The European Prevention of Alzheimer's Dementia (EPAD) MRI Dataset and Processing Workflow

Lorenzini Luigi ^{1*}, Ingala Silvia ^{1*}, Wink Alle Meije ¹, Kuijter Joost PA ¹, Wottschel Viktor ¹,
Dijsselhof Mathijs ¹, Sudre Carole H ^{2,3,4,5}, Haller Sven ^{6,7}, Molinuevo José Luis ^{8,9,10,11}, Gispert Juan
Domingo ^{9,10,11,12}, Cash David M ^{3,13}, Thomas David L ^{3,14,15}, Vos Sjoerd B ^{4,14}, Prados Ferran ^{16,17,18}, Petr
Jan ^{1,19}, Wolz Robin ^{20, 21}, Palombit Alessandro ²⁰, Schwarz Adam J ²², Chételat Gael ²³, Payoux Pierre
^{24,25}, Di Perri Carol ²⁶, Wardlaw Joanna ^{26,27}, Frisoni Giovanni B ^{28,29}, Foley Christopher ³⁰, Fox Nick C ³,
Ritchie Craig ³¹, Pernet Cyril ²⁶, Waldman Adam ^{26,32}, Barkhof Frederik ^{1,33}, Mutsaerts Henk JMM ^{1,34}

for the EPAD consortium

* These two authors contributed equally to this study

1. Dept. of Radiology and Nuclear Medicine, Amsterdam University Medical Center, Amsterdam Neuroscience, Amsterdam, The Netherlands;
2. MRC unit for Lifelong Health and Ageing at UCL, London, UK
3. Dementia Research Centre, Department of Neurodegenerative Disease, UCL Queen Square Institute of Neurology, London, UK ;
4. Centre for Medical Image Computing, University College London, London, UK;
5. School of Biomedical Engineering & Imaging Sciences, King's College London, UK
6. CIRD Centre d'Imagerie Rive Droite, Geneva, Switzerland;
7. Department of Surgical Sciences, Radiology, Uppsala University, Uppsala, Sweden ;
8. Barcelonaβeta Brain Research Center (BBRC), Pasqual Maragall Foundation, Barcelona, Spain;
9. CIBER Fragilidad y Envejecimiento Saludable (CIBERFES), Madrid, Spain;
10. IMIM (Hospital del Mar Medical Research Institute), Barcelona Spain,
11. Universitat Pompeu Fabra, Barcelona, Spain;
12. CIBER Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Madrid, Spain;
13. UK Dementia Research Institute, University College of London, London, UK;
14. Neuroradiological Academic Unit, UCL Queen Square Institute of Neurology London, UK;
15. Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology, London, UK;
16. Nuclear Magnetic Resonance Research Unit, Queen Square Multiple Sclerosis Centre, UCL Queen Square Institute of Neurology, London, United Kingdom;
17. Department of Medical Physics and Biomedical Engineering, Centre for Medical Image Computing, University College London, London, United Kingdom;
18. e-Health Centre, Universitat Oberta de Catalunya, Barcelona, Spain;
19. Helmholtz-Zentrum Dresden-Rossendorf, Institute of Radiopharmaceutical Cancer Research, Dresden, Germany;
20. IXICO, London, UK;
21. Imperial College London, London, UK;
22. Takeda Pharmaceuticals Ltd., Cambridge, MA, USA
23. Université de Normandie, Unicaen, Inserm, U1237, PhIND "Physiopathology and Imaging of Neurological Disorders", Institut Blood-and-Brain @ Caen-Normandie, Cyceron, 14000 Caen, France.
24. Department of Nuclear Medicine, Toulouse CHU, Purpan University Hospital, Toulouse, France
25. Toulouse NeuroImaging Center, University of Toulouse, INSERM, UPS, Toulouse, France
26. Centre for Clinical Brain Sciences, The University of Edinburgh, Edinburgh, UK;
27. UK Dementia Research Institute at Edinburgh, University of Edinburgh, UK

28. *Laboratory Alzheimer's Neuroimaging & Epidemiology, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy*
 29. *University Hospitals and University of Geneva, Geneva, Switzerland*
 30. *GE Healthcare, Amersham, UK*
 31. *Centre for Dementia Prevention, The University of Edinburgh, Scotland, UK*
 32. *Department of Brain Sciences, Imperial College London, London, UK;*
 33. *Institute of Neurology and Healthcare Engineering, University College London, London, UK;*
 34. *Ghent Institute for Functional and Metabolic Imaging (GfMI), Ghent University, Ghent, Belgium*

Corresponding author:

Luigi Lorenzini, MSc

l.lorenzini@amsterdamumc.nl

Dep. of Radiology and Nuclear Medicine, PK -1

De Boelelaan 1117

1081 HV Amsterdam, The Netherlands

Abstract

The European Prevention of Alzheimer Dementia (EPAD) is a multi-center study that aims to characterize the preclinical and prodromal stages of Alzheimer's Disease. The EPAD imaging dataset includes core (3D T1w, 3D FLAIR) and advanced (ASL, diffusion MRI, and resting-state fMRI) MRI sequences.

Here, we give an overview of the semi-automatic multimodal and multisite pipeline that we developed to curate, preprocess, quality control (QC), and compute image-derived phenotypes (IDPs) from the EPAD MRI dataset. This pipeline harmonizes DICOM data structure across sites and performs standardized MRI preprocessing steps. A semi-automated MRI QC procedure was implemented to visualize and flag MRI images next to site-specific distributions of QC features — i.e. metrics that represent image quality. The value of each of these QC features was evaluated through comparison with visual assessment and step-wise parameter selection based on logistic regression. IDPs were computed from 5 different MRI modalities and their sanity and potential clinical relevance were ascertained by assessing their relationship with biological markers of aging and dementia.

The EPAD v1500.0 data release encompassed core structural scans from 1356 participants 842 fMRI, 831 dMRI, and 858 ASL scans. From 1356 3D T1w images, we identified 17 images with poor quality and 61 with moderate quality. Five QC features — Signal to Noise Ratio (SNR), Contrast to Noise Ratio (CNR), Coefficient of Joint Variation (CJV), Foreground-Background energy Ratio (FBER) , and Image Quality Rate (IQR) — were selected as the most informative on image quality by comparison with visual assessment. The multimodal IDPs showed greater impairment in associations with age and dementia biomarkers, demonstrating the potential of the dataset for future clinical analyses.

Keywords: Magnetic resonance imaging; EPAD; Image analysis pipeline; Multi-modal data integration; Quality control

List of abbreviations

ASL	Arterial spin labeling	LST	Lesion Segmentation Toolbox
BIDS	Brain Imaging Data Structure	MRI	Magnetic Resonance Imaging
CSF	Cerebrospinal Fluid	NIFTI	Neuroimaging Informatics Technology Initiative
DICOM	Digital Imaging Communication in Medicine	QC	Quality Control
dMRI	Diffusion MRI	RDM	Research Data Management
EPAD	European Prevention of Alzheimer's Dementia	rs-fMRI	Resting-state functional MRI
EPI	Echo-planar imaging	SPM	Statistical Parametric Mapping
FLAIR	Fluid-attenuated inversion recovery	SWI	Susceptibility weighted imaging
FSL	FMRIB Software Library	TE	Echo Time
GIF	Geodesic Information Flows	TR	Repetition Time
GM	Gray Matter	WMH	White Matter Hyperintensities
IDP	Image-derived phenotype	WM	White Matter
LEAP	Learning Embedding for Atlas Propagation		

1. Introduction

In recent years, data sharing in neuroimaging research communities has become increasingly common, with multiple collaborative efforts for pooling data to form large, diverse samples (Thompson et al. 2014; Jack et al. 2008). Advantages of clinical multicenter imaging studies include obtaining larger samples of subjects from potentially diverse demographic populations, increasing statistical power, generalizability of sophisticated analyses (Friedman et al. 2006), and allowing the development of site- and scanner-independent imaging biomarkers (Pomponio et al., n.d.).

Research Data Management (RDM) is a complex and laborious process in large multisite neuroimaging studies (Nourani, Ayatollahi, and Dodaran 2019), requiring well-defined practice to ensure the accessibility and organization of data, and the provenance of processing steps (Borghi and Van Gulick 2018). The procedure of processing MRI data is highly flexible, and decisions made at early stages can lead to substantial variability in analysis outcomes (Carp 2012). Moreover, there is no standard systematic procedure for MRI quality control (QC) in large multicenter studies. While individual visual inspection is often too laborious, automated procedures can be promising, but highly dependent on study design and the type of sequences acquired (Esteban et al. 2017; Fidel Alfaro-Almagro et al. 2018).

The European Prevention of Alzheimer Dementia (EPAD) study is a prospective, multi-center, European cohort study that aims to characterize the prodromal stages of Alzheimer's Disease (AD) and create a pool of well-characterized individuals for recruitment in potential pharmacological trials (Solomon et al. 2019). Multimodal imaging data are acquired at each center and centrally stored and processed.

This overview documents the methods and implementation details of the MRI data processing, QC procedures, and computation of several imaging-derived phenotypes (IDPs), as developed for the EPAD neuroimaging dataset. We then explored the sanity of IDPs by reporting their relationship with other biomarkers of neurodegeneration. The described pipeline is publicly available online (<https://github.com/ExploreASL/ExploreASL/tree/EPAD>, n.d.).

2. Methods

2.1 The European Prevention of Alzheimer's Dementia Longitudinal Cohort Study (EPAD LCS)

EPAD eligibility criteria were age above 50 years and no history of dementia (clinical dementia rating (CDR) < 1). After providing written informed consent, participants underwent an extensive multimodal test battery including five outcome measurements: cognitive tests, demographics, cerebrospinal fluid (CSF) biomarkers, genetics, and brain MRI. Participants were followed up after 6 months, and after 12, 24, or 36 months, depending on their CDR score at baseline. Details on the EPAD rationale and study protocol are provided elsewhere (Solomon et al. 2019).

Here, we considered the EPAD LCS v1500.0 data release, which consists of the baseline data from the first 1500 participants included in the study.

2.2 The EPAD MRI acquisition protocol

The v1500.0 baseline data were acquired at 21 EPAD sites, including seven different scanner models from Siemens Healthineers, Philips Healthcare, and GE Healthcare. A common scanning protocol was developed during the preparation phase to keep between-site differences as small as possible while accommodating differences in scanner hardware and software limitations.

The EPAD LCS imaging protocol was composed of core and advanced sequences ([Table 1](#)). The core sequences provided structural information and confirmed participants' eligibility status through baseline radiological assessment, and the advanced sequences were designed to investigate brain structure and function in greater detail. Whereas the core sequences were conducted at all sites ($n=21$), the advanced sequences were performed in a subset of EPAD sites ($n=13$) with 3T scanners.

Table 1. Core and Advanced MRI scan protocols

	IDPs or radiological assessment	Siemens (n=13)	Philips (n = 7)	GE Healthcare (n = 1)
Core Sequences				
3D T1w	Regional and global GM volumes, regional GM thickness	1.2 x 1.05 x 1.05, 176 x 256 x 240, Sagittal, TE = 2.95, TR = 2300	1.1 x 1.1 x 1.2, 176 x 256 x 256, Sagittal, TE = 3.11, TR = 1672.6/1526.6	1.2 x 1.2 x 1.05, 196 x 256 x 256, Sagittal, TE = 3.09, TR = 7.65
3D FLAIR	Global WM lesions volume	1.0 x 1.0 x 1.0, 192 x 256 x 256, Sagittal, TE = 393, TR = 5000	1.0 x 1.0 x 1.0, 192 x 256 x 256, Sagittal, TE = 395, TR = 5000	1.0 x 1.0 x 1.0, 188 x 256 x 256, Axial, TE = 155, TR = 5000
2D T2w	Radiological assessment of vascular pathology	0.9 x 0.9 x 3.0, 232 x 256 x 47, Axial, TE = 78, TR = 4510	0.9 x 0.9 x 3.0, 232 x 256 x 47, Axial, TE = 80, TR = 3000	0.9 x 0.9 x 3.0, 256 x 256 x 47, Axial, TE = 82.4, TR = 4000
2D T2*w	Radiological assessment of cerebral microbleeds	0.9 x 0.9 x 3.0, 256 x 256 x 47, Axial, TE = 20, TR = 640	0.9 x 0.9 x 3.0, 256 x 256 x 47, Axial, TE = 20, TR = 640	0.9 x 0.9 x 3.0, 256 x 256 x 47, Axial, TE = 20, TR = 640
Advanced Sequences		n=7	n=6	
3D SWI/SWIp	Radiological assessment of hemorrhage/vascular pathology.	0.5 x 0.5 x 2.0, 384 x 312 x 60, Axial, TE = 23.7, TR = 29.0	0.5 x 0.5 x 3.0, 384 x 384 x 160, Axial, TE = 25, TR = 28	NA
rs-fMRI	Resting State Networks connectivity strength	3.3 x 3.3 x 3.3, 64 x 64 x 38, Axial, TE = 30.0, TR = 2020.0, v = 204	3.3 x 3.3 x 3.3, 64 x 64 x 43, Axial, TE = 30, TR = 1640, v = 202	NA
dMRI	Global and local WM microstructure integrity	2.0 x 2.0 x 2.0, 112 x 112 x 60, Axial, TE = 81.0, TR = 7400, N_B0 = 1, N_B1000 = 54	2.0 x 2.0 x 2.0, 128 x 128 x 56, Axial, TE = 70, TR = 6836, N_B0 = 1, N_B1000 = 48	NA
ASL	Cerebral blood flow (CBF) and Spatial Coefficient of Variation (sCoV)	3D GRASE PASL, 3,75 x 3,75 x 4.5, Axial, TI ₁ = 800, TI ₂ = 2000, v = 20	2D EPI PCASL, 3.4 x 3.4 x 4.5, 64 x 64 x 36, Axial, label duration = 1650, post-labeling delay = 2025	NA

For each sequence, derived data, common uses and average vendor's parameters are given. The shown acquisition parameters are acquisition voxel size (in mm), matrix size, echo time (TE, in ms), repetition time (TR, in ms), number of volumes (v), number of B=0 volumes (only dMRI, N_B0), number of B=1000 volumes (only dMRI, N_B1000), label duration and post-labeling delay (only for Philips ASL, in ms). IDPs = image-derived phenotypes; *N* = number of sites; *GM* = gray Matter; *FLAIR* = fluid attenuated inversion recovery; *SWI* = susceptibility weighted imaging; *SWIp* = SWI-phase; *rs-fMRI* = resting-state functional MRI; *dMRI* = diffusion MRI; *ASL* = arterial spin labeling; *GRASE* = gradient and spin echo; *PASL* = pulsed ASL; *EPI* = echo planar imaging; *PCASL* = pseudo continuous ASL; *w* = weighted; *NA* = not applicable.

2.3 The EPAD Imaging Pipeline

The EPAD image analysis pipeline consisted of four modules ([Figure 1](#)): 1) Curation of raw DICOM files, including harmonization of DICOM structure among sites, initial DICOM quality control (QC), and conversion to NIfTI; 2) Image preprocessing for core and advanced sequences; 3) Semi-automatic QC of processed data through an in-house code-based toolbox; 4) Computation of image-derived phenotypes (IDP) (Gong, Beckmann, and Smith 2020), i.e. extraction of numeric derivatives from images. Data sharing procedures are described in supplementary material section 2 and elsewhere (<http://ep-ad.org/erap/>).

All preprocessing steps were implemented in ExploreASL (Mutsaerts et al. 2020), an SPM-based toolbox designed to harmonize image processing for multi-center structural MRI and arterial spin labeling (ASL) studies. The toolbox was extended to include resting-state functional MRI (rs-fMRI) and diffusion MRI (dMRI) preprocessing routines based on SPM12 r7771 and FSL 6.0.2 (Penny et al. 2011; Jenkinson et al. 2012). Table S1 in supplementary material summarizes the main processing steps and their software implementations in the EPAD pipeline.

Figure 1. Image processing workflow in the EPAD study. *DICOM* = Digital Imaging and Communications in Medicine; *NIfTI* = Neuroimaging Informatics Technology Initiative; *QC* = quality control; *IDP* = Image-derived phenotypes.

2.3.1 DICOM Curation

EPAD data were initially collected in DICOM format in the */sourcedata/* folder and each subject was saved under its EPAD ID. DICOM headers were loaded with ExploreASL's wrapper around DICOM ToolKit (DCMTK v 1.18) and DICOM fields SeriesDescription, or ProtocolName and ImageType were used to recognize the scan type using scanner-specific regular expressions provided in a TSV-file (Figure S1). The assorted zipped files and directories with DICOM files were sorted by scan type. DICOM header information was also used for DICOM QC, e.g., to verify that the StudyID is concordant across scans within an MRI session, to exclude duplicates, and verify completeness of DICOM series and consistency of its parameters.

Following the Brain Imaging Data Structure (BIDS) (Gorgolewski et al. 2016) convention, dcm2niiX r20190902 was used to convert DICOM images to NIfTI format (Li et al. 2016), along with an accompanying JavaScript Object Notation (JSON) sidecar storing relevant metadata. Additional modifications to convert the dcm2niiX output to BIDS then included: sorting and splitting SWI magnitude and phase NIFTIs, splitting the ADC image from the dMRI NIfTI, sorting the phase encoding polarity (PEPolar) scans, obtaining PEPolar parameters and adding them to the JSON sidecars, managing ASL-specific conversion issues (Clement et al. 2019), and manage vendor- and scanner-specific conversion issues.

2.3.2 Image Preprocessing

An overview of the preprocessing steps for the core and advanced sequences is shown in [Figure 2](#) and Table S1.

2.3.2.1 Core Sequences Preprocessing Pipeline

The structural module of ExploreASL v1.0.2, described in (Mutsaerts et al. 2020), was used in combination with the Bayesian Model Selection (BaMoS) for WMH segmentation (Carole H. Sudre et al. 2015) to preprocess 3D T1 and 3D FLAIR images. Other 2D core sequences were only used for radiological assessment of patient eligibility and not preprocessed.

Preprocessing began with the registration of the 3D FLAIR to the 3D T1w (rigid-body) and the 3D T1w to the MNI center of mass (rigid-body). WMH segmentation was computed using BaMoS (Carole H. Sudre et al. 2015), a hierarchical unsupervised model selection framework simultaneously accounting for healthy tissue and unexpected observations. The resulting WMH segmentations were then used with the Lesion Segmentation Toolbox (LST) v2.0.15 (Gaser 2009) to fill these lesion areas on the T1w image, which can be present as hypointensities and affect subsequent segmentation and non-linear registration (Schmidt et al. 2019).

Tissue segmentation was then performed with the Computational Anatomy Toolbox (CAT) 12 (r1363), which estimates and corrects the bias field inhomogeneity in 3D T1w images, and iteratively improves the non-linear registration to MNI standard space and the creation of partial volume maps of gray matter (GM), white matter (WM) and CSF. All the described transformations were combined in a single transformation, to avoid multiple interpolations.

2.3.2.2 Advanced Sequences Preprocessing Pipeline

Rs-fMRI, dMRI, and ASL were handled by different preprocessing submodules and underwent three common steps: geometric distortion correction, motion correction, and registration with the structural reference image T1w. Two additional steps were performed for dMRI (see below). SWI was only used for radiological assessment as no automated processing routines exist for this sequence.

A) *Geometric Distortion Correction*. Echo-planar images (EPI) suffered from geometric distortion due to B_0 field inhomogeneity induced by magnetic susceptibility variability (Holland, Kuperman, and Dale 2010). For this reason, EPI scans were accompanied by short acquisitions with reversed phase-encoding gradient polarity, which have an opposite distortion pattern. From this pair of images, the geometric distortion was estimated using a previously described method (Andersson and Sotiropoulos 2016) as implemented in FSL topup (Stephen M. Smith et al. 2004), and used to correct the geometric distortion of the fMRI, dMRI, and ASL with a 2D EPI readout.

B) *Motion Correction*. Head motion within fMRI and ASL was estimated with rigid-body transformations using the SPM12 realign function (Friston et al. 1995), where the ASL images were

combined with the threshold-free outlier exclusion method ENhancement of Automated BLood flow Estimates (ENABLE) (Mutsaerts et al. 2020; Shirzadi et al. 2018). In dMRI images, an additional off-resonance source is caused by the rapidly changing magnetic field inducing eddy currents (EC) within conductors (Zhuang et al. 2006). Head motion and eddy current-induced geometrical EPI distortions were estimated and corrected by the FSL Eddy tool (Andersson and Sotiropoulos 2016).

C) Structural registration. Advanced sequences were registered to the 3D T1w images using rigid-body transformations. Similar to the core preprocessing, all transformation fields are combined and applied simultaneously, to avoid multiple cumulative interpolations.

D) dMRI Tensor Fitting. In the dMRI preprocessing submodule the registration output was fed into the FSL Brain Extraction Toolbox (BET) (Stephen M. Smith 2002) and then into FSL DTIFIT, to fit the diffusion tensor model to the data and produce diffusion tensor imaging (DTI) scalars maps (fractional anisotropy (FA), and mean (MD), axial (AD) and radial (RD) diffusivity).

Figure 2. Schematic diagram of preprocessing steps. *Left:* the core sequences preprocessing pipeline performed on 3D T1w and 3D FLAIR scans; *Right:* the three common steps of the advanced sequences preprocessing pipeline. *Abbreviations:* EPI = Echo-Planar Imaging.

2.3.3 Semi-automatic QC

To control the quality of the EPAD imaging cohort, we created an in-house workflow to perform semi-automatic QC of MRI data. This set of QC functionalities was written as an extension to ExploreASL called ExploreQC. The semi-automated QC procedure was based on two steps: feature estimation and visualization ([Figure 3](#)). ExploreQC code availability and software specifications are listed in section 4 of the supplementary material.

2.3.3.1 QC - Feature Estimation

Image quality features were computed from five image feature domains: motion, noise, inhomogeneity, asymmetry, and descriptives, in line with recent MRI QC studies (Esteban et al. 2017; Fidel Alfaro-Almagro et al. 2018; Zarrar et al. 2015; Bastiani et al. 2019)). Definitions of domains and individual features are provided in table S2. All QC features were included in the MRI data release and can be referred to for study-specific inclusion/exclusion criteria.

2.3.3.2 QC - Visualization

The visualization module consists of an interactive dashboard with violin and scatter plots for observing variation between and within sites, respectively ([Figure 3](#)). Individual scans can be visually inspected by selecting their data points on the scatter plots, allowing to visualize the scans themselves together with the QC features.

As a proof-of-concept for our semi-automated QC strategy, 3D T1w images were passed on to the visualization module. For all QC features, the difference with the site-specific mean was calculated as Z-score. Each scan was then sorted site-wise based on the sum of their absolute within-site Z-score values of all QC features. Scans with the 15% highest summed deviations were then automatically flagged for “needing visual QC”. To control for possible false-negative cases, an equivalent subset of non-flagged images was randomly selected and visually checked.

Visual QC was performed within the same pane by a single rater (LL), blinded to whether the image was an outlier or flagged as a random inlier. The image quality was categorized as “good” — desired MRI contrast visible and no artifacts or quality degradation detected, “moderate” — desired MRI

contrast visible but some quality degradation, or “poor” — no desired MRI contrast visible and/or clear artifacts are present. These flags were added to the EPAD data release as a QC category advice for external researchers, together with the estimated features.

2.3.3.3 QC Statistical Analysis

After visual inspection, we explored the association of the QC features with participants’ characteristics and with visual QC judgments. First, we used linear models to assess whether each QC feature distribution was related to the scanning site, age, sex, MMSE, amyloid, and APOE status of the participant. P-values were Bonferroni corrected.

To further investigate the features’ informative character in relation to the visual inspection, we built a logistic regression model with QC features as predictors of images showing quality issues among visual inspection. A stepwise backward parameter selection based on the Akaike Information Criteria was then performed to remove non-informative QC features from the final model.

Explore QC

1. Feature Estimation

Motion

Noise

Inhomogeneity

Asymmetry

Descriptives

2. Visualization

Figure 3. Overview of the quality control workflow. QC features are computed in the feature estimation module and cover 5 image features domains. Feature distributions can then be interactively inspected between-sites (2A) and within-sites (2B). Single-subject scans can be opened by clicking on the scatterplots (2C).

2.3.4 Image-derived phenotypes (IDPs)

IDPs are image-specific summary statistics that provide a quantitative way to investigate structural and functional brain characteristics (Gong, Beckmann, and Smith 2020).

2.3.4.1 Core IDPs

Regional GM volume and cortical thickness are established phenotypes in neurodegenerative diseases (Scheltens et al. 2016). For 3D T1w sequences, we computed the volumetrics of several pipelines with different segmentation strategies. Template-based tissue volumetrics were computed from the CAT12-SPM tissue segmentation pipeline (Gaser 2009) described above. FreeSurfer v6.0.0 (Fischl 2012), one of the most widely used packages for measuring GM volumes and cortical thickness, was run on 3D T1w scans and included in the core IDPs release. A third method, the Learning Embeddings for Atlas Propagation (LEAP), was used to compute atlas-propagation-based regional volumes (Wolz et al. 2010). However, LEAP was run as a part of a separate and parallel pipeline and its results are not discussed here. WMH regional volumes were calculated from BaMoS segmentations and constitute 3D FLAIR derived data (C. H. Sudre et al. 2018).

2.3.4.2 Advanced IDPs

rs-fMRI

Temporally correlated low-frequency ($<0.1\text{Hz}$) fluctuations in the rsfMRI signal are defined as functional resting-state networks (RSNs) (S. M. Smith et al. 2009). To identify RSNs in fMRI time series, a group-level independent component analysis (ICA) was performed by FSL Melodic (Beckmann and Smith 2004) on the preprocessed fMRI datasets. Two atlases of RSNs with a predefined number of independent components were generated: a low dimensional atlas with 20 and a higher dimensional atlas with 50 independent components. A dual regression approach was then used to obtain subject-specific RSNs (Nickerson et al. 2017). First, each RSN's summary time course was estimated at the participant level by spatial regression of the full set of independent components from the high and low dimensional Melodic analysis against each participant's fMRI data. Second, the

resulting time courses were regressed into the same participants' fMRI data to obtain subject-specific RSN maps. fMRI IDPs were computed as the mean within-network connectivity strength per subject.

dMRI

Tract-based spatial statistics (TBSS) is an automated, observer-independent approach for assessing voxel-wise fractional anisotropy in white matter tracts across groups of dMRI scans (Stephen M. Smith et al. 2006). The brain-extracted fractional anisotropy (FA) images, obtained after tensor fitting, were aligned into a common space using nonlinear registration. Next, the mean FA image was thinned to create a mean FA skeleton representing the center of all tracts common to the group and use it as a mask to compute individual FA values. Diffusion MRI IDPs were computed as global and regional FA features from the JHU ICBM-DTI-81 atlas (Wakana et al. 2004).

ASL

Arterial spin labeling perfusion MRI acquires cerebral perfusion *in vivo* in a non-invasive manner. Recent findings have shown that the 2D EPI readout on previous Philips software releases exhibits fat-saturation-related artifacts that considerably alter the quality of these scans, even to the point of being unusable (Mutsaerts et al. 2020). Therefore, we derived data only on a subset of ASL images that did not suffer from this artifact. Mean cerebral blood flow (CBF) and spatial coefficient-of-variation were computed as described in (Mutsaerts et al. 2020).

2.3.4.3 IDPs relationship with AD markers

We finally explored the sanity and relevance of IDPs by assessing their relationship with age and other non-imaging data whose association with brain phenotypes has been established in the Alzheimer literature. Gray matter volume association with age was assessed through the Pearson correlation coefficient of both global and regional values. Increase in volume of WMH per additional year of age was evaluated as previously described (C. H. Sudre et al. 2018) in males and females separately, to account for sex differences in cerebrovascular disease (Haberman, Capildeo, and Rose 1981). As nonlinear trajectories of functional connectivity have been observed with age (Persson et al. 2014), mean within RSN connectivity quadratic relationship with age and differences between CDR = 0 and CDR = 0.5 participants were evaluated. dMRI and ASL IDPs correlation with age was computed,

interacting age with amyloid status as defined in (Ingala et al. 2021) and Apolipoprotein E (APOE) e4 allele carrier status, respectively. As an initial exploratory data analysis, we here only report several preliminary associations. We did not explore the statistical significance level or effect of covariate correction on these correlations.

3. Results

3.1 The EPAD LCS Baseline Imaging Dataset

Of the 1500 screened participants, 144 did not fulfill the EPAD LCS eligibility criteria and were excluded from the analysis (Solomon et al. 2019). The resulting dataset has a complete set of core sequences for 1356 participants. Advanced sequences were performed at thirteen sites, with 756 SWI, 842 fMRI, 831 dMRI, and 858 ASL scans acquired. The mean age was 65.46 ± 7.14 and 775 (55.71 %) were female. All participants were without dementia at inclusion with an average MMSE of 28.58 ± 1.66 , and 259 (19.11%) had a CDR of 0.5. A total of 1246 participants had CSF measurements, of which 845 (67.8 %) individuals were CSF amyloid negative and 401 (32.2 %) CSF amyloid positive, following previously defined study-specific cutoffs (Ingala et al. 2021). A detailed description of the baseline clinical and demographic characteristics of the EPAD cohort can be found in (Ingala et al. 2021).

Pre-processing steps were successful for all the T1w images, while failed for 8.6 % of fMRI and 8.6% of dMRI data, mostly due to high head motion or lack of necessary files (e.g. reverse-phase). A high percentage (62.3%) of the ASL sequences showed fat-saturation artifacts and are currently excluded from the dataset; preprocessing failure rate for the remaining ASL sequences was 0%. The final preprocessed EPAD dataset resulted in 1356 T1w and FLAIR, 770 fMRI, 759 dMRI, and 237 ASL scans. Details on the number of scanned and available processed data are provided in [Figure 4](#).

3.2 Quality Control

49 QC features from 3 different image modalities (3D T1w = 12, fMRI = 17, dMRI = 20) were computed and are included in the current release.

Linear models showed that the QC feature distribution of T1w, fMRI, and DTI significantly differed between scanning sites. Participants' demographic characteristics, such as age and sex, were related to the motion- and noise-related QC features for the 3D T1w and fMRI images ($P < 0.05$), but not for the dMRI images ($P > 0.05$). Amyloid status and MMSE were associated with 3D T1w motion- and noise-related QC features. APOE $\epsilon 4$ status was not related to any QC feature (Table S3).

Based on within-site distributions of the 12 QC features extracted from 3D T1w images, 197 scans (15% of the whole sample) were flagged as “needing-QC” for the visualization module. Of those, 16 were categorized as “poor quality” on visual inspection, while 51 were labeled as “moderate quality”. In the same number of visually inspected non-flagged scans, only one was labeled “poor” while 10 were judged to have “moderate” quality.

In the stepwise logistic regression analysis, backward parameter elimination removed four out of the 12 structural QC features in the final reduced model (Table 3). Five of those — Signal to Noise Ratio (SNR), Contrast to Noise Ratio (CNR), Coefficient of Joint Variation (CJV), Foreground-Background energy Ratio (FBER) , and Image Quality Rate (IQR) — were significantly associated with the QC visual judgment and thus considered the most informative. When comparing the reduced model (i.e., after backward elimination) with the initial full model (i.e. before backward elimination), there was no significant loss of fitting. Adding site, sex, and age as covariates in the model did not significantly affect our findings (data not shown). A detailed description of the presented QC parameters can be found in section 5 of the supplementary material.

Figure 4. Consort diagram representing number of scanned and successfully processed sequences. Abbreviations: EPAD = European Prevention of Alzheimer’s Dementia; T1w = T1 weighted; FLAIR = Fluid attenuated inversion recovery; fMRI = functional magnetic resonance imaging; dMRI = diffusion magnetic resonance imaging; ASL = Arterial spin labeling.

Table 3. Results of stepwise backward parameter elimination in QC logistic regression.

<i>Parameter</i>	<i>Domain</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)		1.47	0.6014 – 2.3455	0.001
Signal to Noise Ratio (SNR)	Noise	-0.015	-0.0210 – -0.0091	<0.001
Contrast to Noise Ratio (CNR)	Noise	-0.089	-0.1246 – -0.0544	<0.001
Coefficient of Joint Variation (CJV)	Inhomogeneity	-0.062	-0.1160 – -0.0073	0.026
Foreground-Background energy Ratio (FBER)	Inhomogeneity/Motion	0.003	0.0013 – 0.0045	<0.001
Asymmetry Index percentage (AI_perc)	Asymmetry	-0.001	-0.0013 – 0.0001	0.107

Image Quality Rate (IQR)	Inhomogeneity/Noise	0.317	0.2660 – 0.3675	<0.001
Kurtosis in the CSF (CSF_k)	Descriptives	-0.014	-0.0295 – 0.0022	0.092
White Matter to Maximum Intensity ratio (WM2MAX)	Inhomogeneity/Descriptives	-0.761	-1.7109 – 0.1890	0.116

The reduced model included 8 parameters, 5 of which showed a p-value of less than 0.05 in the association with the visual QC judgment. Estimates represent the change in the log odds of the outcome for a one-unit increase in the predictor variable. P-values < 0.05 are shown in bold

3.3 Image-derived phenotypes

358 IDPs per subject were computed from core sequences and provided information about total and regional GM volumes, cortical thickness, and white matter lesions. Example 3D T1w-IDPs are shown to correlate with age in Figure 5 and Figure S2. The FreeSurfer results showed negative associations of regional GM volumes with age, most markedly in the hippocampus ($r=-0.34$). A similar association is shown for the global GM volume computed with the CAT12 segmentation. From the set of 36 regional WMH volumes quantified on the 3D FLAIR scans, we found a frontal and parietal prevalence of WMH increase with aging, with no clear differences between sexes (Figure 6).

The mean network connectivity was computed for 70 ICA components for the low and high dimensional ICA resulting in 70 fMRI IDPs. Figure 7A shows six RSNs commonly found in the literature (S. M. Smith et al. 2009), identified in the low-dimensional ICA. Complete sets of extracted components from both high and low dimensional ICA are shown in supplementary material Figure S3 and S4. RSN's non-linear relationship with age and cognitive status was assessed and is shown in Figure 7B. Non-linear relationship of mean within-network connectivity with age differed by CDR groups, more strongly in the default mode and visual networks.

Global and local skeletonized FA values for the 48 regions from the JHU ICBM-DTI-81 atlas (Wakana et al. 2004) were calculated as dMRI IDP. Negative correlations of age with both global and regional

WM integrity were found, with more negative coefficients in the amyloid positive participants (Figure 8).

Usable ASL data (without fat-saturation artifact) were derived from 237 participants. Mean CBF maps across participants showed considerable regional variability across the brain with high perfusion in the cingulate and precuneus GM and lower perfusion in basal ganglia (Figure 9A). An atypical relationship between CBF and age was observed in APOE-e4 allele carriers, showing a lack of CBF decline with age. In contrast, non-carrier participants showed a reduction of CBF with advancing age (Figure 9B).

Figure 5. 3D T1w derived phenotypes. Example association between core sequences derived data and age. A) FreeSurfer surface reconstruction of one 3D T1w image; B) Association of eight cortical regional volumes with age; C) CAT12 tissue segmentation of one 3D T1 scan output: green = cerebrospinal fluid, blue = white matter, red = gray matter; D) Association of total gray matter volume, as computed with CAT12 segmentation, with age.

Figure 6. FLAIR derived phenotypes. A) Example FLAIR scan from one EPAD participant with relatively high lesion volume; B) Result of the white matter hyperintensities (WMH) segmentation using BaMoS; C, D) Lobes and layer atlases, respectively, used for regional WMH volume computation, methodological details are given in (C. H. Sudre et al. 2018); E, F) Effect of age on WMH frequency (expressed in percentage of increase in frequency, i.e. the proportion of lesion in a given region, per additional year of age) for male and female respectively.

Figure 7. Resting-state fMRI derived phenotypes. A) Six group resting-state networks spatial maps from a low dimensional (20 independent components) melodic ICA; B) Scatter plots showing the non-linear relationship of mean within-network functional connectivity with age, grouped by clinical dementia rating (CDR) score. R values are computed as the Pearson correlation coefficients between the quadratic age term (age^2) and mean within network connectivity values.

Figure 8. Diffusion MRI-derived phenotypes. A) The FA skeleton as computed in the TBSS pipeline (upper row) and the skeletonized white matter atlas used to extract local FA values (bottom row); B) Association of mean global FA values with age and amyloid status (as defined in (Ingala et al. 2021)). C) Association of 4 regional FA values with age and amyloid status. Abbreviations: FA = Fractional anisotropy; TBSS = Tract based spatial statistics; WM = White Matter; Sp. = Superior; CC = Corpus Callosum; r = Pearson correlation coefficient.

Figure 9. Arterial spin labeling IDPs. A) Mean CBF in the gray matter across 237 participants. B) CBF in the GM relationship with age and APOE ε4 carriership. Abbreviations: CBF = Cerebral blood flow; GM = Gray matter; APOE = Apolipoprotein E.

4. Discussion and Future Directions

Here, we provided a detailed description of the EPAD MRI dataset and the semi-automatic pipeline developed to process raw multimodal multicenter imaging data from the v1500 baseline data release, and illustrate its IDP feature extraction. We proposed to use a combination of distribution-based statistics and subpopulation visual assessments to help identify low-quality images in EPAD. Finally, we described the computation of MRI summary measures (IDPs) from the core and advanced sequences and provided evidence of meaningful associations between brain phenotypes and other non-imaging metadata. We anticipate that this work may benefit both EPAD investigators by providing a structured, organized, and well-characterized neuroimaging dataset.

4.1 MRI Cohorts Research Data Management

The proposed pipeline for the EPAD study represents an attempt to deliver standardized data in a multicenter MRI study, in a similar way to what has been previously done in other cohorts. Among publicly available neuroimaging cohorts, extensive work on the documentation of MRI data management strategies has been recently published by two large MRI cohort studies (Sudlow et al.

2015; Bookheimer et al. 2019). These efforts entail fully automated pipelines for MRI processing, QC, and computation of IDPs from multimodal single-scanner imaging datasets (F. Alfaro-Almagro and Jenkinson 2016; Glasser et al. 2013). Compared to these studies, a unique feature of our pipeline is that it caters multiple scanners from multiple vendors and harmonizes differences between different scanner outputs to allow for automated processing and semi-automated QC. Multisite neuroimaging RDM represents a more complex procedure due to the variety of DICOM output structures from different manufacturers and scanners. The multisite Alzheimer Disease Neuroimaging Initiative (ADNI, (Mueller et al. 2005)) has focused on MRI standardization across sites, mainly during the preparation phase (Jack et al. 2008) where QC and preprocessing were done centrally and also provided extracted MRI metrics to investigators. The Preprocessed Connectome Project (Cameron et al. 2013) is a community effort to systematically process data from several available multicenter datasets including structural and functional MRI workflows and comparing outputs of different processing choices and toolboxes. As an example, the Autism Brain Imaging Data Exchange (ABIDE) dataset (Di Martino et al. 2014), which includes 1114 participants with structural (3D T1w) and rs-fMRI data from 16 sites, was processed using three different structural and four functional pipelines, and corresponding derivatives were made available. While both these studies and our EPAD pipeline managed to deal with such diverse datasets, we specifically aimed at creating a unified workflow by delivering a uniform and QCed shared outcome, promoting reproducibility, and avoiding redundancy and variability of results.

4.2 Quality Control

Our data-informed QC procedure was focused on selecting a handful of most informative image features in a semi-automatic fashion. On the other hand, most recent MRI studies focused on predicting the scan quality automatically using a fairly large number of potentially helpful image features through unsupervised or semi-supervised methods (F. Alfaro-Almagro and Jenkinson 2016; Pizarro et al. 2016). The application of machine learning classifiers has proved its efficiency in recent QC efforts on classifying 3D T1w image-quality from QC features distribution, both the UKBiobank (F. Alfaro-Almagro and Jenkinson 2016) and ABIDE (Esteban et al. 2017; Di Martino et al. 2014)

datasets. Similarly, random forest classifiers trained on FreeSurfer QC output showed good accuracy in scanning site identification, supporting the use of multivariate approaches for QC metrics' importance evaluation (Raamana et al. 2021). However, while these works aimed at the fully automatic prediction of image quality from unseen scans/sites, we focused on identifying a set of informative QC features as a pre-selection and guide for visual inspection. Similar semi-automated procedures have also been proposed in the literature. Previously ((Bastiani et al. 2019)), dMRI QC features were used to create individual and group reports and, similar to our work, interactively inspect automatically flagged problematic scans. MRIQC (Esteban, Gorgolewski, and Poldrack 2017) provides interactive individual reports created for straightforward low-quality image visualization. Likewise, well-established processing pipelines for different MRI modalities (Esteban et al. 2019; Mutsaerts et al. 2020) produce single-subject visual QC reports for the quality assurance of specific processing steps.

We showed a valuable pragmatic and intuitive solution for identifying problematic acquisitions and reducing the number of scans for visual inspection in large multicenter cohorts. Moreover, considering the recent efforts to relate quality metrics to the output of expert visual rating (Esteban et al., n.d.), another advantage of our QC procedure was the possibility of using statistical regression models to evaluate the informative character of single QC features and their agreement with a human visual inspection, as well as their relationship with scanning site and other participants' demographic data.

In line with previous work (Esteban et al. 2017), all QC features were strongly dependent on scanners and their sites, advocating for within-site normalization of quality metrics in multicenter MRI cohorts, as similar features' values might be related to different quality of images from different scanners. Moreover, we also showed that demographic and clinical participants' characteristics are related to image quality metrics. Former studies had already shown that age and sex can relate to impact the quality of structural derived measures of brain atrophy (Gilmore, Buser, and Hanson 2021), and fMRI derived functional connectivity (Hodgson et al. 2017). Together with our results, these findings suggest that scan quality might confound effects attributed to clinical variables and, consequently, that fully automated QC procedures might be more prone to exclude scans from selected groups of participants (e.g. older participants or clinical groups).

4.3 Image-derived Phenotypes

In agreement with current literature (Damoiseaux 2017), we showed an overall effect of age in modulating brain IDPs across different MRI modalities in the EPAD-LCS data set. However, different from more clinically oriented papers, we here only reported basic association without investigating statistical significance or the effect of covariates correction on the studied associations. Similar to our findings using FreeSurfer and CAT12 segmentations, studies on structural brain converge on a gradual loss of brain volume with advancing age, specifically in the hippocampus (Salat et al. 2004). Moreover, a similar frontal and parietal prevalence of vascular disease with older age, as demonstrated with the 3D FLAIR IDPs, has been shown in previous works (Chabriat and Jouvent 2020). Similar to our findings, a nonlinear age-related decline in resting-state fMRI networks has previously been demonstrated. In (Persson et al. 2014), while participants below 66 years showed an increase in DMN connectivity over time, while participants older than 74 years showed a decline. Additionally, in (Brier et al. 2012) early cognitive changes modulated default mode and frontoparietal network connectivity, possibly altering normal aging trajectories. As observed with the dMRI IDPs, alterations of white matter integrity are a typical sign of aging brains (Barrick et al. 2010). Previous publications agree with our result of reduced TBSS values in amyloid-positive individuals (Collij et al. 2020). Finally, A decline of CBF in the GM has been widely reported in recent studies of elderly individuals (Juttukonda et al. 2021). Previous studies on CBF with APOE genotype have shown higher brain perfusion related to worse cognitive impairment in older adults carrying the APOE e4 allele (Zlata et al. 2016), and higher regional perfusion in e4 carriers in the left cingulate and lateral frontal and parietal regions (McKiernan et al. 2020). Moreover, similar group regional variability, showing high perfusion in the cingulate, precuneus, and frontal cortices and low perfusion in basal ganglia, has been previously reported (Pfefferbaum et al. 2010).

4.4 Limitations

Although we acknowledge the heterogeneity of existing MRI processing methodologies and implementations, we focused here on a purely descriptive overview of the procedures implemented for the EPAD study. As we followed generally accepted and standard pipelines, a potential limitation of this study could be the lack of more novel and AI-based processing routines. For example, performing denoising of signal drift correction is becoming a preferable procedure for dMRI preprocessing (Vos et al. 2017), even if few implementations exist. Another weakness of the proposed pipeline is the lack of longitudinal routines. Challenges of longitudinal MRI processing and QC entail the necessity of additionally taking into account within-subject variability, which could be added to this pipeline in future extensions (Mills and Tamnes 2014). Moreover, one main limitation of the validation approach used for the QC workflow is the partial circularity of constructing and testing the visual QC assessment based on the estimated QC features. Nonetheless, while our first aim was to automatically flag poor quality images for visual inspection, we then elaborated on the informative characters of QC features, studying their association with visual judgment. Furthermore, in contrast to more systematic procedures, in which visual inspection is performed on the whole sample (Waber et al. 2007), we only focused on automatically flagged scans and on a subset of non-flagged. The present work is a pragmatic approach of combining image feature-based QC with visual QC of a limited number of scans, which may be more feasible for large imaging cohorts. Reflective of EPAD's multi-center and multi-vendor design, ExploreQC was tailored to this dataset and requires further validation and testing before it can be generalized to other studies. Eventually, the lack of visual QC standards for advanced MRI sequences motivates the development of more standardized approaches in future works. The IDPs found in the EPAD v1500.0 data sample cover a broad range of structural and functional brain phenotypes. However, several possibilities for summary brain measures exist (Gong, Beckmann, and Smith 2020). Future efforts should focus on widening the present number of IDPs to entail a new range of brain phenotypes, including measures of longitudinal changes where applicable.

Conclusion

We provide a detailed description of the baseline EPAD LCS MRI dataset including the processing and QC procedures and details of the computation of derived data from core and advanced MRI sequences yielding biologically plausible IDPs. The introduced procedures and results may serve as a reference point for future developments and promote replicability and stability of results on the EPAD cohort. We made the pipeline available for external investigators aiming at the comparability of outcomes between different cohorts. We anticipate that this work will help both imaging and non-imaging researchers working on EPAD for an easier understanding and use of the shared data.

Acknowledgments

This work is part of the EPAD LCS (European Prevention of Alzheimer's Dementia Longitudinal Cohort Study). The authors would like to express their gratitude to the EPAD-LCS participants, without whom this research would have not been possible.

Funding

EPAD is supported by the EU/EFPIA Innovative Medicines Initiative Joint Undertaking EPAD grant agreement 115736.

FB is supported by the NIHR UCLH Biomedical Research Centre

NCF is supported by the NIHR UCLH Biomedical Research Centre and the UK Dementia Research Institute at UCL

DLT is supported by the UCLH NIHR Biomedical Research Centre, the Wellcome Trust (Centre award 539208), and Alzheimer's Research UK (ARUK-NAS2016B-2)

DMC is supported by the UK Dementia Research Institute which receives its funding from DRI Ltd, funded by the UK Medical Research Council, Alzheimer's Society, and Alzheimer's Research UK, as

well as Alzheimer's Research UK (ARUK-PG2017-1946) and the UCL/UCLH NIHR Biomedical Research Centre.

GBF is funded by: A.P.R.A. - Association Suisse pour la Recherche sur la Maladie d'Alzheimer, Genève; Fondation Segré, Genève; Ivan Pictet, Genève; Fondazione Agusta, Lugano; Fondation Chmielewski, Genève; Swiss National Science Foundation; and VELUX Foundation.

HM is supported by the Dutch Heart Foundation (2020T049), by the Eurostars-2 joint program with co-funding from the European Union Horizon 2020 research and innovation program, provided by the Netherlands Enterprise Agency (RvO), and by the EU Joint Program for Neurodegenerative Disease Research, provided by the Netherlands Organisation for Health Research and Development and Alzheimer Nederland.

JMW is supported by the UK Dementia Research Institute Ltd (funded by the UK MRC, ARUK, and Alzheimer Society) and the Fondation Leducq (16 CVD 05).

JDG holds the “Ramón y Cajal” fellowship RYC-2013-13054.

Declaration of competing interest

None.

REFERENCES

- Alfaro-Almagro, Fidel, Mark Jenkinson, Neal K. Bangerter, Jesper L. R. Andersson, Ludovica Griffanti, Gwenaëlle Douaud, Stamatis N. Sotiropoulos, et al. 2018. "Image Processing and Quality Control for the First 10,000 Brain Imaging Datasets from UK Biobank." *NeuroImage* 166 (February): 400–424.
- Alfaro-Almagro, F., and M. Jenkinson. 2016. "UK Biobank Brain Imaging: Automated Processing Pipeline and Quality Control for 100,000 Subjects." *For Human Brain* https://www.researchgate.net/profile/Fidel_Alfaro-Almagro/publication/311572037_UK_Biobank_Brain_Imaging_Automated_Processing_Pipeline_and_Quality_Control_for_100000_subjects/links/584dcf7a08aeb9892526466f/UK-Biobank-Brain-Imaging-Automated-Processing-Pipeline-and-Quality-Control-for-100-000-subjects.pdf.
- Andersson, Jesper L. R., and Stamatis N. Sotiropoulos. 2016. "An Integrated Approach to Correction for off-Resonance Effects and Subject Movement in Diffusion MR Imaging." *NeuroImage* 125 (January): 1063–78.
- Barrick, Thomas R., Rebecca A. Charlton, Chris A. Clark, and Hugh S. Markus. 2010. "White Matter Structural Decline in Normal Ageing: A Prospective Longitudinal Study Using Tract-Based Spatial Statistics." *NeuroImage* 51 (2): 565–77.
- Bastiani, Matteo, Michiel Cottaar, Sean P. Fitzgibbon, Sana Suri, Fidel Alfaro-Almagro, Stamatis N. Sotiropoulos, Saad Jbabdi, and Jesper L. R. Andersson. 2019. "Automated Quality Control for within and between Studies Diffusion MRI Data Using a Non-Parametric Framework for Movement and Distortion Correction." *NeuroImage* 184 (January): 801–12.
- Beckmann, Christian F., and Stephen M. Smith. 2004. "Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging." *IEEE Transactions on Medical Imaging* 23 (2): 137–52.
- Bookheimer, Susan Y., David H. Salat, Melissa Terpstra, Beau M. Ances, Deanna M. Barch, Randy L. Buckner, Gregory C. Burgess, et al. 2019. "The Lifespan Human Connectome Project in Aging: An Overview." *NeuroImage*. <https://doi.org/10.1016/j.neuroimage.2018.10.009>.
- Borghi, John A., and Ana E. Van Gulick. 2018. "Data Management and Sharing in Neuroimaging: Practices and Perceptions of MRI Researchers." *PloS One* 13 (7): e0200562.
- Brier, Mathew R., Jewell B. Thomas, Abraham Z. Snyder, Tammie L. Benzinger, Dongyang Zhang, Marcus E. Raichle, David M. Holtzman, John C. Morris, and Beau M. Ances. 2012. "Loss of Intranetwork and Internetwork Resting State Functional Connections with Alzheimer's Disease Progression." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 32 (26): 8890–99.
- Cameron, Craddock, Benhajali Yassine, Chu Carlton, Chouinard Francois, Evans Alan, Jakab András, Khundrakpam Budhachandra, et al. 2013. "The Neuro Bureau Preprocessing Initiative: Open Sharing of Preprocessed Neuroimaging Data and Derivatives." *Frontiers in Neuroinformatics*. <https://doi.org/10.3389/conf.fninf.2013.09.00041>.
- Carp, Joshua. 2012. "On the Plurality of (methodological) Worlds: Estimating the Analytic Flexibility of fMRI Experiments." *Frontiers in Neuroscience* 6 (October): 149.
- Chabriat, H., and E. Jouvent. 2020. "Imaging of the Aging Brain and Development of MRI Signal Abnormalities." *Revue Neurologique* 176 (9): 661–69.
- Clement, Patricia, Marco Castellaro, Thomas Okell, D. Thomas, C. Gorgolewski, S. Appelhoff, Jan Petr, M. Chappell, and H-J Mutsaerts. 2019. "ASL-BIDS, the Brain Imaging Data Structure Extension for Arterial Spin Labeling." *Magnetic Resonance Materials in Physics, Biology and Medicine*. <https://www.hzdr.de/publications/Publ-29979>.
- Collij, L. E., H. Top, S. Ingala, K. Stickney, J. Tomassen, E. Konijnenberg, M. Ten Kate, et al. 2020. "White Matter Integrity Disruption in Early Amyloid Accumulators." *Age* 70: 7–14.
- Damoiseaux, Jessica S. 2017. "Effects of Aging on Functional and Structural Brain Connectivity." *NeuroImage* 160 (October): 32–40.
- Di Martino, A., C-G Yan, Q. Li, E. Denio, F. X. Castellanos, K. Alaerts, J. S. Anderson, et al. 2014. "The Autism Brain Imaging Data Exchange: Towards a Large-Scale Evaluation of the Intrinsic Brain Architecture in Autism." *Molecular Psychiatry* 19 (6): 659–67.
- Esteban, Oscar, Daniel Birman, Marie Schaer, Oluwasanmi O. Koyejo, Russell A. Poldrack, and Krzysztof J. Gorgolewski. 2017. "MRIQC: Advancing the Automatic Prediction of Image Quality in MRI from Unseen Sites." *PloS One* 12 (9): e0184661.
- Esteban, Oscar, Ross W. Blair, Dylan M. Nielson, Jan C. Varada, Sean Marrett, Adam G. Thomas, Russell A. Poldrack, and Krzysztof J. Gorgolewski. n.d. "Crowdsourced MRI Quality Metrics and Expert Quality Annotations for Training of Humans and Machines." <https://doi.org/10.1101/420984>.
- Esteban, Oscar, Krzysztof Gorgolewski, and Russell Poldrack. 2017. "MRIQC: Automatic Prediction of Quality and Visual Reporting of MRI Scans." *F1000Research* 6.
- Esteban, Oscar, Christopher J. Markiewicz, Ross W. Blair, Craig A. Moodie, A. Ilkay Isik, Asier Erramuzpe, James D. Kent, et al. 2019. "fMRIPrep: A Robust Preprocessing Pipeline for Functional MRI." *Nature Methods* 16 (1): 111–16.
- Fischl, Bruce. 2012. "FreeSurfer." *NeuroImage* 62 (2): 774–81.
- Friedman, Lee, Gary H. Glover, Diana Krenz, Vince Magnotta, and FIRST BIRN. 2006. "Reducing Inter-Scanner Variability of Activation in a Multicenter fMRI Study: Role of Smoothness Equalization." *NeuroImage* 32 (4): 1656–68.
- Friston, Karl J., J. Ashburner, C. D. Frith, J-B Poline, J. D. Heather, and R. S. J. Frackowiak. 1995. "Spatial Registration and Normalization of Images." *Human Brain Mapping* 3 (3): 165–89.
- Gaser, C. 2009. "Partial Volume Segmentation with Adaptive Maximum A Posteriori (MAP) Approach." *NeuroImage Supplement* 1 (47): S121.
- Gilmore, Alysha D., Nicholas J. Buser, and Jamie L. Hanson. 2021. "Variations in Structural MRI Quality Significantly Impact Commonly Used Measures of Brain Anatomy." *Brain Informatics* 8 (1): 7.
- Glasser, Matthew F., Stamatis N. Sotiropoulos, J. Anthony Wilson, Timothy S. Coalson, Bruce Fischl, Jesper L. Andersson, Junqian Xu, et al. 2013. "The Minimal Preprocessing Pipelines for the Human Connectome Project." *NeuroImage* 80 (October): 105–24.
- Gong, W., C. F. Beckmann, and S. M. Smith. 2020. "Phenotype Discovery from Population Brain Imaging." *bioRxiv*. <https://www.biorxiv.org/content/10.1101/2020.03.05.973172v1.abstract>.
- Gorgolewski, Krzysztof J., Tibor Auer, Vince D. Calhoun, R. Cameron Craddock, Samir Das, Eugene P. Duff, Guillaume Flandin, et al. 2016. "The Brain Imaging Data Structure, a Format for Organizing and Describing Outputs of Neuroimaging Experiments." *Scientific Data*. <https://doi.org/10.1038/sdata.2016.44>.
- Haberman, S., R. Capildeo, and F. C. Rose. 1981. "Sex Differences in the Incidence of Cerebrovascular Disease." *Journal of Epidemiology and Community Health* 35 (1): 45–50.
- Hodgson, Karen, Russell A. Poldrack, Joanne E. Curran, Emma E. Knowles, Samuel Mathias, Harald H. H. Göring, Nailin Yao, et al. 2017. "Shared Genetic Factors Influence Head Motion During MRI and Body Mass Index." *Cerebral Cortex* 27 (12): 5539–46.
- Holland, Dominic, Joshua M. Kuperman, and Anders M. Dale. 2010. "Efficient Correction of Inhomogeneous Static Magnetic Field-Induced Distortion in Echo Planar Imaging." *NeuroImage* 50 (1): 175–83.
- "<https://github.com/ExploreASL/ExploreASL/tree/EPAD>." n.d. <https://github.com/ExploreASL/ExploreASL/tree/EPAD>.

- <https://github.com/ExploreASL/ExploreASL/tree/EPAD>.
- Ingala, Silvia, Casper De Boer, Larissa A. Masselink, Ilaria Vergari, Luigi Lorenzini, Kaj Blennow, Gaël Chételat, et al. 2021. "Application of the ATN Classification Scheme in a Population without Dementia: Findings from the EPAD Cohort." *Alzheimer's & Dementia*. <https://doi.org/10.1002/alz.12292>.
- Jack, Clifford R., Jr, Matt A. Bernstein, Nick C. Fox, Paul Thompson, Gene Alexander, Danielle Harvey, Bret Borowski, et al. 2008. "The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI Methods." *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* 27 (4): 685–91.
- Jenkinson, Mark, Christian F. Beckmann, Timothy E. J. Behrens, Mark W. Woolrich, and Stephen M. Smith. 2012. "FSL." *NeuroImage*. <https://doi.org/10.1016/j.neuroimage.2011.09.015>.
- Juttukonda, Meher R., Binyin Li, Randa Almaktoom, Kimberly A. Stephens, Kathryn M. Yochim, Essa Yacoub, Randy L. Buckner, and David H. Salat. 2021. "Characterizing Cerebral Hemodynamics across the Adult Lifespan with Arterial Spin Labeling MRI Data from the Human Connectome Project-Aging." *NeuroImage* 230 (January): 117807.
- Li, Xiangrui, Paul S. Morgan, John Ashburner, Jolinda Smith, and Christopher Rorden. 2016. "The First Step for Neuroimaging Data Analysis: DICOM to NIFTI Conversion." *Journal of Neuroscience Methods* 264 (May): 47–56.
- McKiernan, Elizabeth Frances, Elijah Mak, Maria-Eleni Dounavi, Katie Wells, Craig Ritchie, Guy Williams, Li Su, and John O'Brien. 2020. "Regional Hyperperfusion in Cognitively Normal APOE ε4 Allele Carriers in Mid-Life: Analysis of ASL Pilot Data from the PREVENT-Dementia Cohort." *Journal of Neurology, Neurosurgery, and Psychiatry* 91 (8): 861–66.
- Mills, Kathryn L., and Christian K. Tamnes. 2014. "Methods and Considerations for Longitudinal Structural Brain Imaging Analysis across Development." *Developmental Cognitive Neuroscience*. <https://doi.org/10.1016/j.dcn.2014.04.004>.
- Mueller, Susanne G., Michael W. Weiner, Leon J. Thal, Ronald C. Petersen, Clifford R. Jack, William Jagust, John Q. Trojanowski, Arthur W. Toga, and Laurel Beckett. 2005. "Ways toward an Early Diagnosis in Alzheimer's Disease: The Alzheimer's Disease Neuroimaging Initiative (ADNI)." *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* 1 (1): 55–66.
- Mutsaerts, Henk J. M. M., Jan Petr, Paul Groot, Pieter Vandemaële, Silvia Ingala, Andrew D. Robertson, Lena Václavů, et al. 2020. "ExploreASL: An Image Processing Pipeline for Multi-Center ASL Perfusion MRI Studies." *NeuroImage* 219 (October): 117031.
- Nickerson, Lisa D., Stephen M. Smith, Döst Öngür, and Christian F. Beckmann. 2017. "Using Dual Regression to Investigate Network Shape and Amplitude in Functional Connectivity Analyses." *Frontiers in Neuroscience*. <https://doi.org/10.3389/fnins.2017.00115>.
- Nourani, Aynaz, Haleh Ayatollahi, and Masoud Soleymani Dodaran. 2019. "Clinical Trial Data Management Software: A Review of the Technical Features." *Reviews on Recent Clinical Trials* 14 (3): 160–72.
- Penny, William D., Karl J. Friston, John T. Ashburner, Stefan J. Kiebel, and Thomas E. Nichols. 2011. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier.
- Persson, Jonas, Sara Pudas, Lars-Göran Nilsson, and Lars Nyberg. 2014. "Longitudinal Assessment of Default-Mode Brain Function in Aging." *Neurobiology of Aging* 35 (9): 2107–17.
- Pfefferbaum, Adolf, Sandra Chanraud, Anne-Lise Pitel, Ajit Shankaranarayanan, David C. Alsop, Torsten Rohlfing, and Edith V. Sullivan. 2010. "Volumetric Cerebral Perfusion Imaging in Healthy Adults: Regional Distribution, Laterality, and Repeatability of Pulsed Continuous Arterial Spin Labeling (PCASL)." *Psychiatry Research* 182 (3): 266–73.
- Pizarro, Ricardo A., Xi Cheng, Alan Barnett, Herve Lemaitre, Beth A. Verchinski, Aaron L. Goldman, Ena Xiao, et al. 2016. "Automated Quality Assessment of Structural Magnetic Resonance Brain Images Based on a Supervised Machine Learning Algorithm." *Frontiers in Neuroinformatics* 10 (December): 52.
- Pomponio, Raymond, Guray Erus, Mohamad Habes, Jimit Doshi, Dhivya Srinivasan, Elizabeth Mamourian, Vishnu Bashyam, et al. n.d. "Harmonization of Large Multi-Site Imaging Datasets: Application to 10,232 MRIs for the Analysis of Imaging Patterns of Structural Brain Change throughout the Lifespan." <https://doi.org/10.1101/784363>.
- Raamana, P. R., A. Theyers, T. Selliah, P. Bhati, and S. R. Arnott. 2021. "Visual QC Protocol for Freesurfer Cortical Parcellations from Anatomical MRI." *BioRxiv*. <https://www.biorxiv.org/content/10.1101/2020.09.07.286807.abstract>.
- Salat, David H., Randy L. Buckner, Abraham Z. Snyder, Douglas N. Greve, Rahul S. R. Desikan, Evelina Busa, John C. Morris, Anders M. Dale, and Bruce Fischl. 2004. "Thinning of the Cerebral Cortex in Aging." *Cerebral Cortex* 14 (7): 721–30.
- Scheltens, Philip, Kaj Blennow, Monique M. B. Breteler, Bart de Strooper, Giovanni B. Frisoni, Stephen Salloway, and Wiesje Maria Van der Flier. 2016. "Alzheimer's Disease." *The Lancet* 388 (10043): 505–17.
- Schmidt, Paul, Viola Pongratz, Pascal Küster, Dominik Meier, Jens Wuerfel, Carsten Lukas, Barbara Bellenberg, et al. 2019. "Automated Segmentation of Changes in FLAIR-Hyperintense White Matter Lesions in Multiple Sclerosis on Serial Magnetic Resonance Imaging." *NeuroImage. Clinical* 23 (May): 101849.
- Shirzadi, Zahra, Bojana Stefanovic, Michael A. Chappell, Joel Ramirez, Graeme Schwindt, Mario Masellis, Sandra E. Black, and Bradley J. MacIntosh. 2018. "Enhancement of Automated Blood Flow Estimates (ENABLE) from Arterial Spin-Labeled MRI." *Journal of Magnetic Resonance Imaging*. <https://doi.org/10.1002/jmri.25807>.
- Smith, S. M., P. T. Fox, K. L. Miller, D. C. Glahn, P. M. Fox, C. E. Mackay, N. Filippini, et al. 2009. "Correspondence of the Brain's Functional Architecture during Activation and Rest." *Proceedings of the National Academy of Sciences*. <https://doi.org/10.1073/pnas.0905267106>.
- Smith, Stephen M. 2002. "Fast Robust Automated Brain Extraction." *Human Brain Mapping* 17 (3): 143–55.
- Smith, Stephen M., Mark Jenkinson, Heidi Johansen-Berg, Daniel Rueckert, Thomas E. Nichols, Clare E. Mackay, Kate E. Watkins, et al. 2006. "Tract-Based Spatial Statistics: Voxelwise Analysis of Multi-Subject Diffusion Data." *NeuroImage* 31 (4): 1487–1505.
- Smith, Stephen M., Mark Jenkinson, Mark W. Woolrich, Christian F. Beckmann, Timothy E. J. Behrens, Heidi Johansen-Berg, Peter R. Bannister, et al. 2004. "Advances in Functional and Structural MR Image Analysis and Implementation as FSL." *NeuroImage* 23 Suppl 1: S208–19.
- Solomon, Alina, Miia Kivipelto, José Luis Molinuevo, Brian Tom, Craig W. Ritchie, and EPAD Consortium. 2019. "European Prevention of Alzheimer's Dementia Longitudinal Cohort Study (EPAD LCS): Study Protocol." *BMJ Open* 8 (12): e021017.
- Sudlow, Cathie, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, et al. 2015. "UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age." *PLoS Medicine* 12 (3): e1001779.
- Sudre, Carole H., M. Jorge Cardoso, Willem H. Bouvy, Geert Jan Biessels, Josephine Barnes, and Sebastien Ourselin. 2015. "Bayesian Model Selection for Pathological Neuroimaging Data Applied to White Matter Lesion Segmentation." *IEEE Transactions on Medical Imaging* 34 (10): 2079–2102.
- Sudre, C. H., B. Gomez Anson, I. Davagnanam, A. Schmitt, A. F. Mendelson, F. Prados, L. Smith, et al. 2018. "Bullseye's Representation of Cerebral White Matter Hyperintensities." *Journal of Neuroradiology. Journal de Neuroradiologie* 45 (2): 114–22.
- Thompson, Paul M., Jason L. Stein, Sarah E. Medland, Derrek P. Hibar, Alejandro Arias Vasquez, Miguel E. Renteria, Roberto Toro, et al. 2014. "The ENIGMA Consortium: Large-Scale Collaborative Analyses of Neuroimaging and Genetic Data." *Brain Imaging and Behavior* 8 (2): 153–82.

- Vos, Sjoerd B., Chantal M. W. Tax, Peter R. Luijten, Sebastien Ourselin, Alexander Leemans, and Martijn Froeling. 2017. "The Importance of Correcting for Signal Drift in Diffusion MRI." *Magnetic Resonance in Medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* 77 (1): 285–99.
- Waber, Deborah P., Carl De Moor, Peter W. Forbes, C. Robert Almli, Kelly N. Botteron, Gabriel Leonard, Denise Milovan, Tomas Paus, Judith Rumsey, and Brain Development Cooperative Group. 2007. "The NIH MRI Study of Normal Brain Development: Performance of a Population Based Sample of Healthy Children Aged 6 to 18 Years on a Neuropsychological Battery." *Journal of the International Neuropsychological Society: JINS* 13 (5): 729–46.
- Wakana, Setsu, Hangyi Jiang, Lidia M. Nagae-Poetscher, Peter C. M. van Zijl, and Susumu Mori. 2004. "Fiber Tract-Based Atlas of Human White Matter Anatomy." *Radiology* 230 (1): 77–87.
- Wolz, Robin, Paul Aljabar, Joseph V. Hajnal, Alexander Hammers, Daniel Rueckert, and Alzheimer's Disease Neuroimaging Initiative. 2010. "LEAP: Learning Embeddings for Atlas Propagation." *NeuroImage* 49 (2): 1316–25.
- Zarrar, Shehzad, Giavasis Steven, Li Qingyang, Benhajali Yassine, Yan Chaogan, Yang Zhen, Milham Michael, Bellec Pierre, and Craddock Cameron. 2015. "The Preprocessed Connectomes Project Quality Assessment Protocol - a Resource for Measuring the Quality of MRI Data." *Frontiers in Neuroscience*. <https://doi.org/10.3389/conf.fnins.2015.91.00047>.
- Zhuang, Jiancheng, Jan Hrabec, Alayar Kangarlou, Dongrong Xu, Ravi Bansal, Craig A. Branch, and Bradley S. Peterson. 2006. "Correction of Eddy-Current Distortions in Diffusion Tensor Images Using the Known Directions and Strengths of Diffusion Gradients." *Journal of Magnetic Resonance Imaging*. <https://doi.org/10.1002/jmri.20727>.
- Zlata, Zvinka Z., Amanda Bischoff-Grethe, Chelsea C. Hays, Thomas T. Liu, M. J. Meloy, Robert A. Rissman, Mark W. Bondi, and Christina E. Wierenga. 2016. "Higher Brain Perfusion May Not Support Memory Functions in Cognitively Normal Carriers of the ApoE ε4 Allele Compared to Non-Carriers." *Frontiers in Aging Neuroscience* 8 (June): 151.