unprocessed files, automatically detects the order of control and label images from the image intensities, and checks the Dicom tags of repetition and echo time and scale factors/slopes across individuals.

3.2. Structural module

This module processes the structural images by the following steps: 2.1) segments the white matter (WM) hyperintensities (WMH) on fluid-attenuated inversion recovery (FLAIR) images and uses them to fill the corresponding WM hypointensities on the T1w images, 2.2) the structural images are subsequently segmented into gray matter (GM), white matter, and cerebrospinal fluid (CSF) maps, and 2.3) normalized to the
MNI standard space (Evans et al., 2012). The segmentations are used to obtain tissue partial volume (PV) fractions for computation of CBF (Asllani et al., 2008). The registration transformations are used to bring ASL images acquired from different sessions and/or different subjects in the same space and thus facilitate visual comparison in the same space, automatic QC, as well as group analysis.

3.2.1. WMH correction

The presence of WMH can affect the GM/WM classification of T1w images in two ways: 1) WMH themselves can be incorrectly segmented as GM, 2) image intensities of WMH bias global modeling of GM and WM intensity distributions (Pareto et al., 2016; Battaglini et al., 2012). ExploreASL alleviates these complications by lesion-filling the T1w image before initiating the segmentation (Battaglini et al., 2012); voxel intensities in the hypointense WMH regions on the T1w images are replaced by bias field-corrected values from the surrounding, normal-appearing WM (Chard et al., 2010) (Fig. 1). The Lesion Segmentation Toolbox (LST, version 2.0.15) is used because of its empirically proven robustness, scanner independence, and non-reliance on the requirement of a training set (de Sitter et al., 2017a). LST detects outliers in the FLAIR WM intensity distribution and assesses their likelihood of being WMH (Schmidt et al., 2012). While ExploreASL offers the option of both LST lesion growing and lesion prediction algorithms, the default is set to the latter, which has been shown to be more robust (de Sitter et al., 2017a). This WHM correction described here is only performed when FLAIR images are available.

3.2.2. Segmentation

To segment the 3 main tissue classes GM, WM, and CSF, ExploreASL uses the Computational Anatomy Toolbox 12, release 1363 (CAT12, the successor of VBM8) (Gaser et al., 2009) for SPM12. CAT12 allows local variations in the tissue intensity distributions, making it more robust to the presence of pathology such as tumors, edema, and WM lesions (Battaglini et al., 2012; Petr et al 2018b) (Supplementary Fig. 2). CAT12 has been shown to outperform other available methods such as FreeSurfer v5.3.0, FSL v5.0, and SPM12 (Mendrik et al., 2015). The CAT12 segmentation algorithm is based on improvements of Unified Segmentation (Ashburner et al., 2005), two essential improvements being that it allows spatially varying GM-WM intensity distributions, and provides PV maps rather than posterior probability maps (Tohka et al., 2004).

3.2.3. Spatial normalization

For non-linear registration to MNI space ExploreASL uses Geodesic Shooting (Ashburner et al., 2011) - the successor of Diffeomorphic Anatomical RegisTration using Exponentiated Lie algebra (DARTEL) (Ashburner et al., 2007) - within the CAT12 toolbox (Gaser et al., 2009). The reason for this choice is that CAT12 has a single subject implementation using the IXI adult template, brain-development.org/ixi-dataset. Optionally, new templates can be created by these SPM toolboxes on a population level, e.g. for populations where an adult template is not sufficient. Although alternative methods (Klein et al., 2010) may outperform DARTEL/GS in specific populations, the default settings of DARTEL and Geodesic Shooting are sufficiently tested in clinical studies to provide adequate performance across different populations and scanners (Ripollès et al., 2012).

We adapted the CAT12 segmentation algorithm to offer the possibility to input customized segmentations of structural lesions such as space-occupying lesions or cerebral infarcts such that the lesion region is ignored by the non-linear registration (Crinion et al., 2007) (Supplementary Fig. 3).

ExploreASL offers the option to register longitudinal ASL studies with the SPM12 module for longitudinal registration (Ashburner and Ridgway, 2012), which takes the similarity between structural images from the same subjects into account. The first time point is used as a reference for both within- and between-subject registration. However, this requires further validation in the presence of large brain deformations between sessions, such as tumors, resections, or infarcts (Petr et al 2018b).

3.3. ASL module

This module processes the ASL images by 3.1) correcting for motion, 3.2) removing outliers, 3.3) registering with the structural data, and by 3.4) processing the M0 images. Then, 3.5) the CBF is quantified with correction for hematocrit and vascular artifacts, after which 3.6) the PV effects are corrected for. All image processing described below is performed in native space, unless stated otherwise. All intermediate and final images are also transformed into standard space for QC and group analyses.

3.3.1. Motion correction

The adverse effects of head motion can be partly alleviated by
Table 1
Overview of image processing steps and implementation in ExploreASL.

<table>
<thead>
<tr>
<th>Processing step</th>
<th>ExploreASL implementation</th>
<th>Specifics, optional features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Import module</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1 Data import</td>
<td>dcm2niix</td>
<td>Converts DICOM to NIfTI, supports MOSAIC, PAR/REG, BIDS</td>
</tr>
<tr>
<td>2. Structural module</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1 WMH correction</td>
<td>LST2 (±r2.0.15) - LPA (LGA optional)</td>
<td>Segments WMH and fills lesions on T1w; improves T1w segmentation</td>
</tr>
<tr>
<td>2.2 Segmentation</td>
<td>CAT12 (±1363)</td>
<td>Outputs partial volume map; supports lesion cost function masking</td>
</tr>
<tr>
<td>2.3 Spatial normalization</td>
<td>Geodesic Shooting</td>
<td>Uses CAT12 template, supports creation of study-specific templates</td>
</tr>
<tr>
<td>3. ASL module</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1 Motion correction</td>
<td>SPMI2 realign</td>
<td>Realigns ASL control/label images to mean position, uses a zig-zag control-label regressor</td>
</tr>
<tr>
<td>3.2 Outlier exclusion</td>
<td>ENABLE</td>
<td>Removes motion peaks, uses tSNR optimization</td>
</tr>
<tr>
<td>3.3 Registration</td>
<td>SPMI2 rigid-body</td>
<td>Registers ΔM-PM or M0-T1w time course masks, smooths, extrapolates M0 to avoid division artifacts</td>
</tr>
<tr>
<td>3.4 M0 processing</td>
<td>M0 image</td>
<td></td>
</tr>
<tr>
<td>3.5 CBF quantification</td>
<td>Consensus paper model</td>
<td>Computes CBF based on the single compartment model, single PLD; supports dual compartment</td>
</tr>
<tr>
<td>3.6 PVC</td>
<td>Linear regression</td>
<td>Performs PVC on kernel or ROI basis, optionally estimates the effective spatial resolution/PSF of ASL</td>
</tr>
<tr>
<td>3.7 Analysis mask creation</td>
<td>Combine individual masks</td>
<td>Combines FoV, susceptibility artifacts and vascular artifacts, use $p &gt; 0.95$ of population masks</td>
</tr>
<tr>
<td>4. Population module</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.1 Template creation</td>
<td>Final images</td>
<td>Calculates population mean, SD, CBF, SNR, also for intermediate images</td>
</tr>
<tr>
<td>4.2 Multi-sequence equalization</td>
<td>Remove residual sequence-specific effects</td>
<td>Equalizes bias fields, spatial CBF, and smoothness, uses sequence-specific templates</td>
</tr>
<tr>
<td>4.3 ROI statistics</td>
<td>CBF and spatial CBF, with or without PVC</td>
<td>Uses MNI structural, Harvard-Oxford, Hammers, and custom atlases</td>
</tr>
<tr>
<td>4.4 Quality control</td>
<td>Single-subject PDF report</td>
<td>Performs QC of images, DICOM values, volumetrics, motion etc. Outputs population report in TSV files</td>
</tr>
</tbody>
</table>

ASL = arterial spin labelling, BIDS = Brain Imaging Data Structure, CAT = Computational Anatomic Toolbox, CBF = cerebral blood flow, C0V = coefficient of variation, ΔM = perfusion-weighted difference image, dcm2niix (Li et al., 2016), DICOM = Digital Imaging and Communications in Medicine, ENABLE = ENhancement of Automated Blood flow Estimates, FoV = field-of-view, LGA = Lesion Growth Algorithm, LPA = Lesion Prediction Algorithm, LST = Lesion Segmentation Toolbox, MNI = Montreal Neurological Institute, NIfTI = Neuro-Imaging Informatics Technology Initiative, QC = quality control, pGM = gray matter partial volume, PLD = post-labeling delay, PSF = point spread function, PVC = partial volume correction, r = release, ROI = region of interest, SD = standard deviation, SNR = signal-to-noise ratio, SPM = Statistical Parametric Mapping, tSNR = temporal SNR, tsv = tab-separated value, WMH = white matter hyperintensity, Zig-zag = "zig-zag" regressor.

correcting for motion using image processing (Alsop et al., 2015). Traditionally, head motion is estimated assuming a 3D rigid-body transformation with a sum-of-squares cost function (Wang et al., 2009; Mato Abad et al., 2016). However, because the average control-label intensity difference can be partly interpreted by the algorithm as motion, some investigators perform motion estimation separately for the control and labeled images (Wang et al., 2009). Instead, in ExploreASL, an adaptation of the SPM12 motion correction is used, which minimizes apparent motion attributable to the control-label intensity difference from the estimated motion parameters using a “zig-zag” regressor (Wang et al., 2012) (Supplementary Fig. 4).

3.3.2. Outlier exclusion

Despite motion correction, large motion spikes can still have a significant negative effect on the ASL image quality, especially when they occur between control and label images (Wang et al., 2008). In fMRI literature, peak motion relative to mean individual motion is often excluded based on a set threshold, e.g. RMS of 0.5 of the voxel size (Power et al., 2012). ExploreASL uses a threshold-free method named ENHancement of Automated BLOOD flow Estimates (ENABLE) (Shirzadi et al., 2015), which sorts control-label pairs by motion and cumulatively averages them until the addition of further pairs significantly decreases the temporal voxel-wise signal stability (Supplementary Fig. 5). The ExploreASL implementation of ENABLE employs the median GM voxel-wise temporal SNR (tSNR) as the criterion for signal stability (Shirzadi et al., 2018), regularized by an empirically defined minimum tSNR improvement of 5%. ENABLE can also remove non-motion-related outliers, since other acquisition artifacts can be picked up by the motion estimation algorithm (Supplementary Fig. 5). ExploreASL also relies on the fact that ENABLE (Shirzadi et al., 2018) also partly removes outliers, as it operates relatively independent of (patho-)physiological changes of the signal intensity in the pairwise subtracted images (Robertson et al., 2017; Li et al., 2018b).

3.3.3. Registration

Accurate registration between the ASL and structural space is a critical step as registration errors are propagated to subsequent stages and analyses of CBF data. Specifically, the relatively large CBF differences between GM, WM, and CSF, mean that small misalignments can have a large impact on the accuracy of tissue-specific CBF quantification (Mutsaerts et al., 2018).

The image registration steps implemented in ExploreASL are based on a previous study in which the performance of several registration options were compared (Mutsaerts et al., 2018). Briefly, the registration of ΔM to gray matter partial volume (pGM) outperformed the registration of M0 to T1w, except for cases where the ΔM contrast was dissimilar to the pGM contrast (e.g. vascular artifacts, labeling artifacts, perfusion pathology). Rigid-body transformation proved to be a robust default choice (Mutsaerts et al., 2018), especially in the presence of pathology (Wang et al., 2008; Macintosh et al., 2010). Therefore, ExploreASL initializes the registration with a M0-T1w registration, after which it performs a ΔM-pGM rigid-body registration by default. The latter is disabled when macrovascular signal predominates tissue signal (spatial CoV above 0.67) (Mutsaerts et al., 2018). However, using the M0-T1w-only option is recommended for 2D PASL without background suppression due to a possible presence of additional artifacts (Supplementary Fig. 6). Note that such images are typically excluded from CBF statistics and only included when analyzing vascular parameters, such as the spatial CoV.

The rigid-body transformation does not account for the geometric distortion typical for 2D echo-planar imaging (EPI) or 3D Gradient And Spin Echo (3D GRASE) ASL images (Gai et al., 2017). Such deformations can be partially corrected with B0 field maps or M0 images with reversed phase-encoding direction (Madali et al., 2016) - which is implemented as option in ExploreASL by calling FSL TopUp (Andersson et al., 2003). Affine and uniform non-linear transformations, such as FNIRT or SPM’s ‘unified segmentation’ (Klein et al., 2009) can outperform the rigid-body transformation in the ΔM-pGM registration (Petr et al. 2018a), although this remains to be validated in the presence of pathology.

3.3.4. M0 processing

The difference in ASL control-label signal is proportional to CBF with the equilibrium magnetization (M0) of blood acting as a scale factor. Ideally, blood M0 would be measured in voxels containing only arterial blood, but that is not usually possible due to the relatively low spatial resolution of ASL images. Instead, M0 is calculated from either the brain
tissue or CSF signal intensity (Cavusoglu et al., 2009). The use of the tissue-based M0 is recommended (Alsop et al., 2015) because of its ability to account for acquisition-specific effects such as variations in receive coil inhomogeneity or T2* weighting. For these reasons, ExploreASL by default processes an M0 image, and optionally supports the use of a single CSF M0 value (Cavusoglu et al., 2009; Pinto et al., 2020).

ExploreASL aims to deliver consistent M0 quantification for multicenter populations with M0-scans acquired at different repetition time and different effective resolutions. ExploreASL smooths the M0 image with a 16 mm FWHM Gaussian (Beaumont, 2015) after it has been masked for WM (Supplementary Figs. 7-8) and rescaled to the mean GM M0 to account for B1 differences between GM and WM. This approach reduces the M0 image into a smooth bias field with the same smoothness/effective resolution for all ASL sequences and participants, and optimal SNR, while still canceling out acquisition-specific B1-field related intensity inhomogeneity. This makes the M0 image more robust and less sensitive to misalignment, and thus more consistent between ASL sequences (Mutsaerts et al., 2018) and individuals (Deisler et al., 2008). ExploreASL has the option to additionally mask the M0 bias-field for lesions that affect the M0 - e.g. brain tumors - and interpolate the M0 signal from the relatively unaffected brain regions (Croal et al., 2019).

3.3.5. CBF quantification

An in-depth overview of ASL CBF quantification has been provided previously (Alsop et al., 2015; Chappell et al., 2018). The previously recommended single compartment model assumes that the label decays with arterial blood T1 only (Alsop et al., 2015). Although a two-compartment model can provide CBF values that are in closer agreement with arterial blood T1 only (Alsop et al., 2015). Several CBF algorithms have been proposed (Chappell et al., 2010; Zhao et al., 2017; Aslani et al., 2008; Liang et al., 2013), which assume locally homogeneous GM and WM CBF. Instead in some studies, GM volume is used as a covariate in the statistical analysis (Chen et al., 2011a). Note that while PVC, in theory, corrects only for the PV effects and takes into account the intra- and inter-subject variability of the GM-WM CBF ratio, GM covariation can additionally affect the estimated physiological correlation between GM CBF and GM volume (Petr et al. 2018).

ExploreASL employs two versions of PVC, both based upon the most frequently used PVC, i.e. linear regression (Aslani et al., 2008): 1) a 3D Gaussian instead of a 2D flat kernel (default, referred to as “voxel-wise”) (Oliver, 2015), or 2) computing PV-corrected CBF within each anatomical or functional region of interest (ROI) separately instead of using a kernel. Whereas the voxel-wise option allows further voxel-based analysis (VBA), the ROI-based PVC is in theory beneficial for a ROI-based analysis as effectively the kernel-size is selected based upon the anatomical ROI, which should be less sensitive to local segmentation errors. Moreover, this avoids cross-talk between ROIs. It still needs to be investigated how to define regions of optimal shape with respect to PVC performance, which depends on the spatial uniformity and SNR of the GM and WM CBF, and partial volume distributions within the ROI. To evaluate the effects of PVC, ExploreASL exports CBF maps and ROI values both with and without PVC.

For proper PVC or ROI definition, the true acquisition resolution - which often differs from the reconstructed voxel size - needs to be taken into account (Petr et al., 2018). This is especially important for 3D readouts, where the through-plane PSF can be up to 1.9 times the nominal voxel-size (Vidorreta et al. 2013, 2014). Effects such as motion (Petr et al., 2016) and scanner reconstruction filters can contribute to further widening of the PSF of the final image. ExploreASL by default uses previously estimated true acquisition resolutions (Vidorreta et al. 2013, 2014; Petr et al., 2018) and can optionally perform a data-driven spatial resolution estimation (Petr et al., 2018) that is generalizable to all ASL acquisitions. Contrary to alternative PSF estimations based on temporal noise autocorrelation (Cox, 2012) or simulations of the acquisition PSF (Vidorreta et al. 2013, 2014), this method does not require time series and inherently accounts for other sources of blurring (e.g. smoothing by motion and/or image processing) and is applicable without having detailed information about the sequence parameters needed to calculate the resolution from the k-space trajectory. However, this method requires further validation, especially in the presence of ASL image artifacts.

Lastly, the GM/WM maps obtained from the high-resolution structural images need to be downsampled to the ASL resolution before they are used for PVC or for ROI delineation in native space. A trivial interpolation to lower resolution may introduce aliasing, which can be addressed by applying a Gaussian filter - or a convolution with the PSF, if the PSF is known - prior to downsampling (Cardoso et al., 2015). This is important to note that the ASL image often has an anisotropic resolution and may be acquired at a different orientation compared to the structural image. To correct for this effect, ExploreASL pre-smooths the structural images with a Gaussian kernel of which the covariance matrix takes the orientation and PSF differences between the ASL and structural images into account (Cardoso et al., 2015).

3.3.7. Analysis mask creation

For the statistics performed in section 4.3 - as well as for any voxel-based group statistics - an analysis mask aims to exclude voxels outside the brain or voxels with artifactual signal (e.g. macro-vascular, signal dropout) and restrict the analysis to regions with sufficient SNR and/or statistical power. This also avoids over-penalizing statistical power by family-wise error corrections. The susceptibility and field-of-view (FoV) masks are combined in section 4.3 into a group mask. The vascular masks influence on the ASL measurement (Supplementary Fig. 9). For these reasons, PVC (Aslani et al., 2009) is essential in studies that aim to differentiate structural changes (e.g. atrophy) from perfusion changes (e.g. related to neurovascular coupling) (Stekete et al., 2016). Several PVC algorithms have been proposed (Chappell et al., 2010; Zhao et al., 2017; Aslani et al., 2008; Liang et al., 2013), which assume locally homogeneous GM and WM CBF. In some studies, GM volume is used as a covariate in the statistical analysis (Chen et al., 2011a). Note that while PVC, in theory, corrects only for the PV effects and takes into account the intra- and inter-subject variability of the GM-WM CBF ratio, GM covariation can additionally affect the estimated physiological correlation between GM CBF and GM volume (Petr et al. 2018a).

ExploreASL employs two versions of PVC, both based upon the most frequently used PVC, i.e. linear regression (Aslani et al., 2008): 1) a 3D Gaussian instead of a 2D flat kernel (default, referred to as “voxel-wise”) (Oliver, 2015), or 2) computing PV-corrected CBF within each anatomical or functional region of interest (ROI) separately instead of using a kernel. Whereas the voxel-wise option allows further voxel-based analysis (VBA), the ROI-based PVC is in theory beneficial for a ROI-based analysis as effectively the kernel-size is selected based upon the anatomical ROI, which should be less sensitive to local segmentation errors. Moreover, this avoids cross-talk between ROIs. It still needs to be investigated how to define regions of optimal shape with respect to PVC performance, which depends on the spatial uniformity and SNR of the GM and WM CBF, and partial volume distributions within the ROI. To evaluate the effects of PVC, ExploreASL exports CBF maps and ROI values both with and without PVC.

For proper PVC or ROI definition, the true acquisition resolution - which often differs from the reconstructed voxel size - needs to be taken into account (Petr et al., 2018). This is especially important for 3D readouts, where the through-plane PSF can be up to 1.9 times the nominal voxel-size (Vidorreta et al. 2013, 2014). Effects such as motion (Petr et al., 2016) and scanner reconstruction filters can contribute to further widening of the PSF of the final image. ExploreASL by default uses previously estimated true acquisition resolutions (Vidorreta et al. 2013, 2014; Petr et al., 2018) and can optionally perform a data-driven spatial resolution estimation (Petr et al., 2018) that is generalizable to all ASL acquisitions. Contrary to alternative PSF estimations based on temporal noise autocorrelation (Cox, 2012) or simulations of the acquisition PSF (Vidorreta et al. 2013, 2014), this method does not require time series and inherently accounts for other sources of blurring (e.g. smoothing by motion and/or image processing) and is applicable without having detailed information about the sequence parameters needed to calculate the resolution from the k-space trajectory. However, this method requires further validation, especially in the presence of ASL image artifacts.

Lastly, the GM/WM maps obtained from the high-resolution structural images need to be downsampled to the ASL resolution before they are used for PVC or for ROI delineation in native space. A trivial interpolation to lower resolution may introduce aliasing, which can be addressed by applying a Gaussian filter - or a convolution with the PSF, if the PSF is known - prior to downsampling (Cardoso et al., 2015). This is important to note that the ASL image often has an anisotropic resolution and may be acquired at a different orientation compared to the structural image. To correct for this effect, ExploreASL pre-smooths the structural images with a Gaussian kernel of which the covariance matrix takes the orientation and PSF differences between the ASL and structural images into account (Cardoso et al., 2015).
are applied subject-wise to reflect the individual differences in vascular anatomy.

First, regions outside of the ASL FoV are identified, as whole brain coverage is not always achieved (Supplementary Fig. 10). Second, a mask is created to remove voxels with intravascular signal. While intravascular signal - resulting from an incomplete tissue arrival of labeled spins - can be clinically useful (Mutsaerts et al., 2017, 2020), it biases regional CBF estimates. The relatively large local temporal variability of such vascular artifacts can be detected in time series, in multi-post-labeling delay (PLD) acquisitions (Chappell et al., 2010) or by an independent component analysis (ICA) (Hao et al., 2018). ExploreASL uses a pragmatic vascular artifact detection approach that is suitable for both single and multi-PLD ASL images. It identifies clusters of negative apparent CBF (Maumet et al., 2012) and voxels with extreme positive apparent CBF (Supplementary Fig. 11). First, spatially connected subzero voxels are grouped into clusters. The average CBF of each cluster is obtained, and clusters with significant negative mean CBF are isolated (median - 3 median absolute difference (MAD)) within all subzero CBF voxels (Supplementary Fig. 11b). Likewise, voxels with extreme positive signal are detected by having intensity of more than median +3 MAD within all positive CBF voxels (Supplementary Fig. 11b). One potential caveat of masking out vascular voxels is the violation of the stationarity parameter of parametric voxel-wise statistics. While excluding voxels with high signal can violate the stationarity criterion of the ASL signal, there is currently no validated method that would be able to reliably estimate the perfusion and vascular signal contribution in such voxels from single-PLD data. Note that the detected voxels with negative or extreme positive signal may also stem from non-vascular artifacts such as head motion.

Second, regions with susceptibility signal-dropout artifacts are removed. Regions frequently having low SNR for 2D EPI and 3D GRASE ASL include the orbitofrontal cortex near the nasal sinus and the inferior-medial temporal gyrus near the mastoid air cavities. Therefore, the option implemented in ExploreASL is to use sequence-specific template masks obtained from previous population analyses, after which individual masks are restricted to (pGM + pWM) > 0.5 to remove voxels outside the brain. Further development is needed to create masks that take individual anatomical differences in skull and air cavities into account. Noteworthy, ExploreASL applies this analysis mask only for analyses, not for visual QC.

3.4. Population module

This module prepares output for visual QC and creates group-level results for statistical analyses. Whereas the above-described Structural and ASL modules perform image processing on the individual level, this module performs its analysis on multiple-subjects and/or multi-time point level. For this purpose, ASL images are transformed into standard space using the T1w transformation fields smoothed to the effective spatial resolution of ASL. For transformation of all intermediate and final images, all previous spatial transformations are merged into a single combined transformation to minimize accumulation of interpolation artifacts through the pipeline. Partial volumes of GM and WM obtained from anatomical images are multiplied by the Jacobian determinants of the deformation fields - a.k.a. modulation - to account for voxel-volume changes when transforming to standard space (Ashburner and Friston, 1999). The standard space used by ExploreASL is the 1.5 × 1.5 × 1.5 mm3 IXI555-MNI152 space (Gaser, 2009), which is a refined version of the MN152 space with additional geodesic shooting-based template creation for the IXI555 population (Ashburner and Friston, 2011).

3.4.1. Template creation

Population templates can reveal population- or sequence-specific perfusion patterns that are not visible on the individual level. ExploreASL generates the mean and between-subject standard deviation (SD) images for the total study population and, optionally, for different sets (e.g. different centers/sequences/cohorts) within the study (Fig. 2). In addition to CBF itself, auxiliary images (e.g. MO), intermediate images (e.g. mean control images), or QC images (e.g. temporal SD) can provide a valuable overview of the data, for example when comparing data originating from different centers.

3.4.2. Multi-sequence equalization

Quantitative CBF images can differ between centers because of a number of hardware, labeling, and readout choices implemented by different MRI vendors and/or laboratories (Deibler et al., 2008; Heijtje et al., 2014; Alsop et al., 2015; Jack et al., 2010). Some of these differences can be accounted for, as detailed in the previous sections. However, until a more robust procedure is devised - e.g. the use of a flow phantom (Oliver-Taylor et al., 2017) - a pragmatic approach is required to remove the remaining CBF quantification differences between sequences, scanner types, and centers (Mutsaerts et al., 2018, 2019). ExploreASL optionally performs spatially varying intensity normalization by computing a smooth average CBF bias field for each ASL sequence (Supplementary Fig. 12). Noteworthy, this step assumes that the demographics and expected (patho-)physiological effects are equally distributed across the subjects scanned with each sequence, scanner, and/or site. To fulfill this assumption, it is advisable to estimate these biasfields based on images of healthy controls (Mutsaerts et al., 2018). However, this is often not feasible due to relative high physiological variability of CBF and relative small size of the control groups. Using ASL images from all participants including patients is a viable alternative (Mutsaerts et al., 2019), provided that the distribution of demographics and expected patho-(physiology) effects on perfusion are comparable between sequences. The final CBF images in the standard space are smoothed with an 8 × 8 × 8 mm full-width at half-maximum Gaussian, and averaged to create a sequence/scanner type/site-specific mean CBF image. These site-specific CBF images averaged over all subjects are intensity normalized to GM CBF of 60 mL/100g/min and further averaged to create a general mean CBF image of all sites. The site-specific bias field is calculated by dividing the general mean with the site-specific mean CBF image. The individual CBF images of each site are multiplied by their site-specific bias to adjust for between-site differences (Mutsaerts et al., 2018).

3.4.3. ROI statistics

In ExploreASL, ROI masks are created by combining existing atlases with individual GM and WM masks. The GM atlases currently implemented are: (i) MNI structural (Mazziotta et al., 2001), (ii) Harvard-Oxford (Desikan et al., 2006), and (iii) Hammers (Hammers et al., 2002). A deep WM atlas is created by eroding the SPM12 WM tissue class by a 4 voxel sphere (i.e. 6 mm), to avoid signal contamination from the GM (Mutsaerts et al., 2014a). Other existing, or custom, atlases can be easily applied. The Online Brain Atlas Reconciliation Tool (OBART) at obart.brainarchitecture.org (Bohland et al., 2009) provides an overview of the overlap and differences between atlases. For each ROI, statistics are calculated separately within the left and right hemisphere, as well as for the full ROI, both with and without PVC, and both in the standard space and subject’s native space. The same CBF statistics are also calculated for user-provided ROIs and lesion masks, as well as for the 25 mm margin around the ROI/lesion (Moghaddasi et al., 2015), for the ipsilateral hemisphere excluding the lesion, and for the same three masks at the contralateral side. Subject-specific ROI and lesion masks are treated the same, except for the fact that lesion masks are also used for the cost function masking (see section 3.3). Individual vascular masks are used to exclude regions with intra-vascular signal (see section 3.7) from CBF statistics, but not from spatial CoV statistics.

Finally, all masks are intersected with a group-level analysis mask, created from the individual analysis masks created in section 3.7. Individual differences of these analysis masks can be caused by differences in head position, FoV, and nasal sinus size. To limit the effects of this mask heterogeneity on statistical analyses, ExploreASL creates a group-level analysis mask from standard-space voxels present in at least 95% of the
individuals masks (Supplementary Fig. 10).

3.4.4. Quality control

On a participant level, ExploreASL outputs QC parameters in a JSON file and provides unmasked images in standard space for visual QC, for both intermediate and final images (Supplementary Fig. 13) to detect technical failure, outliers and artifacts. QC parameters are also obtained by comparing individual ASL images with an atlas, a group average, or an average from a previous study. Whole-brain and regional differences larger than 2–3 SD are indicated and should be visually inspected. Deviations can hint to software updates or different scanners and, if not accounted for, can lead to low power of the statistical analyses (Chenevert et al., 2014). All QC parameters and images are also collected in a PDF file (Fig. 3, Supplementary Table 2). While these QC parameters can be helpful in detecting artifacts and/or protocol deviations, their use has not yet been validated, and the normal and abnormal range for each of the parameters still need to be determined.

4. Methods

We illustrate the ExploreASL image processing results and reproducibility for three populations with similar 2D-EPI PCASL protocols: perinatally infected HIV children, healthy adults, and elderly with mild cognitive complaints, from the NOVICE (Blokhuis et al., 2017), the Sleep (Elvsåshagen et al., 2019), and the European Prevention of Alzheimer’s Dementia (EPAD) studies (Ritchie et al., 2016), respectively (Supplementary Table 3). All three studies adhered to the Declaration of Helsinki and were approved by the local ethics committees (Academic Medical Center (AMC) in Amsterdam, Norwegian South East Regional Ethics Committee, and VU Medical Center Amsterdam and University of Edin-burgh, respectively). Written informed consent was obtained from all participants (or parents of children younger than 12 years for NOVICE). Each participant of the Sleep study received NOK 500 for participation.

ASL studies processed at different centers typically use different OS or software versions possibly affecting the ability to compare CBF values between studies. Additionally, longitudinal studies may be subject to updates in analysis hardware or software. The performance of image processing should thus be comparable irrespective of the used analysis software version and/or OS to allow data pooling and comparison between studies. Here, we investigated the image processing reproducibility between OSes and Matlab versions for the intermediate and final pipeline results of the three previously acquired datasets, without and with the ExploreASL-specific modifications of the SPM12, CAT12, and LST source code (modifications described below). To this end, a single participant from each study was analyzed: with the lowest GM volume from NOVICE and EPAD (GM/ICV ratio 0.41 and 0.33, respectively), and the highest GM volume (GM/ICV ratio 0.55) from the Sleep study. These three datasets were processed at two centers with the following combinations of OS and MATLAB version, twice at each center for each of the combinations: Linux-2018b (HZDR, Dresden, Germany; Linux server, 2.1 GHz Intel Xeon 6130, Ubuntu 5), Windows-2015a and 2018b (Amsterdam UMC, The Netherlands; Dell Alienware laptop, 2.9–4.3 GHz Intel I7-7820HK, Windows 10 Version, 1903). After each pipeline step, the between-system reproducibility was obtained as a difference of the image.
xASL report: Philips2DEPI_EPAD-001 ASL_1

### Structural
- FLAIR CNR GM WM Ratio: 0.002365
- FLAIR EFC bits: 1.866e+04
- FLAIR FBER WMref Ratio: 5.876
- FLAIR Mean Al Perc: 7.012
- FLAIR SD Al Perc: 5.998
- FLAIR SD WMref: 147.7
- FLAIR SNR GM Ratio: 6.526
- FLAIR WMH_v
- FLAIR WMH_vol mL: 94.406
- FLAIR WMref_vol mL: 10.58
- ID: EPAD-001
- T1w CNR GM WM Ratio: 20.2782
- T1w CSF_vol mL: 603.8
- T1w EFC_bits: 1.37e+04
- T1w FBER WMref Ratio: 178.3
- T1w GM JCV Ratio: 0.335
- T1w GM_vol mL: 924
- T1w JCV_vol mL: 1805
- T1w JOR Perc: 83.57
- T1w LR flip YesNo: 0
- T1w Mean Al Perc: 7.012
- T1w SD Al Perc: 5.998
- T1w SD WMref: 31.64
- T1w SNR GM Ratio: 25.06
- T1w WM_vol mL: 597
- T1w WMref_vol mL: 15.09
- Version CAT12: 1363
- Version ExplotaASL: 1.00
- Version LST: 2.0.15
- Version Matlab: 8.5 R2015a
- Version SPM12: 7219

### ASL
- Al Perc: 40.42
- ASL CoveragePerc: 92.02
- ASL tSNR CSF Ratio: 0.0403
- ASL tSNR GM WM Ratio: 0.1819
- ASL tSNR GM WM WMref Ratio: 0.2626
- ASL tSNR GM Ratio: 0.3238
- ASL tSNR PhysioC2Ratio: 0.7212
- ASL tSNR Slope Corr: 0.9999346
- ASL tSNR WM Ratio: 0.748
- ASL tSNR WMref Ratio: 0.1205
- Acquisition Time h:mm:ss: 14:04:49.29
- BackGrSupPulses: 2
- CBF GM Median mL/100g/min: 32.06
- CBF GM PVC2 mL/100g/min: 51.24
- CBF GM WM Ratio: 4.751
- CBF WM PVC2 mL/100g/min: 10.79
- Echo Time ms: 10.49
- ID: EPAD-001 ASL_1
- Initial PLD ms: 2055
- LR flip YesNo: 0
- Labeling Duration ms: 1850
- Labeling Type: CASL
- MRSlope Slope: 0.0007
- Matrix: 64 64 36 60
- Mean SSIM Perc: 69.13
- Motion Excl Perc: 0
- Motion Mean mm: 0.1479
- Peak SNR Ratio: 674.7
- RMSE Perc: 59.37
- Repetition Time ms: 4800
- Rescale Lin Intercept: 0
- Rescale Slope: 2.712
- Rescale Slope-Origin: 2.712
- Rigid Body Control mm: 150.7
- Stoner Relaxation Time ms: 36.53
- Spatial CoV GM Perc: 96.69
- Vendor: Philips
- Version FSL: 6.0.1
- Voxel Size mm: 3.4375 3.4375 4.5
- nRMSE Perc: 23.26

(caption on next page)
intensities and orientation between the NIfTIs of the two compared systems. The image intensity reproducibility was calculated as the median voxel-wise relative intensity difference (Kurth et al., 2015), whereas the image orientation reproducibility was calculated as the mean voxel-wise net displacement vector in real-world coordinates (Power et al., 2012). These were calculated for T1w with GM segmentation, FLAIR with WMH segmentation, M0, quantified CBF, GM partial volume in ASL native space (pGMASL), and PV-corrected GM CBF.

More complex calculations involving floating-point arithmetic operations, e.g. matrix inversions, can produce different results between OSes and MATLAB versions in the last digits. As randomly seeded pseudo-random number generators are used for some optimization processes, results can differ upon re-run even on the same system. These minimal differences can accumulate in iterative algorithms such as segmentation and registration, and propagate across the pipeline. To mitigate these effects, during the process of implementing and using the pipeline for previous clinical studies, we modified parts of the SPM12, CAT12, and LST toolboxes: e.g. using the MATLAB '/' operator for solving...
a system of linear equations instead of calculating a matrix inversion, providing a separate C++ implementation for convolutions, and/or rounding some calculations to 15 significant digits.

Finally, we reran the full pipeline for the three example populations, but with the Explore/ASL-specific features disabled. This would be similar to running a pipeline based on SPM12 and DARTEL without lesion filling and ENABLE, with M0-T1w registration only and the default M0 image processing, which comprise the basic image processing steps that are recommended in the ASL consensus paper. Using the standard space partial volume corrected GM CBF values obtained with the two pipelines, we compare the between-group difference and the correlation of GM CBF with age and sex.

5. Results

Running time for a single EPAD participant took 22:11, 4:34, and 0:40 min for the Structural, ASL, and Population modules, respectively (27:25 min in total). On the ‘low quality’ setting, the same processing...
took 7:30 min, 2:24, and 0:34 respectively (10:28 min in total) (Windows-2018b). Fig. 4 shows differences between populations or sequences on the ExploreASL population-specific parametric maps. While the GM CBF was highest in the pediatric and lowest in the geriatric population (Fig. 4c), both the between-subject CoV and within-scan temporal SD were comparable in these populations and lowest in the healthy adults (Fig. 4d-e). The temporal SD (Fig. 4e) was high in vascular regions and highest around the ventricles in the pediatric dataset, due to a 2D EPI fat-saturation related artifact. Despite these differences, the temporal SNR appeared relatively comparable (Fig. 4f), albeit slightly higher for the pediatric population. The average mean control images (Fig. 4g) showed subtle differences in background suppression efficacy, as different tissue contrast and inferior-superior background suppression efficiency gradient. Only slight differences in ventricle and sulci size were visible between the pediatric and geriatric population (Fig. 4a) confirming satisfactory performance of spatial normalization.

All three datasets showed zero difference when the pipeline was repeated twice on the same system. When comparing OSes only - Linux-2018b vs Windows-2018b - the structural module showed final voxel-wise differences of 0.77% pGM in NOVICE and 1.74% WMH in EPAD (2018b vs Windows-2018b - the structural module showed satisfactory performance of spatial normalization.

When comparing ExploreASL with the typical consensus pipeline - i.e. ExploreASL with all ExploreASL-features turned off - the within-cohort coefficients of variation (SD/mean) were 39.6% vs. 41.4% (Novice, n = 28), 11.0% vs. 15.4% (Sleep, n = 38), and 16.7% vs. 22.5% (EPAD, n = 75) for the maps computed by ExploreASL and the typical consensus pipeline, respectively. Subsequently, the between-cohort GM CBF difference for the maps processed by the ExploreASL pipeline (one-way ANOVA, F = 66.2, p = 10^-21, n = 141) was stronger compared to the typical consensus pipeline (F = 28.0, p = 10^-11, n = 141). The explained variance of age and sex for the three combined populations was larger for the CBF maps processed by ExploreASL (R^2 = 0.379, F = 60.1, p = 10^-21, n = 141) than with the typical pipeline (R^2 = 0.270, F = 37.1, p = 10^-14, n = 141).

6. Discussion and future directions

Here, we reviewed many of the most salient ASL image processing choices, and their implementation in ExploreASL version 1.0.0. We demonstrated the software’s functionality to review individual cases as well as population-average images for quality control. Our findings show that between-system computing differences can lead to voxel-wise CBF quantification differences of up to 5.7% on average for the total GM, which were reduced to 1.7% by addressing implementation differences of complex floating-point operations between MATLAB versions and OSes. This may especially be beneficial for multi-center studies or for pooling multiple ASL studies to attain sample sizes required for the discovery of subtle (patho-)physiological perfusion patterns.

On the MRI scanner consoles of GE, Philips, and Siemens, CBF quantification is available according to the consensus recommendation (Alsop et al., 2015), although they differ in their image processing and quantification parameters. While this suffices for visual reading, offline image processing is recommended to optimize the image quality and extract regional statistics with respect to an anatomical reference. Several other ASL image processing pipelines are publicly available and free for academic use, each providing specific features. The first publicly available pipeline ASLtbx quantifies CBF of various ASL sequences (Wang et al., 2008) and features customized motion correction and advanced outlier detection (Dolui et al., 2017); ASL contains a graphical user interface (GUI) with an interface for population analyses, and generates statistical reports (Mato Abad et al., 2016); the ASL toolbox is a MATLAB- and SPM-based command-line tool (Homan et al., 2010), ASL-MRiCloud features a web interface with an automated cloud solution (Li et al., 2019); ASL-QC handles multiple vendors and provides QC metrics (not published); BASIL (Chappell et al., 2009) uses a Bayesian approach for the quantification and PVC of multi-TI (Chappell et al., 2011), QUASAR, and time-encoded ASL data, thus offering the most comprehensive quantification (Chappell et al., 2010); CBFBIR offers an online data repository with online image processing (Shin et al., 2016); Functional ASL (Functional MRI Laboratory, University of Michigan) and fMRI Grocer (Center for Functional Neuroimaging, University of Pennsylvania)) (Zhu et al., 2010) are SPM toolboxes that process both functional ASL and BOLD MRI; GIN MRI performs separate control and label realignment and automatically excludes outliers and volumes with strong motion (unpublished); MilxASL features spatial and temporal denoising (Fazlollahi et al., 2015); MJD-ASL is implemented into ‘cranial cloud’, addresses noise concerns and processes cerebral blood volume (Manus Donahue, Vanderbilt University Medical Center); NiftyFit supports quantification of other MRI sequences as IVIM, NODDI, and relaxometry (Melbourne et al., 2016); VANDIRE is Python-based, has a scanner console plugin and allows flow territory mapping from vessel-encoded ASL (VU e- Innovations) (Arteaga et al., 2017).

ExploreASL has focused on optimizing the processing for clinical studies that have diverse clinical populations, hardware, and sequences used by allowing the import and processing of different sequences of different vendors in a single study. Providing an integrated module for structural image processing, population statistics, and QC is, among other things, essential for processing large multi-center studies. Other strengths of ExploreASL include compatibility with the most-used OSs and its tested between-system reproducibility of image processing. Moreover, ExploreASL is available through GitHub with a growing team of international contributors. We follow the recommendations of the Committee on Best Practices in Data Analysis and Sharing (COBIDAS) (Nichols et al., 2017) by improving the between-system reproducibility, including complete reporting of all facets of a study and the provenance, having standardized source-code headers, and following the best coding practice. ExploreASL was built upon freely available Matlab-based toolboxes that perform well in a wide array of cases, rather than opting for solutions with optimal performance in specific cases but not applicable in general.

Although ExploreASL allows custom labeling efficiency values and global CBF calibration, it does not estimate labeling efficiency. While literature values for labeling efficiency (Dai et al., 2008) may suffice in many clinical cases (Heijtel et al., 2014), individual correction can be beneficial for specific populations (Václaví et al., 2019). For these, phase-contrast MRI can improve the CBF quantification by a) calibrating CBF based on total flow through the brain-feeding arteries (Aslan and Lu, 2010; Ambarki et al., 2015), or b) modeling the labeling efficiency based on the velocity in the labeling plane (Václaví et al., 2016). Compared to ASL, drawbacks of the phase-contrast MRI include its lower reproducibility for whole brain CBF estimates (Dolui et al., 2016), and its lower agreement with PET (Puig et al., 2019). Moreover, an automatic implementation requires good data quality, perpendicular placement of the labeling plane to the vessel, and the absence of vessel tortuosity, conditions that are rarely met in clinical datasets. Future solutions may be provided by new sequences under development, which allow direct labeling efficiency measurements during the ASL acquisition (Chen et al., 2018; Lorenze et al., 2018).

Several additional features are scheduled for future releases, including full BIDS support (Gorgolewski et al., 2016); support for Hitachi and Canon datasets; unit testing to ensure stability of the pipeline through the continuous development; inclusion of WM atlases for
extended WM analysis; a GUI for easier configuration and execution; quantification of advanced ASL schemes such as velocity- and acceleration-selective ASL (Schmid et al., 2015) and integration of the BASIL toolbox to allow multi-PLD and time-encoded sequence quantification (Chappell et al., 2009); and support for individual-center calibration, e.g., using the recently introduced Quantitative ASL Perfusion Reference (QASPER) (Oliver-Taylor et al., 2017) phantom (Gold Standard Phantoms, London, UK). Although ExploreASL’s computation times are moderate for research purposes, a clinical scanner implementation would benefit from parallelization on graphical processing units (GPUs) to provide robust automatic QC within clinical scanning time (e.g. <5 min). Another improvement would be the investigation of the effect of image processing choices, as well as the availability of physiological and quantification parameters for different populations (Fazlollahi et al., 2015). This would allow for the incorporation of quantification confidence intervals in the output of ExploreASL. For anonymization purposes, the face can be removed from the structural scans (Nichols et al., 2017; Leung et al., 2015) using a defacing algorithm such as the one implemented in SPM12, but further testing is required to verify that the analysis is not affected (de Sitter et al., 2017b). Statistical analyses can be biased for populations with large inter-subject differences in their deformations, e.g., developing brains or a wide range of atrophy. Dedicated templates are typically used for infants to ensure proper segmentation and normalization (Shi et al., 2011). For older children, the use of a dedicated template is still advised (Sanchez et al., 2012), although adult templates are often sufficient. Further errors in deformations and volume changes can be encountered when stretching pediatric brains to an adult standard space. Either, such deformations need to be accounted in both the volumes and CBF maps, or the analysis has to be performed in standard space, knowing that the pGM thresholds can be different with different total brain volumes as discussed in Section 3.6. A more optimal solution is to use the CerebroMatic toolbox (Wilke et al., 2017) is a tool that accounts for this bias and will be incorporated in future releases of ExploreASL. Finally, we intend to implement ExploreASL as a cloud solution and a plugin for scanner workstation to allow seamless ExploreASL image processing in clinical routine.

There are several methods that we decided not to include in ExploreASL: thresholding the M0 or the mean control image to identify signal dropout (Wang, 2014) or masking them with FSL BET (Smith, 2002), as this may fail with background suppression, blurred 3D acquisitions, poor ASL-M0 registration, or a strong bias field (Mutsaerts et al., 2018). If multiple individual unsubtracted control-label images are available, a mask could be created based on the tSNR of the control or label images. However, time series are not always available, and the tSNR may be biased by the presence of (patho-)physiological signal changes and head motion. Currently, no consensus exists on whether the M0 should be quantified separately for GM and WM tissue types, especially for longer repetition time with a distinct GM-WM contrast. The M0 quantification can potentially be improved by using tissue specific quantification parameters - such as blood-brain partition coefficients λ and tissue relaxation times (Çavuşoğlu et al., 2009), and/or partial volume correction (Ahlgren et al., 2018). However, this can induce quantification errors in cases of suboptimal ASL-M0 registration. While ExploreASL is designed for analyses in the human brain, many of its methods and principles could be translated to pre-clinical ASL studies and/or studies outside the brain (Nery et al., 2020). The main challenge in applying ExploreASL outside the brain is the dependency of the structural module on human brain templates. Currently, when running only the ASL module on non-brain ASL data, only motion correction, rigid-body registration, and quantification can be performed. Note also that the quantification model is not adapted for pre-clinical or non-brain imaging.

Image processing techniques that require validation include: using the UNWARP toolbox for simultaneous motion and susceptibility deformation correction (Andersson et al., 2001), using temporal information for artifact removal through the application of an independent-component analysis (ICA) (Wells et al., 2010; Hao et al., 2018; Zhu et al., 2018) or using respiratory and cardiac signal (Restom et al., 2006) or data-driven outlier rejection (Maumet et al., 2014; Tan et al., 2009; Dohui et al., 2017), spatial denoising - once validated under realistic conditions (Spann et al., 2017; Wells et al., 2010; Bibic et al., 2010; Liang et al., 2015), obtaining GM-WM segmentations from fractional signal modeling of the magnetization recovery profile acquired with a Look-Locker readout (Petit et al., 2013; Ahlgren et al., 2014), and using the BBR method for motion correction or registration (Greve and Fischl, 2009). ExploreASL offers the option to register to the atlas of spatially normalized mean M0 images, CBF images, or CBF images with a high number of vascular artifacts created for different vendors and ASL sequences from previously processed large ASL datasets. Finally, we aim to improve the inter-center reproducibility even further.

7. Conclusion

ExploreASL is a versatile pipeline that performs well on a wide-range of diseases, including datasets with lesions, allows flexible parameter definition, and a quick exploration of datasets and QC images of each pipeline step in the same space. We made the pipeline available at www.ExploreASL.org. We anticipate that ExploreASL will allow for more flexible collaboration amongst clinicians and scientists, help to achieve the consensus standards for ASL processing sought by the OSIPI, facilitate validation of ASL image processing approaches, and accelerate translation to clinical research and practice.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Henk J.M.M. Mutsaerts: Conceptualization, Formal analysis, Methodology, Resources, Software, Validation, Visualization, Writing - original draft, Writing - review & editing. Jan Petr: Conceptualization, Formal analysis, Methodology, Software, Validation, Visualization, Writing - original draft, Writing - review & editing. Paul Groot: Data curation, Methodology, Software, Writing - review & editing. Pieter Vandemaele: Data curation, Methodology, Validation, Software, Writing - review & editing. Andrew D. Robertson: Data curation, Writing - review & editing. Lena Vaclavici: Methodology, Writing - review & editing. Ingel Groote: Conceptualization, Resources, Writing - review & editing. Hugo Kuijf: Conceptualization, Data curation, Writing - review & editing. Fernando Zelaya: Resources, Writing - review & editing. Owen O’Daly: Investigation, Resources, Writing - review & editing. Saima Hilal: Data curation, Investigation, Resources, Writing - review & editing. Alle Meije Wink: Resources, Software, Writing - review & editing. Ilse Kant: Conceptualization, Writing - review & editing. Matthew W.A. Caan: Investigation, Software, Writing - review & editing. Catherine Morgan: Writing - review & editing. Jeroen de Bresser: Conceptualization, Writing - review & editing. Elisabeth Lysvik: Investigation, Writing - review & editing. Anouk Schrantee: Investigation, Writing - review & editing. Astrid Bjørnebekk: Data curation, Resources, Writing - review & editing. Patricia Clement: Data curation, Resources, Writing - review & editing. Zahra Shirzadi: Conceptualization, Data curation, Methodology, Software, Writing - review & editing. Joost P.A. Kuijer: Conceptualization, Methodology, Software, Writing - review & editing. Viktor Wootschel: Data curation, Resources, Writing - review & editing. Udunna C. Anazodo: Investigation, Writing - review & editing. Dassja Pajkrt: Investigation, Resources, Writing - review & editing. Edo Richard: Resources, Writing - review & editing. Reinoud P.H. Bokkers: Conceptualization, Funding acquisition, Writing - review & editing. Liesbeth Reneman: Investigation, Writing - review & editing. Mario
Acknowledgments

This project has received support from the following EU/EFPIA Innovative Medicines Initiatives (1 and 2) Joint Undertakings: EPAD grant no. 115736, AMYPAD grant no. 115952. Additionally, this work received support from the EU-EFPIA Innovative Medicines Initiatives Joint Undertaking (grant No 115952). HM is supported by Amsterdam Neuroscience funding. FB and XG are supported by NIHR funding Joint Undertaking (grant No 115952). HM is supported by Amsterdam Neuroscience funding. MC received funding from NWO. MC received funding from the Netherlands Organization for Scientiﬁc Research (NWO). MC received funding from the Engineering and Physical Sciences Research Council UK (EP/P012361/1), and is a shareholder of Nico.lab BV, Amsterdam, The Netherlands. The Wellcome Centre for Integrative Neuroimaging is funded by the Netherlands Organisation for Scientiﬁc Research (NWO).}

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2020.117031

References


Vidorreta, Marta, Balteau, Evelyne, Wang, Ze, De Vita, Enrico, Pastor, María a., Thomas, David L., Detre, John A., Fernández-Seara, María a., 2014. Evaluation of segmented 3D acquisition schemes for whole-brain high resolution arterial spin labeling at 3 T. NMR Biomed. 27 (11), 1387–1396.


