What is a "preprocessing"?

- Preprocessing is a series of steps that, in brief,

  - Corrects acquisition artefacts

  - Co-registers functional and anatomical volumes

  - (Normalises your data to a template)

  - Improves your SNR

- Divided into Anatomical and Functional

Is it important (like, for real)?

Is it important (like, for real)?

YES.
(like, really)

Is it important (like, for real)?

- What you do (or don't do) during preprocessing will affect your data analysis, your results and the way you can interpret it.

- For this reason, (some) preprocessing steps are always highly debated.

- It is good habit to *report* your preprocessing when you present results.

- It is even better habit to *report* it in advance (e.g. during project presentations or preregistration)!

Myth: standard preproc

- (Almost) every step is optional, sometimes even the order might change.

- Different tools perform different preproc.

- There is no such thing as a standard preprocessing. (There might be a standard preprocessing in ${tool}).

- There is no "better" preprocessing per se.

- But there are good and bad ways to do it.

How to choose a preprocessing

- Option 1: use your standard tool preprocessing
  (but know what it is doing and what limitations it implies).

- Option 2: refer to previous literature for comparison
  (and use their data acquisition).

- Option 3: refer to a consortium e.g. HCP
  (but then use their sequences to acquire).

- Option 4: choose your steps depending on theoretical reasons
  and your data.

Reproducible preprocessing

- The fact that there's no standard preprocessing, and the fact that it depends on your OS and libraries, makes this step hard to reproduce.

- In order to improve the reproducibility of your process, you can use community tools (e.g. fmriprep).

- Even better, you can "freeze" the environment you do your computation in using BIDS App.

- If you only use python, you can use virtualenv.

Where can I find more about it?

Suggestion: always have at least one of these books around!

- Bijsterbosch, J., Smith, S. M., & Beckmann, C. F. (2017). **Introduction to Resting State FMRI Functional Connectivity.** Oxford University Press.

- Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2014). **Handbook of Functional MRI Data Analysis.** Cambridge University Press.

- Jenkinson, M., & Chappell, M. (2018). **Introduction to Neuroimaging Analysis.** Oxford University Press.

- Ashby, F. G. (2011). **Statistical Analysis of FMRI Data.** MIT Press.

- Wager, T. D., & Lindquist, M. A. (2015). **Principles of fMRI.** Leanpub.

Where can I find more about it?

More bibliography, if you are interested (and a bit masochist)

- Bright, M. G., & Murphy, K. (2017). Cleaning up the fMRI time series: **Mitigating noise with advanced acquisition and correction strategies.** NeuroImage, 154(March), 1–3. http://doi.org/10.1016/j.neuroimage.2017.03.056

- Caballero-Gaudes, C., & Reynolds, R. C. (2017). **Methods for cleaning the BOLD fMRI signal.** NeuroImage, 154(December 2016), 128–149. http://doi.org/10.1016/j.neuroimage.2016.12.018

- Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Ayse, I., Erramuzpe, A., ... Gorgolewski, K. J. (2018). **FMRIPrep : a robust preprocessing pipeline for functional MRI**, 5, 1–20. http://doi.org/10.1101/306951

- Liu, T. T. (2016). **Noise contributions to the fMRI signal: An overview.** NeuroImage, 143, 141–151. http://doi.org/10.1016/j.neuroimage.2016.09.008

- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). **Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion.** NeuroImage, 59(3), 2142–2154. http://doi.org/10.1016/j.neuroimage.2011.10.018

- Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). **Methods to detect, characterize, and remove motion artifact in resting state fMRI.** NeuroImage, 84, 320–341. http://doi.org/10.1016/j.neuroimage.2013.08.048

- Power, J. D., Plitt, M., Gotts, S. J., Kundu, P., Voon, V., Bandettini, P. A., & Martin, A. (2018). **Ridding fMRI data of motion-related influences: Removal of signals with distinct spatial and physical bases in multiecho data.** Proceedings of the National Academy of Sciences, 201720985. http://doi.org/10.1073/PNAS.1720985115

First step of a good preprocessing

First step of a good preprocessing

LOOK AT YOUR DATA

First step of a good preprocessing

- Check your data while you acquire: some artefacts or mistakes can be corrected at the moment!

- Preprocessing can help, but it can't do miracles.

- Sometimes it is just better to acquire your data again.

Anatomical preprocessing

- Bias Field Correction

- Brain extraction

- Tissue segmentation

- Normalisation to template

Functional preprocessing: Correction

- Discard Trs

- Despike

- Slice Interpolation

- Blip aka Topup (susceptibility induced distortions correction)

- Motion computation

Functional preprocessing: Normalisation

- Co-registration to anatomical volume(s)

- Normalisation to template

Functional preprocessing: Noise correction

- Smoothing

- Scrubbing (censoring)

- Detrending

- Nuisance regression

  - Average tissue compartment signal regression (WM and/or CSF)

  - Global Signal Regression

  - CompCorr

  - ICA-based Clean-up

- Physiological denoising

- Bandpass

Discard TRs

- Sometimes, the magnetic field has to "stabilise"

- If the scanner does not do it, it is a good idea to eliminate the first 10-20 seconds

Despike

- Sometimes hardware artefacts can cause spikes in your signal

- It is possible to interpolate the signal to get rid of such spikes

- Might be skipped if the data analysis is frequency-dependent (or strictly time dependent)

Slice Interpolation

- During a TR, the slices of a volume are acquired at a (slightly) different time

- It might be a good idea to shift such slices to take that into account

- Not suggested with TR under 1" (the correction can be ignored) or above 2" (due to the BOLD response, you might introduce artifacts)

- In task-fMRI, it might be better to interpolate your convolved regressors instead (FSL does it automatically)

Blip (topup)

- The phase encoding direction in which you acquire your data has an impact on the signal that you record

- If you acquire a volume in the inverse direction, you can correct this impact by warping the two images together (but use a "toward middle point" warping, not a "source to destination" one)

Motion computation

- It's inevitable. Even the best subject moves
  (Unless it's dead. But then fMRI doesn't really make sense.
  And even then you might have induced movements.)

- It is possible to register different volumes together, and
  measure the displacement to later eliminate the related data
  variance

- With particular groups of interest, normal motion
  computation might introduce biases

Anatomical co-registration

- Needed for nuisance regression and normalisation (due to better resolution and tissue contrast in anatomical volumes)

- As normally anatomical volumes have higher resolution and better contrast (and they are less interesting in fMRI), it is a good idea to coregister the anatomical to the functional volume

- Linear (affine) registration (or BBR if you have enough contrast)

Normalisation (1)

- Necessary to compare different subjects or different studies or to extract ROIs' signal

- Not necessary in single subject voxel-wise analyses

- If you're interested in the average signal of a ROI, register the ROI to the functional space (not the functional to the template space)

- Use non-linear registration (if possible) between anatomical and template

Normalisation (2)

- Can be a multiple-step transformation (from 2 to 5):

    1) Functional (Native) space

    2) Functional (Native) reference

    3) T2w anatomical (T2highres)

    4) T1w anatomical (highres)

    5) Template

- Most of the time, it is a 3-step transformation (1,4,5)

- Generally computed after motion correction, you can apply it in a second moment

- SPM applies it straight away, AFNI at the end of the preproc and FSL even after data analysis!

- Pay attention to your template resolution: don't oversample your data, nor undersample too much!

Smoothing (1)

- Decreases your spatial accuracy, but improves your statistics (as it improves SNR)

- Avoid if the data analysis is ReHo or if you expect activation in little structures of the brain (e.g. olfactory bulbs, sub-thalamic nuclei, …)

- Normally, it is an average of neighbouring voxels, weighted by a Gaussian function, expressed by its Full Width at Half Maximum

- It's safe to use FWHM that are equivalent to 1.5-2 times your resolution

Smoothing (2)

- It violate the GLM pre-requisite of independence (but everyone does it anyway)

- There are post-processing corrections to take that into account

- Pay attention to it (you need it, maybe even more than 2 times your resolution!) if you are using Gaussian Random Field Theory (e.g. SPM)

Scrubbing

- Some (sub)volumes might be just *too* noisy (due to motion or artifacts)

- Scrubbing (or censoring) corrects for that

- It is possible to remove such timepoints, or to set them to 0, or to "interpolate" them

- Avoid if the data analysis is frequency-dependent, or strictly time dependent (e.g. Granger Causality or some Entropy measures), or the interest is in effective connectivity

Demeaning and Detrending

- Some noises have a specific shape, or trend (e.g. heat noise)

- It is possible to remove such noises regressing polynomials.

- The polynomial of 0 order is the mean (normally, timeseries are demeaned at least)

Nuisance

Check dr. Cesar Caballero's slides!

Bandpass filter

- Remove some frequencies of no interest (if they exist)

- It is possible to do just a "lowpass" or "highpass" filter

- If properly corrected for, it reduces DoFs (and worsens your statistics) (hint: a few are properly correcting for it)

- Unless required by your data analysis, it can be skipped with fast TRs

- Absolutely to skip when computing ALFF-derived measures

Final recommendations

- Reduce as much as you can your interpolation steps (best: merge from slice interpolation to normalisation)

- Skip non required steps

- Pay attention not to reintroduce noises previously corrected (e.g: if you detrend, compute your noise regressors on detrended data)

- If you regress noise, try to regress it in one step

- If you have a task, add the convolved regressors (and any other of your interest) to the noise matrix

Final recommendations

AND ALWAYS LOOK AT YOUR DATA

(and think about them)