fMRI Course, Day 12: Reproducibility

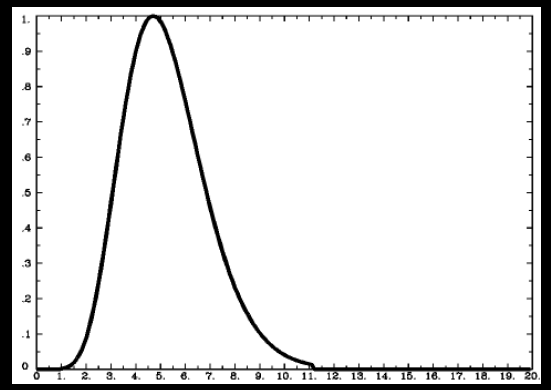
August 11th, 2023

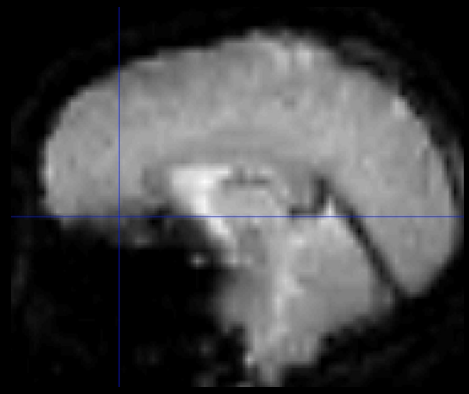
Review So Far

From stimulus to the BOLD response

How tissue properties, blood flow, and magnetic properties interact

Creating contrast images from T1- and T2-weightings



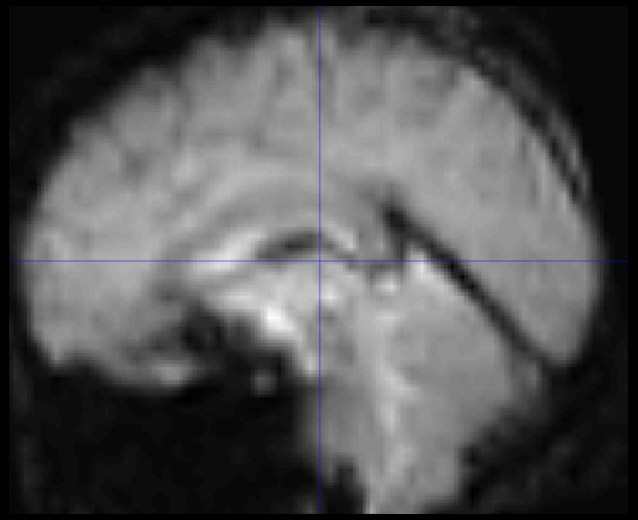


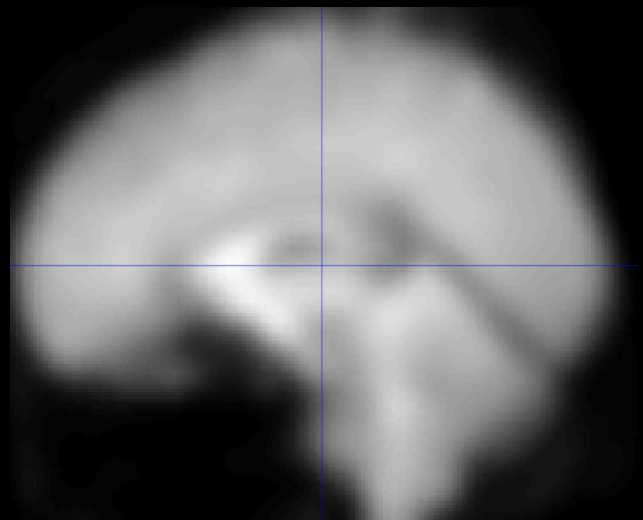
Review So Far

Image artifacts: How to preprocess and why

Quality assurance checks after each step

Parameters that stay the same, versus those you can modify





Review So Far

When would you want to use a smaller smoothing kernel? A larger smoothing kernel? Why?

When would you not want to use slice-timing correction? Or should it always be used, no matter what?

What is the argument for using the mean functional image as the Reference and the anatomical image as the Source? Would you ever want to swap them?

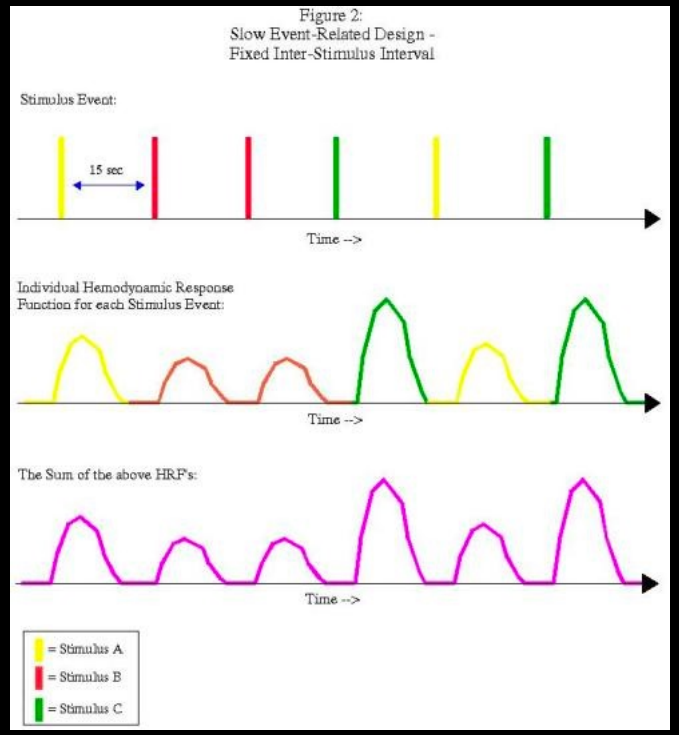
Review So Far

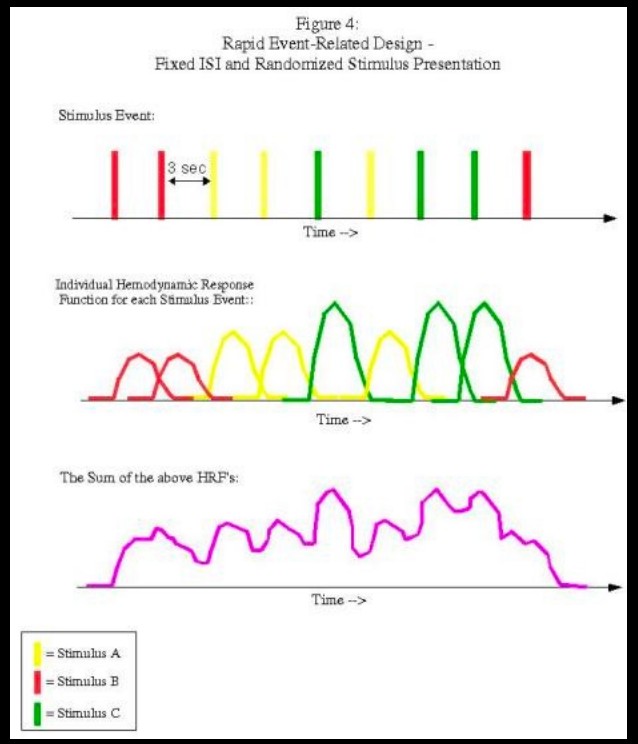
Experimental Design: Block vs. event-related

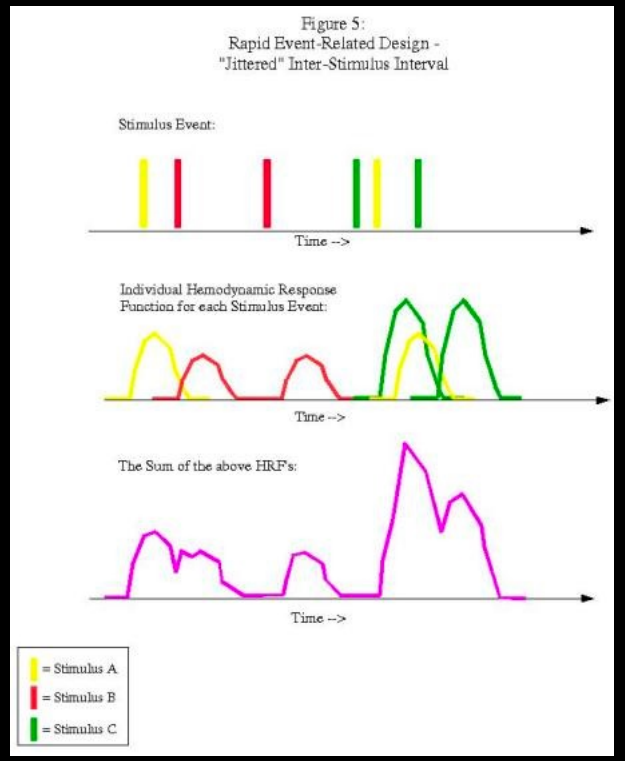
Slow vs. fast event-related

Jitter, collinearity, and power

Review So Far







What are the main advantages and disadvantages of each?

What is Reproducibility?

Replication: Arriving at the same result, using an independent dataset

Reproducibility: Ability to obtain the same result, using the same data and methods

Today: Use an open-access website to reproduce the results of that study

What is Reproducibility?

Replication and reproducibility have become more talked about in the past decade

Methods such as pre-registration can help to increase the likelihood of a successful replication

In other words: Minimize the researcher degrees of freedom

John Ioannides' 2005 Paper

PLOS MEDICINE

OPEN ACCESS

ESSAY

Why Most Published Research Findings Are False

John P. A. Ioannidis

Published: August 30, 2005 • https://doi.org/10.1371/journal.pmed.0020124

Like we discussed previously, controlling for Type I Error rates isn't enough to guarantee that results are real

John Ioannides' 2005 Paper

The crux of the paper rests on a formula called Positive Predictive Value (PPV)

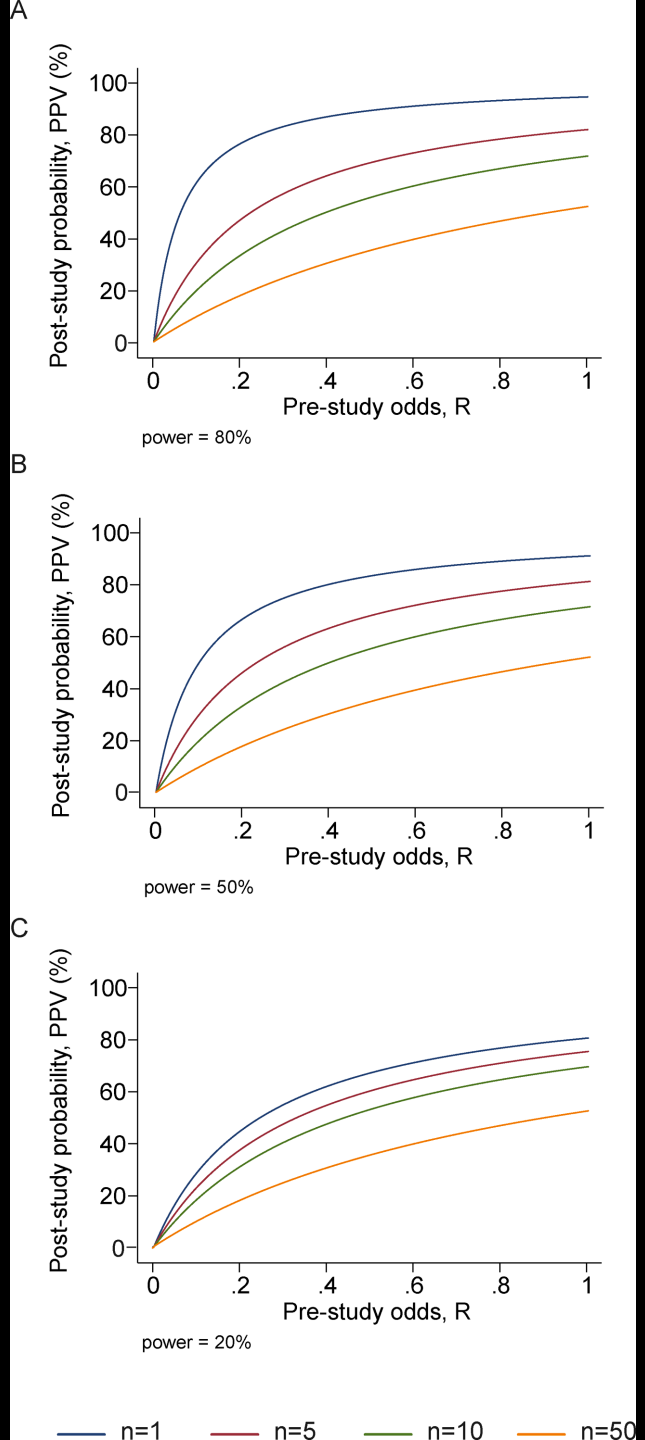
PPV is the number of true positives, divided by the number of rejected tests

John Ioannides' 2005 Paper

Let R be the ratio of the number of "true relationships" to "no relationships" (i.e., #HA/#H0)

Let α=alpha level, and (1-β)=Power

$$PPV = \frac{R(1-\beta)}{\alpha + R(1-\beta)}$$


A
Post-study probability, PPV (%)
100
80
60
40
20
0
0
.2
.4
.6
.8
1
Pre-study odds, R
power = 80%
B
Post-study probability, PPV (%)
100
80
60
40
20
0
0
.2
.4
.6
.8
1
Pre-study odds, R
power = 50%
C
Post-study probability, PPV (%)
100
80
60
40
20
0
0
.2
.4
.6
.8
1
Pre-study odds, R
power = 20%
n=1
n=5
n=10
n=50

John Ioannides' 2005 Paper

Lastly, let u=bias, the amount that a study is affected by biased practices (e.g., p-hacking); anything that tends to generate a positive results when it shouldn't

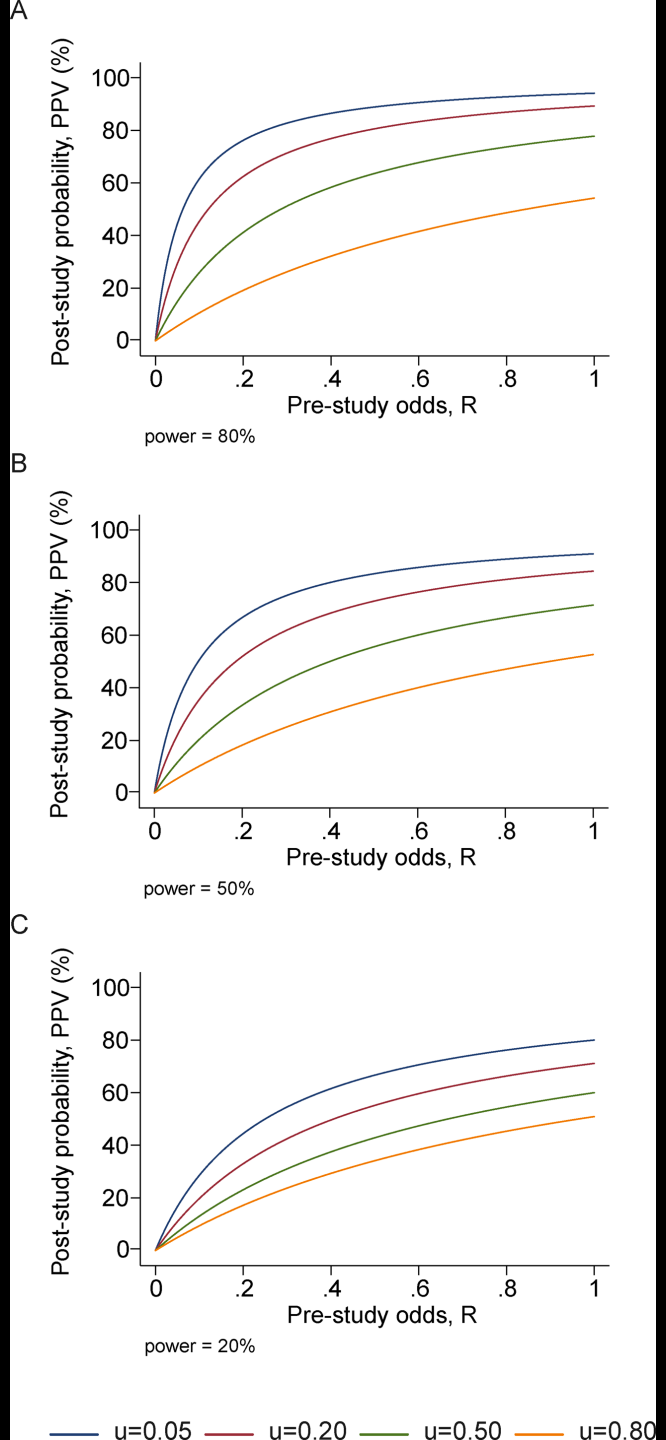

A
Post-study probability, PPV (%)
100
80
60
40
20
0
0
.2
.4
.6
.8
1
Pre-study odds, R
power = 80%
B
Post-study probability, PPV (%)
100
80
60
40
20
0
0
.2
.4
.6
.8
1
Pre-study odds, R
power = 50%
C
Post-study probability, PPV (%)
100
80
60
40
20
0
0
.2
.4
.6
.8
1
Pre-study odds, R
power = 20%
u=0.05
u=0.20
u=0.50
u=0.80

Table 4. PPV of Research Findings for Various Combinations of Power (1 – β), Ratio of True to Not-True Relationships *(R)*, and Bias *(u)*

1 – β	*R*	*u*	Practical Example	PPV
0.80	1:1	0.10	Adequately powered RCT with little bias and 1:1 pre-study odds	0.85
0.95	2:1	0.30	Confirmatory meta-analysis of good-quality RCTs	0.85
0.80	1:3	0.40	Meta-analysis of small inconclusive studies	0.41
0.20	1:5	0.20	Underpowered, but well-performed phase I/II RCT	0.23
0.20	1:5	0.80	Underpowered, poorly performed phase I/II RCT	0.17
0.80	1:10	0.30	Adequately powered exploratory epidemiological study	0.20
0.20	1:10	0.30	Underpowered exploratory epidemiological study	0.12
0.20	1:1,000	0.80	Discovery-oriented exploratory research with massive testing	0.0010
0.20	1:1,000	0.20	As in previous example, but with more limited bias (more standardized)	0.0015

Corollaries

Theory-driven vs. Data-driven

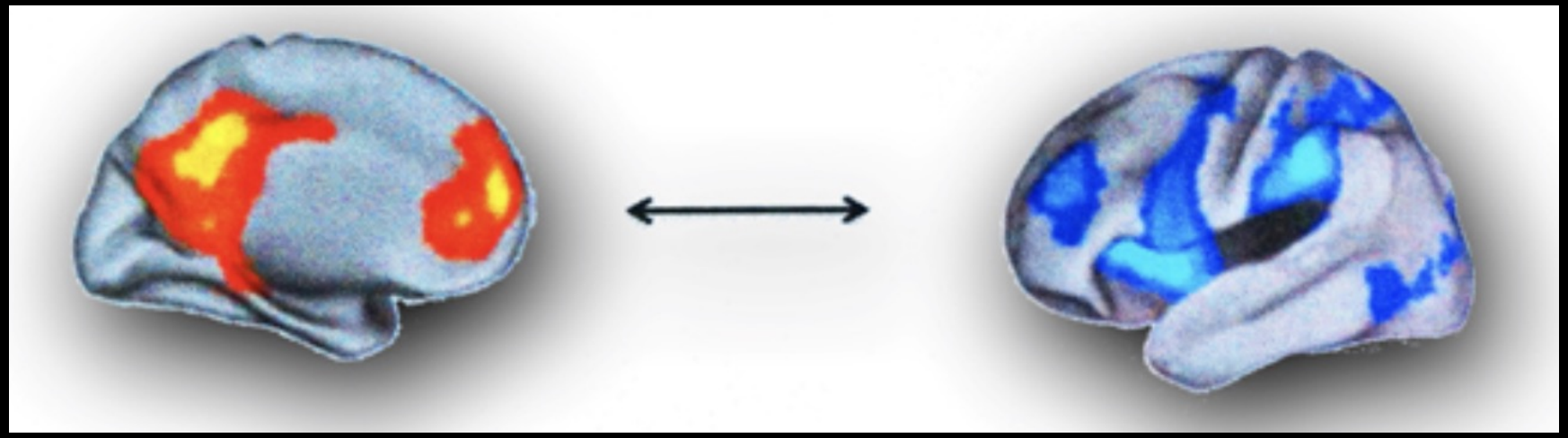
Theory-driven: Based on previous studies, reasonable predictions about what a region does

Data-driven: Uses the data itself to identify patterns, and then possibly create theories

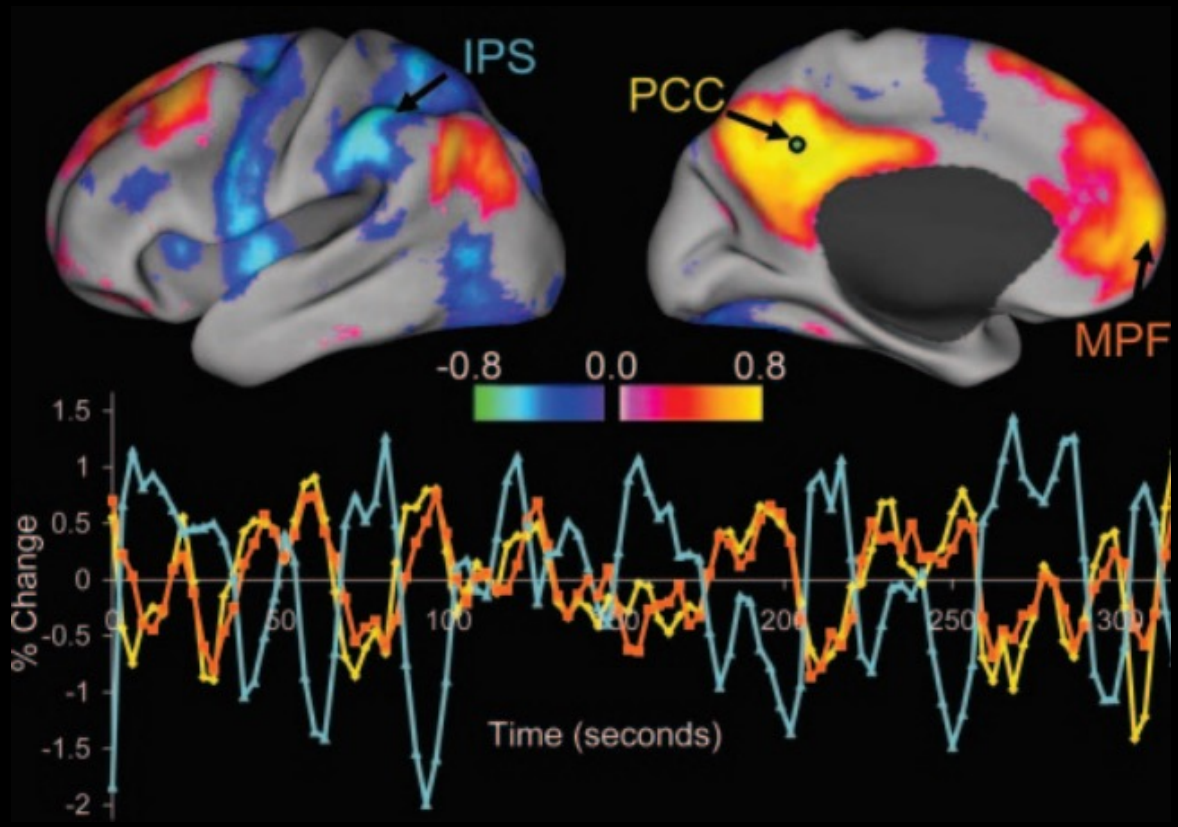
Theory-driven vs. Data-driven

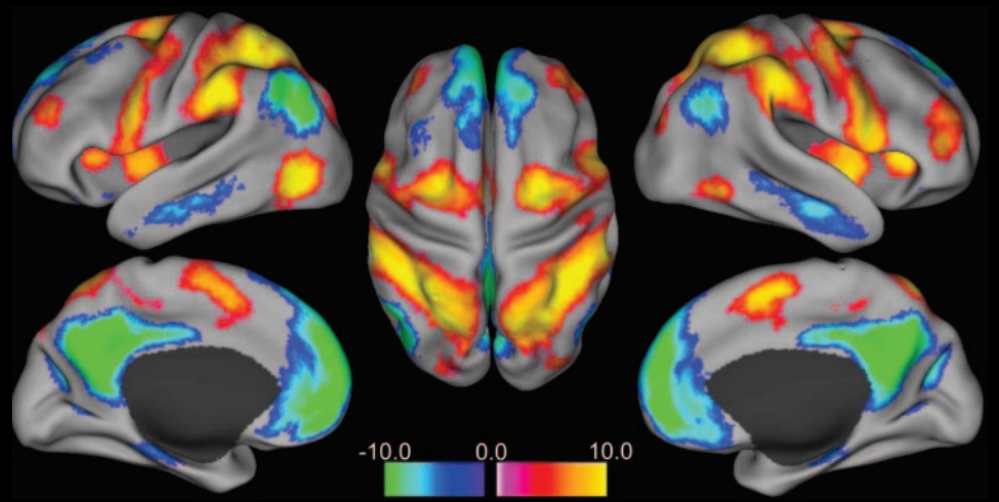
Theory-driven approaches have usually been more popular

Example: Fox et al., 2005



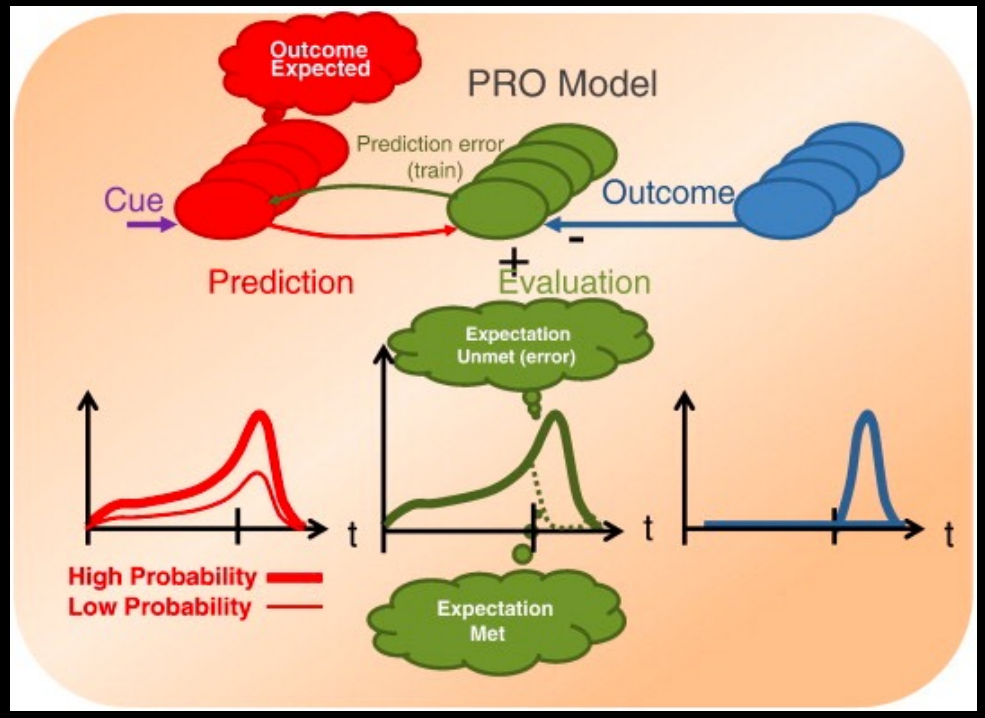
Theory-driven vs. Data-driven





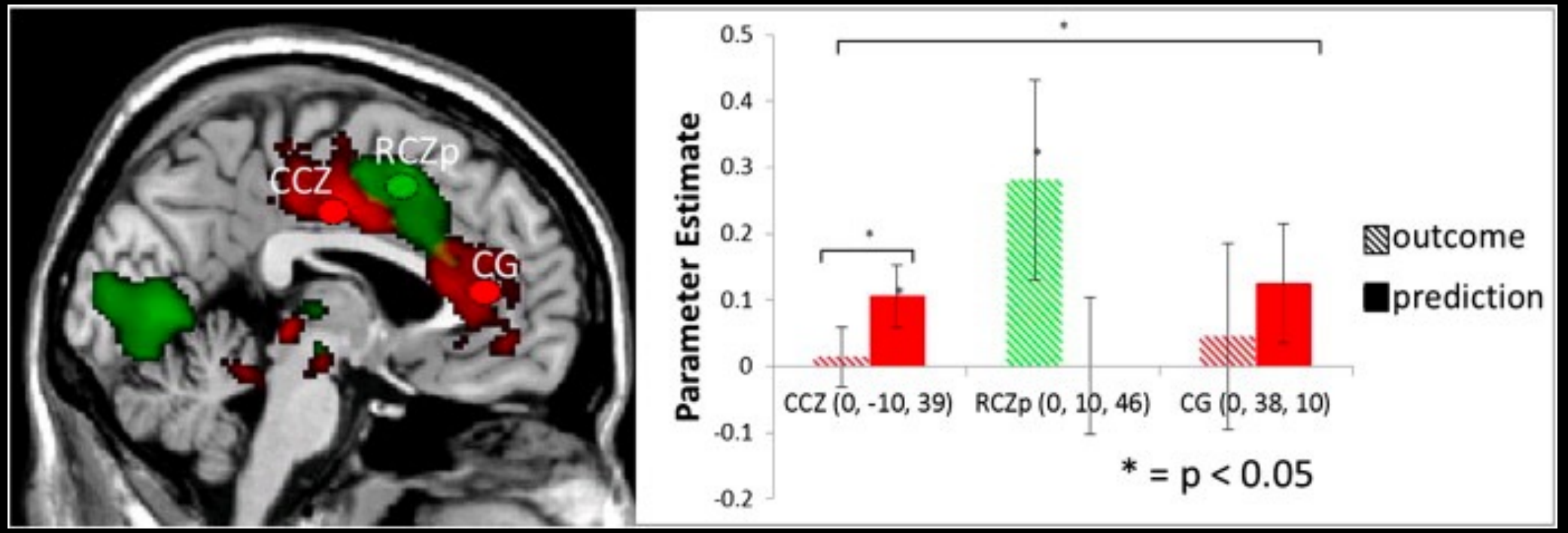
Theory-driven vs. Data-driven

Can also use a theoretical framework to create regressors



Theory-driven vs. Data-driven

$$Activit\,y_t = \sum_i \lfloor \Pr edicted\ Outcom\,e_{i,t} - Actual\ Outcom\,e_{i,t} \rfloor^+$$



Data-driven

Popular data-driven method: Multivariate Pattern Analysis

Usually requires a large number of observations or subjects

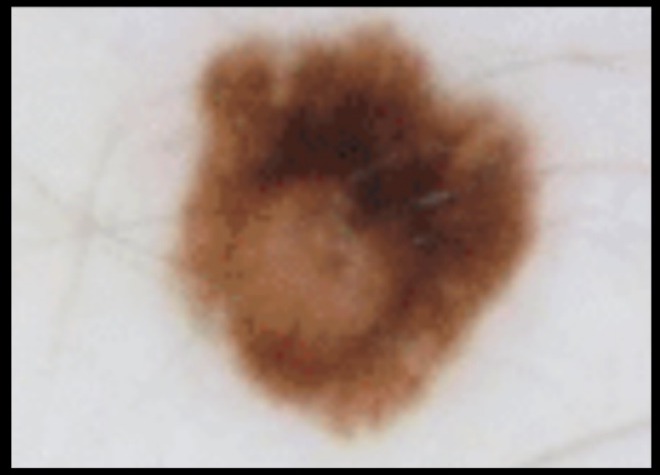
Other methods can be used: e.g., clustering

Machine Learning to Detect Skin Cancer

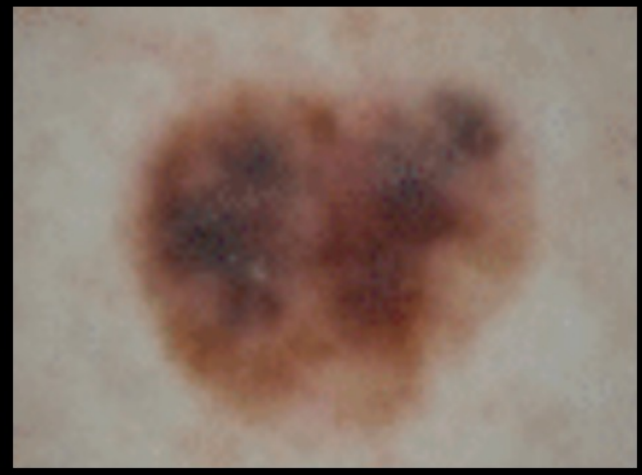
Features

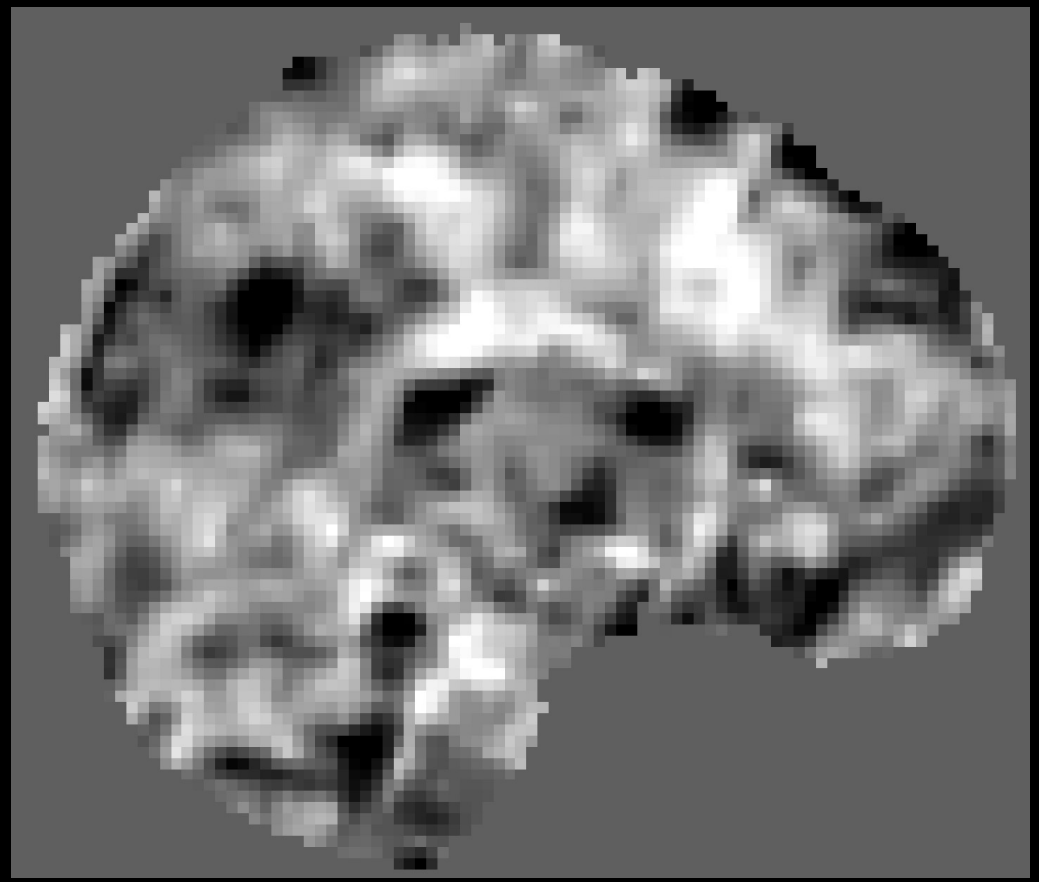
Machine Learning to Detect Skin Cancer

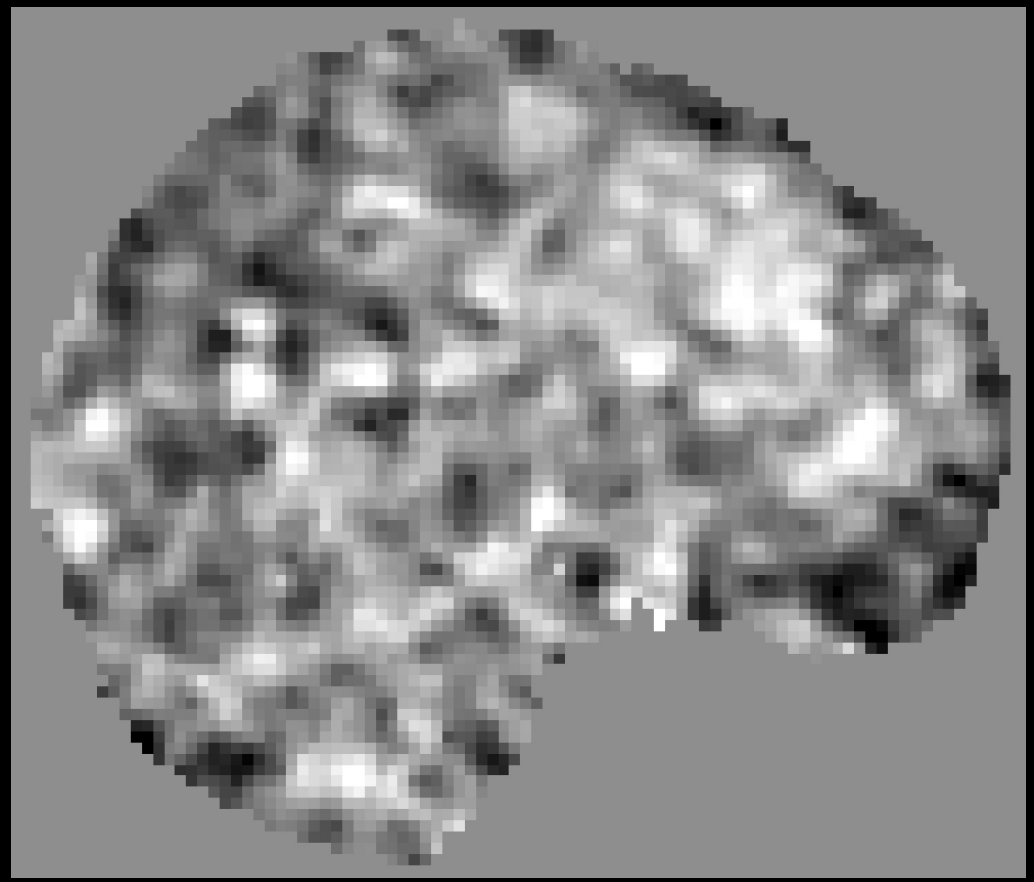
False Positive



False Negative

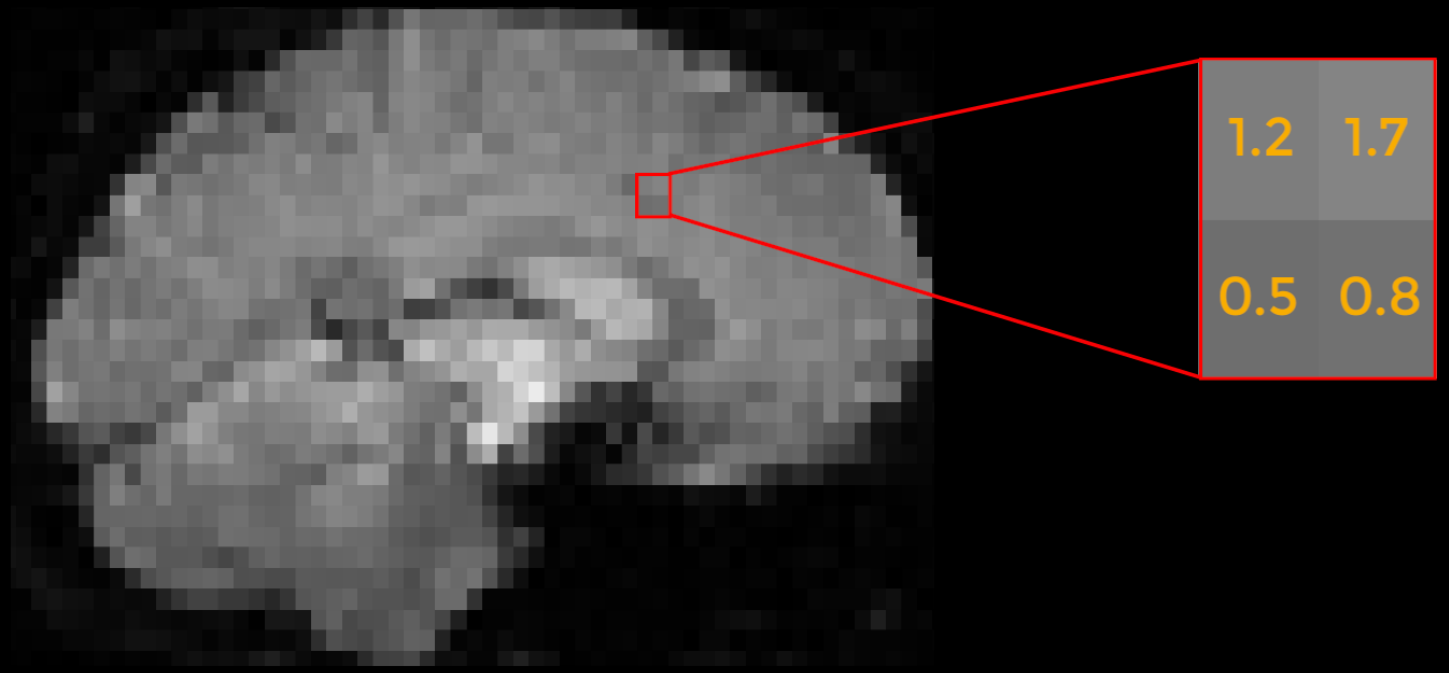


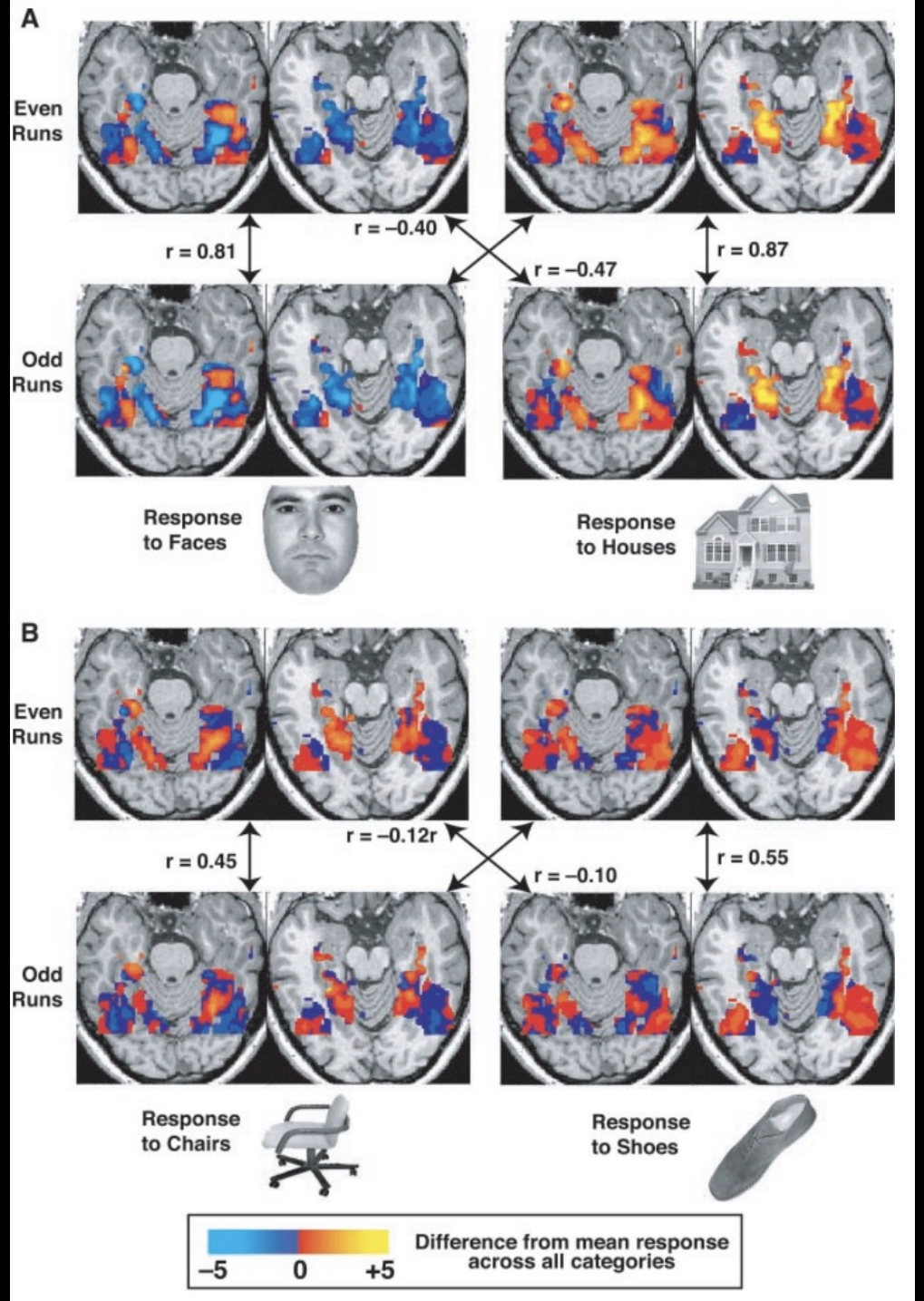




MVPA

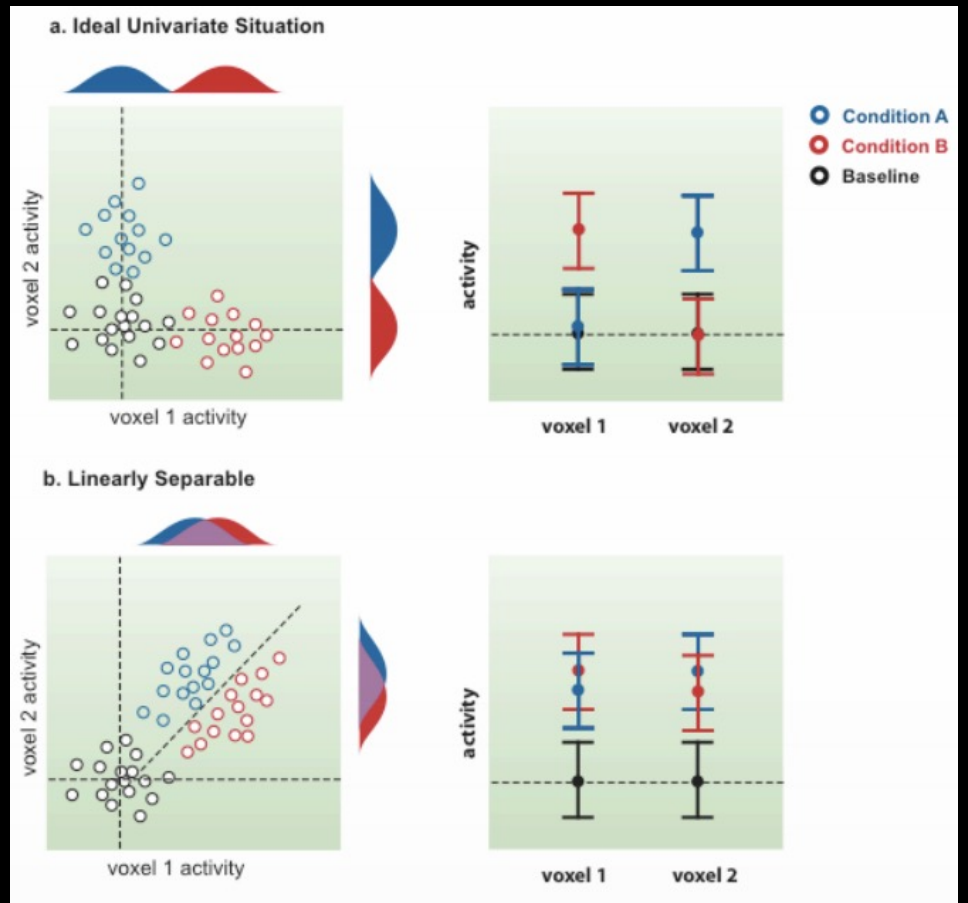
Applied to fMRI Data

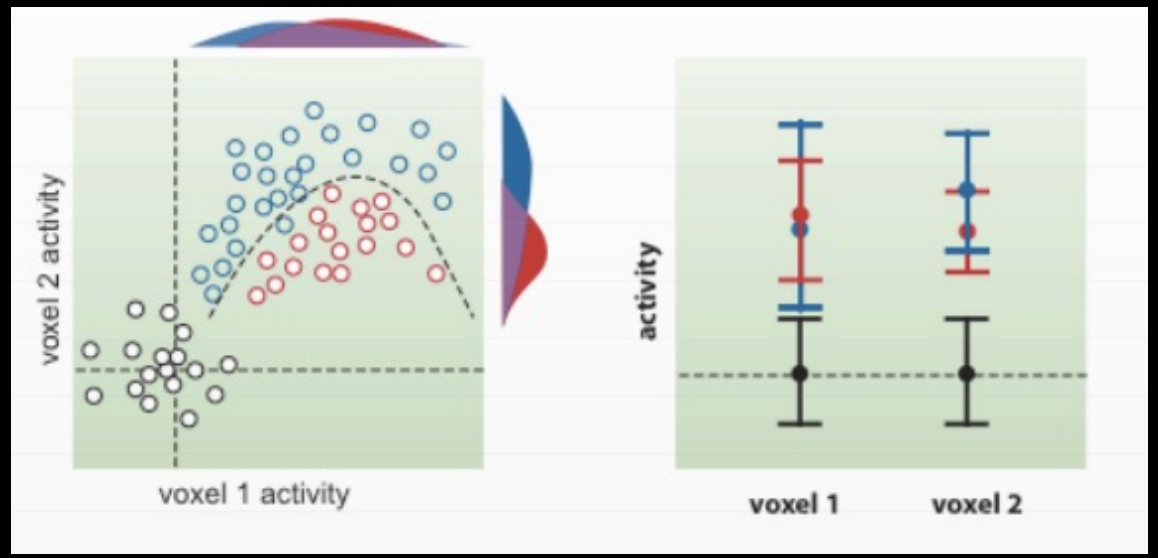



A
Even Runs
Odd Runs
r = 0.81
r = –0.40
r = –0.47
r = 0.87
Response to Faces
Response to Houses
B
Even Runs
Odd Runs
r = 0.45
r = –0.12r
r = –0.10
r = 0.55
Response to Chairs
Response to Shoes
–5
0
+5
Difference from mean response across all categories

MVPA

Use Support Vector Machines to classify beta maps





OpenNEURO

Versions

580166c2cce88d000aa33631 2018-07-13

00001 2018-07-13

Visual object recognition

uploaded by Chris Gorgolewski on 2016-10-14 - almost 4 years ago
last modified on 2018-07-14 - about 2 years ago
authored by Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P.
807 15359

Download Analyze on brainlife.io

OpenNeuro Accession Number: ds000105

Files: 1095, **Size:** 1.75GB, **Subjects:** 6, **Session:** 1

Available Tasks: object viewing

Available Modalities: T1w, bold

README

This dataset was obtained from the OpenfMRI project (http://www.openfmri.org).
Accession #: ds105
Description: Visual object recognition

Please cite the following references if you use these data:

Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science, 293(5539):2425-30

Hanson, S.J., Matsuka, T., Haxby, J.V. (2004). Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area? Neuroimage. 23(1):156-66

O'Toole, A.J., Jiang, F., Abdi, H., Haxby, J.V. (2005). Partially distributed representations of objects and faces in ventral temporal cortex. J Cogn Neurosci, 17(4):580-90

Release history:
10/12/2011: initial release

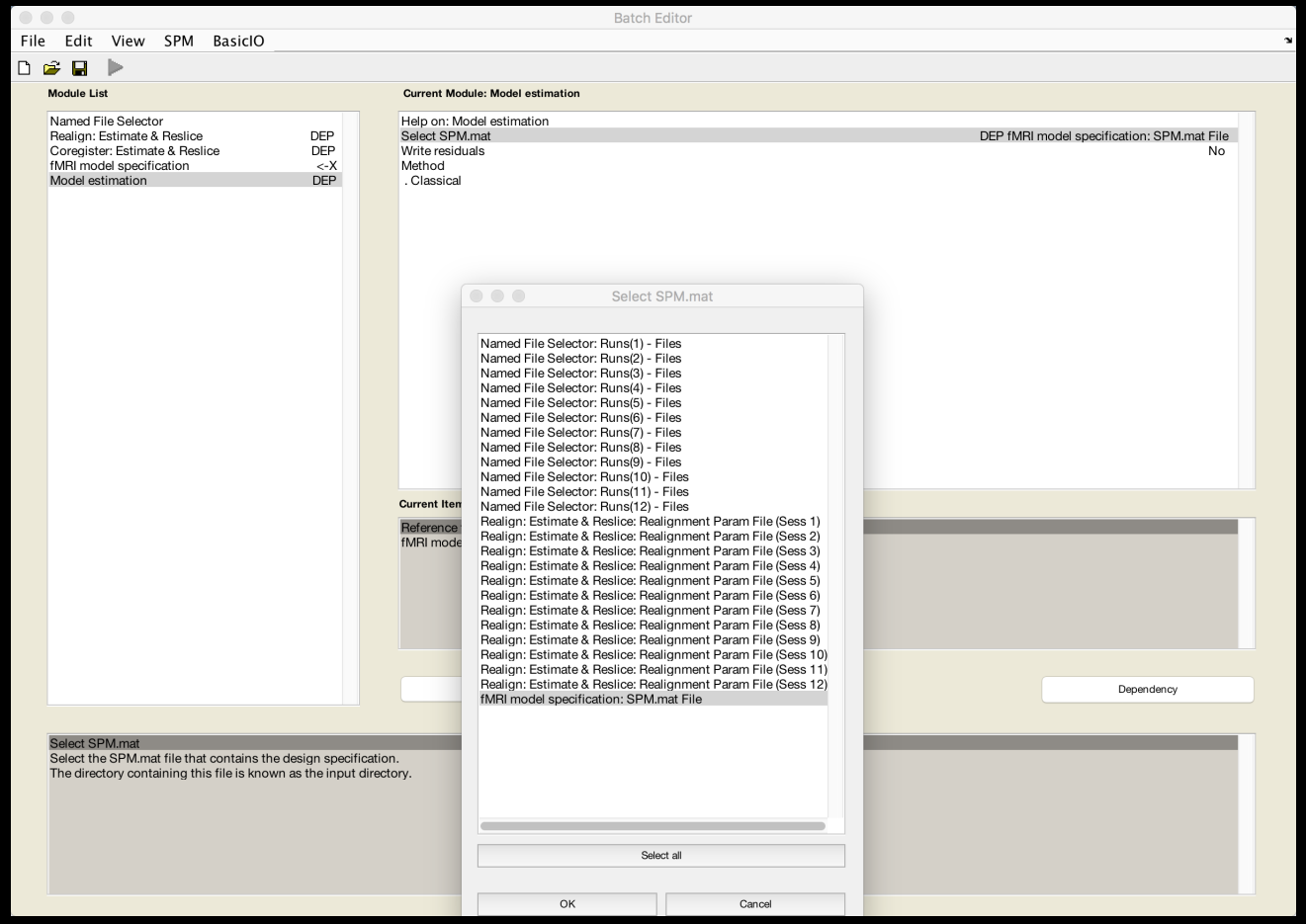
BIDS Validation

Invalid

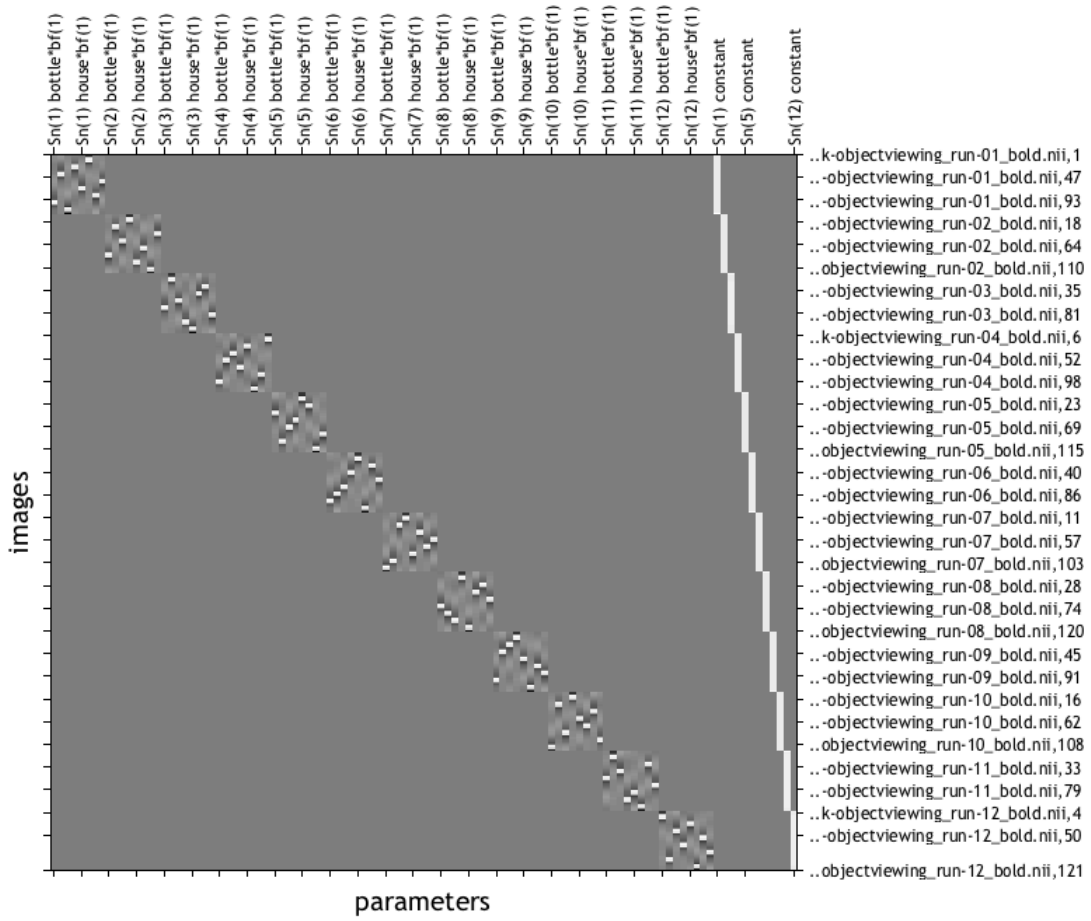
Dataset File Tree

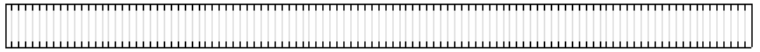
Visual object recognition
- CHANGES — DOWNLOAD VIEW
- dataset_description.json — DOWNLOAD VIEW
- README — DOWNLOAD VIEW
- task-objectviewing_bold.json — DOWNLOAD VIEW
- derivatives
- sub-1
- sub-2
- sub-3
- sub-4
- sub-5
- sub-6

Preprocessing an MVPA experiment is similar to fMRI, with the exception of smoothing



Statistical analysis: Design





(gray → β not uniquely specified)

parameter estimability

Design description...

Basis functions :	hrf
Number of sessions :	12
Trials per session :	8 8 8 8 8 8 8 8 8 8 8 8
Interscan interval :	2.50 {s}
High pass Filter :	[min] Cutoff: 128 {s}
Global calculation :	mean voxel value
Grand mean scaling :	session specific
Global normalisation :	None

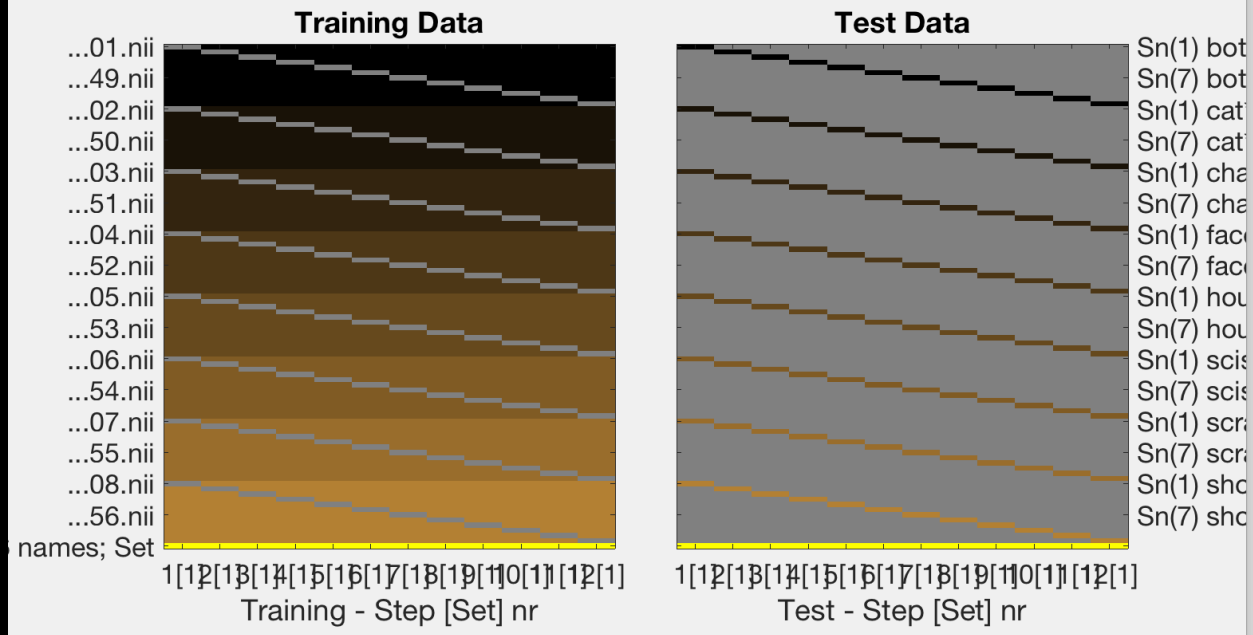
```
labelname1 = 'bottle';
labelname2 = 'cat';
labelname3 = 'chair';
labelname4 = 'face';
labelname5 = 'house';
labelname6 = 'scissors';
labelname7 = 'scrambledpix';
labelname8 = 'shoe';
```

```
% Make sure the decoding toolbox and your favorite software (SPM or AFNI)
% are on the Matlab path (e.g. addpath('/home/decoding_toolbox') )
% addpath('$ADD FULL PATH TO TOOLBOX AS STRING OR MAKE THIS LINE A COMMENT IF IT IS ALREADY$')
% addpath('$ADD FULL PATH TO TOOLBOX AS STRING OR MAKE THIS LINE A COMMENT IF IT IS ALREADY$')

% Set defaults
cfg = decoding_defaults;
cfg.results.overwrite = 1;

% Set the analysis that should be performed (default is 'searchlight')
cfg.analysis = 'ROI';
cfg.searchlight.radius = 3; % use searchlight of radius 3 (by default in voxels), see more details below

% Set the output directory where data will be saved, e.g. 'c:\exp\results\buttonpress'
cfg.results.dir = [pwd '/SPM_Results_1'];

% Set the filepath where your SPM.mat and all related betas are, e.g. 'c:\exp\glm\model_button'
beta_loc = [pwd '/SPM_Results_1'];

% Set the filename of your brain mask (or your ROI masks as cell matrix)
% for searchlight or wholebrain e.g. 'c:\exp\glm\model_button\mask.img' OR
% for ROI e.g. {'c:\exp\roi\roimaskleft.img', 'c:\exp\roi\roimaskright.img'}
% You can also use a mask file with multiple masks inside that are
% separated by different integer values (a "multi-mask")
cfg.files.mask = [pwd '/Haxby_Masks/sub-1_mask4_vt.nii'];

% Set the label names to the regressor names which you want to use for
% decoding, e.g. 'button left' and 'button right'
% don't remember the names? -> run display_regressor_names(beta_loc)
  labelname1 = 'bottle';
  labelname2 = 'cat';
  labelname3 = 'chair';
  labelname4 = 'face';
  labelname5 = 'house';
  labelname6 = 'scissors';
  labelname7 = 'scrambledpix';
```

Figure 1: Decoding Design

File Edit View Insert Tools Desktop Window Help

TDT - Decoding details
Filestart: /Users/ajahn/Desktop/Haxby_Data/SPM_Results_1/beta_00
Results: /Users/ajahn/Desktop/Haxby_Data/SPM_Results_1
Start: 28-Sep-2020 15:36:29, End: 28-Sep-2020 15:36:32



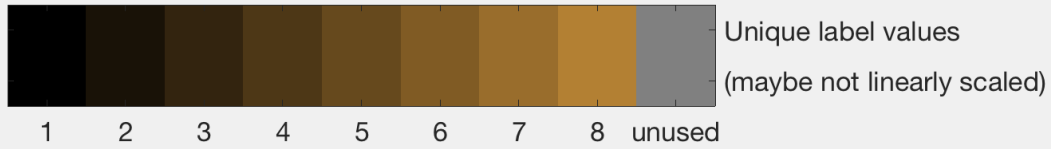
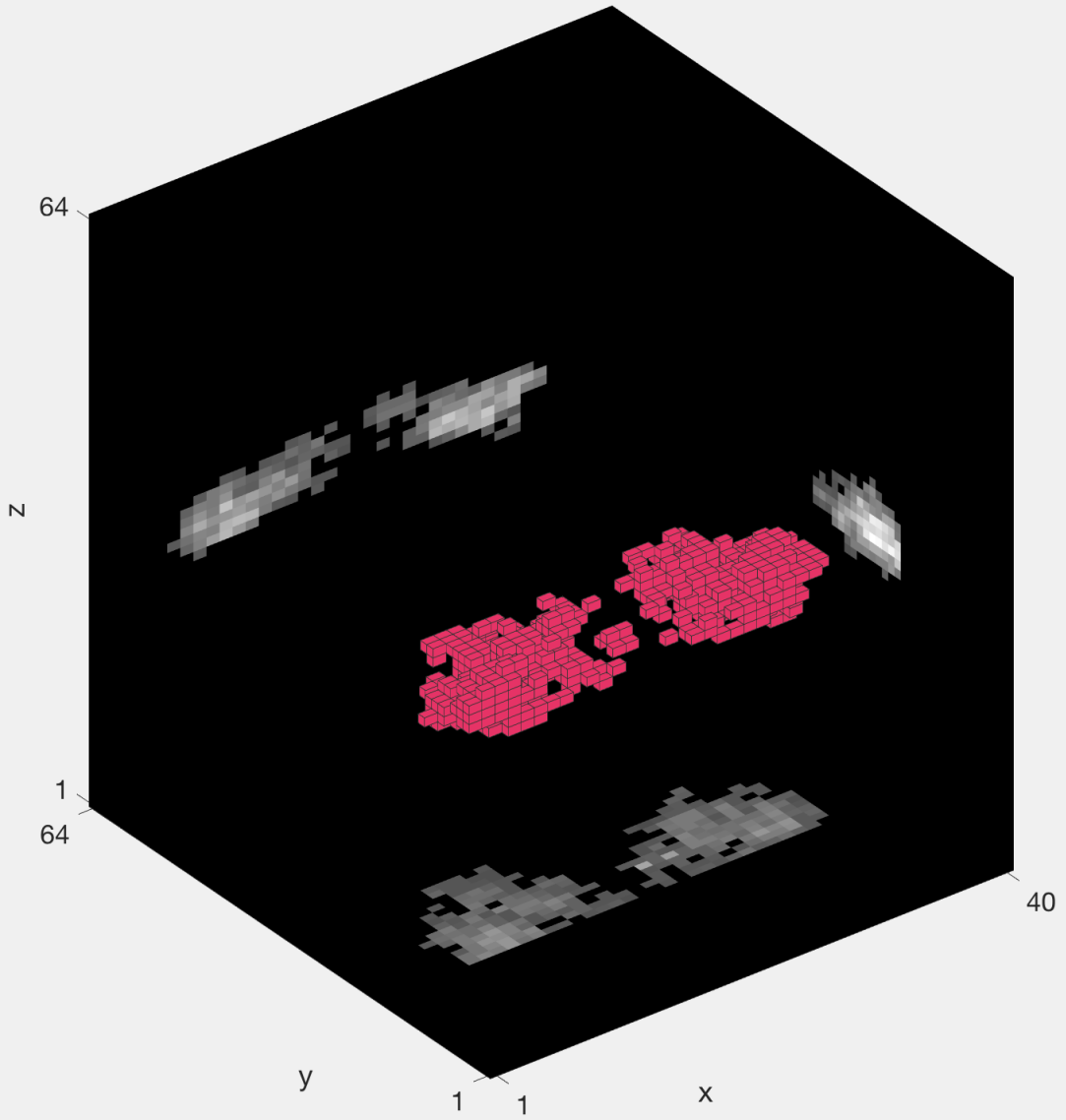
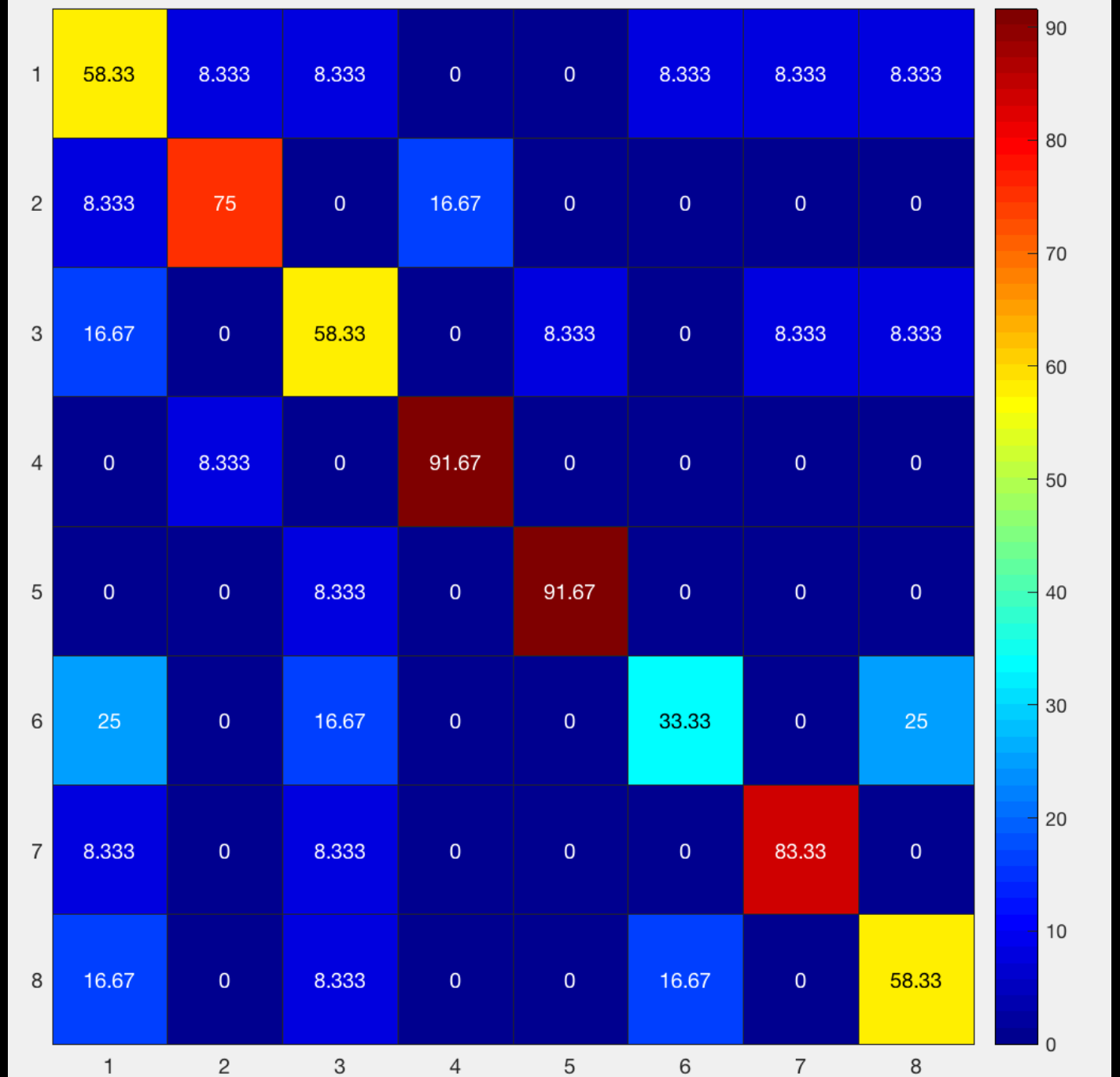


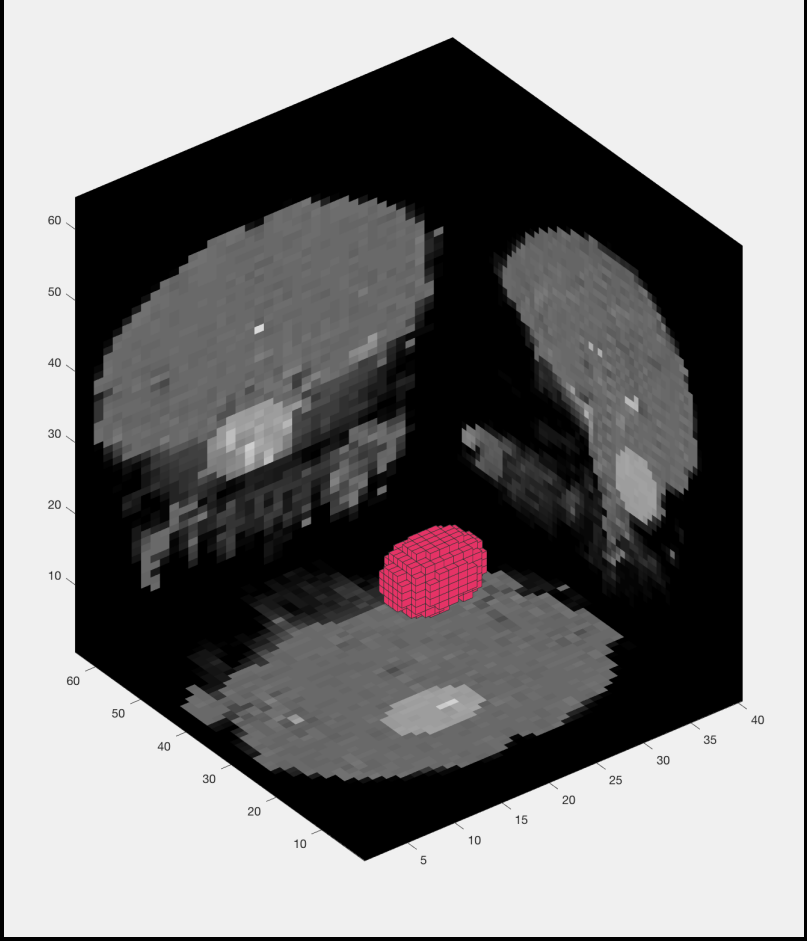
Figure 2: Online ROI, showing 1/500 steps (cfg.plot_selected_voxels=0 for more speed)

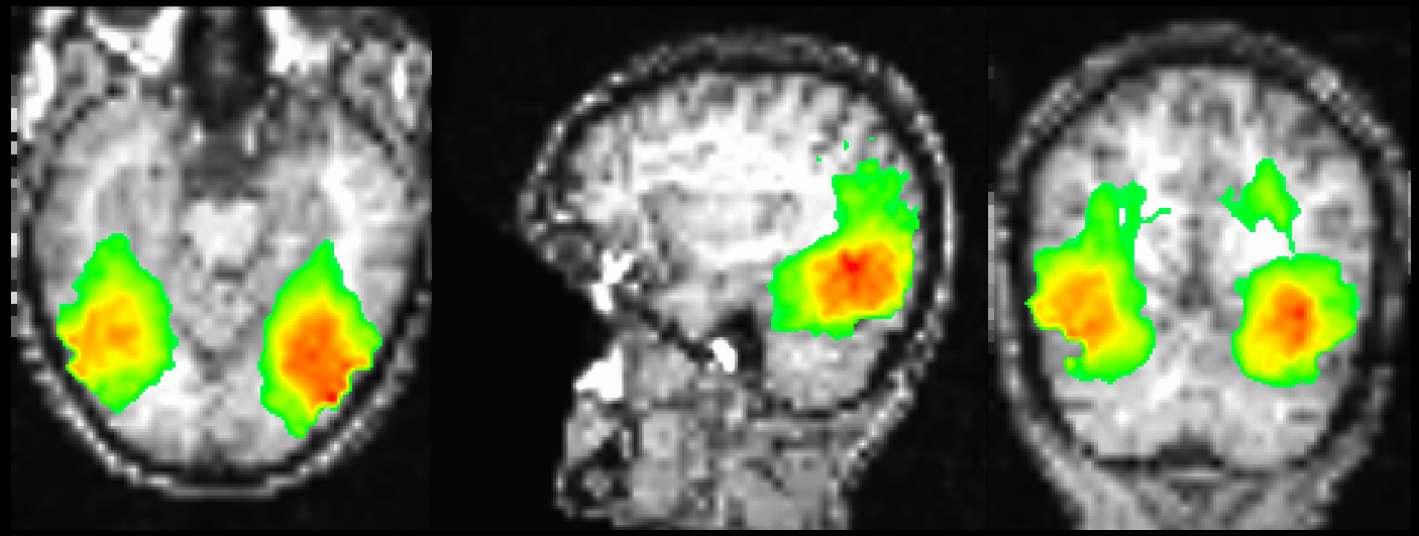
File Edit View Insert Tools Desktop Window Help




1 58.33 8.333 8.333 0 0 8.333 8.333 8.333
2 8.333 75 0 16.67 0 0 0 0
3 16.67 0 58.33 0 8.333 0 8.333 8.333
4 0 8.333 0 91.67 0 0 0 0
5 0 0 8.333 0 91.67 0 0 0
6 25 0 16.67 0 0 33.33 0 25
7 8.333 0 8.333 0 0 0 83.33 0
8 16.67 0 8.333 0 0 16.67 0 58.33
1 2 3 4 5 6 7 8
90 80 70 60 50 40 30 20 10 0

Searchlight Analysis

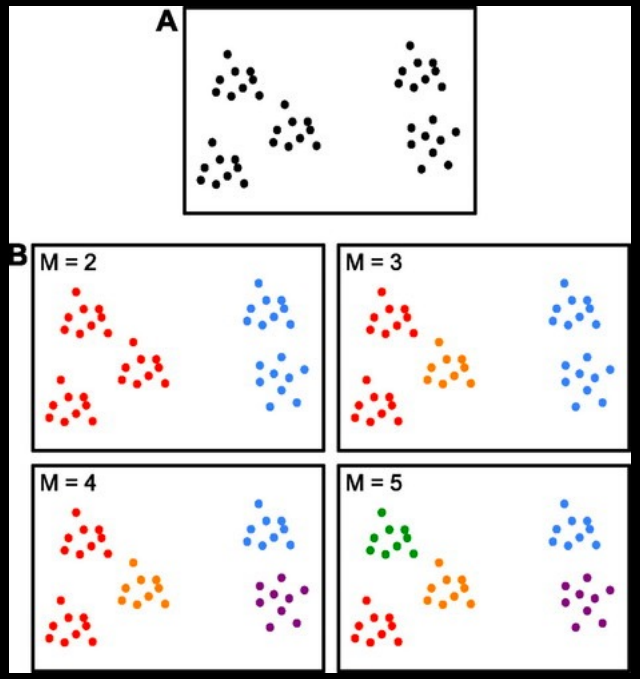




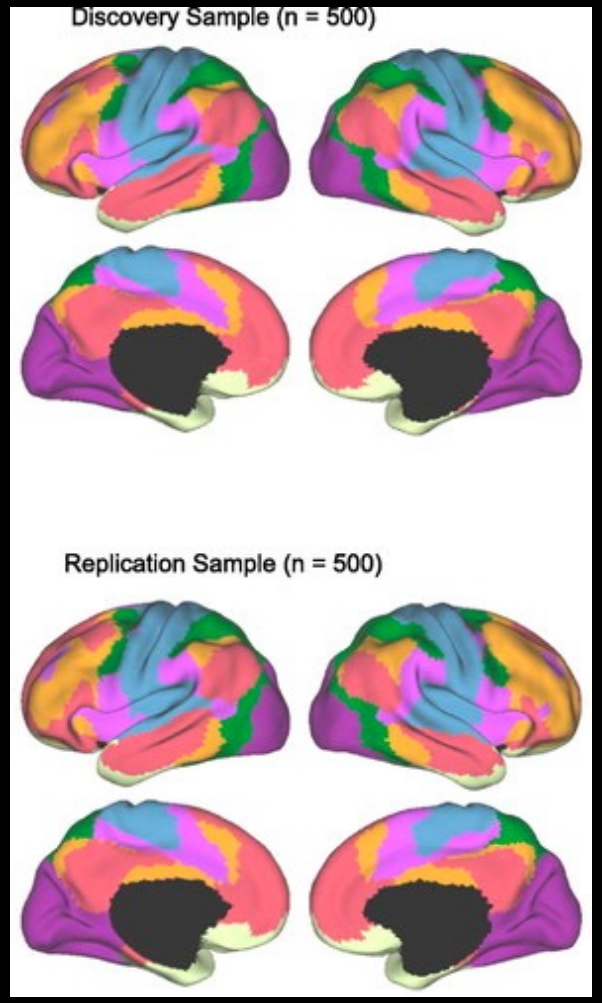
Data-driven

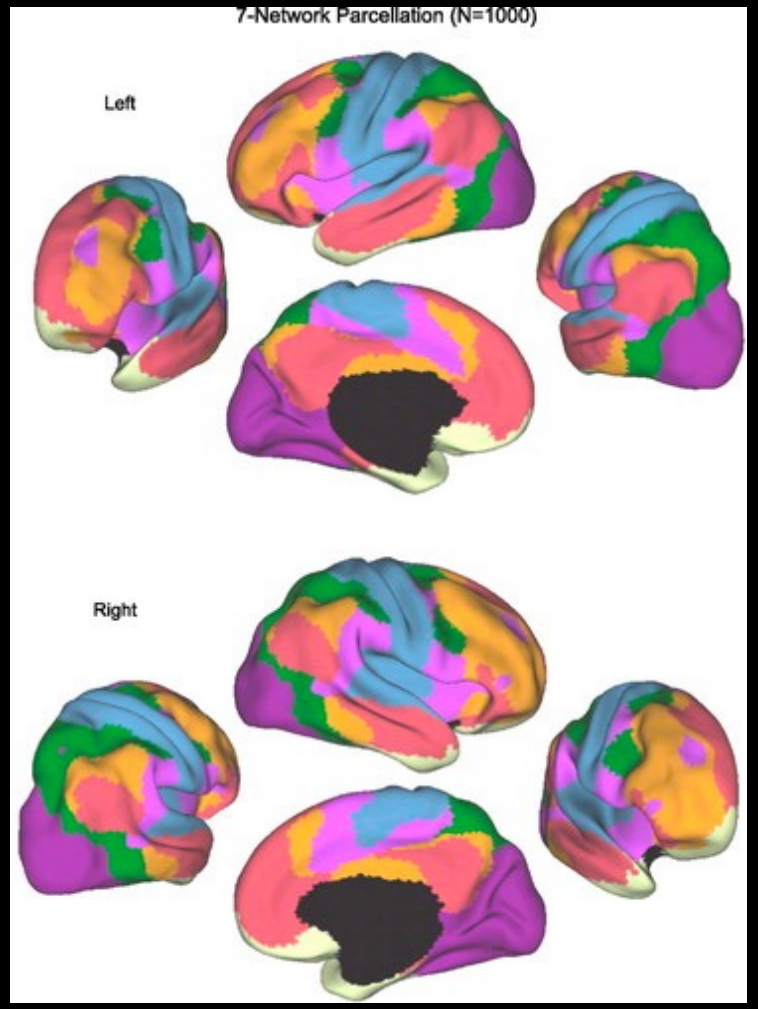
Example of Clustering: Yeo et al., 2011

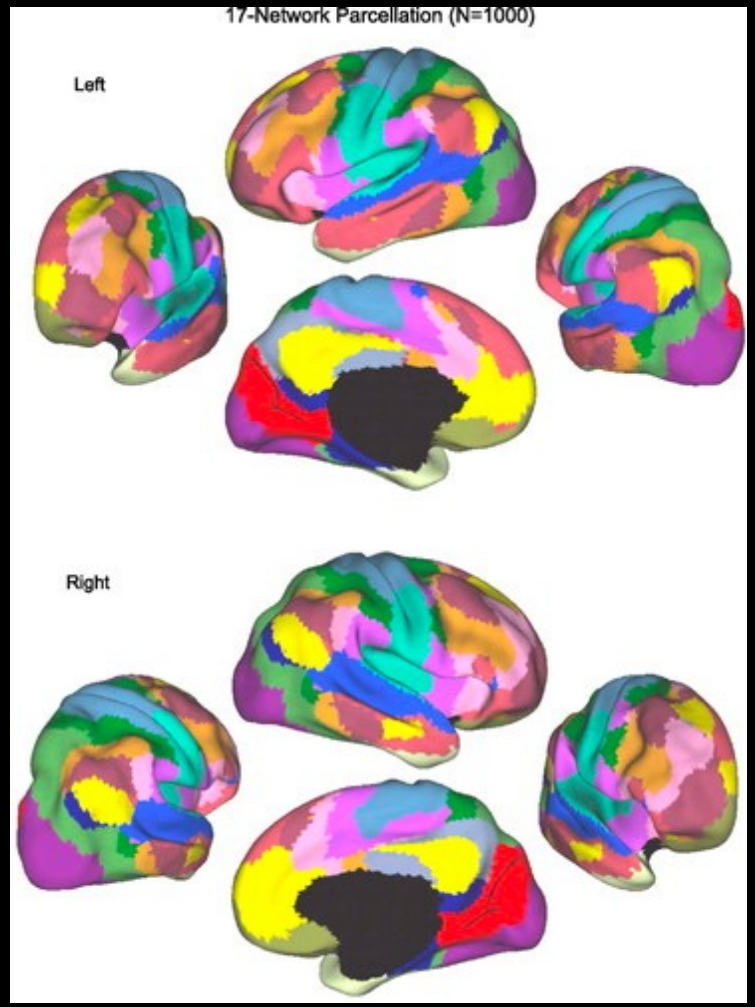
Used clustering to reveal intrinsic FC networks



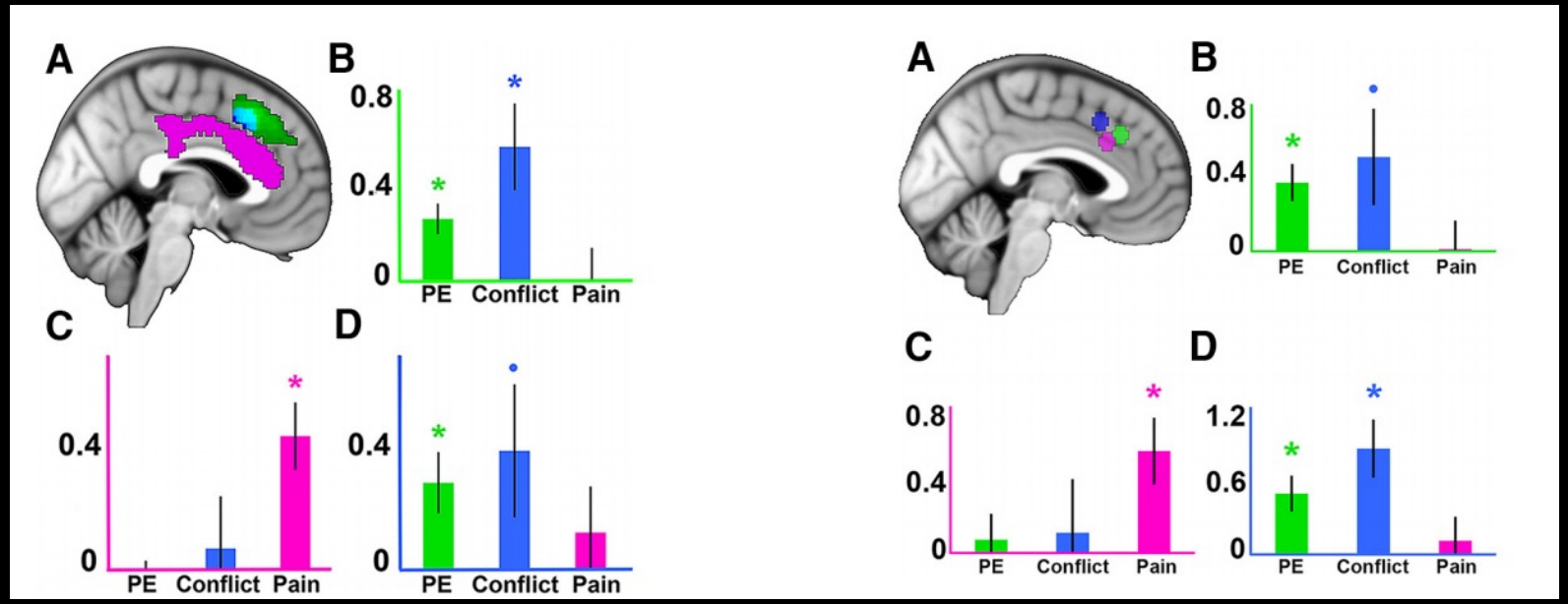
Data-driven

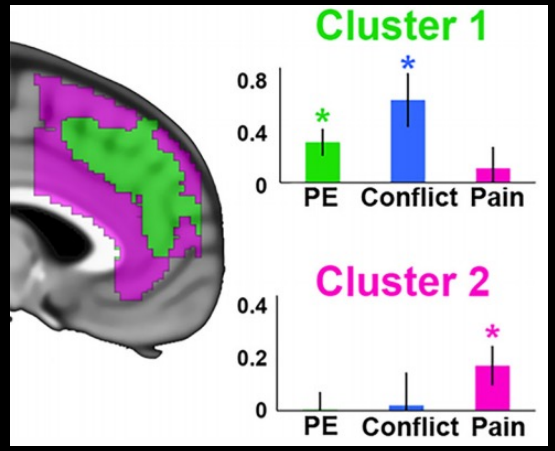

Discovery Sample (n = 500)
Replication Sample (n = 500)


7-Network Parcellation (N=1000)
Left
Right


17-Network Parcellation (N=1000)
Left
Right

Data-driven


A
B
0.8
0.4
0
PE
Conflict
Pain
C
D
0.4
0
PE
Conflict
Pain
A
B
0.8
0.4
0
PE
Conflict
Pain
C
D
0.8
0.4
0
1.2
0.6
0
PE
Conflict
Pain


Cluster 1
0.8
0.4
0
PE
Conflict
Pain
Cluster 2
0.4
0.2
0
PE
Conflict
Pain

Summary: Comparison of each Approach

Theory-driven: Builds upon previous research; logically coherent; consecutive results; can be done with relatively few subjects

Data-driven: Can leverage large open-access datasets to answer new questions; depending on number of subjects and trials, has huge power

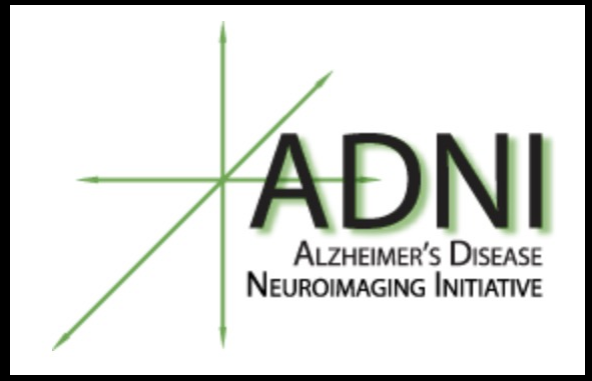
Word of Caution

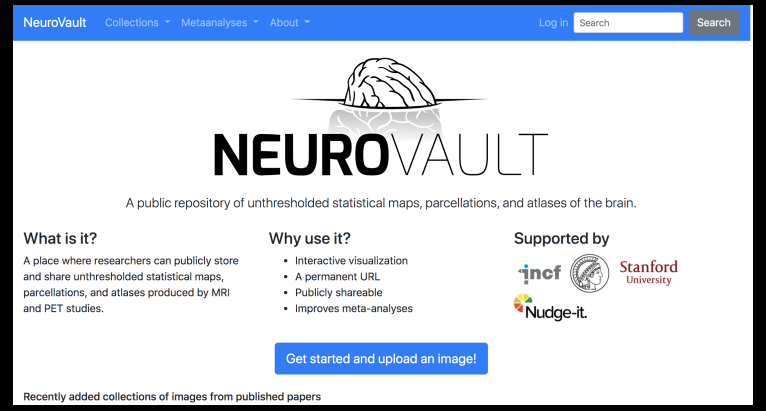
Does not protect you from statistical fallacies discussed earlier

Large datasets provide more power; also,
more opportunities for fishing and p-hacking

Tools for Reproducibility

Open-access repositories





Overview of Openneuro

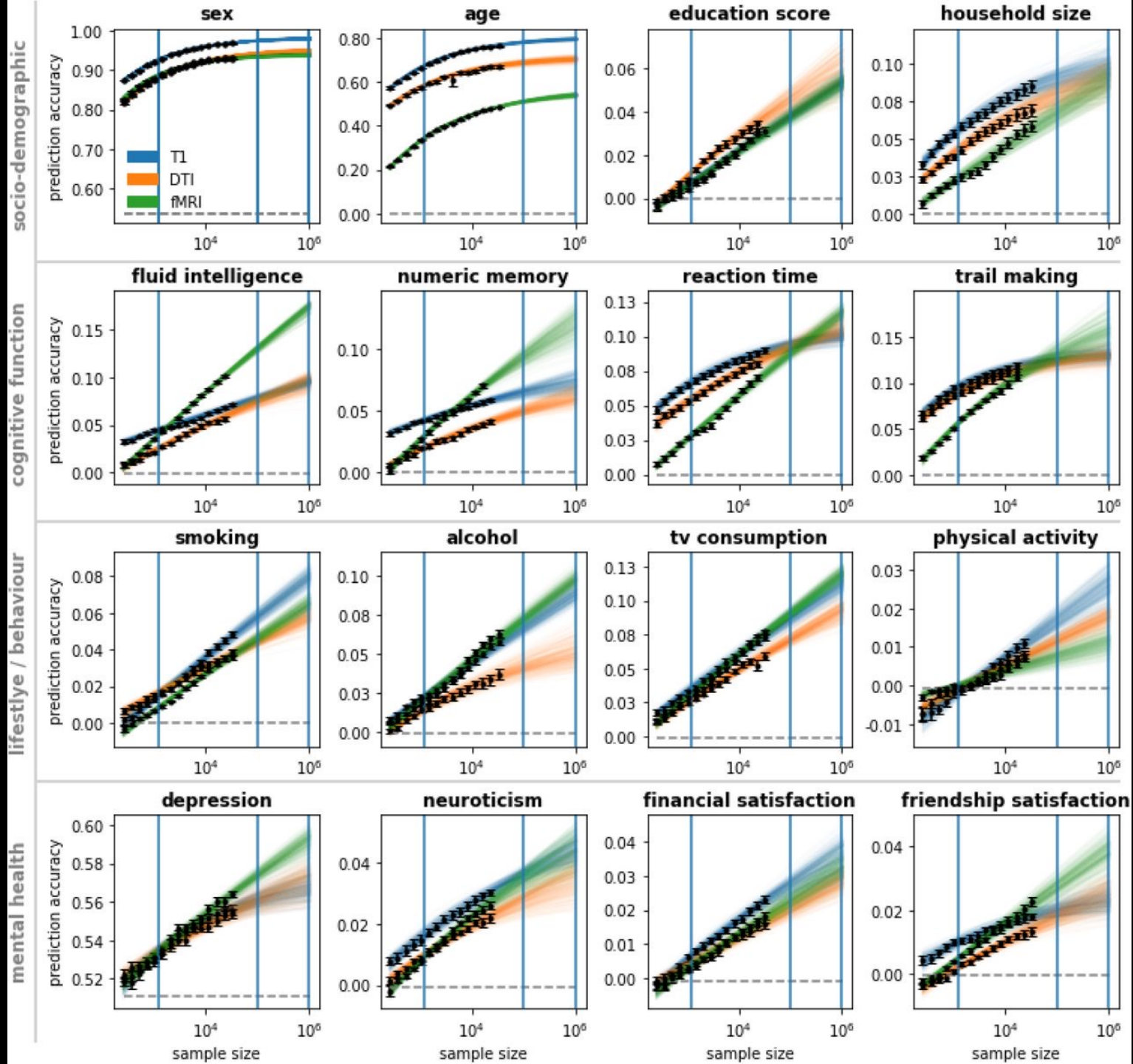
Growing use of large databases

ABCD (n=11,874)

Human Connectome Project (n=1,200)

UK Biobank (n=36,735)

Typical fMRI study n = ~25

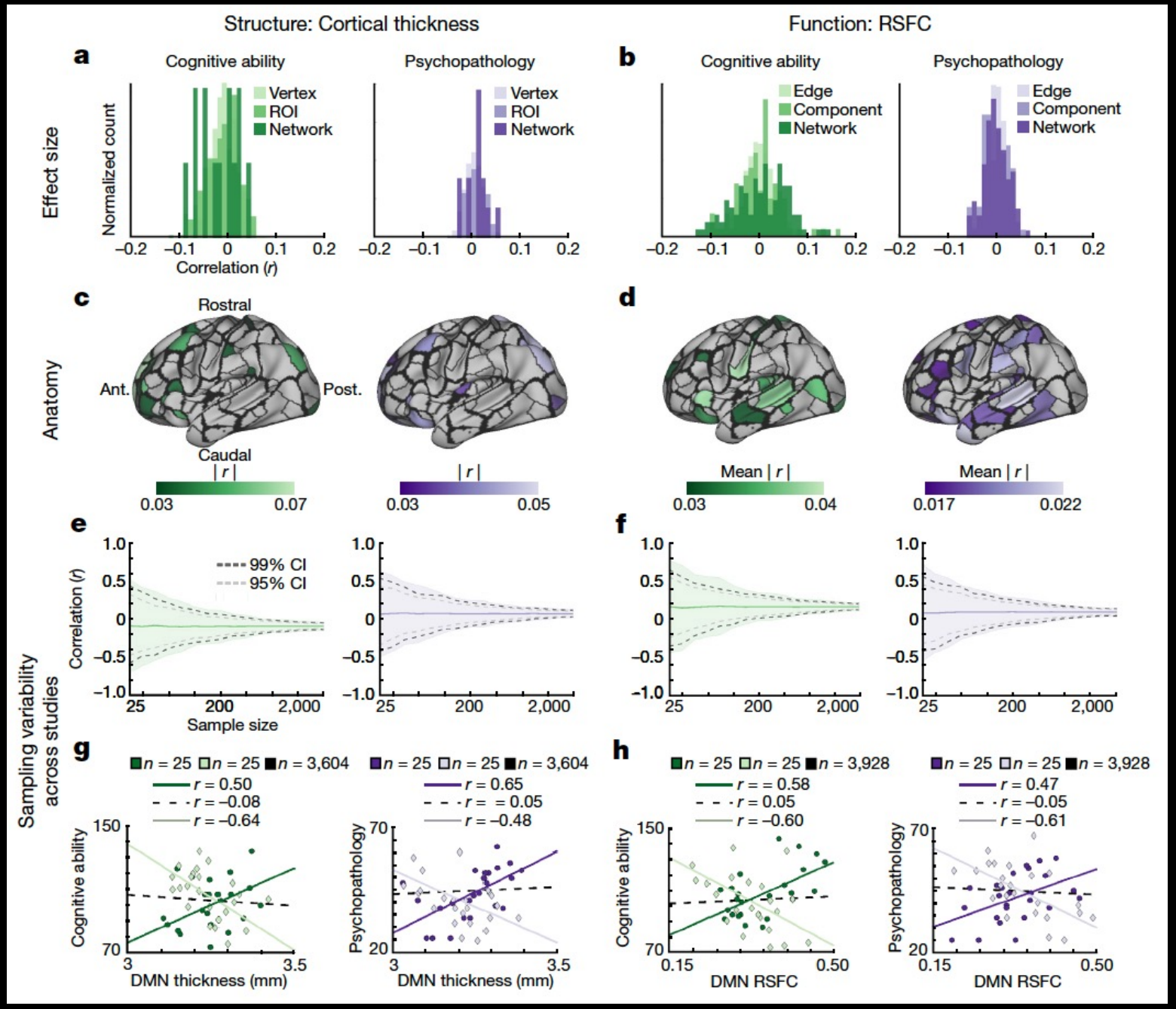

socio-demographic
cognitive function
lifestlye / behaviour
mental health
sex
age
education score
household size
fluid intelligence
numeric memory
reaction time
trail making
smoking
alcohol
tv consumption
physical activity
depression
neuroticism
financial satisfaction
friendship satisfaction
prediction accuracy
sample size
T1
DTI
fMRI

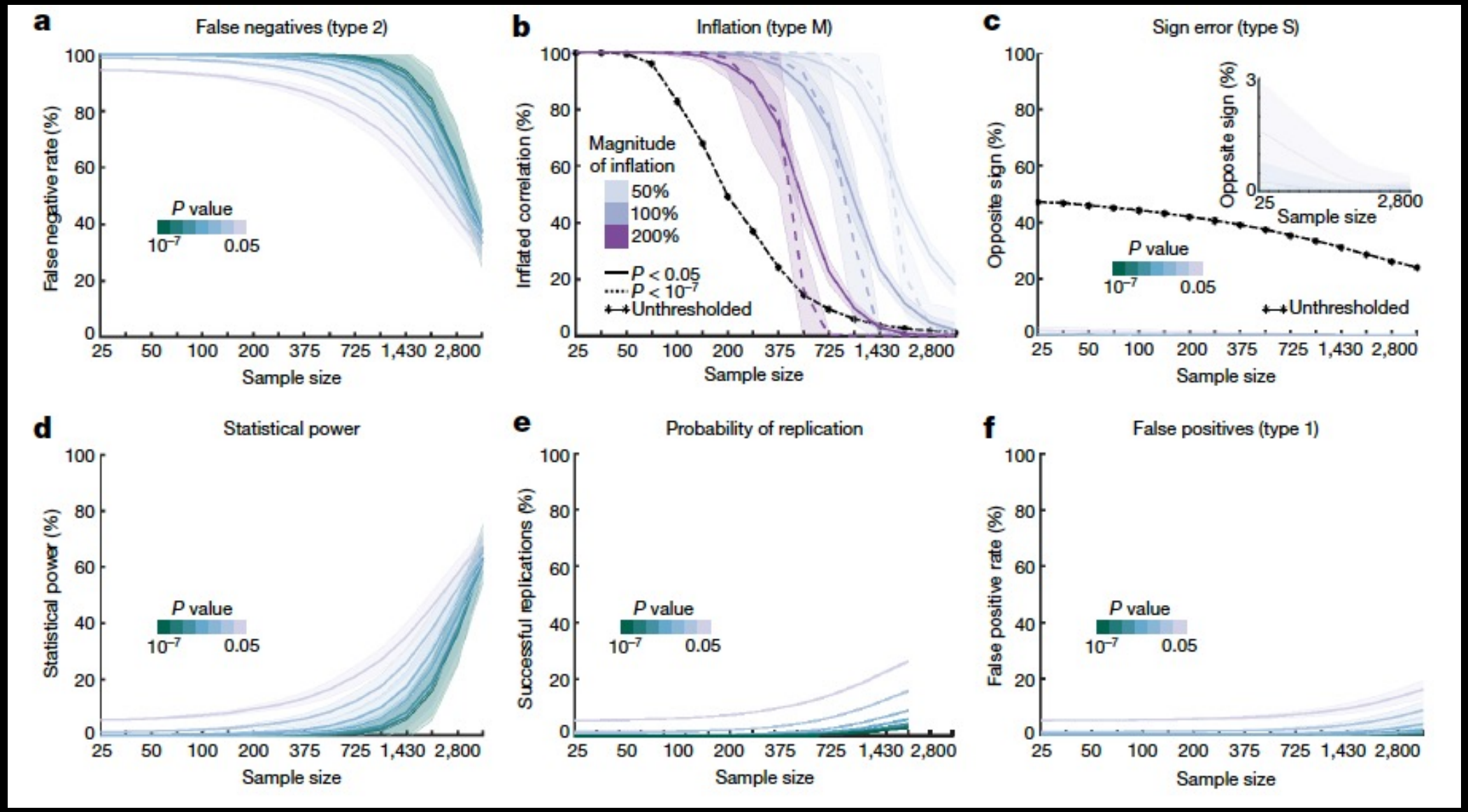
Large database issues

Can be collected from different sites

Changes in acquisition protocols over time
(e.g., ADNI phases 1-3)

Current study: Look at ABCD, HCP, and UK Biobank


Structure: Cortical thickness
Function: RSFC
a
b
Cognitive ability
Psychopathology
Effect size
Normalized count
Vertex
ROI
Network
Edge
Component
Network
Correlation (r)
c
d
Anatomy
Rostral
Ant.
Post.
Caudal
| r |
Mean | r |
0.03
0.07
0.05
0.04
0.017
0.022
e
f
Sampling variability across studies
Correlation (r)
99% CI
95% CI
Sample size
g
h
n = 25 n = 25 n = 3,604
n = 25 n = 25 n = 3,928
r = 0.50
r = –0.08
r = –0.64
r = 0.65
r = 0.05
r = –0.48
r = 0.58
r = 0.05
r = –0.60
r = 0.47
r = –0.05
r = –0.61
DMN thickness (mm)
DMN RSFC


a
False negatives (type 2)
False negative rate (%)
100
80
60
40
20
0
P value
10−7
0.05
25 50 100 200 375 725 1,430 2,800
Sample size
b
Inflation (type M)
Inflated correlation (%)
Magnitude of inflation
50%
100%
200%
P < 0.05
P < 10−7
Unthresholded
Sample size
c
Sign error (type S)
Opposite sign (%)
Opposite sign (%)
3
0
Sample size
P value
Unthresholded
d
Statistical power
Statistical power (%)
e
Probability of replication
Successful replications (%)
f
False positives (type 1)
False positive rate (%)

Summary

Although not all measures were studied, we can assume the effect sizes are similar (e.g., with EEG)

Compared to GWAS, BWAS requires fewer subjects

Ways to boost power: Within-study designs, Multivariate methods, interventions vs. observations

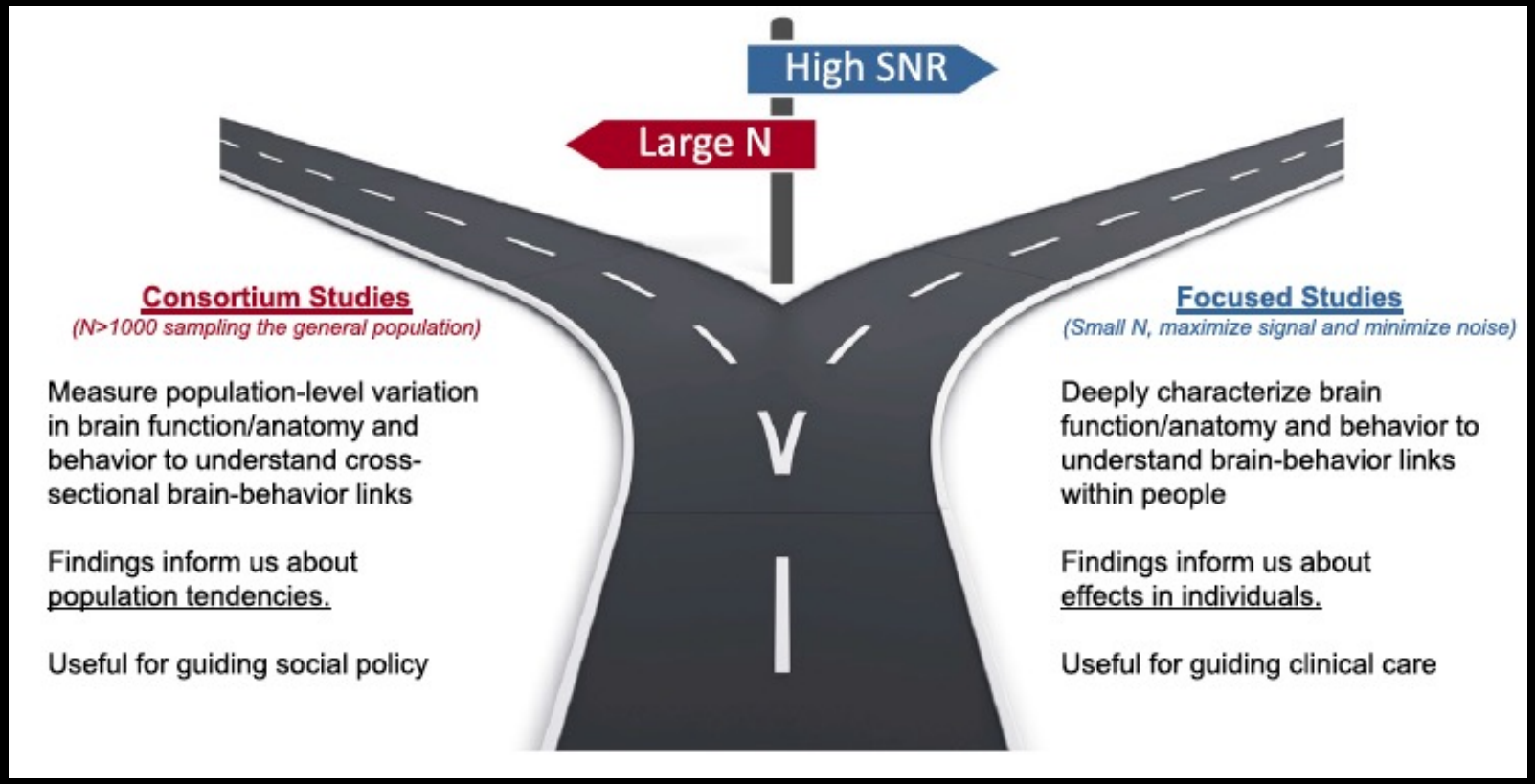
Gratton's Response

Cross-sectional studies with small N are useless

Consortia studies may have small effect sizes, but they are comparable to others that are useful

Nevertheless, they usually avoid novel experimental questions and designs

Gratton's Response

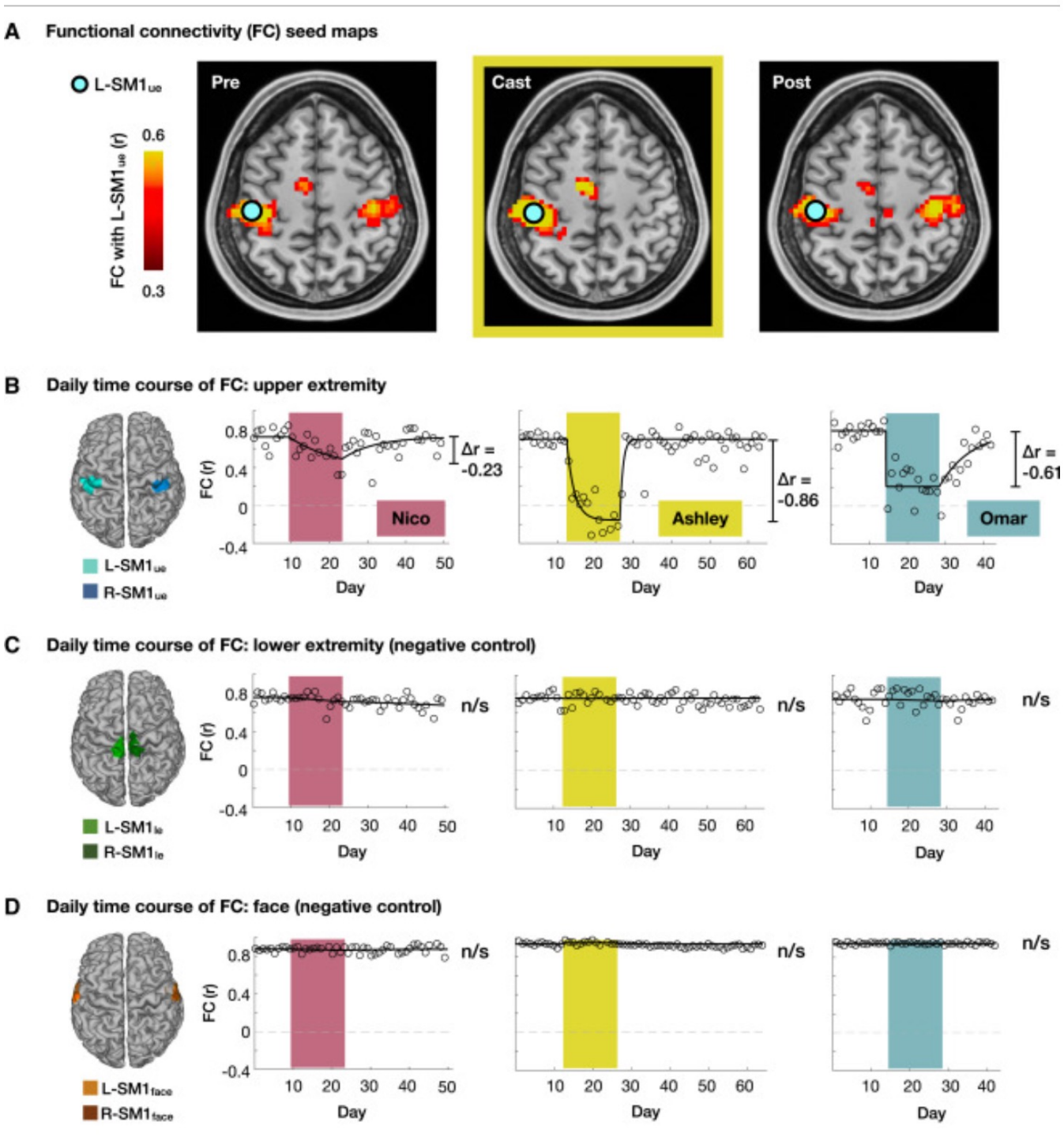


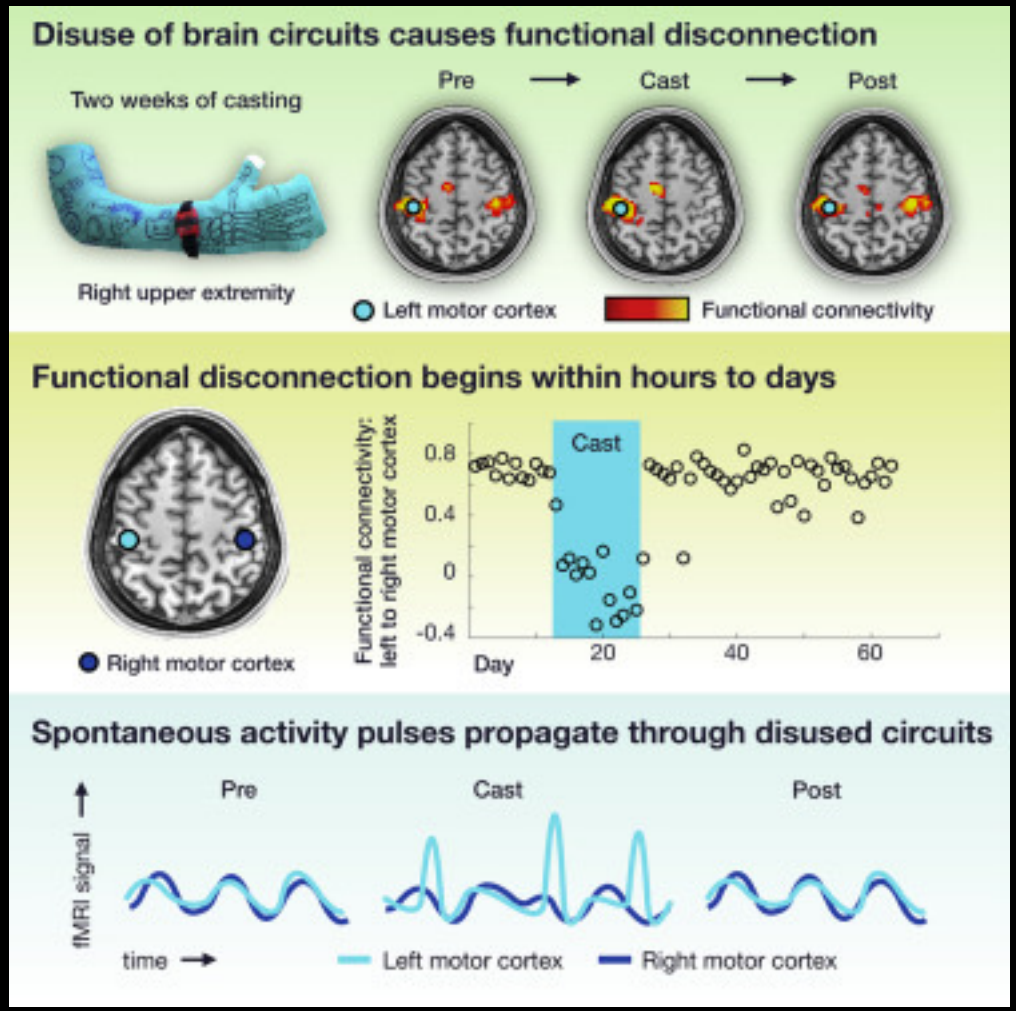
Other possibilities

Over-reliance on large-scale studies can limit funding opportunities for junior researchers

Smaller studies can still yield useful results through higher signal and lower noise designs

e.g., within-subjects designs, using designs that induce large alterations in behavior





Newbold et al., 2020

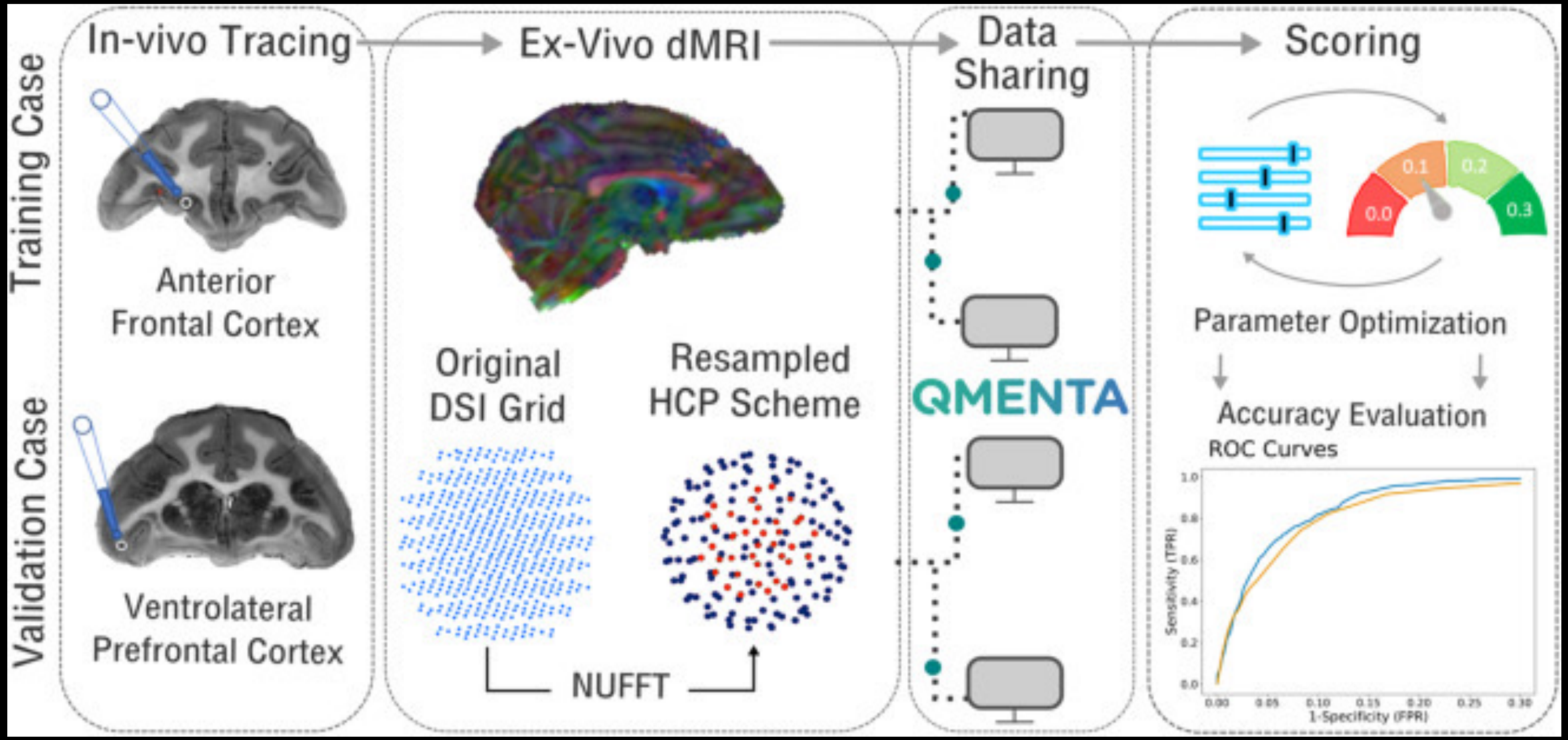
Background to the IronTract Challenge

Historically, several issues with reconstructing pathways

Demands for higher spatial and angular resolution

Advanced acquisition parameters were developed by the Human Connectome Project (HCP)

Need for comparing different methods


Training Case
Validation Case
In-vivo Tracing
Ex-Vivo dMRI
Data Sharing
Scoring
Anterior Frontal Cortex
Ventrolateral Prefrontal Cortex
Original DSI Grid
Resampled HCP Scheme
NUFFT
QMENTA
Parameter Optimization
Accuracy Evaluation
ROC Curves
Sensitivity (TPR)
1-Specificity (FPR)

Round 1

Allowed to use analysis methods of choice

Both probabilistic and deterministic tractography were used

Training case: Could repeat analysis any number of times

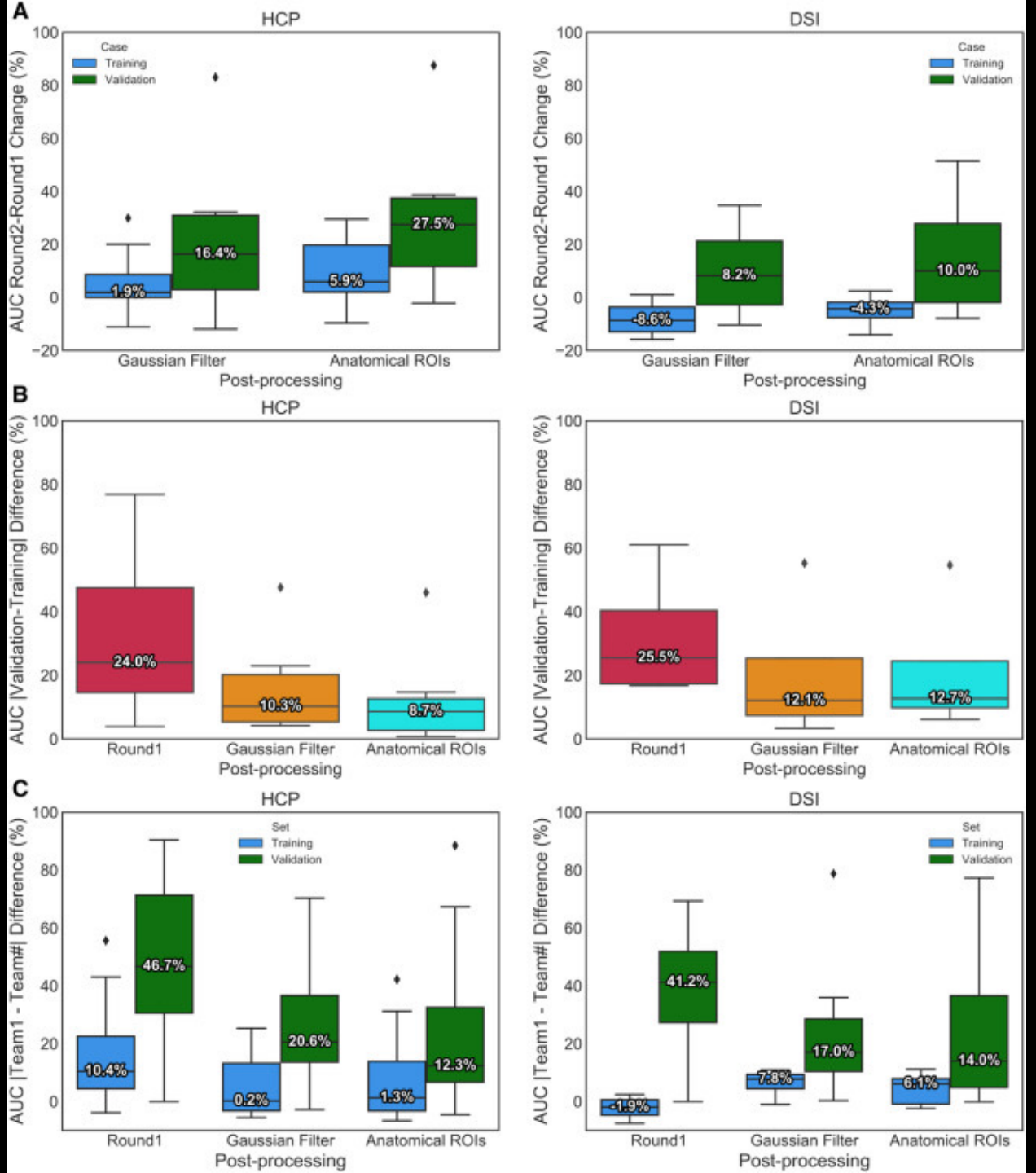
Round 2

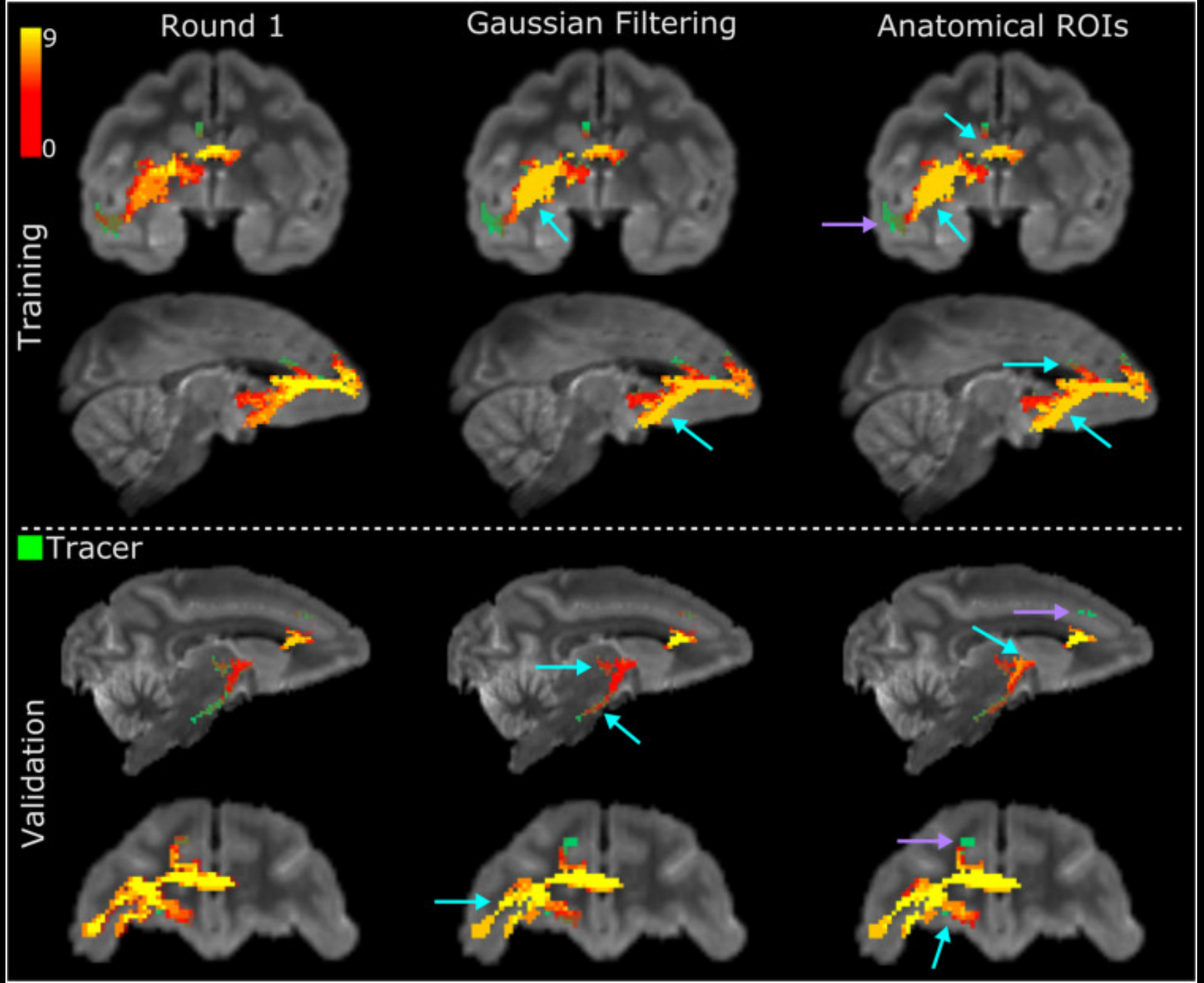
Pre- and post-processing was standardized

Choice of orientation reconstruction and tractography

ROIs analyzed: Cingulum, CC genu, external capsule, internal capsule, and uncinate fasciculus

Round 2 Results




Round 1
Gaussian Filtering
Anatomical ROIs
9
0
Training
Tracer
Validation

Questions?

Pre-registration

Posit hypotheses before collecting data

Specify parameters such as sample size, analysis options, dependent variables, and exclusion criteria

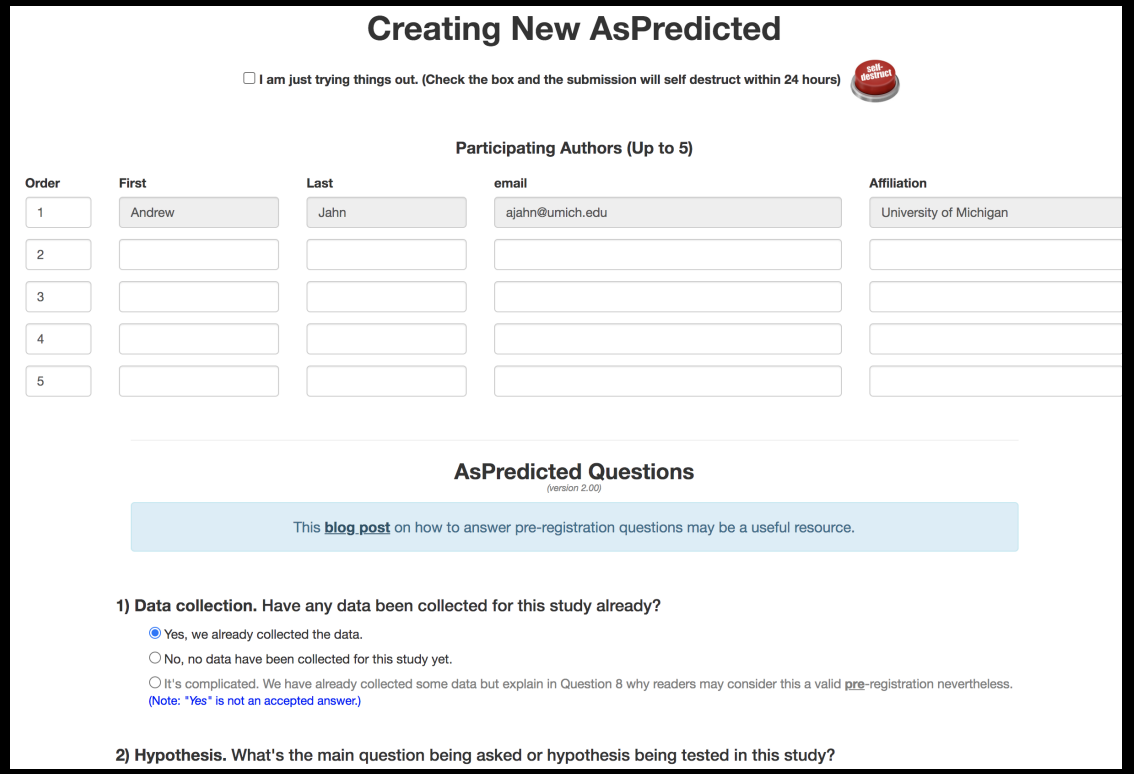
Pre-registration

Example: Open Science Framework

Improve your research with preregistration. By writing out specific details such as data collection methods, analysis plans, and rules for data exclusion, you can make important decisions early on and have a clear record of these choices. This can help reduce biases that occur once the data are in front of you.

Pre-registration

Alternative: AsPredicted.org



Creating New AsPredicted

☐ I am just trying things out. (Check the box and the submission will self destruct within 24 hours)

Participating Authors (Up to 5)

Order	First	Last	email	Affiliation
1	Andrew	Jahn	ajahn@umich.edu	University of Michigan
2				
3				
4				
5				

AsPredicted Questions

(version 2.00)

This **blog post** on how to answer pre-registration questions may be a useful resource.

1) Data collection. Have any data been collected for this study already?

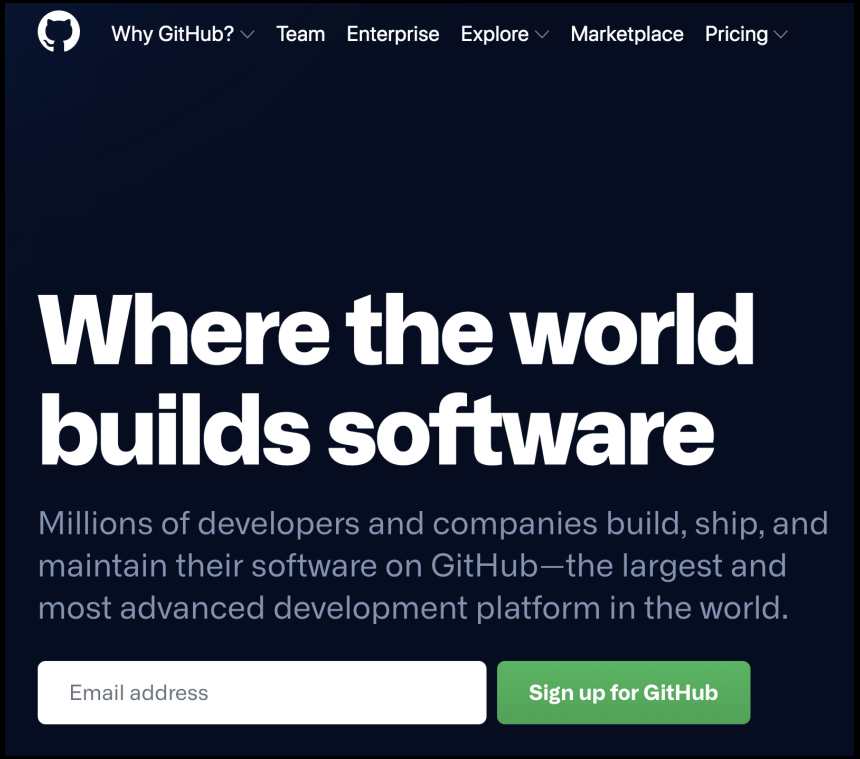
- ◉ Yes, we already collected the data.
- ○ No, no data have been collected for this study yet.
- ○ It's complicated. We have already collected some data but explain in Question 8 why readers may consider this a valid **pre**-registration nevertheless.

(Note: "Yes" is not an accepted answer.)

2) Hypothesis. What's the main question being asked or hypothesis being tested in this study?

Other Tools

Github



Github

Andrew Jahn
andrewjahn

Follow

Neuroimaging consultant, working primarily with AFNI, FSL, SPM, FreeSurfer, and MRtrix.

108 followers · **1** following · ☆ **0**

University of Michigan
Ann Arbor, MI
https://andysbrainbook.readthedocs.io...

Achievements

Overview | Repositories 15 | Projects | Packages

Popular repositories

AndysBrainBook
This repository contains the files that generate Andy's Brain Book on ReadTheDocs.
☆ 55 ⑂ 23

OpenScience_Scripts
Scripts to use with Open Science materials such as fMRIPrep
Shell ☆ 6 ⑂ 7

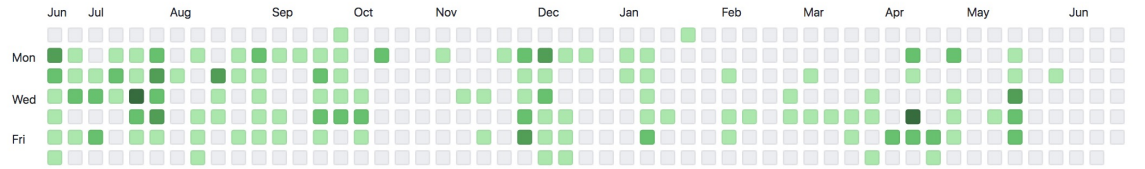
AFNI_Scripts
Scripts used for fMRI data analysis in AFNI
Shell ☆ 5 ⑂ 5

FSL_Scripts
Scripts for analyzing fMRI data using FSL
Shell ☆ 4 ⑂ 8

MRtrix_Analysis_Scripts
Scripts for analyzing diffusion data with MRtrix
Shell ☆ 3 ⑂ 3

CONN_Scripts
MATLAB ☆ 3 ⑂ 3

440 contributions in the last year



Downloading the “git” command

Installer for Macintosh: https://git-scm.com/download/mac

Installer for Windows: **https://git-scm.com/download/win**

Vocabulary

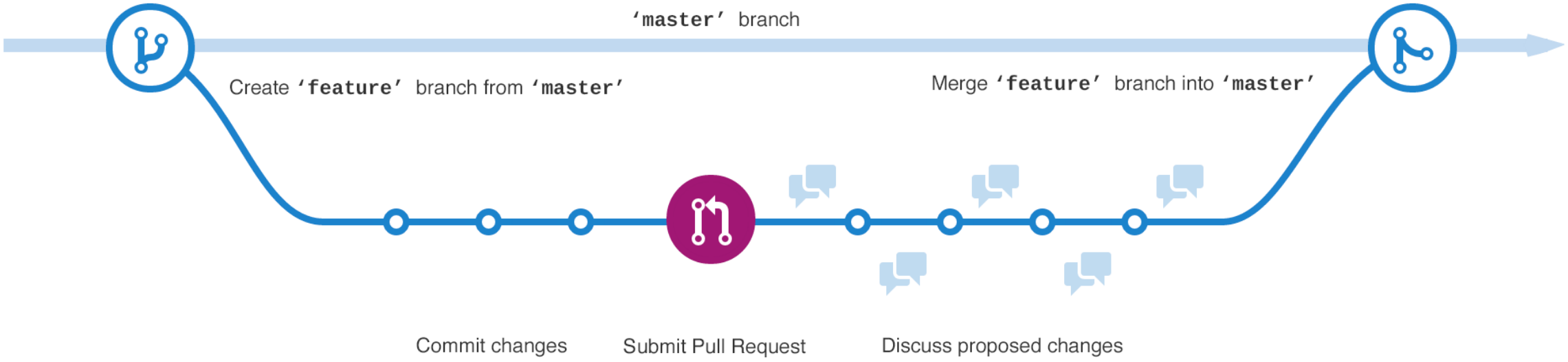
Github has a technical (and sometimes confusing!) vocabulary

Repository:
contains all of the project files (including documentation), and stores each file's revision history

Vocabulary

Branch: Current copy of the finalized edits

Each repository by default has a “main” branch



Vocabulary

Fork: Create a copy of a repository

Any edits made to this copy will not be seen by the public until the changes are merged

Vocabulary

Commit: Snapshot of an edit that can be later merged into the main branch

Commits can be labeled with messages describing what the change was

```
23
24  .. figure:: Github_Repositories.png
25
26    Example of repositories on a Github page.
27
28  * Clone: Copying a repository to your local machine. For example, if I want to clone the repository ``SPM_Scripts`` from Andy's Github page, I would need to know
    the link to the page (i.e., https://github.com/andrewjahn/SPM_Scripts), and then use it with the ``git`` command:
29
30    ..
31
32      git clone https://github.com/andrewjahn/SPM_Scripts
33
34  This will clone the SPM_Scripts repository to my local machine, from where I ran the ``git`` command.
35
36  * Branching: Each repository has a ``main`` branch, which contains all of the final edits that are seen by the public. A new branch is created to make edits, and
    can be called anything you like; when the edits are approved by whoever owns the repository, they are merged into the main branch.
37
38  .. figure:: Github_Branch.png
39
```

Commit changes

Add Branching and Clone Definitions

This edit defines what Branching and Cloning mean in Github.

- (•) Commit directly to the `master` branch.
- ( ) Create a **new branch** for this commit and start a pull request. Learn more about pull requests.

Commit changes | Cancel

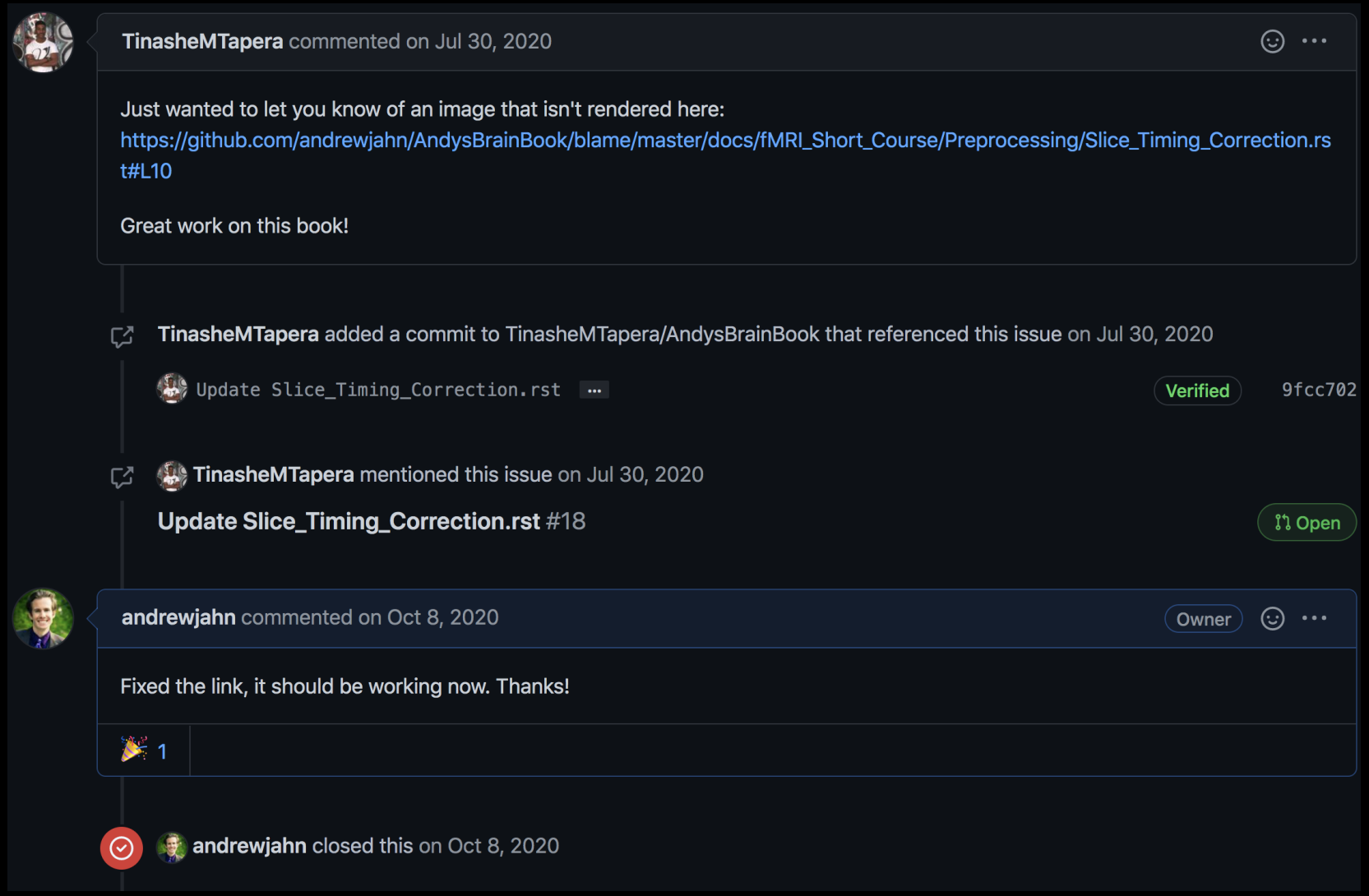
Vocabulary

Push and pull

Push: Send changes to your repository, even if they are created locally (need permissions)

Pull request: Ask for a review of your commits before they are merged into the main branch

Opening an Issue



TinasheMTapera commented on Jul 30, 2020

Just wanted to let you know of an image that isn't rendered here: https://github.com/andrewjahn/AndysBrainBook/blame/master/docs/fMRI_Short_Course/Preprocessing/Slice_Timing_Correction.rst#L10

Great work on this book!

TinasheMTapera added a commit to TinasheMTapera/AndysBrainBook that referenced this issue on Jul 30, 2020

`Update Slice_Timing_Correction.rst` Verified `9fcc702`

TinasheMTapera mentioned this issue on Jul 30, 2020

Update Slice_Timing_Correction.rst #18 Open

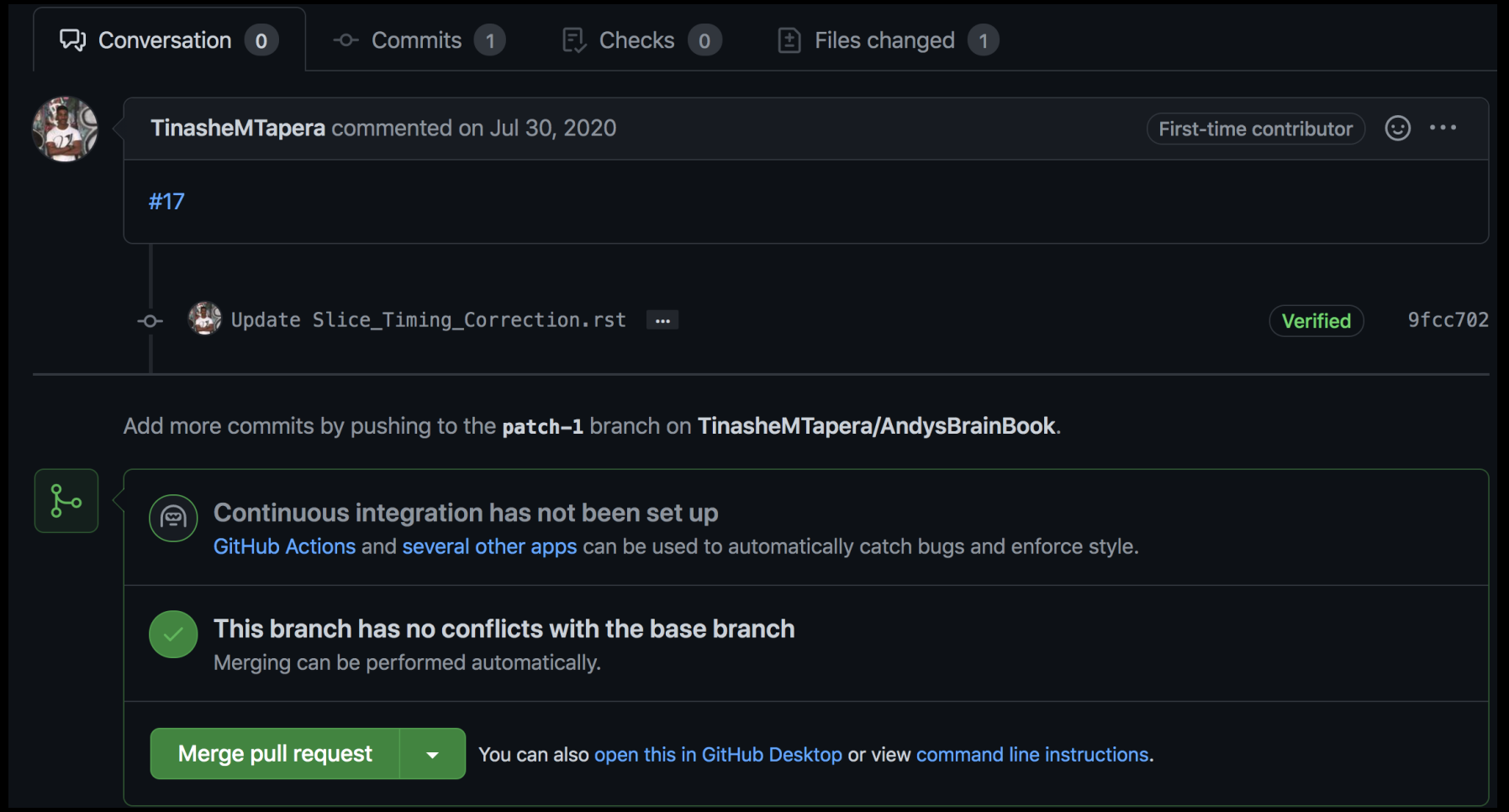
andrewjahn commented on Oct 8, 2020 Owner

Fixed the link, it should be working now. Thanks!

1

andrewjahn closed this on Oct 8, 2020

Closing an Issue



Update Slice_Timing_Correction.rst #18

Edit | Open with

Merged **andrewjahn** merged 1 commit into `andrewjahn:master` from `TinasheMTapera:patch-1` now

Conversation 0 | Commits 1 | Checks 0 | Files changed 1 | +1 −1

Changes from all commits | File filter | Conversations | Jump to | 0 / 1 files viewed | Review changes

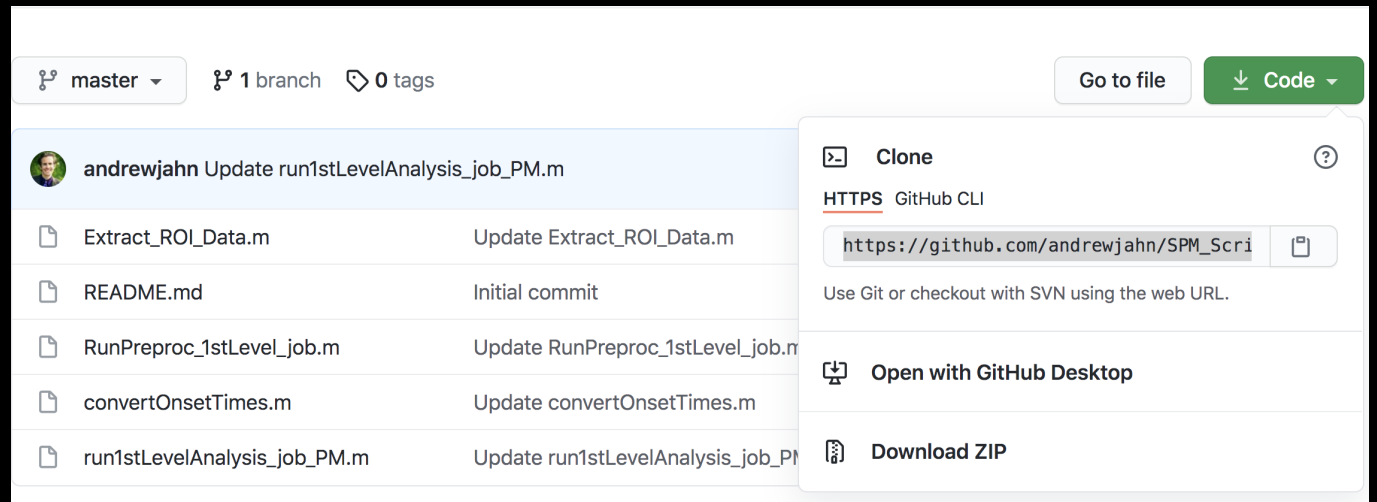
2 docs/fMRI_Short_Course/Preprocessing/Slice_Timing_Correction.rst | Viewed

```
@@ -7,7 +7,7 @@ Unlike a photograph, in which the entire picture is taken in a single moment, an
 7  7
 8  8   The two most commonly used methods for creating volumes are sequential and interleaved slice acquisition. Sequential slice acquisition acquires each
        adjacent slice consecutively, either bottom-to-top or top-to-bottom. Interleaved slice acquisition acquires every other slice, and then fills in the gaps
        on the second pass. Both of these methods are illustrated in the video below.
 9  9
10     - .. figure:: SliceTImingCorrection_Demo.gif
    10 + .. figure:: SliceTimingCorrection_Demo.gif
11 11
12 12   As you'll see later on, when we model the data at each voxel we assume that all of the slices were acquired simultaneously. To make this assumption
        valid, the :ref:`time-series <Time_Series>` for each slice needs to be shifted back in time by the duration it took to acquire that slice. `Sladky et al.
        (2011) <https://www.sciencedirect.com/science/article/pii/S1053811911007245>`__ also demonstrated that slice-timing correction can lead to significant
        increases in statistical power for studies with longer TRs (e.g., 2s or longer), and especially in the dorsal regions of the brain.
13 13
```

Vocabulary

Clone: Copy of a repository that lives on your computer

Sample usage: git clone <repository address>

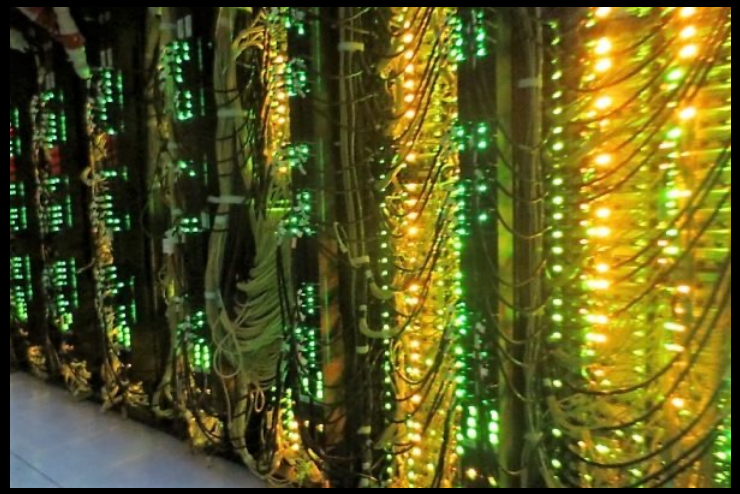


```
(base) ajahn:~/Desktop$ git clone https://github.com/andrewjahn/SPM_Scripts.git
Cloning into 'SPM_Scripts'...
remote: Enumerating objects: 40, done.
remote: Counting objects: 100% (40/40), done.
remote: Compressing objects: 100% (39/39), done.
remote: Total 40 (delta 19), reused 0 (delta 0), pack-reused 0
Unpacking objects: 100% (40/40), done.
(base) ajahn:~/Desktop$ ls
Archived                Flanker_fMRIPrep    SPM_Scripts
BTC_preop               Gambles             ds002422-download
CONN_Demo               Haxby_Data          network_TDA_tutorial
FSL_Flanker             Haxby_Data_Umich
Flanker                 NeuroNav
(base) ajahn:~/Desktop$ cd SPM_Scripts/
(base) ajahn:~/Desktop/SPM_Scripts$ ls
Extract_ROI_Data.m              convertOnsetTimes.m
README.md                       run1stLevelAnalysis_job_PM.m
RunPreproc_1stLevel_job.m
(base) ajahn:~/Desktop/SPM_Scripts$
```

Supercomputing

What is a supercomputer?

Great Lakes is a supercomputing cluster,
i.e., a large collection of computers



Supercomputing at the University of Michigan

LSA students can apply for a supercomputing account through Michigan's Advanced Research Computing (ARC) center

Usually requires a shortcode from the PI

Can apply for a Umich Research Computing Package (UMRCP)

80,000 CPU hours and 10TB of storage per year,
100TB archive storage

Supercomputing at the University of Michigan

Uses Batch computing

i.e., you specify the resources for a job or several jobs

Jobs are run by a job manager, which is told when to run by a job scheduler

These are run with a computing language called SLURM

```
#!/bin/bash

#-----------------------------------------------------------------------
#---------- HEADER ----------#
#SBATCH --job-name=mri_prep_0000004_01_01_T1

#---------- log file
#SBATCH -o /scratch/precisionhealth_project_root/precisionhealth_project1/shared_data/brainmri/slogs/mri_prep_0000004_01_01_T1.log

#---------- Cancel job after d-hh:mm:ss
#SBATCH --time=09:00:00

#---------- Number of cores
#SBATCH --nodes=1
#SBATCH --ntasks-per-node=1
#SBATCH --cpus-per-task=5

#---------- GB Memory
#SBATCH --mem=10g

#---------- Account will pay job
#SBATCH --account=precisionhealth_project1

#---------- Partition where job "lives"?
#SBATCH --partition=standard

#------ No e-mail notifications of job start/end/error
#SBATCH --mail-type=NONE

echo "Working in dir ${PWD}:"


#---------- MODULES ----------#
module purge


module load fsl/6.0.5.1
module load afni/18.0.27
module load freesurfer
```

EXAMPLE JOBS AND THEIR CHARGES[1]

To help illustrate how the job charges work, here are some examples of differently-sized jobs.

Partition	Total CPU Cores Used	Total Memory Used	Total GPUs Used	Cost Per Minute
standard	1	1 GB	N/A	$0.000250
standard	1	10 GB	N/A	$0.000500
standard	36	5 GB	N/A	$0.009000
standard	1	50 GB	N/A	$0.002000
largemem	1	180 GB	N/A	$0.003852
gpu	1	20 GB	1	$0.002739

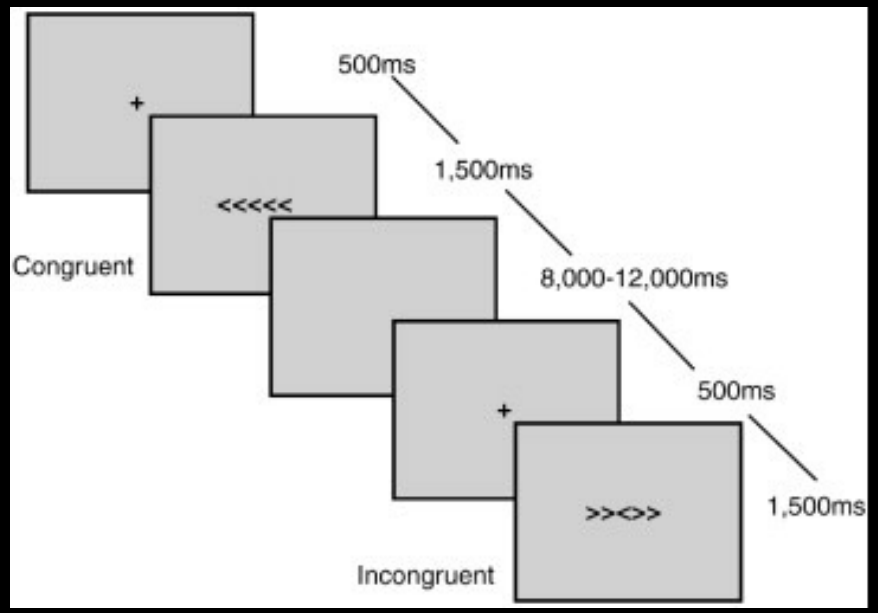
Example: Analyzing a dataset from openneuro.org

The Dataset

Flanker Task (Kelly et al., 2008)

Cognitive Control

Filtering out irrelevant stimuli to perform a task



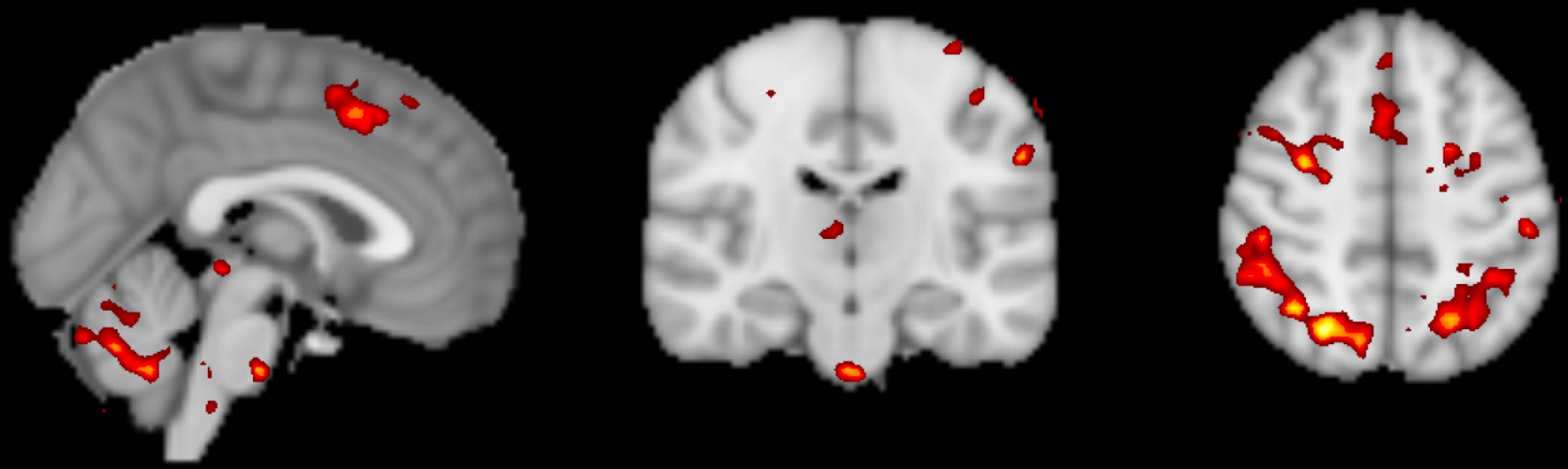
Kelly et al., 2008

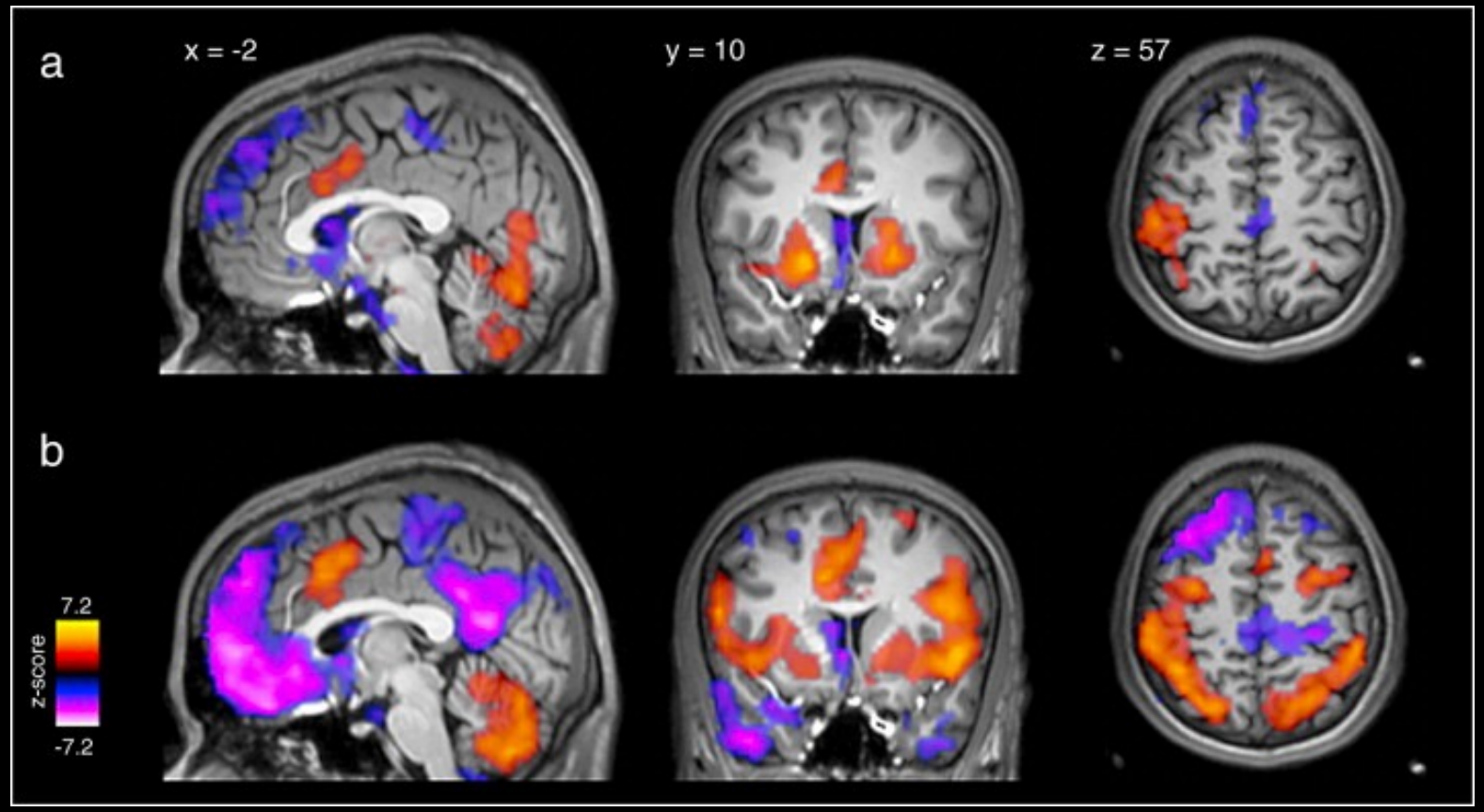
>>>><>>>>

Incongruent

>>>>>>>>>>

Congruent




a
x = -2
y = 10
z = 57
b
7.2
z-score
-7.2

Demonstration: Download the data and analyze it

General Q & A Session

Any questions about the material covered since last Friday?

PSY808 Course in the Fall and Winter will build upon what you've learned in the course

Future trends in neuroimaging analysis

Difficult to predict

What I think will happen:

Greater emphasis on standardized pipelines, software

More labs using large open-access datasets

Wider use of supercomputers

Concluding Remarks