PREPROCESSING OF FMRI IN A NUTSHELL

Head-motion correction (HMC). Functional MRI (fMRI) data are usually acquired with echo-planar imaging (EPI) or other fast MR acquisition techniques in order to estimate the timecourse of blood oxygenation signals. During that time, the blood-oxygen-level dependent (BOLD) signal is measured at a pre-specified frequency called repetition time (TR), typically selected in a range of 300ms for the fastest sequences and 3s for the legacy, traditional acquisitions. Even at the shortest TRs and scanning the most cooperative participants, the position of the head with respect to the scanner will change to some degree. Head motion occurs as a slow drift from the original position or suddenly when triggered, for example, by swallowing, task responses or involuntary movements of both healthy and pathological subjects. Motion causes a spatial misalignment in the source of the BOLD signal measured in all voxels at a certain TR and its neighbors in the time-series. The effect is particularly noticeable at the brain edges, where the signal may fall inside or outside the brain mask computed on the spatial average of the time-series.

Susceptibility-derived distortion correction (SDC). Magnetic susceptibility of tissues locally disturbs the B_0 field. Even though the tissue susceptibility is quite homogeneous across the brain parenchyma, those regions near air-filled cavities such as sinuses or the ear-canals may induce variations on the field. In addition, one limitation of fast MR acquisitions like EPI is the reduced bandwidth on the phase-encoding (PE) direction. The so-called *susceptibility distortion* renders as a warping along the PE direction of the registered BOLD signal, and deformation is proportional in magnitude to the inhomogeneity of the field at that point in space. Several techniques to estimate the variance of the field that induces the distortion have been proposed, such as field mapping techniques[1], the acquisition of several reference images with different PE directions that allow for estimating the field map through intra-modal image registration[2,3], using multi-modal image registration to a structural, "anatomically-correct" scan of the same subject (e.g. a T2-weighted (T2w) image[4], a T1-weighted (T1w) image[5]), or surface-driven registration[6]. Once the field map is estimated, the distortion can be computed[7] as a nonlinear map of displacements along the PE direction (oftentimes called voxel-shift map).

Slice-timing correction (STC). The MR schemes commonly utilized to acquire BOLD signal work with staggered scanning of 2D slices. Although all 2D slices are acquired very closely to one another, the hemodynamic response is sampled at different points in time. This effect is alleviated but still present in multi-band acquisitions that are able to acquire several 2D slices in parallel. To correct for slice timing it is necessary to record the relative time each slice is registered with respect to the first one and fit the hemodynamic response to then interpolate the signal at the reference TR[8]. Slice timing effects interact with head motion, raising the question of which one should be accounted for. Some correction methods are able to tackle both issues simultaneously by merging both interpolation processes[9].

Intra-subject registration. Head motion also occurs between the acquisition of the different MR sequences defined in the protocol, especially when the experiment involves running different sessions. In order to allow fusing information from the different runs in the protocol, an image registration process between them is necessary. A further complication to the image registration process is the fact that the different MR sequences utilized in the protocol pose a multi-modal registration problem.

Spatial normalization. In order to enable group comparisons, it is necessary to find a spatial mapping that aligns the subject-space into a common space, generally defined using an atlas. After spatial normalization, the information given by the different images in subject-space (after all intra-subject registrations have been calculated) can be analyzed *as if all the participants shared the same underlying anatomy*.

Surface sampling. Although most studies have analyzed fMRI analysis in the three-dimensional volume, it is increasingly common to analyze the data on the cortical surface e.g.[10,11]. The cortical surface of the brain is reconstructed from the anatomical (T1w and optionally T2w) images. Using the alignment described previously (intra-subject registration), these surfaces are mapped to the BOLD scan space and
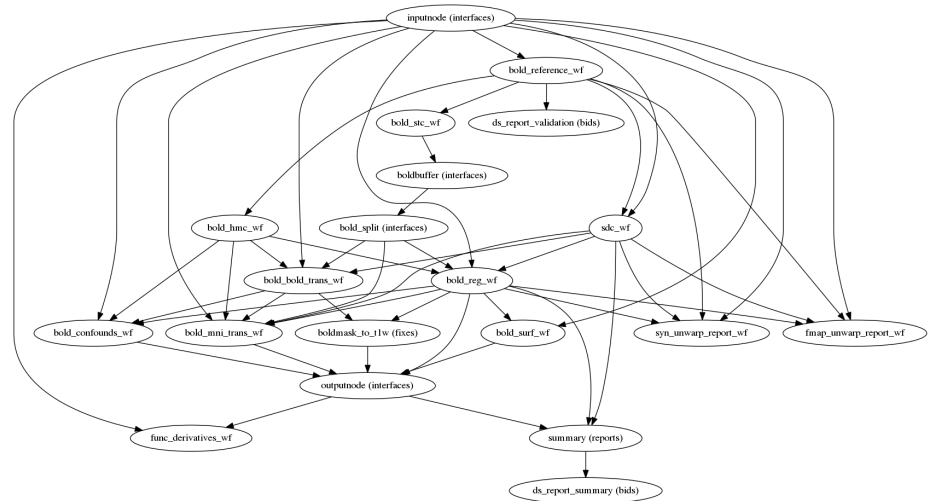
Figure S1. The overview of *fMRIPrep*'s workflow for functional preprocessing. The graph corresponding to the BOLD processing workflow of a dataset with slice-time information, one of the possible choices of field maps and surface reconstruction enabled in the anatomical processing branch (not shown in this graph representation). The graph also shows that BOLD data will be resampled into both MNI space and the subject's corresponding T1w space.

the signal is sampled at the vertices of the surface.

Other preprocessing elements. Other preprocessing tasks include the **spatio-temporal smoothing** of the signal, the **detection of spikes** and their removal, the identification of **non-steady state** (strong T1 contrast) at the beginning of the time-series, **frame scrubbing or censoring** in resting-state fMRI, the **calculation of masks and regions-of-interest (ROIs)**, the **calculation of confounds**, etc. see [12,13].

THE PREPROCESSING WORKFLOW OF *FMRIPREP*

The general overview and details of processing are thoroughly documented

During the workflow set-up, *fMRIPrep* applies a set of heuristics to determine the appropriate processing operations for the input dataset (see Main Document, *Figure 1*). All available sub-workflows implementing individual processing steps that can be combined in the final processing pipeline are described in the documentation website[1]. *Figure S1* shows how the right-hand side of *Figure 1* (Main Document) translates into the pipeline graph, as implemented with Nipype.

Particular processing elements of *fMRIPrep*

Generating BOLD references for image registration. Although generating realignment and co-registration references may seem like a clerical task, doing so reliably for any dataset is not trivial. *FMRIPrep* integrates a BOLD reference generation workflow that adapts to the contrast of the input image. The workflow was tested to generate over ten thousand binary brain masks, of which two thousand were screened to ensure they did not exclude extensive areas of the brain nor included substantial background volumes. *Figure S2* describes the workflow and provides further details on the generation of BOLD references.

Fieldmap-less susceptibility-derived distortion correction (SDC). This workflow takes a skull-stripped T1w image and reference BOLD image, and estimates a field of displacements that compensates for the warp caused by susceptibility distortion. The tool uses ANTs' antsRegistration configured with symmetric normalization (SyN) to align a fieldmap template [18] and applies the template as prior information to regularize a follow-up registration process. The follow-up registration process also uses antsRegistration with SyN deformation, with displacements restricted to the PE direction. If no PE direction is specified, anterior-posterior PE is assumed. Based on the fieldmap atlas, the displacement field is
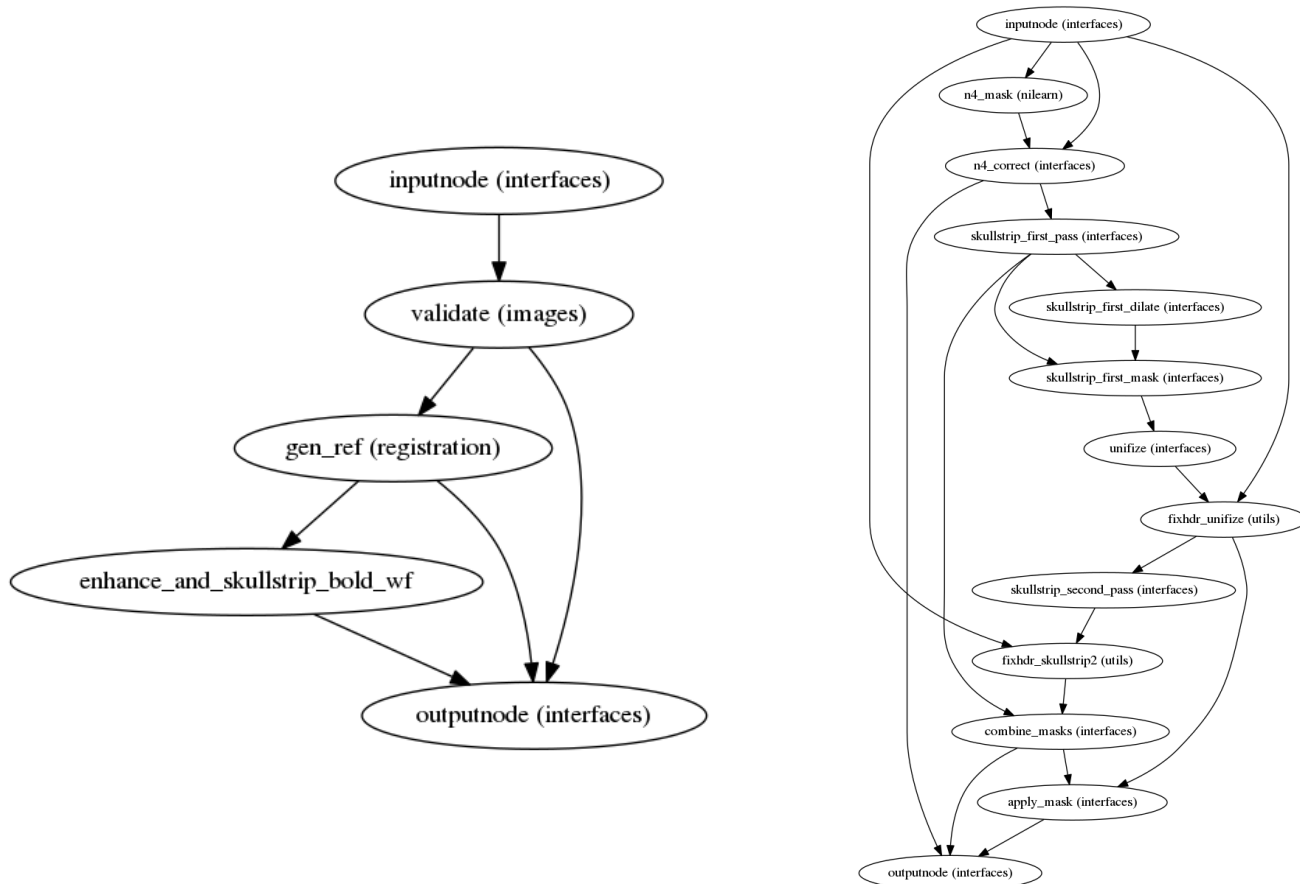
---

[1] https://fmriprep.readthedocs.io/en/latest/workflows.html

Figure S2. Left | Workflow for the generation of a BOLD reference. The input BOLD image is first validated and the *xform* information (q-form and s-form matrices) are checked. Then, volumes with substantial T1 contrast derived from nonsteady states of the scanner (typically found in the beginning of EPI schemes) are identified, realigned, and averaged to generate a reference (within the node labeled gen_ref). That reference is then contrast-enhanced and a brain mask is calculated by the enhance_and_skullstrip_bold_wf workflow, presented on the right. **Right | Workflow for enhancement and skull-stripping of the BOLD reference**. The reference undergoes several interim masking and contrast enhancement processes to robustly find the outline of the brain from the BOLD average. The process involves the use of ANT's N4BiasFieldCorrection[14], FSL's bet[15], Nilearn's compute_epi_mask[16], AFNI's 3dAutomask[17], and *in-house* image enhancement heuristics.

optimized only within regions that are expected to have a >3mm (approximately 1 voxel) warp. This technique is a variation on previous work[5,19].

Organization of preprocessing results. The results from preprocessing with *fMRIPrep* follow the "Derivatives Extension" of BIDS. BIDS-Derivatives is inspired by the original BIDS for *raw* data. Therefore, results are sorted by subject in folders under the `fmriprep/` folder. In each subject folder, two folders contain the derivatives from processing the anatomical reference (under the `anat/` folder) and the functional data (`func/` folder). Results are generally 3D or 4D brain data objects (represented with brain icons in the filesystem tree of *Figure S4*). They also contain one "`<derivative_prefix>_confounds.tsv`" corresponding to each fMRI run. These *confounds files* contain specific regressors for different noise sources that can be included in analysis. The full description of *fMRIPrep*'s outputs is documented in fmriprep.readthedocs.io/en/latest/outputs.html.

EVALUATION OF *FMRIPREP*

Iterative quality and robustness assurance

Data. Participants were drawn from a multiplicity of studies available in OpenfMRI, accessed on September 30, 2017. Studies were sampled uniformly (four participants each), except for *DS000031* that consists of only one participant. Data selection criteria are described below. Magnetic resonance imaging (MRI) data were acquired at multiple scanning centers, with the following frequencies of vendors: ~70% SIEMENS, ~14% PHILIPS, ~14% GE. Data were acquired by 1.5T and 3T systems running varying software versions. Acquisition protocols, as well as the particular acquisition parameters (including relevant BOLD settings such as the TR, the TE, the number of TRs and the resolution) also varied with each study. However, only datasets including at least one T1w and one BOLD per subject run were included. Datasets containing BIDS errors (*DS000210*), and degenerate data (many T1w images of *DS000223* are skull-stripped) at the time of access were discarded. Similarly, very-narrow field-of-view (FoV) BOLD datasets (*DS000172*, *DS000217*, and *DS000232*) were also excluded. In total, 54 datasets (46 single-session datasets, 8 multi-session) were included in this assessment (Main Document, *Table 2*). Data access, download and management were performed using datalad[21]. The selection of OpenfMRI as the data source ensures large heterogeneity in terms of acquisition protocols, settings, instruments and parameters that is necessary to demonstrate the robustness of *fMRIPrep* against the variability in input data features. Main Document, *Table 2* overviews the particular properties of each dataset, summarizing the large heterogeneity of the resource.

Data coverage. Data coverage is a metric used in software engineering that measures the area that a given test or test-set covers with respect to the full domain of possible input data. This evaluation covered 54 studies out of a total of 58 studies in OpenfMRI that included the two required imaging modalities (T1w and BOLD). Therefore, these tests covered 93% of the studies in OpenfMRI.

Methodology and test plan. To ensure that *fMRIPrep* fulfills the specifications on reliability and scientific-software standards, the tool undergoes a thorough acceptance testing plan. The plan is structured in three phases: the first was aimed at the discovery of faults, the second at the evaluation of the robustness, and the final phase at the full coverage of OpenfMRI. To note, an early test **Phase 0** was conducted as a proof of concept for the tool[22]. During **Phase I**, a total of 120 subjects from 30 different datasets were manually identified as low-quality using MRIQC[23]. This sub-sample of OpenfMRI underwent pre-processing in the *Stampede2* supercomputer of the Texas Advanced Computer Center (TACC), Austin, TX. Results were visually inspected and failures reported in the GitHub repository. Once the faults were fixed, **Phase II** was launched. In this second phase, the coverage of OpenfMRI was extended to all 60 available datasets, 326 participants randomly (with replacement of the participants of the previous phase) selecting four participants per dataset (except for the dataset with accession numbers *ds000031*, which has a sample size of one). In Phase II, the inspection protocol was more thorough, as specified in the Testing Plan document. Finally, with the submission of this work **Phase III** will be kicked-off. This phase will end when the full coverage of OpenfMRI is reached.
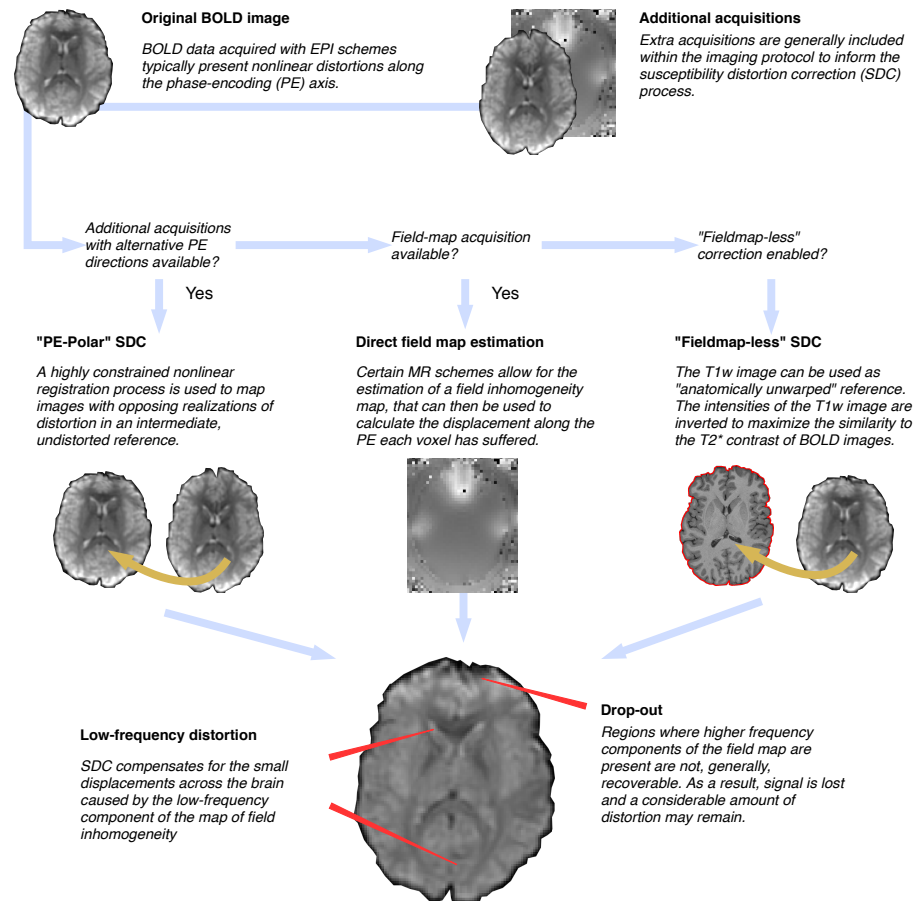
Original BOLD image

BOLD data acquired with EPI schemes typically present nonlinear distortions along the phase-encoding (PE) axis.

Additional acquisitions

Extra acquisitions are generally included within the imaging protocol to inform the susceptibility distortion correction (SDC) process.

Additional acquisitions with alternative PE directions available?

Field-map acquisition available?

"Fieldmap-less" correction enabled?

Yes

Yes

"PE-Polar" SDC

A highly constrained nonlinear registration process is used to map images with opposing realizations of distortion in an intermediate, undistorted reference.

Direct field map estimation

Certain MR schemes allow for the estimation of a field inhomogeneity map, that can then be used to calculate the displacement along the PE each voxel has suffered.

"Fieldmap-less" SDC

The T1w image can be used as "anatomically unwarped" reference. The intensities of the T1w image are inverted to maximize the similarity to the T2 contrast of BOLD images.*

Low-frequency distortion

SDC compensates for the small displacements across the brain caused by the low-frequency component of the map of field inhomogeneity

Drop-out

Regions where higher frequency components of the field map are present are not, generally, recoverable. As a result, signal is lost and a considerable amount of distortion may remain.

Figure S3. The overview of *fMRIPrep*'s workflow for susceptibility-derived distortion correction (SDC). The BIDS [20] (Brain Imaging Data Structure) structure is hierarchically queried to discover whether extra acquisitions containing field map information are available. The highest priority is given to the "PE-Polar" (phase-encoding POLARity) approach. In this case, in addition to the BOLD time-series, an extra EPI image is acquired with all parameters matching those of the BOLD data, except for the PE direction. Changing either the axis or (more often) the polarity of the PE results in a different realization of the distortion. "PE-Polar" methods establish a highly constrained registration framework between the two images of the same object with different distortions to estimate the underlying field map. If extra EPI scans for "PE-Polar" approaches are not available, then *fMRIPrep* queries for actual field mapping acquisitions. Typically, two gradient-recalled echoes can be acquired. Measuring the phase drift between the two echoes, and knowing the time delta between them, it is possible to map the inhomogeneity of the field. Once the field map is estimated, distortion is proportional to the inhomogeneity. In the last resort, when no field map information was acquired, *fMRIPrep* may run the "fieldmap-less" correction option (when enabled). The three methodologies will compensate for the low-frequency components of the distortion. However, current BOLD acquisitions are very fast at the cost of accepting the complete deletion (drop-out) of signal in particular regions. These regions are typically located around the orbitofrontal lobe (due to the proximity of the sinuses) and the temporal lobes (near the ear canals).
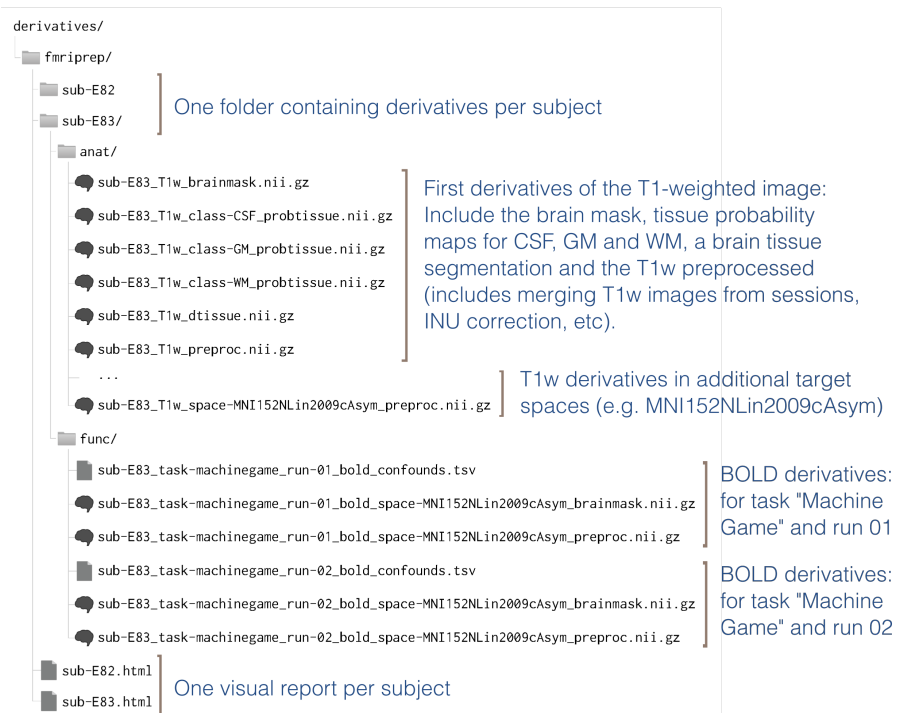
Figure S4. Example of the organization of *fMRIPrep*'s processing outcomes. The BIDS-Derivatives draft describes how the results from preprocessing can be organized. In this figure, imaging data are represented with brain icons, while additional text metadata and the visual reports are represented with a paper sheet.
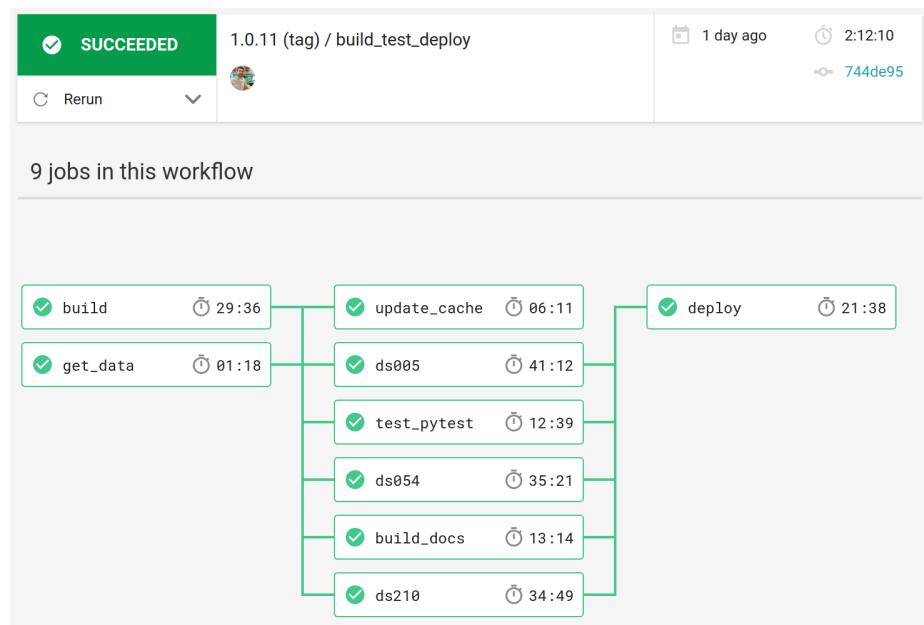

Figure S5. Continuous Integration testing and deployment. Every code iteration adding features or fixing defects triggers a battery of automated tests. The full documentation website is regenerated, some unit tests are performed and *fMRIPrep* is also run on three datasets from OpenfMRI (i.e. *DS000005*, *DS000054*, and *DS000210*). In order to ease the computational load on the continuous integration service, the three datasets were fixed for BIDS errors (in particular *DS000210*), reduced to a single-participant and heavily down-sampled, after downloading from OpenfMRI. The figure shows the current deployment workflow at CircleCI. Only after all tests (middle column) passed successfully, the "deploy" job pushes the Docker image to the Docker Hub repository and the Python package is uploaded to the Python Package Index (PyPI).
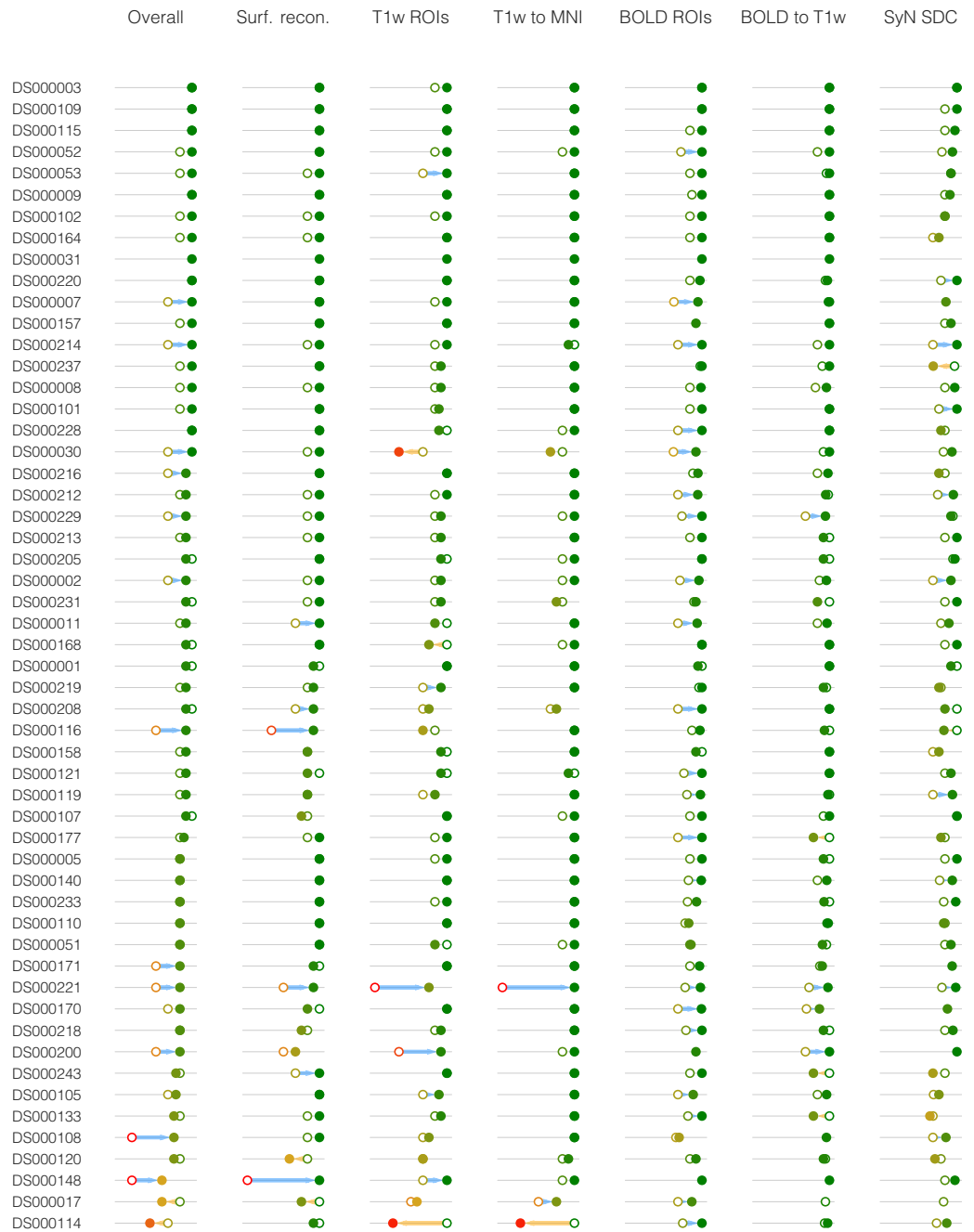
Figure S6. Visual assessment of all datasets under test, arranged by the mean overall rating. Each column of the diagram represents one out of the seven quality control points in the Phase II assessment: i) overall performance; ii) surface reconstruction from anatomical MRI; iii) T1w brain mask and tissue segmentation; iv) spatial normalization; v) brain mask and ROIs for CompCor application in native BOLD space ("BOLD ROIs"); vi) intra-subject BOLD-to-T1w co-registration; and vii) "fieldmap-less" SDC. Experts are instructed to assign a score on a scale from 1 (poor) to 4 (excellent) at each quality control point. A special rating score of 0 (unusable) could be assigned to cases where a particular step failed in a critical way that would completely hamper further processing.

Comparison to FSL `feat`

Data

We reuse the UCLA Consortium for Neuropsychiatric Phenomics LA5c Study[24], a dataset that is publicly available on OpenfMRI under data accession *DS000030*. During the experiment, subjects performed six tasks, a block of rest, and two anatomical scans. The study includes imaging data of a large group of healthy individuals from the community, as well as samples of individuals diagnosed with schizoprenia, bipolar disorder, and ADHD. A total of N=257 participants were available. For this experiment, we will only use the T1w, anatomical scans (for coregistration) and the functional scans from the Stop Signal (referred to as "stopsignal") task described below.

As described in their data descriptor[24], MRI data were acquired on one of two 3T Siemens Trio scanners, located at the Ahmanson-Lovelace Brain Mapping Center (syngo MR B15) and the Staglin Center for Cognitive Neuroscience (syngo MR B17). FMRI data were collected using an EPI sequence (slice thickness=4mm, 34 slices, TR=2s, TE=30ms, flip angle=90deg, matrix 64×64, FoV=192mm, oblique slice orientation). Additionally, a T1w image is available per participant (MPRAGE, TR=1.9s, TE=2.26ms, FoV=250mm, matrix=256×256, sagittal plane, slice thickness=1mm, 176 slices).

Stop signal task. During this stopsignal task, participants were instructed to respond quickly to a 'go' stimulus. During some of the trials, at unpredictable times, a stop signal would appear after the stimulus is presented. During those trials, the subject has to inhibit any planned response. In this experiment, we specifically look into the difference between the brain activation during a successful stop trial and a go trial (contrast: Go - StopSuccess). We expect to see brain regions responsible for response inhibition (negative) and motor response (positive). A detailed description of this particular task is available with the dataset descriptor[24].

Methods

Data were preprocessed with two alternate pipelines: *fMRIPrep* v1.0.8 and FSL's `feat` v5.0.10. We then performed identical analyses on each dataset preprocessed with either pipeline. The data are analysed on the first level: the results are obtained for each subject. We perform group level analyses in a specific resampling scheme to allow a statistical comparison between the pipelines: two random (non-overlapping) subsets of *n* subjects are repeatedly entered into a second level analysis. The first step is the experimental manipulation resulting in two conditions: (1) the data are preprcessed with *fMRIPrep*, and (2) the data are preprocessed using `feat`. The next two steps are identical for both conditions. We describe the details on the three analysis steps in more detail below.

Preprocessing. Preprocessing with *fMRIPrep* is described using the corresponding *citation boilerplate*, *Box S1*. We configured `feat` using its graphical user interface (GUI) and generated a `template.fsf` file, which can be found in GitHub [2]. We manually extended execution to all participants in our sample creating the script `fsl_feat_wrapper.py` that accompanies the `template.fsf` file in GitHub. As it can be seen on the `template.fsf` file, we disabled band-pass filtering and spatial smoothing to make results of preprocessing comparable. Both processing steps (temporal filtering and spatial smoothing) were implemented in a common, subsequent analysis workflow described below. Additionally, we manually configured the ICBM 152 Nonlinear Asymmetrical template version 2009c as target for spatial normalization. Finally, we manually resampled the preprocessed BOLD files into template space using FSL's `flirt`.

First-level analysis. We analyzed the task data using FSL and AFNI tools, integrated in a workflow using Nipype. Spatial smoothing was applied using AFNI's `3dBlurInMask` with a Gaussian kernel with FWHM=5mm. Activity was estimated using a general linear model (GLM) with FSL's `feat`. For the one condition under comparison (go - successful) one task regressor was included with a fixed duration of 1.5s and an extra regressor was added with equal amplitude, but the duration equal to the reaction time. Again, these regressors were orthogonalized with respect to the fixed duration regressor of the same

---

[2] https://github.com/oesteban/misc/tree/16660df9fe80d20107b6abd7fc8ce1f4946791e6/fsl-feat

Box S1. The citation boilerplate

Results included in this manuscript come from preprocessing performed using *fMRIPrep* version 1.0.8, a Nipype [25,26] based tool. Each T1w volume was corrected for intensity non-uniformity (INU) using `N4BiasFieldCorrection` v2.1.0 [14] and skull-stripped using `antsBrainExtraction.sh` v2.1.0 (using the OASIS template). Brain surfaces were reconstructed using recon-all from FreeSurfer v6.0.0 [5], 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 [20]. Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c [6] was performed through nonlinear registration with the antsRegistration tool of ANTs v2.1.0 [7], using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast [16] (FSL v5.0.9). Fieldmap-less" distortion correction was performed by co-registering the functional image to the same-subject T1w image with intensity inverted [12,13] constrained with an average fieldmap template [14], implemented with antsRegistration (ANTs). This was followed by co-registration to the corresponding T1w using boundary-based registration [15] with 9 degrees of freedom, using bbregister (FreeSurfer v6.0.0). Motion correcting transformations, field distortion correcting warp, BOLD-to-T1w transformation and T1w-to-template (MNI) warp were concatenated and applied in a single step using antsApplyTransforms (ANTs v2.1.0) using Lanczos interpolation.

Physiological noise regressors were extracted applying CompCor [17]. Principal components were estimated for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). A mask to exclude signal with cortical origin was obtained by eroding the brain mask, ensuring it only contained subcortical structures. Six tCompCor components were then calculated including only the top 5% variable voxels within that subcortical mask. For aCompCor, six components were calculated within the intersection of the subcortical mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run. Frame-wise displacement [18] was calculated for each functional run using the implementation of Nipype. Many internal operations of FMRIPREP use Nilearn [21], principally within the BOLD-processing workflow. For more details of the pipeline see https://fmriprep.readthedocs.io/en/latest/workflows.html.

condition. Predictors were convolved with a double-gamma canonical hemodynamic response function. Temporal derivatives were added to all task regressors to compensate for variability on the hemodynamic response function. Furthermore, the six rigid-motion parameters (translation in 3 directions, rotation in 3 directions) were added as regressors to avoid confounding effects of head-motion. We included a high-pass filter (100Hz) in FSL's `feat`.

The resampling scheme and group level analysis. Subsequent to the single subject analyses, two random (non-overlapping) subsamples of *n* subjects were taken and entered into a second level analysis. We vary the sample size *n* between 10 and 90 (there are in total of 257 subjects). This process is repeated 100 times. We analyzed the group data using ordinary least squares (OLS) mixed modeling using FSL's `flame`. Subsequently, we threshold the statistical maps, ensuring control of the False Discovery Rate (FDR) using FSL's `fdr` command.

Mapping the BOLD variability on standard space. To investigate the spatial consistency of the average BOLD across participants, we calculated standard deviation maps in MNI space[27] for the temporal average map derived from preprocessing with both alternatives.

Activation-count maps. The statistical map for each subject is binarized at $z=\pm1.65$. For each contrast, the average of these maps is computed over subjects. The average negative map (percentage of subjects showing a negative effect with $z < -1.65$) is subtracted from the average positive map to indicate the direction of effects. These maps show for each voxel the percentage of subjects that have values exceeding the threshold. High values in certain regions and low values in other regions show a good overlap of activation between subjects. A z-value of 1.65 corresponds to a two-sided test of $p<0.1$.

Smoothness. We used AFNI's `3dFWHMx` to estimate the (average) smoothness of the data at two checkpoints: i) before the first-level analysis workflow, and ii) after applying a 5.0mm full-width half-maximum (FWHM) spatial smoothing, which was the first step of the analysis workflow.

Extended results

***FMRIPrep* achieved higher spatial accuracy.** *Figure S7* shows the variability maps obtained with both preprocessing pipelines, and suggests that results from `feat` are smoother. These maps reveal a superior anatomical accuracy of *fMRIPrep* over `feat`, likely reflecting the combined effects of a more precise spatial normalization scheme, the single-shot interpolation that uses the more accurate sinc kernel (Lanczos interpolation), and the application of the "fieldmap-less" SDC methodology. *FMRIPrep* outcomes are particularly better aligned with the underlying anatomy in regions typically warped by susceptibility distortions such as the orbitofrontal lobe, as demonstrated by the close-ups on the lower panel of *Figure S7*.

First-level and group-level modeling did not show substantial differences. The results and figures associated to this experiment are fully reported using a Jupyter Notebook[3]

RELATED FMRI PREPROCESSING WORKFLOWS

The most comprehensive neuroimaging packages –AFNI, FSL and SPM in particular– include GUIs (see *Figure S8*) and workflows for fMRI preprocessing. AFNI (*Analysis of Functional NeuroImages*) offers a general purpose GUI and `afni_proc.py`, a single-subject, single-run preprocessing workflow for fMRI. The `afni_proc.py` pipeline is complemented by a GUI (`uber_subject.py`) and `align_epi_anat.py`, an image alignment tool (including several intra-subject registration schemes: fMRI-to-anatomical, anatomical-to-anatomical, fMRI-to-fMRI and anatomical-to-atlas). AFNI is completely open-source, and user feedback is utilized to improve the tool and users can contribute with bugfixes and features through https://github.com/afni/afni. FSL (*FMRIB's Software Library*) possesses the most extensive GUI, covering a number of MRI processing and analysis workflows, not limited to fMRI. One of the available tools is *FEAT*

---

[3]https://github.com/poldracklab/fmriprep-notebooks/blob/b3ada9425785c744884087937314e90610be1c22/02_Evaluation%20-%20FEAT%20Comparison.ipynb
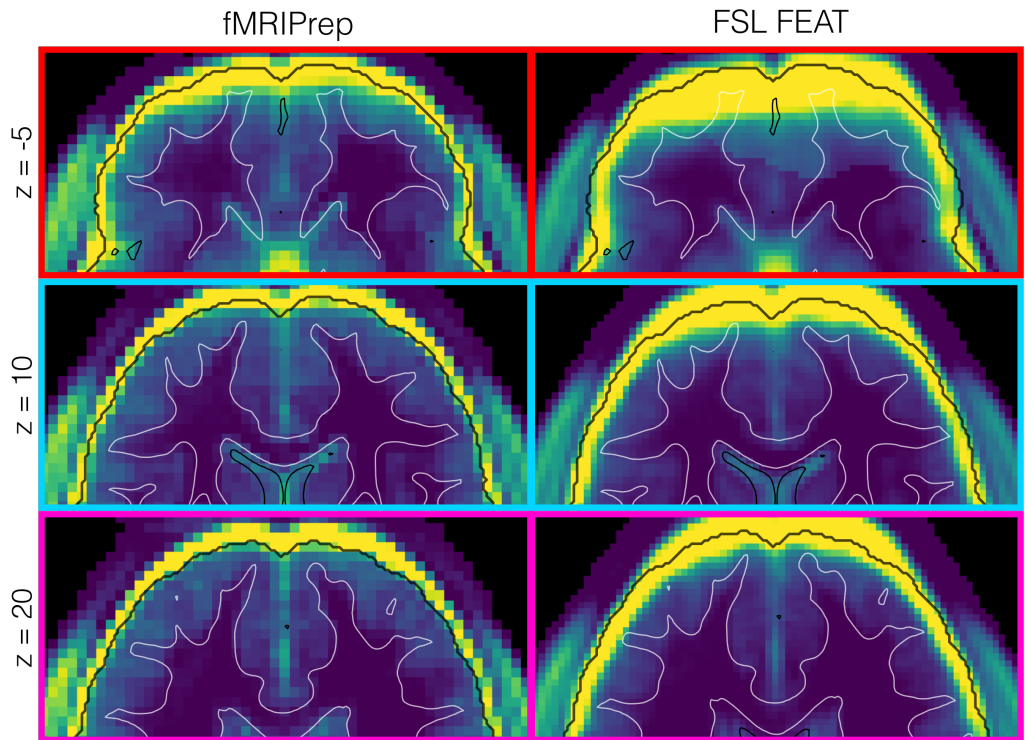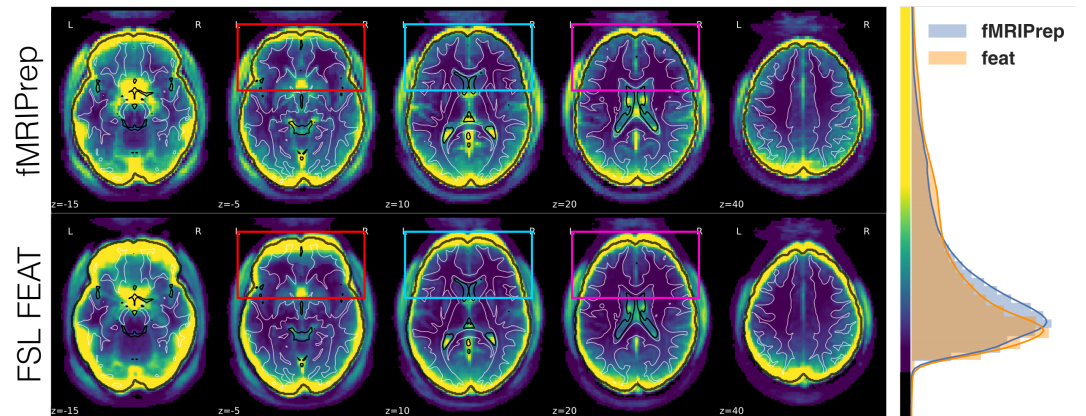
Figure S7. Maps of between-subjects variability of the averaged BOLD time-series resampled into MNI space. We preprocess DS000030 (N=257) with *fMRIPrep* and FSL `feat`. This figure shows greater between-subject variability of the averaged BOLD series obtained with `feat`, in MNI space. The top box of the panel shows these maps at different axial planes of the image grid, with reference contours from the MNI atlas. The map summarizing `feat`-derived results displays greater variability outside the brain mask delineated with the black contour. This effect is generally associated with a lower performance of spatial normalization[27]. The histogram at the right side plots the normalized frequency of variability (arbitrary units) for both maps, within the brain mask. The distribution corresponding to FSL `feat` shows a heavier tail. As demonstrated in the lower box of the panel, *fMRIPrep*-preprocessed images keep a higher spatial detail, especially on those regions affected by susceptibility-derived distortions.

FMRI Analysis, a GUI that creates a single-subject, single-run preprocessing and analysis workflow. Execution of the workflow can be delayed by writing one script file in the FSL internal format, generally referred to as `design.fsf`. Users can manually modify the file to analyze different subjects and runs. The user is responsible for manually executing other elements of the pipeline, such as anatomical preprocessing or susceptibility-derived distortion corrections (SDCs) other than fieldmap-based. SPM (*Statistical Parametric Mapping*) offers a GUI for the different processing and analysis tasks in the workflow and the users can combine them manually. Particular processing elements can be introduced using *plug-ins*, for instance to include a SDC step. Finally, SPM allows *plug-outs* to create workflows and batch scripts to run over several subjects. Both FSL and SPM are open-source, although they follow a more traditional development model less open to third-party contributions.

Interoperability between these three main software packages is low, since their internal processing and analysis elements are highly coupled. Although *Nipype*[25] provides a homogeneous interface that enables the decoupling these components (and effectively allows one to mix-and-match these tools), most researchers have been historically attached to one of the options. Their custom processing workflows are generally based off of one single package. In other words, their *ad hoc* pipelines are variations of `afni_proc.py`, FSL `feat` or SPM's batch scripts.

Human Connectome Project (HCP[28]) Pipelines constitute a different type of solution to this challenge. They propose a very comprehensive ensemble of workflows for preprocessing and analysis of "HCP-like" datasets. HCP Pipelines are built combining (mainly) FSL and FreeSurfer to perform fMRI surface analysis (among other anatomical and diffusion MRI analyses). The development of HCP Pipelines is open-source and collaborative through http://github.com.

Finally, C-PAC[29] (configurable pipeline for the analysis of connectomes) integrates AFNI and FSL tools through Nipype for preprocessing and analysis of resting-state fMRI connectomes. C-PAC offers a GUI that generates an extremely detailed configuration file where all the processing steps and all the necessary parameters are defined. Additional strengths of C-PAC are the open-source and collaborative development, and up-to-date and thorough documentation.

CURRENT USAGE OF *FMRIPREP*

FMRIPrep has been adopted by a large number of scientists, as suggested by the number of unique, worldwide visitors to the documentation website (see *Figure S9*) and more than ten thousand *pulls* of *fMRIPrep*'s Docker image[4]. On OpenNeuro.org, *fMRIPrep* has accumulated more than 240 run requests, accounting for the ~40% of all analyses requested so far in the platform. Some of these requests were executed on data uploaded by users – not found in OpenfMRI.

We have incorporated *fMRIPrep* in the processing workflow of all datasets in our laboratory. One example of the robust performance of *fMRIPrep* on idiosyncratic datasets is presented in *Figure S10*. *FMRIPrep* performed with high-accuracy on a challenging dataset with simultaneous electrocorticography (ECoG) recordings.
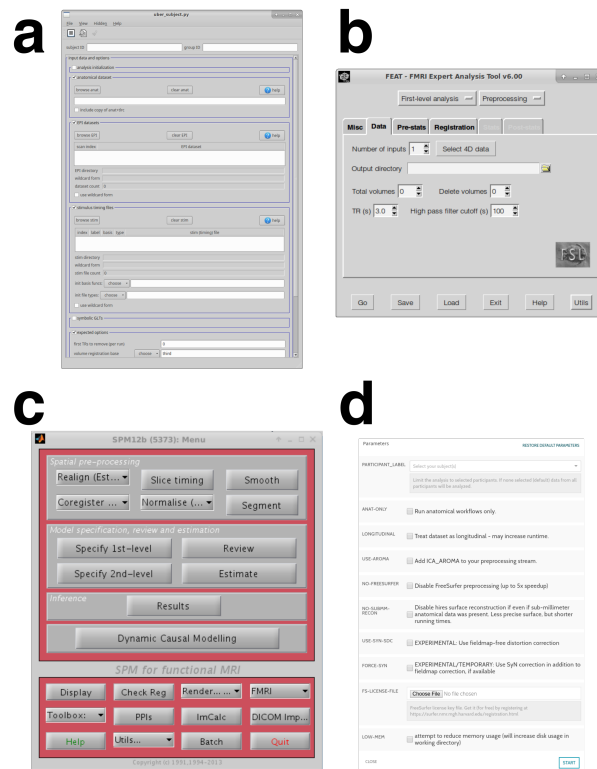
---

[4]https://hub.docker.com/r/poldracklab/fmriprep/

Figure S8. a | AFNI's uber_subject.py GUI. **b** | FSL's feat GUI. **c** | SPM's fMRI analysis GUI. **d** | *FMRIPrep*'s GUI at OpenNeuro.org.
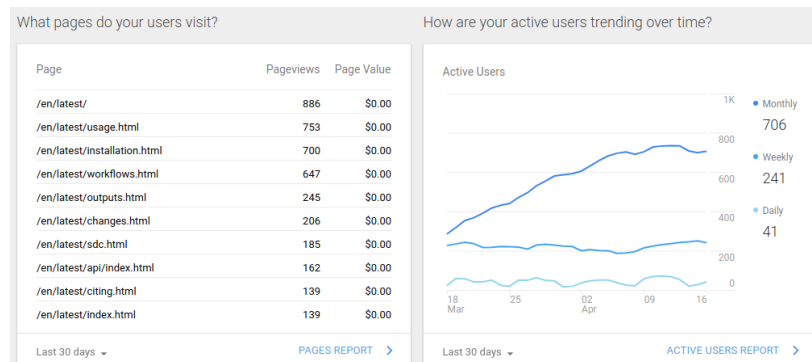

Figure S9. Unique visitors to the *fMRIPrep*'s documentation site. Clip from Google Analytics Panel showing tracking results for the documentation website http://fmriprep.org.
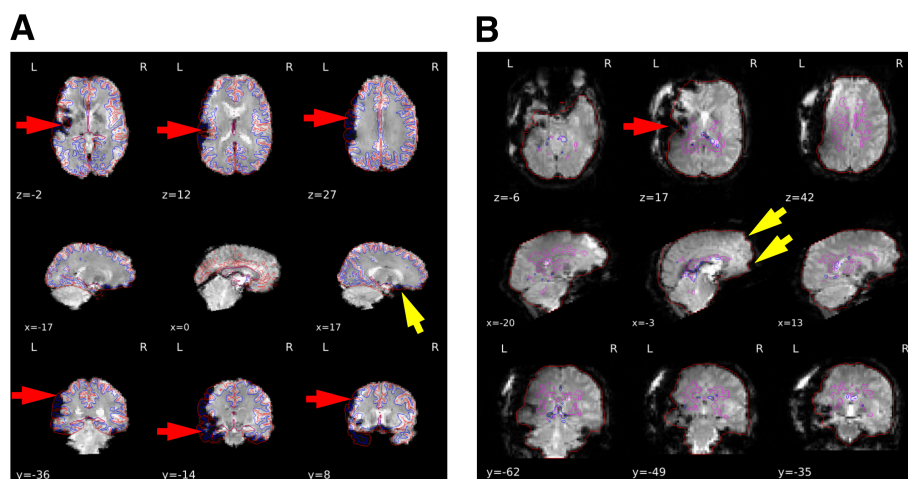
Figure S10. Successfully processing electrocorticography (ECoG) data. *FMRIPrep* successfully processed a some datasets acquired with simultaneous ECoG recordings. For ECoG, electrodes are placed directly on surfaces of the brain surgically exposed. As described previously, the exposed areas introduce large steps of magnetic susceptibility across the imaged volume, that lead to large distortions and signal drop-out (marked by red arrows). Yellow arrows point at distortions typically found in regular BOLD data. **A | Performance of co-registration**. Despite the wide areas of signal completely lost, boundary based registration performed accurately. **B | Delineation of regions-of-interest (ROIs)**. *FMRIPrep*'s heuristics and workflow for BOLD skull-stripping performed correctly (red contour). Moreover, regions where data are sampled to calculate *CompCor* [30] (blue and magenta contours) do not include any cortical regions.

REFERENCES

1 Jezzard, P. & Balaban, R. S. Correction for geometric distortion in echo planar images from B0 field variations. *Magnetic Resonance in Medicine* **34**, 65–73 (1995). doi:10.1002/mrm.1910340111.

2 Cordes, D., Arfanakis, K., Haughton, V. & Meyerand, M. Geometric distortion correction in EPI using two images with orthogonal phase-encoding directions. In *Proc. Intl. Soc. Mag. Reson. Med.*, vol. 8, 1712 (Denver, USA, 2000).

3 Reber, P. J., Wong, E. C., Buxton, R. B. & Frank, L. R. Correction of off resonance-related distortion in echo-planar imaging using EPI-based field maps. *Magnetic Resonance in Medicine* **39**, 328–330 (1998). doi:10.1002/mrm.1910390223.

4 Kybic, J., Thevenaz, P., Nirkko, A. & Unser, M. Unwarping of unidirectionally distorted EPI images. *IEEE Transactions on Medical Imaging* **19**, 80–93 (2000). doi:10.1109/42.836368.

5 Wang, S. *et al.* Evaluation of Field Map and Nonlinear Registration Methods for Correction of Susceptibility Artifacts in Diffusion MRI. *Frontiers in Neuroinformatics* **11** (2017). doi:10.3389/fninf.2017.00017.

6 Esteban, O. *et al.* Surface-driven registration method for the structure-informed segmentation of diffusion MR images. *NeuroImage* **139**, 450–461 (2016). doi:10.1016/j.neuroimage.2016.05.011.

7 Hutton, C. *et al.* Image Distortion Correction in fMRI: A Quantitative Evaluation. *NeuroImage* **16**, 217–240 (2002). doi:10.1006/nimg.2001.1054.

8 Sladky, R. *et al.* Slice-timing effects and their correction in functional MRI. *NeuroImage* **58**, 588–594 (2011). doi:10.1016/j.neuroimage.2011.06.078.

9 Roche, A. A Four-Dimensional Registration Algorithm With Application to Joint Correction of Motion and Slice Timing in fMRI. *IEEE Transactions on Medical Imaging* **30**, 1546–1554 (2011). doi:10.1109/TMI.2011.2131152.

10 Andrade, A. *et al.* Detection of fMRI activation using Cortical Surface Mapping. *Human Brain Mapping* **12**, 79–93 (2001). doi:10.1002/1097-0193(200102)12:2 extless79::AID-HBM1005 extgreater3.0.CO;2-I.

11 Laumann, T. O. *et al.* Functional System and Areal Organization of a Highly Sampled Individual Human Brain. *Neuron* **87**, 657–670 (2015). doi:10.1016/j.neuron.2015.06.037.

12 Strother, S. C. Evaluating fMRI preprocessing pipelines. *IEEE Engineering in Medicine and Biology Magazine* **25**, 27–41 (2006). doi:10.1109/MEMB.2006.1607667.

13 Caballero-Gaudes, C. & Reynolds, R. C. Methods for cleaning the BOLD fMRI signal. *NeuroImage* **154**, 128–149 (2017). doi:10.1016/j.neuroimage.2016.12.018.

14 Tustison, N. J. *et al.* N4ITK: Improved N3 Bias Correction. *IEEE Transactions on Medical Imaging* **29**, 1310–1320 (2010). doi:10.1109/TMI.2010.2046908.

15 Smith, S. M.   Fast robust automated brain extraction.   *Human Brain Mapping* **17**, 143–155 (2002). doi:10.1002/hbm.10062.

16 Abraham, A. *et al.* Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics* **8** (2014). doi:10.3389/fninf.2014.00014.

17 Cox, R. W. & Hyde, J. S. Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine* **10**, 171–178 (1997). doi:10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L.

18 Treiber, J. M. *et al.* Characterization and Correction of Geometric Distortions in 814 Diffusion Weighted Images. *PLOS ONE* **11**, e0152472 (2016). doi:10.1371/journal.pone.0152472.

19 Huntenburg, J. M. *Evaluating nonlinear coregistration of BOLD EPI and T1w images*. Master's thesis, Freie Universität, Berlin (2014). URL http://hdl.handle.net/11858/00-001M-0000-002B-1CB5-A.

20 Gorgolewski, K. J. *et al.* The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data* **3**, 160044 (2016). doi:10.1038/sdata.2016.44.

21 Halchenko, Y. *et al.* datalad/datalad 0.9.1. Tech. Rep., Zenodo (2017). URL https://zenodo.org/record/1000098#.Wnj4YHXwbmE.

22 Esteban, O. *et al.*   FMRIprep: a robust preprocessing pipeline for task-based and resting-state fMRI data.   In *Organization for Human Brain Mapping*, vol. 23, 1820 (Vancouver, Canada, 2017). doi:10.7490/f1000research.1114420.1.

23 Esteban, O. *et al.* MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PLOS ONE* **12**, e0184661 (2017). doi:10.1371/journal.pone.0184661.

24 Poldrack, R. A. *et al.* A phenome-wide examination of neural and cognitive function. *Scientific Data* **3**, 160110 (2016). doi:10.1038/sdata.2016.110.

25 Gorgolewski, K. *et al.* Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in Python. *Frontiers in Neuroinformatics* **5**, 13 (2011). doi:10.3389/fninf.2011.00013.

26 Gorgolewski, K. J. *et al.* Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in Python. *Zenodo [Software]* (2017). doi:10.5281/zenodo.581704.

27 Calhoun, V. D. *et al.* The impact of T1 versus EPI spatial normalization templates for fMRI data analyses. *Human Brain Mapping* **38**, 5331–5342 (2017). doi:10.1002/hbm.23737.

28 Van Essen, D. *et al.* The Human Connectome Project: A data acquisition perspective. *NeuroImage* **62**, 2222–2231 (2012). doi:10.1016/j.neuroimage.2012.02.018.

29 Sikka, S. *et al.*   Towards automated analysis of connectomes: The configurable pipeline for the analysis of connectomes (C-PAC).   In *5th INCF Congress of Neuroinformatics*, vol. 117 (Munich, Germany, 2014). doi:10.3389/conf.fninf.2014.08.00117.

30 Behzadi, Y., Restom, K., Liau, J. & Liu, T. T. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage* **37**, 90–101 (2007). doi:10.1016/j.neuroimage.2007.04.042.