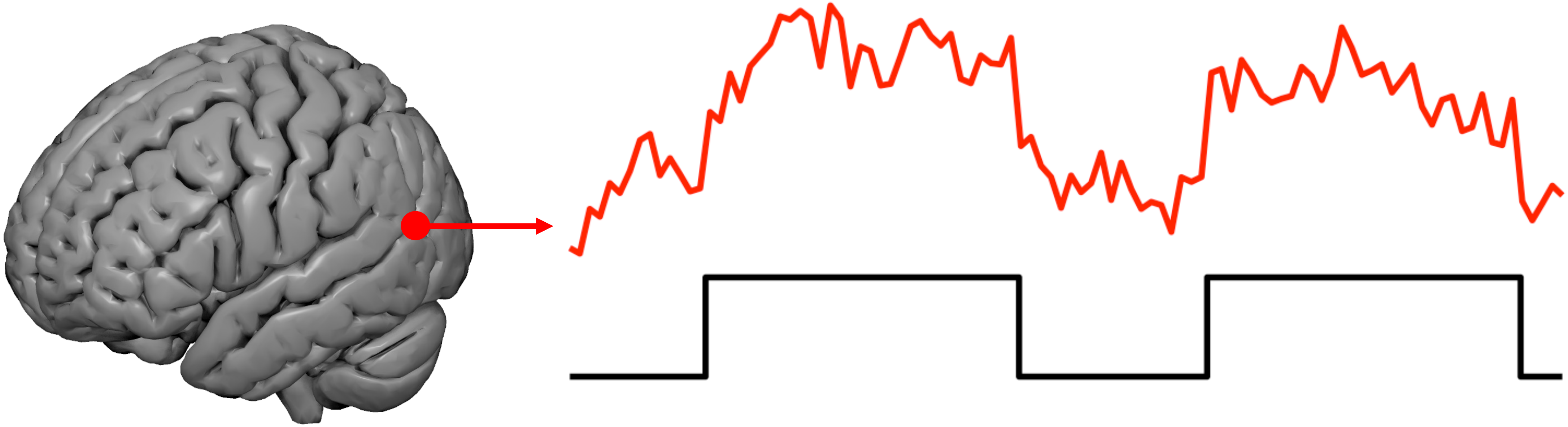
Preprocessing with fMRIPrep

Vesa Putkinen

Academy Research Fellow

Turku PET Centre

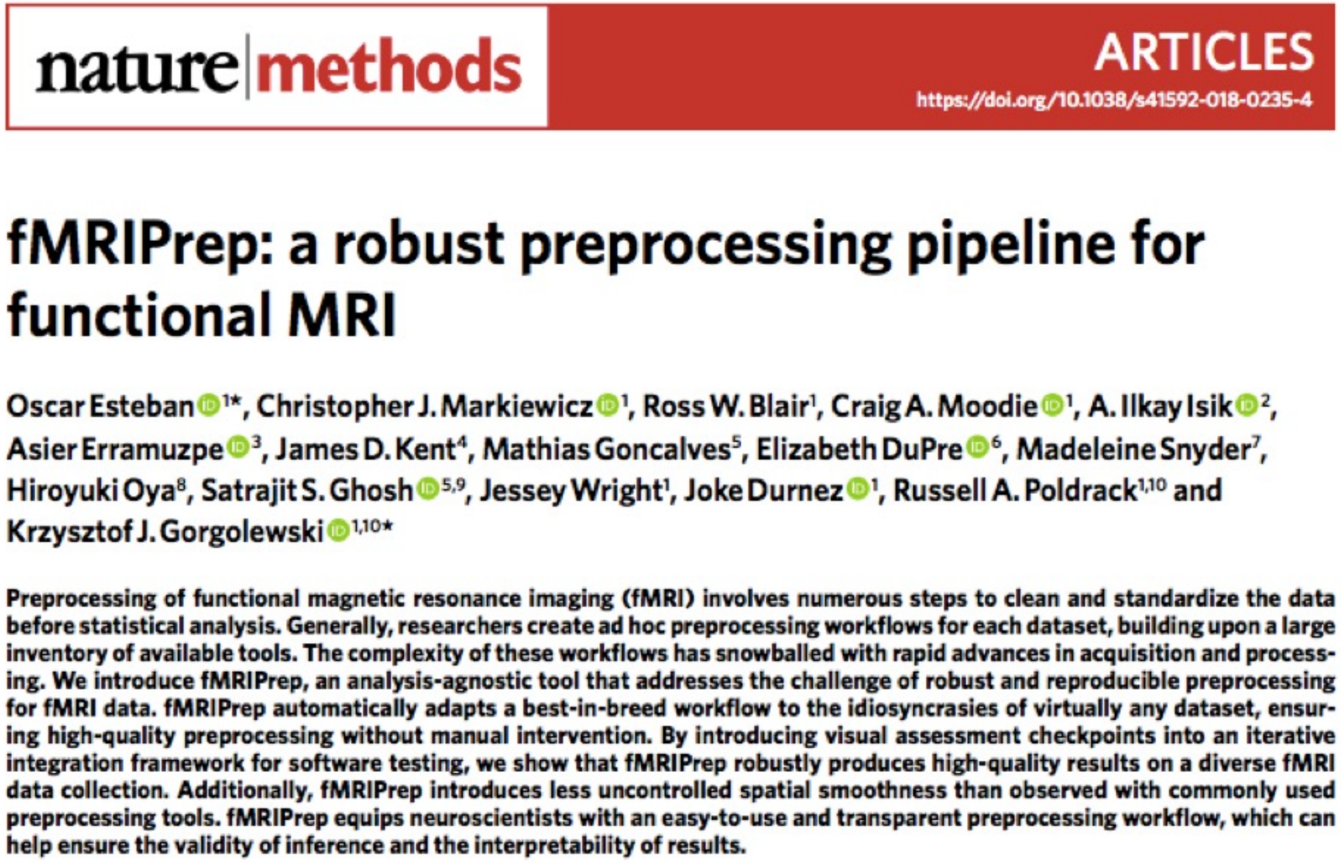


Preparing your data for statistical analysis

I. Dealing with artifacts

II. Normalization

fMRIPrep



nature methods

ARTICLES

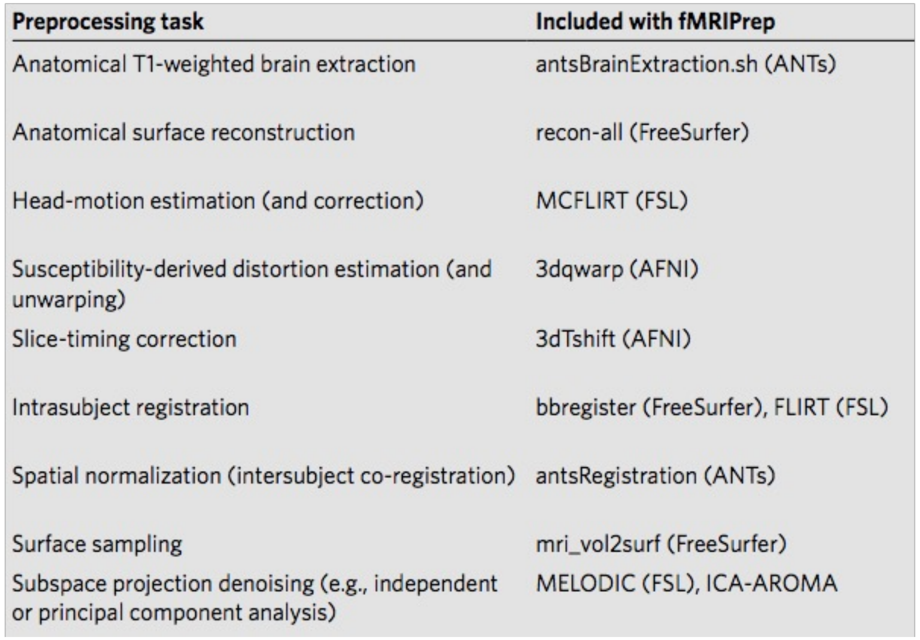
https://doi.org/10.1038/s41592-018-0235-4

fMRIPrep: a robust preprocessing pipeline for functional MRI

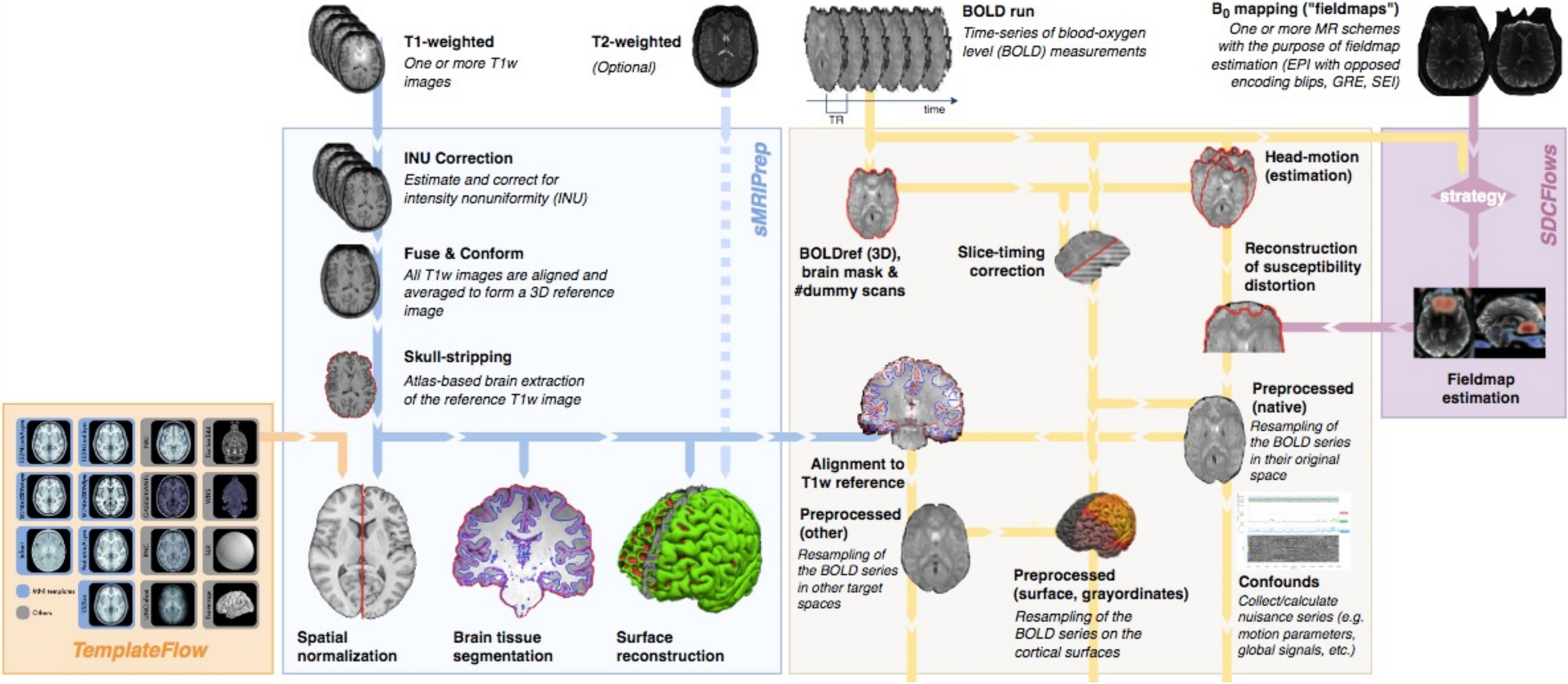
Oscar Esteban[1*], Christopher J. Markiewicz[1], Ross W. Blair[1], Craig A. Moodie[1], A. Ilkay Isik[2], Asier Erramuzpe[3], James D. Kent[4], Mathias Goncalves[5], Elizabeth DuPre[6], Madeleine Snyder[7], Hiroyuki Oya[8], Satrajit S. Ghosh[5,9], Jessey Wright[1], Joke Durnez[1], Russell A. Poldrack[1,10] and Krzysztof J. Gorgolewski[1,10*]

Preprocessing of functional magnetic resonance imaging (fMRI) involves numerous steps to clean and standardize the data before statistical analysis. Generally, researchers create ad hoc preprocessing workflows for each dataset, building upon a large inventory of available tools. The complexity of these workflows has snowballed with rapid advances in acquisition and processing. We introduce fMRIPrep, an analysis-agnostic tool that addresses the challenge of robust and reproducible preprocessing for fMRI data. fMRIPrep automatically adapts a best-in-breed workflow to the idiosyncrasies of virtually any dataset, ensuring high-quality preprocessing without manual intervention. By introducing visual assessment checkpoints into an iterative integration framework for software testing, we show that fMRIPrep robustly produces high-quality results on a diverse fMRI data collection. Additionally, fMRIPrep introduces less uncontrolled spatial smoothness than observed with commonly used preprocessing tools. fMRIPrep equips neuroscientists with an easy-to-use and transparent preprocessing workflow, which can help ensure the validity of inference and the interpretability of results.

Nipype:
Neuroimaging in Python
Pipelines and Interfaces



Preprocessing task	Included with fMRIPrep
Anatomical T1-weighted brain extraction	antsBrainExtraction.sh (ANTs)
Anatomical surface reconstruction	recon-all (FreeSurfer)
Head-motion estimation (and correction)	MCFLIRT (FSL)
Susceptibility-derived distortion estimation (and unwarping)	3dqwarp (AFNI)
Slice-timing correction	3dTshift (AFNI)
Intrasubject registration	bbregister (FreeSurfer), FLIRT (FSL)
Spatial normalization (intersubject co-registration)	antsRegistration (ANTs)
Surface sampling	mri_vol2surf (FreeSurfer)
Subspace projection denoising (e.g., independent or principal component analysis)	MELODIC (FSL), ICA-AROMA



T1-weighted
One or more T1w images

T2-weighted
(Optional)

BOLD run
Time-series of blood-oxygen level (BOLD) measurements

B_0 mapping ("fieldmaps")
One or more MR schemes with the purpose of fieldmap estimation (EPI with opposed encoding blips, GRE, SEI)

INU Correction
Estimate and correct for intensity nonuniformity (INU)

Fuse & Conform
All T1w images are aligned and averaged to form a 3D reference image

Skull-stripping
Atlas-based brain extraction of the reference T1w image

Spatial normalization

Brain tissue segmentation

Surface reconstruction

BOLDref (3D), brain mask & #dummy scans

Slice-timing correction

Head-motion (estimation)

Reconstruction of susceptibility distortion

Fieldmap estimation

Alignment to T1w reference

Preprocessed (native)
Resampling of the BOLD series in their original space

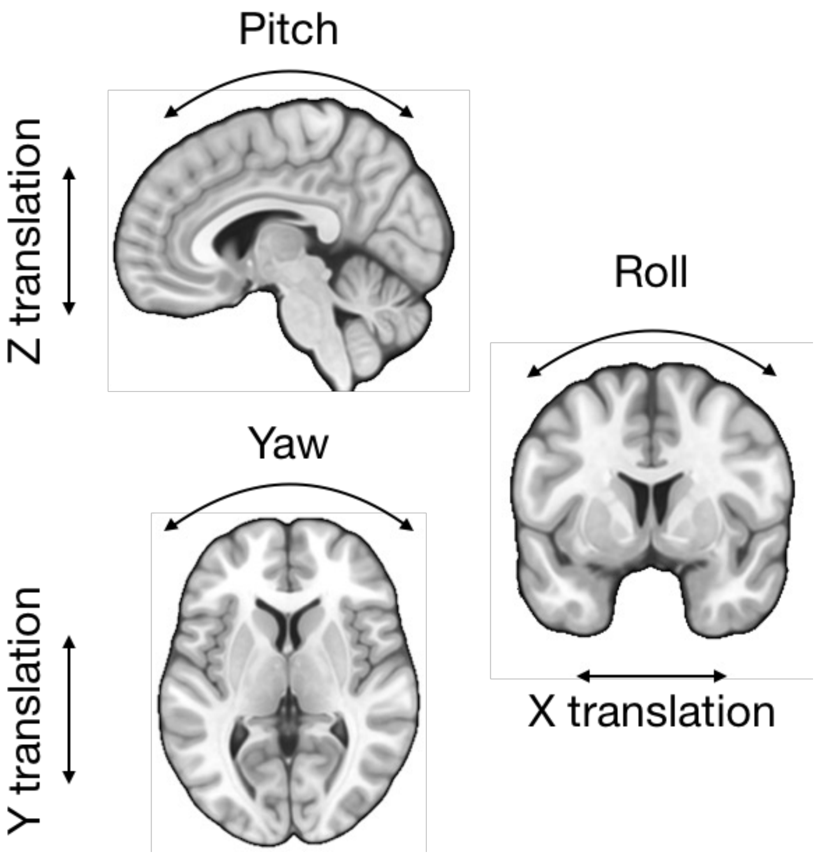
Preprocessed (other)
Resampling of the BOLD series in other target spaces

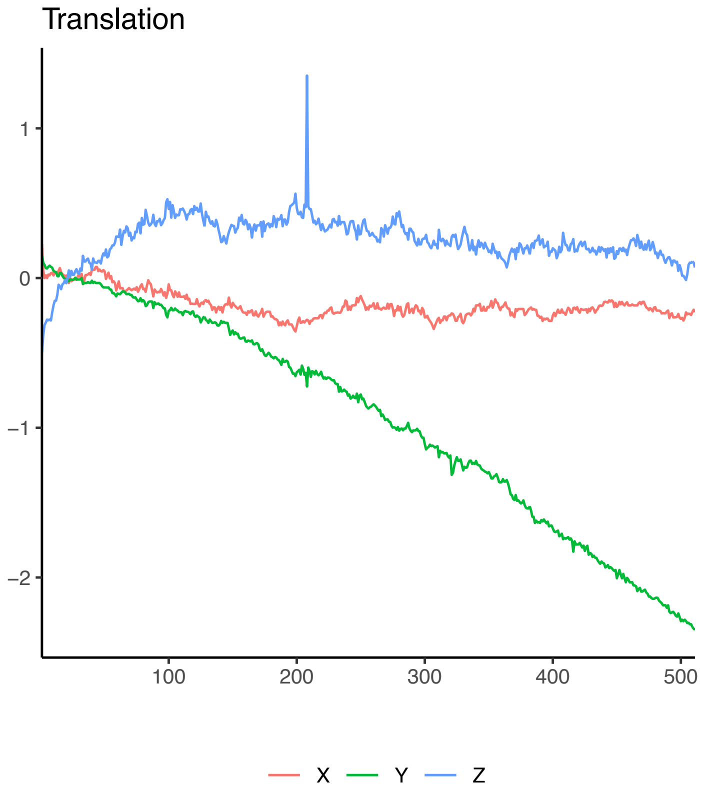
Preprocessed (surface, grayordinates)
Resampling of the BOLD series on the cortical surfaces

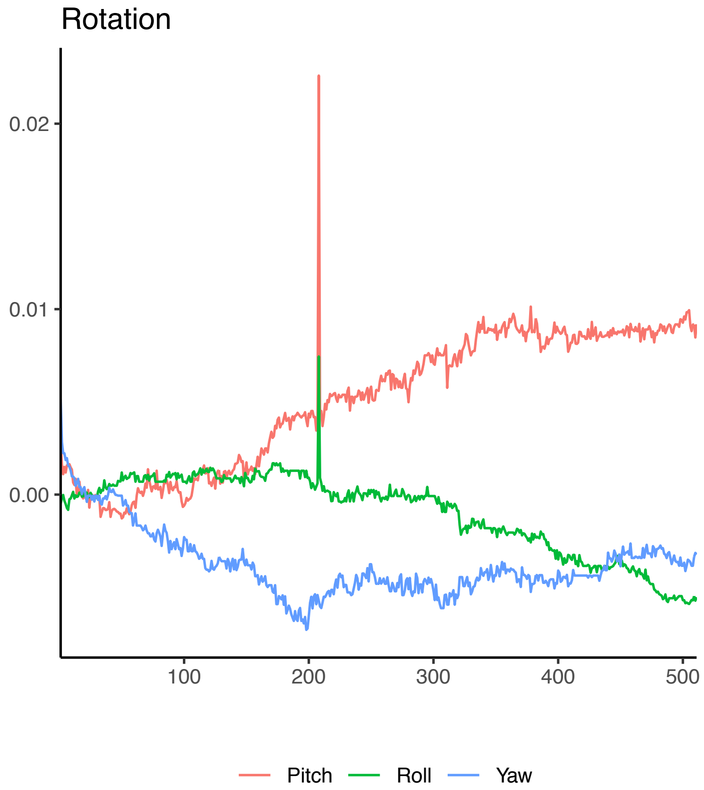
Confounds
Collect/calculate nuisance series (e.g. motion parameters, global signals, etc.)

Head Motion

Head Motion




Translation


Rotation

Head Motion

- fMRIPrep uses FSL’s MCFLIRT to estimate head motion
- Motion correction:

I. Selection of a reference volume

II. Registration of each frame to the reference frame through rigid-body transformations

- The goal is to realign all volumes to the reference volume

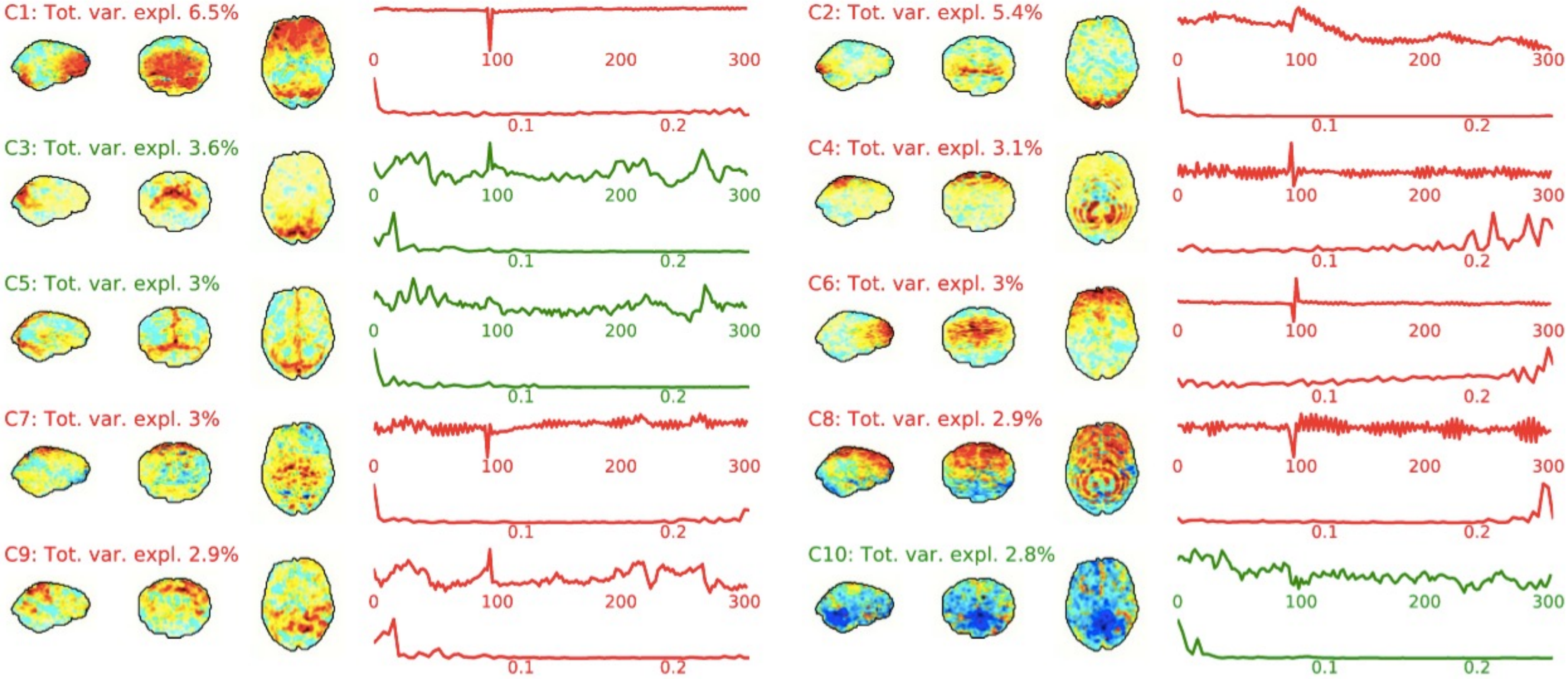
Frame 1 vs. frame 200 without motion correction

Frame 1 vs. frame 200 with motion correction with FSL’s mcflirt

ICA-AROMA Denoising

- **I**CA-AROMA can be used as a part of the preprocessing pipeline
  - **ICA** = Independent Component Analysis
  - **AROMA** = Automatic Removal of Motion Artifacts (Pruim et al., 2015).
- ICA decomposes the BOLD data into spatially independent component maps, and associated time-courses
- ICA-AROMA automatically identifies and removes motion-related independent components from the BOLD time series

ICA-AROMA Denoising

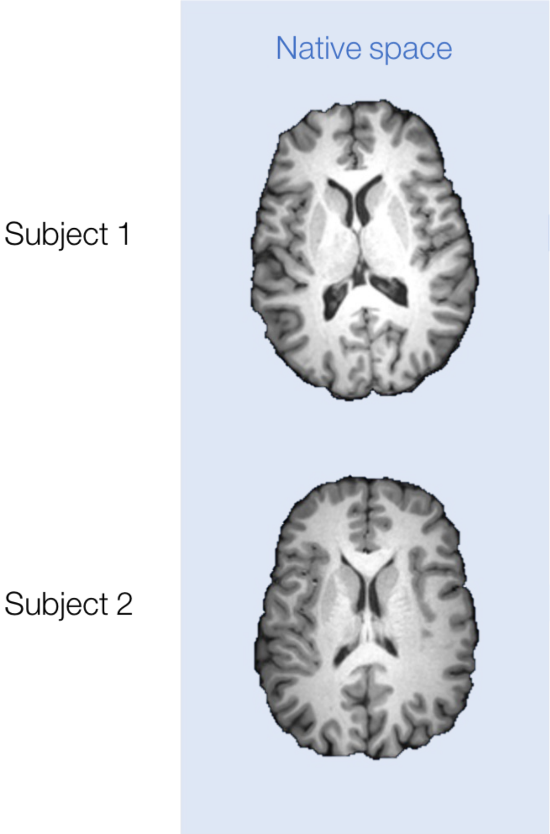


Spatial Normalization

Spatial Normalization

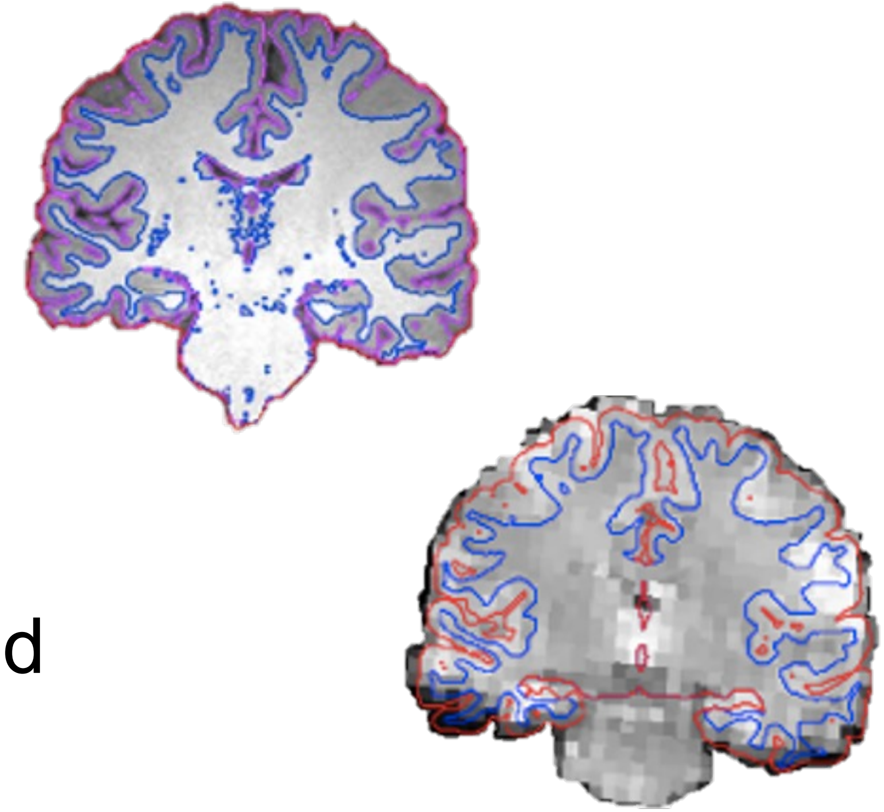
- The goal is to deform the images so that a given voxel corresponds (roughly) to the same region across subjects.
- The most commonly used standard space is the **MNI** (Montreal Neurological Institute) space.
- By default, fMRIPrep uses (Advanced Normalization tools) ANTs nonlinear registration (Avants et al., 2009) with MNI152NLin2009cAsym as the reference.

Spatial Normalization



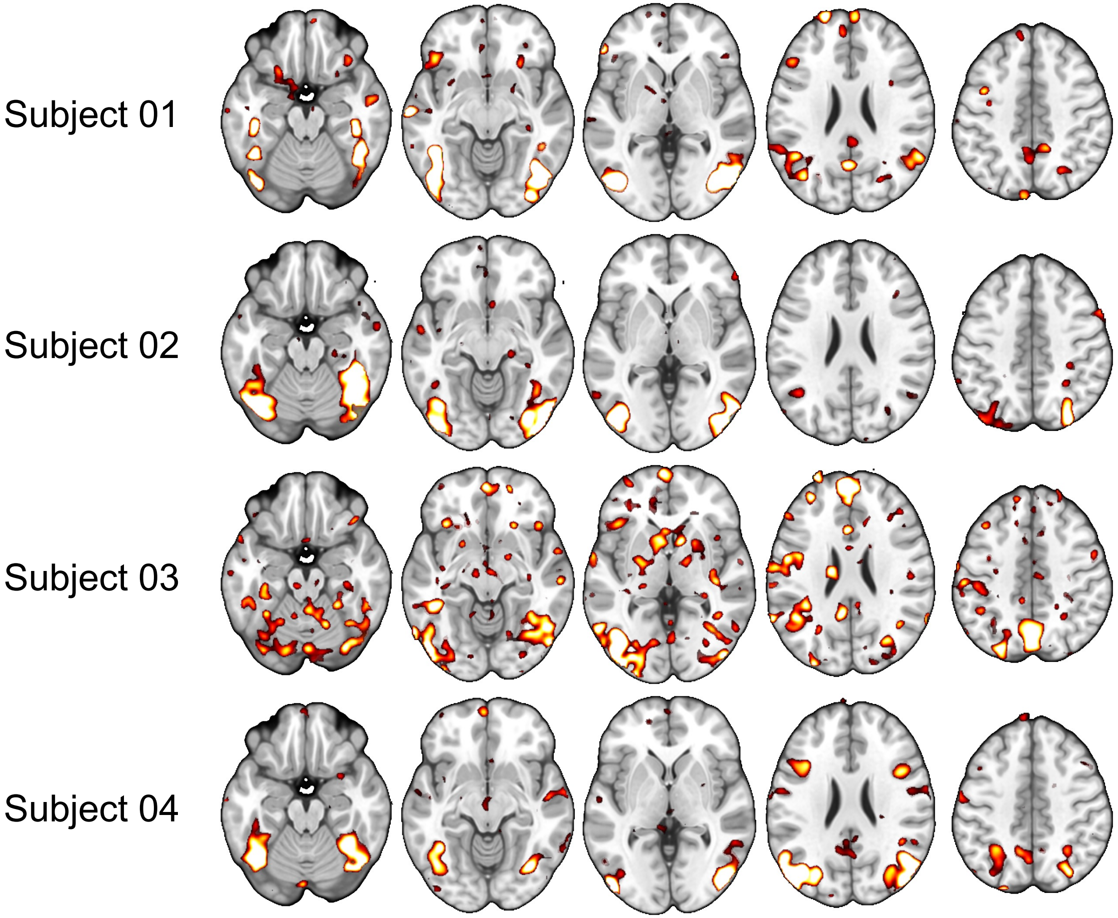
Co-registration

- Alignment of subject’s fMRI data with the separately acquired anatomical image
- fMRIPrep uses Boundary Based Registration (BBR) (Greve & Fischl, 2009) for co-registration
- Gray/white matter boundary is obtained either by Freesurfer or FSL’s FAST



1st level analysis

Subject-wise activation maps in standard space



Spatial Smoothing

Spatial Smoothing

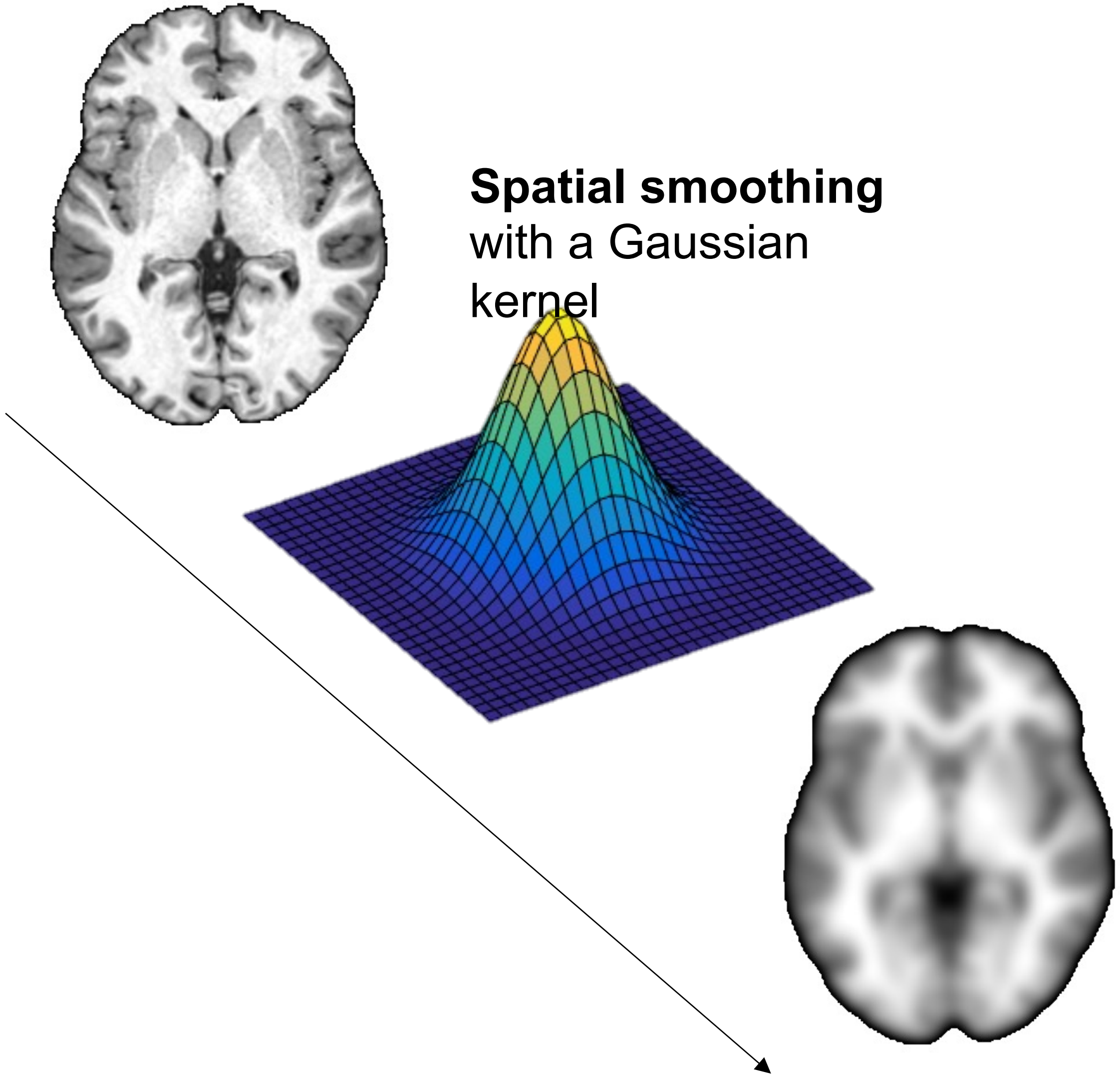
Possible benefits:

I. Improved SNR

II. More normally distributed data

III. Accomodation of intersubject-variability

By default, fMRIPrep does not perform spatial smoothing (however, smoothing is included in ICA-AROMA)

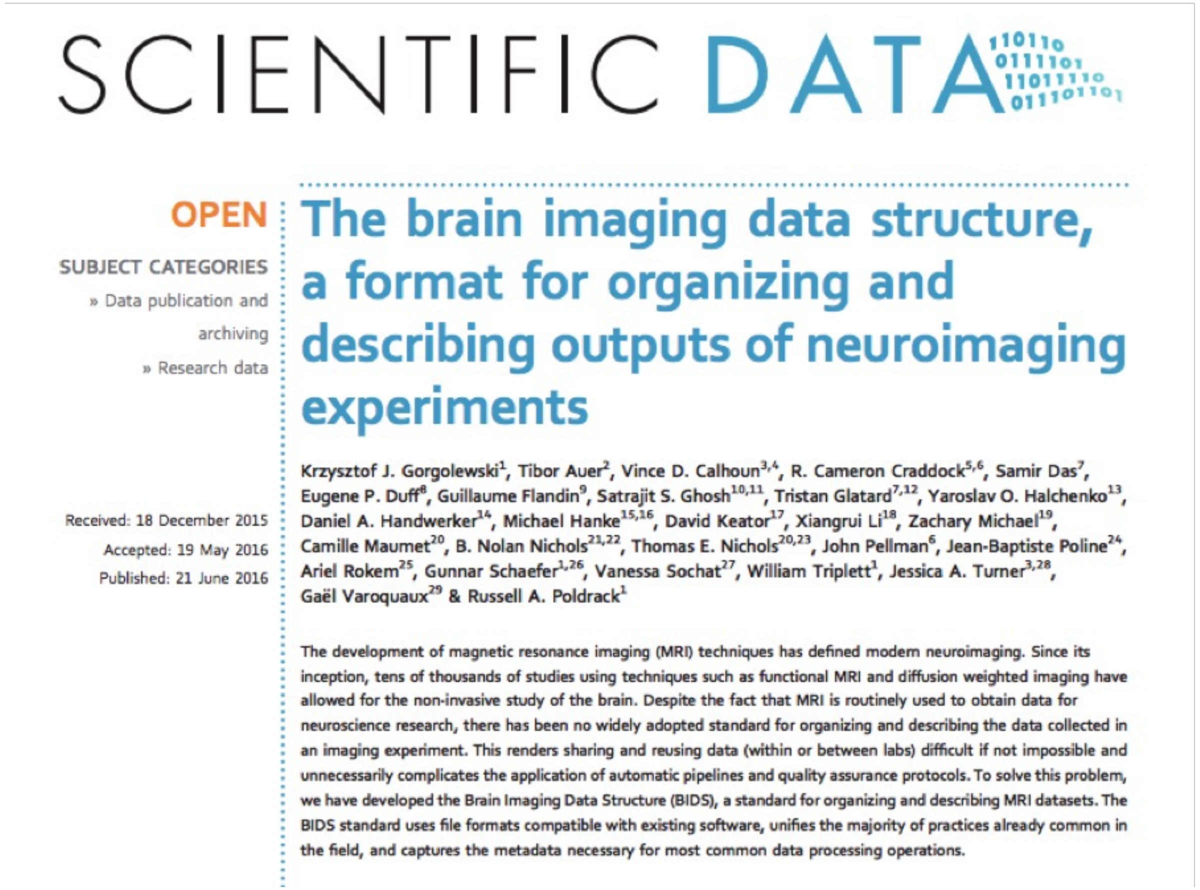


Spatial smoothing with a Gaussian kernel

Running fMRIPrep

Running fMRIPrep

- Convert your images into 4D nifti format
- Save **metadata** to a task-specific JSON file
- Make a **BIDS-compatible** folder structure
- Rename files in a **BIDS-compatible** manner



SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES
» Data publication and archiving
» Research data

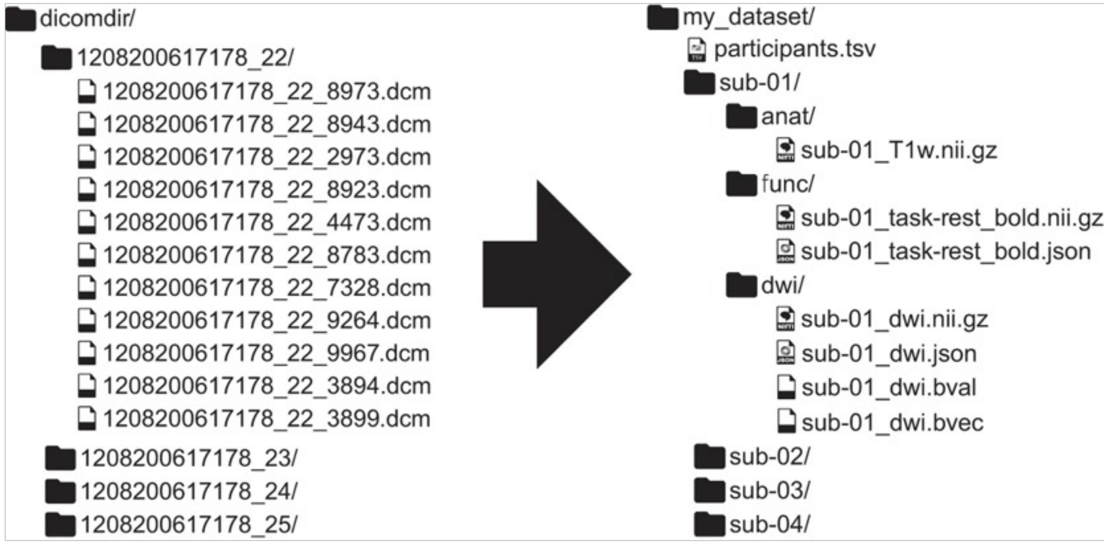
Received: 18 December 2015
Accepted: 19 May 2016
Published: 21 June 2016

The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments

Krzysztof J. Gorgolewski[1], Tibor Auer[2], Vince D. Calhoun[3,4], R. Cameron Craddock[5,6], Samir Das[7], Eugene P. Duff[8], Guillaume Flandin[9], Satrajit S. Ghosh[10,11], Tristan Glatard[7,12], Yaroslav O. Halchenko[13], Daniel A. Handwerker[14], Michael Hanke[15,16], David Keator[17], Xiangrui Li[18], Zachary Michael[19], Camille Maumet[20], B. Nolan Nichols[21,22], Thomas E. Nichols[20,23], John Pellman[6], Jean-Baptiste Poline[24], Ariel Rokem[25], Gunnar Schaefer[1,26], Vanessa Sochat[27], William Triplett[1], Jessica A. Turner[3,28], Gaël Varoquaux[29] & Russell A. Poldrack[1]

The development of magnetic resonance imaging (MRI) techniques has defined modern neuroimaging. Since its inception, tens of thousands of studies using techniques such as functional MRI and diffusion weighted imaging have allowed for the non-invasive study of the brain. Despite the fact that MRI is routinely used to obtain data for neuroscience research, there has been no widely adopted standard for organizing and describing the data collected in an imaging experiment. This renders sharing and reusing data (within or between labs) difficult if not impossible and unnecessarily complicates the application of automatic pipelines and quality assurance protocols. To solve this problem, we have developed the Brain Imaging Data Structure (BIDS), a standard for organizing and describing MRI datasets. The BIDS standard uses file formats compatible with existing software, unifies the majority of practices already common in the field, and captures the metadata necessary for most common data processing operations.

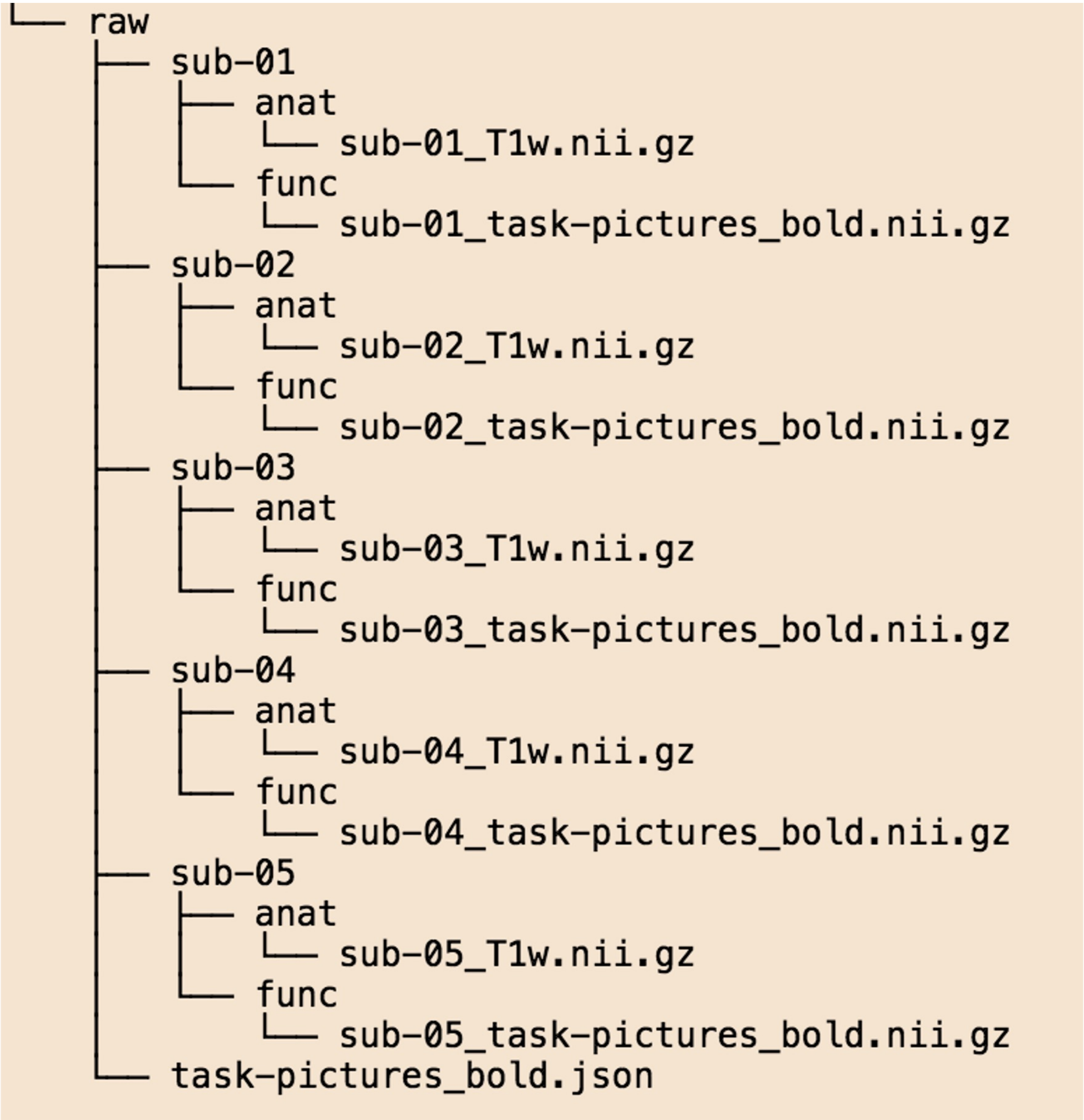
BIDS
BRAIN IMAGING DATA STRUCTURE



```
dicomdir/
  1208200617178_22/
    1208200617178_22_8973.dcm
    1208200617178_22_8943.dcm
    1208200617178_22_2973.dcm
    1208200617178_22_8923.dcm
    1208200617178_22_4473.dcm
    1208200617178_22_8783.dcm
    1208200617178_22_7328.dcm
    1208200617178_22_9264.dcm
    1208200617178_22_9967.dcm
    1208200617178_22_3894.dcm
    1208200617178_22_3899.dcm
  1208200617178_23/
  1208200617178_24/
  1208200617178_25/
```

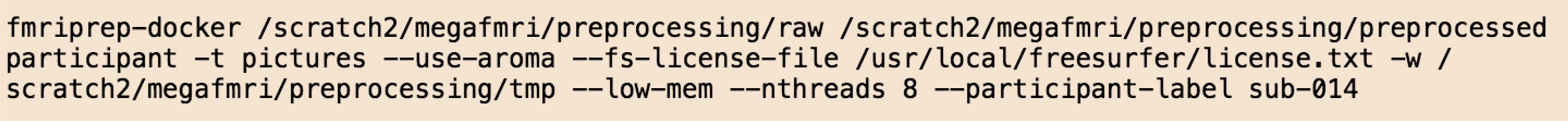
```
my_dataset/
  participants.tsv
  sub-01/
    anat/
      sub-01_T1w.nii.gz
    func/
      sub-01_task-rest_bold.nii.gz
      sub-01_task-rest_bold.json
    dwi/
      sub-01_dwi.nii.gz
      sub-01_dwi.json
      sub-01_dwi.bval
      sub-01_dwi.bvec
  sub-02/
  sub-03/
  sub-04/
```

https://bids.neuroimaging.io/



```
└── raw
    ├── sub-01
    │   ├── anat
    │   │   └── sub-01_T1w.nii.gz
    │   └── func
    │       └── sub-01_task-pictures_bold.nii.gz
    ├── sub-02
    │   ├── anat
    │   │   └── sub-02_T1w.nii.gz
    │   └── func
    │       └── sub-02_task-pictures_bold.nii.gz
    ├── sub-03
    │   ├── anat
    │   │   └── sub-03_T1w.nii.gz
    │   └── func
    │       └── sub-03_task-pictures_bold.nii.gz
    ├── sub-04
    │   ├── anat
    │   │   └── sub-04_T1w.nii.gz
    │   └── func
    │       └── sub-04_task-pictures_bold.nii.gz
    ├── sub-05
    │   ├── anat
    │   │   └── sub-05_T1w.nii.gz
    │   └── func
    │       └── sub-05_task-pictures_bold.nii.gz
    └── task-pictures_bold.json
```

Running fMRIPrep from the command line (docker wrapper)



```
fmriprep-docker /scratch2/megafmri/preprocessing/raw /scratch2/megafmri/preprocessing/preprocessed
participant -t pictures --use-aroma --fs-license-file /usr/local/freesurfer/license.txt -w /
scratch2/megafmri/preprocessing/tmp --low-mem --nthreads 8 --participant-label sub-014
```

fmriprep-docker <bids_dir> <output_dir> participant

-t: task to be processed

--use-aroma: use ICA_AROMA to remove movement related components

--fs-license-file: path to FreeSurfer license key file. Get it from https://surfer.nmr.mgh.harvard.edu/registration.html

-w: path where intermediate results are stored

--low-mem: reduce memory usage (will increase disk usage in working directory)

--nthreads: maximum number of threads across all processes

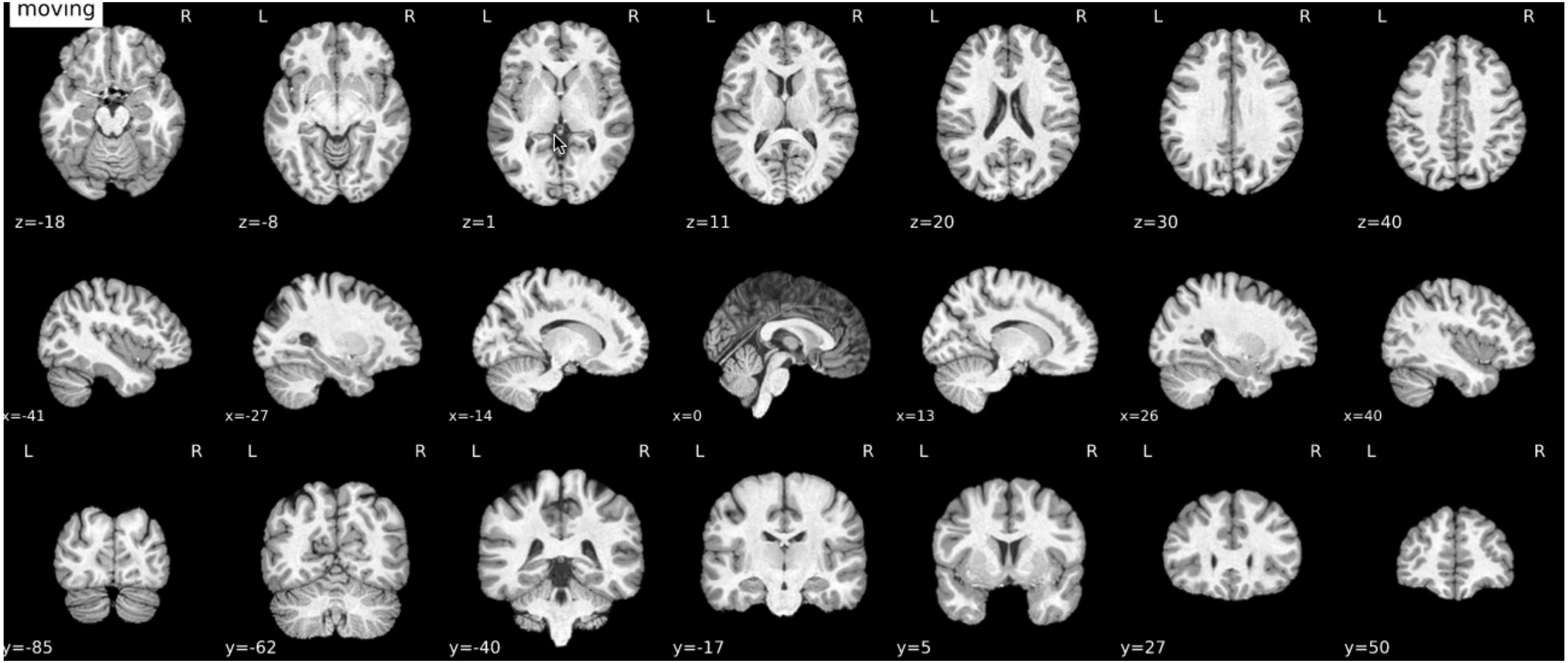
--participant-label: participant identifier

fMRIPrep Outputs

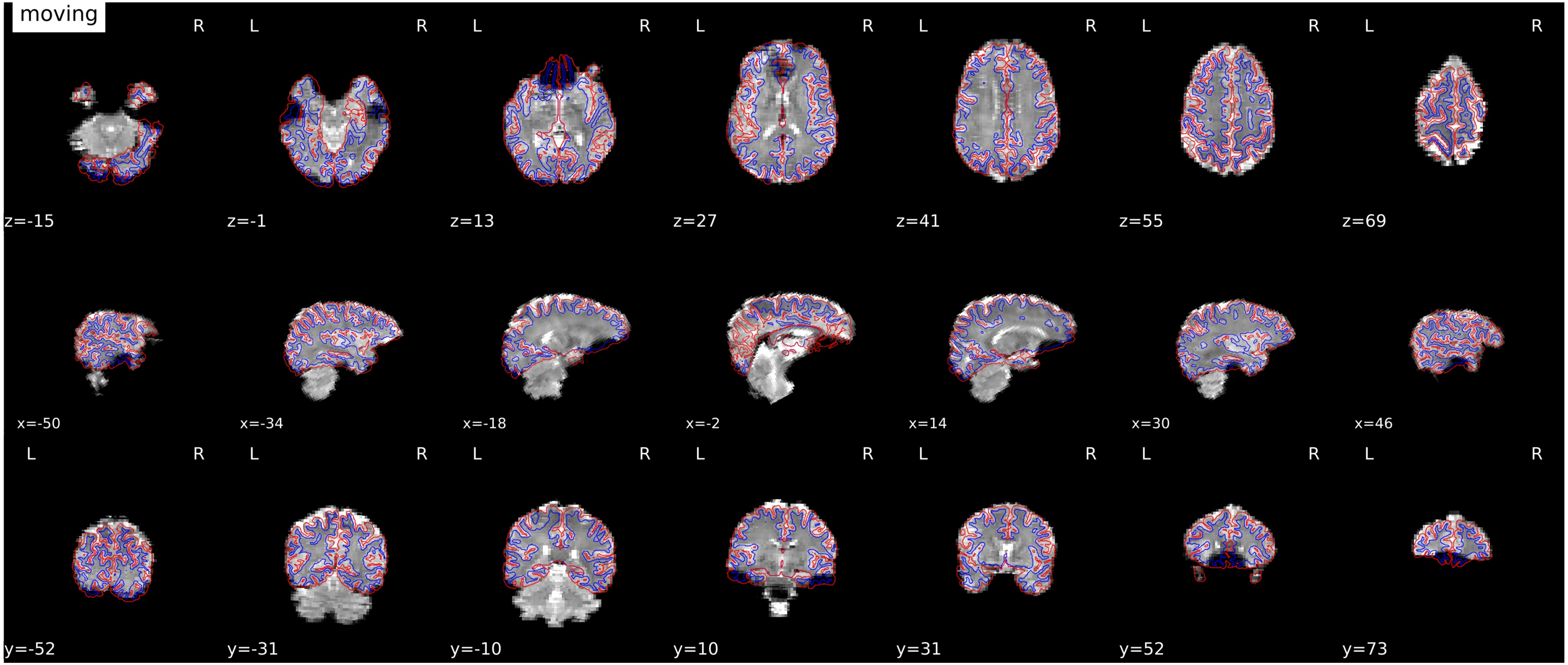
fRMIPrep derivates

```
./func/
├── sub-014_task-pictures_AROMAnoiseICs.csv
├── sub-014_task-pictures_desc-confounds_regressors.tsv
├── sub-014_task-pictures_desc-MELODIC_mixing.tsv
├── sub-014_task-pictures_space-fsaverage5_hemi-L.func.gii
├── sub-014_task-pictures_space-fsaverage5_hemi-R.func.gii
├── sub-014_task-pictures_space-MNI152NLin2009cAsym_boldref.nii.gz
├── sub-014_task-pictures_space-MNI152NLin2009cAsym_desc-aparcaseg_dseg.nii.gz
├── sub-014_task-pictures_space-MNI152NLin2009cAsym_desc-aseg_dseg.nii.gz
├── sub-014_task-pictures_space-MNI152NLin2009cAsym_desc-brain_mask.nii.gz
├── sub-014_task-pictures_space-MNI152NLin2009cAsym_desc-preproc_bold.nii.gz
└── sub-014_task-pictures_space-MNI152NLin2009cAsym_desc-smoothAROMAnonaggr_bold.nii.gz
```

QA: T1 Normalization



QA: Co-registration



https://fmriprep.org/

References

- Esteban, et al (2020). Analysis of task-based functional MRI data preprocessed with fMRIPrep. Nature Protocols, 15(7), 2186–2202.
- Esteban, et al. (2019). FMRIPrep: A robust preprocessing pipeline for functional MRI. Nature Methods, 16(1), 111.
- Gorgolewski, et al. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Scientific Data, 3(1), 160044.
- Pruim, et al. (2015). ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. Neuroimage, 112, 267-277.