

# The NiPreps ecosystem for reproducible neuroimaging

Russ Poldrack  
Stanford University

# The standard design (anti-)pattern for fMRI preprocessing

- Pick a single software package
  - Usually based on considerations other than performance



# The standard design (anti-)pattern for fMRI preprocessing

- Pick a single software package
  - Usually based on considerations other than performance
- String together the tools from that package into a script to run the preprocessing workflow

```
#!/bin/bash

# Generate the subject list to make modifying this script
# to run just a subset of subjects easier.

for id in `seq -w 1 26` ; do
    subj="sub-$id"
    echo "====> Starting processing of $subj"
    echo
    cd $subj

    # If the brain mask doesn't exist, create it
    if [ ! -f anat/${subj}_T1w_brain_f02.nii.gz ] ; then
        echo "Skull-stripped brain not found, using bet with a fractional intensity thresh
# Note: This fractional intensity appears to work well for most of the subjects in
# Flanker dataset. You may want to change it if you modify this script for your own
bet2 anat/${subj}_T1w.nii.gz \
anat/${subj}_T1w_brain_f02.nii.gz -f 0.2
    fi

    # Copy the design files into the subject directory, and then
    # change "sub-08" to the current subject number
    cp ..//design_run1.fsf .
    cp ..//design_run2.fsf .

    # Note that we are using the | character to delimit the patterns
    # instead of the usual / character because there are / characters
    # in the pattern.
    sed -i '' "s|sub-08|${subj}|g" \
        design_run1.fsf
    sed -i '' "s|sub-08|${subj}|g" \
        design_run2.fsf

    # Now everything is set up to run feat
    echo "====> Starting feat for run 1"
    feat design_run1.fsf
    echo "====> Starting feat for run 2"
    feat design_run2.fsf
    echo

    # Go back to the directory containing all of the subjects, and repeat the loop
    cd ..
done
```

# The standard design (anti-)pattern for fMRI preprocessing

- Pick a single software package
  - Usually based on considerations other than performance
- String together the tools from that package into a script to run the preprocessing workflow
  - Generally written by a student or postdoc with little software engineering experience
- Hope it keeps working over time...
- Trust that it does the right thing...

```
#!/bin/bash

# Generate the subject list to make modifying this script
# to run just a subset of subjects easier.

for id in `seq -w 1 26` ; do
    subj="sub-$id"
    echo "====> Starting processing of $subj"
    echo
    cd $subj

    # If the brain mask doesn't exist, create it
    if [ ! -f anat/${subj}_T1w_brain_f02.nii.gz ] ; then
        echo "Skull-stripped brain not found, using bet with a fractional intensity thresh
# Note: This fractional intensity appears to work well for most of the subjects in
# Flanker dataset. You may want to change it if you modify this script for your own
bet2 anat/${subj}_T1w.nii.gz \
anat/${subj}_T1w_brain_f02.nii.gz -f 0.2
    fi

    # Copy the design files into the subject directory, and then
    # change "sub-08" to the current subject number
    cp ..//design_run1.fsf .
    cp ..//design_run2.fsf .

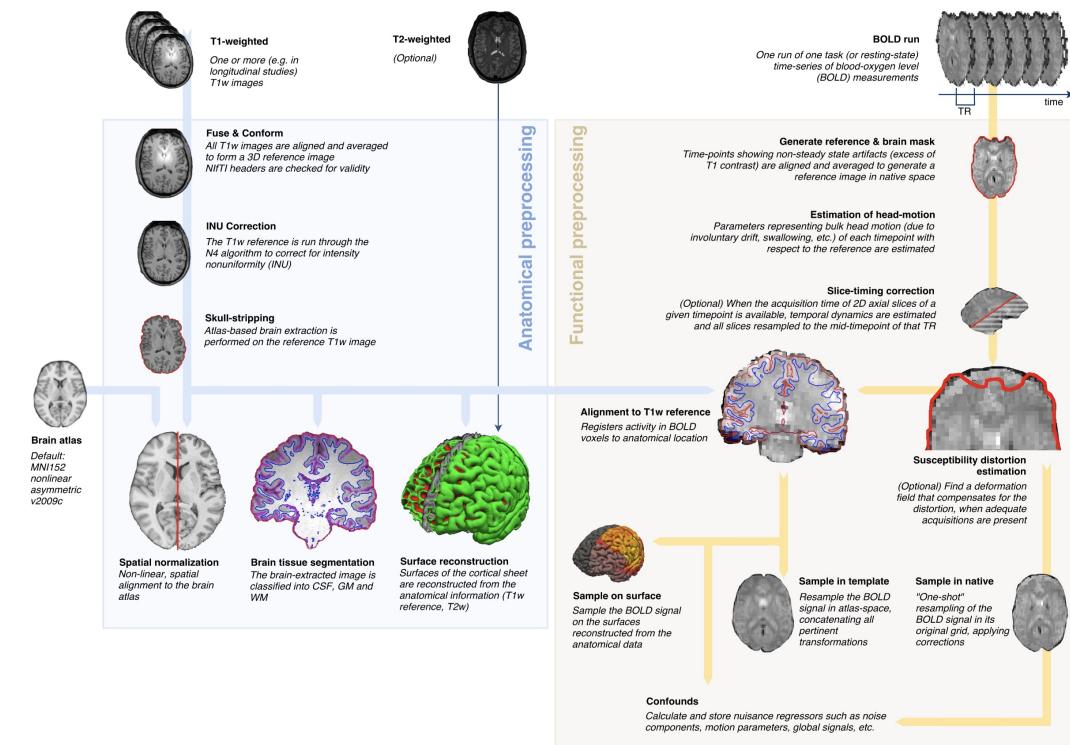
    # Note that we are using the | character to delimit the patterns
    # instead of the usual / character because there are / characters
    # in the pattern.
    sed -i '' "s|sub-08|${subj}|g" \
    design_run1.fsf
    sed -i '' "s|sub-08|${subj}|g" \
    design_run2.fsf

    # Now everything is set up to run feat
    echo "====> Starting feat for run 1"
    feat design_run1.fsf
    echo "====> Starting feat for run 2"
    feat design_run2.fsf
    echo

    # Go back to the directory containing all of the subjects, and repeat the loop
    cd ..
done
```

# Towards a new design pattern for preprocessing: fMRIPrep

- A robust workflow for preprocessing fMRI data
  - Tested on a large number of fMRI datasets from OpenNeuro
  - Adapt to each dataset and processes it in the best way possible
  - Outputs to BIDS-Derivatives format
  - Provides powerful visualizations :
    - to help identify problems
    - allowing glass-box access to training researchers
  - Currently averaging ~5000 successful runs per week



Esteban et al., 2019, *Nature Methods*

# RF1MH121867: NiPreps (NeuroImaging PREProcessing tools)

*... the overarching **objective** of this project is to develop NiPreps, a software framework to perform standardized preprocessing of diverse neuroimaging data.*

- Aim 1: solidify the foundations of the NiPreps integration.
- Aim 2: enable integrative analysis approaches of heterogeneous data.
- Aim 3: accelerate the dissemination of NiPreps to the neuroscience community through hackathons and “docusprints”.

## ⓘ Why?

In developing fMRIprep, we learned critical aspects of the fMRI neuroimaging workflows. NiPreps is envisioned as a generalization of fMRIprep.

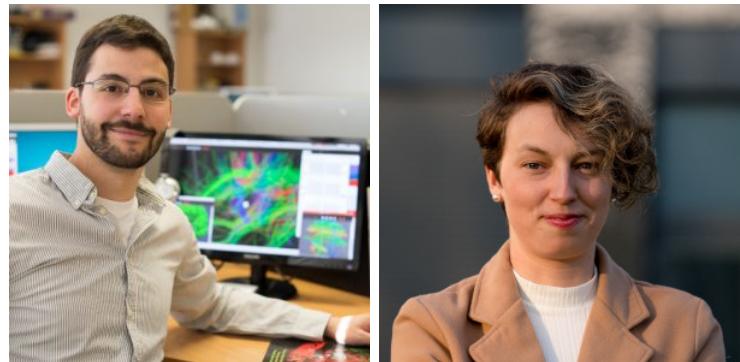
# RF1MH121867: Sites



Poldracklab (Stanford)



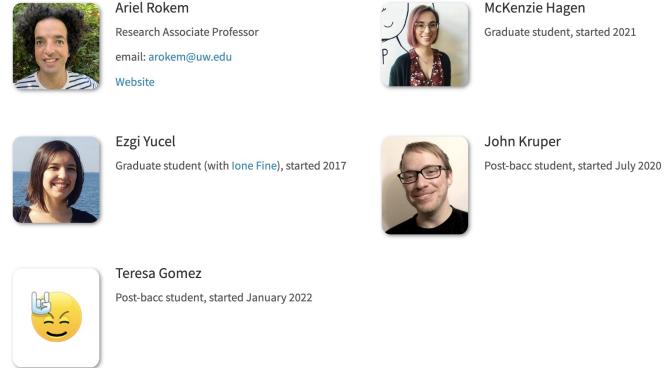
Satterthwaite Lab (Penn)



Esteban Lab (CHUV Lausanne)

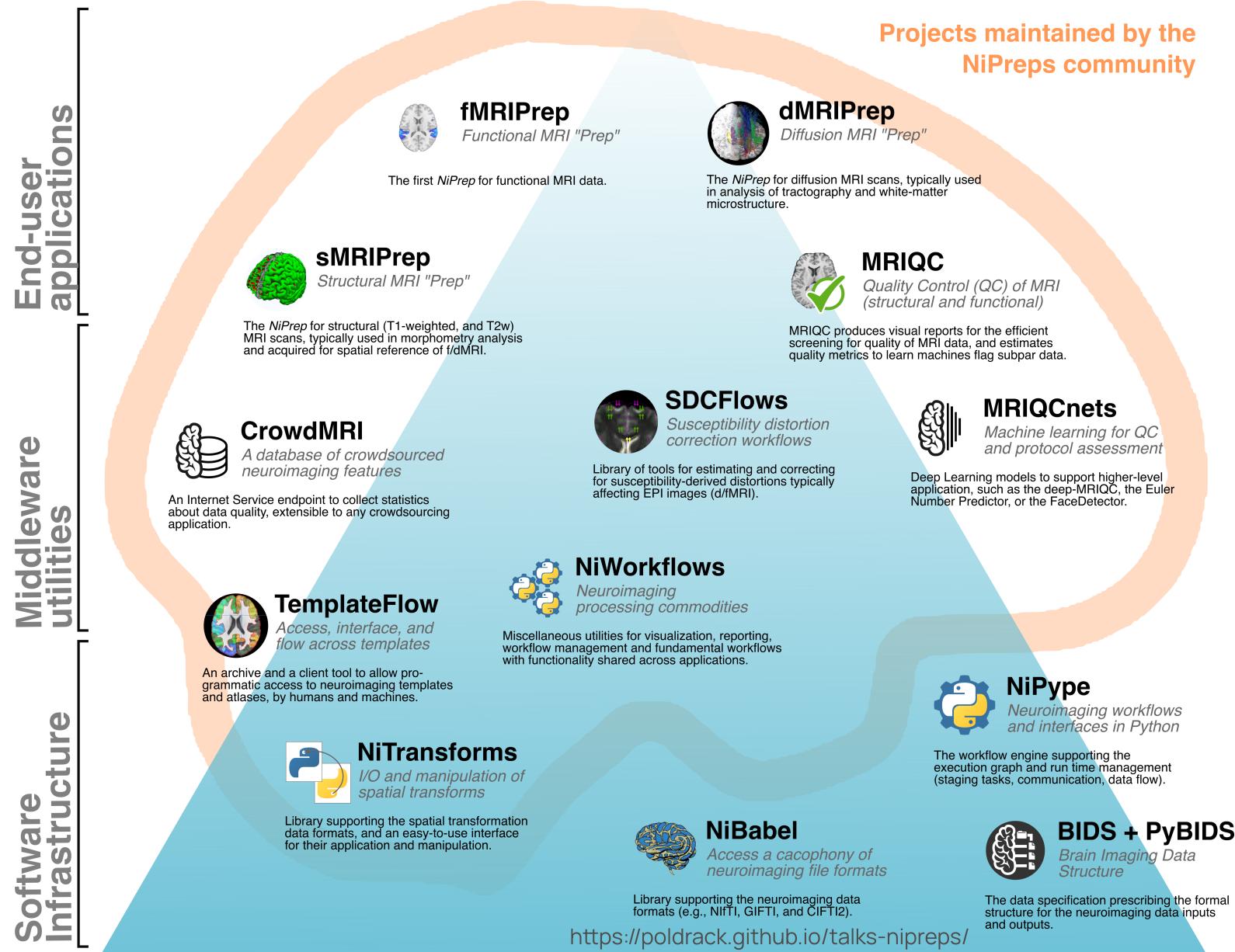


Milham Lab (Child Mind Inst)



Rokem Lab (UW)

# NiPreps: deconstructing fMRIPrep in reusable modules



# Nurturing a new community

## Welcome

NiPreps

Home

NiPreps

BIDS-Apps

Community

Welcome

Membership

New features

Contributing

Code of Conduct

Licensing

Developers

### Current members of the GitHub organization

- > A total of 90 neuroimagers have already joined us. Becoming a member will give you access to additional forums for discussion, subscribing notifications for events and meetings, etc. You can request you are added to the organization by [creating a new issue here](#).



[www.nipreps.org](http://www.nipreps.org)

# Aim 1: Solidify the foundations of the NiPreps integration.

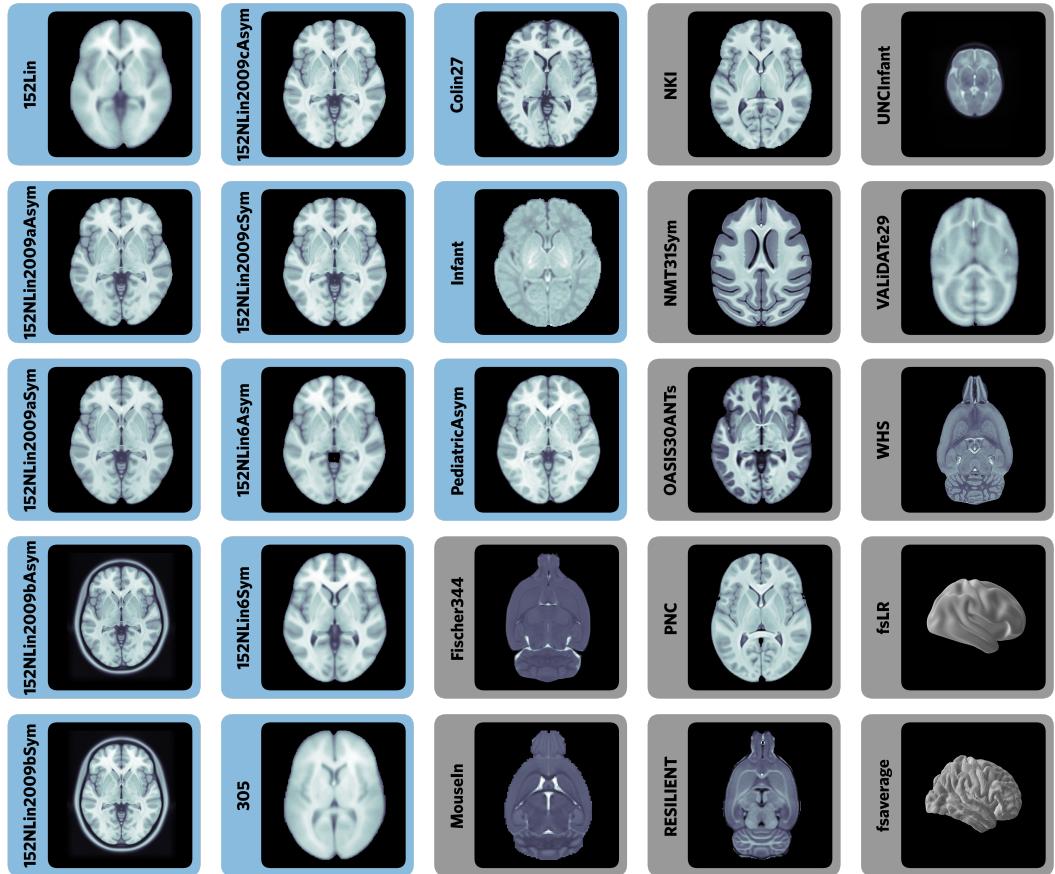
In addition to maintenance across the board and infrastructure (e.g., website), this aim focuses in four specific elements:

- TemplateFlow: FAIR Sharing and management of neuroimaging templates and atlases
- SDCflows: Integrating susceptibility distortion correction (SDC)
- Modularization, and generalization of the Visual Reports System (NiReports)
- Integrating preclinical imaging into NiPreps (NiRodents)
- Spatial transforms integration (NiTransforms)

# Result YR1: TemplateFlow is mature

- Templates and atlases are commonly used in neuroimaging research
- There is significant lack of clarity in the use of these templates
  - There are numerous versions of the widely used “MNI template”
- Templateflow provides programmatic access to a database of templates and mappings between them
- Easy to use for humans and machines:

```
>>> from templateflow import api as tf  
>>> tf.get("MNI152NLin2009cAsym",  
...           suffix="T1w")
```



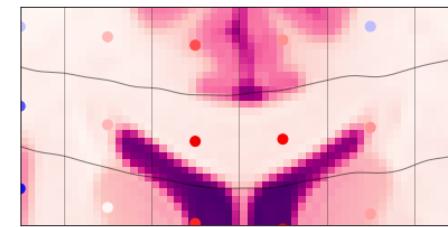
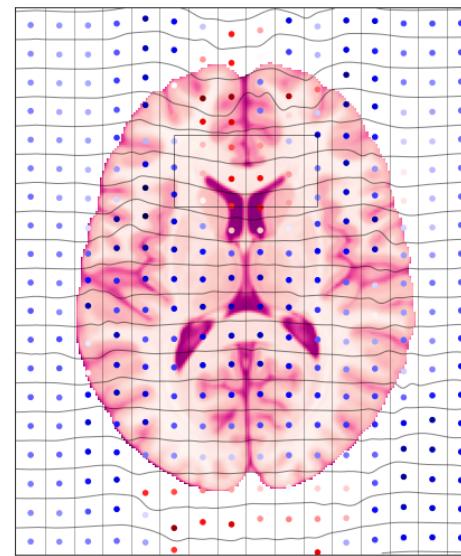
Ciric et al., 2022, *Nature Methods*

# Result YR1: SDCFlows overhaul

## SDC is a significant source of analytical flexibility in functional and diffusion MRI processing

- Multiple fieldmap sequences (spin echo, gradient echo, scanner-reconstructed  $B_0$  maps), and combinations thereof
- Multiple correction implementations with different data structures

- Overhaul started early 2021 ([Esteban et al., OHBM 2021](#))
- Faced many technical challenges → numerous bugfixes
  - vast amount of “edge” cases
- Opacity of the problem itself → educational materials
  - Jupyter notebooks

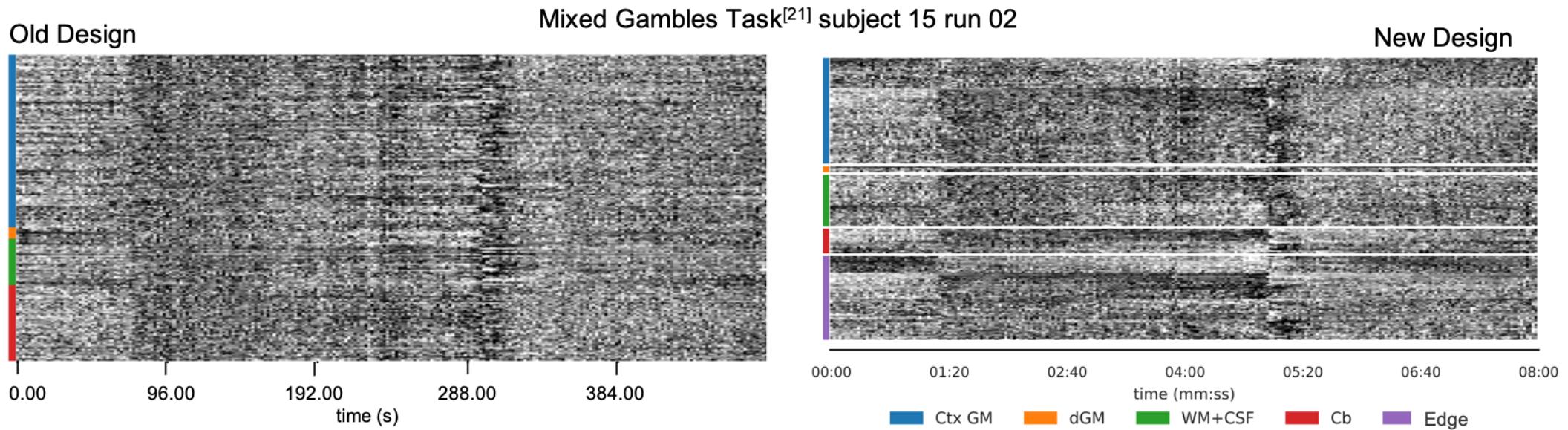


Jupyter Notebook  
on GitHub:  
[nipreps/sdcflows](#)

# SDCFlows in a nutshell

- **What?** Wraps or implements methods for **estimating** field maps in various scenarios:
  - covering (i) so-called *TOPUP*, (ii) phase-difference fieldmaps and their variants, (iii) *fMRIPrep*'s “fieldmapless”; but
  - it does not cover point-spread-function approaches (extremely marginal).
- **How?**
  - Defines a shared representation model (B-Spline) for the field map (comparability ↑↑, methodological variability ↓↓)
  - Technical perk: “decouples” estimation and application steps (modularity ↑↑, methodological variability ↑).
- **Why?**
  - A single tool can be applied to correct for distortion, no matter how it was estimated (method comparability ↑↑)
  - The model coefficients can easily be moved with head motion (e.g., dynamic fieldmap)
  - The model coefficients can easily be integrated in spatial transformation chains (one-shot interpolation)
  - The model coefficients can easily be integrated within other software (e.g., eddymotion)

# Result YR1: new visualizations in MRIQC and fMRIPrep



(Provins et al., ISMRM 2022)

- Added visualization of voxels at the edge of the brain ("crown")

# Result YR1: NiRodents

- MRIQC-rodents saw a first release presented in EMIM 2022
- fMRIPrep-rodents is currently integrating new SDCFlows' API
- NiRodents has stimulated several improvements of the reporting system (see next)
- NiRodents has stimulated the inclusion (and revision of existing) rodent templates in TemplateFlow

## Extending MRIQC to rodents: image quality metrics for rat MRI

Eilidh MacNicol<sup>1</sup>, McKenzie Hagen<sup>2,3</sup>, Céline Provins<sup>4</sup>, Eugene Kim<sup>1</sup>, Diana Cash<sup>1</sup>, Oscar Esteban<sup>3,4</sup>

Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK

Department of Psychology, University of Washington, WA, USA

Department of Psychology, Stanford University, CA, USA

Department of Radiology, Lausanne University Hospital and University of Lausanne, Switzerland



nipreps/mriqc

<https://poldrack.github.io/talks-nipreps/>

# Results YR1: Other areas

- NiReports: provides standard mechanisms to build “reportlets” and full reports.
  - A repository has been initiated
  - Code already exists, but it is scattered across tools (e.g., MRIQC, fMRIPrep)
  - In the OHBM Hackathon, work initiated for moving MRIQC visual components into it (fMRIPrep will follow).
- NiTransforms:
  - The component is in a stable status
  - The component is key to achieving a redesign of fMRIPrep (described next)

# Infrastructure: Architectural redesign

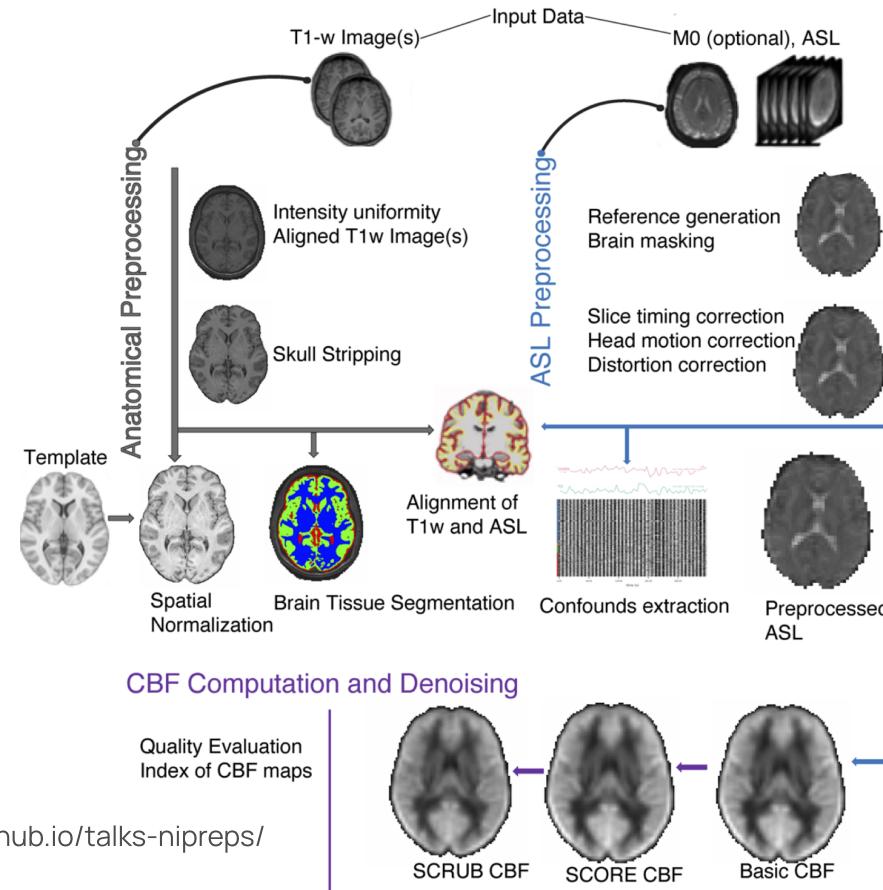
- Problem: fMRIPrep's "one size fits all" design has limitations for emerging use cases
  - Archiving preprocessing results requires balancing storage costs against possible use cases.
  - Including alternative algorithms requires custom code to integrate.
- Solution: Accept pre-computed derivatives and defer computationally cheap operations
  - E.g., Deep learning segmentations and masks can be accepted, skipping fMRIPrep defaults.
  - Multiple template registrations can be archived, analysts may resample BOLD series to different spaces on demand.
- This approach is implemented in SDCFlows and is being generalized to other components.

# Aim 2: Expand the portfolio of end-user NiPreps

- ASLPrep
- dMRIPrep
- PETPrep
- fMRIPrep-infants (based on *NiBabies*)

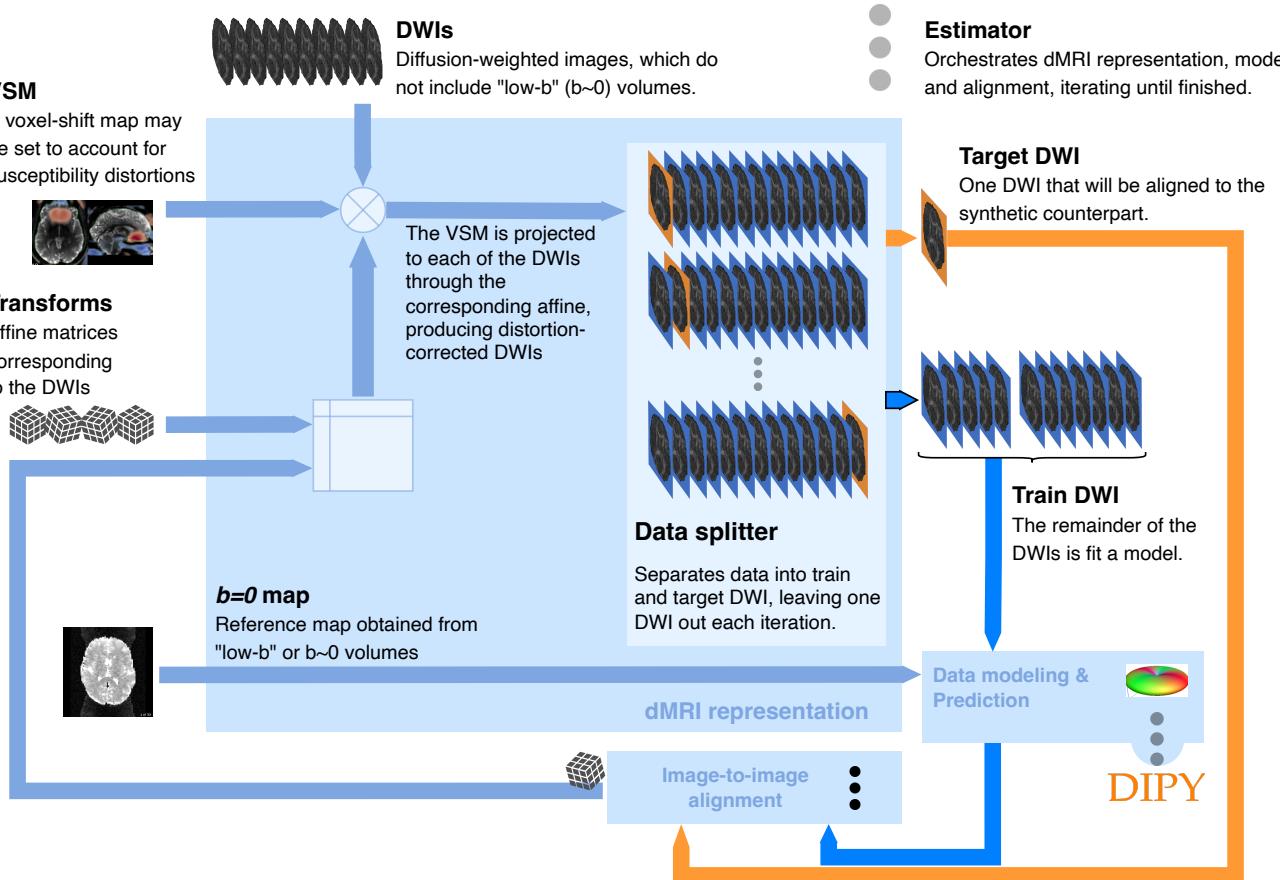
# Workflows: ASLPrep (cerebral blood flow quantification)

- A robust workflow for preprocessing arterial spin labeling (ASL) data
  - Including cerebral blood flow (CBF) quantification
  - Provides quality evaluation for CBF maps
  - Provides CBF quantification at the regional level using atlases



# Workflows: dMRIPrep (diffusion MRI)

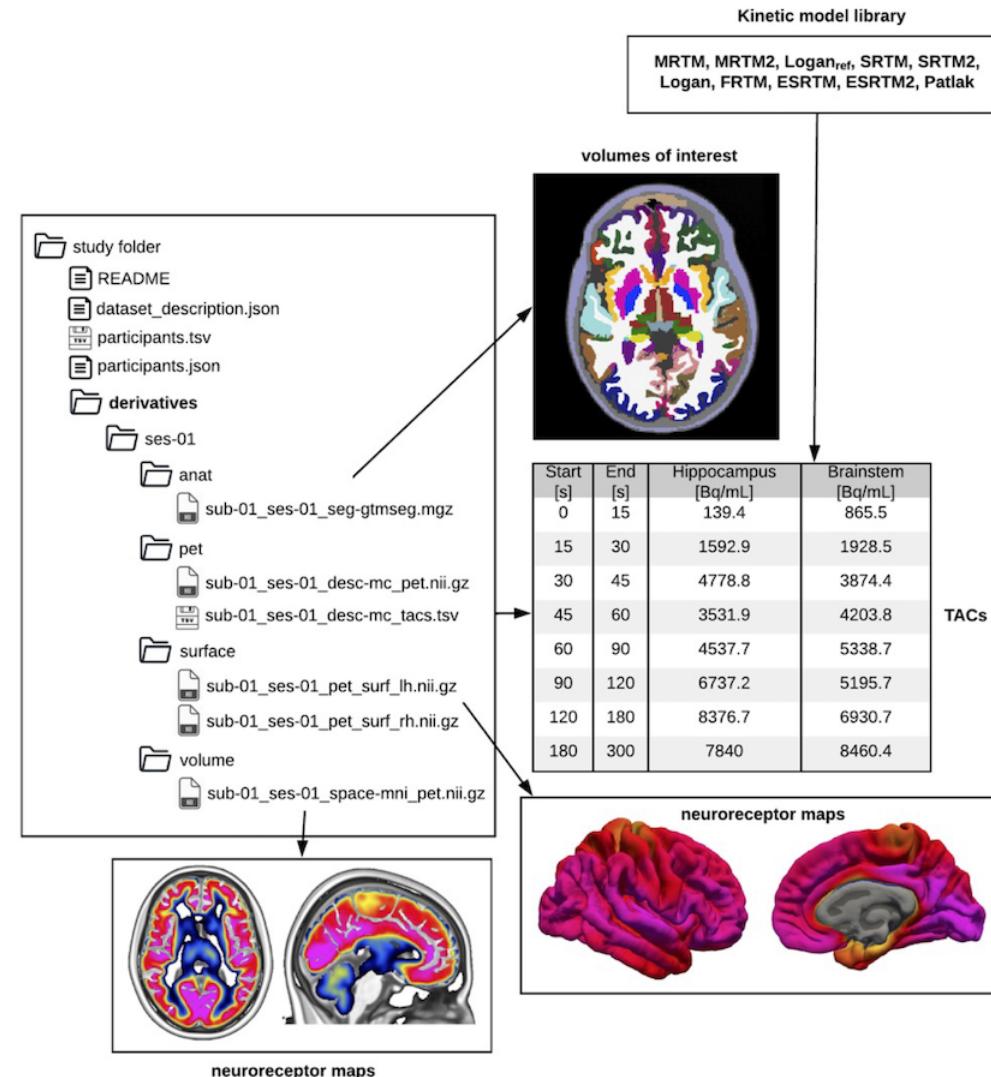
Development directed toward *eddy motion*, an algorithm to estimate head-motion (modality-agnostic) and modality-specific artifacts (eddy currents in the case of dMRI):



Preprint: Pisner et al., 2022.

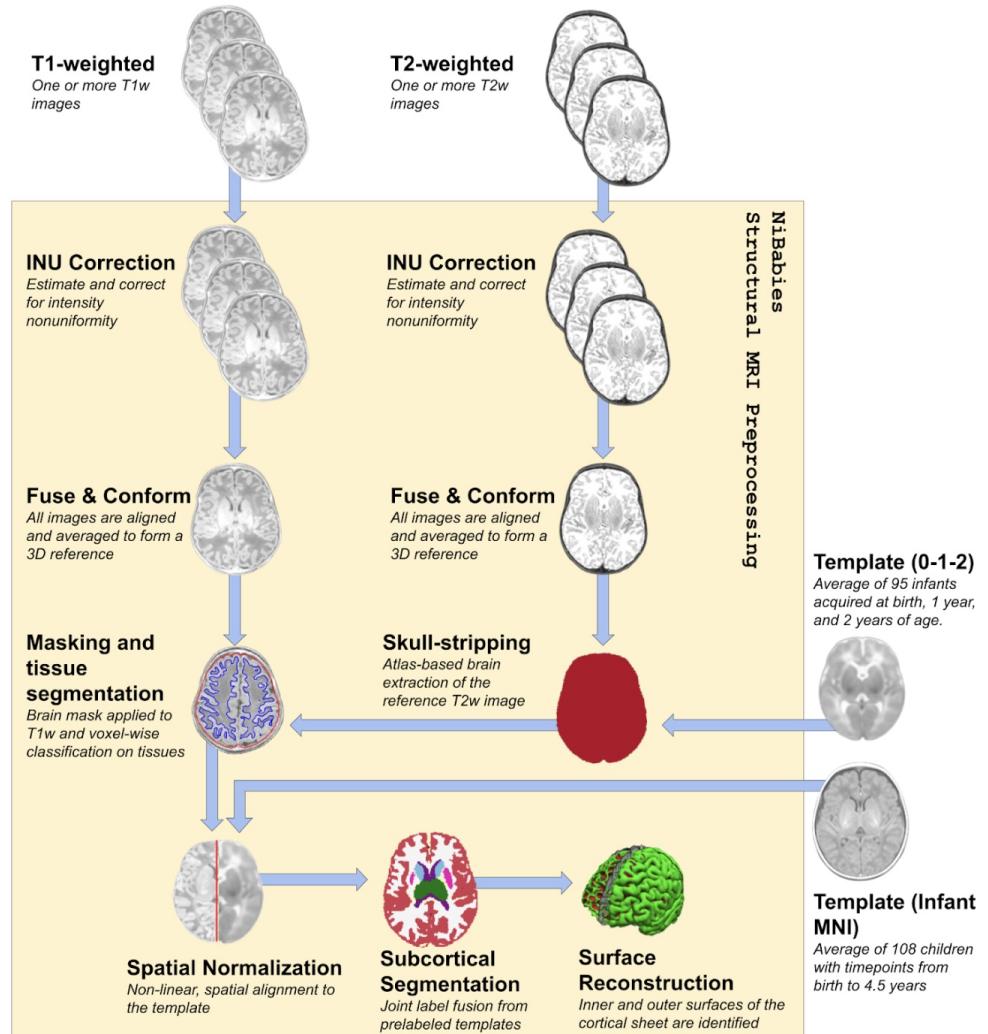
# Workflows: PETPrep (positron emission tomography)

- A NiPreps workflow for PET preprocessing
  - Successfully merged petsurfer into nipype (1.8.0)
  - Incorporated nipype implementation of a robust head motion correction workflow (petprep\_hmc)
  - Developing a BIDS-Derivatives standard for PET derivatives
  - Model-based head-motion correction leveraging *eddy motion* in progress



# Workflows: fMRIprep for infants

- Collaboration with Damien Fair & HBCD team
- **New developments**
  - Support for pre-computed derivatives (mask, segmentations).
  - Improved robustness and validity of CIFTI-2 outputs.
- **Upcoming developments**
  - Morphometric outputs (cortical thickness, curvature)
  - Improvements to susceptibility distortion correction versatility
  - T2 assisted surface generation



# Aim 3: Consolidate the NiPreps community

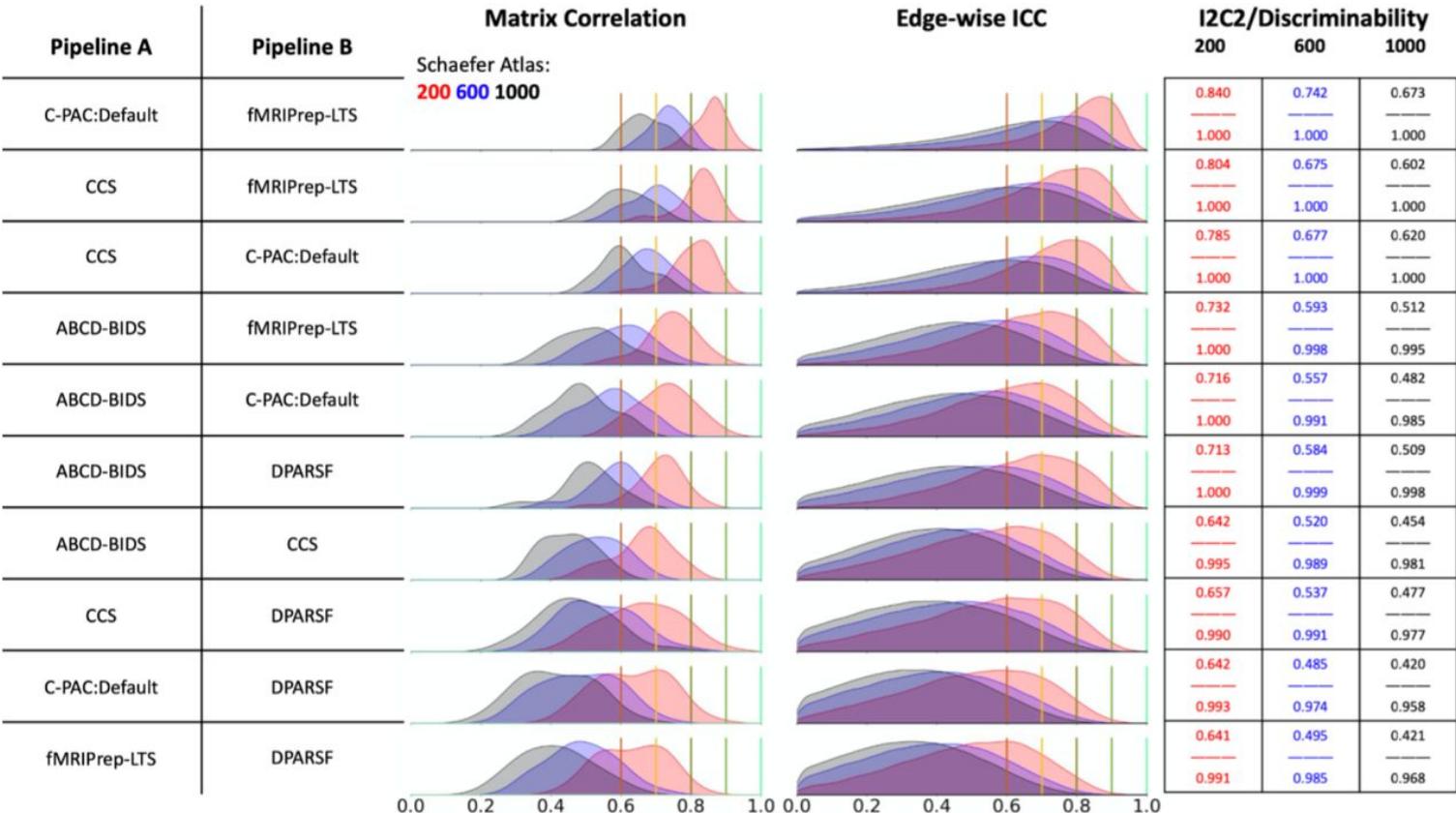
- Project monitoring infrastructure: MIGAS
- Evaluation of cross-workflow reproducibility
- Hackathons and documentation
- Best practices and educational resources

# Project monitoring: MIGAS

- An open-source, customizable telemetry solution
- Allows collecting usage information, errors, and status throughout a process's lifetime
- Easy to deploy with various cloud providers (Heroku / GCP / AWS)
- Available as a Python package: <https://pypi.org/project/migas/>

# Reproducibility: Cross-workflow evaluation

- CMI team developed a CPAC implementation of fMRIPrep
  - Able to achieve high levels of reproducibility in connectivity metrics between harmonized workflows
  - Helped identify causes of divergence, such as use of different versions of MNI template



# Hackathons and documentation

- Held a hackathon/documentation sprint in Glasgow following OHBM 2022
- Will participate in Brainhack Global 2022
- Planning to hold a hackathon/documentation sprint in Montreal in association with OHBM 2023

# Results YR1: Best practices and educational resources

- Collaborative QC-Book (educational, ISMRM 2021): <https://nipreps.org/qc-book>
- MRIQC-SOPs (standard operating procedures)
  - A GitHub template-repository to create and maintain versioned SOPs documentation and checklists.
  - Example: <https://nipreps.org/mriqc-sops/>
- MRIQC Protocol report (Hagen et al., in preparation)
- Frontiers' research topic on QC of fMRI (Provins et al., under review)
- Biases introduced by defacing in QC (Provins et al., pre-registered report under review)
- MRIQC extension to dMRI - under development

# Thank you!

# fMRIPrep usage

- Usage tracked using an opt-out telemetry system
  - Allows quick identification of bugs and usage patterns
- Currently averaging ~5000 successful runs per week

