



THE  
**NYSPI MRI**  
RESEARCH PROGRAM



XNAT



CLUSTER

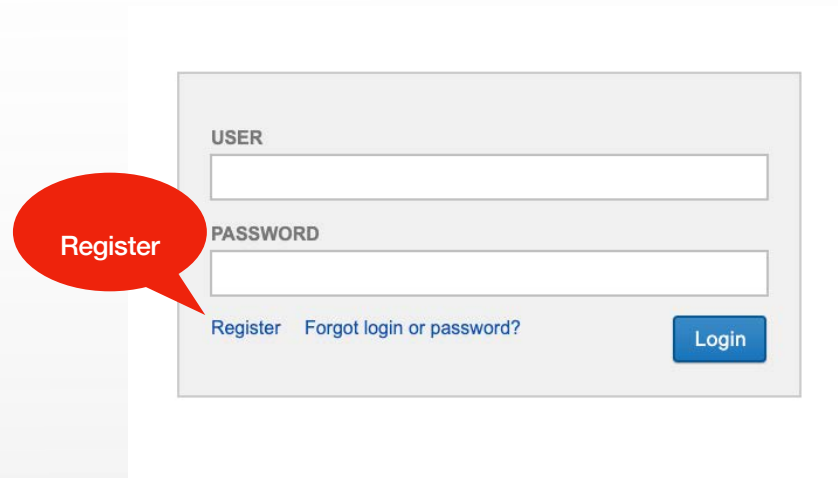


AWS

The MRI Research program provides NYSPI users the framework to pre-process supported fMRI, DTI, and anatomical data. This is a tutorial on how to access and utilize the resources.

- XNAT stores the raw DICOM images acquired during a scan. XNAT can also convert raw MRI data into BIDS format and provide mriqc outputs through the XNAT webpage.
- The cluster can be accessed through a terminal and provides a program which pulls all BIDS data and supported non-BIDS data from XNAT and validates those data for processing pipelines.
- AWS is accessed through a website that launches the available processing pipelines to process the validated BIDS data.

## XNAT account creation



The image shows a web form for XNAT account creation. It has two input fields: 'USER' and 'PASSWORD'. Below the 'PASSWORD' field, there are two links: 'Register' and 'Forgot login or password?'. A blue 'Login' button is located to the right of these links. A red speech bubble with the word 'Register' inside points to the 'Register' link.

- XNAT accounts are created by visiting <https://xnat.nyspi.org/> and registering for an account.
- For access to a protocol to be granted, the user must email the PI of the study, cc'ing [Matthew.Riddle@nyspi.columbia.edu](mailto:Matthew.Riddle@nyspi.columbia.edu) requesting access to the protocol on XNAT.

# Taking Notes at Scan Time

```
EXAM #: 20418
MRI TECH: Frank
RA: Nicole
DATE: 7/20/21
SCAN DAY: GBU
1 17:32 loc ssfse
2 17:33 T1 -unusable, motion
3 17:38 T1
4 17:44 T2
5 17:45 HOS
6 17:46 topup fpe
7 17:47 topup rpe
8 17:48 RS1
9 17:56 RS2
10 18:13 RS3 -unusable, scan stopped
11 18:15 RS3
12 18:24 audiotest
13 18:39 topupe fpe
14 18:41 topup rpe
15 18:44 piemanET -unusable, eye tracking did not work
```

Unusable

Unusable

Unusable

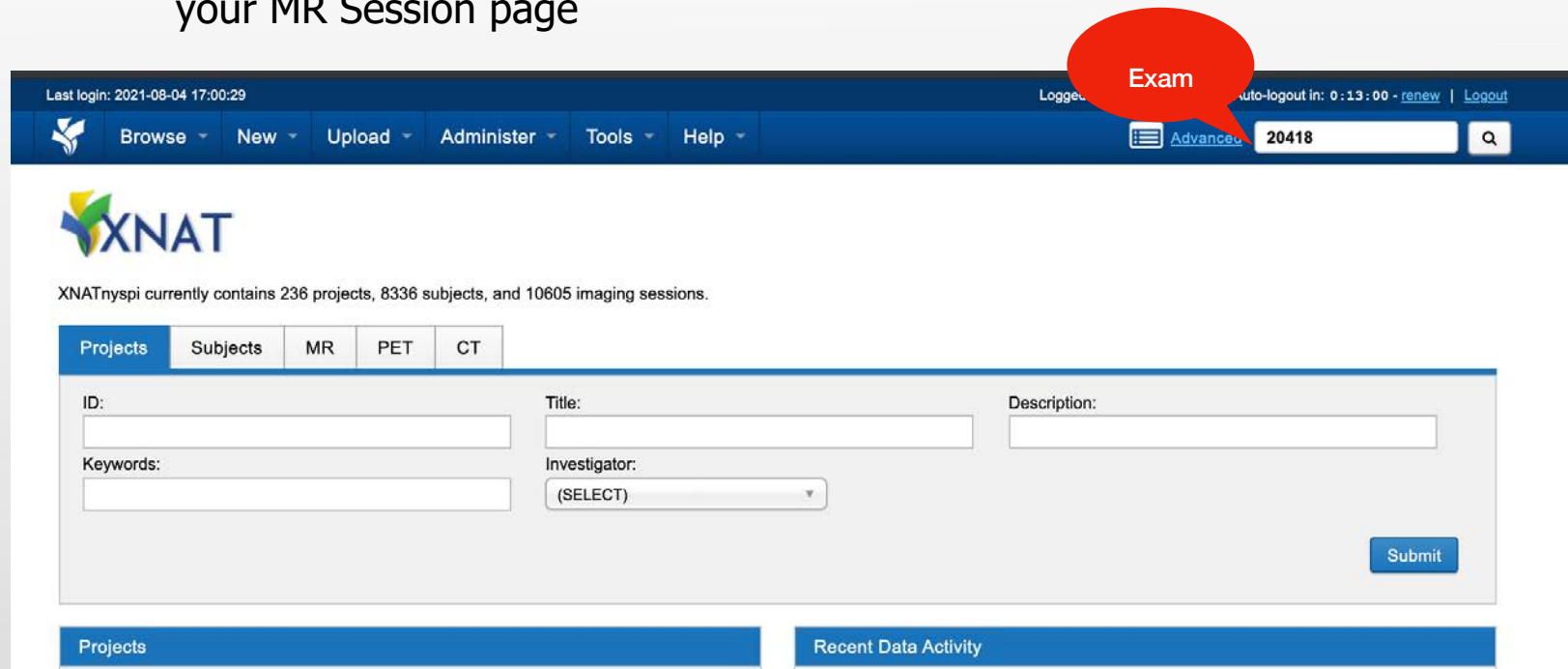
- It is crucial to take notes during the acquisition of MR data.
- Each MR scan has an associated and unique exam number that is created by the MR console.

These notes should include:

- date
- exam number in addition each series should be noted. In the above example, Series 2, Series 10, and Series 15 are unusable. If the data going into processing pipelines is bad it could lead to errors in processing or bad data coming out.

# XNAT Interface

- Once you have your notes and your exam number, we can now find and update your exam on XNAT. There are many ways to navigate on the XNAT website. The easiest way to find your data is by using your exam number.
- Data is automatically transferred to XNAT from the MRI console within 2 days of acquisition.
- Enter the exam number into the search bar on XNAT to go to your MR Session page



The screenshot displays the XNAT web interface. At the top, a dark blue navigation bar contains the XNAT logo, a menu with 'Browse', 'New', 'Upload', 'Administer', 'Tools', and 'Help', and a search bar. The search bar has a red speech bubble with the word 'Exam' pointing to it and contains the text '20418'. To the right of the search bar are links for 'Advanced', 'Logout', and 'auto-logout in: 0:13:00 - renew'. Below the navigation bar, the main content area features the XNAT logo and a status message: 'XNATnyspi currently contains 236 projects, 8336 subjects, and 10605 imaging sessions.' Below this is a tabbed interface with 'Projects', 'Subjects', 'MR', 'PET', and 'CT' tabs. The 'Projects' tab is active, showing a form with fields for 'ID:', 'Title:', 'Description:', 'Keywords:', and 'Investigator:' (with a dropdown menu). A 'Submit' button is located at the bottom right of the form. At the bottom of the page, there are two blue bars: 'Projects' and 'Recent Data Activity'.

Last login: 2021-08-04 17:00:29

Logged in: auto-logout in: 0:13:00 - [renew](#) | [Logout](#)

[Browse](#) [New](#) [Upload](#) [Administer](#) [Tools](#) [Help](#) [Advanced](#)  [Q](#)

**XNAT**

XNATnyspi currently contains 236 projects, 8336 subjects, and 10605 imaging sessions.

[Projects](#) [Subjects](#) [MR](#) [PET](#) [CT](#)

ID:

Title:

Description:

Keywords:

Investigator:

[Submit](#)

[Projects](#) [Recent Data Activity](#)

# XNAT Session Page

Before using the XNAT interface to qc download or proceed with data flow, it is important to familiarize yourself with the following XNAT terms.

- XNAT ID
  - Provided to MRI operator via Metal Screener.
  - Must be BIDS formatted (no special characters or spaces)
- XNAT PROJECT ID
  - Shorthand project ID automatically assigned in XNAT
- EXAM number
  - Unique session ID automatically assigned by scanner
- ACCESSION ID
  - Unique session ID automatically assigned by scanner
- SERIES / SERIES DESCRIPTION
  - Each exam is comprised of several Series, determined in the protocol for each project. Each Series has a Series Description assigned by Feng.
- FILES / SIZE
  - The number of files and their size should be the same for each series type.

PROJECT ID: [pateprfsz](#) > SUBJECT: [prfsz080002](#) > [20418](#)

MR Session: 20418

Details	Projects	Exams	Series	Scans	Actions
<b>Accession #:</b> XNATnyspi19_E00856 <b>Date Added:</b> 2021-07-20 15:42:46 (admin) 2021-07-20 13:18:49 GEOCGEOC GE MEDICAL SYSTEMS SIGNA Premier Site: New York State Psychiatric Inst					<b>Actions</b> Edit View Download Email Processing Manage Files View Images Delete Run Containers

**Scans**

Bulk Actions: Download

Scan	Type	Series Desc	Usability	Files	Note
1	LOC_3-P_FGRE	LOC_3-P_FGRE	usable	3.8 MB in 27 files	
2	STRUC_T1MPRAGE_n oPROMO	STRUC_T1MPRAGE_n oPROMO	usable	68.0 MB in 222 files	
3	STRUC_T1MPRAGE_n oPROMO	STRUC_T1MPRAGE_n oPROMO	usable	67.8 MB in 222 files	
4	STRUC_T2CUBE_PRO MO	STRUC_T2CUBE_PRO MO	usable	58.2 MB in 218 files	
5	HOS_WB_HRBRAIN	HOS_WB_HRBRAIN	usable	3.2 MB in 128 files	
6	FUNC_TOPUP_FPE	FUNC_TOPUP_FPE	usable	7.2 MB in 134 files	
7	FUNC_TOPUP_RPE	FUNC_TOPUP_RPE	usable	7.2 MB in 134 files	
8	FUNC_RS_1_MUX6	FUNC_RS_1_MUX6	usable	1.3 GB in 24224 files	
9	FUNC_RS_2_MUX6	FUNC_RS_2_MUX6	usable	1.3 GB in 24224 files	

# XNAT Series Usability

- The first step in managing your data on XNAT is to indicate which Series, if any, are unusable so that they may be excluded from data flow.

- IMPORTANT:** Verify that your subject ID is COMPLETELY accurate in XNAT and that your exam is in the correct XNAT project. DO NOT continue if either of these are incorrect, contact Matt Riddle:

[Matthew.Riddle@nyspi.columbia.edu](mailto:Matthew.Riddle@nyspi.columbia.edu)

[PROJECT: catniprfz](#) > [SUBJECT: prfsz080002](#) > **20418**

MR Session: 20418

Details	Projects	Synchronization	Audit Trail
<b>Accession #:</b> XNATnyspi19_E00856 <b>Date Added:</b> 2021-07-20 15:42:46 (admin) <b>Date:</b> 2021-07-20 <b>Time:</b> 13:18:49 <b>Scanner Name:</b> GEOGEOC <b>Scanner Type:</b> GE MEDICAL SYSTEMS SIGNA Premier <b>Acquisition Site:</b> New York State Psychiatric Inst			
<b>Subject:</b> <a href="#">prfsz080002</a> <b>Gender:</b> <b>Handedness:</b> <b>Age:</b> --		<b>Actions</b> Edit View Download Email Processing Manage Files View Images Delete Run Containers	

- From The MR Session page select the Edit tab.
- For example, from the Scan notes, it is clear that Series 15 needs to be marked as unusable
- Select unusable from the dropdown menu for Series 15 and fill in any notes.

13	18:39	topup fpe
14	18:41	topup rpe
15	18:44	piemanET -unusable, eye tracking did not work

13	FUNC_TOPUP_FPE	usable	7.2 MB in 134 files
14	FUNC_TOPUP_RPE	usable	7.2 MB in 134 files
15	FUNC_piemanET_MUX6	unusable	eye tracker did not work 213.1 MB in 5346 files
40002	ORIG_STRUC_T1MPRAGE	usable	41.4 MB in 220 files
40003	ORIG_STRUC_T1MPRAGE	usable	41.4 MB in 220 files
40004	ORIG_STRUC_T2CUBE_PR	usable	40.6 MB in 216 files

Total: 5.9 GB in 108673 files

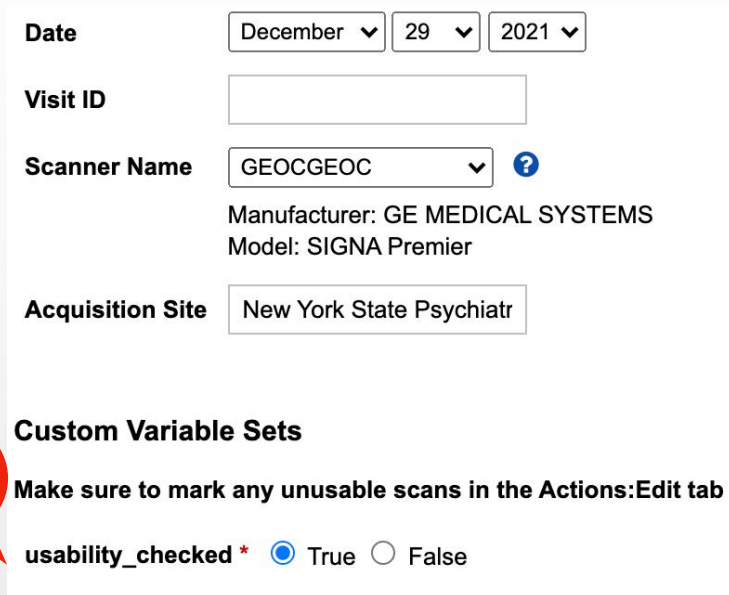
Additional Notes

Submit

Back Submit

## XNAT Series Usability cont.

- **IMPORTANT:** Once you are done confirming the usability of each series and marking any unusable series as “unusable”, you need to set the “usability\_checked” variable to “True” for the exam session.
- This step **MUST** be done whether the exam is all usable series or has any unusable series.
- Marking this variable “True” will flag the exam as being ready for further processing in the supported pipelines.
- To do this, find the radio button for the “usability\_checked” variable on the Edit page and be sure it is set to True, and finally hit Submit at the bottom of the page.



The screenshot shows a web form for XNAT Series Usability. It includes fields for Date (December 29, 2021), Visit ID, Scanner Name (GEOCGEOC), Acquisition Site (New York State Psychiatr), and a section for Custom Variable Sets. A red callout bubble with the text 'set to True' points to the 'usability\_checked' radio button, which is currently selected.

**Date** December ▼ 29 ▼ 2021 ▼

**Visit ID**

**Scanner Name** GEOCGEOC ▼ ?  
Manufacturer: GE MEDICAL SYSTEMS  
Model: SIGNA Premier

**Acquisition Site** New York State Psychiatr

**Custom Variable Sets**  
Make sure to mark any unusable scans in the Actions:Edit tab

usability\_checked \* ☒ True ☐ False



## XNAT Series Usability cont.

In addition to the hardware failure indicated in the previous example of usability, there are other factor which may cause a series to be deemed unusable.

1. If the subject had to go to the bathroom and interrupted a Series
2. If the MR tech had to restart a Series (they should always indicate this)
3. If the Head Movement was severe
4. If the reconstruction of the Series failed (Hardware)

### TIPS!

- It is important to check that the size of the series is consistent and accurate. The same Series will usually have the same number of files and total size from run to run, as seen on the right of the XNAT session page.
- It is often helpful if the person who ran the scan and collected the MRI data also does the usability check on XNAT!
- Being mindful of recording these Series errors is crucial for avoiding problems with data flow.

## XNAT BIDS and mriqc

The XNAT platform has the ability to produce BIDS and run mriqc.

- This must be requested for each protocol by sending an email to [Juan.Sanchez@nyspi.columbia.edu](mailto:Juan.Sanchez@nyspi.columbia.edu) cc'ing the protocol PI.
- BIDS data are automatically generated from XNAT exam sessions that have had their series Usability Checks completed and the XNAT variable "usability\_checked" marked "True" in the XNAT session page. Data will usually process within a few hours of being marked depending on system utilization.
- Once the BIDS data has been generated, you will be able to run mriqc on the exam session.

# Introduction to mriqc

## Quality control of MRI data

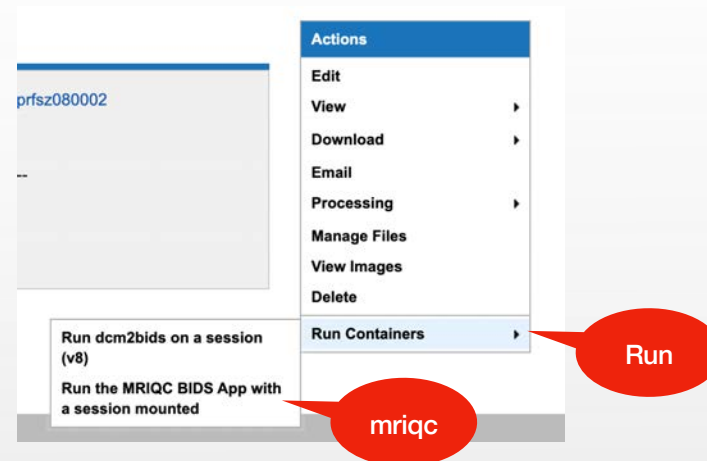
Here, we propose MRIQC for the assessment of the acquired (unprocessed) data. We also recommend specifying clear exclusion criteria before assessment. To our knowledge, there is no consensus on a data curation protocol; laboratories currently address the problem by applying their internal know-how and subjective assessments, or by skipping the data-curation step altogether. An ultimate curation protocol remains the subject of active discussion in the field. Describing such a protocol to check the quality of unprocessed data is beyond the scope of this article.

<https://www.nature.com/articles/s41596-020-0327-3>

## XNAT mriqc usage

After XNAT generates BIDS data for your session, you may run mriqc.

1. From the Session page, select the Run Containers Tab.
2. ONLY AFTER dcm2bids completes can you launch mriqc.
3. You will receive email confirmation once mriqc completes (up to an hour).



## XNAT mriqc usage cont.

- After MRIQC is run you will be able to access the outputs as html files on the MR Session Page

### MRI QC

#### Report Files

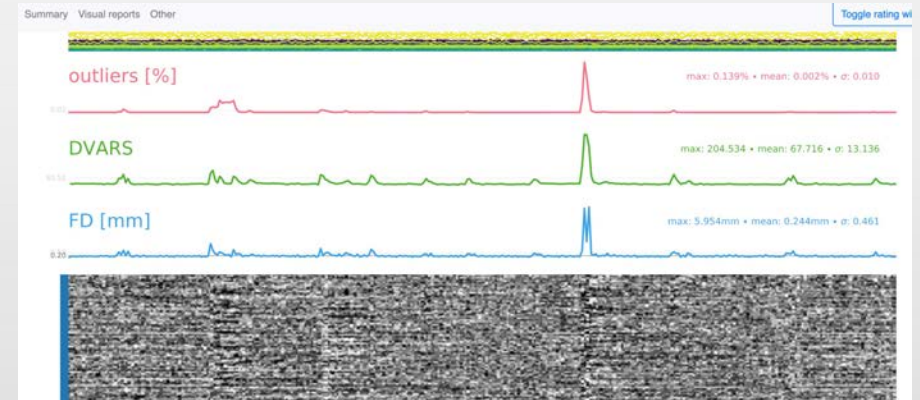
sub-prfsz080002_ses-20418_task-rest_run-03_bold.html
sub-prfsz080002_ses-20418_run-02_T1w.html
sub-prfsz080002_ses-20418_T2w.json
sub-prfsz080002_ses-20418_run-02_T1w.json
sub-prfsz080002_ses-20418_run-01_T1w.json
sub-prfsz080002_ses-20418_task-rest_run-02_bold.json
sub-prfsz080002_ses-20418_task-rest_run-04_bold.json
sub-prfsz080002_ses-20418_task-rest_run-01_bold.json
sub-prfsz080002_ses-20418_task-rest_run-03_bold.json
sub-prfsz080002_ses-20418_task-rest_run-04_bold.html
dataset_description.json
sub-prfsz080002_ses-20418_T2w.html
sub-prfsz080002_ses-20418_task-rest_run-01_bold.html
sub-prfsz080002_ses-20418_run-01_T1w.html
sub-prfsz080002_ses-20418_task-rest_run-02_bold.html

html

### History

Reports

- Mriqc provides Visual reports for anatomy
- Mriqc provides FD DVARS and visual reports for BOLD



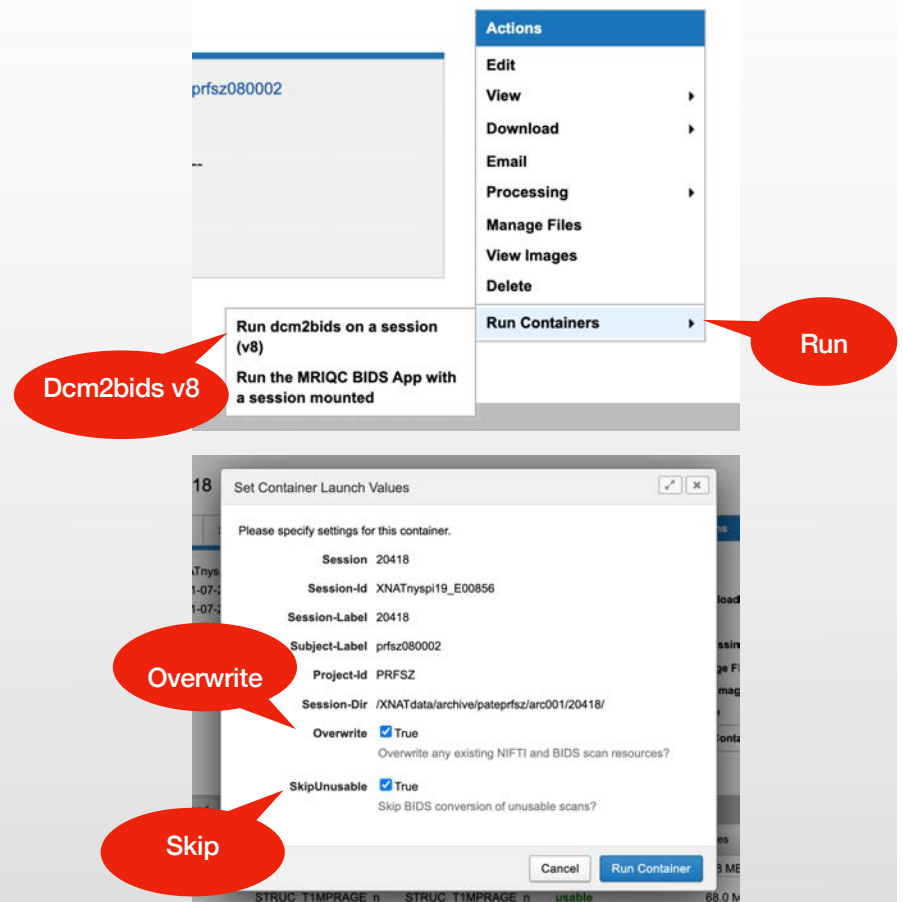
For further explanation of mriqc outputs visit <https://mriqc.readthedocs.io/en/0.9.0-0/mriqc.reports.html>

# Running XNAT BIDS Manually

Running the XNAT BIDS pipeline on a session manually is supported in circumstances where BIDS data needs to be corrected

1. From the Session page, select the Run Containers Tab.
2. IMPORTANT: Only run one exam through the BIDS pipeline at a time in order to limit resource utilization on the XNAT system.
3. Select "dcm2bids" v8:
  - Once selected make sure to select Skip Unusable and Overwrite, unless you don't need to reconvert existing bids data.
4. You will receive email confirmation once dcm2bids completes (up to an hour).
5. NOTE: If you are NOT an "owner" for the project, please contact Juan or David to rerun dcm2bids on a session for you as you will not have permissions to update the session information.

[David.Semanek@nyspi.columbia.edu](mailto:David.Semanek@nyspi.columbia.edu)



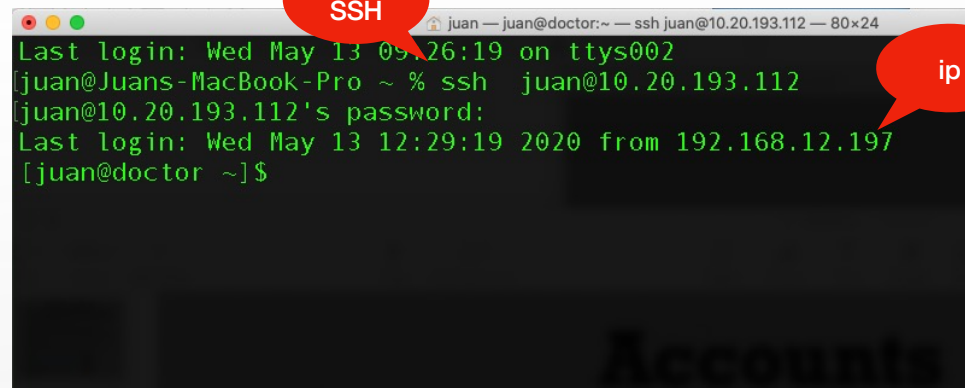
## Preparing your data for post processing on the CLUSTER

Once your data has been stored, marked useable and unusable, and converted to BIDS format on XNAT, The MRI computing cluster will provide the infrastructure for preparing your data for pre-processing.

- Pulling all BIDS data from XNAT and creating a BIDS-compliant, "rawdata" folder structure
- Editing the json sidecars for proper use of FieldMaps
- Editing multi volume field map acquisitions
- Validating BIDS data
- Pulling some non-BIDS data from XNAT to the cluster

\*The cluster refers to the array of different computers that work together to allow for parallel processing and is accessed through a single managing computer called "the doctor"

## CLUSTER account creation and access



A terminal window titled 'juan — juan@doctor:~ — ssh juan@10.20.193.112 — 80x24'. The window shows the following text: 'Last login: Wed May 13 09:26:19 on ttys002', '[juan@Juans-MacBook-Pro ~ % ssh juan@10.20.193.112', '[juan@10.20.193.112's password:', 'Last login: Wed May 13 12:29:19 2020 from 192.168.12.197', and '[juan@doctor ~]\$'. A red speech bubble labeled 'SSH' points to the terminal window, and another red speech bubble labeled 'ip' points to the IP address '10.20.193.112' in the command line.

```
juan — juan@doctor:~ — ssh juan@10.20.193.112 — 80x24
Last login: Wed May 13 09:26:19 on ttys002
[juan@Juans-MacBook-Pro ~ % ssh juan@10.20.193.112
[juan@10.20.193.112's password:
Last login: Wed May 13 12:29:19 2020 from 192.168.12.197
[juan@doctor ~]$
```

- CLUSTER accounts are created by emailing [Juan.Sanchez@nyspi.columbia.edu](mailto:Juan.Sanchez@nyspi.columbia.edu)
- Once your account has been created you will have a username and password associated with your “cluster account”
- To access the cluster from your computer open terminal and ssh into “the doctor” (ip 10.20.193.112) using your username and password.
- A basic knowledge of terminal is needed to navigate the cluster as there is no graphical user interface available.
  - <https://www.youtube.com/watch?v=BMGixkvJ-6w>
- A basic knowledge of vi is needed to manage text files
  - <https://www.youtube.com/watch?v=HMpB28l1sLc>



# BIDS Directory Structure on cluster

- Once you have access to the protocol, you can navigate to your project's home directory.

```
[juan@doctor ~]$ cd /MRI_DATA/nyspi/xnatprojectname
```

- The project name can be found on the MR Session page
- IMPORTANT: Do not create any new, arbitrary folders within a BIDS directory as this can disrupt the supported pipelines!
- The directory structure on the cluster is in BIDS format and will have no MR data (yet).
  - **rawdata** will be the destination for all BIDS data pulled from XNAT.
  - **derivatives** will be the destination for all post processing outputs and non-BIDS compliant data.
    - Abcd, neuromelanin, fmripreg, ciftify ... etc
  - **scripts** will be the location of the **xnatproject\_working.lst** file, which you will populate with the XNAT exam numbers for the data you wish you pull to the cluster.

```
xnatprojectname/derivatives:
abcd          ciftify          fmripreg

xnatprojectname/rawdata:
sub-xnatsubject1      sub-xnatsubject3
sub-xnatsubject2      sub-xnatsubject4

xnatprojectname/scripts:
xnatprojectname_working.lst
```

## BIDS SPEC explanations

<https://bids-specification.readthedocs.io/en/stable/01-introduction.html>

<https://bids-specification.readthedocs.io/en/stable/02-common-principles.html#source-vs-raw-vs-derived-data>

## Using the `sub-xgrab` program to prepare data for post processing

- BIDS data downloading and preparation is performed through the `sub-xgrab` program on the cluster.
- Before running the program, it is necessary to populate the `xnatproject_working.lst` for any subjects that need to be prepped.
- The program reads the `xnatproject_working.lst` present in the "scripts" directory list which the user populates in vi using data from XNAT.

```
/MRI_DATA/nyspi/patensasc/scripts/patensasc_working.lst
```

```
20286  
20071  
20185  
20218  
20126
```

Exam Number

- Each row on this list will represent a single subject session that needs to be prepped.
- The only information needed for `sub-xgrab` to find the data on XNAT is the unique exam number for the session.

## Running the **sub-xgrab** program

- Now that the **xnatproject\_working.lst** list is populated we can use the **sub-xgrab** program, which will run on all sessions in the list.
- The program is accessed through terminal and can be run from any location on the doctor:

```
[semanekD@doctor ~]$ sub-xgrab -d department -p project_id -a arguments -u username_for_xnat
```

- In order to use the program, 4 arguments must be passed after the above command
  - -d directory in MRI\_DATA where your project lives (if acquired at NYSPI, -d will always be nyspi)
  - -p the XNAT project name
  - -a the arguments that you want to run to prepare your data
  - -u your XNAT username

```
usage: test/sub-xgrab [ -d department -p project_id -a arguments -u username_for_xnat]

Where department is before project in the path

[arguments] can be:

Individual:
    [pull] - Download bids into rawdata
    [split] - Split multivolume data.
    [intend] - Apply intendedFor and taskName
    [nm] - Download neuromelanin
    [check] - Check if sessions exist in rawdata
Combined:
    [pulltosplit] - pull + split
    [pulltointend] - pull + split + intend
    [splittointend] - split + intend
```

- While -d and -p will always be the same for your project, -a can be any of these options

## sub-xgrab arguments for BIDS data converted on XNAT

You can input any of the single three arguments to run on the exams in the working list. You can also combine them.

- -a pull
  - Downloads converted BIDS data from XNAT to "rawdata". This argument by itself will release the terminal and run in the background so you can continue to use your SSH session. It's best for situations where you are pulling several exams from XNAT as this step takes some time to run. Logs can be checked using [sub-xmanage](#) (more on this in the next slide).
- -a split
  - Edits multivolume field maps if applicable, produces two field map series for each multivolume field map, with the data from each encoding direction in a separate series. This argument runs quickly, even on multiple exams, in the foreground of the terminal and prints its log information in the terminal's normal outputs.
- -a intend
  - Populates fieldmap json sidecars with information linking them to their corresponding bold data. This argument runs quickly, even on multiple exams, in the foreground of the terminal and prints its log information in the terminal's normal outputs.
- -a check
  - Runs a check of sessions in the working list against session folders in rawdata. The script will report any missing sessions or confirm all sessions exist in rawdata. This is useful if you have pulled many sessions and want to confirm they have all transferred, or if split or intend report a missing session and you want to find if there are any other sessions missing. You can also put all your session numbers for your project in your working list and run a comprehensive report of what sessions are available on the cluster.
- Combinations
  - -a pulltosplit
  - -a pulltointend (unlike "pull" by itself, the combined commands will run in the foreground of the terminal, and could take fair amount of time if you are pulling multiple exams)
  - -a splittointend

## Managing **sub-xgrab** workflows with **sub-xmanage**

- You will see the Active status replaced with Inactive when the sub-xgrab program has completed.
  - If you need to troubleshoot any errors, you can look at the log files by typing 1.
  - Make sure to Remove the job once completed by entering 3!
- When submitting jobs that include the **-a pull** argument, you can check on the progress and troubleshoot any errors from the logs by using **sub-xmanage**.
  - In this example we ran the following command:  
**sub-xgrab -d nyspi -p spanint -a pull -u semanekD**
  - You can check on the progress and see the output by typing the following in the terminal from any location: **sub-xmanage**.
  - After entering this command, you will see the following screen. NOTE: You MUST use **sub-xmanage** to remove your **sub-xgrab** job after it is done as the cluster only allows one job per user on the system (you may of course run multiple exams from the same XNAT project in a single job).

```
XGRAB MANAGEMENT

CURRENT ACTIVE & INACTIVE XGRAB JOBS FOR USER semanekD

  Status      Job Name
  Active  -->  sub-xgrab_semanekD

Available management options:
1) Check xgrab Log    3) Remove xgrab Job    5) Exit
2) Save xgrab Log     4) Refresh

(Press Ctrl+C to abort)

Select an option: █
```

Job status

Job logs

Remove job

# BIDS validation and data correction

- Once the data are prepped, they must be validated.
- All subjects/exams in "rawdata" will be validated regardless of who is in the working list.
- To validate your data run the following command in the terminal from any location: `validate xnatprojectname`

WOO HOO!

```
[semanekD@doctor ~]$ validate spanint
bids-validator@1.8.9
bids-specification@disable
(node:1) Warning: Closing directory handle on garbage collection
(Use `node --trace-warnings ...` to show where the warning was created)
This dataset appears to be BIDS compatible.
      Summary:                Available Tasks:                Available Modalities:
      46 Files, 2.41GB         FUNC_MUX6_GNG                MRI
      2 - Subjects            FUNC_MUX6_Resting
      2 - Sessions            FUNC_MUX6_SST

If you have any questions, please post on https://neurostars.org/tags/bids.
```

- If the BIDS subject ID for a session needs to be changed, we offer `sub-xfix_modbidsid` with the following usage:  
`sub-xfix_modbidsid [ -d department -p project_id -s BIDS_session_number -n new_ID ]`

## sub-xgrab and non-BIDS compatible DICOM data

sub-xgrab can pull some supported DICOM-only data from XNAT for data that cannot be converted to BIDS using the XNAT tool. Currently, the cluster supports the pulling of neuromelanin series using the following argument:

- -a nm
  - Downloads neuromelanin DICOM data for exams on the working list (in the "scripts" folder) from XNAT to the /derivatives/neuromelanin folder in your project folder on the cluster in "{exam\_number}\_{series\_number}.tar.gz" format. This argument by itself will release the terminal and run in the background so you can continue to use your SSH session. Logs can be checked, and jobs managed using sub-xmanage, the same as for -a pull, and only one job can be run at a time. Please remove your job after it completes.

## Managing Your Saved XNAT Password

If your XNAT password changes or the logs report an issue connecting to XNAT and you know your account is active and XNAT is reachable, you can reset your own XNAT password on the cluster. The pipeline securely stores your password in an encrypted file that is accessible to the user who owns it. Once removed, sub-xgrab will ask you again to enter your XNAT password. You can run the following command to reset your password:

```
rm -f /MRI_DATA/.xnatauth/{DOCTOR_username}_token.enc
```

## Post Processing in AWS

- Once you have your account and you have validated all of your raw data you are now ready to run the post processing pipelines by accessing <http://mrijob.nyspi.local>
- Choose between fmripred/ciftify , abcd or freesurfer7
- Please be sure to consult with [Gaurav.Patel@nyspi.columbia.edu](mailto:Gaurav.Patel@nyspi.columbia.edu) to help decide which pipeline is best for your data.
- In order to launch the post processing pipelines you will need to email [Juan.Sanchez@nyspi.columbia.edu](mailto:Juan.Sanchez@nyspi.columbia.edu) to create aws account
- The username and password will be the same as your [nyspi.columbia.edu](mailto:nyspi.columbia.edu) email



## Post Processing in AWS

1. Select your pipeline
2. Enter the project acronym or XNAT project
3. Enter the input (location of your verified rawdata on the cluster)
4. Enter the output (location of the derivatives on the cluster)
  1. Abcd requires the abcd directory in the path, fmriprep does not
5. Enter the subjects you wish to be processed
6. Click I agree

### MRI PREPROCESSING

#### MRI Processing Actions

Launch a job

Job status

#### MRI Pre-processing

All use of this system is logged. Your NYSPI account is associated to this workload and any activity will be individually attributed to you. You are responsible for any activity performed on this system.

Any costs associated with processing will be attributed to your group / project. Contact the MRI unit for current details.

When the server has started, you will be able to log using the account information

#### Existing Sessions

No active sessions detected for your account yet. But you've come to the right place to fix that!

#### New Session Details

Pipeline:

☐ Multiband FMRIprep + CIFTIFY

☐ Singleband FMRIprep + CIFTIFY

☐ CIFTIFY Only

☒ ABCD DCAN

☐ Freesurfer recon-all

\*FMRIprep is the standard. Did you need to use the legacy FMRIprep 1.5.10?

☐ Use FMRIprep 1.5.10

Project acronym (e.g. horgbell):

Input and Output must be using standard shared volumes MRI\_DATA or XNATDATA

Input path (e.g. /MRI\_DATA/myid/project/rawdata):

Output path (e.g. /MRI\_DATA/myid/project/derivatives):

Subject ID(s) (e.g sub-horgbell8001) one per line: 

sub-patensasc4044  
sub-patensasc4045  
sub-patensasc4047  
sub-patensasc4051

Click below to acknowledge your use of this service for all activity from these services, including costs associated with their use.

pipeline

Project name

Rawdata

Derivatives

Subjects

Agree

## Accessing Post Processed Data

- Once AWS is complete the data will be automatically synced from the cloud to the doctor and in the chosen derivatives directory
- The amount of time to run depends on the pipeline and the size of the data being processed
  - Abcd - 6-8 hours
  - Fmriprep - 4-6 hours
- You can monitor the status of the job(s) by selecting the Job Status tab
- You are responsible for rsyncing your data if you need it locally
- You may access the data on the cluster using the Virtual Machine
  - If this is your plan you will need to email [Juan.Sanchez@nsypi.columbia.edu](mailto:Juan.Sanchez@nsypi.columbia.edu)

## ABCD - derivatives organization

### Outputs

The outputs are organized in the following structure:

```
output_dir/
|__ sub-id
    |__ ses-session
        |__ files
            |__ executive_summary
            |__ MNINonLinear
                |__ fsaverage_LR32k
                |__ Results
            |__ T1w
                |__ id
                |__ task-taskname
        |__ logs
```

### files

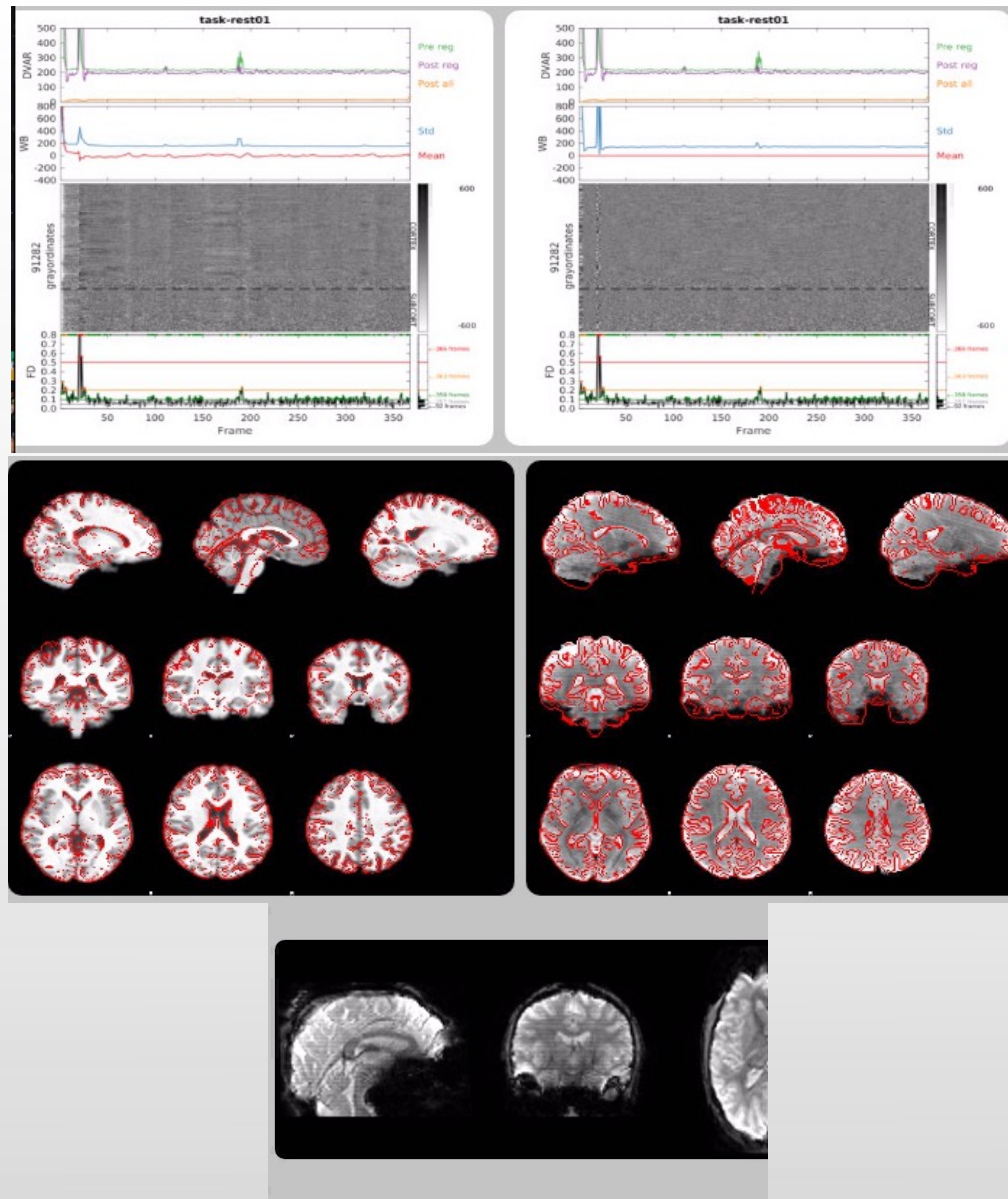
- **executive\_summary** : The .html file within can be opened for quality inspection of pipeline results.
- **MNINonLinear** : Contains the final space results of anatomy in 164k resolution.
- **MNINonLinear/fsaverage\_LR32k** : Final space anatomy in 32k resolution, where functional data is ultimately projected.
- **MNINonLinear/Results** : Final space functional data.
- **T1w** : Contains native space anatomical data as well as intermediate preprocessing files.
- **T1w/id** : The participant ID folder within T1w is the FreeSurfer subject folder.
- **task-taskname** : These folders contain intermediate functional preprocessing files.

- Please visit the [https://www.humanconnectome.org/storage/app/media/documentation/tutorials/Connectome\\_WB\\_Tutorial\\_v1.5.pdf](https://www.humanconnectome.org/storage/app/media/documentation/tutorials/Connectome_WB_Tutorial_v1.5.pdf) for an in-depth review of file types and how to use workbench to visualize data. In addition for more information on DCAN signal processing please visit: [https://github.com/DCAN-Labs/dcan\\_bold\\_processing](https://github.com/DCAN-Labs/dcan_bold_processing)

ABCD - derivatives  
MNINonLinear/Results/task/DCAN

	<b>BOLD</b>
<b>Volume MNI</b>	MNINonLinear/Results/task-rest01/task-rest01.nii.gz
<b>Surface fs_LR_32k (HCP Surface Atlas)</b>	MNINonLinear/Results/task-rest01/task-rest01_Atlas.dtseries.nii
<b>Surface Motion Artifact Cleaned</b>	MNINonLinear/Results/task-rest01/DCANBOLDProc_v4.0.0/task-rest01_DCANBOLDProc_v4.0.0_Atlas.dtseries.nii
<b>Surface Motion Artifact Cleaned concatenated (use this for voxelwise RS analyses)</b>	MNINonLinear/Results/task-rest_DCANBOLDProc_v4.0.0_Atlas.dtseries.nii
<b>Surface Motion Artifact Cleaned concatenated and parcellated (use this for parcel wise RS analyses)</b>	MNINonLinear/Results/task-rest_DCANBOLDProc_v4.0.0_Gordon.ptseries.nii

## ABCD - Executive Summary/QC



## FM RIPREP - derivatives (anatomy)

### Anatomical derivatives

Anatomical derivatives are placed in each subject's `anat` subfolder:

```
sub-<subject_label>/
  anat/
    sub-<subject_label>[_space-<space_label>]_desc-preproc_T1w.nii.gz
    sub-<subject_label>[_space-<space_label>]_desc-brain_mask.nii.gz
    sub-<subject_label>[_space-<space_label>]_dseg.nii.gz
    sub-<subject_label>[_space-<space_label>]_label-CSF_probseg.nii.gz
    sub-<subject_label>[_space-<space_label>]_label-GM_probseg.nii.gz
    sub-<subject_label>[_space-<space_label>]_label-WM_probseg.nii.gz
```

### FreeSurfer derivatives

A FreeSurfer subjects directory is created in `<output_dir>/freesurfer`, or the directory indicated with the `--fs-subjects-dir` flag.

```
<output_dir>/
  fmriprep/
  ...
  freesurfer/
    fsaverage{,5,6}/
      mri/
      surf/
      ...
    sub-<subject_label>/
      mri/
      surf/
      ...
  ...
```

Please visit <https://fmriprep.org/en/stable/outputs.html> to learn more details about the fmriprep outputs.

# FMRIprep - derivatives (func)

## Functional derivatives

Functional derivatives are stored in the `func/` subfolder. All derivatives contain `task-<task_label>` (mandatory) and `run-<run_index>` (optional), and these will be indicated with `[specifiers]` :

```
sub-<subject_label>/
  func/
    sub-<subject_label>_[specifiers]_space-<space_label>_boldref.nii.gz
    sub-<subject_label>_[specifiers]_space-<space_label>_desc-brain_mask.nii.gz
    sub-<subject_label>_[specifiers]_space-<space_label>_desc-preproc_bold.nii.gz
```

**Grayordinates files.** `CIFTI` is a container format that holds both volumetric (regularly sampled in a grid) and surface (sampled on a triangular mesh) samples. Sub-cortical time series are sampled on a regular grid derived from one MNI template, while cortical time series are sampled on surfaces projected from the [\[Glasser2016\]](#) template. If CIFTI outputs are requested (with the `--cifti-outputs` argument), the BOLD series are also saved as `dtseries.nii` CIFTI2 files:

```
sub-<subject_label>/
  func/
    sub-<subject_label>_[specifiers]_bold.dtseries.nii
```

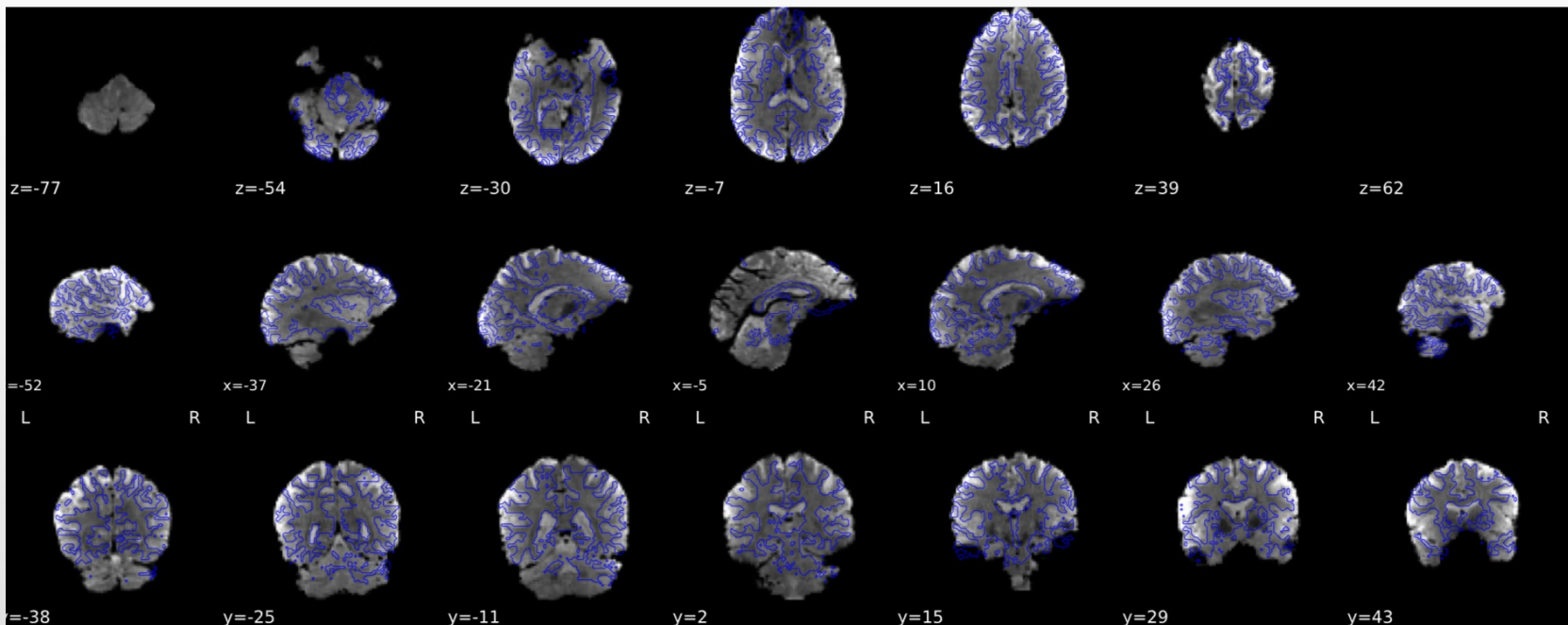


# FM RIPREP - QC

## Visual Reports

*fMRI*Prep outputs summary reports, written to `<output_dir>/fmripred/sub-<subject_label>.html`.

These reports provide a quick way to make visual inspection of the results easy. Each report is self contained and thus can be easily shared with collaborators (for example via email). [View a sample](#)



Please visit <https://fmripred.org/en/stable/outputs.html> to learn more details about the fmripred outputs.