Network modelling analysis

- Resting state data characteristics
- Preprocessing
- Network modelling analysis
- Methods comparisons and considerations
Data characteristics
Replicable networks

Large-scale inherent organisation is reproducibly found across studies and approaches

Grey matter networks

Resting state network structure is localised in grey matter
Resting state networks are similar to task activation patterns at group and single subject level.

Functional vs structural connectivity

Functional connectivity is related to structural connectivity

Honey et al (2009), Damoiseaux & Greicius (2009)
Low frequency fluctuations?

power spectra for 5 RSNs in low-TR data (mean of all 5 in black)

Cordes et al (2001)
Low frequency fluctuations?

- BOLD decreases as $1/f$
- Degrees of freedom increase as $\sqrt{f}$
Low frequency fluctuations?

- BOLD decreases as $1/f$
- Degrees of freedom increase as $\sqrt{f}$
- Combined effect contributes to RSN estimation across frequency range!
Electrophysiology of BOLD connectivity

Static versus dynamic connectivity

- Most connectivity measures are static (based on the full resting state scan)
- Dynamic connectivity is like to occur (changes over time)
- Static connectivity measures reflect average across dynamic states
- Dynamic connectivity measures are challenging (in terms of noise influences, significance testing)

Arousal

• Subjects fall asleep
• Changes in BOLD amplitude
• Related changes in correlation

Resting state big picture
Generic study blueprint

1. Data acquisition
2. Data preprocessing
3. Single-subject analysis
4. Group-level analysis
5. Statistical inference

ICA+dual regression
Connectome modeling

Seed-based correlation
Dynamic causal modeling
Why more than one tool?

“Brain representations”

Bijsterbosch et al (2020)
Why more than one tool?

“Brain representations”

Which tool to use?

What parts of the brain are interesting in your study?

What type of change do you expect (e.g., strength/shape/connection)?

How much power do you have?

Bijsterbosch et al (2020)
Preprocessing
Careful cleanup required

- Structured artefacts much more of a problem for rfMRI than task-fMRI
  - No model of expected activation
  - Instead based on correlating timeseries with each other

Van Dijk et al (2012)
Noise sources

- Head motion
- Cardiac & breathing cycles
- Scanner artefacts
Regressing out noise

- Head motion parameters
- White-matter / CSF
- Use GLM to remove nuisance timeseries
- Perform analysis on residuals
- “CompCor” method (PCA-based)

\[ X = \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 + \beta_7 + \beta_8 + \text{Residual “clean” data} \]

Data after standard preprocessing

Muschelli et al (2014)
Preprocessing overview

<table>
<thead>
<tr>
<th>Conventional preprocessing steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motion &amp; distortion correction</td>
</tr>
<tr>
<td>Slice timing correction</td>
</tr>
<tr>
<td>High pass temporal filtering</td>
</tr>
<tr>
<td>Spatial smoothing</td>
</tr>
<tr>
<td>Registration</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Noise reduction steps (use at least one of these)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuisance regression</td>
</tr>
<tr>
<td>Low pass temporal filtering</td>
</tr>
<tr>
<td>Volume censoring</td>
</tr>
<tr>
<td>Global signal regression</td>
</tr>
<tr>
<td>ICA-based clean-up</td>
</tr>
<tr>
<td>Physiological noise regression</td>
</tr>
</tbody>
</table>
Lowpass temporal filtering

- E.g., common to remove frequencies > 0.1Hz
- May remove useful signal
- Not guaranteed to remove much artefact
Global signal regression

- Regress out mean timeseries across all voxels (or all grey matter voxels)
- Shifts connectivity values to be zero mean
- Therefore, more negative correlations
- Not necessary if using partial correlation

Murphy et al (2009)
GSR effects & alternative

PCC correlation
spatial ICA cleanup

PCC correlation
spatial ICA cleanup + GSR

PCC correlation
spatial ICA cleanup + temporal ICA cleanup

Glasser et al (2018)
Clean-up comparison

- no additional correction
- 24RP-regression
- 24RP + volume censoring
- ICA-AROMA

FMRI data

DVARS
Clean-up comparison

Ciric et al (2017)
Preprocessing advice

- Read up on the latest literature
- Nuisance regression is not enough
- Low-pass filtering is not enough & often not necessary when using other approaches
- Use ICA-based methods and/or volume censoring
- Use physiological noise regression when interested in brainstem or other vulnerable brain regions
- Don’t use global signal regression
Data acquisition advice

• Just a guide, may vary depending on study aims!
• Whole brain coverage, voxelsize: 2 - 3 mm
• Scan duration:
  • 10-15 minutes per scan
  • Potentially multiple scans
• Repetition time: ideally close to 1 second (multiband/multiplexed imaging)
• Paradigm: eyes open, fixation cross
• Auxiliary data: physiology, sleep
Analysis method advice

• Don’t do the same thing that your lab always does without further consideration

• Do think about your study and hypotheses
  • Brain areas will inform spatial summary
  • Expected change will inform feature type

• Ok to test multiple dimensionalities (e.g., ICA) without looking at final statistical results

• Could be interesting to look at multiple brain representations, but only if it can be done robustly
Time for a break!
Network modelling analysis
Glossary

• Node = functional brain region
  • Contiguous nodes = interconnected ‘blobs’
  • Non-contiguous nodes = e.g. bilateral
• Parcellation = separation of all voxels into a set of nodes
  • Hard parcellation = binary regions
  • Soft parcellation = weighted regions
• Edge = connection between nodes
• Connectomics = mapping all connections between all brain regions
Analysis steps

1. Node definition
2. Timeseries extraction
3. Edge calculation
4. Network matrix
5. Group analysis
Node definition

Anatomical atlases

Functional atlases

Data-driven parcellation

Node definition

Anatomical atlases
- Harvard-Oxford/ AAL
- Avoid if possible because typically based on small number of subjects and not a good estimation of functional boundaries

Functional atlases
- Data-driven parcellation

## Node definition

<table>
<thead>
<tr>
<th>Anatomical atlases</th>
<th>Functional atlases</th>
<th>Data-driven parcellation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avoid if possible because typically based on small number of subjects and not a good estimation of functional boundaries</td>
<td>Many good functional atlases available, though few comparison studies</td>
<td></td>
</tr>
<tr>
<td></td>
<td>How to map onto individuals is very important</td>
<td></td>
</tr>
</tbody>
</table>
Node definition

**Anatomical atlases**
- Harvard-Oxford/ AAL
- Avoid if possible because typically based on small number of subjects and not a good estimation of functional boundaries

**Functional atlases**
- Yeo 2011/ Glasser 2016
- Many good functional atlases available, though few comparison studies
- How to map onto individuals is very important

**Data-driven parcellation**
- ICA/ Clustering/ Gradients
- Estimate parcellation from the same dataset used for further analyses
- How to map group parcellation onto individuals very important

ICA for parcellation
Timeseries extraction

Hard parcellation:

• Masking (mean timeseries)
• Eigen timeseries (PCA)
• Using multilayer classifier

ICA (soft parcellation):

• **Thresholded** dual regression/ back projection

Alternative:

• Hierarchical estimation of group & subject
• e.g. PROFUMO

Edge calculation

- Presence/absence of edges
- Strength of edges
- Directionality of edges
Direct versus indirect connections

- Correlation between 2 and 3 will exist
- Therefore full correlation will incorrectly estimate connection 2-3
- 2-3 is an indirect connection
Partial correlation

• Before correlating 2 and 3, first regress 1 out of both (“orthogonalise wrt 1”)
  • If 2 and 3 are still correlated, a direct connection exists
• More generally, first regress all other nodes’ timecourses out of the pair in question
  • Equivalent to the inverse covariance matrix
Regularisation

- Urgh! If you have 200 nodes and 100 timepoints, this is impossible!
- A problem of DoF - need large #timepoints - #nodes
- When inverting a “rank-deficient” matrix it is common to aid this with some mathematical conditioning, e.g. force it to be sparse (force low values that are poorly estimated to zero)
- Regularised partial correlation (such as ICOV, Ridge)
- But still important to maximise temporal degrees of freedom
Need to carefully define nodes

Berkson’s paradox = false positive (2-3)

Over-splitting = false negative (1-2)
Directionality of edges

• Directionality is hard to estimate in BOLD data

• Don’t use lag-based methods such as Granger causality

• Perhaps directionality is oversimplistic view of neural connectivity (particularly in resting-state)?

Smith et al (2011)
Network matrix
Hierarchical clustering
Partial correlation is sparser than full correlation matrix.
Group analysis

• Calculate network matrix for each subject

• Combine all network matrices into one

• Perform group-level comparisons:
  • Univariate tests for each edge (GLM)
  • Multivariate prediction methods (SVM)
FSLnets

• Currently uses Matlab or Octave

• Therefore this practical will be a bit different from other practicals

• More information and download here: https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLNets
Example: positive-negative mode

Picture vocabulary test
Fluid intelligence (number of correct responses)
Delay discounting (area under the curve for discounting of $200)
Years of education completed
Life satisfaction
List sorting working memory test
Oral reading recognition test
Sustained attention continuous performance test (true positives)
Sustained attention continuous performance test (specificity)
Delay discounting (area under the curve for discounting of $40,000)
Picture sequence memory test
Years since smoked last cigarette
Financial income (eight bands)
Peg-board dexterity test (time taken)
Visual acuity (ratio)
No history of psychiatric or neurologic disorders – father
Pattern comparison processing speed
Two-minute walking endurance test

Included in CCA
Variance explained:
4%
17%

Excluded
Age first smoked (smokers only)
Thought problems score (self-report)
Still smoking
Perceived stress score
Regional taste intensity score
Rule breaking behavior score (self-report)
Anger-physical aggression score
Times used any tobacco today
Pittsburgh sleep quality index (higher is worse)
Drug or alcohol problems – father
Total weekdays with any tobacco in last week
Sustained attention continuous performance test (false positives)
Positive test for THC (cannabis)
Fluid intelligence (number of skipped responses)

Example: connectivity fingerprint

Finn et al (2015)
Comparison of methods
Overview of resting state methods

**Voxel-based**
- Seed-based correlation analysis
- Independent component analysis
- Amplitude of low frequency fluctuations
- Regional homogeneity

**Node-based**
- Network modelling analysis
- Graph theory analysis
- Dynamic causal modelling
- Non-stationary methods
Seed-based correlation

- Easy to interpret
- No correspondence problem
- Seed-selection bias
- Only models seed-effect (ignoring complex structure & noise)
Seed-based correlation results are strongly influenced by small changes in seed location.
ICA

- Multivariate: decompose full dataset
- Test for shape & amplitude
- Can be hard to interpret
- No control over decomposition (may not get breakdown you want)
Graph theory

- Simple summary measures (derived from network matrix)

- Network matrix often binarised

- Difficult to meaningfully interpret (abstract and far removed from data)

Rubinov et al (2010)
Dynamic causal modelling

- Directional interpretation (effective connectivity)
- Biophysical model
- Assumes HRF homogeneity
- Limited model comparisons

Daunizeau et al (2011)
Overview of node-based methods

**Functional Connectivity**
- simpler, less meaningful
- network “discovery”, better conditioned
- can handle more nodes

**Effective Connectivity**
- more complex, more meaningful
- pre-specify (constrain) network model, harder to estimate
- can handle fewer nodes

**Graph Theory**
- clusters / hierarchies, network hubs
- network summary statistics (e.g. small-worldness, efficiency)

**Network Modelling from FMRI Data**
- full correlation
- partial correlation
- regularised partial correlation
- Bayes nets
- non-biological dynamic Bayes nets
- biophysical neural-groups to FMRI-signal forward model, fit to data using Bayes (e.g. DCM)

**Bottom-Up Neural Network Simulations**
- network of individual neurons simulated
- network of groups of neurons simulated (e.g. neural mass model)

**Closeness to (interaction with) real FMRI data**
- full correlation
- partial correlation
- regularised partial correlation
- Bayes nets
- non-biological dynamic Bayes nets
- biophysical neural-groups to FMRI-signal forward model, fit to data using Bayes (e.g. DCM)
Which method to choose?

- Relationship to RSNs
  - Dual regression
- Summary values
- Connections in system
  - Network modelling
- Biophysical system
  - DCM
- Connectomics
  - Network modelling
That’s all folks