

FMRIPrep: a robust preprocessing pipeline for functional MRI

Oscar Esteban^{1*}, Christopher J. Markiewicz¹, Ross W. Blair¹, Craig A. Moodie², A. Ilkay Isik³, Asier Erramuzpe⁴, James D. Kent⁵, Mathias Goncalves⁶, Elizabeth DuPre⁷, Madeleine Snyder⁸, Hiroyuki Oya⁹, Satrajit S. Ghosh^{6,10}, Jessey Wright¹, Joke Durnez¹, Russell A. Poldrack^{1‡}, Krzysztof J. Gorgolewski^{1‡*}

*For correspondence:

phd@oscaresteban.es (OE);
krzysztof.gorgolewski@gmail.com (KG)

‡Contributed equally to this work

¹Department of Psychology, Stanford University, California, USA; ²Medical School Center, Stanford University, California, USA; ³Max Planck Institute for Empirical Aesthetics, Hesse, Germany; ⁴Computational Neuroimaging Lab, Biocruces Health Research Institute, Bilbao, Spain; ⁵Neuroscience Program, University of Iowa, USA;

⁶McGovern Institute for Brain Research, Massachusetts Institute of Technology: MIT, Cambridge, MA, USA;

⁷Montreal Neurological Institute, McGill University; ⁸Department of Psychiatry, Stanford Medical School, Stanford University, California, USA; ⁹Department of Neurosurgery, University of Iowa Health Care, Iowa City, Iowa;

¹⁰Department of Otolaryngology, Harvard Medical School, Boston, MA, USA

1 Preprocessing of functional MRI (fMRI) involves numerous steps to clean and standardize data
2 before statistical analysis. Generally, researchers create *ad hoc* preprocessing workflows for
3 each new dataset, building upon a large inventory of tools available for each step. The
4 complexity of these workflows has snowballed with rapid advances in MR data acquisition and
5 image processing techniques. We introduce *fMRIPrep*, an analysis-agnostic tool that
6 addresses the challenge of robust and reproducible preprocessing for task-based and resting
7 fMRI data. *fMRIPrep* automatically adapts a best-in-breed workflow to the idiosyncrasies of
8 virtually any dataset, ensuring high-quality preprocessing with no manual intervention. By
9 introducing visual assessment checkpoints into an iterative integration framework for
10 software-testing, we show that *fMRIPrep* robustly produces high-quality results on a diverse
11 fMRI data collection comprising participants from 54 different studies in the OpenfMRI
12 repository. We review the distinctive features of *fMRIPrep* in a qualitative comparison to other
13 preprocessing workflows. We demonstrate that *fMRIPrep* achieves higher spatial accuracy as
14 it introduces less uncontrolled spatial smoothness than one commonly used preprocessing
15 tool. *fMRIPrep* has the potential to transform fMRI research by equipping neuroscientists with
16 a high-quality, robust, easy-to-use and transparent preprocessing workflow which can help
17 ensure the validity of inference and the interpretability of their results.

18 Functional magnetic resonance imaging (fMRI) is a commonly used technique to map human brain
19 activity¹. However, the blood-oxygen-level dependent (BOLD) signal measured by fMRI is typically
20 mixed with many non-neural sources of variability². Preprocessing identifies the nuisance sources and
21 reduces their effect on the data³. Other major preprocessing steps⁴ deal with particular imaging arti-
22 facts and the anatomical location of signals. For instance, slice-timing⁵ correction (STC), head-motion
23 correction (HMC), and susceptibility distortion correction (SDC) address particular artifacts; while co-
24 registration, and spatial normalization are concerned with signal location (see Online Methods, sec.
25 Preprocessing of fMRI in a nutshell, for a summary). Extracting a signal that is most faithful to the
26 underlying neural activity is crucial to ensure the validity of inference and interpretability of results⁶.
27 Faulty preprocessing may lead to the interpretation of noise patterns as signals of interest. For example,
28 Power et al. demonstrated that unaccounted-for head-motion can generate spurious and systematic cor-
29 relations in resting-state fMRI⁷, which would be interpreted as functional connectivity. An illustration
30 of failed spatial normalization familiar to most researchers is finding significant activation outside of the

31 brain. Other preprocessing choices may result in the removal of signal originating from brain activity.
32 The ongoing debate on the need for regressing out global signals^{2,8,9} reflects just such concerns. Thus,
33 a primary goal of preprocessing is to reduce sources of Type I errors without inducing excessive Type II
34 errors.

35 Workflows for preprocessing fMRI produce two broad classes of outputs: *preprocessed* data (as op-
36 posed to *raw*, original data) and measurements of experimental *confounds* for use in later modeling.
37 Preprocessed data generally include new fMRI time-series after the application of retrospective signal
38 correction and filtering algorithms. In addition, these data are typically resampled onto a target space
39 appropriate for analysis, such as a standardized anatomical reference. The *confounds* are additional
40 time-series such as physiological recordings and estimated noise sources that are useful for analysis (e.g.
41 they can be applied as nuisance regressors). Some commonly used confounds include: motion param-
42 eters, framewise displacement (FD⁷), spatial standard deviation of the data after temporal differencing
43 (DVARS⁷), global signals, etc. Preprocessing may include further steps for denoising and estimation
44 of confounds. For instance, dimensionality reduction methods based on principal components analysis
45 (PCA) or independent components analysis (ICA), such as component-based noise correction (*Comp-*
46 *Cor*¹⁰) or automatic removal of motion artifacts (ICA-AROMA¹¹).

47 The neuroimaging community is well equipped with tools that implement the majority of the individ-
48 ual steps of preprocessing described so far. These tools are readily available within software packages
49 including AFNI¹², ANTs¹³, FreeSurfer¹⁴, FSL¹⁵, Nilearn¹⁶, or SPM¹⁷. Despite the wealth of accessible
50 software and multiple attempts to outline best practices for preprocessing^{2,4,6,18}, the large variety of data
51 acquisition protocols have led to the use of *ad hoc* pipelines customized for nearly every study; for exam-
52 ple, Carp¹⁹ found 223 unique analysis workflows across 241 fMRI studies. Thus, current preprocessing
53 workflows offer a poor trade-off between the quality of results and robust, consistent performance on
54 datasets other than those that they were built for. Alternatively, researchers can adopt the acquisition
55 protocols defined by large neuroimaging consortia like the Human Connectome Project (HCP²⁰) or the
56 UK Biobank²¹, which then allows the use of their preprocessing pipelines^{22,23} developed for those stud-
57 ies. Since these pipelines are optimized for particular data acquisition protocols, they are not applicable
58 to datasets acquired using different protocols. In practice, the neuroimaging community lacks a prepro-
59 cessing workflow that reliably provides high-quality and consistent results on arbitrary datasets.

60 Here we introduce *fMRIprep*, a preprocessing workflow for task-based and resting-state fMRI. *fMRI-*
61 *Prep* is built around four driving principles: 1) **robustness** to the idiosyncrasies of the input dataset; 2)
62 **quality** of preprocessing outcomes; 3) **transparency** to encourage the scrutiny of preprocessing results
63 for quality, and to facilitate accurate communication of the methods; and 4) **ease-of-use** with the min-
64 imization of manual intervention. *fMRIprep* is robust by virtue of a flexible, self-adapting architecture
65 that combines tools from existing neuroimaging analysis packages. Tools for each processing operation
66 are selected through an evidence-driven and community-informed optimization process. Here we also
67 report a comprehensive evaluation of the workflow on a large and heterogeneous subsample of the
68 OpenfMRI repository, to quantify robustness and quality of the results. This evaluation leverages the
69 comprehensive visual reports generated by *fMRIprep*, which facilitate assessment and curation of the re-
70 sults. These reports exemplify the “glass-box” philosophy with which the software was developed; rather
71 than hiding a complex set of operations within a monolithic black box, *fMRIprep* exposes interim results
72 at multiple steps to encourage active engagement by the scientist.

RESULTS

73 *fMRIprep* is a robust and convenient tool for researchers and clinicians to prepare both task-based
74 and resting-state fMRI for analysis. Its outputs enable a broad range of applications, including within-
75 subject analysis using functional localizers, voxel-based analysis, surface-based analysis, task-based
76 group analysis, resting-state connectivity analysis, and many others. In the following, we describe the
77 overall architecture, software engineering principles, and a comprehensive validation of the tool.

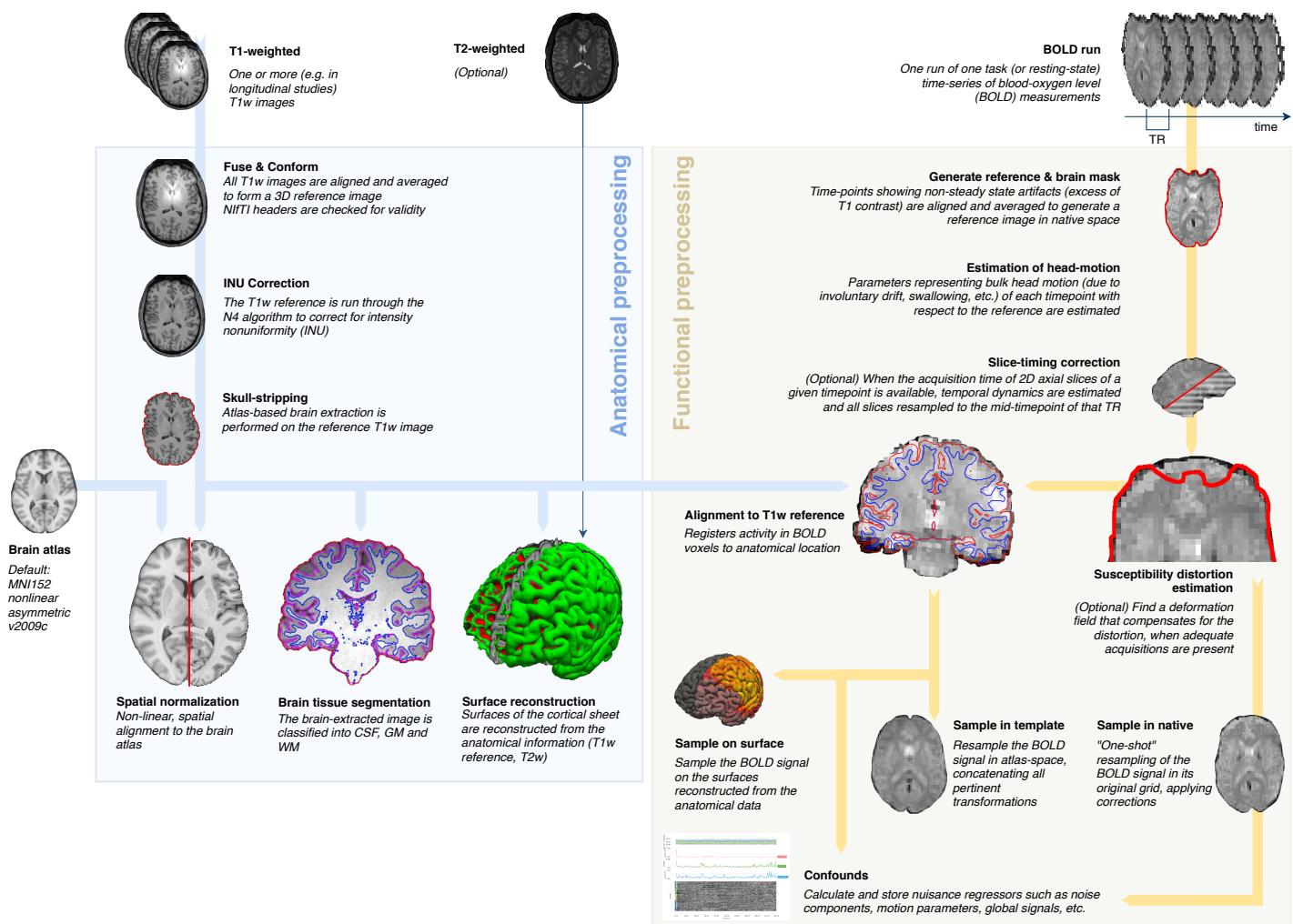


Figure 1. fMRIPrep is a fMRI preprocessing tool that adapts to the input dataset. Leveraging the Brain Imaging Data Structure (BIDS²⁴), the software self-adjusts automatically, configuring the optimal workflow for the given input dataset. Thus, no manual intervention is required to locate the required inputs (one T1-weighted image and one BOLD series), read acquisition parameters (such as the repetition time –TR– and the slice acquisition-times) or find additional acquisitions intended for specific preprocessing steps (like field maps and other alternatives for the estimation of the susceptibility distortion). Outputs are easy to navigate due to compliance with the BIDS Extension Proposal for derived data (see Online Methods, Figure S4).

78 **A modular design allows for a flexible, adaptive workflow**

79 The foundation of *fMRIPrep* is presented in *Figure 1*. The workflow is composed by sub-workflows
80 that are dynamically assembled into different configurations depending on the input data. These building
81 blocks combine tools from widely-used, open-source neuroimaging packages (see *Table 1* for a summary). *Nipype*²⁵ is used to stage the workflows and to deal with execution details (such as resource
82 management). As presented in *Figure 1*, the workflow comprises two major blocks, separated into
83 anatomical and functional MRI processing streams.

85 **Automatically understanding the input dataset.** The Brain Imaging Data Structure (BIDS²⁴) allows
86 *fMRIPrep* to precisely identify the structure of the input data and gather all the available metadata (e.g.
87 imaging parameters). *fMRIPrep* reliably adapts to dataset irregularities such as missing acquisitions or
88 runs through a set of heuristics. For instance, if only one participant of a sample lacks field-mapping
89 acquisitions, *fMRIPrep* will by-pass the correction step for that one participant.

Table 1. State-of-art neuroimaging offers a large catalog of readily available software tools. *fMRIPrep* integrates best-in-breed tools for each of the preprocessing tasks that its workflow covers.

Preprocessing task	<i>fMRIPrep</i> includes	Alternatives (not included within <i>fMRIPrep</i>)
Anatomical T1w brain-extraction	<code>antsBrainExtraction.sh</code> (ANTs)	<code>bet</code> (FSL), <code>3dSkullstrip</code> (AFNI), <code>MRTOOL</code> (SPM Plug-in)
Anatomical surface reconstruction	<code>recon-all</code> (FreeSurfer)	CIVET, BrainSuite, Computational Anatomy (SPM Plug-in)
Head-motion estimation (and correction)	<code>mcflirt</code> (FSL)	<code>3dvolreg</code> (AFNI), <code>spm_realign</code> (SPM), <code>cross_realign_4dfp</code> (4dfp), <code>antsBrainRegistration</code> (ANTs)
Susceptibility-derived distortion estimation (and unwarping)	<code>3dqwarp</code> (AFNI)	<code>fugue</code> and <code>topup</code> (FSL), FieldMap and HySCO (SPM Plug-ins)
Slice-timing correction	<code>3dTshift</code> (AFNI)	<code>slicetimer</code> (FSL), <code>spm_slice_timing</code> (SPM), <code>interp_4dfp</code> (4dfp)
Intra-subject registration	<code>bbregister</code> (FreeSurfer), <code>flirt</code> (FSL)	<code>3dvolreg</code> (AFNI), <code>antsRegistration</code> (ANTs), <code>Coregister</code> (SPM GUI)
Spatial normalization (inter-subject co-registration)	<code>antsRegistration</code> (ANTs)	<code>@auto_tlrc</code> (AFNI), <code>fnirt</code> (FSL), <code>Normalize</code> (SPM GUI)
Surface sampling	<code>mri_vol2surf</code> (FreeSurfer)	MNE, Nilearn
Subspace selection methods	<code>melodic</code> (FSL), ICA-AROMA	Nilearn, LMGS (SPM Plug-in)
Confounds	<i>in-house</i> implementation	TAPAS PhysIO (SPM Plug-in)
Steady-state detection	<i>in-house</i> implementation	<i>Ad hoc</i> implementations

90 Preprocessing anatomical images. The T1-weighted (T1w) image is corrected for intensity non-
91 uniformity (INU) using `N4BiasFieldCorrection`²⁶ (ANTs), and skull-stripped using `antsBrainExtrac-`
92 `tion.sh` (ANTs). Skull-stripping is performed through coregistration to a template, with two options
93 available: the OASIS template²⁷ (default) or the NKI template²⁸. Using visual inspection, we have found
94 that this approach outperforms other common approaches, which is consistent with previous reports²².
95 When several T1w volumes are found, the INU-corrected versions are first fused into a reference T1w
96 map of the subject with `mri_robust_template`²⁹ (FreeSurfer). Brain surfaces are reconstructed from
97 the subject's T1w reference (and T2-weighted images if available) using `recon-all`³⁰ (FreeSurfer). The
98 brain mask estimated previously is refined with a custom variation of a method (originally introduced in
99 `Mindboggle`³¹) to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray matter
100 (GM). Both surface reconstruction and subsequent mask refinement are optional and can be disabled
101 to save run time when surface-based analysis is not needed. Spatial normalization to the ICBM 152
102 Nonlinear Asymmetrical template³² (version 2009c) is performed through nonlinear registration with
103 `antsRegistration`³³ (ANTs), using brain-extracted versions of both the T1w reference and the standard
104 template. ANTs was selected due to its superior performance in terms of volumetric group level over-
105 lap³⁴. Brain tissues –cerebrospinal fluid (CSF), white matter (WM) and GM– are segmented from the
106 reference, brain-extracted T1w using `fast`³⁵ (FSL).

107 Preprocessing functional runs. For every BOLD run found in the dataset, a reference volume and its
108 skull-stripped version are generated using an *in-house* methodology (reported in Online Methods, sec.
109 Particular processing elements of *fMRIPrep*). Then, head-motion parameters (volume-to-reference trans-
110 form matrices, and corresponding rotation and translation parameters) are estimated using `mcflirt`³⁶
111 (FSL). Among several alternatives (see *Table 1*), `mcflirt` is used because its results are comparable
112 to other tools³⁷ and it stores the estimated parameters in a format that facilitates the composition of
113 spatial transforms to achieve one-step interpolation (see below). If slice timing information is available,
114 BOLD runs are (optionally) slice time corrected using `3dTshift` (AFNI¹²). When field map information is

available, or the experimental “fieldmap-less” correction is requested (see Highlights of *fMRIPrep* within the neuroimaging context), SDC is performed using the appropriate methods (see Online Methods, *Figure S3*). This is followed by co-registration to the corresponding T1w reference using boundary-based registration³⁸ with nine degrees of freedom (to minimize remaining distortions). If surface reconstruction is selected, *fMRIPrep* uses *bbregister* (FreeSurfer). Otherwise, the boundary based coregistration implemented in *flirt* (FSL) is applied. In our experience, *bbregister* yields the better results³⁸ due to the high resolution and the topological correctness of the GM/WM surfaces driving registration. To support a large variety of output spaces for the results (e.g. the native space of BOLD runs, the corresponding T1w, FreeSurfer’s *fsaverage* spaces, the atlas used as target in the spatial normalization step, etc.), the transformations between spaces can be combined. For example, to generate preprocessed BOLD runs in template space (e.g. MNI), the following transforms are concatenated: head-motion parameters, the warping to reverse susceptibility-distortions (if calculated), BOLD-to-T1w, and T1w-to-template mappings. The BOLD signal is also sampled onto the corresponding participant’s surfaces using *mri_volsurf* (FreeSurfer), when surface reconstruction is being performed. Thus, these sampled surfaces can easily be transformed onto different output spaces available by concatenating transforms calculated throughout *fMRIPrep* and internal mappings between spaces calculated with *recon-all*. The composition of transforms allows for a single-interpolation resampling of volumes using *antsApplyTransforms* (ANTs). Lanczos interpolation is applied to minimize the smoothing effects of linear or Gaussian kernels³⁹. Optionally, ICA-AROMA can be performed and corresponding “non-aggressively” denoised runs are then produced.

Extraction of nuisance time-series. *FMRIPrep* is analysis-agnostic and thus, it does not perform any temporal denoising. Nonetheless, it provides researchers with a diverse set of confound estimates that could be used for explicit nuisance regression or as part of higher-level models. This lends itself to decoupling preprocessing and behavioral modeling as well as evaluating robustness of final results across different denoising schemes. A set of physiological noise regressors are extracted for the purpose of performing component-based noise correction (*CompCor*¹⁰). Principal components are estimated after high-pass filtering the BOLD time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (*tCompCor*) and anatomical (*aCompCor*). Six *tCompCor* components are then calculated from the top 5% variable voxels within a mask covering the subcortical regions. Such subcortical mask is obtained by heavily eroding the brain mask, which ensures it does not include cortical GM regions. For *aCompCor*, six components are calculated within the intersection of the aforementioned mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run (using the inverse BOLD-to-T1w transformation). Frame-wise displacement⁴⁰ is calculated for each functional run, using the implementation in Nipype. DVARS are also calculated using Nipype. Three global signals are extracted within the CSF, the WM, and the whole-brain masks using Nilearn¹⁶. If ICA-AROMA¹¹ is requested, the “aggressive” noise-regressors are collected and placed within the corresponding confounds files. In addition, a “non-aggressive” version of preprocessed data is also provided since this variant of ICA-AROMA denoising cannot be performed using only nuisance regressors.

154 Visual reports ease quality control and maximize transparency

155 Users can assess the quality of preprocessing with an individual report generated per participant.
156 *Figure 2* shows an example of such reports and describes their structure. Reports contain dynamic
157 and static mosaic views of images at different quality control points along the preprocessing pipeline.
158 Many visual elements of the reports, as well as some of the figures in this manuscript are generated
159 using Nilearn¹⁶. Only a web browser is required to open the reports on any platform, since they are
160 written in hypertext markup language (HTML). HTML also enables the trivial integration within online
161 neuroimaging services such as OpenNeuro.org, and maximizes shareability between peers. These reports
162 effectively minimize the amount of time required for assessing the quality of the results. They also help
163 understand the internals of processing by visually reporting the full provenance of data throughout the

164 workflow. As an additional transparency enhancement, reports are accompanied by a *citation boilerplate*
165 (see Online Methods, *Box S1*) that follows the guidelines for reporting fMRI studies by Poldrack et
166 al.⁴¹. Meant for its inclusion within the methodological section of papers using *fMRIPrep*, the boilerplate
167 provides a literate description of the processing that includes software versions of all tools involved in
168 the particular workflow and gives due credit to all authors of all of the individual pieces of software used
169 within *fMRIPrep*.

170 **Highlights of *fMRIPrep* within the neuroimaging context**

171 *fMRIPrep* is not the first preprocessing pipeline for fMRI data. The most widely used neuroimaging
172 packages generally provide workflows, such as afni_proc.py (AFNI) or feat (FSL). Other alternatives
173 include C-PAC⁴² (configurable pipeline for the analysis of connectomes), HCP Pipelines or the Batch
174 Editor of SPM. In this section, we highlight some additional features beyond robustness and quality that
175 will likely incline scientists to find in *fMRIPrep* the best fit for their fMRI preprocessing needs.

176 **Analysis-agnostic:** *fMRIPrep* is meant to support all kinds of analysis. To some extent, all alternative
177 workflows limit the possible analyses that can be performed on the preprocessed data. These limitations
178 mostly derive from the coordinates space of the outputs and the regular (volume) vs. irregular (surface)
179 sampling of the BOLD signal. For example, HCP Pipelines supports surface-based analyses on subject
180 or template space. Conversely, afni_proc.py, C-PAC and feat are volume-based only. *fMRIPrep* allows
181 a multiplicity of output spaces including subject-space and atlases for both volume-based and surface-
182 based analyses. While *fMRIPrep* avoids including processing steps that may limit further analysis (e.g.
183 spatial smoothing), other tools are designed to perform preprocessing that supports specific analysis
184 pipelines. For instance, C-PAC performs several processing steps towards the connectivity analysis of
185 resting-state fMRI.

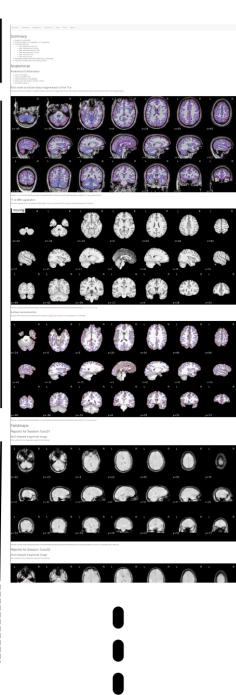
186 **Susceptibility distortion correction (SDC) in the absence of field maps.** Many legacy and current
187 human fMRI protocols lack the MR field maps necessary to perform standard methods for SDC. *fMRIPrep*
188 adapts the “fieldmap-less” correction method for diffusion echo-planar imaging (EPI) images introduced
189 by Wang et al.⁴³. They propose using the same-subject T1w reference as the *undistorted* target in a
190 nonlinear registration scheme. To maximize the similarity between the T2* contrast of the EPI scan
191 and the reference T1w, the intensities of the latter are inverted. To regularize the optimization of the
192 deformation field only displacements along the phase-encoding (PE) direction are allowed, and the
193 magnitude of the displacements is modulated using priors. To our knowledge, no other existing pipeline
194 implements “fieldmap-less” SDC to the BOLD images.

195 ***fMRIPrep* is thoroughly documented, community-driven, and developed with high-standards of**
196 **software engineering.** Preprocessing pipelines are generally well documented, however the extreme
197 flexibility of *fMRIPrep* makes its proper documentation substantially more challenging. As for other large
198 scientific software communities, *fMRIPrep* contributors pledge to keep the documentation thorough and
199 updated along coding iterations. Packages also differ on the involvement of the community: while *fMRI-Prep*
200 includes researchers in the decision making process and invites their suggestions and contributions,
201 other packages have a more closed model where the feedback from users is more limited (e.g. a mailing
202 list). In contrast to other pipelines, *fMRIPrep* is community-driven. This paradigm allows the fast adop-
203 tion of cutting-edge advances on fMRI preprocessing. For example, while *fMRIPrep* initially performed
204 STC before HMC, we adapted the tool to the recent recommendations of Power et al.¹⁸ upon a user’s
205 request*. This model has allowed the user base to grow rapidly and enabled substantial third-party con-
206 tributions to be included in the software, such as the support for processing datasets without anatomical
207 information. The open-source nature of *fMRIPrep* has permitted frequent code reviews that are effective
208 in enhancing the software’s quality and reliability⁴⁴. Finally, *fMRIPrep* undergoes continuous integration
209 testing (see Online Methods, *Figure S5*), a technique that has recently been proposed as a mean to
210 ensure reproducibility of analyses in computational sciences^{45,46}.

*<https://neurostars.org/t/obtaining-movement-estimates-before-slice-time-correction/1007>

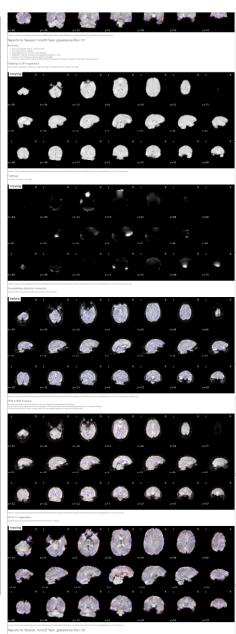
Summary

Reports start with an overview of the dataset, as identified using BIDS.



Anatomical processing

Several panels allow for quality control of the anatomical workflow. Brain tissue segmentation, spatial normalization and surface reconstruction (if requested) can be inspected using these visualization panels.



Fieldmaps processing

When the dataset contains any of the supported alternatives to estimate the deformation map corresponding to susceptibility distortions, these panels help assess these images were correctly processed.

Each BOLD run across the different tasks and sessions will be presented at different quality control points.

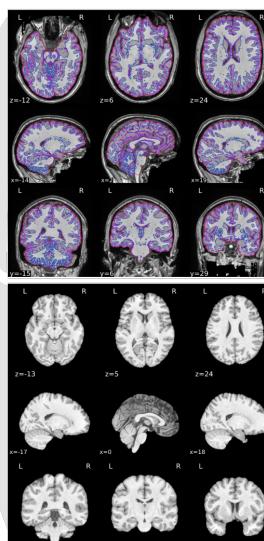
First, when fieldmaps were found, some mosaics will show the alignment of those maps to the BOLD reference. The block ends with a dynamic plot showing how images are unwarped.

The report also shows processing in native BOLD space plotting the brain mask calculated from the functional MR signal and the regions-of-interest (ROIs) where the CompCor confounds are calculated.

Finally, the alignment between same-subject T1-weighted and that specific BOLD run is presented.

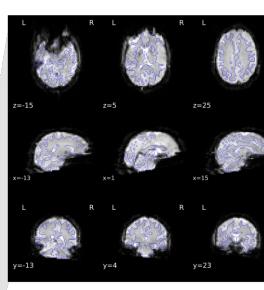
Errors

fMRIprep is explicit about errors, and any problems encountered along the processing will be listed at the end of the report, with collapsible panels containing the specific detail of each error.

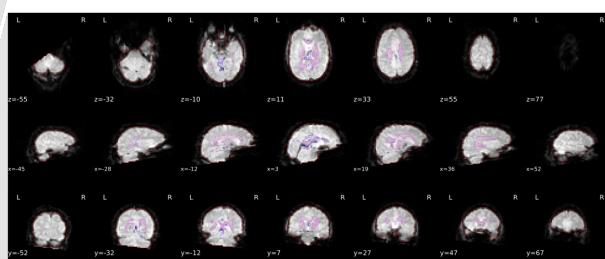


T1-weighted reference, brain mask, intensity inhomogeneity and brain tissue segmentation panel. A static mosaic allows the assessment of these four crucial steps of pre-processing anatomical images.

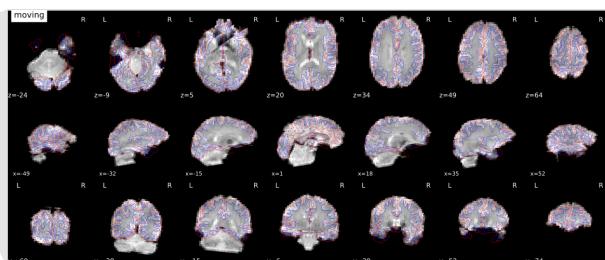
Spatial normalization. A dynamic mosaic that transitions between the target atlas space and the T1w-reference aligned into that space allows checking the accuracy of this image registration process.



Susceptibility distortion correction. If fieldmap information was found or the "fieldmap-less" correction is requested, the step is assessed with a dynamic mosaic that transitions between the unwarped ("after") and original ("before"). Contours of the white-matter are also presented as anatomical cue.



BOLD mask and CompCor ROIs. The final BOLD signal is presented, with contours representing the outline of the brain mask, and two regions-of-interest (ROIs) where CompCor confounds are estimated.



Alignment of BOLD and the T1w reference. The correct alignment to the anatomical reference is assessed with a dynamic mosaic that renders the reconstructed surfaces over the BOLD reference.

Figure 2. Anatomy of the visual reports generated by fMRIprep. The visual reports ease quality control of the results and help understand the preprocessing flow.

211 Ensuring reproducibility with hard versioning and containers. For enhanced reproducibility, *fMRI-Prep*
212 fully supports execution via the Docker (<https://docker.com>) and Singularity⁴⁷ container platforms.
213 Container images are generated and uploaded to a public repository for each new version of *fMRIPrep*.
214 This helps address the widespread lack of reporting of specific software versions and the large variability
215 of software versions, which threaten the reproducibility of fMRI analyses¹⁹. These containers are
216 released with a fixed set of software versions for *fMRIPrep* and all its dependencies, maximizing run-to-run
217 reproducibility in an easy way. Except for C-PAC, alternative pipelines do not provide official
218 support for containers. The adoption of the BIDS-Apps⁴⁵ container model makes *fMRIPrep* amenable to
219 a multiplicity of infrastructures and platforms: PC, high-performance computing (HPC), Cloud, etc.

220 ***fMRIPrep* yields high-quality results on a diverse set of input data**

221 *Figure 3* presents the validation framework that we applied to iteratively maximize the robustness
222 of the tool and validate the quality of the results. The validation framework implements a testing plan
223 elaborated prior the release of the version 1.0 of the software (see Online Methods, sec. Evaluation of
224 *fMRIPrep*). The plan is divided in two validation phases in which different data samples and validation
225 procedures are applied. *Table 2* describes the data samples used on each phase and emphasizes how
226 these data are collected from a large number of different, unrelated studies. In Phase I, we ran *fMRIPrep*
227 on a manually selected sample of participants that are potentially challenging to the tool's robustness,
228 exercising the adaptiveness to the input data. Phase II focused on the visual assessment of the quality of
229 preprocessing results on a large and heterogeneous sample.

230 Validation Phase I – Fault-discovery testing. We tested *fMRIPrep* on a set of 30 datasets from [OpenfMRI](#)
231 (see *Table 2*). Included participants were manually selected for their low quality as visually assessed by
232 two experts using MRIQC¹⁰⁵ (the assessment protocol is further described in in Online Methods, sec.
233 Evaluation of *fMRIPrep*). Data showing substandard quality are known to likely degrade the outcomes
234 of image processing¹⁰⁵, and therefore they are helpful to test software reliability. Phase I concluded with
235 the release of *fMRIPrep* version 1.0 on December 6, 2017.

236 Validation Phase II – Quality assurance and reliability testing. We extended the evaluation data
237 up to 54 datasets from OpenfMRI (see *Table 2*). Participants were selected randomly as described in
238 Online Methods, sec. Evaluation of *fMRIPrep*. Validation Phase II integrated a protocol for the screening
239 of results into the software testing (*Figure 3*). As shown in *Figure 4*, this effectively contributed to
240 substantive improvements on the quality of results. Three raters (authors CJM, KJG and OE) evaluated
241 the 213 visual reports at six quality control points throughout the pipeline, and also assigned an overall
242 score to each participant. Their ratings are made available with the corresponding reports for scrutiny.
243 The scoring scale has three levels: 1 (“poor”), 2 (“acceptable”) and 3 (“excellent”). A special rating of 0
244 (“unusable”) is assigned to critical failures that hamper any further processing beyond the quality control
245 checkpoint. After Phase II, 50 datasets out of the total 54 were rated above the “acceptable” average
246 quality level. The remaining 4 datasets were all above the “poor” level and in or nearby the “acceptable”
247 rating. *Figure 4* illustrates the quality of results, while Online Methods, *Figure S6* shows the individual
248 evolution of every dataset at each of the seven quality control points. Phase II concluded with the release
249 of *fMRIPrep* version 1.0.8 on February 22, 2018.

250 ***fMRIPrep* improves spatial precision through reduced smoothing**

251 We investigate whether the focus on robustness against data irregularity comes at a cost in quality
252 of the preprocessing outcomes by comparing it to the commonly used FSL *feat* workflow. Using all
253 the scans of the “stopsignal” task in DS000030 (N=257 participants) from [OpenfMRI](#), we ran *fMRIPrep*
254 and a standard *feat* workflow. We chose *feat* because DS000030 had successfully been preprocessed
255 and analyzed with FSL tools previously⁵⁵. Smoothing is intentionally excluded from both preprocessing
256 routes with the aim to apply it early within a common (identical) analysis workflow. We calculated
257 standard deviation maps in MNI space¹⁰⁶ for the temporal average map of the “stopsignal” task derived
258 from preprocessing with both alternatives. Visual inspection of these variability maps (*Figure 5*) reveals

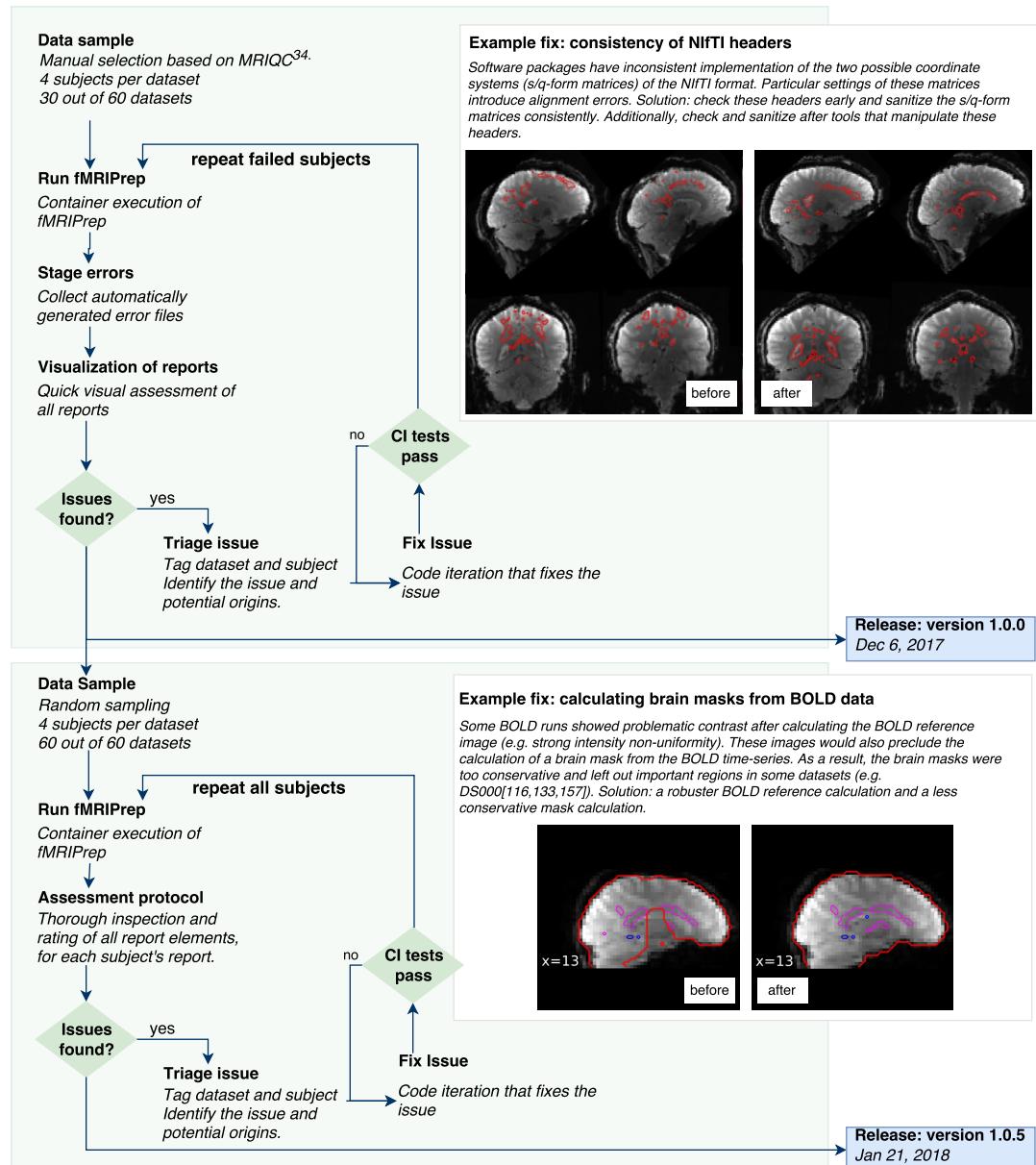


Figure 3. Combining visual assessment within the software testing flow. We complement well-established techniques for software integration testing with manual assessment of the outputs. The evaluation framework is designed with two subsequent testing phases. Phase I focuses on fault-discovery and visual reports are used to better understand the issues found. The top box (Example fix 1) shows an example of defect identified and solved during this testing cycle. After addressing a total of 21 issues affecting 7 datasets, and the release of fMRIprep version 1.0.0, the next testing stage is initiated. Phase II focuses on increasing the overall quality of results as evaluated visually by experts. Following an inspection protocol, reports from 213 participants belonging to 58 different studies were individually assessed. We found 12 additional issues affecting 11 datasets that have been addressed with the release of fMRIprep version 1.0.3 on January 3, 2018. The bottom box (Example fix 2) illustrates one of these issues, which produced errors in the brain extraction process from BOLD data.

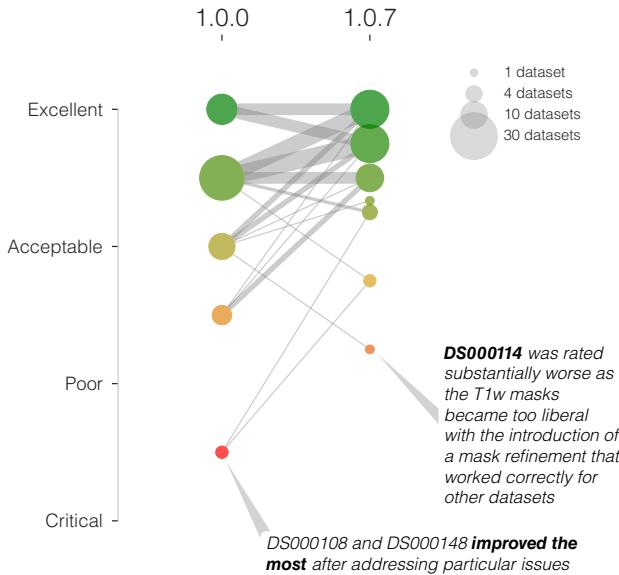


Figure 4. Integrating visual assessment into the software testing framework effectively increases the quality of results. In an early assessment of quality using *fMRIprep* version 1.0.0, the overall rating of two datasets was below the “poor” category and four below the “acceptable” level (left column of colored circles). After addressing some outstanding issues detected by the early assessment, the overall quality of processing is substantially improved (right column of circles), and no datasets are below the “poor” quality level. Only two datasets are rated below the “acceptable” level in the second assessment (using *fMRIprep* version 1.0.7).

259 a higher anatomical accuracy of *fMRIprep* over *feat*, likely reflecting the combined effects of a more
260 precise spatial normalization scheme and the application of “fieldmap-less” SDC. *fMRIprep* outcomes
261 are particularly better aligned with the underlying anatomy in regions typically warped by susceptibility
262 distortions such as the orbitofrontal lobe, as demonstrated by close-ups in Online Methods, *Figure S7*.

263
264 We also compared preprocessing done with *fMRIprep* and FSL’s *feat* in two common fMRI analyses.
265 First, we performed within subject statistical analysis using *feat* –the same tool provides preprocessing
266 and first-level analysis– on both sets of preprocessed data. Second, we perform a group statistical analy-
267 sis using ordinary least squares (OLS) mixed modeling (*flame*¹⁰⁷, FSL). In both experiments, we applied
268 identical analysis workflows and settings to both preprocessing alternatives. Using AFNI’s 3dFWHMx, we
269 estimated the smoothness of data right after preprocessing (unsmoothed), and after an initial smooth-
270 ing step of 5.0mm (full-width half-minimum, FWHM) of the common analysis workflow. As visually
271 suggested by *Figure 5*, we indeed found that *feat* produces smoother data (*Figure 6A*). Although pre-
272 processed data were resampled to an isotropic voxel size of 2.0×2.0×2.0 [mm], the smoothness estima-
273 tion (before the prescribed smoothing step) for *fMRIprep* was below 4.0mm, very close to the original
274 resolution of 3.0×3.0×4.0 [mm] of these data. The first-level analysis showed that the thresholded ac-
275 tivation count maps for the go vs. successful stop contrast in the “stopsignal” task were very similar
276 (*Figure 6B*). It can be seen that the results from both pipelines identified activation in the same regions.
277 However, since data preprocessed with *feat* are smoother, the results from *fMRIprep* are more local and
278 better aligned with the cortical sheet.

279 To investigate the implications of either pipeline on the group analysis use-case, we run the same
280 OLS modeling on two disjoint subsets of randomly selected subjects. We calculate several metrics of
281 spatial agreement on the resulting maps of (uncorrected) *p*-statistical values, and also after binarizing
282 these maps with a threshold chosen to control for the false discovery rate at 5%. The overlap of statistical
283 maps, as well as Pearson’s correlation, were tightly related to the smoothing of the input data. In Online
284 Methods, sec. Comparison to FSL *feat* we report the group-level analysis in full. We ran two variants of

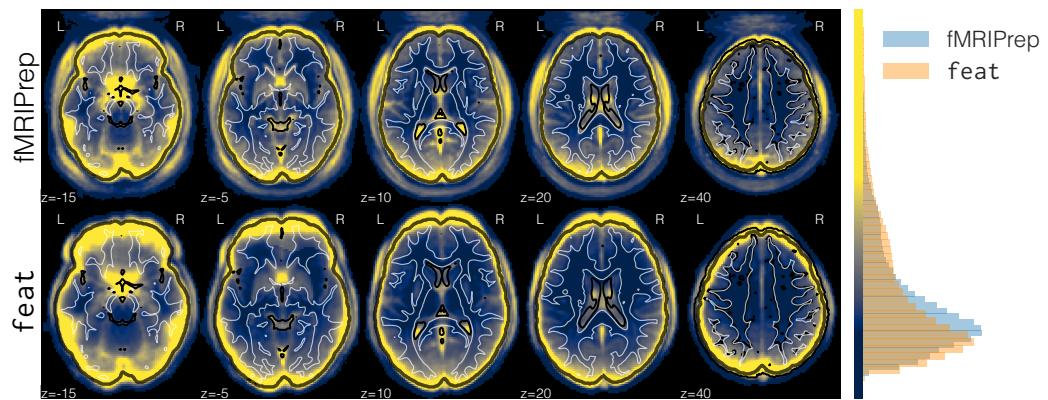


Figure 5. Maps of between-subjects variability of the averaged BOLD time-series resampled into MNI space. We preprocessed DS000030 ($N=257$) with fMRIprep and FSL feat. This figure shows greater between-subject variability of the averaged BOLD series obtained with feat, in MNI space. The top box of the 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 normalization¹⁰⁶. 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. See Online Methods, *Figure S7* for close-ups into regions affected by susceptibility-derived distortions.

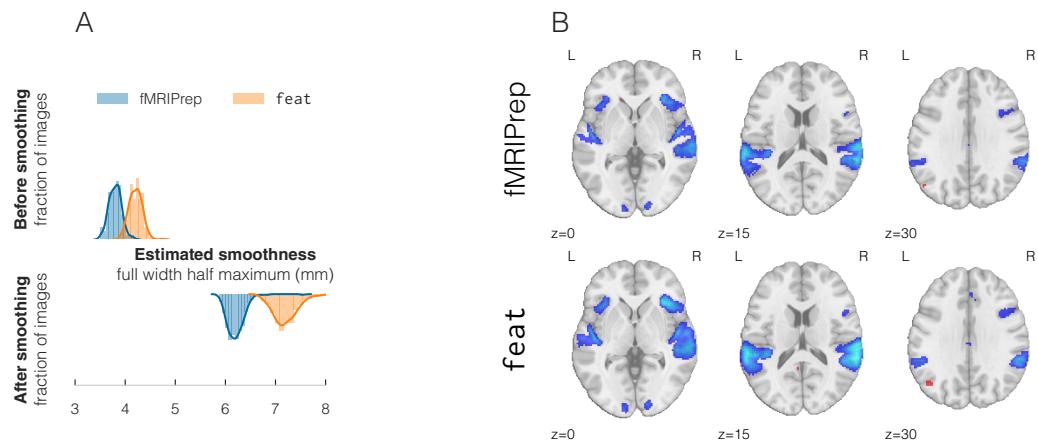


Figure 6. A | Estimating the spatial smoothness of data before and after the initial smoothing step of the analysis workflow confirmed that results of preprocessing with feat are intrinsically smoother. Therefore, fMRIprep allows the researcher for a finer control over the smoothness of their analysis. **B |** Thresholded activation count maps for the go vs. successful stop contrast in the “stopsignal” task after preprocessing using either fMRIprep or FSL’s feat, with identical single subject statistical modeling. Both tools obtained similar activation maps, with fMRIprep results being slightly better aligned with the underlying anatomy.

285 the analysis: with a prescribed smoothing of 5.0mm FWHM, and without smoothing step. These results
286 showed that, at the group-level analysis, *fMRIPrep* and *feat* perform equivalently.

DISCUSSION

287 *fMRIPrep* is a fMRI preprocessing workflow developed to excel at four aspects of scientific software:
288 *robustness* to data idiosyncrasies, high *quality* and consistency of results, maximal *transparency* in the
289 assessment of results and subsequent communication, and *ease-of-use*. We describe how using the Brain
290 Imaging Data Structure (BIDS²⁴) along with a flexible design allows the workflow to self-adapt to the
291 idiosyncrasy of inputs (sec. A modular design allows for a flexible, adaptive workflow). The workflow
292 (briefly summarized in *Figure 1*) integrates state-of-art tools from widely used neuroimaging software
293 packages at each preprocessing step (see *Table 1*). Some other relevant facets of *fMRIPrep* and how
294 they relate to existing alternative pipelines are presented in sec. Highlights of *fMRIPrep* within the neu-
295 roimaging context. To note some, the analysis-agnostic nature of the tool, or the uniqueness of the
296 “fieldmap-less” SDC method. We highlight that *fMRIPrep* is developed with the best software engineer-
297 ing principles, which are fundamental to ensure software reliability. The pipeline is easy to use for
298 researchers and clinicians without extensive computer engineering experience, and produces compre-
299 hensive visual reports (*Figure 2*). These automated reports exemplify the “glass-box” principle, which
300 requires that software allows scientists to understand how it works internally. This is in contrast to
301 typical “black-box” applications that perform valuable services without providing a way to understand
302 how the tool has transformed their data into the desired output. These reports maximize transparency
303 by allowing scientists to critically inspect and better understand the underlying mechanisms of their
304 preprocessing.

305 We demonstrate the robustness of *fMRIPrep* on a data collection from datasets associated with differ-
306 ent studies (*Table 2*), representing the variety of input data in the field (sec. *fMRIPrep* yields high-quality
307 results on a diverse set of input data). We then interrogate the quality of those results with the individual
308 inspection of the corresponding visual reports by experts (sec. Visual reports ease quality control and
309 maximize transparency and the corresponding summary in *Figure 4*). A comparison to FSL’s *feat* (sec.
310 *fMRIPrep* improves spatial precision through reduced smoothing) demonstrates that *fMRIPrep* achieves
311 higher spatial accuracy and introduces less uncontrolled smoothness (*Figures 5, 6*). Group *p*-statistical
312 maps only differed on their smoothness (sharper for the case of *fMRIPrep*). The fact that first-level and
313 second-level analyses resulted in small differences between *fMRIPrep* and our *ad hoc* implementation of
314 a *feat*-based workflow indicates that the individual preprocessing steps perform similarly when they are
315 fine-tuned to the input data. That justifies the need for *fMRIPrep*, which autonomously adapts the work-
316 flow to the data without error-prone manual intervention. To a limited extent, that also mitigates some
317 concerns and theoretical risks arisen from the analytical degrees-of-freedom¹⁹ available to researchers.
318 *fMRIPrep* stands out amongst pipelines because it automates the adaptation to the input dataset without
319 compromising the quality of results.

320 One limitation of this work is the use of visual (the reports) and semi-visual (e.g. *Figure 5* and
321 *Figure 6*) assessments for the quality of preprocessing outcomes. Although some frameworks have been
322 proposed for the quantitative evaluation of preprocessing on task-based (such as NPAIRS¹⁰⁸) and resting-
323 state¹⁰⁹ fMRI, they impose a set of assumptions on the test data and the workflow being assessed that
324 severely limit their suitability. The modular design of *fMRIPrep* defines an interface to each processing
325 step, which will permit the programmatic evaluation of the many possible combinations of software
326 tools and processing steps. That will also enable the use of quantitative testing frameworks to pursue
327 the minimization of Type I errors without the cost of increasing Type II errors.

328 The range of possible applications for *fMRIPrep* also presents some boundaries. For instance, very
329 narrow field-of-view (FoV) images oftentimes do not contain enough information for standard image
330 registration methods to work correctly. Reduced FoV datasets from *OpenfMRI* were excluded from the
331 evaluation since they are not yet fully supported by *fMRIPrep*. Extending *fMRIPrep*’s support for these
332 particular images is already a future line of the development road-map. *fMRIPrep* may also under-
333 perform for particular populations (e.g. infants) or when brains show nonstandard structures, such as

334 tumors, resected regions or lesions. Nonetheless, *fMRIPrep*'s architecture makes it straightforward to
335 extend the tool to support specific populations or new species by providing appropriate atlases of those
336 brains. This future line of work would be particularly interesting in order to adapt the workflow to
337 data collected from rodents and nonhuman primates. By contrast, *fMRIPrep* performed robustly on data
338 from a simultaneous MRI/electrocorticography (ECoG) study, which is extremely challenging to analyze
339 due to the massive BOLD signal drop-out near the implanted cortical electrodes (see Online Methods,
340 *Figure S10*).

341 Approximately 80% of the analysis pipelines investigated by Carp¹⁹ were implemented using either
342 AFNI¹², FSL¹⁵, or SPM¹⁷. *Ad hoc* pipelines adapt the basic workflows provided by these tools to the
343 particular dataset at hand. Although workflow frameworks like Nipype¹¹⁰ ease the integration of tools
344 from different packages, these pipelines are typically restricted to just one of these alternatives (AFNI,
345 FSL or SPM). Otherwise, scientists can adopt the acquisition protocols and associated preprocessing
346 software of large consortia like the Human Connectome Project (HCP) or the UK Biobank. This option
347 allows scientists to shortcut the intricacies of preprocessing by applying a “black-box” that has been
348 validated on similar data by a third party. The *off-the-shelf* applicability of these workflows is contravened
349 by important limitations on the experimental design. Therefore, researchers typically opt to recode
350 their custom preprocessing workflows with nearly every new study¹⁹. That practice entails a “pipeline
351 debt”, which requires the investment on proper software engineering to ensure an acceptable correctness
352 and stability of the results (e.g. continuous integration testing) and reproducibility (e.g. versioning,
353 packaging, containerization, etc.). A trivial example of this risk would be the leakage of *magic numbers*
354 that are hard-coded in the source (e.g. a crucial imaging parameter that inadvertently changed from one
355 study to the next one). Until *fMRIPrep*, an analysis-agnostic approach that builds upon existing software
356 instruments and optimizes preprocessing for robustness to data idiosyncrasies, quality of outcomes, ease-
357 of-use, and transparency, was lacking.

358 The rapid increase in volume and diversity of available data, as well as the evolution of more so-
359 phisticated techniques for processing and analysis, presents an opportunity for significantly advancing
360 research in neuroscience. However, the influx of new data, new analysis methods, and new modeling
361 strategies represents a risk as well as an opportunity. The inferential promises of big data, and the
362 sophisticated analysis tools that can leverage it, incentivize researchers to progressively build on more
363 complex analysis pipelines that rely on more complex and more obscure models of the data to pro-
364 duce interpretable results. This way of moving forward risks producing a future generation of cognitive
365 neuroscientists who have become experts in using sophisticated computational methods, but have little
366 to no working knowledge of the biological processes underlying brain's function¹¹¹. It also obscures
367 important steps in the inductive process mediating between experimental measurements and reported
368 findings. Easy-to-use, *off-the-shelf* tools that function as black boxes –providing scientists with limited
369 insight into how the tool functions, and developed primarily behind closed doors– may only exacerbate
370 this problem. *fMRIPrep* offers a novel “glass-box” approach for the development, maintenance and use
371 of computational tools that mitigates these risks. By standardizing preprocessing, *fMRIPrep* allows re-
372 searchers to focus their attention and expertise on the inferentially significant stages of data production,
373 analysis and interpretation. Additionally, *fMRIPrep* mitigates concerns about black-box processing by
374 being thoroughly documented, producing reports and visualizations at critical quality control points in
375 the workflow, and being developed according to the best practices of open source engineering. These
376 features of *fMRIPrep* make it possible for researchers to learn how the tool works, develop an under-
377 standing of each step in the workflow, and even reconstruct the preprocessing pipeline from first principles.
378 *fMRIPrep* aims to better equip fMRI practitioners to perform reliable, reproducible, statistical analyses
379 with a high-standard, consistent, and adaptive preprocessing instrument.

CONCLUSION

380 Despite efforts to achieve high-quality preprocessing of idiosyncratic fMRI datasets, doing so reliably
381 has remained an open problem. *fMRIPrep* is an analysis-agnostic, preprocessing workflow that yields
382 consistent results across a wide range of input datasets. *fMRIPrep* is built on top of the best neuroimaging

383 tools selected from various software packages. These tools are integrated into workflows that can be
384 dynamically combined to compose a full preprocessing workflow adapted to the input data. The optimal
385 workflow for the input dataset is constructed at runtime, blending a set of heuristics with the Brain
386 Imaging Data Structure (BIDS) to read the inputs. *fMRIprep* excels in four design goals: robustness,
387 high-quality of results, transparency and ease-of-use. To validate and demonstrate these features, we
388 integrate the individual screening of preprocessing results with continuous integration techniques of
389 software testing. The process is aided by comprehensive, portable reports that inform the scientist about
390 the workflow, ease the quality control of results and maximize the shareability of research outcomes.
391 We highlight the aspects that justify the development of *fMRIprep* with respect to currently available
392 preprocessing workflows. We quantitatively demonstrate that *fMRIprep* does not introduce uncontrolled
393 smoothing as compared to one alternative software. *fMRIprep* aims to better equip fMRI practitioners
394 to perform reliable, reproducible statistical analyses with a high-standard, transparent, and verifiable
395 instrument.

ACKNOWLEDGMENTS

396 This work was supported by the Laura and John Arnold Foundation, NIH R01 EB020740, NIH
397 1R24MH114705-01, and NINDS grant 1U01NS103780-01. JD has received funding from the Euro-
398 pean Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant
399 agreement No 706561.

REFERENCES

- 1 Poldrack, R. A. & Farah, M. J. Progress and challenges in probing the human brain. *Nature* **526**, 371–379 (2015). doi:[10.1038/nature15692](https://doi.org/10.1038/nature15692).
- 2 Power, J. D., Plitt, M., Laumann, T. O. & Martin, A. Sources and implications of whole-brain fMRI signals in humans. *NeuroImage* **146**, 609–625 (2017). doi:[10.1016/j.neuroimage.2016.09.038](https://doi.org/10.1016/j.neuroimage.2016.09.038).
- 3 Lindquist, M. A. The Statistical Analysis of fMRI Data. *Statistical Science* **23**, 439–464 (2008). doi:[10.1214/09-STS282](https://doi.org/10.1214/09-STS282).
- 4 Strother, S. C. Evaluating fMRI preprocessing pipelines. *IEEE Engineering in Medicine and Biology Magazine* **25**, 27–41 (2006). doi:[10.1109/MEMB.2006.1607667](https://doi.org/10.1109/MEMB.2006.1607667).
- 5 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](https://doi.org/10.1016/j.neuroimage.2011.06.078).
- 6 Ashburner, J. Preparing fMRI Data for Statistical Analysis. In Filippi, M. (ed.) *fMRI Techniques and Protocols*, no. 41 in Neuromethods, 151–178 (Humana Press, 2009). doi:[10.1007/978-1-60327-919-2_6](https://doi.org/10.1007/978-1-60327-919-2_6).
- 7 Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. & Petersen, S. E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage* **59**, 2142–2154 (2012). doi:[10.1016/j.neuroimage.2011.10.018](https://doi.org/10.1016/j.neuroimage.2011.10.018).
- 8 Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B. & Bandettini, P. A. The impact of global signal regression on resting state correlations: Are anti-correlated networks introduced? *NeuroImage* **44**, 893–905 (2009). doi:[10.1016/j.neuroimage.2008.09.036](https://doi.org/10.1016/j.neuroimage.2008.09.036).
- 9 Liu, T. T., Nalci, A. & Falahpour, M. The global signal in fMRI: Nuisance or Information? *NeuroImage* **150**, 213–229 (2017). doi:[10.1016/j.neuroimage.2017.02.036](https://doi.org/10.1016/j.neuroimage.2017.02.036).
- 10 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](https://doi.org/10.1016/j.neuroimage.2007.04.042).
- 11 Pruim, R. H. R. *et al.* ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *NeuroImage* **112**, 267–277 (2015). doi:[10.1016/j.neuroimage.2015.02.064](https://doi.org/10.1016/j.neuroimage.2015.02.064).
- 12 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](https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L).
- 13 Avants, B. B. *et al.* A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage* **54**, 2033–44 (2011). doi:[10.1016/j.neuroimage.2010.09.025](https://doi.org/10.1016/j.neuroimage.2010.09.025).

- 14 Fischl, B. FreeSurfer. *NeuroImage* **62**, 774–781 (2012). doi:[10.1016/j.neuroimage.2012.01.021](https://doi.org/10.1016/j.neuroimage.2012.01.021).
- 15 Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W. & Smith, S. M. FSL. *NeuroImage* **62**, 782–790 (2012). doi:[10.1016/j.neuroimage.2011.09.015](https://doi.org/10.1016/j.neuroimage.2011.09.015).
- 16 Abraham, A. *et al.* Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics* **8** (2014). doi:[10.3389/fninf.2014.00014](https://doi.org/10.3389/fninf.2014.00014).
- 17 Friston, K. J., Ashburner, J., Kiebel, S. J., Nichols, T. E. & Penny, W. D. *Statistical parametric mapping : the analysis of functional brain images* (Academic Press, London, 2006).
- 18 Power, J. D., Plitt, M., Kundu, P., Bandettini, P. A. & Martin, A. Temporal interpolation alters motion in fMRI scans: Magnitudes and consequences for artifact detection. *PLOS ONE* **12**, e0182939 (2017). doi:[10.1371/journal.pone.0182939](https://doi.org/10.1371/journal.pone.0182939).
- 19 Carp, J. The secret lives of experiments: Methods reporting in the fMRI literature. *NeuroImage* **63**, 289–300 (2012). doi:[10.1016/j.neuroimage.2012.07.004](https://doi.org/10.1016/j.neuroimage.2012.07.004).
- 20 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](https://doi.org/10.1016/j.neuroimage.2012.02.018).
- 21 Miller, K. L. *et al.* Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nature Neuroscience* **19**, 1523–1536 (2016). doi:[10.1038/nn.4393](https://doi.org/10.1038/nn.4393).
- 22 Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80**, 105–124 (2013). doi:[10.1016/j.neuroimage.2013.04.127](https://doi.org/10.1016/j.neuroimage.2013.04.127).
- 23 Alfaro-Almagro, F. *et al.* Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *NeuroImage* (2017). doi:[10.1016/j.neuroimage.2017.10.034](https://doi.org/10.1016/j.neuroimage.2017.10.034).
- 24 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](https://doi.org/10.1038/sdata.2016.44).
- 25 Gorgolewski, K. J. *et al.* Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in Python. *Zenodo [Software]* (2016). doi:[10.5281/zenodo.50186](https://doi.org/10.5281/zenodo.50186).
- 26 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](https://doi.org/10.1109/TMI.2010.2046908).
- 27 Marcus, D. S. *et al.* Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults. *Journal of Cognitive Neuroscience* **19**, 1498–1507 (2007). doi:[10.1162/jocn.2007.19.9.1498](https://doi.org/10.1162/jocn.2007.19.9.1498).
- 28 Noonan, K. B. *et al.* The NCI-Rockland Sample: A Model for Accelerating the Pace of Discovery Science in Psychiatry. *Frontiers in Neuroscience* **6** (2012). doi:[10.3389/fnins.2012.00152](https://doi.org/10.3389/fnins.2012.00152).
- 29 Reuter, M., Rosas, H. D. & Fischl, B. Highly accurate inverse consistent registration: A robust approach. *NeuroImage* **53**, 1181–1196 (2010). doi:[10.1016/j.neuroimage.2010.07.020](https://doi.org/10.1016/j.neuroimage.2010.07.020).
- 30 Dale, A. M., Fischl, B. & Sereno, M. I. Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage* **9**, 179–194 (1999). doi:[10.1006/nimg.1998.0395](https://doi.org/10.1006/nimg.1998.0395).
- 31 Klein, A. *et al.* Mindboggling morphometry of human brains. *PLOS Computational Biology* **13**, e1005350 (2017). doi:[10.1371/journal.pcbi.1005350](https://doi.org/10.1371/journal.pcbi.1005350).
- 32 Fonov, V., Evans, A., McKinstry, R., Almlie, C. & Collins, D. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* **47**, Supplement 1, S102 (2009). doi:[10.1016/S1053-8119\(09\)70884-5](https://doi.org/10.1016/S1053-8119(09)70884-5).
- 33 Avants, B., Epstein, C., Grossman, M. & Gee, J. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis* **12**, 26–41 (2008). doi:[10.1016/j.media.2007.06.004](https://doi.org/10.1016/j.media.2007.06.004).
- 34 Klein, A. *et al.* Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage* **46**, 786–802 (2009). doi:[10.1016/j.neuroimage.2008.12.037](https://doi.org/10.1016/j.neuroimage.2008.12.037).

- 35 Zhang, Y., Brady, M. & Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging* **20**, 45–57 (2001). doi:[10.1109/42.906424](https://doi.org/10.1109/42.906424).
- 36 Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage* **17**, 825–841 (2002). doi:[10.1006/nimg.2002.1132](https://doi.org/10.1006/nimg.2002.1132).
- 37 Oakes, T. R. *et al.* Comparison of fMRI motion correction software tools. *NeuroImage* **28**, 529–543 (2005). doi:[10.1016/j.neuroimage.2005.05.058](https://doi.org/10.1016/j.neuroimage.2005.05.058).
- 38 Greve, D. N. & Fischl, B. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage* **48**, 63–72 (2009). doi:[10.1016/j.neuroimage.2009.06.060](https://doi.org/10.1016/j.neuroimage.2009.06.060).
- 39 Lanczos, C. Evaluation of Noisy Data. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis* **1**, 76–85 (1964). doi:[10.1137/0701007](https://doi.org/10.1137/0701007).
- 40 Power, J. D. *et al.* Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage* **84**, 320–341 (2014). doi:[10.1016/j.neuroimage.2013.08.048](https://doi.org/10.1016/j.neuroimage.2013.08.048).
- 41 Poldrack, R. A. *et al.* Guidelines for reporting an fMRI study. *NeuroImage* **40**, 409–414 (2008). doi:[10.1016/j.neuroimage.2007.11.048](https://doi.org/10.1016/j.neuroimage.2007.11.048).
- 42 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](https://doi.org/10.3389/conf.fninf.2014.08.00117).
- 43 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](https://doi.org/10.3389/fninf.2017.00017).
- 44 McIntosh, S., Kamei, Y., Adams, B. & Hassan, A. E. The Impact of Code Review Coverage and Code Review Participation on Software Quality: A Case Study of the Qt, VTK, and ITK Projects. In *Proceedings of the 11th Working Conference on Mining Software Repositories*, MSR 2014, 192–201 (ACM, New York, NY, USA, 2014). doi:[10.1145/2597073.2597076](https://doi.org/10.1145/2597073.2597076).
- 45 Gorgolewski, K. J. *et al.* BIDS Apps: Improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. *PLOS Computational Biology* **13**, e1005209 (2017). doi:[10.1371/journal.pcbi.1005209](https://doi.org/10.1371/journal.pcbi.1005209).
- 46 Beaulieu-Jones, B. K. & Greene, C. S. Reproducibility of computational workflows is automated using continuous analysis. *Nature Biotechnology* **35**, 342 (2017). doi:[10.1038/nbt.3780](https://doi.org/10.1038/nbt.3780).
- 47 Kurtzer, G. M., Sochat, V. & Bauer, M. W. Singularity: Scientific containers for mobility of compute. *PLOS ONE* **12**, e0177459 (2017). doi:[10.1371/journal.pone.0177459](https://doi.org/10.1371/journal.pone.0177459).
- 48 Schonberg, T. *et al.* Decreasing Ventromedial Prefrontal Cortex Activity During Sequential Risk-Taking: An fMRI Investigation of the Balloon Analog Risk Task. *Frontiers in Neuroscience* **6** (2012). doi:[10.3389/fnins.2012.00080](https://doi.org/10.3389/fnins.2012.00080).
- 49 Aron, A. R., Gluck, M. A. & Poldrack, R. A. Long-term test-retest reliability of functional MRI in a classification learning task. *NeuroImage* **29**, 1000–1006 (2006). doi:[10.1016/j.neuroimage.2005.08.010](https://doi.org/10.1016/j.neuroimage.2005.08.010).
- 50 Xue, G. & Poldrack, R. A. The Neural Substrates of Visual Perceptual Learning of Words: Implications for the Visual Word Form Area Hypothesis. *Journal of Cognitive Neuroscience* **19**, 1643–1655 (2007). doi:[10.1162/jocn.2007.19.10.1643](https://doi.org/10.1162/jocn.2007.19.10.1643).
- 51 Tom, S. M., Fox, C. R., Trepel, C. & Poldrack, R. A. The Neural Basis of Loss Aversion in Decision-Making Under Risk. *Science* **315**, 515–518 (2007). doi:[10.1126/science.1134239](https://doi.org/10.1126/science.1134239).
- 52 Xue, G., Aron, A. R. & Poldrack, R. A. Common Neural Substrates for Inhibition of Spoken and Manual Responses. *Cerebral Cortex* **18**, 1923–1932 (2008). doi:[10.1093/cercor/bhm220](https://doi.org/10.1093/cercor/bhm220).
- 53 Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J. & Poldrack, R. A. Triangulating a Cognitive Control Network Using Diffusion-Weighted Magnetic Resonance Imaging (MRI) and Functional MRI. *Journal of Neuroscience* **27**, 3743–3752 (2007). doi:[10.1523/JNEUROSCI.0519-07.2007](https://doi.org/10.1523/JNEUROSCI.0519-07.2007).
- 54 Foerde, K., Knowlton, B. J. & Poldrack, R. A. Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences* **103**, 11778–11783 (2006). doi:[10.1073/pnas.0602659103](https://doi.org/10.1073/pnas.0602659103).

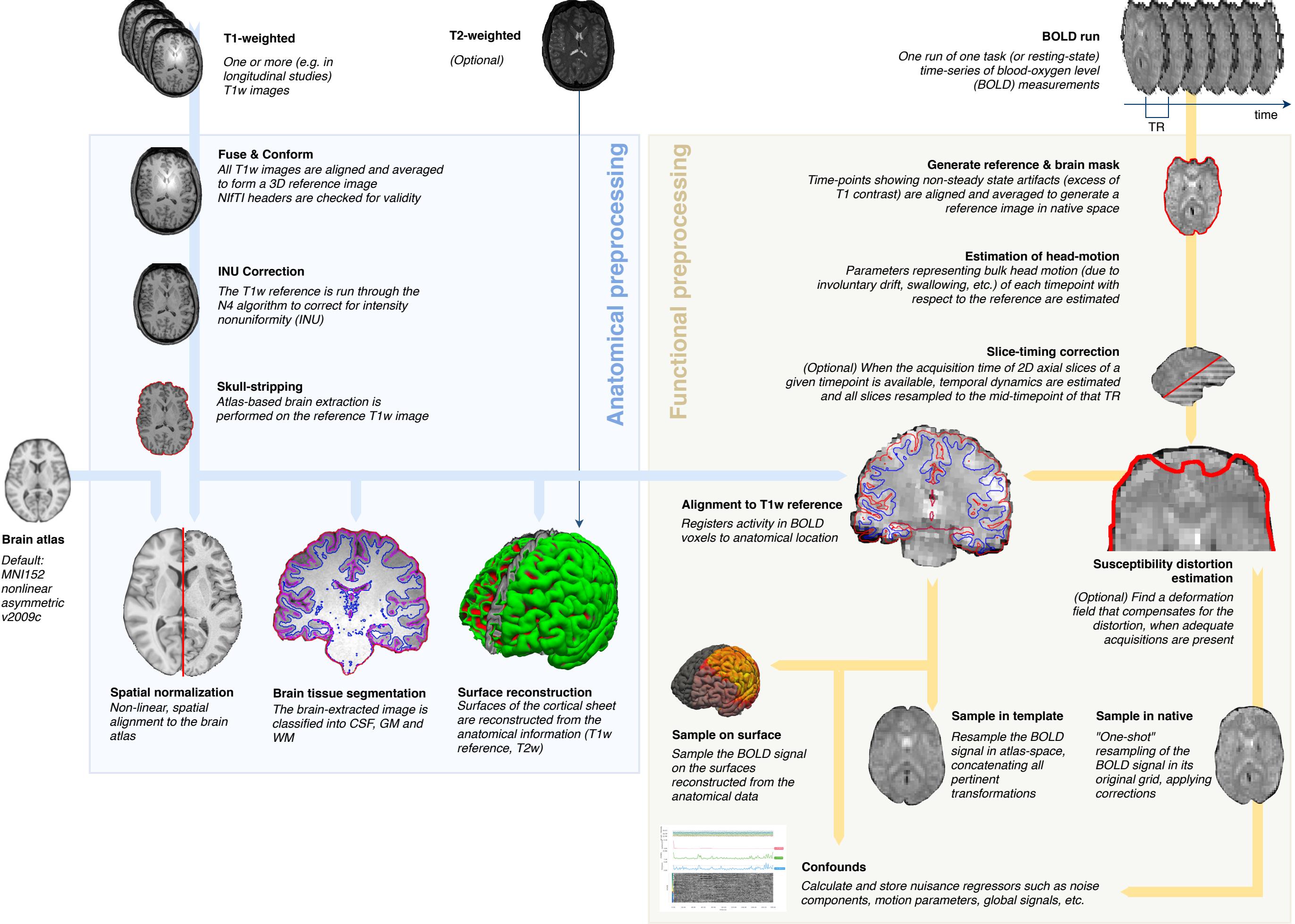
- 55 Poldrack, R. A. *et al.* A phenome-wide examination of neural and cognitive function. *Scientific Data* **3**, 160110 (2016). doi:[10.1038/sdata.2016.110](https://doi.org/10.1038/sdata.2016.110).
- 56 Gorgolewski, K. J., Durnez, J. & Poldrack, R. A. Preprocessed Consortium for Neuropsychiatric Phenomics dataset. *F1000Research* **6**, 1262 (2017). doi:[10.12688/f1000research.11964.2](https://doi.org/10.12688/f1000research.11964.2).
- 57 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](https://doi.org/10.1016/j.neuron.2015.06.037).
- 58 Alvarez, R., Jaszewski, G. & Poldrack, R. A. Building memories in two languages: an fMRI study of episodic encoding in bilinguals. In *SfN Neuroscience* (Orlando, FL, US, 2002). URL <http://www.sfn.org/annual-meeting/past-and-future-annual-meetings/abstract-archive/abstract-archive-detail>.
- 59 Poldrack, R. A. *et al.* Interactive memory systems in the human brain. *Nature* **414**, 546–550 (2001). doi:[10.1038/35107080](https://doi.org/10.1038/35107080).
- 60 Kelly, A. M. C., Uddin, L. Q., Biswal, B. B., Castellanos, F. X. & Milham, M. P. Competition between functional brain networks mediates behavioral variability. *NeuroImage* **39**, 527–537 (2008). doi:[10.1016/j.neuroimage.2007.08.008](https://doi.org/10.1016/j.neuroimage.2007.08.008).
- 61 Mennes, M. *et al.* Inter-individual differences in resting-state functional connectivity predict task-induced BOLD activity. *NeuroImage* **50**, 1690–1701 (2010). doi:[10.1016/j.neuroimage.2010.01.002](https://doi.org/10.1016/j.neuroimage.2010.01.002).
- 62 Mennes, M. *et al.* Linking inter-individual differences in neural activation and behavior to intrinsic brain dynamics. *NeuroImage* **54**, 2950–2959 (2011). doi:[10.1016/j.neuroimage.2010.10.046](https://doi.org/10.1016/j.neuroimage.2010.10.046).
- 63 Haxby, J. V. *et al.* Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. *Science* **293**, 2425–2430 (2001). doi:[10.1126/science.1063736](https://doi.org/10.1126/science.1063736).
- 64 Hanson, S. J., Matsuka, T. & Haxby, J. V. Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a face area? *NeuroImage* **23**, 156–166 (2004). doi:[10.1016/j.neuroimage.2004.05.020](https://doi.org/10.1016/j.neuroimage.2004.05.020).
- 65 Duncan, K. J., Pattamadilok, C., Knierim, I. & Devlin, J. T. Consistency and variability in functional localisers. *NeuroImage* **46**, 1018–1026 (2009). doi:[10.1016/j.neuroimage.2009.03.014](https://doi.org/10.1016/j.neuroimage.2009.03.014).
- 66 Wager, T. D., Davidson, M. L., Hughes, B. L., Lindquist, M. A. & Ochsner, K. N. Prefrontal-Subcortical Pathways Mediating Successful Emotion Regulation. *Neuron* **59**, 1037–1050 (2008). doi:[10.1016/j.neuron.2008.09.006](https://doi.org/10.1016/j.neuron.2008.09.006).
- 67 Moran, J. M., Jolly, E. & Mitchell, J. P. Social-Cognitive Deficits in Normal Aging. *Journal of Neuroscience* **32**, 5553–5561 (2012). doi:[10.1523/JNEUROSCI.5511-11.2012](https://doi.org/10.1523/JNEUROSCI.5511-11.2012).
- 68 Uncapher, M. R., Hutchinson, J. B. & Wagner, A. D. Dissociable Effects of Top-Down and Bottom-Up Attention during Episodic Encoding. *Journal of Neuroscience* **31**, 12613–12628 (2011). doi:[10.1523/JNEUROSCI.0152-11.2011](https://doi.org/10.1523/JNEUROSCI.0152-11.2011).
- 69 Gorgolewski, K. J. *et al.* A test-retest fMRI dataset for motor, language and spatial attention functions. *GigaScience* **2**, 1–4 (2013). doi:[10.1186/2047-217X-2-6](https://doi.org/10.1186/2047-217X-2-6).
- 70 Repovs, G. & Barch, D. M. Working memory related brain network connectivity in individuals with schizophrenia and their siblings. *Frontiers in Human Neuroscience* **6** (2012). doi:[10.3389/fnhum.2012.00137](https://doi.org/10.3389/fnhum.2012.00137).
- 71 Repovs, G., Csernansky, J. G. & Barch, D. M. Brain Network Connectivity in Individuals with Schizophrenia and Their Siblings. *Biological Psychiatry* **69**, 967–973 (2011). doi:[10.1016/j.biopsych.2010.11.009](https://doi.org/10.1016/j.biopsych.2010.11.009).
- 72 Walz, J. M. *et al.* Simultaneous EEG-fMRI Reveals Temporal Evolution of Coupling between Supramodal Cortical Attention Networks and the Brainstem. *Journal of Neuroscience* **33**, 19212–19222 (2013). doi:[10.1523/JNEUROSCI.2649-13.2013](https://doi.org/10.1523/JNEUROSCI.2649-13.2013).
- 73 Walz, J. M. *et al.* Simultaneous EEG-fMRI reveals a temporal cascade of task-related and default-mode activations during a simple target detection task. *NeuroImage* **102**, 229–239 (2014). doi:[10.1016/j.neuroimage.2013.08.014](https://doi.org/10.1016/j.neuroimage.2013.08.014).
- 74 Conroy, B. R., Walz, J. M. & Sajda, P. Fast Bootstrapping and Permutation Testing for Assessing Reproducibility and Interpretability of Multivariate fMRI Decoding Models. *PLOS ONE* **8**, e79271 (2013). doi:[10.1371/journal.pone.0079271](https://doi.org/10.1371/journal.pone.0079271).

- 75 Walz, J. M. *et al.* Prestimulus EEG alpha oscillations modulate task-related fMRI BOLD responses to auditory stimuli. *NeuroImage* **113**, 153–163 (2015). doi:[10.1016/j.neuroimage.2015.03.028](https://doi.org/10.1016/j.neuroimage.2015.03.028).
- 76 Velanova, K., Wheeler, M. E. & Luna, B. Maturational Changes in Anterior Cingulate and Frontoparietal Recruitment Support the Development of Error Processing and Inhibitory Control. *Cerebral Cortex* **18**, 2505–2522 (2008). doi:[10.1093/cercor/bhn012](https://doi.org/10.1093/cercor/bhn012).
- 77 Padmanabhan, A., Geier, C. F., Ordaz, S. J., Teslovich, T. & Luna, B. Developmental changes in brain function underlying the influence of reward processing on inhibitory control. *Developmental Cognitive Neuroscience* **1**, 517–529 (2011). doi:[10.1016/j.dcn.2011.06.004](https://doi.org/10.1016/j.dcn.2011.06.004).
- 78 Geier, C. F., Terwilliger, R., Teslovich, T., Velanova, K. & Luna, B. Immaturities in Reward Processing and Its Influence on Inhibitory Control in Adolescence. *Cerebral Cortex* **20**, 1613–1629 (2010). doi:[10.1093/cercor/bhp225](https://doi.org/10.1093/cercor/bhp225).
- 79 Cera, N., Tartaro, A. & Sensi, S. L. Modafinil Alters Intrinsic Functional Connectivity of the Right Posterior Insula: A Pharmacological Resting State fMRI Study. *PLOS ONE* **9**, e107145 (2014). doi:[10.1371/journal.pone.0107145](https://doi.org/10.1371/journal.pone.0107145).
- 80 Woo, C.-W., Roy, M., Buhle, J. T. & Wager, T. D. Distinct Brain Systems Mediate the Effects of Nociceptive Input and Self-Regulation on Pain. *PLOS Biology* **13**, e1002036 (2015). doi:[10.1371/journal.pbio.1002036](https://doi.org/10.1371/journal.pbio.1002036).
- 81 Smeets, P. A. M., Kroese, F. M., Evers, C. & de Ridder, D. T. D. Allured or alarmed: Counteractive control responses to food temptations in the brain. *Behavioural Brain Research* **248**, 41–45 (2013). doi:[10.1016/j.bbr.2013.03.041](https://doi.org/10.1016/j.bbr.2013.03.041).
- 82 Pernet, C. R. *et al.* The human voice areas: Spatial organization and inter-individual variability in temporal and extra-temporal cortices. *NeuroImage* **119**, 164–174 (2015). doi:[10.1016/j.neuroimage.2015.06.050](https://doi.org/10.1016/j.neuroimage.2015.06.050).
- 83 Verstynen, T. D. The organization and dynamics of corticostriatal pathways link the medial orbitofrontal cortex to future behavioral responses. *Journal of Neurophysiology* **112**, 2457–2469 (2014). doi:[10.1152/jn.00221.2014](https://doi.org/10.1152/jn.00221.2014).
- 84 Bursley, J. K., Nestor, A., Tarr, M. J. & Creswell, J. D. Awake, Offline Processing during Associative Learning. *PLOS ONE* **11**, e0127522 (2016). doi:[10.1371/journal.pone.0127522](https://doi.org/10.1371/journal.pone.0127522).
- 85 Ella, G., David, M. & Avi, K. Learning from the other limb's experience: sharing the trained M1 representation of the motor sequence knowledge. *The Journal of Physiology* **594**, 169–188 (2015). doi:[10.1113/JP270184](https://doi.org/10.1113/JP270184).
- 86 Gabitov, E., Manor, D. & Karni, A. Patterns of Modulation in the Activity and Connectivity of Motor Cortex during the Repeated Generation of Movement Sequences. *Journal of Cognitive Neuroscience* **27**, 736–751 (2014). doi:[10.1162/jocn_a_00751](https://doi.org/10.1162/jocn_a_00751).
- 87 Gabitov, E., Manor, D. & Karni, A. Done That: Short-term Repetition Related Modulations of Motor Cortex Activity as a Stable Signature for Overnight Motor Memory Consolidation. *Journal of Cognitive Neuroscience* **26**, 2716–2734 (2014). doi:[10.1162/jocn_a_00675](https://doi.org/10.1162/jocn_a_00675).
- 88 Lepping, R. J., Atchley, R. A. & Savage, C. R. Development of a validated emotionally provocative musical stimulus set for research. *Psychology of Music* **44**, 1012–1028 (2016). doi:[10.1177/0305735615604509](https://doi.org/10.1177/0305735615604509).
- 89 Park, C.-A. & Kang, C.-K. Sensing the effects of mouth breathing by using 3-tesla MRI. *Journal of the Korean Physical Society* **70**, 1070–1076 (2017). doi:[10.3938/jkps.70.1070](https://doi.org/10.3938/jkps.70.1070).
- 90 Iannilli, E. *et al.* Effects of Manganese Exposure on Olfactory Functions in Teenagers: A Pilot Study. *PLOS ONE* **11**, e0144783 (2016). doi:[10.1371/journal.pone.0144783](https://doi.org/10.1371/journal.pone.0144783).
- 91 Kim, J., Wang, J., Wedell, D. H. & Shinkareva, S. V. Identifying Core Affect in Individuals from fMRI Responses to Dynamic Naturalistic Audiovisual Stimuli. *PLOS ONE* **11**, e0161589 (2016). doi:[10.1371/journal.pone.0161589](https://doi.org/10.1371/journal.pone.0161589).
- 92 Tétreault, P. *et al.* Brain Connectivity Predicts Placebo Response across Chronic Pain Clinical Trials. *PLOS Biology* **14**, e1002570 (2016). doi:[10.1371/journal.pbio.1002570](https://doi.org/10.1371/journal.pbio.1002570).
- 93 Chakroff, A. *et al.* When minds matter for moral judgment: intent information is neurally encoded for harmful but not impure acts. *Social Cognitive and Affective Neuroscience* **11**, 476–484 (2016). doi:[10.1093/scan/nsv131](https://doi.org/10.1093/scan/nsv131).
- 94 Koster-Hale, J., Saxe, R., Dungan, J. & Young, L. L. Decoding moral judgments from neural representations of intentions. *Proceedings of the National Academy of Sciences* **110**, 5648–5653 (2013). doi:[10.1073/pnas.1207992110](https://doi.org/10.1073/pnas.1207992110).

- 95 Gao, X. *et al.* My Body Looks Like That Girls: Body Mass Index Modulates Brain Activity during Body Image Self-Reflection among Young Women. *PLOS ONE* **11**, e0164450 (2016). doi:[10.1371/journal.pone.0164450](https://doi.org/10.1371/journal.pone.0164450).
- 96 Romaniuk, L., Pope, M., Nicol, K., Steele, D. & Hall, J. Neural correlates of fears of abandonment and rejection in borderline personality disorder. *Wellcome Open Research* **1**, 33 (2016). doi:[10.12688/wellcomeopenres.10331.1](https://doi.org/10.12688/wellcomeopenres.10331.1).
- 97 Cohen, A. D., Nencka, A. S., Lebel, R. M. & Wang, Y. Multiband multi-echo imaging of simultaneous oxygenation and flow timeseries for resting state connectivity. *PLOS ONE* **12**, e0169253 (2017). doi:[10.1371/journal.pone.0169253](https://doi.org/10.1371/journal.pone.0169253).
- 98 Dalenberg, J. R., Weitkamp, L., Renken, R. J., Nanetti, L. & Horst, G. J. t. Flavor pleasantness processing in the ventral emotion network. *PLOS ONE* **12**, e0170310 (2017). doi:[10.1371/journal.pone.0170310](https://doi.org/10.1371/journal.pone.0170310).
- 99 Roy, A. *et al.* The evolution of cost-efficiency in neural networks during recovery from traumatic brain injury. *PLOS ONE* **12**, e0170541 (2017). doi:[10.1371/journal.pone.0170541](https://doi.org/10.1371/journal.pone.0170541).
- 100 Gordon, E. M. *et al.* Precision Functional Mapping of Individual Human Brains. *Neuron* **95**, 791–807.e7 (2017). doi:[10.1016/j.neuron.2017.07.011](https://doi.org/10.1016/j.neuron.2017.07.011).
- 101 Veldhuizen, M. G. *et al.* Integration of Sweet Taste and Metabolism Determines Carbohydrate Reward. *Current Biology* **27**, 2476–2485.e6 (2017). doi:[10.1016/j.cub.2017.07.018](https://doi.org/10.1016/j.cub.2017.07.018).
- 102 Greene, D. J. *et al.* Behavioral interventions for reducing head motion during MRI scans in children. *NeuroImage* **171**, 234–245 (2018). doi:[10.1016/j.neuroimage.2018.01.023](https://doi.org/10.1016/j.neuroimage.2018.01.023).
- 103 Nastase, S. A. *et al.* Attention Selectively Reshapes the Geometry of Distributed Semantic Representation. *Cerebral Cortex* **27**, 4277–4291 (2017). doi:[10.1093/cercor/bhx138](https://doi.org/10.1093/cercor/bhx138).
- 104 Kanazawa, Y. *et al.* Phonological memory in sign language relies on the visuomotor neural system outside the left hemisphere language network. *PLOS ONE* **12**, e0177599 (2017). doi:[10.1371/journal.pone.0177599](https://doi.org/10.1371/journal.pone.0177599).
- 105 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](https://doi.org/10.1371/journal.pone.0184661).
- 106 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](https://doi.org/10.1002/hbm.23737).
- 107 Beckmann, C. F., Jenkinson, M. & Smith, S. M. General multilevel linear modeling for group analysis in fMRI. *NeuroImage* **20**, 1052–1063 (2003). doi:[10.1016/S1053-8119\(03\)00435-X](https://doi.org/10.1016/S1053-8119(03)00435-X).
- 108 Strother, S. C. *et al.* The Quantitative Evaluation of Functional Neuroimaging Experiments: The NPAIRS Data Analysis Framework. *NeuroImage* **15**, 747–771 (2002). doi:[10.1006/nimg.2001.1034](https://doi.org/10.1006/nimg.2001.1034).
- 109 Karaman, M., Nencka, A. S., Bruce, I. P. & Rowe, D. B. Quantification of the Statistical Effects of Spatiotemporal Processing of Nontask fMRI Data. *Brain Connectivity* **4**, 649–661 (2014). doi:[10.1089/brain.2014.0278](https://doi.org/10.1089/brain.2014.0278).
- 110 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](https://doi.org/10.3389/fninf.2011.00013).
- 111 Marder, E. Understanding Brains: Details, Intuition, and Big Data. *PLOS Biology* **13**, e1002147 (2015). doi:[10.1371/journal.pbio.1002147](https://doi.org/10.1371/journal.pbio.1002147).

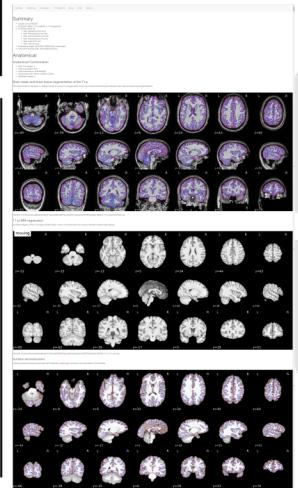
Table 2. Data from OpenfMRI used in evaluation. S: number of sessions; T: number of tasks; R: number of BOLD runs; Modalities: number of runs for each modality, per subject (FM indicates acquisitions for susceptibility distortion correction); Part. IDs (phase): participant identifiers included in testing phase; N: total of unique participants; TR: repetition time; #TR: length of time-series (volumes); Resolution: voxel size of BOLD series.

DS000XXX	Scanner	S	T	R	Modalities	Part. IDs (Phase I)	Part. IDs (Phase II)	N	TR	#TR	Resolution
001 ⁴⁸	SIEMENS	1	1	21	1 T1w, 3 BOLD	02, 03, 09, 15	01, 02, 07, 08	7	2.0	6300	3.12x3.12x4.00
002 ⁴⁹	SIEMENS	1	3	48	1 T1w, 6 BOLD	01, 11, 14, 15	02, 03, 04, 10	8	2.0	9510	3.12x3.12x5.00
003 ⁵⁰	SIEMENS	1	1	6	1 T1w, 1 BOLD	03, 07, 09, 11	02, 09, 10, 11	6	2.0	956	3.12x3.12x4.00
005 ⁵¹	SIEMENS	1	1	21	1 T1w, 3 BOLD	01, 03, 06, 14	01, 04, 05, 15	7	2.0	5040	3.12x3.12x4.00
007 ⁵²	SIEMENS	1	3	46	1 T1w, 5 BOLD	09, 11, 18, 20	03, 04, 08, 12	8	2.0	8205	3.12x3.12x4.00
008 ⁵³	SIEMENS	1	2	38	1 T1w, 5 BOLD	04, 09, 12, 14	10, 12, 13, 15	7	2.0	6808	3.12x3.12x4.39
009	SIEMENS	1	4	48	1 T1w, 6 BOLD	01, 03, 09, 10	17, 18, 21, 23	8	2.0	10528	3.00x3.00x4.00
011 ⁵⁴	SIEMENS	1	4	41	1 T1w, 5 BOLD	01, 03, 06, 08	03, 09, 11, 14	7	2.0	8041	3.12x3.12x5.00
017	SIEMENS	2	2	48	4 T1w, 9 BOLD	2, 4, 7, 8	2, 5, 7, 8	5	2.0	8736	3.12x3.12x4.00
030 ^{55,56}	SIEMENS	1	8	30	1 T1w, 7 BOLD		10[440,638,668,855]	4	2.2	6254	3.00x3.00x4.00
031 ⁵⁷	SIEMENS	107	9	191	29 T1w, 18 T2w, 46 FM, 191 BOLD		01	1	1.2	79017	2.55x2.55x2.54
051 ⁵⁸	SIEMENS	1	1	54	2 T1w, 7 BOLD	03, 04, 05, 13	02, 04, 06, 09	7	2.0	10800	3.12x3.12x6.00
052 ⁵⁹	SIEMENS	1	2	28	2 T1w, 4 BOLD	06, 08, 12, 14	05, 10, 12, 13	7	2.0	6300	3.12x3.12x6.00
053	SIEMENS	1	3	32	1 T1w, 8 BOLD		002, 003, 005, 006	4	1.2	10712	2.40x2.40x2.40
101	SIEMENS	1	1	16	1 T1w, 2 BOLD	06, 08, 16, 19	05, 11, 17, 20	8	2.0	2416	3.00x3.00x4.00
102 ⁶⁰⁻⁶²	SIEMENS	1	1	16	1 T1w, 2 BOLD	05, 19, 22, 23	08, 10, 16, 20	8	2.0	2336	3.00x3.00x4.00
105 ^{63,64}	GE	1	1	71	1 T1w, 11 BOLD	1, 2, 3, 6	1, 4, 5, 6	6	2.5	8591	3.50x3.75x3.75
107 ⁶⁵	SIEMENS	1	1	14	1 T1w, 2 BOLD	02, 05, 20, 29	05, 36, 39, 47	7	3.0	2315	3.00x3.00x3.00
108 ⁶⁶	GE	1	1	41	1 T1w, 5 BOLD	01, 03, 07, 17	03, 10, 24, 26	7	2.0	7860	3.44x3.44x4.50
109 ⁶⁷	SIEMENS	1	1	12	1 T1w, 2 BOLD	02, 10, 39, 47	02, 11, 15, 39	6	2.0	2148	3.00x3.00x3.54
110 ⁶⁸	GE	1	1	80	1 T1w, 10 BOLD	07, 09, 17, 18	01, 02, 03, 06	8	2.0	14880	3.44x3.44x4.01
114 ⁶⁹	GE	2	5	70	2 T1w, 10 BOLD	01, 05, 07, 08	02, 03, 04, 07	7	5.0	10626	4.00x4.00x4.00
115 ^{70,71}	SIEMENS	1	3	24	1 T1w, 3 BOLD	31, 68, 77, 78	04, 33, 67, 79	8	2.5	3288	4.00x4.00x4.00
116 ⁷²⁻⁷⁵	PHILIPS	1	2	36	1 T1w, 6 BOLD	02, 08, 10, 15	08, 12, 15, 17	6	2.0	6120	3.00x3.00x4.00
119 ⁷⁶	SIEMENS	1	1	31	1 T1w, 3 BOLD	10, 51, 59, 74	11, 26, 56, 58	8	1.5	7564	3.12x3.12x4.00
120 ⁷⁷	SIEMENS	1	1	11	1 T1w, 2 BOLD		04, 05, 08, 24	4	1.5	2376	3.12x3.12x4.00
121 ⁷⁸	SIEMENS	1	1	28	1 T1w, 4 BOLD	01, 04, 05, 20	01, 18, 22, 26	7	1.5	5656	3.12x3.12x4.00
133 ⁷⁹	PHILIPS	2	1	24	2 T1w, 6 BOLD		06, 21, 22, 23	4	N/A	3480	4.00x4.00x4.00
140 ⁸⁰	PHILIPS	1	1	36	1 T1w, 9 BOLD		05, 27, 32, 33	4	2.0	7380	2.80x2.80x3.00
148	GE	1	1	12	1 T1w, 1 T2w, 3 BOLD		09, 26, 28, 33	4	1.8	3162	3.00x3.00x3.00
157 ⁸¹	PHILIPS	1	1	4	1 T1w, 1 BOLD		04, 21, 23, 28	4	1.6	1485	4.00x4.00x3.99
158 ⁸²	SIEMENS	1	1	4	1 T1w, 1 BOLD		064, 081, 122, 149	4	2.0	1240	3.00x3.00x3.30
164 ⁸³	SIEMENS	1	1	4	1 T1w, 1 BOLD		006, 012, 019, 027	4	1.5	1480	3.50x3.50x3.50
168 ⁸⁴	SIEMENS	1	1	4	1 T1w, 1 BOLD		08, 27, 30, 49	4	2.5	2112	3.00x3.00x3.00
170 ⁸⁵⁻⁸⁷	GE	1	4	48	1 T1w, 12 BOLD		1700, 1708, 1710, 1713	4	3.0	2160	3.44x3.44x3.40
171 ⁸⁸	SIEMENS	1	2	20	1 T1w, 5 BOLD		control0[4,8,14], mdd03	4	3.0	2066	2.90x2.90x3.00
177 ⁸⁹	SIEMENS	1	1	4	1 T1w, 1 BOLD		04, 07, 10, 11	4	3.0	920	3.00x3.00x3.00
200 ⁹⁰	SIEMENS	1	1	4	1 T1w, 1 BOLD		2004, 2011, 2012, 2014	4	2.5	480	3.28x3.28x4.29
205 ⁹¹	SIEMENS	1	2	12	1 T1w, 3 BOLD		01, 05, 06, 07	4	2.2	4103	3.00x3.00x3.00
208 ⁹²	SIEMENS	1	1	4	1 T1w, 1 BOLD		27, 45, 56, 69	4	2.5	1200	3.44x3.44x3.00
212 ^{93,94}	SIEMENS	1	2	40	1 T1w, 10 BOLD		07, 13, 20, 29	4	3.0	5808	3.12x3.12x4.00
213 ⁹⁵	SIEMENS	1	1	4	1 T1w, 1 BOLD		06, 10, 12, 13	4	2.0	1120	3.00x3.00x3.99
214 ⁹⁶	SIEMENS	1	1	4	1 T1w, 1 BOLD		EESSO[06,31,33,34]	4	1.6	1364	3.44x3.44x5.00
216 ⁹⁷	GE	1	1	16	1 T1w, 4 BOLD (ME)		01, 02, 03, 04	4	3.5	2688	3.00x3.00x3.00
218 ⁹⁸	PHILIPS	1	1	12	1 T1w, 3 BOLD		02, 07, 12, 17	4	1.5	6709	2.88x3.00x2.88
219 ⁹⁸	PHILIPS	1	1	14	1 T1w, 3 BOLD		04, 09, 10, 12	4	1.5	7807	2.88x3.00x2.88
220 ⁹⁹	PHILIPS, SIEMENS	3	1	12	3 T1w, 3 BOLD		tbi[03,05,06,10]	4	N/A	1728	3.00x3.00x4.00
221	SIEMENS	2	1	15	1 MP2RAGE, 9 FM, 3 BOLD		010[016,064,125,251]	4	2.5	9855	2.30x2.30x2.30
224 ¹⁰⁰	SIEMENS	12	6	399	4 T1w, 4 T2w, 10 FM, 79 BOLD	MSC[05,06,08,09]	MSC[05,08,09,10]	5	2.2	88528	4.00x4.00x4.00
228	SIEMENS	1	1	4	1 T1w, 1 BOLD		pixar[001,017,103,132]	4	2.0	672	3.06x3.06x3.29
229 ¹⁰¹	SIEMENS	1	1	12	1 T1w, 3 BOLD		02, 05, 07, 10	4	2.0	4680	3.44x3.44x3.00
231 ¹⁰²	SIEMENS	1	1	12	1 T1w, 3 BOLD		01, 02, 03, 09	4	2.0	4548	2.02x2.02x2.00
233 ¹⁰³	PHILIPS	1	2	80	2 T1w, 10 BOLD	rid0000[12,24,36,41]	rid0000[01,17,31,32]	8	2.0	15680	3.00x3.00x3.00
237 ¹⁰⁴	SIEMENS	1	1	41	1 T1w, 5 BOLD	03, 08, 11, 12	01, 03, 04, 06	7	1.0	19844	3.00x3.00x3.00
243 ⁴⁰	SIEMENS	1	1	13	1 T1w, 1 BOLD	012, 032, 042, 071	023, 066, 089, 094	8	2.5	2884	4.00x4.00x4.00
Total				2176		120	202	304		551769	



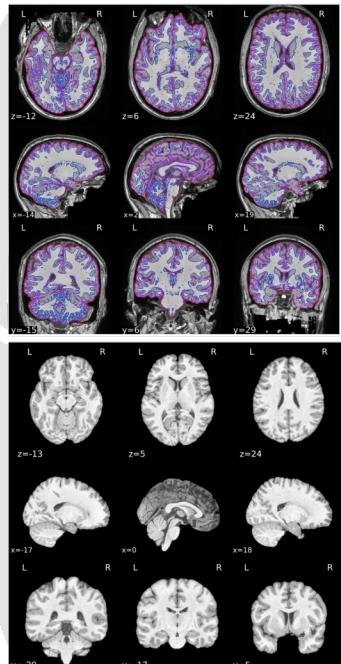
Summary

Reports start with an overview of the dataset, as identified using BIDS.



Anatomical processing

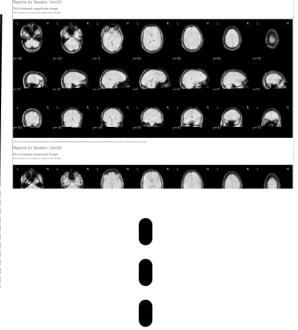
Several panels allow for quality control of the anatomical workflow. Brain tissue segmentation, spatial normalization and surface reconstruction (if requested) can be inspected using these visualization panels.



T1-weighted reference, brain mask, intensity inhomogeneity and brain tissue segmentation panel. A static mosaic allows the assessment of these four crucial steps of pre-processing anatomical images.

Fieldmaps processing

When the dataset contains any of the supported alternatives to estimate the deformation map corresponding to susceptibility distortions, these panels help assess these images were correctly processed.

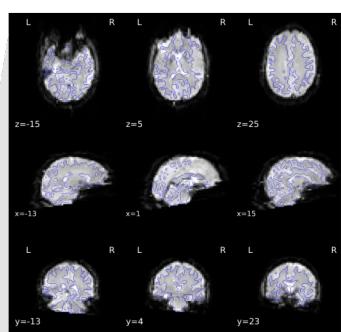
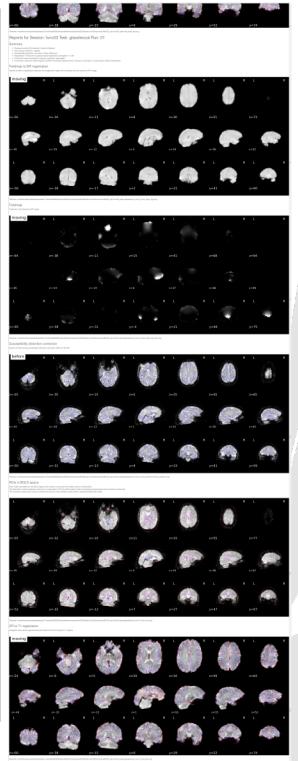


Spatial normalization. A dynamic mosaic that transitions between the target atlas space and the T1w-reference aligned into that space allows checking the accuracy of this image registration process.

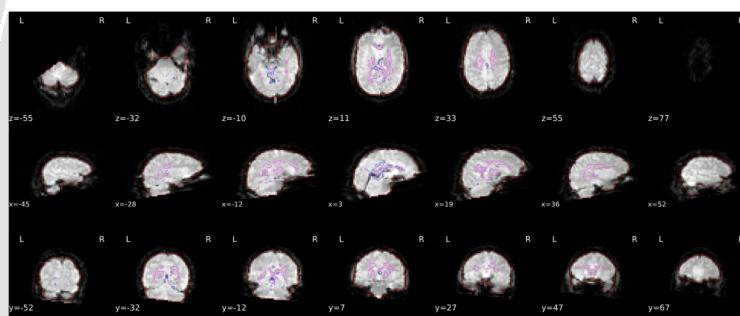
Functional processing

Each BOLD run across the different tasks and sessions will be presented at different quality control points.

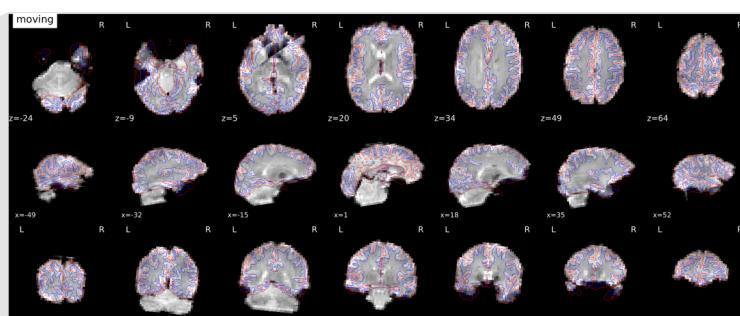
First, when fieldmaps were found, some mosaics will show the alignment of those maps to the BOLD reference. The block ends with a dynamic plot showing how images are unwarped.



Susceptibility distortion correction. If fieldmap information was found or the "fieldmap-less" correction is requested, the step is assessed with a dynamic mosaic that transitions between the unwarped ("after") and original ("before"). Contours of the white-matter are also presented as anatomical cue.



BOLD mask and CompCor ROIs. The final BOLD signal is presented, with contours representing the outline of the brain mask, and two regions-of-interest (ROIs) where CompCor confounds are estimated.



Alignment of BOLD and the T1w reference. The correct alignment to the anatomical reference is assessed with a dynamic mosaic that renders the reconstructed surfaces over the BOLD reference.

Errors

FMRIPrep is explicit about errors, and any problems encountered along the processing will be listed at the end of the report, with collapsible panels containing the specific detail of each error.

Data sample

Manual selection based on MRIQC³⁴.

4 subjects per dataset

30 out of 60 datasets

repeat failed subjects

Run fMRIPrep

bioRxiv preprint first posted online Apr. 25, 2018; doi: <http://dx.doi.org/10.1101/306951>; this version posted Apr. 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Container execution of fMRIPrep

Stage errors

Collect automatically generated error files

Visualization of reports

Quick visual assessment of all reports

Issues found?

yes

Triage issue

Tag dataset and subject
Identify the issue and potential origins.

CI tests pass

Fix Issue

Code iteration that fixes the issue

before

after

Data Sample

Random sampling

4 subjects per dataset

60 out of 60 datasets

repeat all subjects

Run fMRIPrep

Container execution of fMRIPrep

Assessment protocol

Thorough inspection and rating of all report elements, for each subject's report.

Issues found?

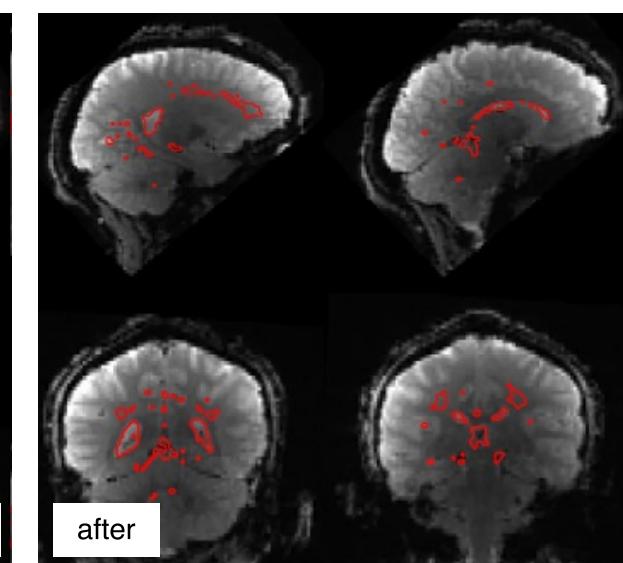
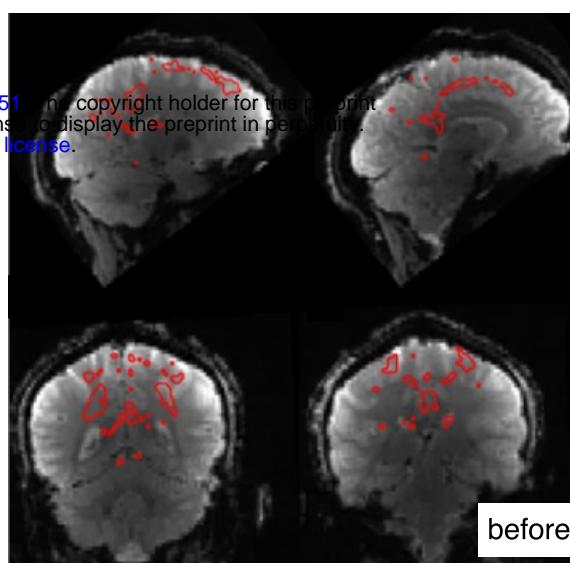
yes

Triage issue

Tag dataset and subject
Identify the issue and potential origins.

Example fix: consistency of NIfTI headers

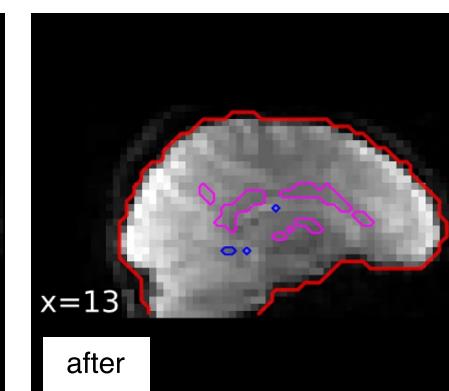
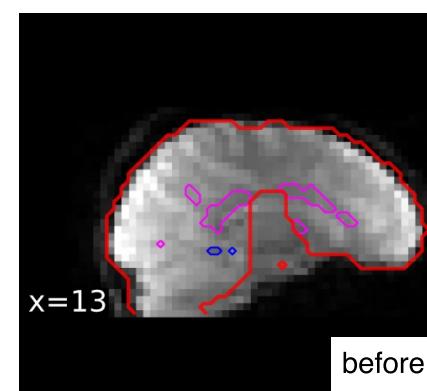
Software packages have inconsistent implementation of the two possible coordinate systems (s/q-form matrices) of the NIfTI format. Particular settings of these matrices introduce alignment errors. Solution: check these headers early and sanitize the s/q-form matrices consistently. Additionally, check and sanitize after tools that manipulate these headers.



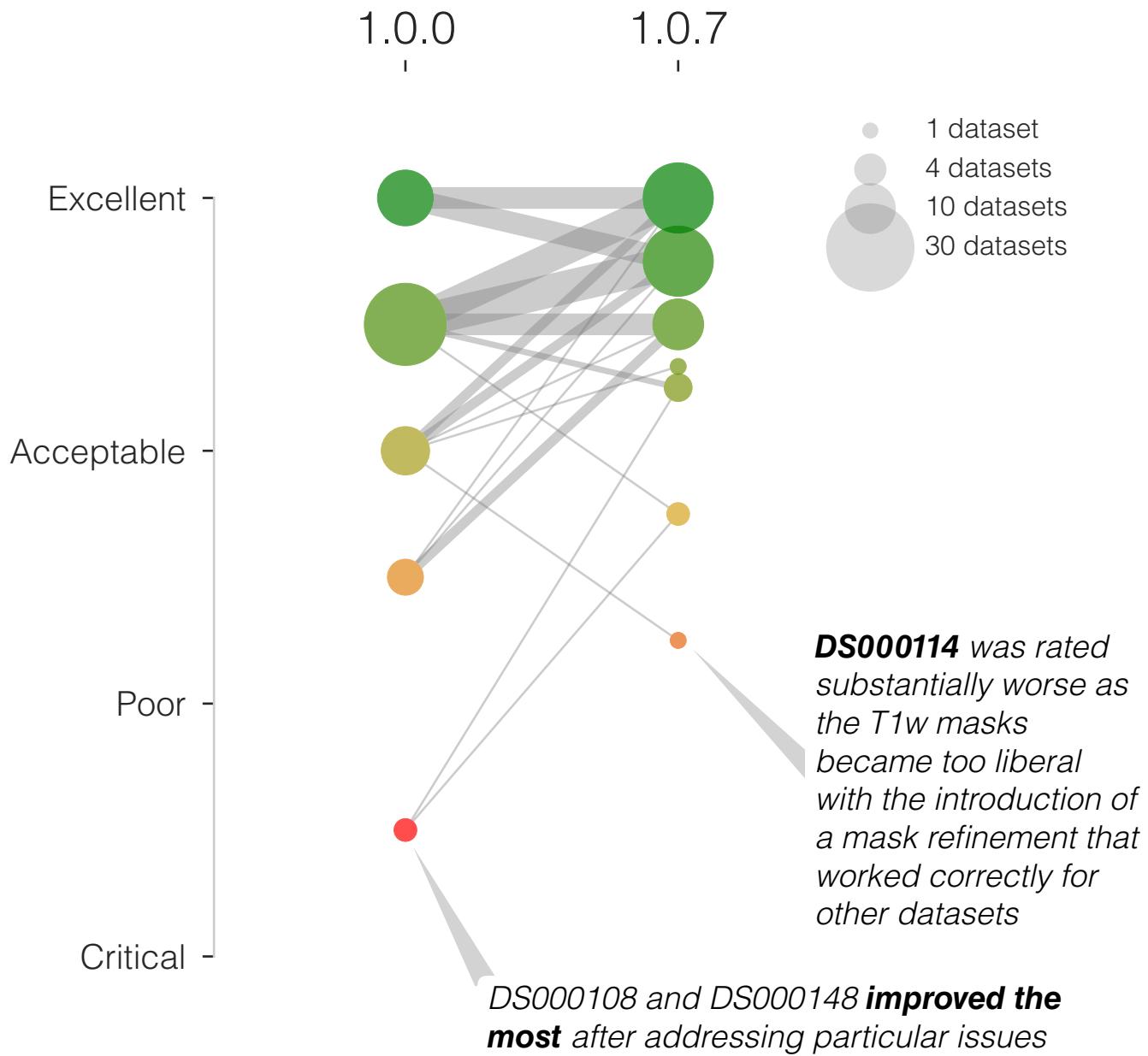
Release: version 1.0.0
Dec 6, 2017

Example fix: calculating brain masks from BOLD data

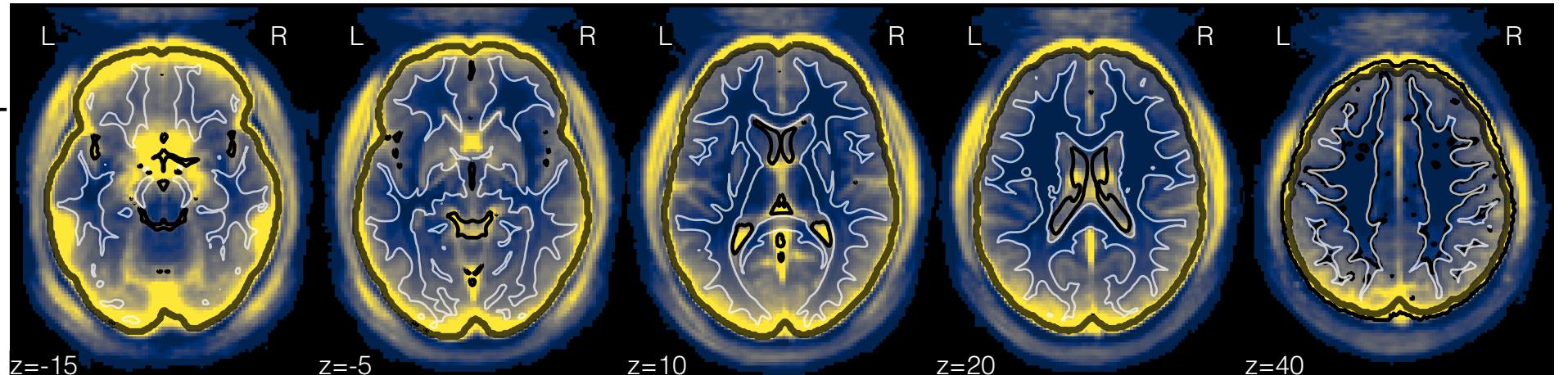
Some BOLD runs showed problematic contrast after calculating the BOLD reference image (e.g. strong intensity non-uniformity). These images would also preclude the calculation of a brain mask from the BOLD time-series. As a result, the brain masks were too conservative and left out important regions in some datasets (e.g. DS000[116, 133, 157]). Solution: a robust BOLD reference calculation and a less conservative mask calculation.



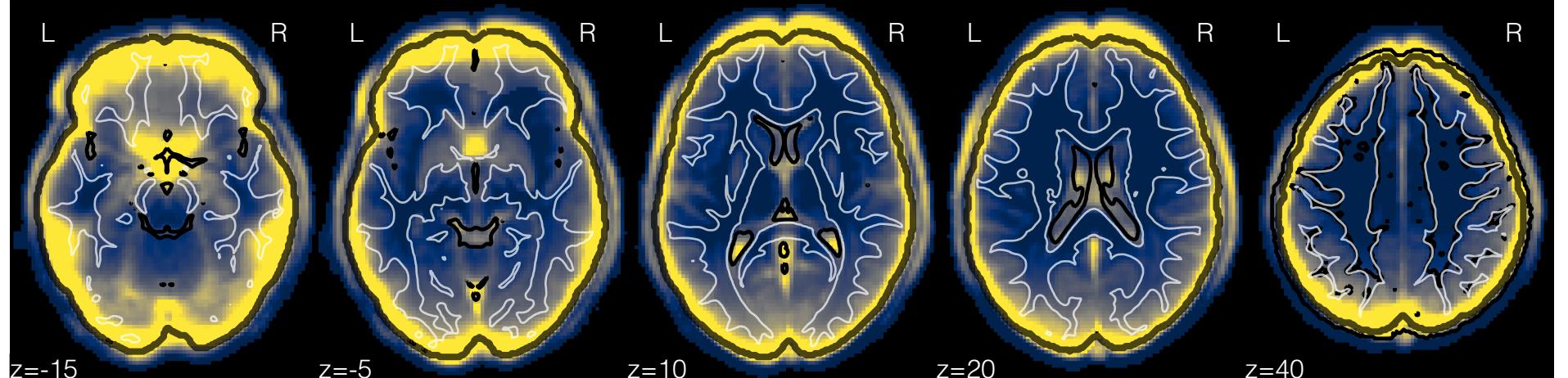
Release: version 1.0.5
Jan 21, 2018

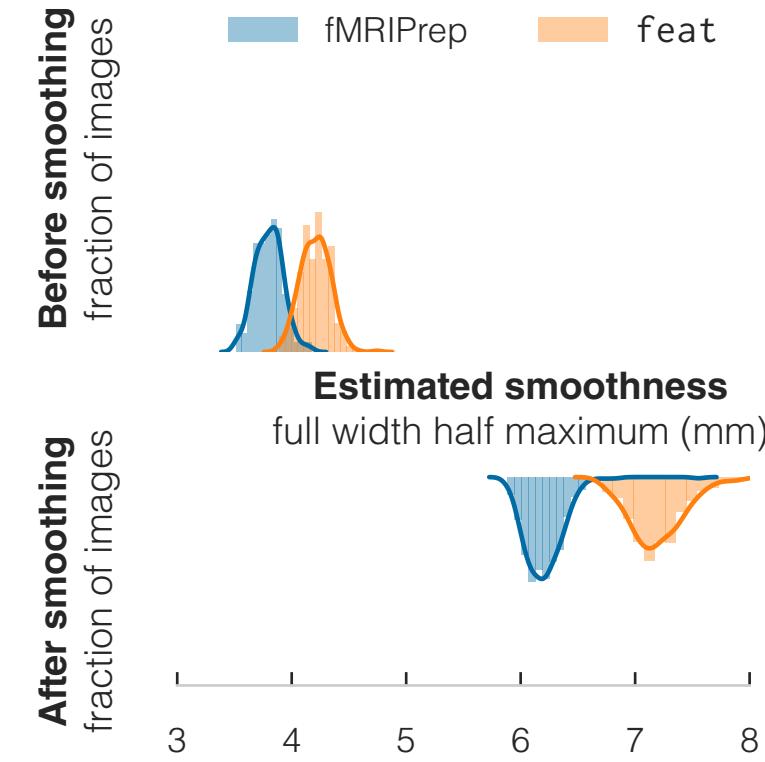


fMRIprep

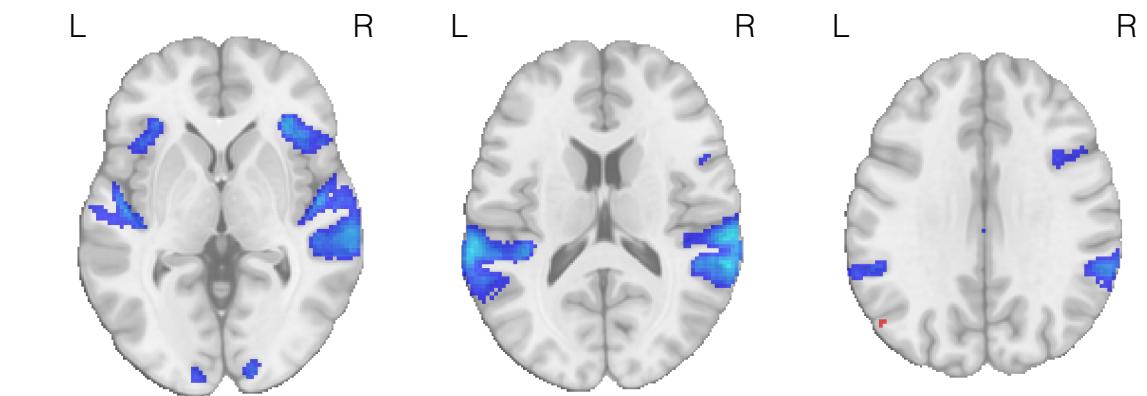


feat

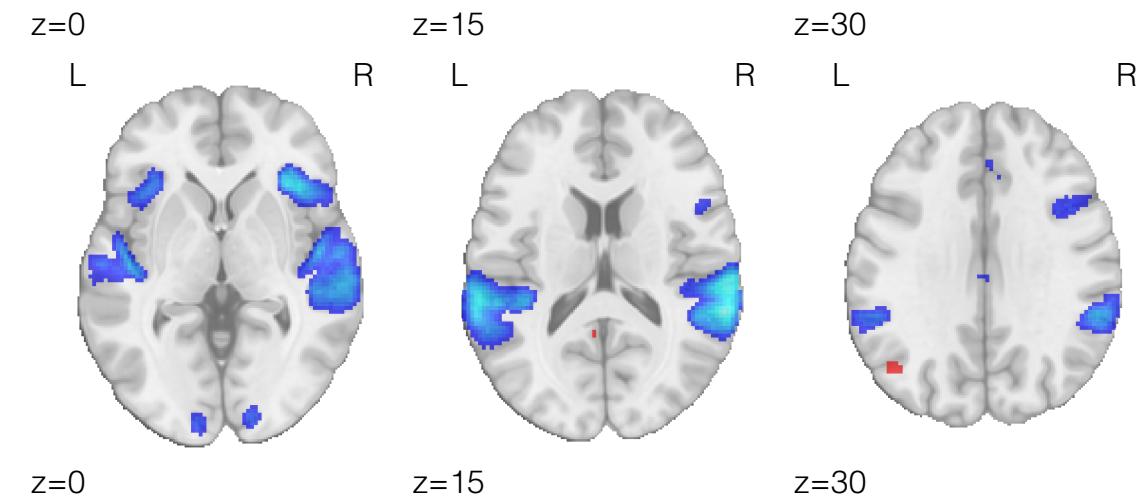


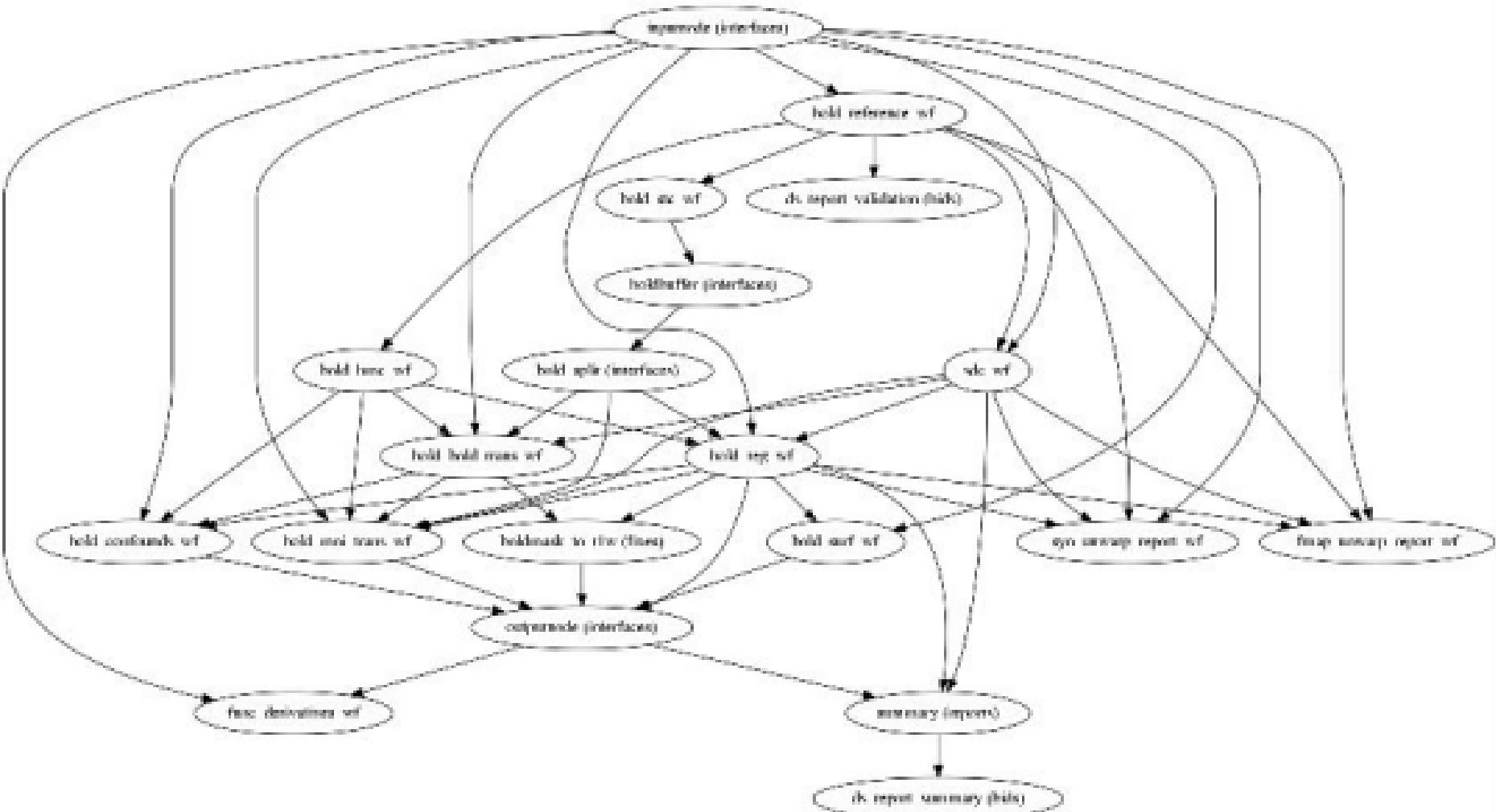
A**B**

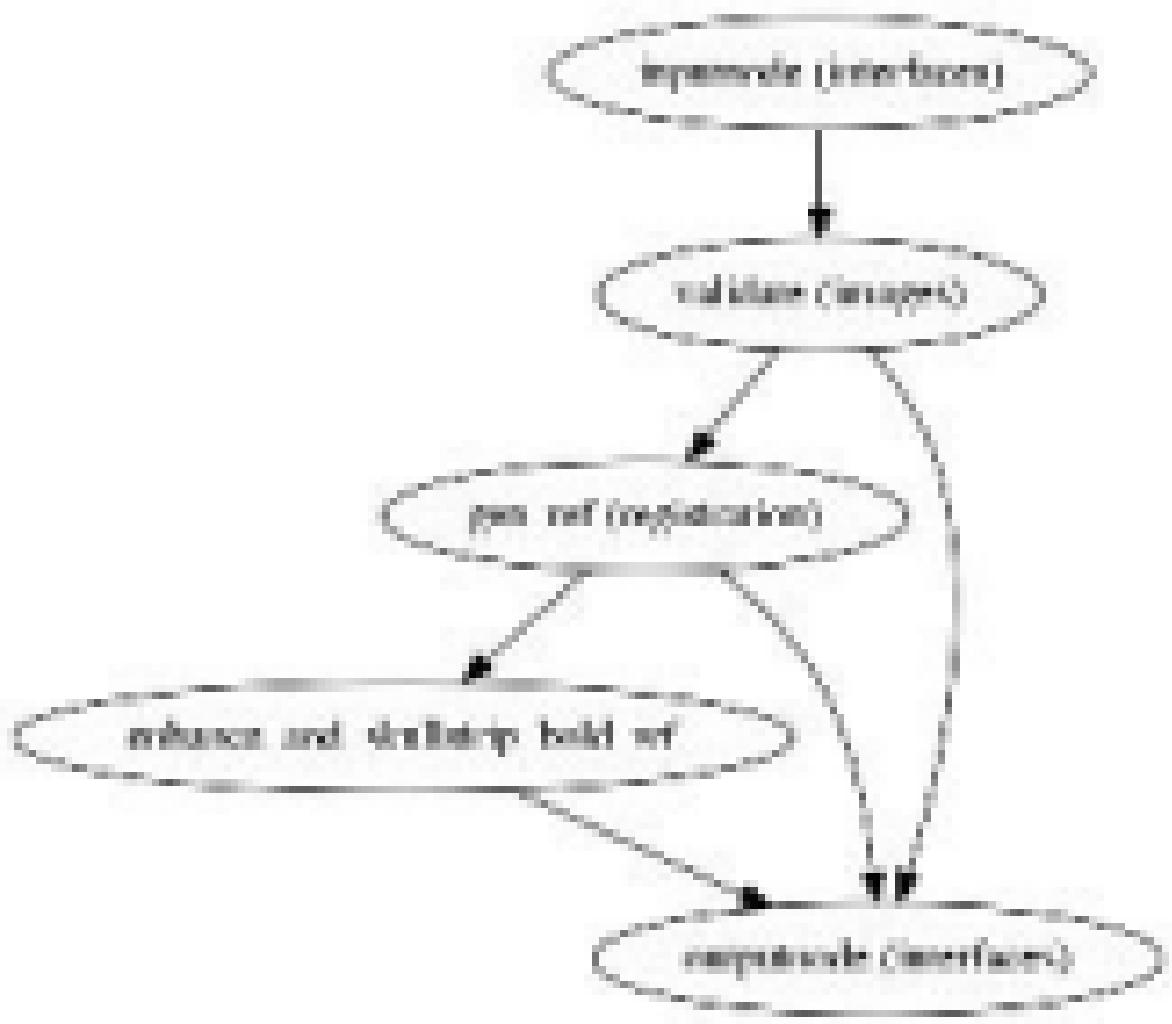
fMRIPrep

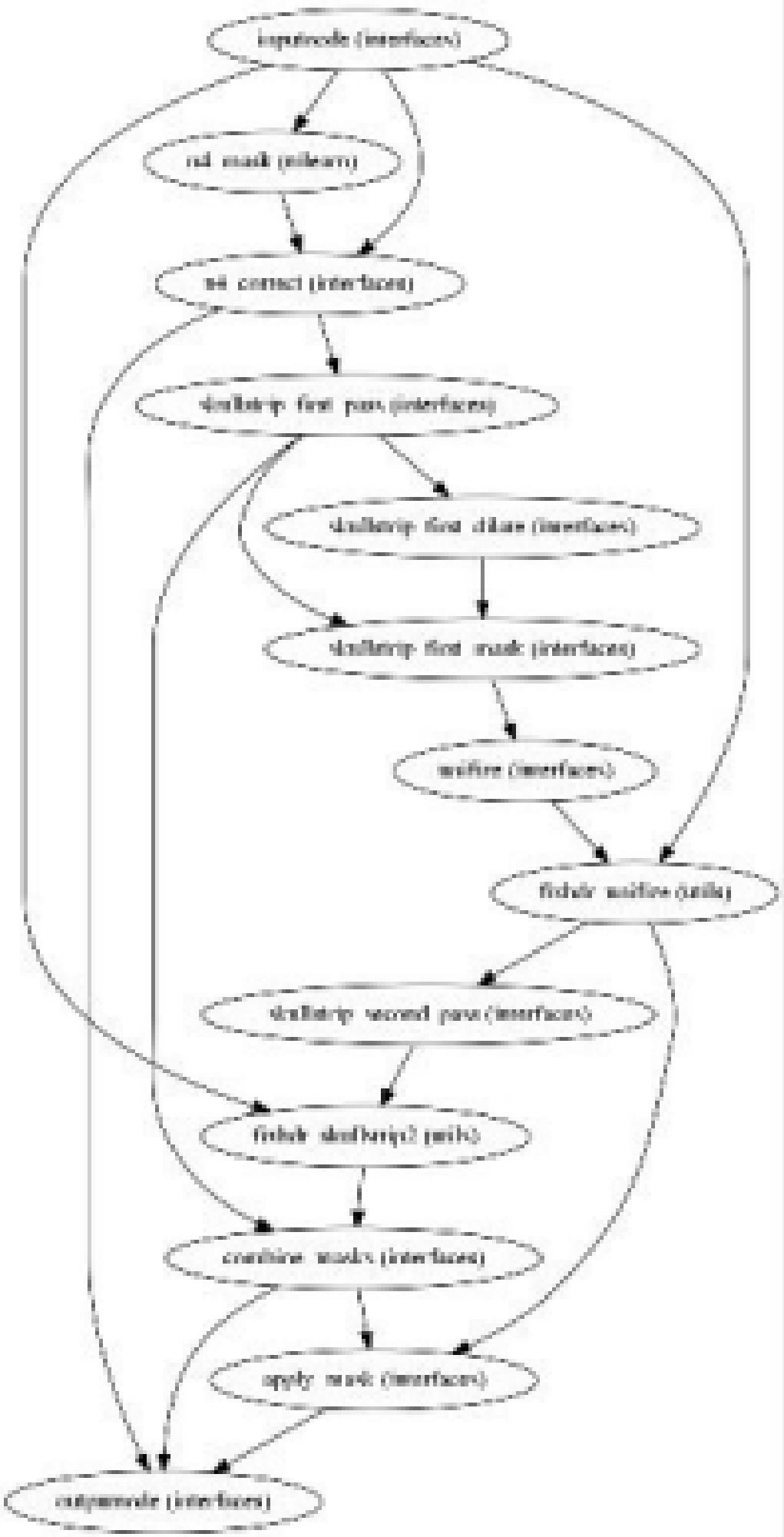


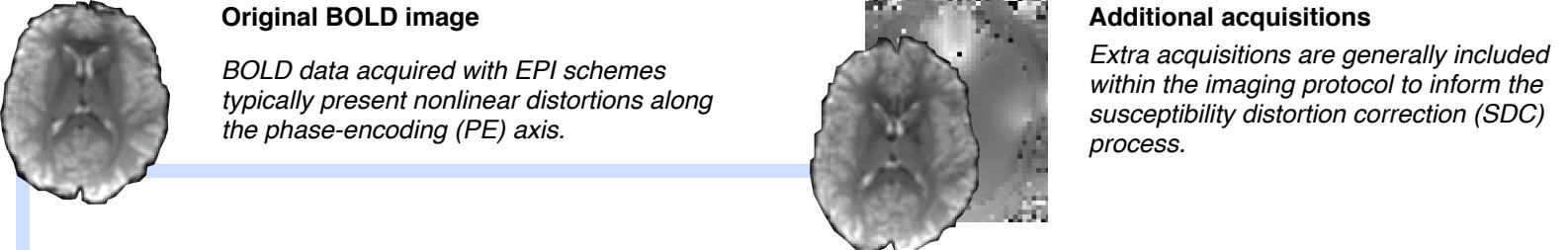
feat







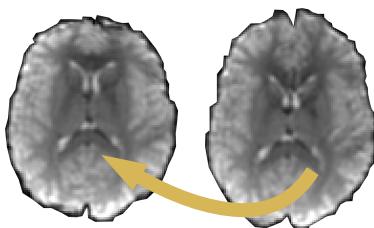




Additional acquisitions with alternative PE directions available? → Field-map acquisition available? → "Fieldmap-less" correction enabled?

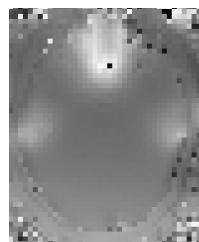
"PE-Polar" SDC

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



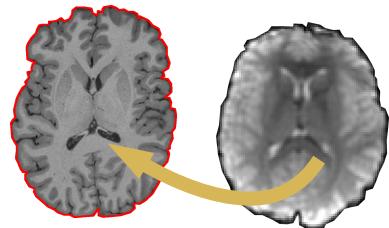
Direct field map estimation

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.



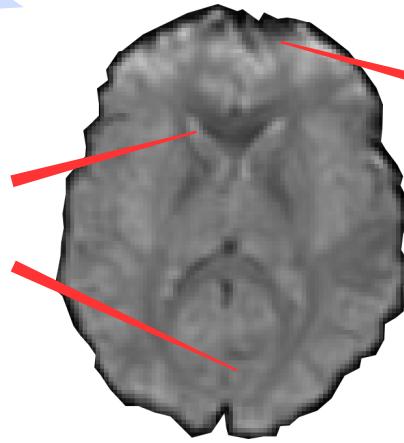
"Fieldmap-less" SDC

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.

```
derivatives/  
  fmrprep/  
    sub-E82/  
    sub-E83/
```

One folder containing derivatives per subject

```
  anat/
```

-  sub-E83_T1w_brainmask.nii.gz
-  sub-E83_T1w_class-CSF_probtissue.nii.gz
-  sub-E83_T1w_class-GM_probtissue.nii.gz
-  sub-E83_T1w_class-WM_probtissue.nii.gz
-  sub-E83_T1w_dtissue.nii.gz
-  sub-E83_T1w_preproc.nii.gz
- ...
-  sub-E83_T1w_space-MNI152NLin2009cAsym_preproc.nii.gz

First derivatives of the T1-weighted image:
Include the brain mask, tissue probability
maps for CSF, GM and WM, a brain tissue
segmentation and the T1w preprocessed
(includes merging T1w images from sessions,
INU correction, etc).

T1w derivatives in additional target
spaces (e.g. MNI152NLin2009cAsym)

```
  func/
```

-  sub-E83_task-machinegame_run-01_bold_confoundtsv
-  sub-E83_task-machinegame_run-01_bold_space-MNI152NLin2009cAsym_brainmask.nii.gz
-  sub-E83_task-machinegame_run-01_bold_space-MNI152NLin2009cAsym_preproc.nii.gz
-  sub-E83_task-machinegame_run-02_bold_confoundtsv
-  sub-E83_task-machinegame_run-02_bold_space-MNI152NLin2009cAsym_brainmask.nii.gz
-  sub-E83_task-machinegame_run-02_bold_space-MNI152NLin2009cAsym_preproc.nii.gz

BOLD derivatives:
for task "Machine
Game" and run 01

BOLD derivatives:
for task "Machine
Game" and run 02

```
  sub-E82.html  
  sub-E83.html
```

One visual report per subject