Description of neuroimaging preprocessing details in fMRIPrep 20.2.5

Anatomical data preprocessing

The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection[37], distributed with ANTs 2.3.3 (RRID:SCR_004757)[38], and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white matter (WM) and gray matter (GM) was performed using fast (FSL 5.0.9, RRID:SCR_002823)[39]. Brain surfaces were reconstructed using recon-all (FreeSurfer 6.0.1, RRID:SCR_001847)[40], and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray- matter of Mindboggle (RRID:SCR_002438)[41]. Volume-based spatial normalization to two standard spaces (MNI152NLin2009cAsym, MNI152NLin6Asym) was performed through nonlinear registration with antsRegistration (ANTs 2.3.3), using brain- extracted versions of both T1w reference and the T1w template. The following templates were selected for spatial normalization: ICBM 152 Nonlinear Asymmetrical template version 2009c [RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym][42], FSL'sMNI ICBM 152 non-linear 6th Generation Asymmetric Average Brain Stereotaxic Registration Model [RRID:SCR_002823; TemplateFlow ID: MNI152NLin6Asym][43].

Functional data preprocessing

For each BOLD run found per subject, the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. Susceptibility distortion correction (SDC) was omitted. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration[44]. Co-registration was configured with six degrees of freedom. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9)[45]. BOLD runs were slice-time corrected to 1.16s (0.5 of slice acquisition range 0s-2.32s) using 3dTshift from AFNI 20160207 (RRID:SCR_005927)[46]. The BOLD time- series (including slice-timing correction when applied) were resampled onto their original, nativespace by applying the transforms to correct for head-motion. These resampled BOLD time-serieswill be referred to as preprocessed BOLD in original space, or just preprocessed BOLD. The BOLD time-series were resampled into several standard spaces, correspondingly generating the following spatially-normalized, preprocessed BOLD runs: MNI152NLin2009cAsym, MNI152NLin6Asym. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. Automatic removal of motion artifacts using independent component analysis (ICA-AROMA)[47] was performed on the preprocessed BOLD on MNI space time-series after removal of non-steady state volumes and

spatial smoothing with an isotropic, Gaussian kernel of 6mm FWHM (full-width half-maximum). Corresponding "non-aggressively" denoised runs were produced after such smoothing. Additionally, the "aggressive" noise-regressors were collected and placed in the corresponding confounds file. Several confounding time-series were calculated based on the preprocessed BOLD: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions)[48] and Jenkinson (relative root mean square displacement betweenaffines)[45]. FD and DVARS are calculated for each functional run, both using their implementations in Nipype (following the definitions by Power et al.[48].). Thethree global signals are extracted within the CSF, the WM, and the whole-brain masks.

Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (CompCor[49]). Principal components are estimated after high-pass filtering the preprocessed BOLD time-series (using a discrete cosine filter with 128s cut- off) for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, the aCompCor masks are subtracted from a mask of pixels that likely contain a volume fraction of GM. This mask is obtained by dilating a GM mask extracted from the FreeSurfer's aseg segmentation, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. Foreach CompCor decomposition, the k components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from headmotion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each[50]. Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardized DVARS were annotated as motion outliers. All resamplings can be performed with a single interpolation step by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co- registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels[51]. Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer).

Results included in this manuscript come from preprocessing performed using fMRIPrep 20.2.5[52,53] (RRID:SCR_016216), which is based on Nipype 1.6.1[54,55] (RRID:SCR_002502).

Many internal operations of fMRIPrep use Nilearn 0.6.2 (RRID:SCR_001362[56]), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in fMRIPrep's documentation.