Overview of (and Some Lessons from) the FMRI Open QC Project

What is quality control in FMRI?

Well, I guess it's the kind of thing usually described in papers as:

"*Data were checked for quality.*"

...

What is quality control in FMRI?

Is it:

- ... finding *good* subject data in a collection, to use?
- ... finding *bad* subject data in a collection, to remove?

What is quality control in FMRI?

Is it:

- ... finding *good* subject data in a collection, to use?
- ... finding *bad* subject data in a collection, to remove?

→ It should probably be viewed as a much larger, holistic process, like:
Being as sure as possible about the contents of the data collection, from acquisition properties to artifact checking to regression model evaluation.

What is quality control in FMRI?

Is it:

- ... finding *good* subject data in a collection, to use?
- ... finding *bad* subject data in a collection, to remove?

→ It should probably be viewed as a much larger, holistic process, like:
Being as sure as possible about the contents of the data collection, from acquisition properties to artifact checking to regression model evaluation.

Things to include : data consistency checks, raw data checks, processing step checks, and more
Things to avoid : wasting data, biasing outcomes, leaving QC until too late :(

What is needed for quality control in FMRI?

Example 1: Is the following EPI data is good or bad?

- Average TSNR is 177, and only 2 of 150 time points were motion-censored.

What is needed for quality control in FMRI?

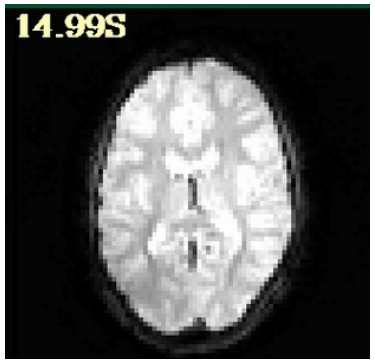
Example 1: Is the following EPI data is good or bad?

- Average TSNR is 177, and only 2 of 150 time points were motion-censored. 😎

What is needed for quality control in FMRI?

Example 1: Is the following EPI data is good or bad?

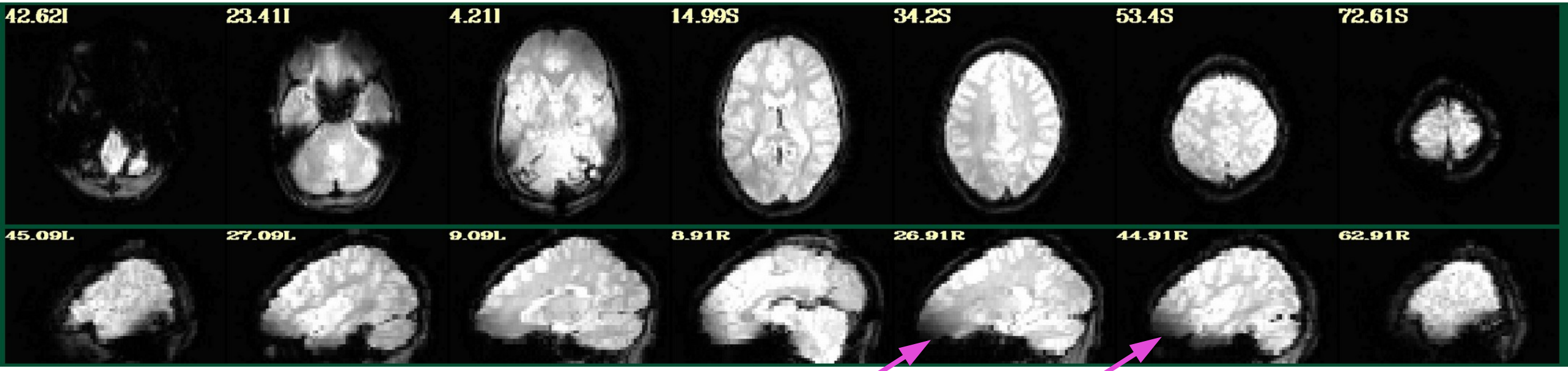
- Average TSNR is 177, and only 2 of 150 time points were motion-censored. 😎
- It looks OK: 😎



What is needed for quality control in FMRI?

Example 1: Is the following EPI data is good or bad?

- Average TSNR is 177, and only 2 of 150 time points were motion-censored. 😎
- It looks OK… or does it?

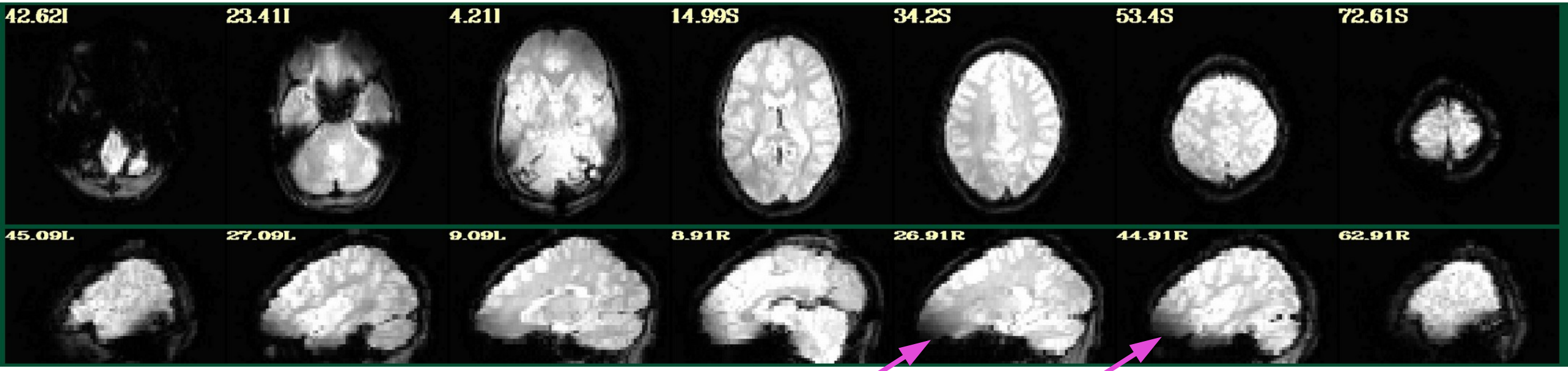


There are locations with severe signal loss. So, this dataset is problematic. 😢

What is needed for quality control in FMRI?

Example 1: Is the following EPI data is good or bad?

- Average TSNR is 177, and only 2 of 150 time points were motion-censored. 😎
- It looks OK… or does it?



There are locations with severe signal loss. So, this dataset is problematic. 😢

... But what if we study only visual/motor corts? *Life with data is complicated!* 🧐

What is needed for quality control in FMRI?

Example 2: Is the following EPI data is good or bad?

- Average TSNR is 194, and only 7 of 150 time points were motion-censored.

What is needed for quality control in FMRI?

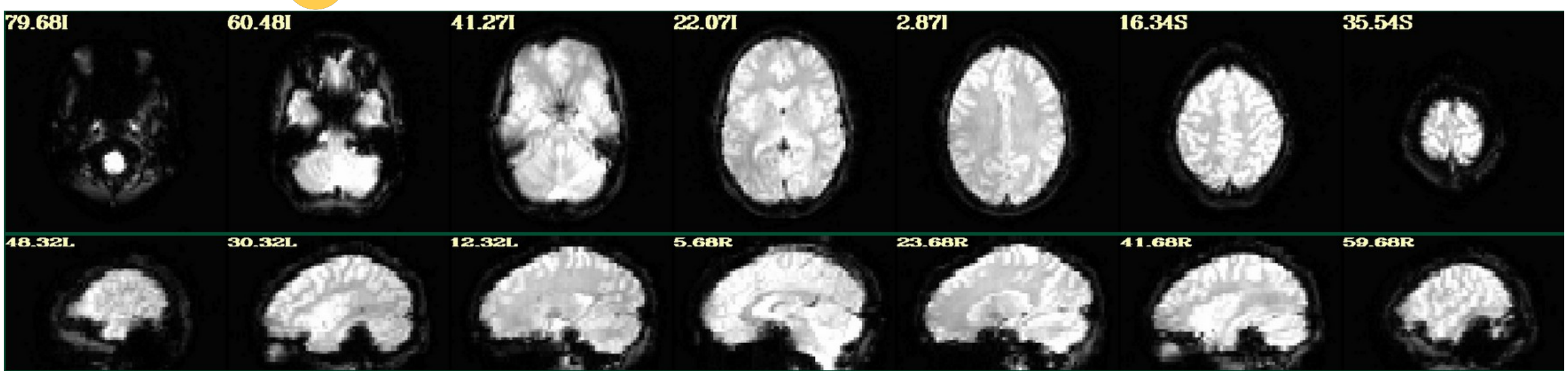
Example 2: Is the following EPI data is good or bad?

- Average TSNR is 194, and only 7 of 150 time points were motion-censored. 😎

What is needed for quality control in FMRI?

Example 2: Is the following EPI data is good or bad?

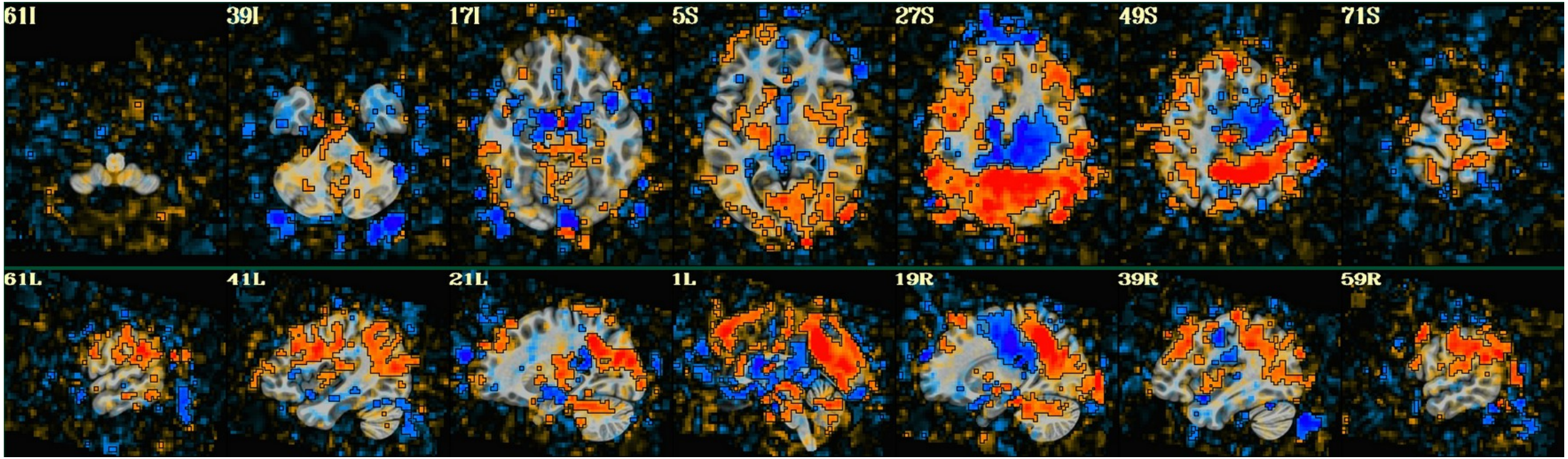
- Average TSNR is 194, and only 7 of 150 time points were motion-censored. 😎
- It looks OK: 😎



What is needed for quality control in FMRI?

Example 2: Is the following EPI data is good or bad?

- Average TSNR is 194, and only 7 of 150 time points were motion-censored. 😎
- It looks OK: 😎
- … but it's "default mode network" correlation map (seed in PCC) is: 😭



→ NB: FMRI datasets are 4D, so checking spatiotemporal properties is key

→ So, the FMRI Open QC Project: what *should* QC in FMRI be?

"Demonstrating Quality Control (QC) Procedures in fMRI"

Organized by Paul Taylor, Jo Etzel, Daniel Glen and Rick Reynolds,
with collaborators Arshitha Basavaraj and Dustin Moraczewski

Important links

- Description and details:
  - https://www.frontiersin.org/research-topics/33922/demonstrating-quality-control-qc-procedures-in-fmri#overview
  - https://osf.io/qaesm/wiki/home/
- Participating teams' article collection:
  - https://www.frontiersin.org/research-topics/33922/demonstrating-quality-control-qc-procedures-in-fmri#articles
- The data collections used:
  - https://osf.io/qaesm/files/osfstorage
- Editorial (overview, summary and notes):
  - Taylor PA, Glen DR, Reynolds RC, Basavaraj A, Moraczewski D and Etzel JA (2023) *Editorial: Demonstrating quality control (QC) procedures in fMRI.* Front. Neurosci. 17:1205928.
  - https://www.frontiersin.org/articles/10.3389/fnins.2023.1205928/full

Why was the Project setup?

1. **To promote the broader adoption of quality control practices in the FMRI field.**
   Many packages already contain QC tools/protocols (e.g., those in AFNI, CONN, DPARSF, fMRIprep, MRIQC, pyfMRIqc and SPM were all used here). Check ’em out!

2. **To facilitate the inclusion of more details in QC protocol descriptions.**
   Each Project contribution contained an explicit list of QC criteria, plus examples.

3. **To share QC criteria across researchers and developers, improving available QC.**
   Increase clarity and potentially broaden the homogeneity of QC methods.

4. **To promote viewing QC as more than "just" vetting datasets, but rather as deeply understanding the contents of the collection and analysis as a whole.**
   Have greater confidence in results. (Maybe even improve reproducibility?)

How was the Project setup?

- Made a collection of real task-based and resting state FMRI datasets
- Invited anyone interested to process+QC the data, and categorize subject data:
  Include: passes QC criteria, have high confidence to use;
  Exclude: fails one or more QC criteria, have high confidence to remove;
  Uncertain: there is a question about whether to include.
- Each team shared detailed list of all QC criteria (qualitative and quantitative)
- Each team shared examples and images of all categories of subject data.

What was the Project data? (All real data, unaltered)

- **Group 1: ABIDE-1, KKI (Barber et al., 2012; Nebel et al., 2014), N = 20 subjects used (of 55 total)**
  Philips Achieva 3T, EPI axial slice acq. with fat saturation and SENSE (factor=3), flip angle = 75°, TE = 30 ms, TR = 2.5 s, voxel size = 2.67x2.67x3.0 mm, slice timing provided in JSON sidecar, PE direction = j-; subjects instructed to focus on a crosshair on black screen.
- **Group 2: ABIDE-1, Trinity (Delmonte et al., 2012), N = 20 subjects used (of 49 total)**
  Philips Achieva 3T, EPI axial slice acquisition with fat saturation and SENSE (factor=2), flip angle = 90°, TE = 28 ms, TR = 2.0 s, voxel size = 3.0x3.0x3.841 mm, slice timing provided in JSON sidecar, PE direction = j-, subjects instructed to close eyes during scan.
- **Group 3: ABIDE-2, KUL-3 (Bernaerts et al., 2016), N = 16 subjects used (of 28 total)**
  Philips Achieva Ds 3T, EPI axial slice acq. with fat saturation and with SENSE (factor=2), flip angle = 90°, TE = 30 ms, TR = 2.5 s, voxel size = 1.562x1.562x3.1 mm, slice timing in JSON sidecar, PE direction = j-, subjects instructed to focus on a white fixation cross on black background.
- **Group 4: FCP, Baltimore (Pekar and Mostofsky, 2010), N = 23 subjects used (of 23 total)**
  3T scanner (unspecified type), TR = 2.5s, voxel size = 2.667x2.667x3.0 mm, subjects instructed to keep eyes open and fixate (target unspecified) during scan.
- **Group 5: OpenNeuro, ds000220 (Roy et al., 2017), N = 20 subjects used (of 26 total)**
  Philips Achieva and Siemens Trio 3T, EPI axial slice acq. with segmented k-space (no SENSE), flip angle = 90°, TE = 34 ms, TR = 2 s, voxel size = 1.85x1.85x4.0 mm, instructions to subjects undescribed.
- **Group 6: OpenNeuro, ds000243 (Peterson et al., 2018), N = 20 subjects used (of 120 total)**
  Siemens Magnetom Trio 3T, 12 ch. head coil, flip angle = 90°, TE = 34 ms, TR = 2.5 s, voxel size = 4.0x4.0x4.0 mm, subject instructions undescribed.
- **Group 7: OpenNeuro, ds000245 (Yoneyama et al., 2018), N = 20 subjects used (of 45 total)**
  Siemens Verio 3T , 12 ch. head coil, flip angle = 80°, TE = 30 ms, TR = 2.5 s, voxel size = 3.0x3.0x3.51 mm, slice timing provided in JSON sidecar, subjects instructed to close eyes during scan.
- **Group 0: OpenNeuro, ds000030, "task-pamenc" (Poldrack et al., 2016; Bilder et al., 2018), N = 30 subjects used (of 272 total)**
  Siemens TrioTim 3T, EPI acq. with segmented k-space and fat saturation (acceleration factor PE = 2), flip angle = 90°, TE = 30 ms, TR = 2 s, slice timing provided in JSON sidecar, PE direction = j-.

Participating teams: list of papers

Quality control procedures and metrics for resting-state functional MRI

Rasmus M. Birn[1,2]*

Quality control in resting-state fMRI: the benefits of visual inspection

Rebecca J. Lepping[1,2]*, Hung-Wen Yeh[3,4], Brent C. McPherson[5], Morgan G. Brucks[2,6], Mohammad Sabati[2,7], Rainer T. Karcher[2], William M. Brooks[1,2], Joshua D. Habiger[8], Vlad B. Papa[2] and Laura E. Martin[2,6]

Quality control in functional MRI studies with MRIQC and fMRIPrep

Céline Provins[1]*, Eilidh MacNicol[2], Saren H. Seeley[3], Patric Hagmann[1] and Oscar Esteban[1]*

A functional MRI pre-processing and quality control protocol based on statistical parametric mapping (SPM) and MATLAB

Xin Di* and Bharat B. Biswal*

Efficient evaluation of the Open QC task fMRI dataset

Joset A. Etzel*

Demonstrating quality control procedures for fMRI in DPABI

Bin Lu[1,2]* and Chao-Gan Yan[1,2,3,4]*

Quality control practices in FMRI analysis: Philosophy, methods and examples using AFNI

Richard C. Reynolds*, Paul A. Taylor and Daniel R. Glen

Functional connectivity MRI quality control procedures in CONN

Francesca Morfini[1], Susan Whitfield-Gabrieli[1,2,3] and Alfonso Nieto-Castañón[4,5]*

The art and science of using quality control to understand and improve fMRI data

Joshua B. Teves[1], Javier Gonzalez-Castillo[1], Micah Holness[1], Megan Spurney[1], Peter A. Bandettini[1,2] and Daniel A. Handwerker[1]*

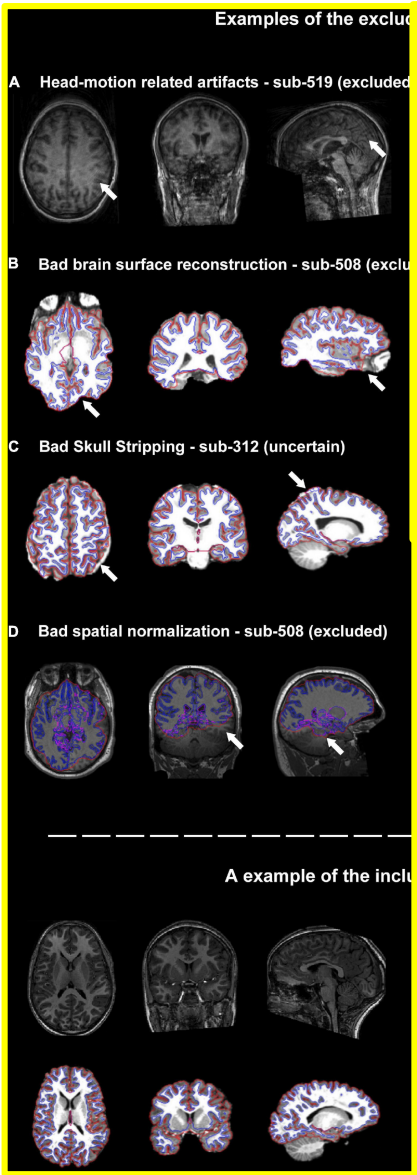
Inter-rater reliability of functional MRI data quality control assessments: A standardised protocol and practical guide using pyfMRIqc

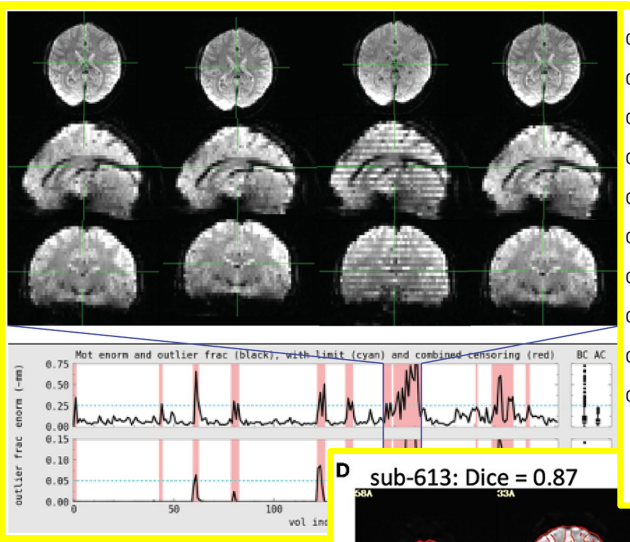
Brendan Williams[1,2]*, Nicholas Hedger[1,2†], Carolyn B. McNabb[3†], Gabriella M. K. Rossetti[1,2†] and Anastasia Christakou[1,2]

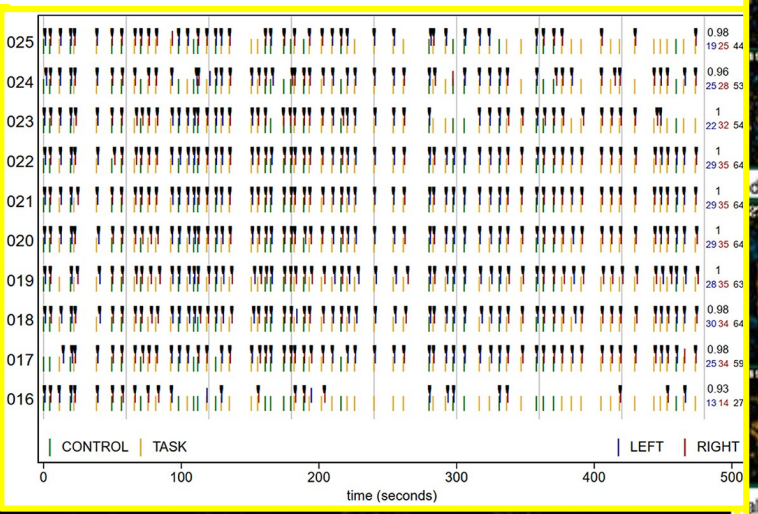
Participating teams: list of software used

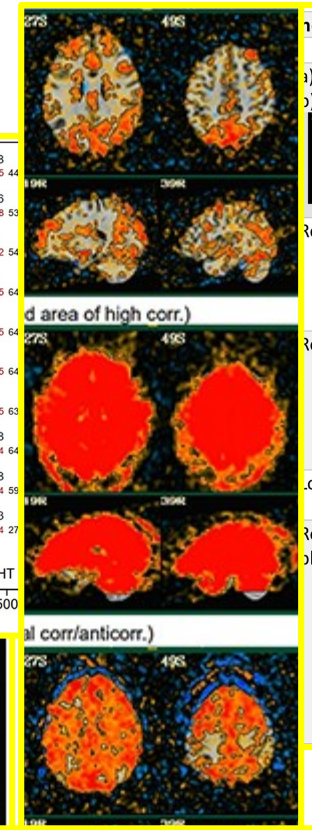
Team	Software for processing	Software for QC
A) Birn	AFNI, FSL, ANTs	AFNI
B) Di and Biswal	SPM, Matlab	SPM, Matlab
C) Etzel	fMRIPrep (with ANTs, AFNI, FreeSurfer, FSL, Nipype)	R (with knitr, RNifti and fields), AFNI
D) Lepping et al.	AFNI	AFNI, REDCap
E) Lu and Yan	DPABI, DPABISurf, DPARSF, fMRIPrep FreeSurfer, ANTs, FSL, AFNI, SPM, PALM, Matlab, DARTEL	DPABISurf, DPARSF, fMRIPrep, Matlab
F) Morfini et al.	CONN (with ART), SPM12, Matlab	CONN, SPM12, Matlab, FSLeyes
G) Provins et al.	MRIQC (with ANTs, AFNI, FreeSurfer, FSL, Nipype, SynthStrip), fMRIPrep (with ANTs, AFNI, FreeSurfer, FSL, Nipype)	MRIQC (with ANTs, AFNI, FreeSurfer, FSL, Nipype, SynthStrip), fMRIPrep (with ANTs, AFNI, FreeSurfer, FSL, Nipype)
H) Reynolds et al.	AFNI, FreeSurfer	AFNI
I) Teves et al.	FreeSurfer, AFNI	AFNI
J) Williams et al.	FSL, cinnqc (with FSL, pyfMRIqc)	pyfMRIqc

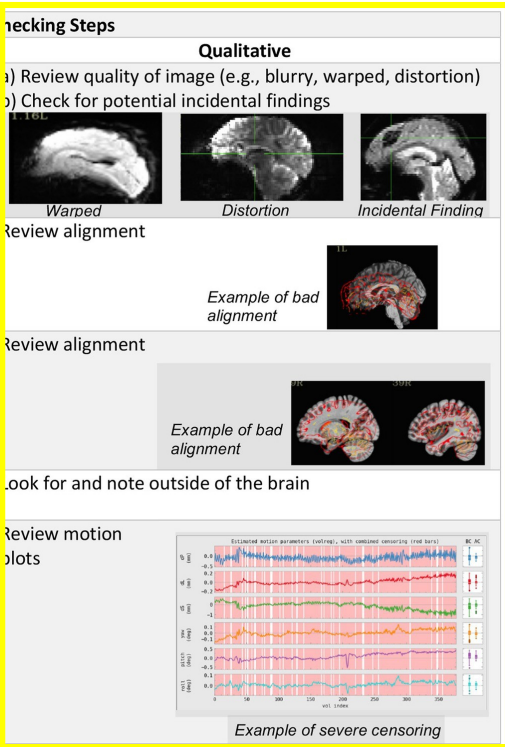
Participating teams: QC example images

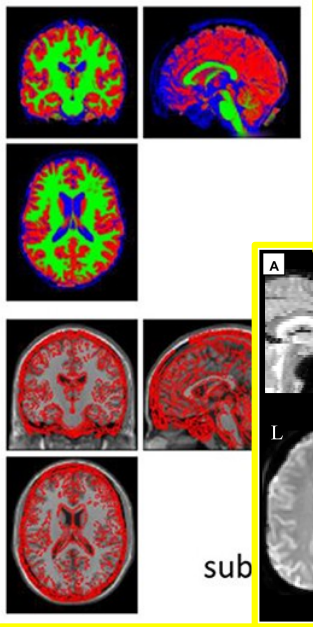


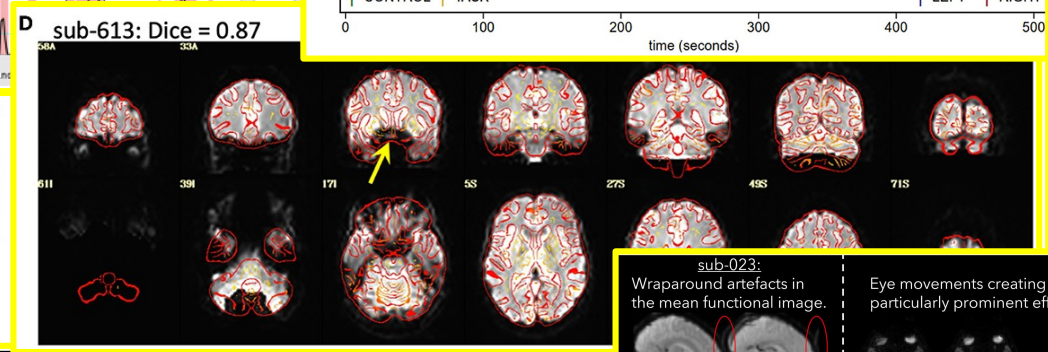


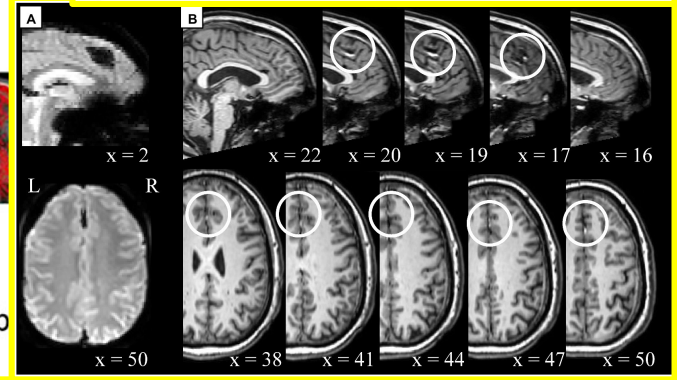


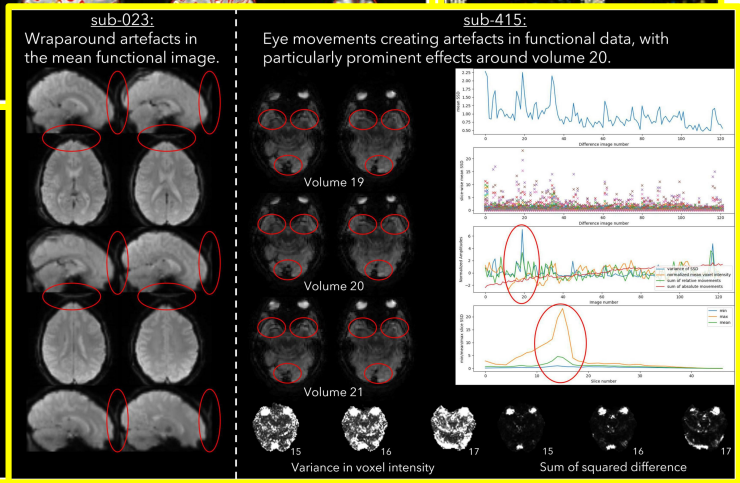


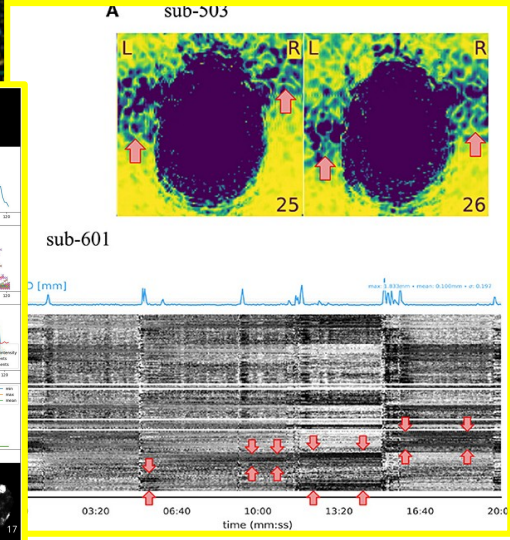




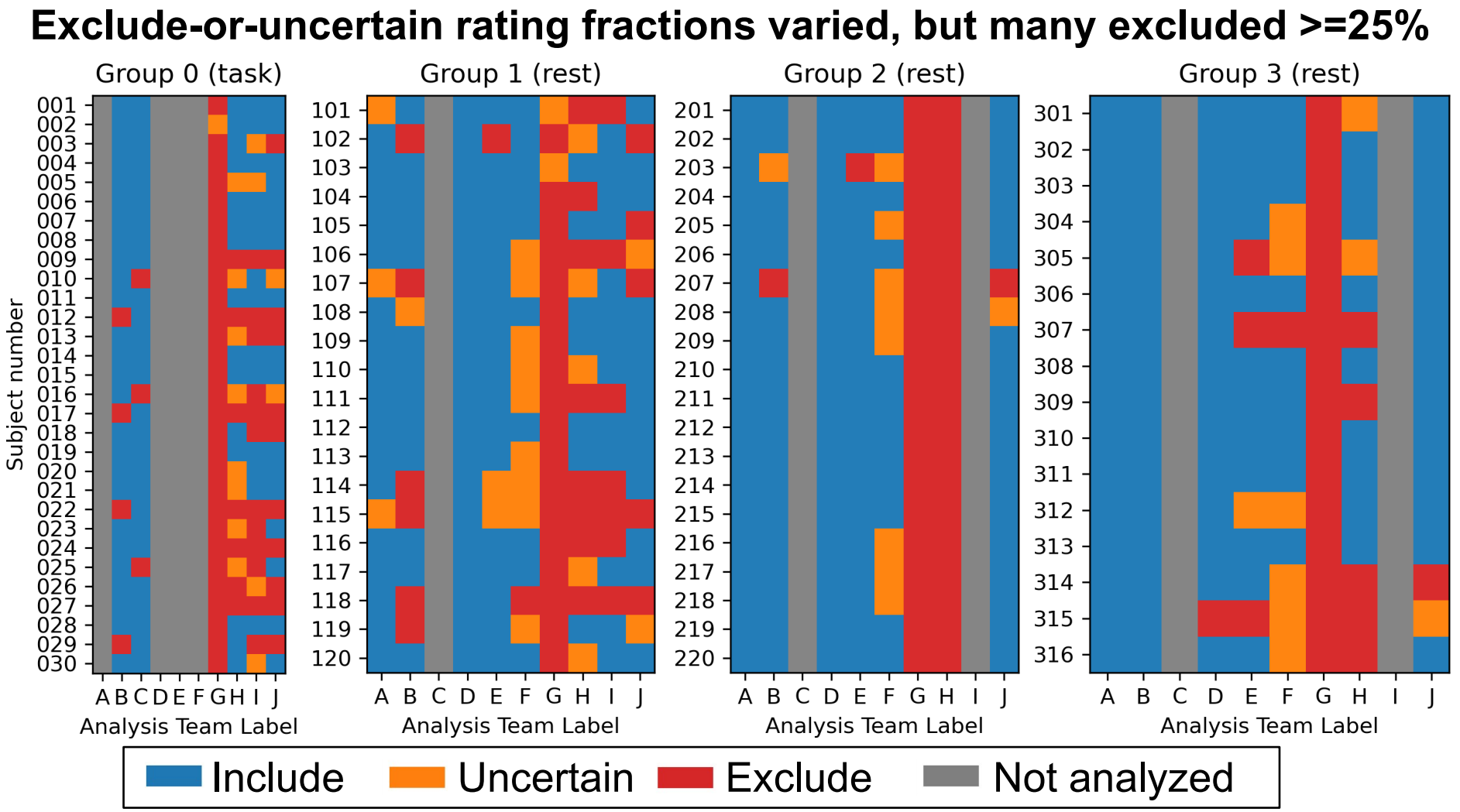




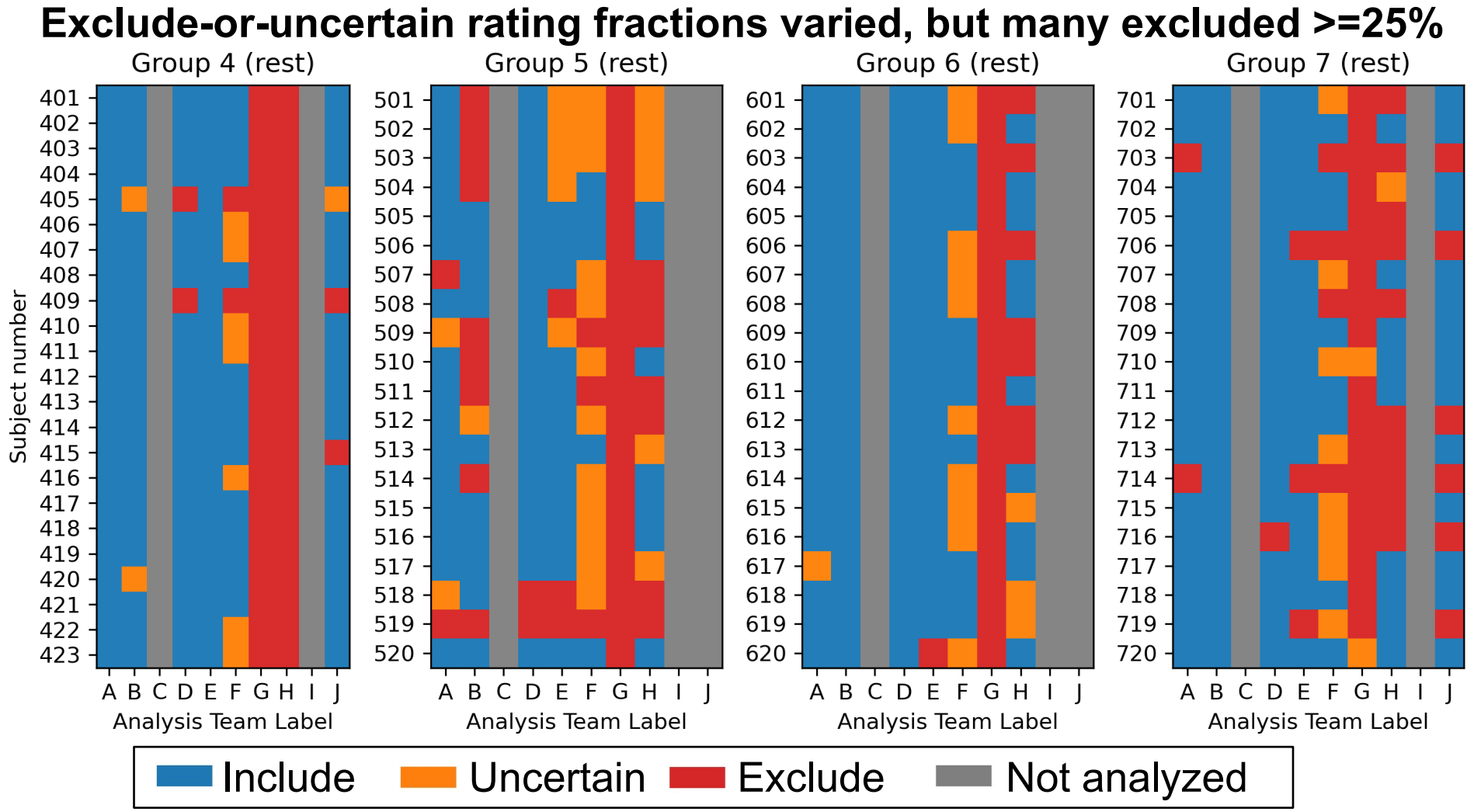




Participating teams: Team evaluations (1/2)



Participating teams: Team evaluations (2/2)



One output: Common themes across teams (1/2)

1. Each team found subjects to exclude based on data quality.
… often *many*. Check all data (downloaded or acquired) for appropriateness to a study.

2. Each team evaluated one or more subject's datasets as "uncertain."
This would typically lead to background checks (and maybe corrective measures).

3. Nearly all QC protocols checked the raw data's consistency and "metadata".
Alterations in scanner settings, software version, DICOM field conversion and more easily occur. These can hurt data compability, suitability and utility within a study.

4. Each team identified consistency, reliability or mismatch errors within datasets.
All teams found two upside-down EPIs. Some identified EPI-anatomical left/right flips, and even a likely subject dataset mismatch.

One output: Common themes across teams (2/2)

5. Each QC protocol used qualitative criteria and visual inspection of datasets.
Checking: raw data quality, and derived images to evaluate processing (or artifacts).

6. Most, but not all, protocols included quantitative/automatic checks.
Visualization still seems key to evaluating data features and processing steps.
Quantitative criteria typically originate as useful extensions of such understanding.

7. QC parameters were closely tied to a specific analysis and research goal.
Datasets may be appropriate for one particular analysis but for not another.

8. Non-EPI items can affect FMRI analysis, too.
All the input data used for analysis (anatomicals, timing files, ...) needs to be checked.

9. Each team made their processing and QC pipelines publicly available.
Hopefully this encourages having more detailed QC protocols and reporting.

Acknowledgments

Thanks to:

- the participating researchers, for spending their time and effort in processing, performing QC, and carefully presenting their results.

  Richard C Reynolds, Daniel R Glen, Joset A Etzel, Arshitha Basavaraj, Dustin Moraczewski, Rasmus M Birn, Xin Di, Bharat B Biswal, Rebecca J Lepping, Brent C McPherson, Hung-Wen Yeh, Morgan G Brucks, Mohammad Sabati, Rainer T Karcher, William Miles Brooks, Joshua D Habiger, Vlad B Papa, Laura E Martin, Bin Lu, Chao-Gan Yan, Francesca Morfini, Susan Whitfield-Gabrieli, Alfonso Nieto-Castanon, Céline Provins, Eilidh MacNicol, Saren H. Seeley, Patric Hagmann, Oscar Esteban, Joshua Teves, Javier Gonzalez-Castillo, Micah Holness, Megan Spurney, Peter A Bandettini, Daniel A Handwerker, Brendan Williams, Nicholas Hedger, Carolyn B McNabb, Gabriella MK Rossetti, Anastasia Christakou

- the reviewers and guest editors, who provided such useful feedback for the work in this Research Topic.