

Preprocessing

First steps on the long road from raw data to results

Tomas Knapen 2022

fMRI analysis pipeline

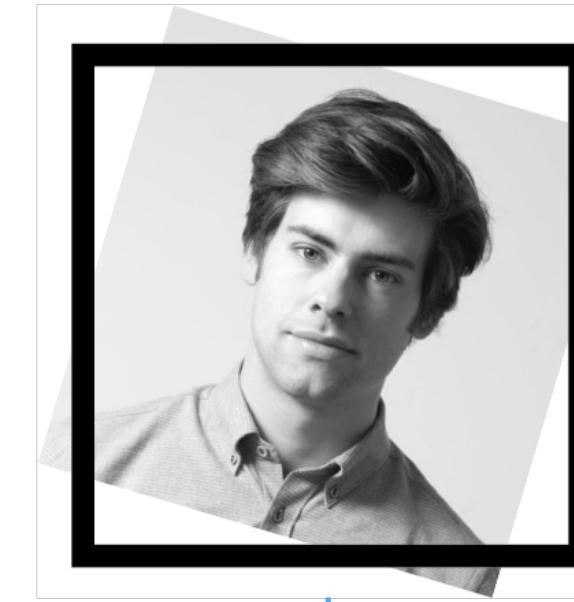
1. Image reconstruction
2. Distortion correction
3. Motion correction
4. Slice timing correction
5. Temporal filtering
6. Spatial smoothing
7. Statistical analysis
8. Spatial normalisation - different ways

linear transformations

original



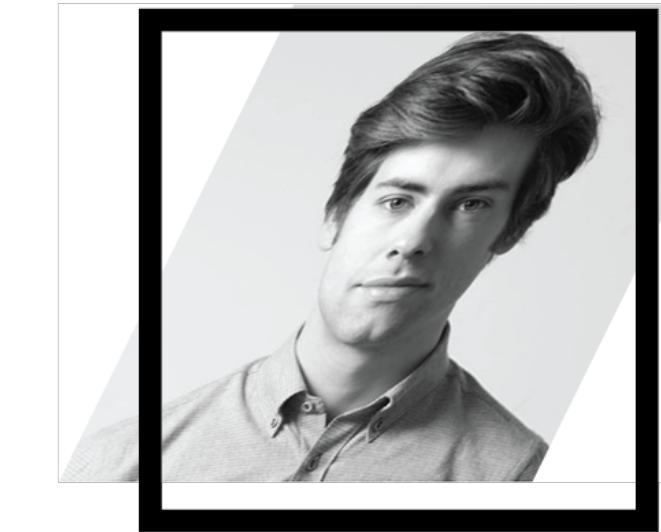
rotation



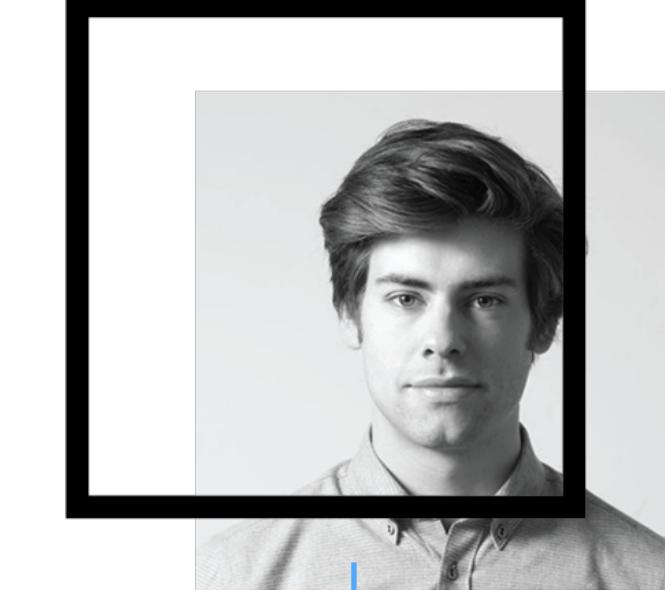
scale



shear



translation



expressed in
transformation
matrices

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & Shear_x & 0 \\ Shear_y & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} Scale_x & 0 & 0 \\ 0 & Scale_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & Trans_x \\ 0 & 1 & Trans_y \\ 0 & 0 & 1 \end{bmatrix}$$

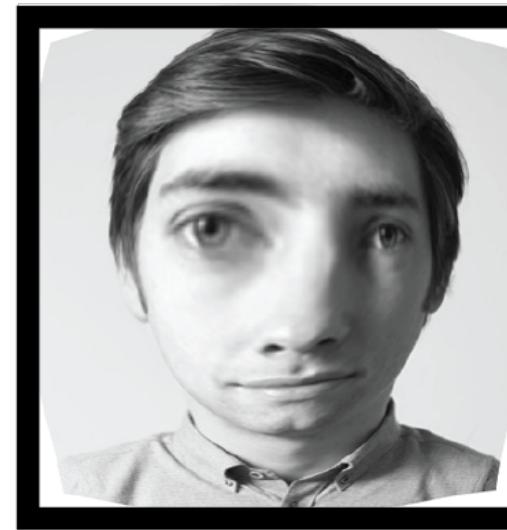
linear transformations
assume a *rigid body*

non-linear transformations

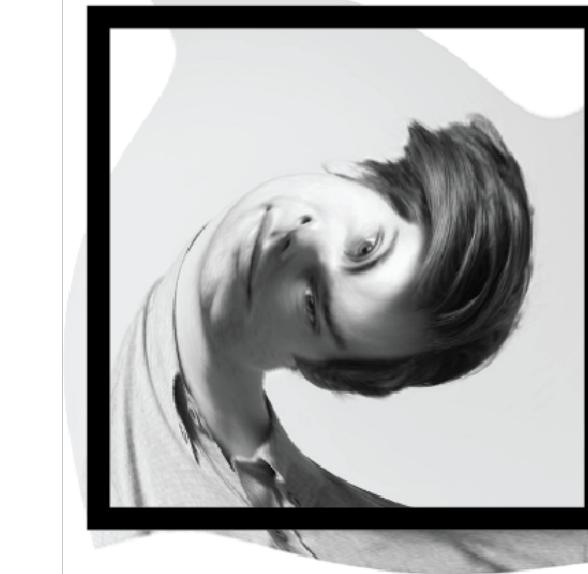
original



pinch



twirl



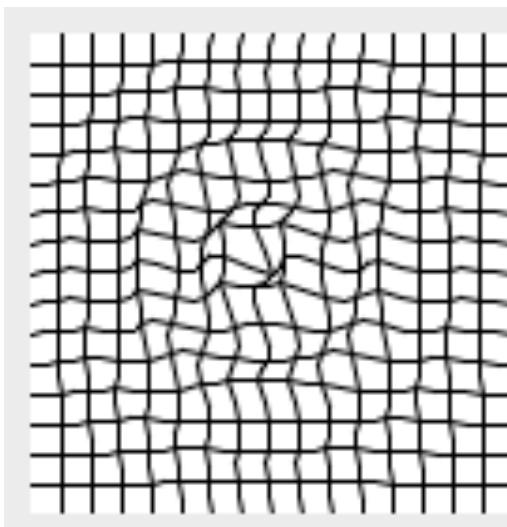
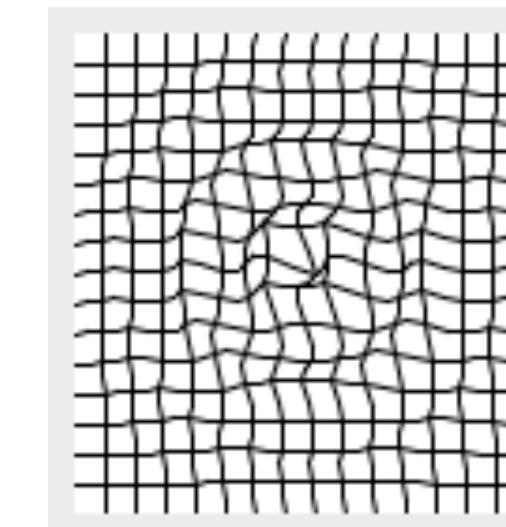
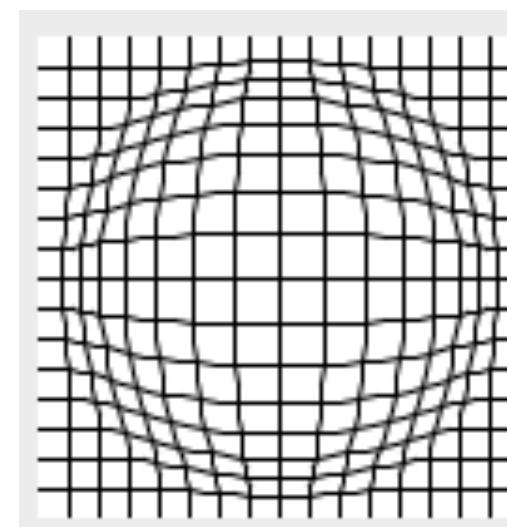
wave



zigzag



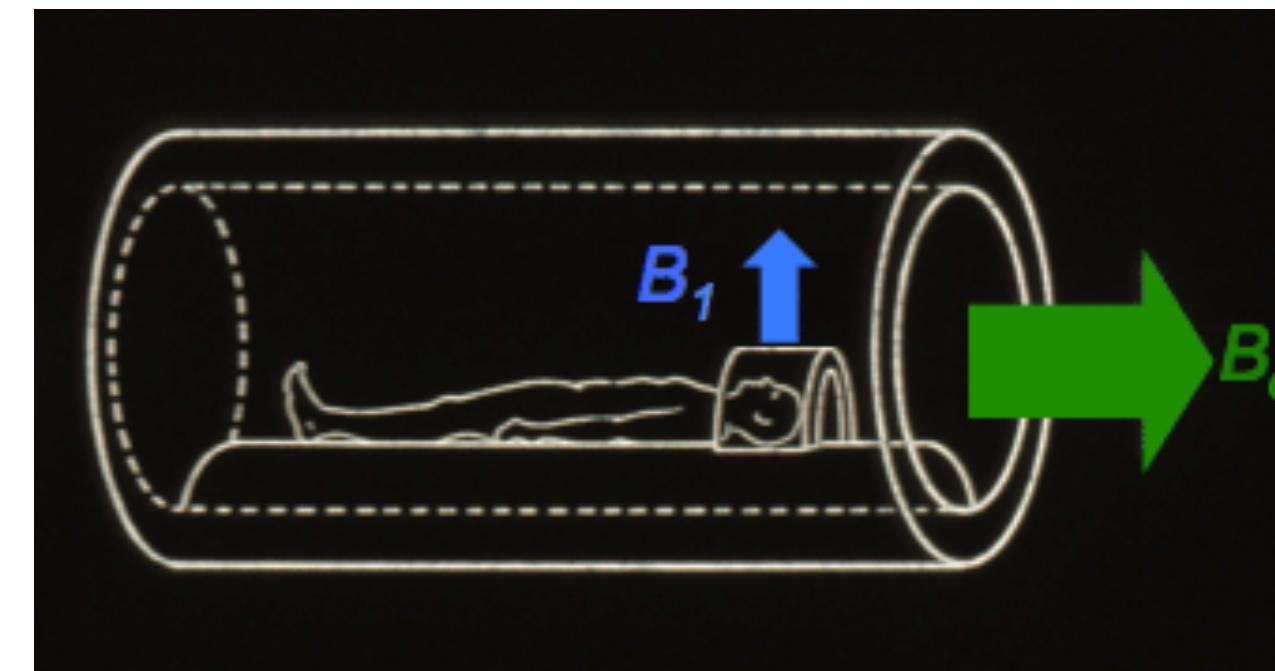
expressed in
warp field



careful with these -
weird things can happen!

Distortion correction

Different methods for correcting distortions

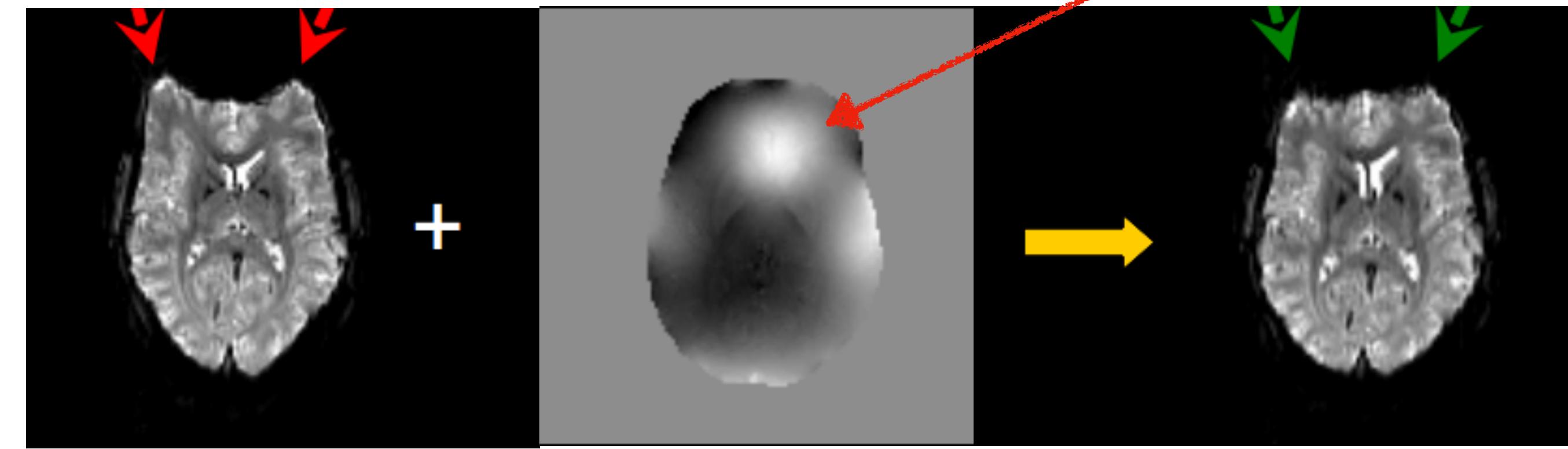


B0 field is distorted
by objects within it (body)

Only one of many possible methods

the distortion can be recorded ('B0-map') ~~inverting this warp-field solves the distortion~~

this causes
warping of
images



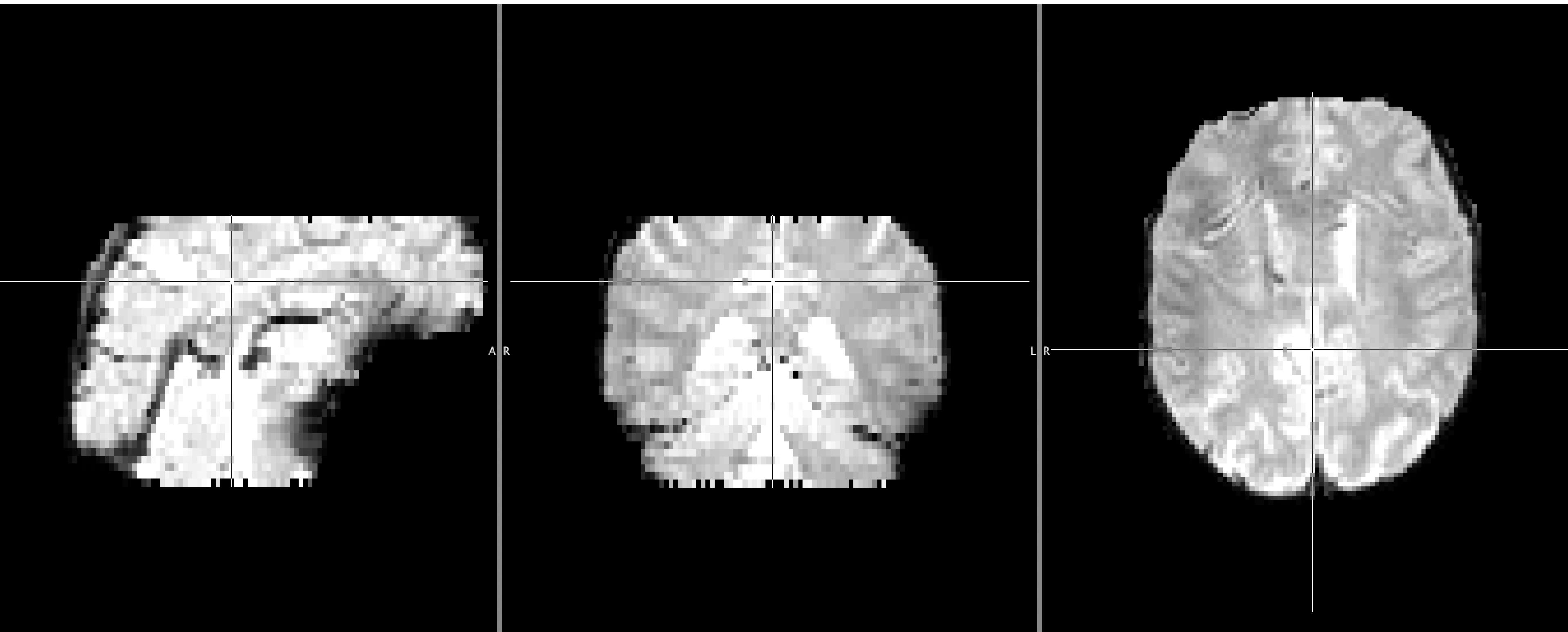
original image

B0 map

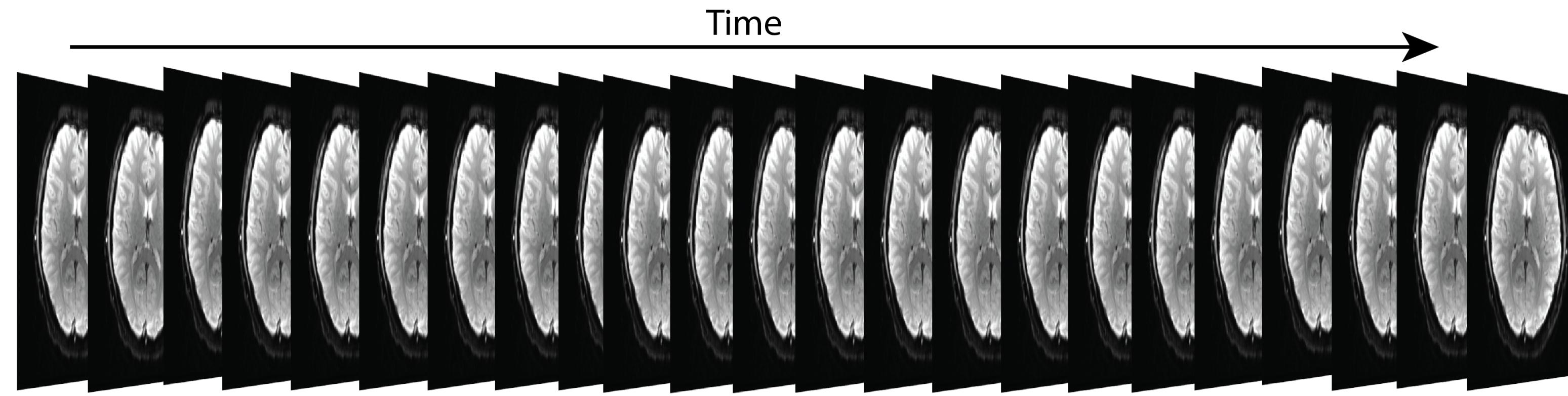
unwarped image

before and after

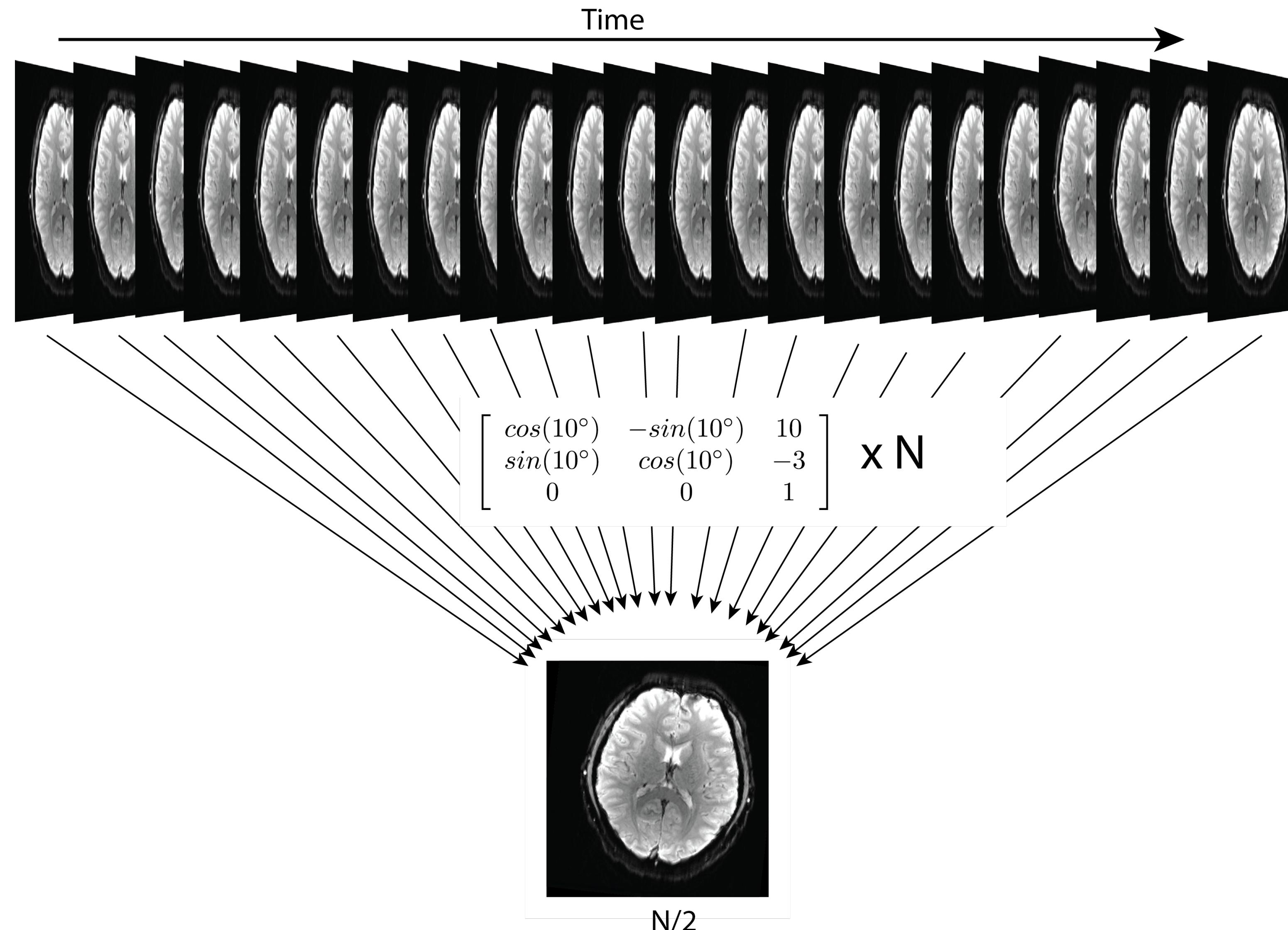
Ginormous distortions are fixed



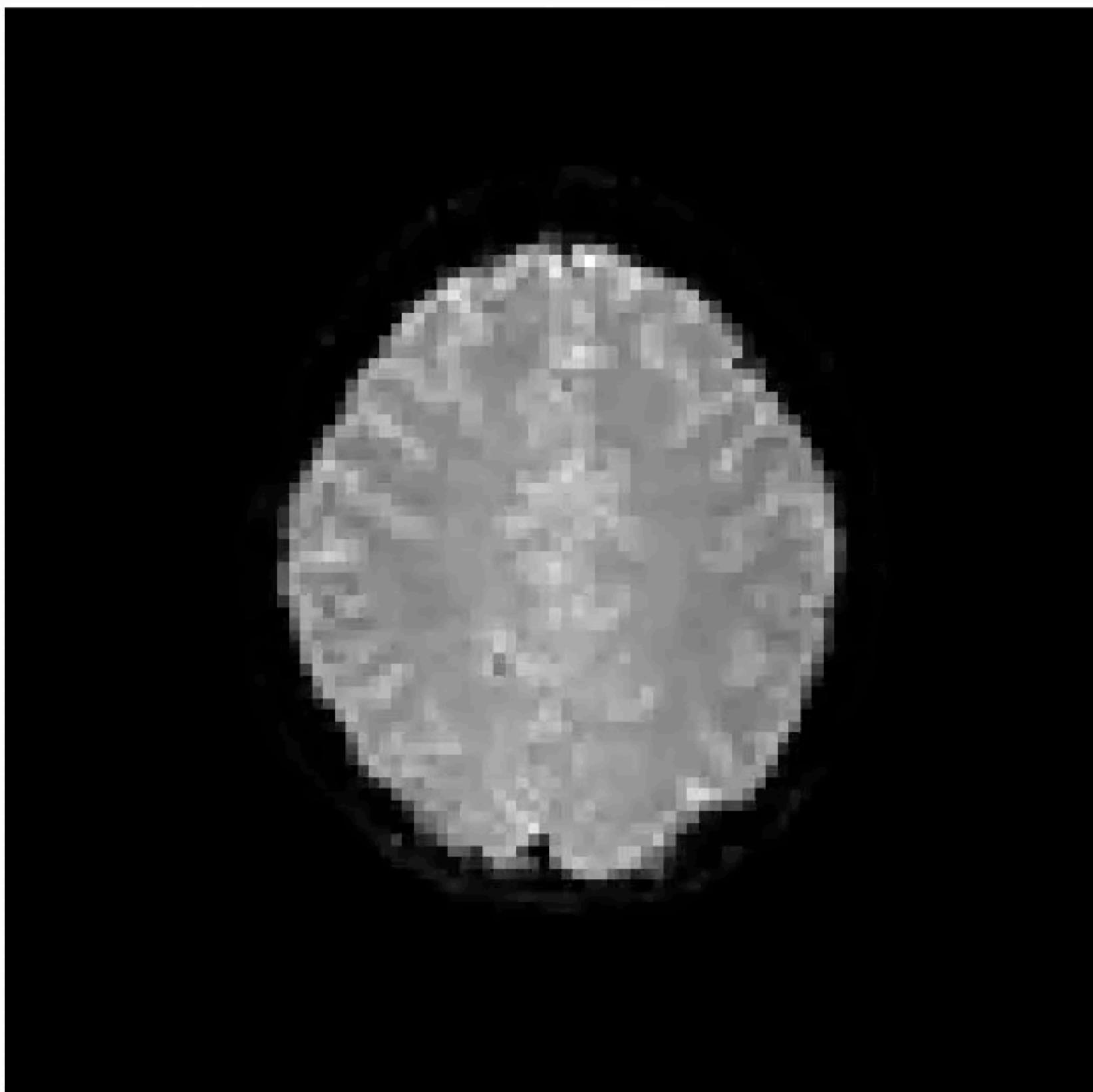
Live human beings: Subject motion!



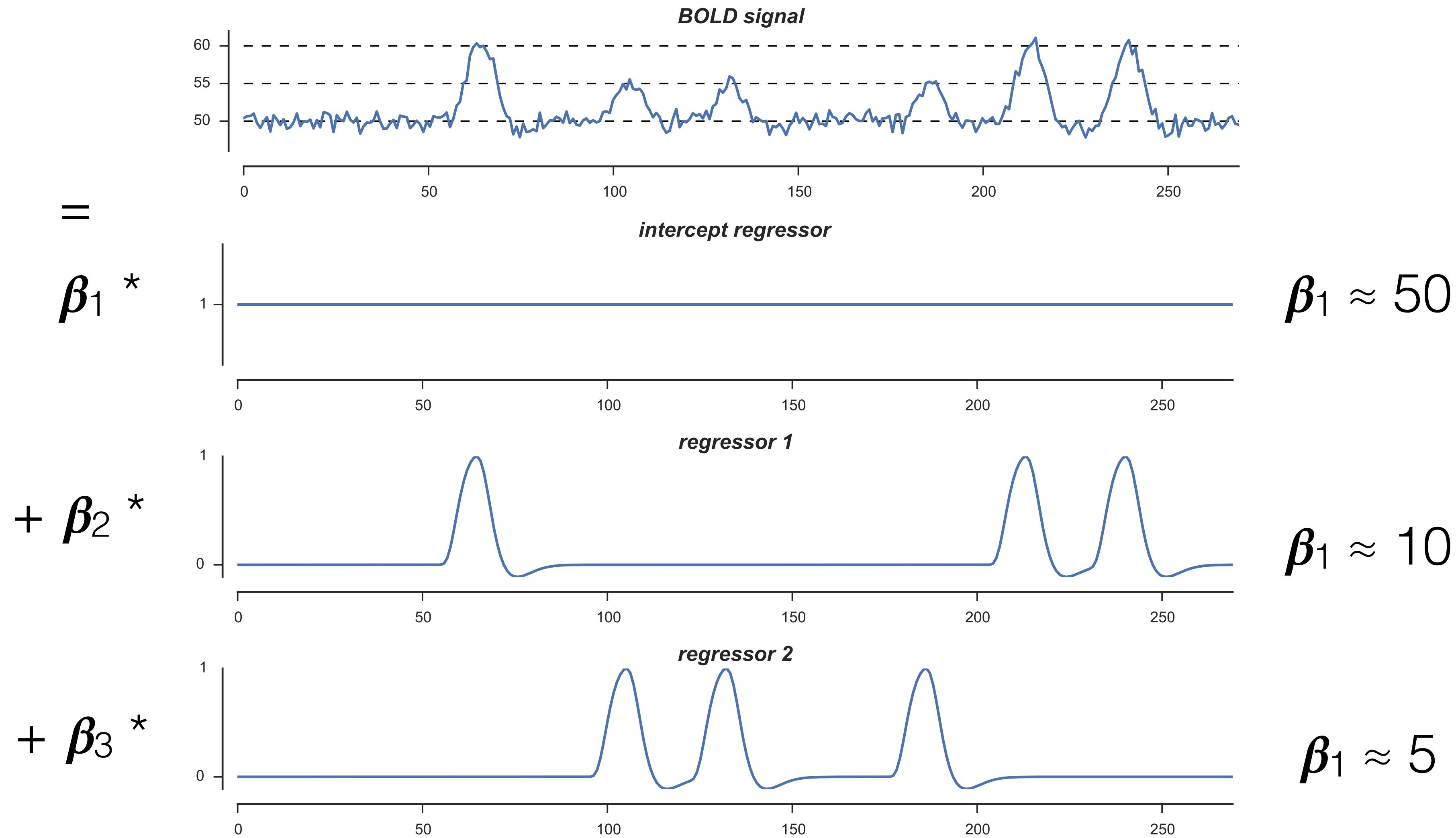
Live human beings: Subject motion!



before and after



GLM recap

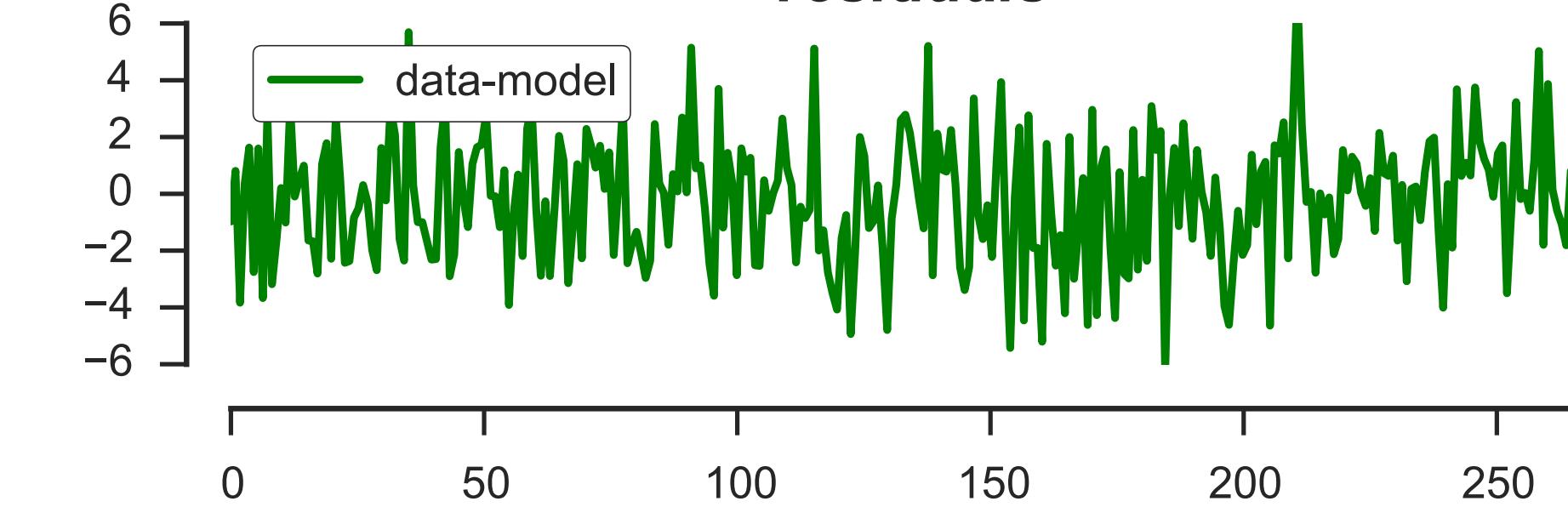
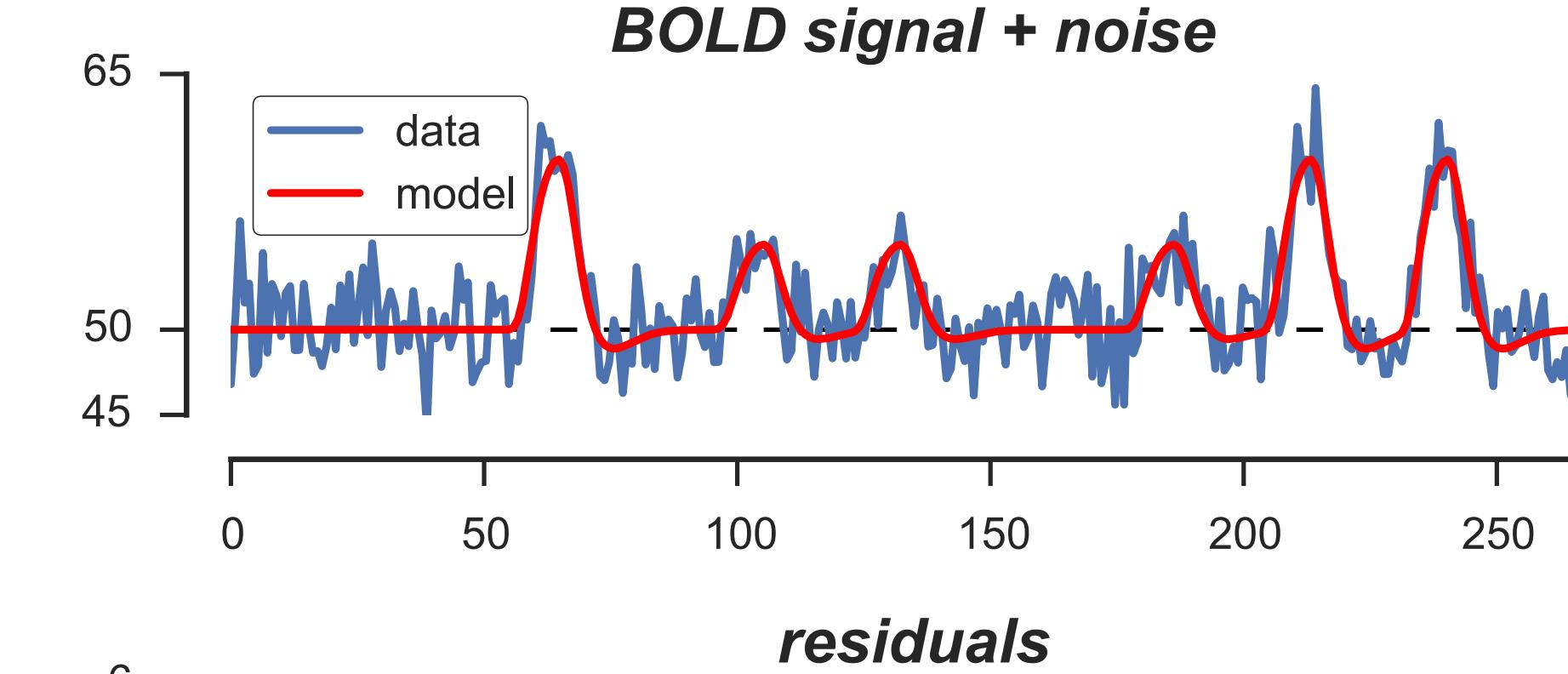
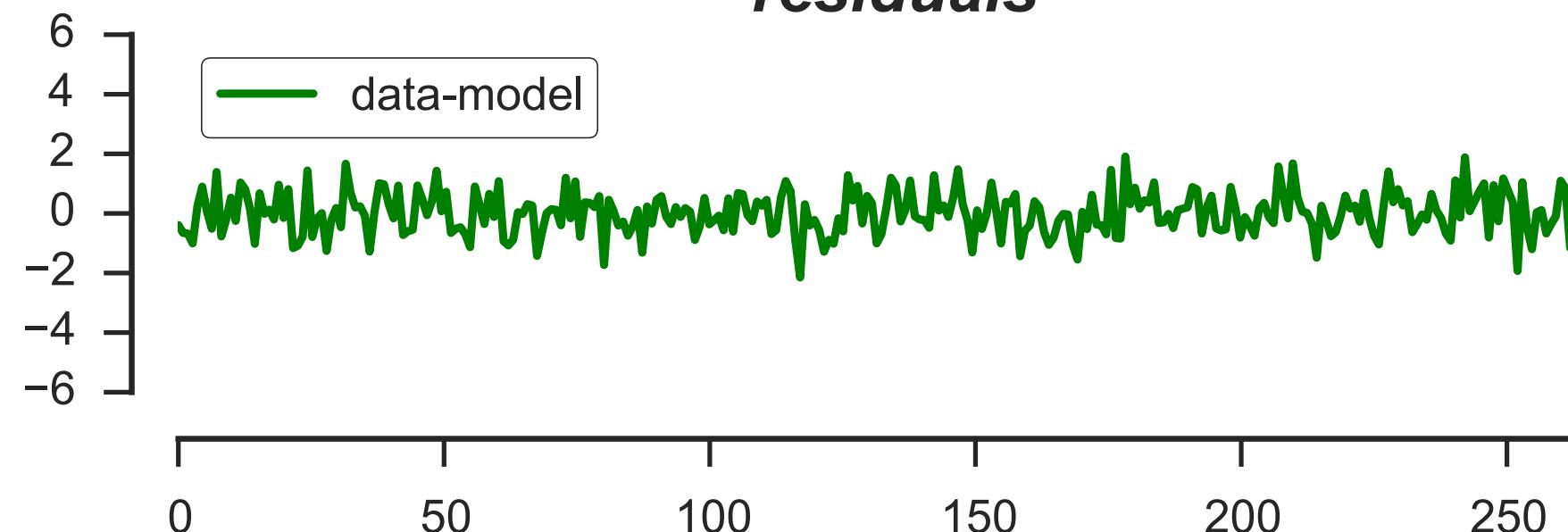
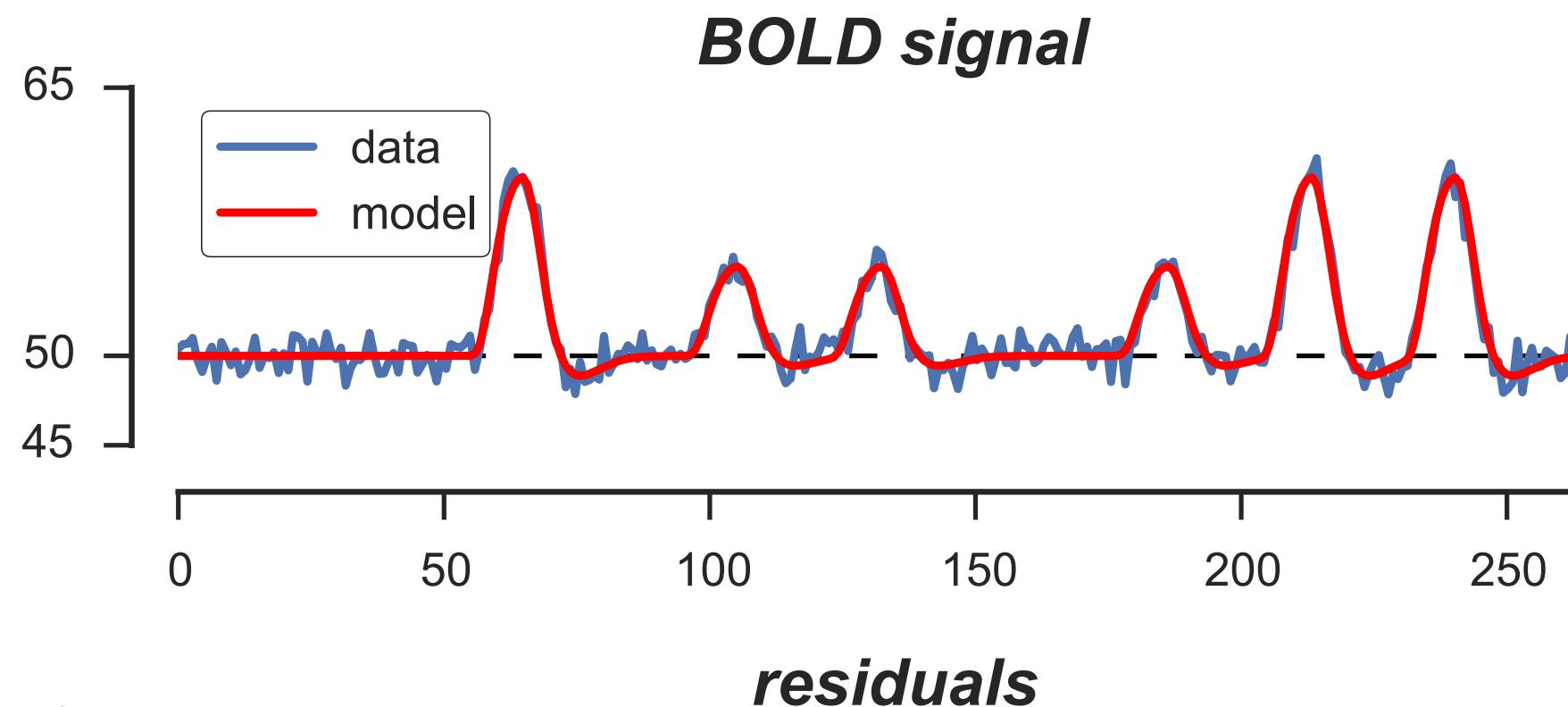


noise deflates t -vals

$$(1) \quad t = \frac{\beta}{SE_R}$$

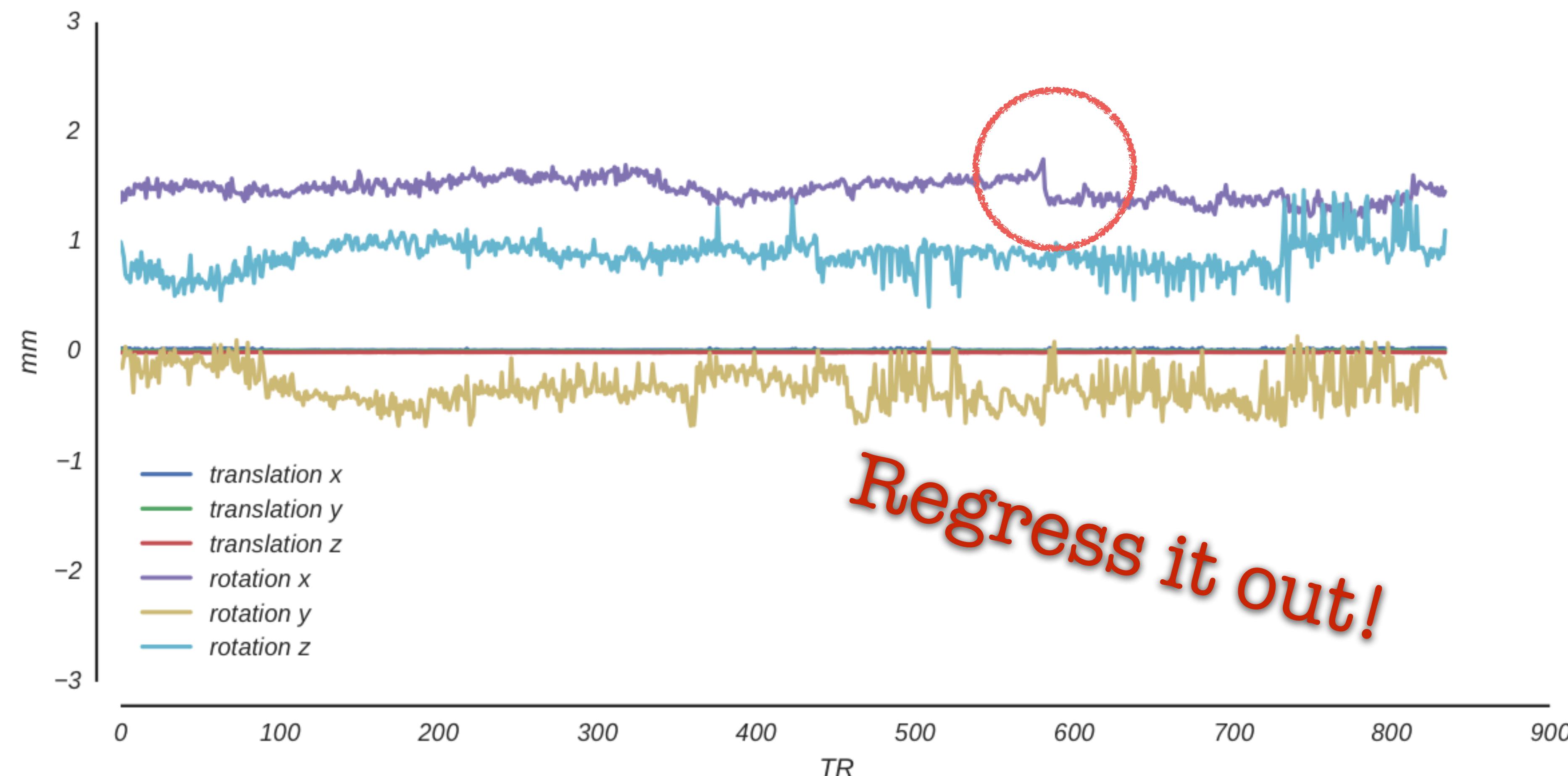
$$(2) \quad SE_R = \sum (data - model)^2$$

equations are oversimplification
for educational purpose!



motion parameters as nuisance regressors

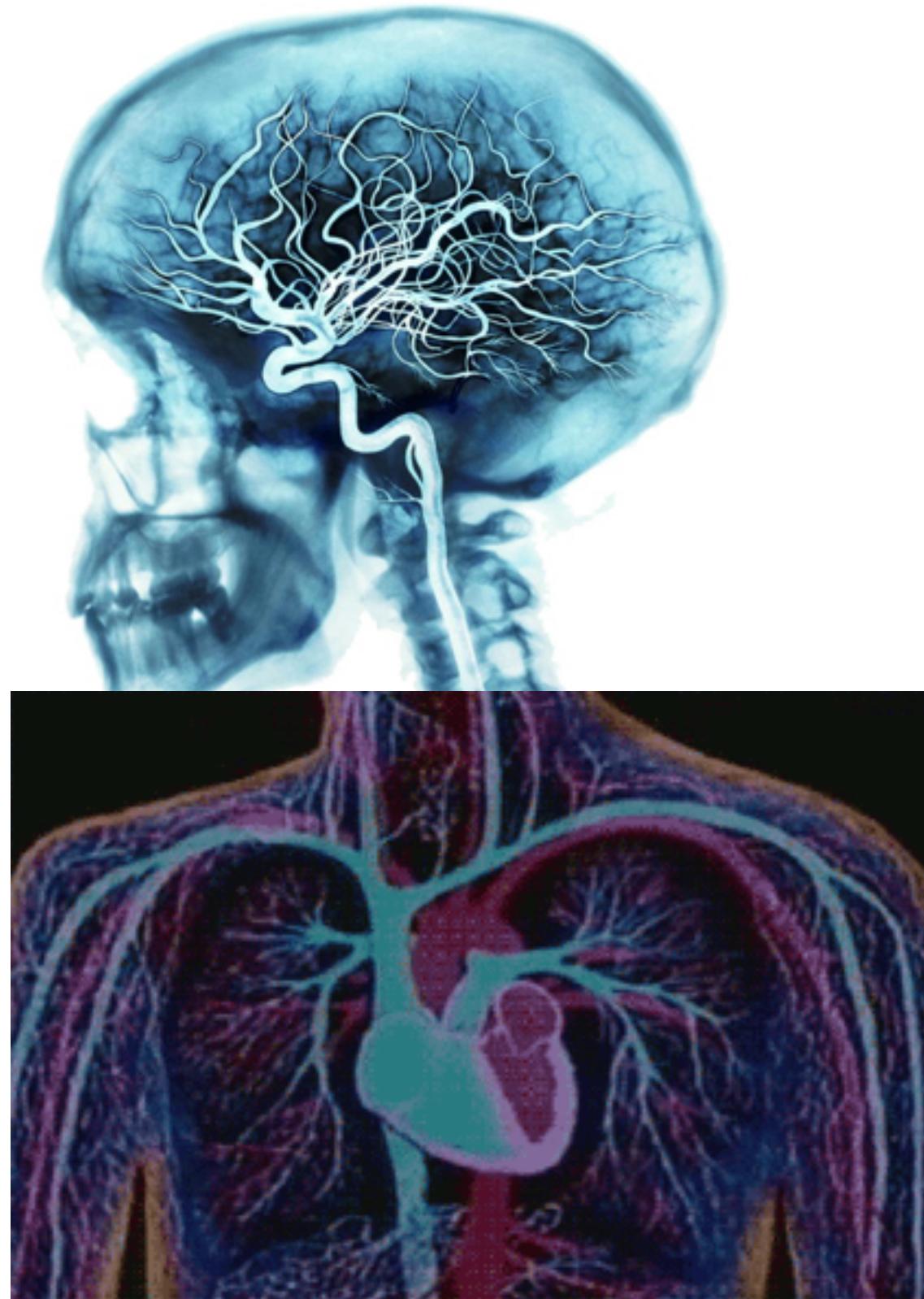
Motion leads to residual signal variance in time-courses



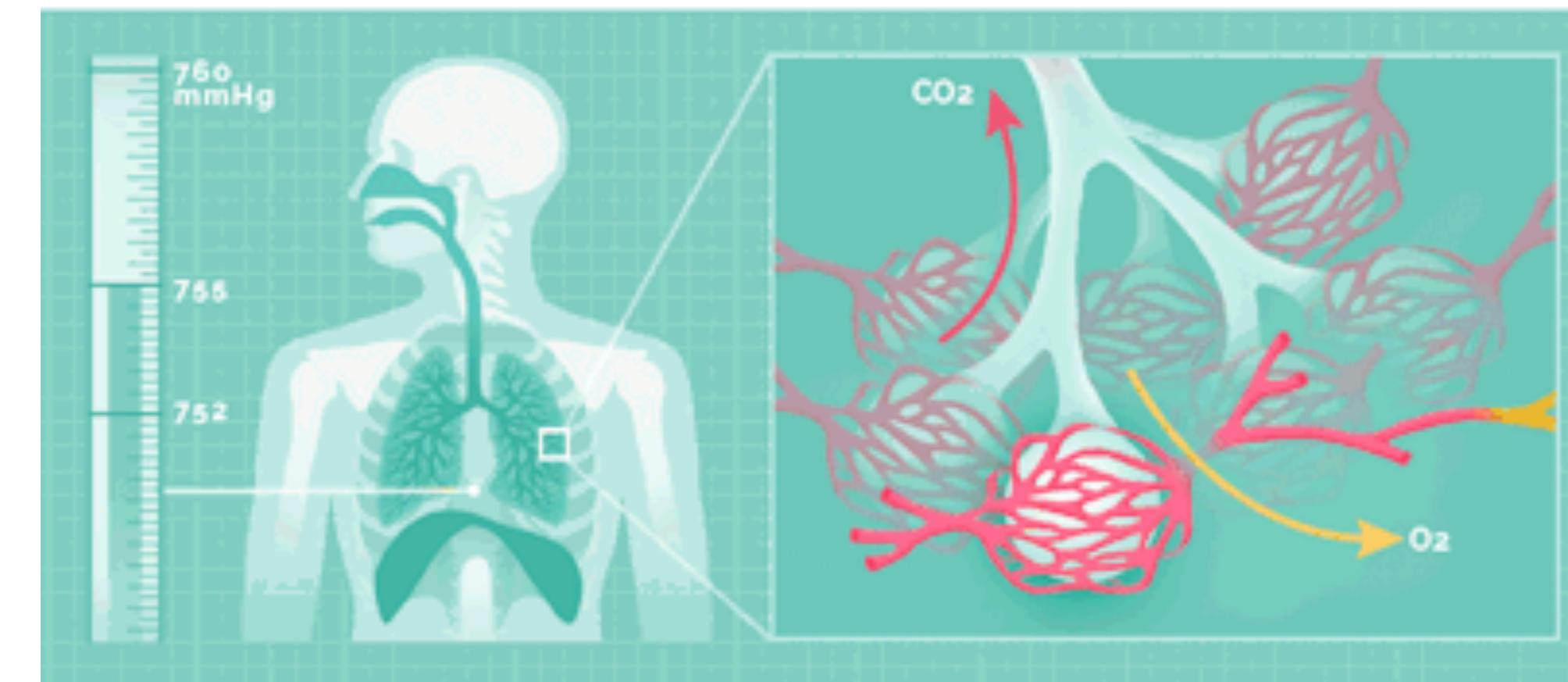
Physiology

Remember, we're measuring blood

heartbeat



respiration



Physiology

We can measure these factors independently

heartbeat



respiration



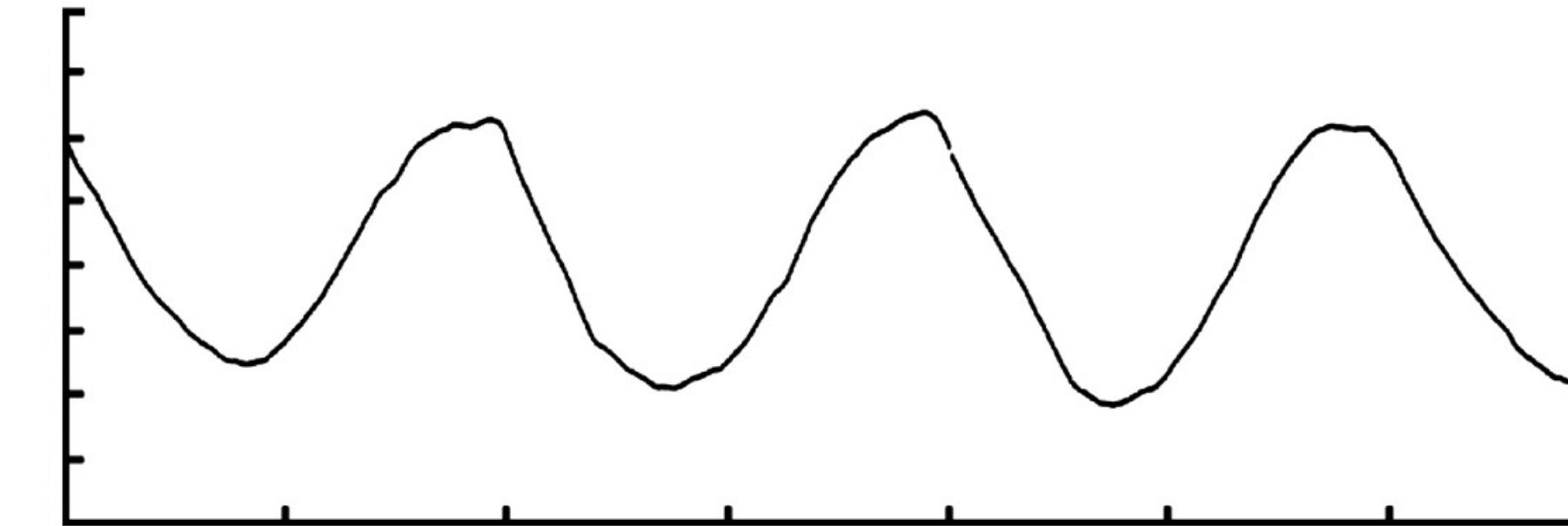
Physiology

Sampled signals

heartbeat

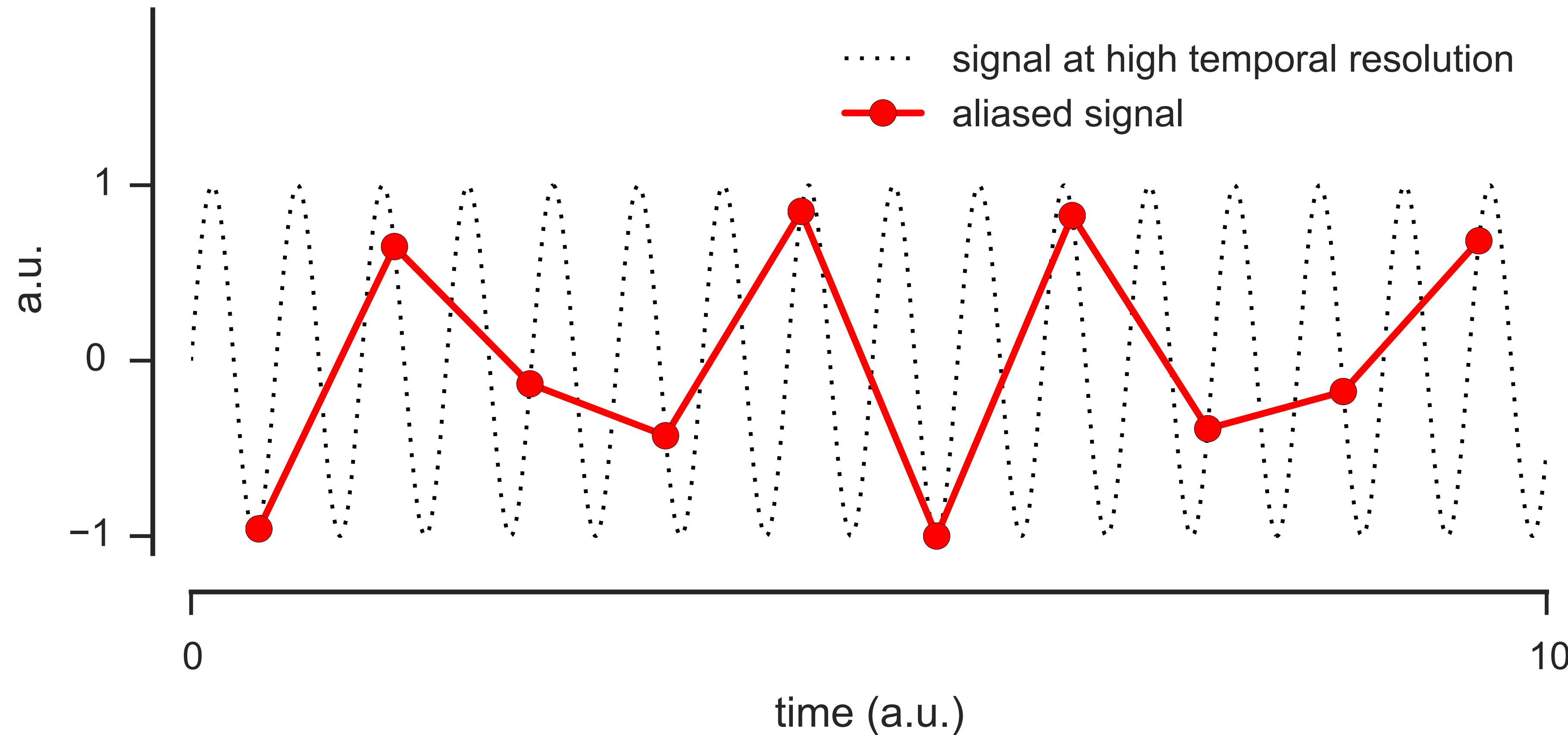


respiration



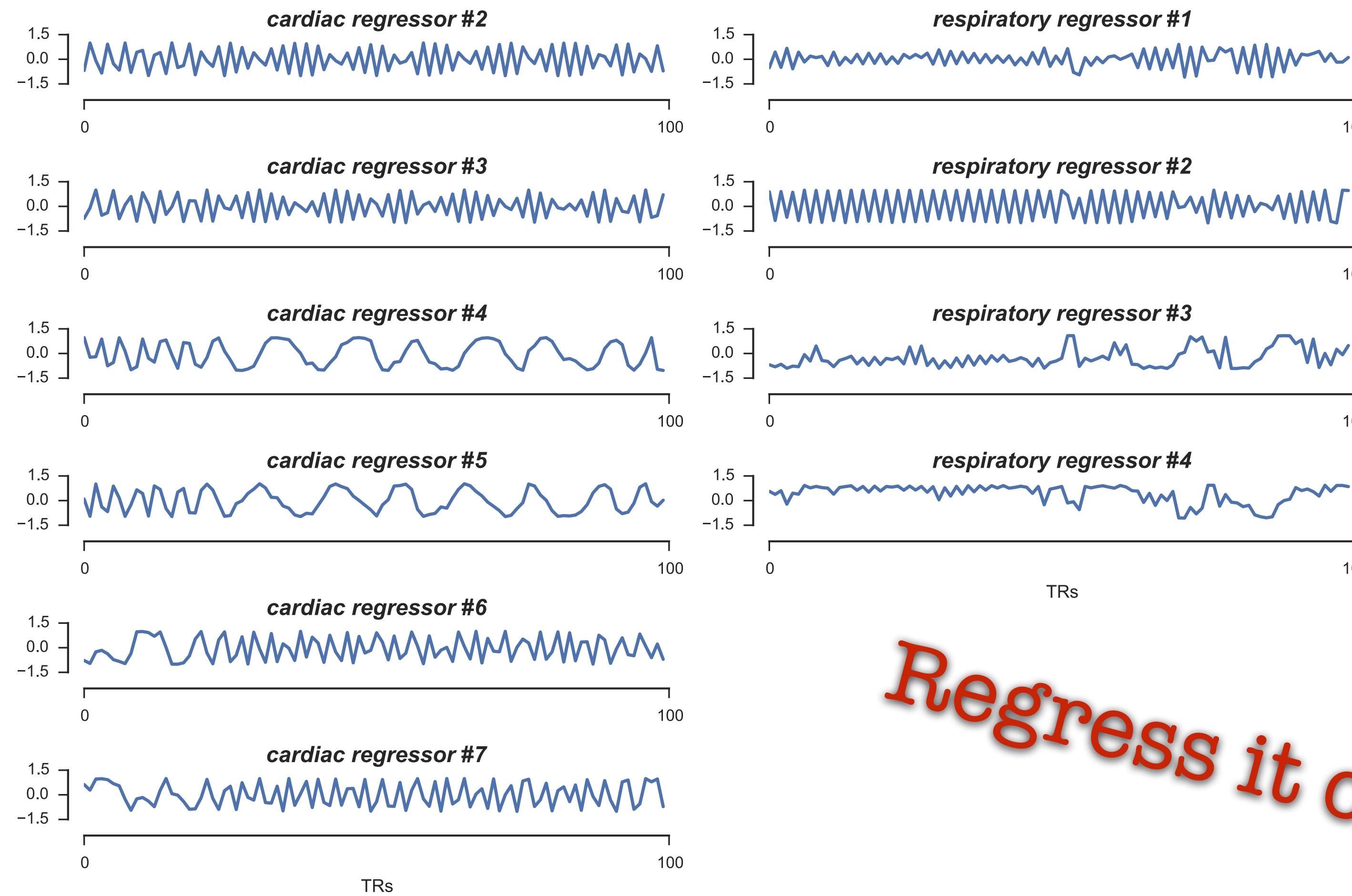
Aliasing

These factors are sampled into the slow fMRI acquisition



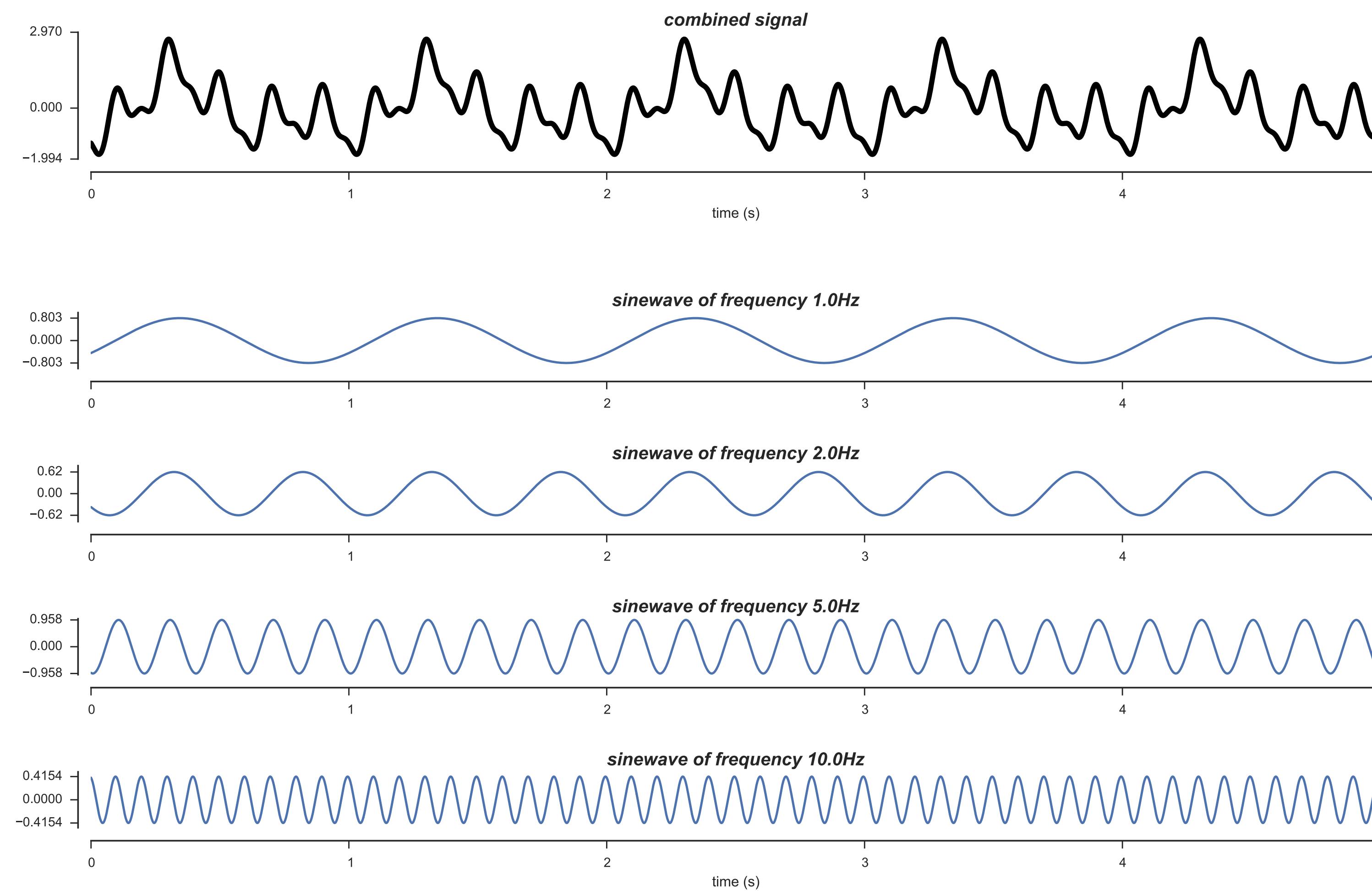
“RETROICOR” nuisance regressors

RETROspective Image CORrection

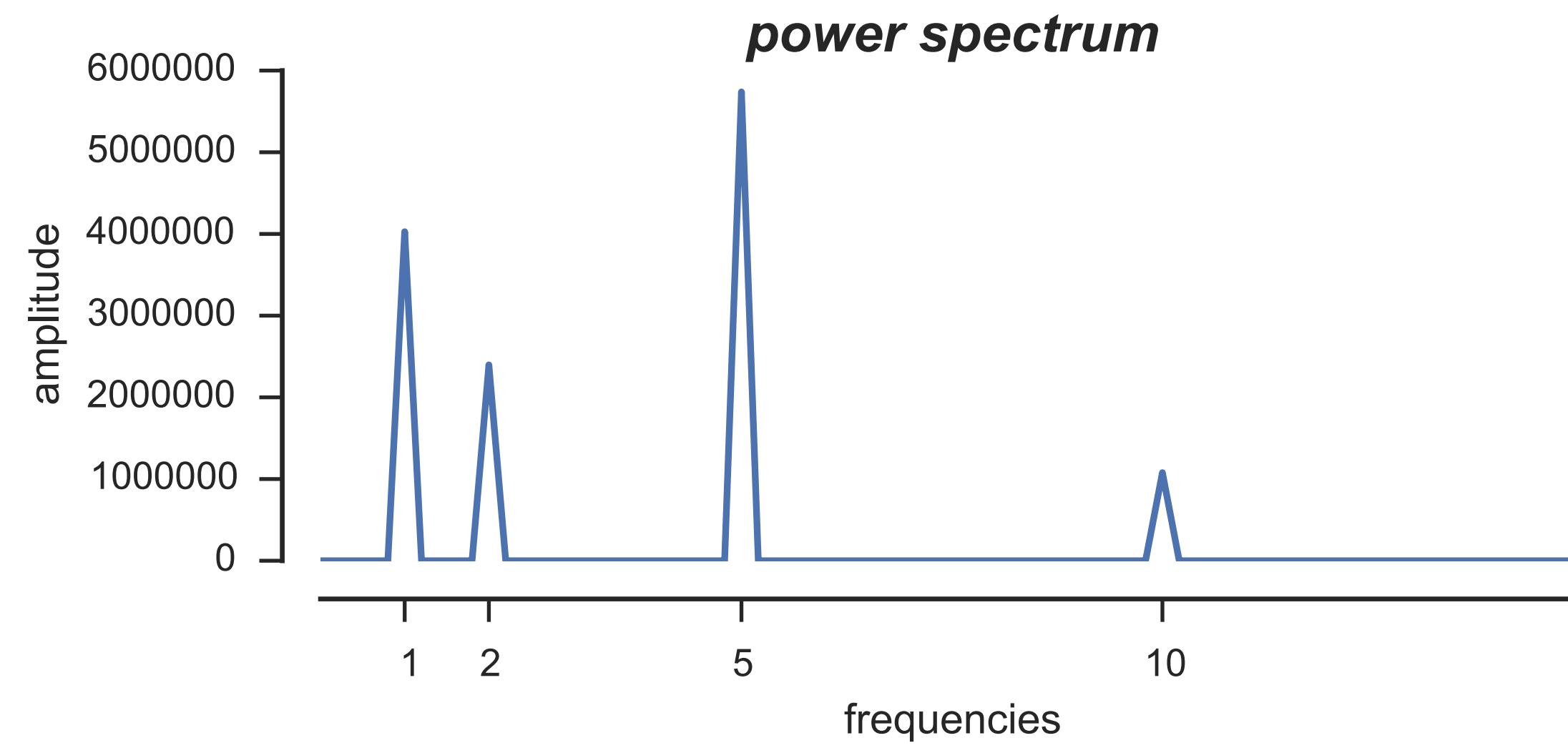
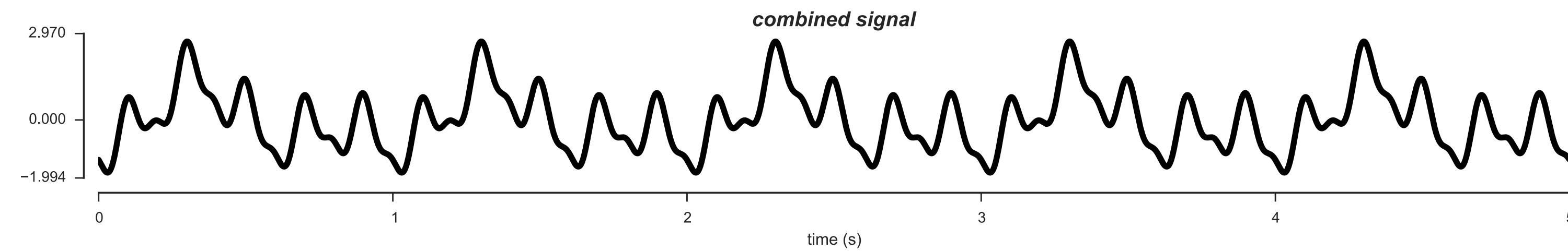


Fourier Transform:

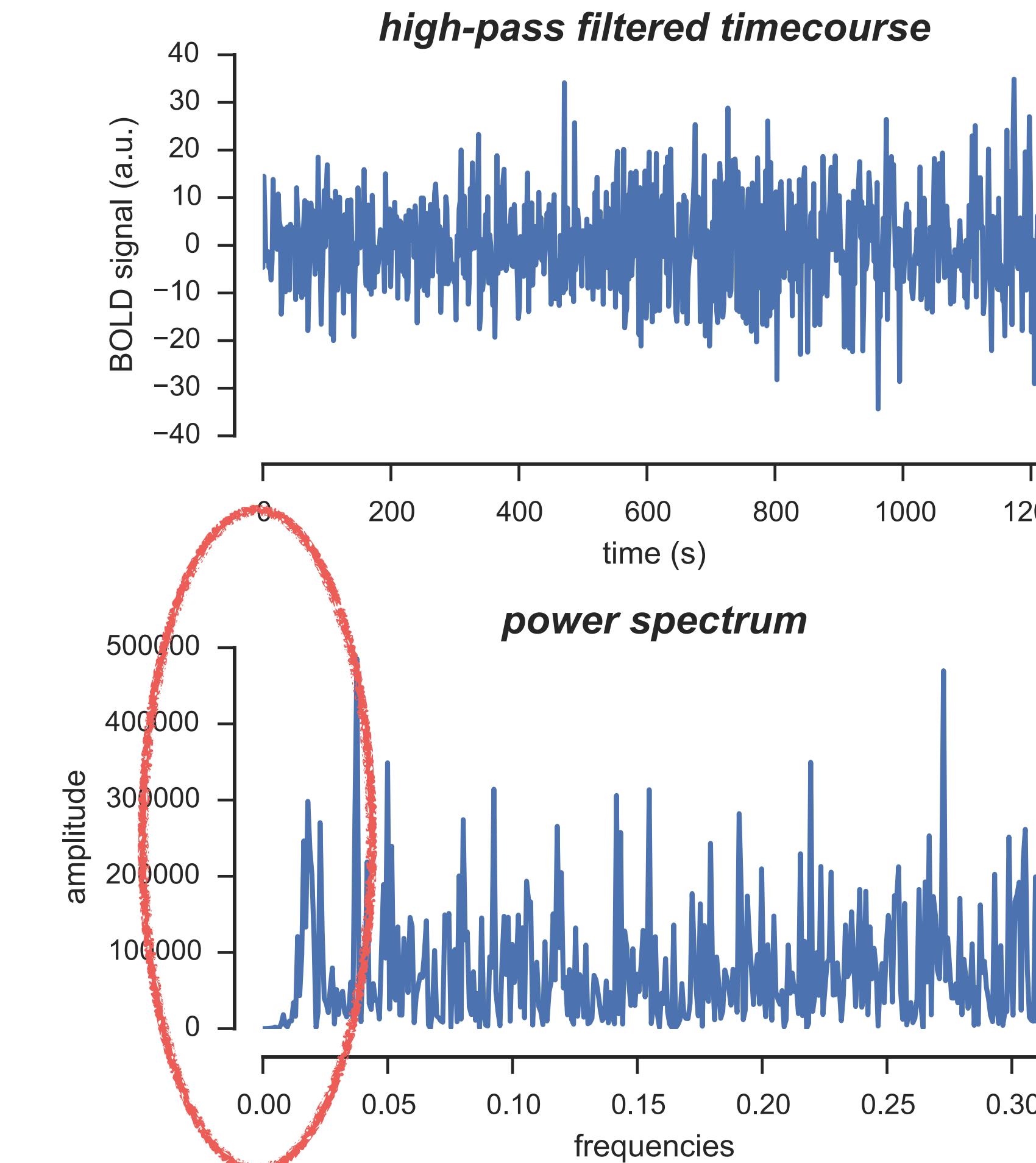
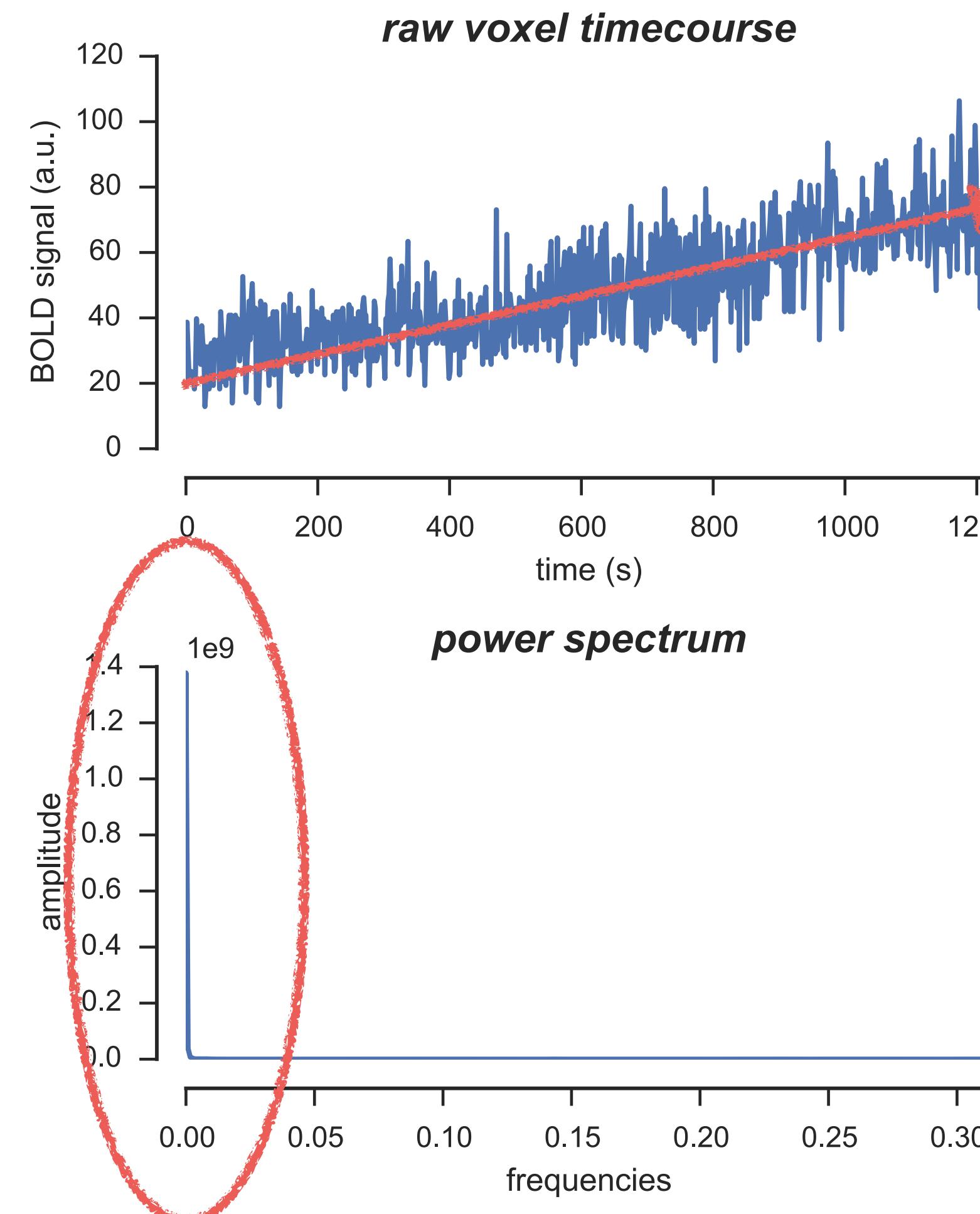
sine-waves as building blocks for continuous signal



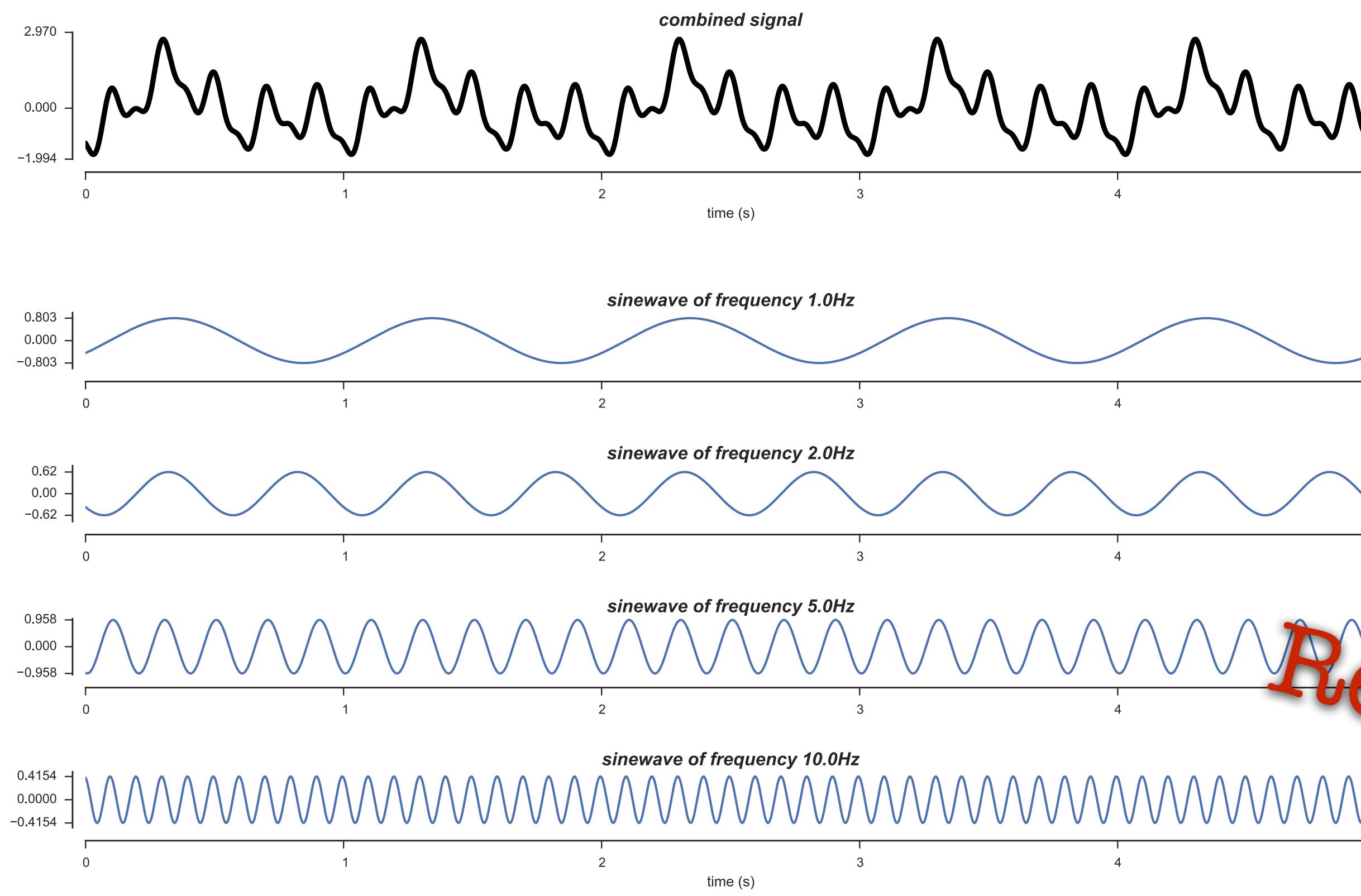
Fourier decomposition



real fMRI data



Use GLM for high-pass filtering



Discrete cosine
basis set
as
regressors
of slow drifts

*R*egress it out!

More and more regressors

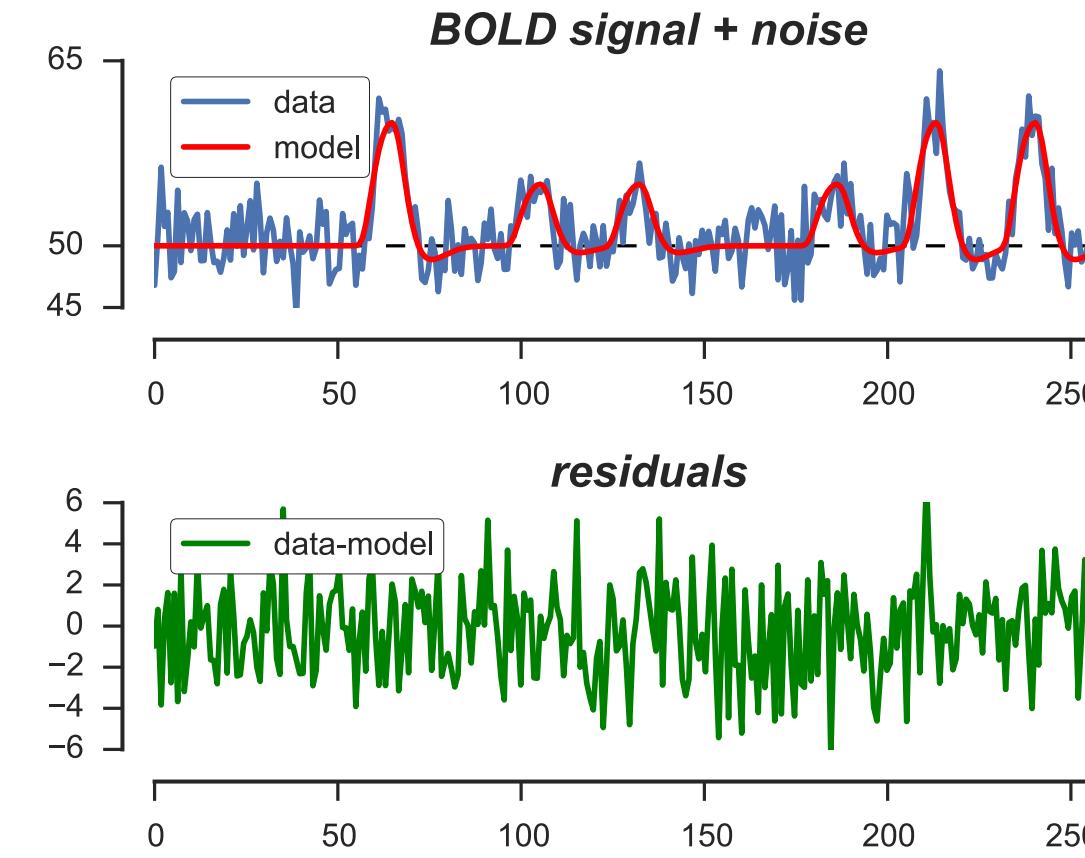
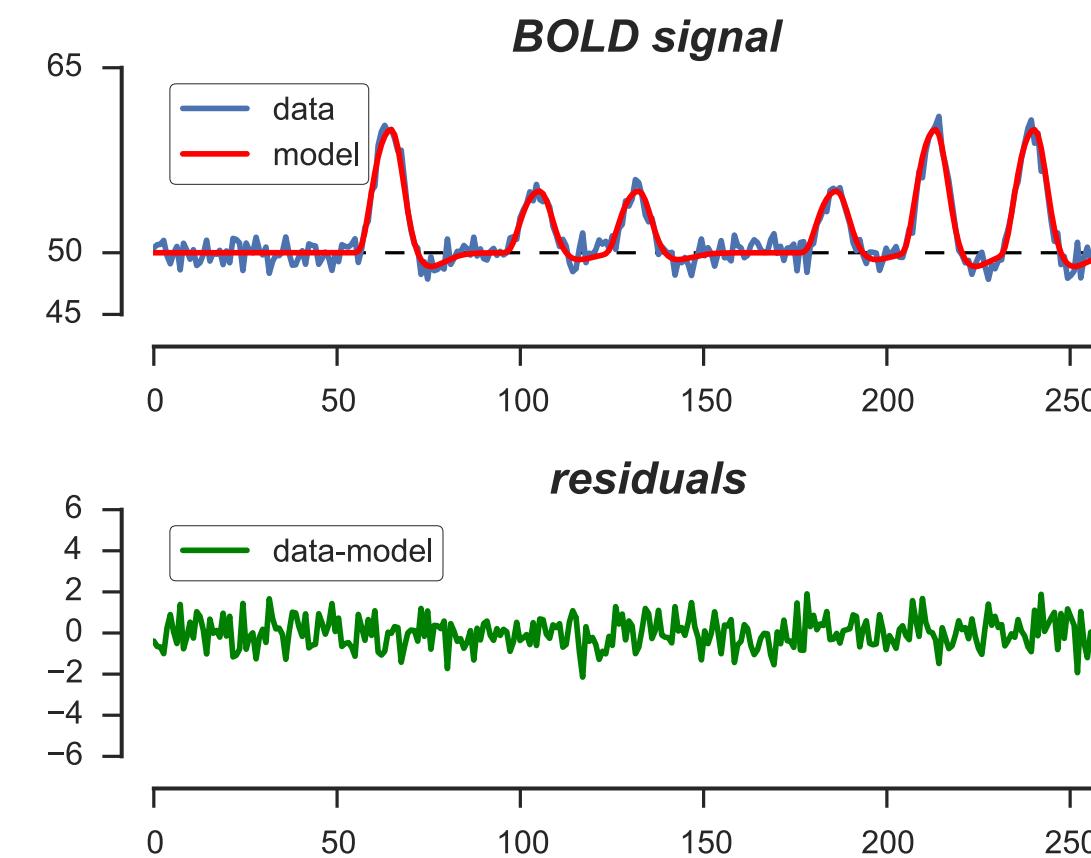
You cannot keep adding regressors to your design matrix.

Some is good - if they capture variance.

But adding more will cause the degrees of freedom in your analysis go down.

This decreases your t-statistic.

In ML context: **overfitting?**



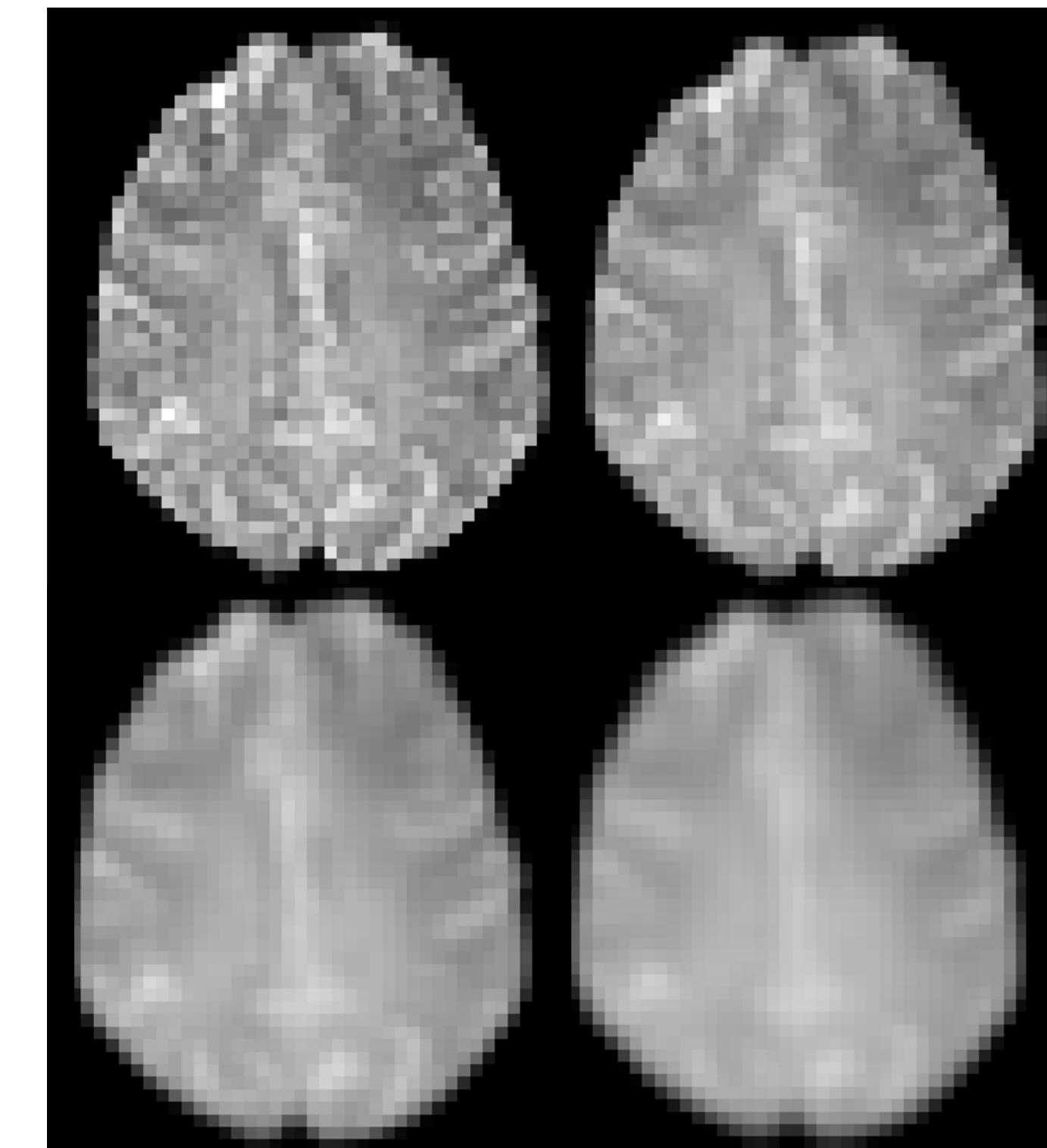
$$\frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{df}$$

spatial interpolation

Interpolate as little as possible, and only with sinc method

30	13	15	28
24	10	24	16
14	22	7	12
11	20	8	5

original



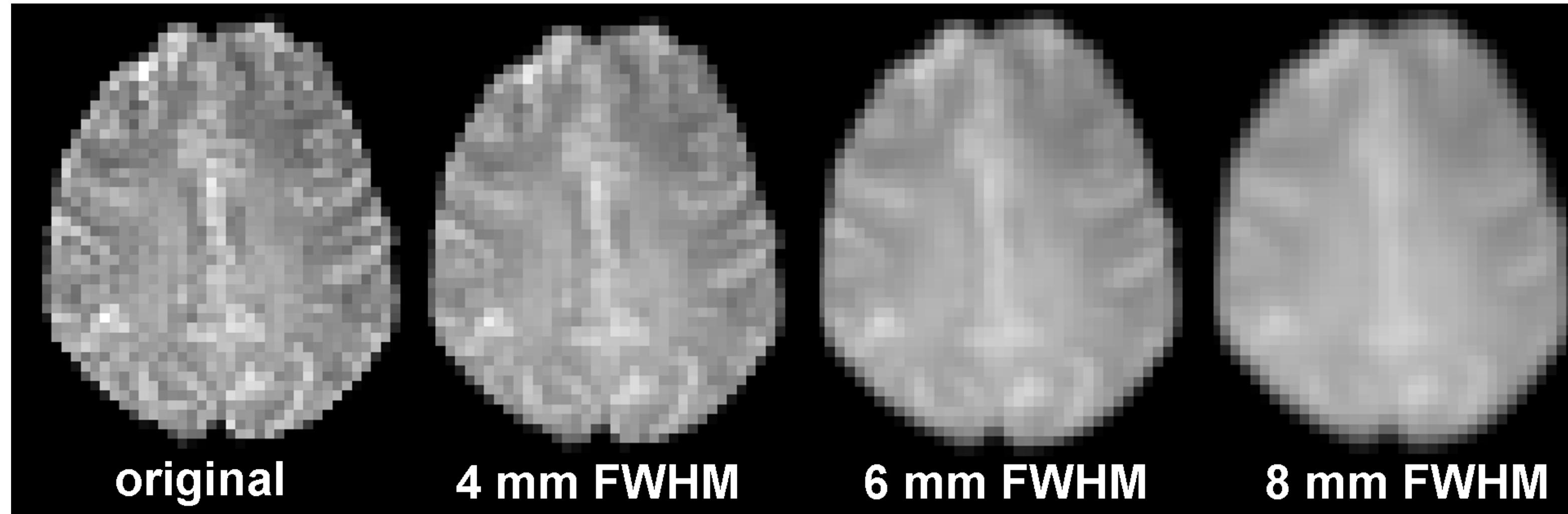
2 times
interpolation

1 time
interpolation

4 times
interpolation

spatial smoothing

Increases SNR - decreases specificity!



Registration

What is where?

fMRI

Where are we measuring things?

We register EPI to T1, within-subject correspondences. Now how about across observers?

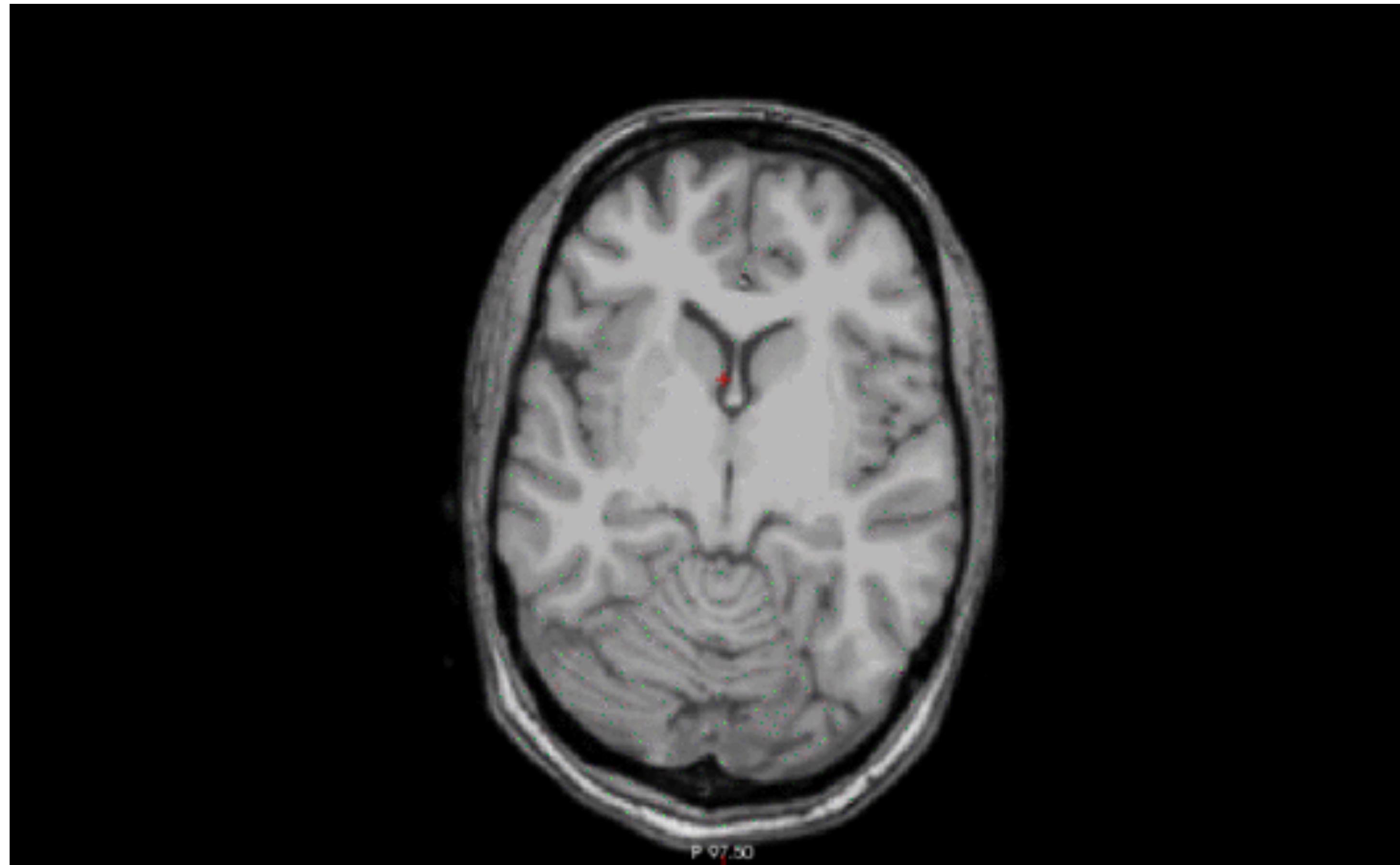
The correspondence problem:

how do we relate a voxel in one person's brain to a voxel location in another person's brain?

What could be possible problems?

registration across participants

We are all unique

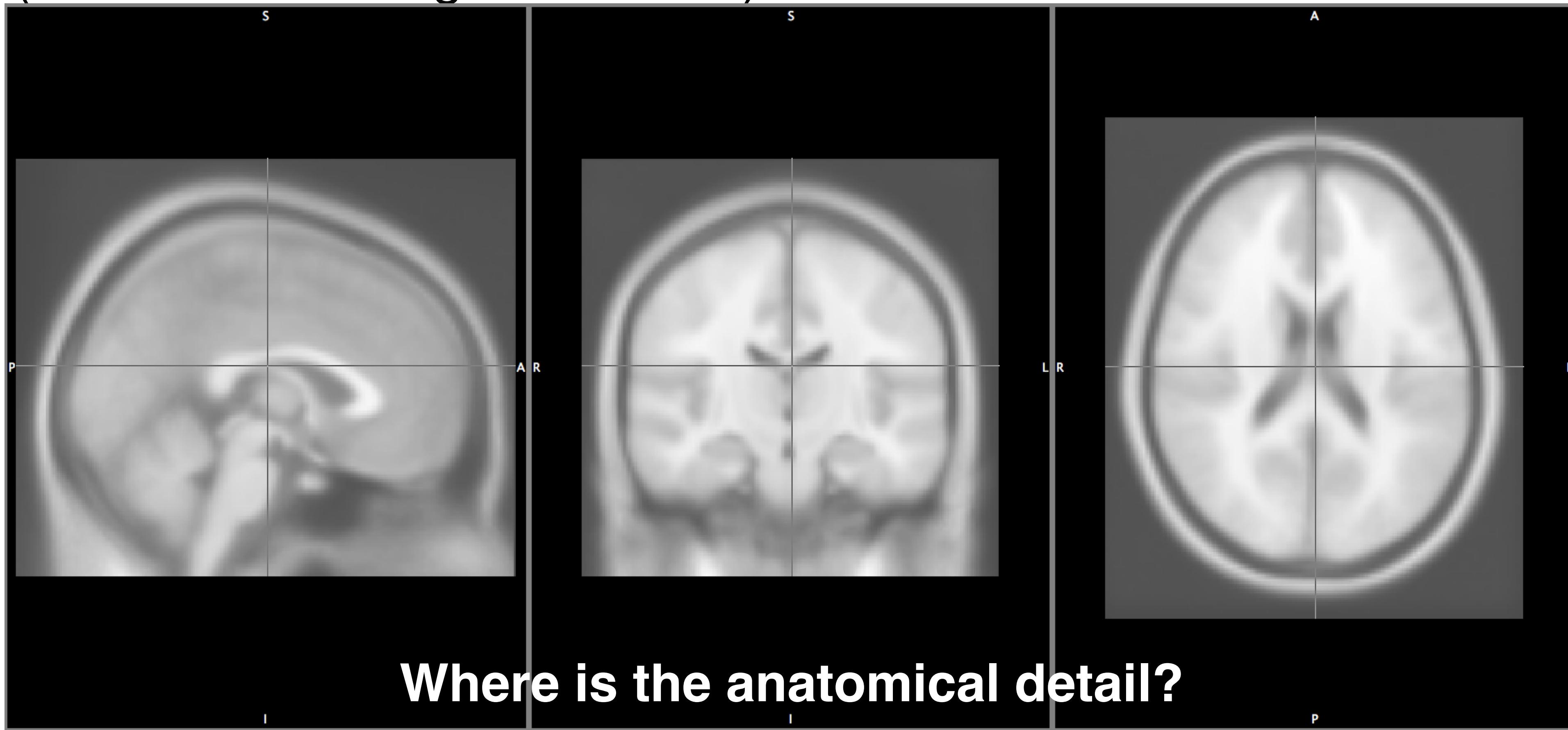


registration across participants

- Now, that sucks: people's anatomies are variable, with sulci and gyri at completely different locations.
- How are we going to work this out?
- Separate solutions:
 - Nonlinear registration to **standard brain**
 - **Surface-based registration**
 - Others; Region Of Interest (**ROI**) based analysis

registration across participants

- MNI (Montreal Neurological Institute) standard brain:



Standard ‘OLD’ recipe

1. First, linear registration of subject T1 anatomy to MNI standard brain template
2. Second, Non-linear registration to MNI, compensating for differences in anatomy

1.

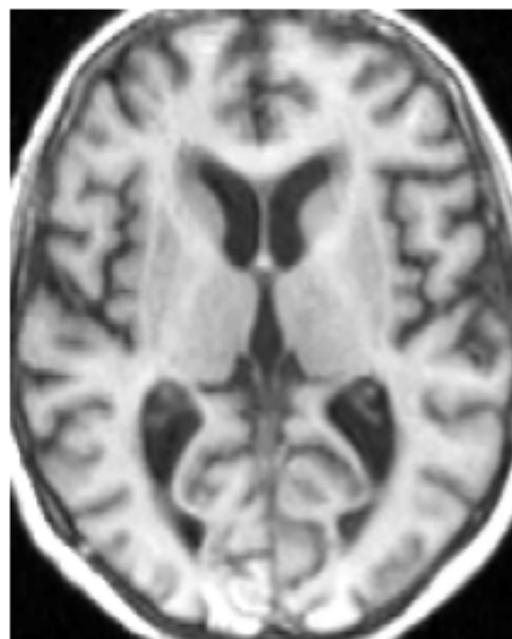
$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$



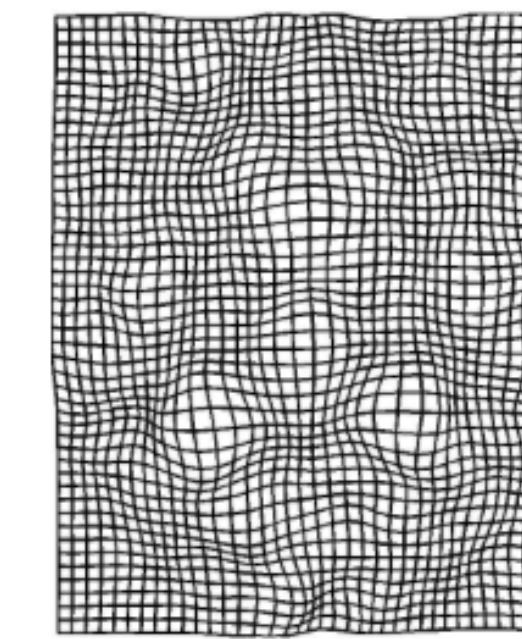
Standard ‘OLD’ recipe

Multi-stage registration from subject anatomy to MNI:

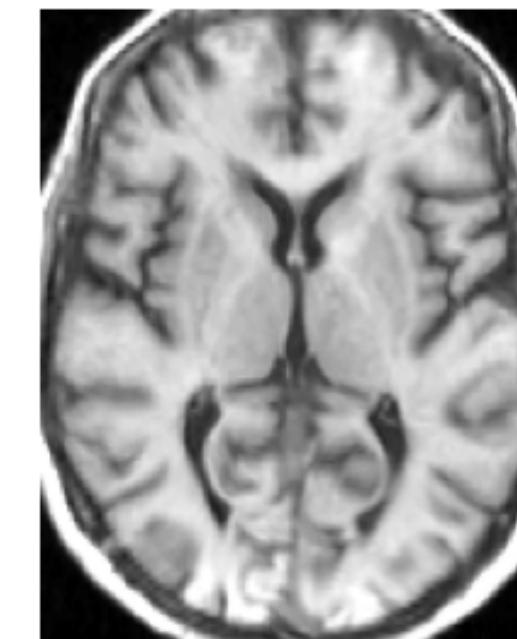
Linear
Affine
(FLIRT)



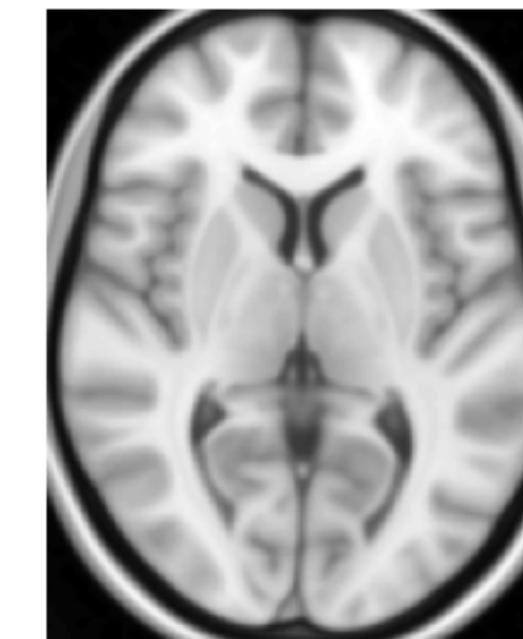
Deformation field



Nonlinear
warped
(FNIRT)



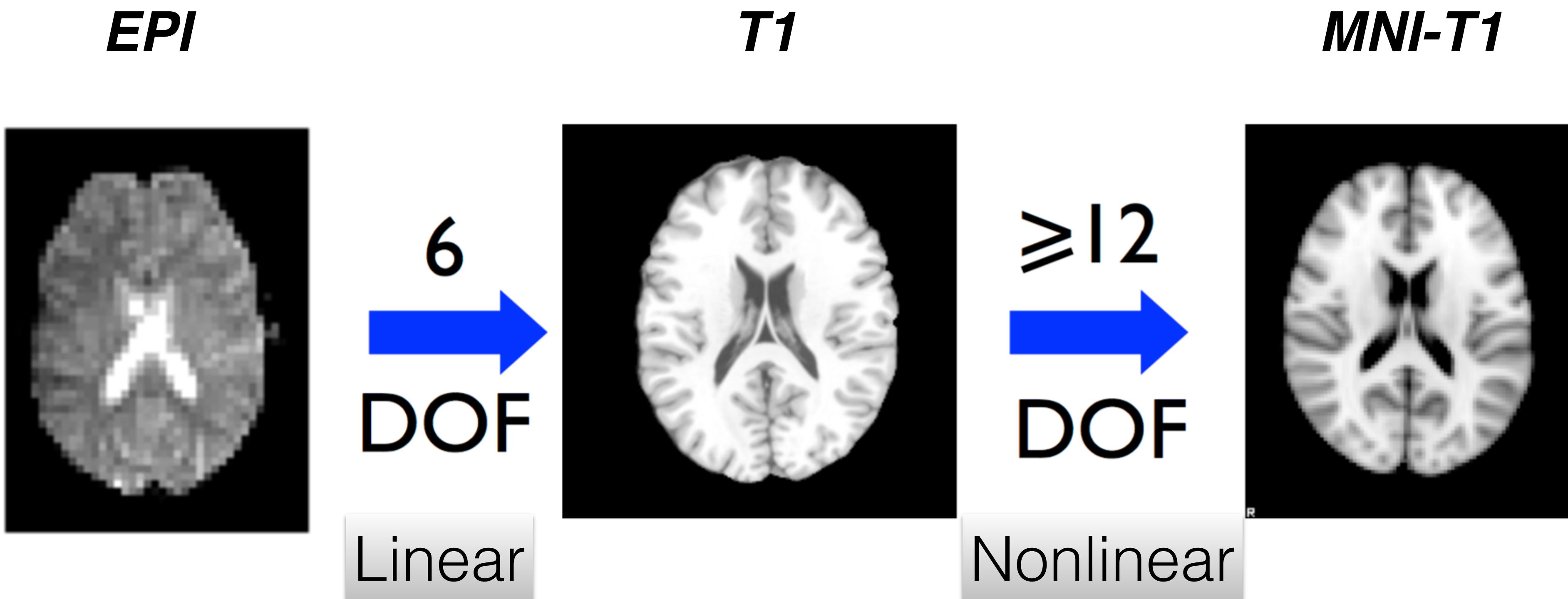
MNI
template



Pros: easy, forgiving
Cons: gyrus/sulcus patterns are warped into one another - **Ugly!**

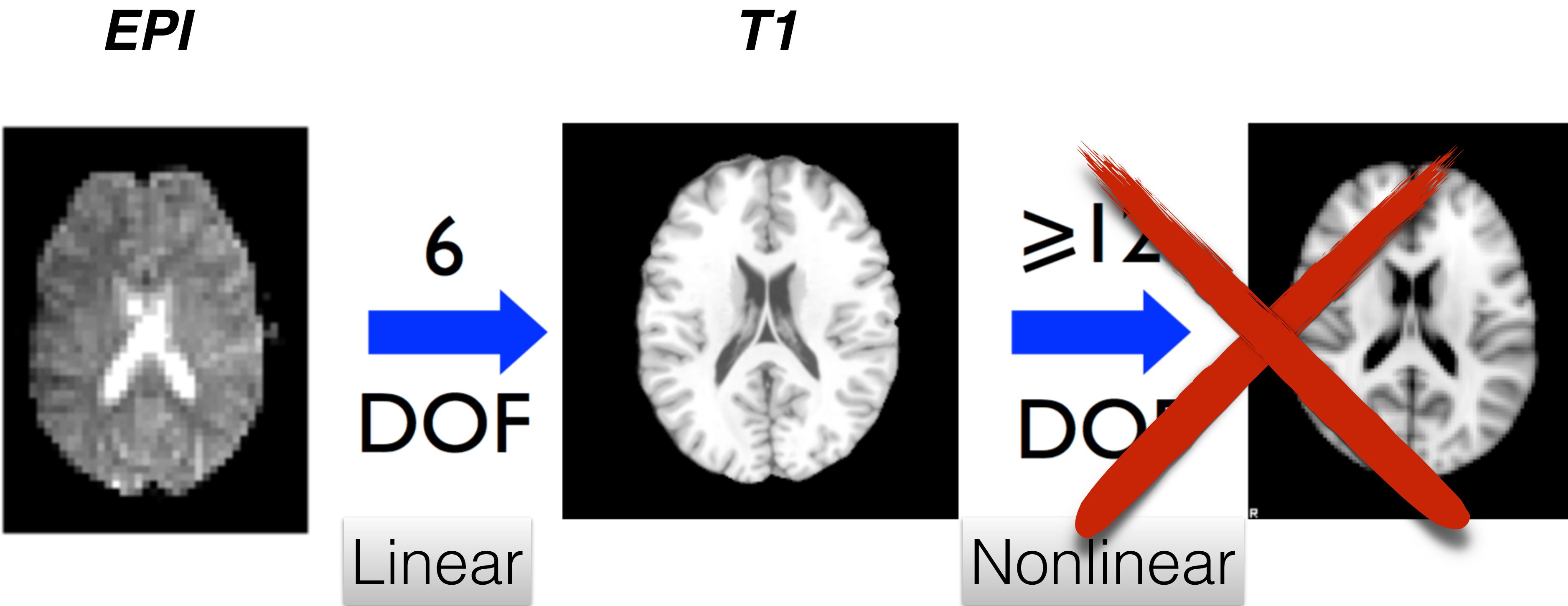
Standard *fMRI* registration

Multi-multi-stage registration from subject anatomy to MNI:



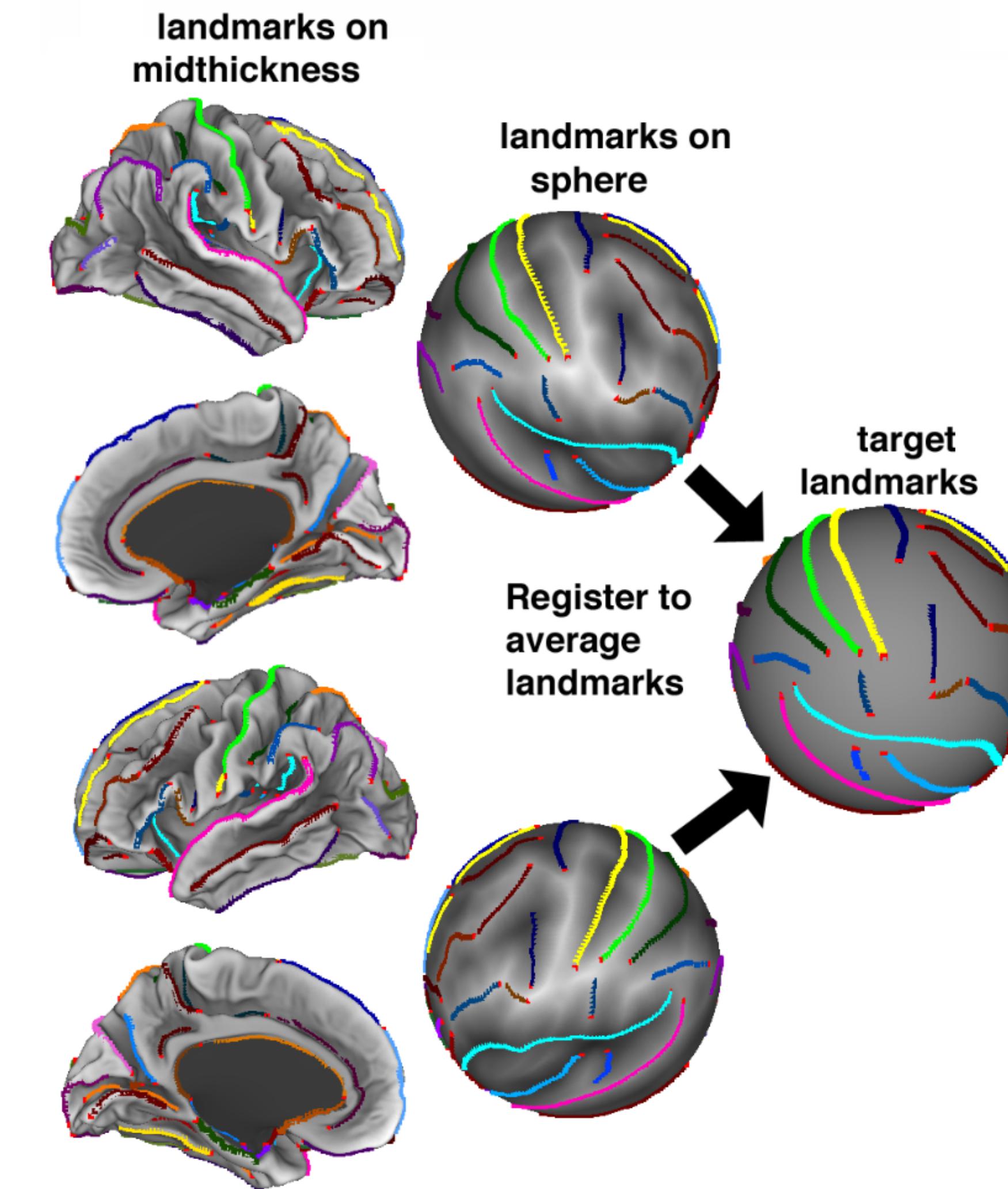
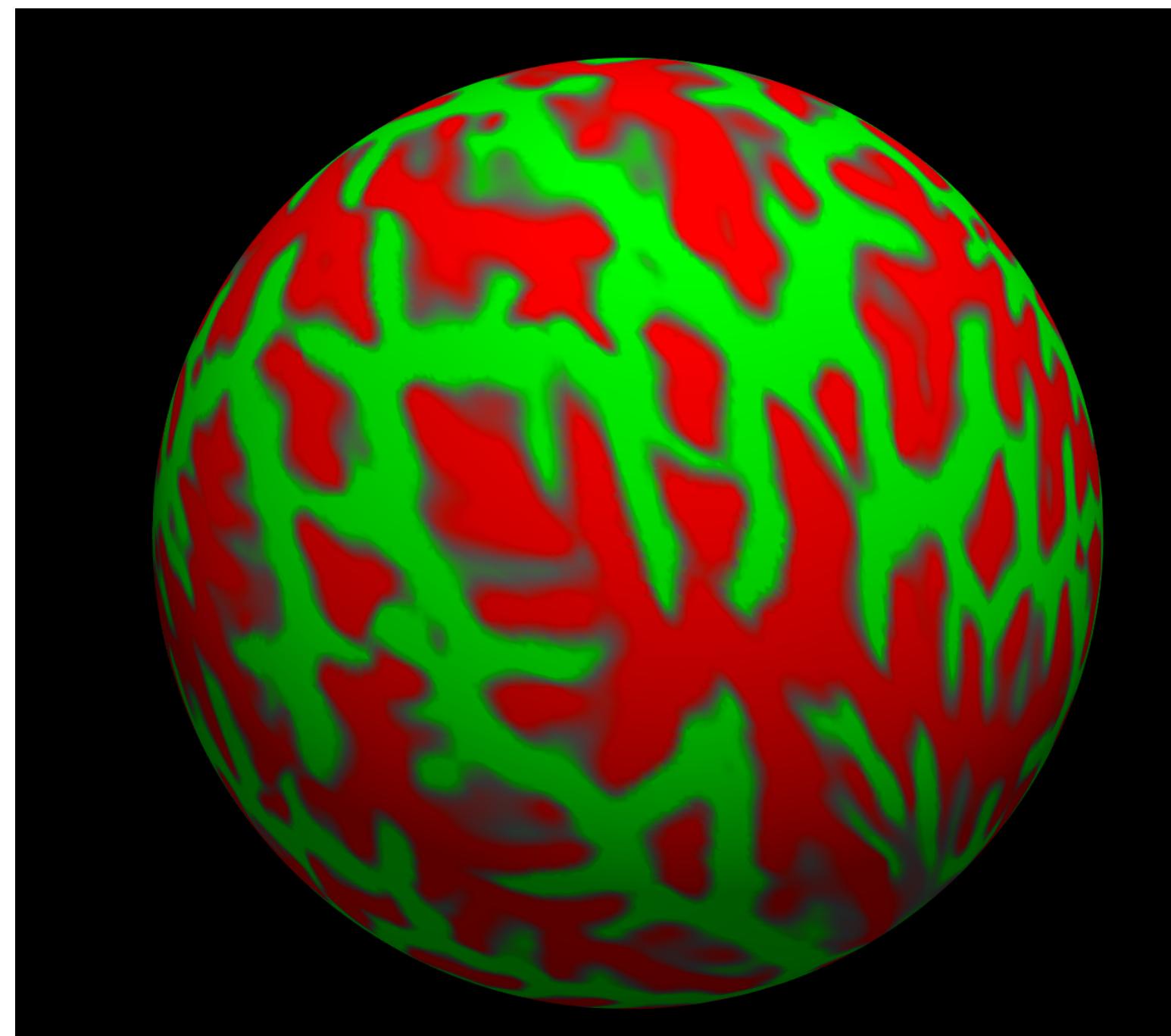
Same pros and cons as for anatomy.

Surface-based registration



Surface-based registration registers the subjects' anatomy to a standard surface, not a standard volume.

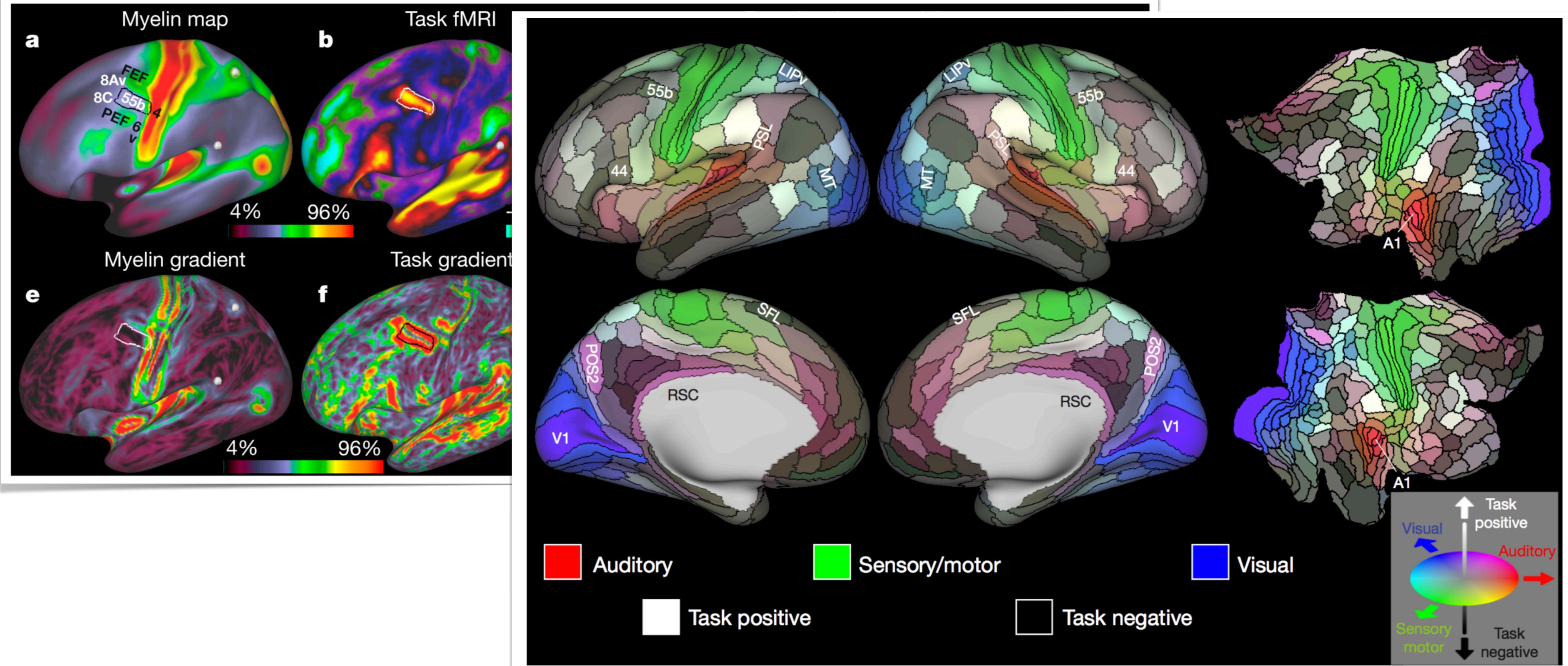
Surface-based registration



Surface-based registration registers the subjects' anatomy to a standard surface, not a standard volume: uses folding information - FNIRT does not!

Multimodal registration

Using as much information as possible



Reproducible fMRI

New methods for ensuring reproducibility

Processing step	Reason	Options [suboptions]	Number of possible options
Motion correction	Correct for head motion during scanning	• ‘Interpolation’ [linear or mean], ‘Registration’ [rigid or affine]	2 × 2 = 4
Slice timing correction	Correct for differences in acquisition timing of different slices	‘Linear’ or ‘Non-linear’	2
Field map correction	Correct for differences in magnetic field across the brain	‘Linear’ or ‘Non-linear’	2
Spatial smoothing	Smooth the data to reduce noise	‘Gaussian’ [FWHM] or ‘Mean’ [SNR]	2
Regional regression	Remove confounds from the data	‘Linear’ or ‘Non-linear’	2
Haemodynamic response	Model the blood oxygen level-dependent signal	‘Linear’ or ‘Non-linear’	2
Temporal autocorrelation model	Model the temporal correlation between time points	‘Linear’ or ‘No’	2
Multiple-comparison correction	Correct for multiple comparisons across the brain	‘Voxel-based GRF’, ‘cluster-based GRF’, ‘FDR’ or ‘non-parametric’	4
Total possible workflows			69,120

BIDS: brain imaging data structure
 Everyone saves their files in exactly the same way

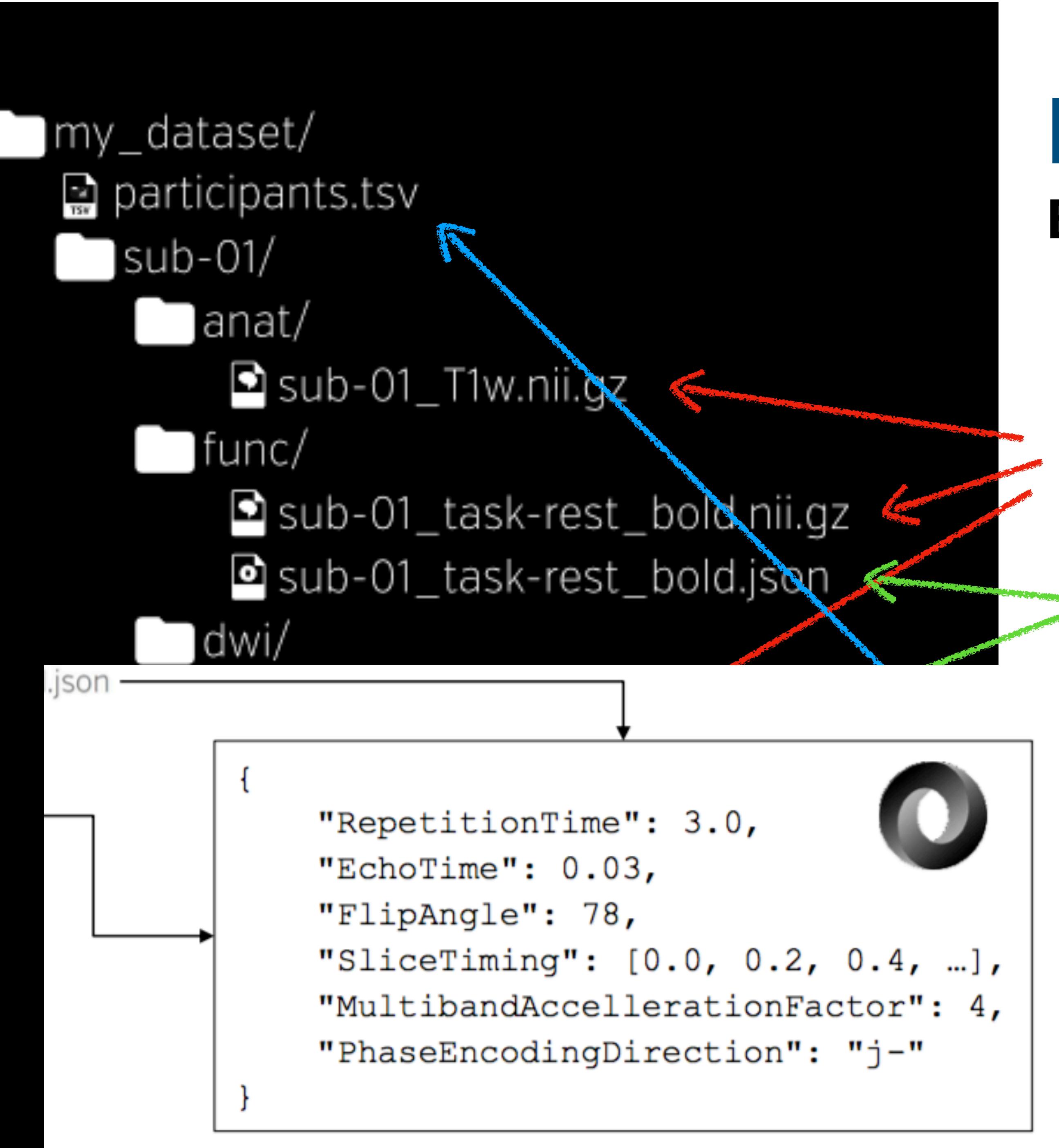
Fmriprep:
 Reproducible preprocessing,
 best of all worlds,
 dedicated programmers,
 completely open science/source

Brain Imaging Data Structure



BIDS

Brain Imaging Data Structure



- Nifti MRI files: standard file type
- Metadata in .json files
- Behavioral, etc. data in .tsv files

BIDS bids.neuroimaging.io

Ensure compatibility

- Once you create a BIDS data format, you need to make sure it ***conforms to the specification***
- Run it through the validator!

incf.github.io/bids-validator

The screenshot shows the BIDS Validator interface. At the top, there are three sections: 'Summary', 'Available Tasks', and 'Available Modalities'. The 'Summary' section lists: • 40 Files, 18.42kB • 13 - Subjects • 1 - Session. The 'Available Tasks' section lists: • rhyme judgment. The 'Available Modalities' section lists: • bold • T1w. Below these sections, a message states: Your dataset is not a valid BIDS dataset. A red button below the message says: view 1 error in 23 files. A yellow button below the red one says: view 1 warning in 4 files.

Summary	Available Tasks	Available Modalities
<ul style="list-style-type: none">• 40 Files, 18.42kB• 13 - Subjects• 1 - Session	<ul style="list-style-type: none">• rhyme judgment	<ul style="list-style-type: none">• bold• T1w

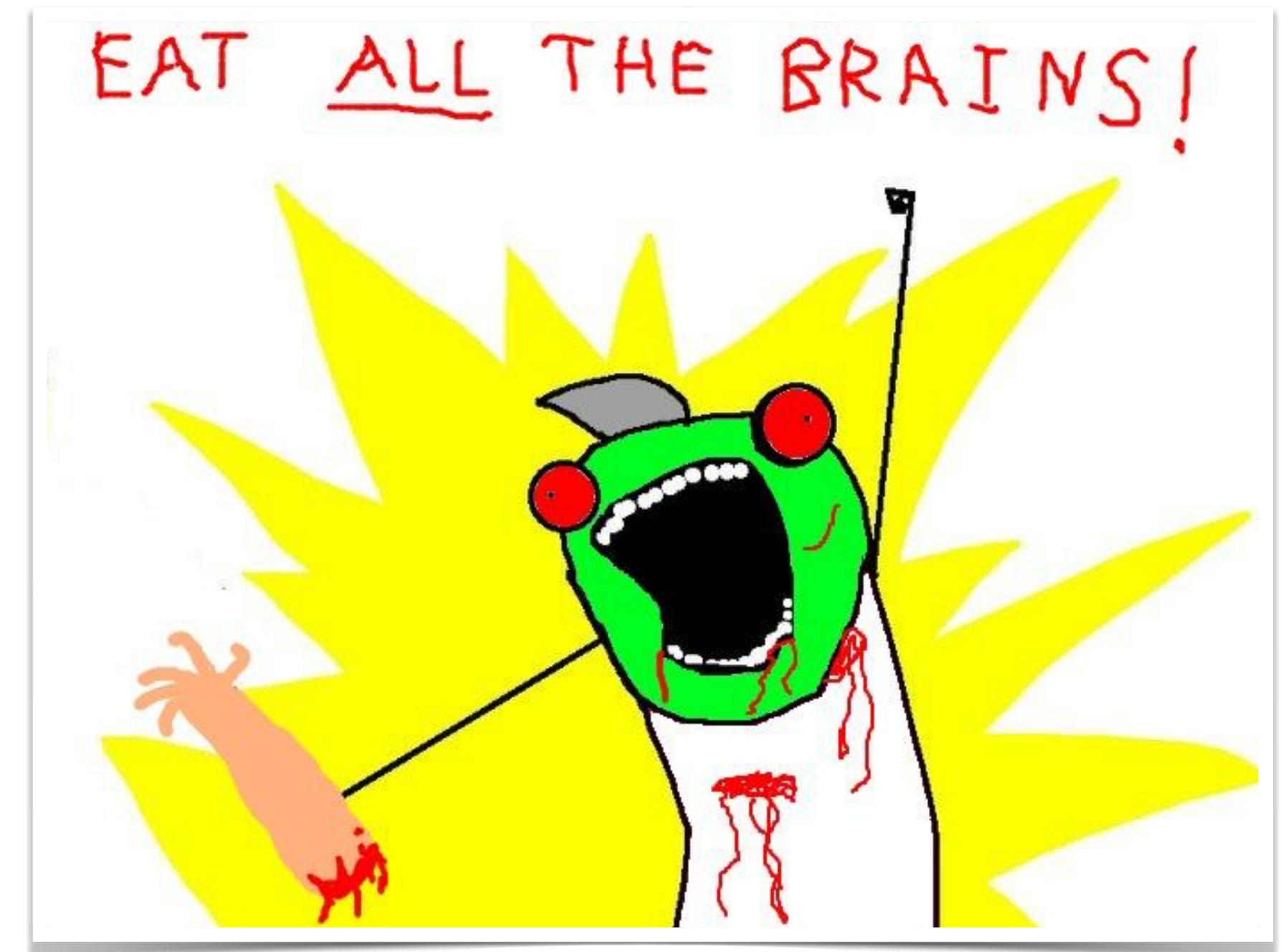
Your dataset is not a valid BIDS dataset.

view 1 error in 23 files

view 1 warning in 4 files

Now that we've got BIDS what are we going to do with it?

- Automate everything!
- BIDS apps:
 - Operating system containers
(no need to install software, run anywhere).
 - Single standardised command line interface.



Available BIDS Apps

BIDS-Apps/example	version 0.0.7	open bug issues 0	build passing	open bug pull requests 0	docker pulls 8k	439.5MB 23 layers
BIDS-Apps/freesurfer	version v6.0.1-3	open bug issues 0	build failed		docker pulls 3k	2.6GB 52 layers
BIDS-Apps/ndmg	version v0.1.0	open bug issues 0	build passing		docker pulls 7k	920.9MB 31 layers
BIDS-Apps/BROCCOLI	version v1.0.1	open bug issues 1	build passing	open bug pull requests 0	docker pulls 256	3GB 21 layers
BIDS-Apps/FibreDensityAndCrosssection	version v0.0.1	open bug issues 0	build passing	open bug pull requests 0	docker pulls 72	576.8MB 31 layers
BIDS-Apps/SPM	version v0.0.14	open bug issues 0	build passing	open bug pull requests 0	docker pulls 872	1.6GB 24 layers
poldracklab/mriqc	version 0.10.3		build passing	open bug pull requests 0	docker pulls 16k	2.6GB 37 layers
BIDS-Apps/QAP	Image not found	open bug issues 0	build passing		docker pulls 7	Image not found
BIDS-Apps/CPAC	version v1.0.2_d...		build passing	open bug pull requests 0	docker pulls 2k	1.4GB 38 layers
BIDS-Apps/hyperalignment	Image not found		build passing		docker pulls 3	Image not found
BIDS-Apps/mindboggle	version 0.0.4-1	open bug issues 2	build passing		docker pulls 376	1.9GB 81 layers
BIDS-Apps/MRtrix3_connectome	version 0.2.2	open bug issues 0	build passing	open bug pull requests 0	docker pulls 345	3.4GB 56 layers
BIDS-Apps/rs_signal_extract	version 0.1		build passing	open bug pull requests 0	docker pulls 75	240MB 17 layers
BIDS-Apps/aa	version enh_vario...		build failed		docker pulls 61	3.8GB 57 layers
BIDS-Apps/niak	version latest		build passing		docker pulls 110	2.7GB 103 layers
BIDS-Apps/oppni	version v0.7.0-1		build passing		docker pulls 137	2.9GB 41 layers
poldracklab/fmriprep	version 1.0.8		build passing	open bug pull requests 0	docker pulls 30k	4.4GB 48 layers
BIDS-Apps/brainiak-srm	version latest		build failed		docker pulls 79	559.3MB 13 layers
BIDS-Apps/nipypipelines	version 0.3.0		build passing		docker pulls 67	478.1MB 20 layers
BIDS-Apps/HCPPipelines	version v3.17.0-15		build passing		docker pulls 495	2.5GB 62 layers
BIDS-Apps/MAGeTbrain	Image not found		build passing		docker pulls 149	Image not found
BIDS-Apps/tracula	version v6.0.0-4		build passing	open bug pull requests 0	docker pulls 345	3.4GB 57 layers
BIDS-Apps/baracus	Image not found		build passing		docker pulls 809	Image not found
BIDS-Apps/antsCorticalThickness	Image not found		build passing	open bug pull requests 0	docker pulls 18	Image not found
BIDS-Apps/DPARSF	version v4.3.12		build passing		docker pulls 109	1.4GB 28 layers
BIDS-Apps/afni_proc	Image not found		build passing	open bug pull requests 0	docker pulls 48	Image not found

http://bids-apps.neuroimaging.io/tutorials/

Quality problems

Data are never perfect

- Do you go through all your scans for visual inspection?
- (f)MRI data can suffer from lots of problems that can be hard to catch.
- MRIQC automates this to a large extent.

