

NEUROINFORMATICS AND LARGE-SCALE ANALYSIS

Lauri Nummenmaa, Turku PET Centre



Turku PET Centre Brain Imaging Course 2024

- sizes are compromised
- at minimum
- experimental design and measurements are otherwise sufficient
- replicability.

More is more

• Imaging is often financially prohibitive (>5000 \in / PET scan) thus sample

Potential harm to subject needs to be minimised —> sample sizes kept

Underpowered studies can result in experimental failure even when

Poor statistical power increases Type 1 & 2 error rates and lead to poor



Cremers et al (2017)



Basic problems

Data storage

- Where are my data?
- What if my lab members
 leave?
- How can i reaccess my data?

Data processing

- What are my postdocs
 doing all day long?
- Are you sure the files are good?

Data analysis

- Is everything done lege artis?
- How can we reproduce
 our results?

Data synthesis

- How can we combine
 data?
- What can we combine?

- Meta-analysis: Pooling standardised effect sizes to estimate population effect location and distribution
- For neuroscience, three main approaches
 - ROI level data and classic univariate meta-analysis
 - Coordinate-based data and volumetric meta-analysis
 - Combination of statistical maps from original studies

Solution 1: Meta-analysis





Classic mixed effects meta-analysis



The observed effect T₁ is sampled from a distribution with true effect θ_1 , and variance σ^2 . **This true effect** θ_1 , in turn, is sampled from a distribution with mean μ and variance τ^2 .

Classical meta-analysis

- Easy to perform even with limited statistical knowledge
- Most required functions available in R package metafor:
 - Effect size caluculation / conversion
 - Model estimation
 - Plotting ullet



Standardized Mean Difference

| 07. | -5.74] | |
|-----------|--------|-----|
| 28. | -5.041 | i |
| 08 | -4.95 | i |
| 24 | -4.30 | i i |
| 48 | -4 41 | i i |
| 10 | -4 64 | |
| 42' | 4.04 | |
| 92, 72 | 2 661 | |
| 13, | -3.00 | |
| 20, | -3.90 | |
| 41, | -3.92 | |
| 15, | -3.33 | |
| 42, | -3.29 | |
| 36, | -3.25 | |
| 28, | -3.30] | |
| 29, | -3.25] | |
| 83, | -3.56] | |
| 97, | -3.40] | |
| 68, | -3.64] | ĺ |
| 80. | -2.46 | İ. |
| 12 | -3.11 | i i |
| 15 | -3.00 | 1 |
| 08 | -2.85 | |
| 15 | -2 72 | ł |
| <u>60</u> | -2.72 | |
| 99, 24 | 2.00 | |
| 24, | -3.59 | |
| 03, | -2.57 | |
| 92, | -2.58 | |
| 66, | -2.35 | |
| 29, | -2.40] | |
| 43, | -3.09] | |
| 99, | -2.38] | |
| 79, | -2.23] | |
| 58. | -2.21] | ĺ |
| 77. | -1.97 | i |
| 34. | -2.28 | i i |
| 79 | -1.79 | 1 |
| 05 | -1.44 | |
| 65 | -1 70 | 1 |
| 60, | -2.36 | ł |
| 24 | -1 70 | |
| 42 42 | 2 23 | |
| 42, | 1 601 | |
| 40 | -1.00 | |
| 49, | -1.87 | |
| 76, | -1.56 | |
| 90, | -1.43 | |
| 91, | -1.18] | |
| 54, | -1.48] | |
| 13, | -1.30] | |
| 20, | -0.71] | |
| 04, | -0.80] | |
| 95, | -0.62 | ĺ |
| 08. | -0.19 | ĺ |
| | | |
| | | |

Approach 1: Regional analysis





Spies et al (2015)

Regional meta-analysis: Pros and cons

- Easy to analyse and interpret
- Data comparable in statistical terms
- No need to worry about normalization etc.
- Laborious
- Anatomical nomenclature not consistent
- Misses effects outside chosen ROIs

Approach 2: Peak-based analysis

Individual foci Permutations





Convergence of activation locations at given threshold

Activation Likelihood Estimation (Eickhoff et al 2015)

Thresholding



Right lateral

Positive emotions





Negative emotions











0.005

Nummenmaa, Putkinen & Sams (COiBS 2021)

ALE value

0.02

ALE Pros and cons

- Relatively easy to analyze and interpret
- Full-volume analysis
- No need to worry about normalization
- Effect direction can be accounted for •
- Effect sizes scaled only by sample size
- Requires coordinate-levels data
- Data modelled per peaks —> cluster size not taken into consideration

Approach 3: Automated data mining



Select highest pro

NeuroSynth (Yarkoni et al, 2011)

Meta-analysis

| | Semantic | 72 | | | | | | | | | | | | | | | | | | | | | | |
|-------------------|----------------|-------|------|------|----|-------|-------|---------|------|-----|-------|-------|-------|------|-----|-------|------|------|------|-------|--------|------|-------|------|
| | Encoding | 62 | 68 | | | | | | | | | | | | | | | | | | | | | |
| | Executive | 70 | 61 | 68 | | | | | | | | | | | | | | | | | | | | ĩ |
| The second second | Language | 62 | 68 | 70 | 73 | | | | | | | | | | | | | | | | | | | |
| | Verbal | 69 | 67 | 53 | 69 | 69 | | | | | | | | | | | | | | Cla | ssifi | ər | | |
| | Phonological | 68 | 72 | 79 | 64 | 62 | 76 | | | | | | | | | | | (| perc | enta | ige o | orre | ect) | |
| | Visual | 73 | 67 | 63 | 73 | 70 | 67 | 71 | | | | | | | | | | | | | | | | |
| | Interference | 65 | 62 | 52 | 66 | 57 | 74 | 55 | 66 | | | | | | | | | | | | | | | |
| pain activation) | Working memory | 72 | 68 | 54 | 71 | 60 | 75 | 68 | 56 | 70 | | | | | | | | | | | | | | |
| 200 | Conflict | 77 | 67 | 64 | 74 | 67 | 77 | 75 | 61 | 63 | 73 | | | | | | | | | | | | | |
| nce | Spatial | 77 | 68 | 63 | 76 | 67 | 76 | 67 | 62 | 64 | 71 | 69 | | | | | | | | | | | | |
| rking memory? | Attention | 74 | 65 | 64 | 74 | 69 | 73 | 69 | 56 | 65 | 67 | 54 | 69 | | | | | | | | | | | |
| Emotion? | Imagery | 69 | 65 | 61 | 68 | 64 | 77 | 61 | 62 | 56 | 72 | 53 | 54 | 67 | | | | | | | | | | |
| | Action | 75 | 70 | 71 | 77 | 69 | 73 | 65 | 64 | 68 | 72 | 63 | 57 | 54 | 71 | | | | | | | | | |
| Pain? | Sensory | 74 | 73 | 73 | 74 | 71 | 81 | 76 | 69 | 72 | 73 | 65 | 63 | 56 | 60 | 72 | | | | | | | | |
| ? | Perception | 73 | 69 | 74 | 70 | 72 | 75 | 70 | 68 | 73 | 79 | 69 | 67 | 67 | 65 | 59 | 72 | | | | | | | |
| | Auditory | 74 | 78 | 73 | 72 | 72 | 74 | 84 | 70 | 76 | 77 | 75 | 77 | 68 | 72 | 76 | 71 | 76 | | | | | | |
| | Pain | 85 | 81 | 86 | 84 | 83 | 89 | 83 | 78 | 84 | 78 | 80 | 81 | 80 | 83 | 76 | 77 | 80 | 82 | | | | | |
| | Reward | 79 | 74 | 76 | 84 | 75 | 88 | 78 | 76 | 79 | 80 | 76 | 74 | 76 | 83 | 76 | 74 | 82 | 80 | 77 | | | | |
| | Arousal | 75 | 67 | 76 | 79 | 74 | 82 | 74 | 68 | 79 | 76 | 75 | 74 | 74 | 76 | 73 | 74 | 77 | 74 | 65 | 73 | | | |
| Daint | Emotion | 75 | 70 | 81 | 81 | 78 | 83 | 78 | 76 | 80 | 83 | 78 | 80 | 78 | 80 | 76 | 77 | 83 | 83 | 70 | 54 | 76 | | |
| Pain | Social | 74 | 68 | 72 | 75 | 73 | 78 | 75 | 74 | 77 | 78 | 75 | 72 | 72 | 74 | 71 | 72 | 77 | 80 | 72 | 58 | 64 | 72 | |
| | Episodic | 65 | 61 | 76 | 76 | 75 | 84 | 71 | 75 | 76 | 75 | 72 | 74 | 77 | 75 | 78 | 75 | 79 | 88 | 75 | 73 | 70 | 62 | 72 |
| | Retrieval | 71 | 71 | 66 | 75 | 69 | 76 | 75 | 70 | 72 | 75 | 71 | 73 | 71 | 77 | 79 | 76 | 79 | 85 | 75 | 75 | 76 | 72 | 48 |
| | Recognition | 69 | 60 | 66 | 75 | 68 | 74 | 71 | 65 | 70 | 73 | 70 | 69 | 71 | 75 | 75 | 71 | 80 | 84 | 76 | 70 | 71 | 65 | 57 |
| obability | 4 | inc i | no . | Nº . | S | 60 | (a) | Jal. | , co | 100 | int | ila . | lor . | d is | 05 | or is | or. | of o | air | big . | sal il | 30 | 101 | Sil |
| | Serra | CUCOC | tecu | angu | Je | noloc | 2. 1. | allele. | mon | Cor | ંદર્શ | Atter | mag | PC | Ser | stcop | AUDI | X | Ren | PION | fuo. | So | EPIS' | 20th |
| | | | | | 8% | Э, | 10 | orkin | 2 | | | | | | 8 | 2 | | | | | | | | 9 |
| | | | | | | | 2 | - | | | | | | | | | | | | | | | | |





neurosynth.org

Neurosynth is a platform for large-scale, automated synthesis of functional magnetic resonance imaging (fMRI) data.

It takes thousands of published articles reporting the results of fMRI studies, chews on them for a bit, and then spits out images that look like this:





An automated meta-analysis of 516 studies of pain



Database Status

507891 activations reported in 14371 studies

Interactive, downloadable meta-analyses of 1334 terms

Functional connectivity and coactivation maps for over 150,000 brain locations

Keyword-based search

 Precompiled meta-analytic maps based on automatically parsed data

Result maps can be downloaded as 3D nifti files for further analysis

Custom analyses and queries possible

Memory



Attention



Face

Motor





A) Distribution of µ-opioid receptors in the brain as measured with [11C]carfentanil PET



B) Overlap between human emotion circuit and the µ-opioid receptor system





*r*_{all} = 0.38 $r_{\text{pleasure}} = 0.44$

Nummenmaa & Tuominen (2018 Br J Pharmac); Kantonen et al (2020 Neurolmage)



Neurosynth: Pros and cons

- Very easy to analyze and interpret
- Data readily available, allows custom analyses
- Full-volume analysis
- If large n, correrlates well with ALE on manual extracted peaks
- Quality contingent on the parser & reporting in studies
- Currently distinguishing activation / condition direction difficult

Solution 2: Large-scale synthesis of old datasets

- than raw data) lower the power of meta-analysis
- can be reanalysed (in Finland with "Rekisteritutkimus")
- needs to be extracted

Between-study variability and reliance on statistical estimates (rather

Existing data are often available and cheap to use given permissions

Data however have to be extracted, reprocessed and the metadata

Integrated approach at PET Centre

Preprocessed BRAIN data

Hospital **PACS**



Automated and supervised quality control









AIVO database



EXTERNAL register sources

SOMATIC WELL BEING

- 1. ICD codes for diagnoses
- 2. Laboratory results
- 3. Frequency of hospitalisation & sickness leaves

PSYCHOLOGICAL FUCNTIONING

- 1. Psychopathology (ICD codes)
- 2. Personality structure
- 3. Disorders of cognition (ICD codes)
- 4. Fluid intelligence and school achievement

SOCIOECONOMIC WELL BEING

- 1. Earning and income transfers
- 2. Education and social status
- 3. Labour market attachment
- 4. Developmental socioeconomic stressors

SOCIAL ATTACHMENT BEHAVIOUR

- 1. Marriage and cohabitation
- 2. Family establishment
- 3. Reproduction and family size





Karjalainen et al (2020)



Malen et al (2022 Neuroimage)

A public repository of unthresholded statistical maps, parcellations, and atlases of the brain.

What is it?

A place where researchers can publicly store and share unthresholded statistical maps, parcellations, and atlases produced by MRI and PET studies.

Why use it?

- A permanent URL
- Publicly shareable

Recently added collections of images from published papers

Name

Genetic, cellular, and connectomic characterization of the brai Principal component analysis reveals multiple consistent respo Attention- versus significance-driven memory formation: Taxo A dynamic gradient architecture generates brain activity state **Towards the Interpretability of Deep Learning Models for Multi Ageing Brain** Whole-brain functional correlates of memory formation in mes HIV infection is linked with reduced error-related default mode management abilities What Executive Function Network is that? An Image-Based Me Reinforcement learning with associative or discriminative gene Residential green space and air pollution are associated with b Log in Search

Search



Interactive visualization

Improves meta-analyses



Get started and upload an image!

| | Number of images |
|---|---------------------|
| in regions commonly plagued by glioma | 1 |
| onses to naturalistic stimuli in children and adults | 22 |
| onomy, neural substrates, and meta-analyses | 3 |
| es a la companya de l | 12 |
| i-modal Neuroimaging: Finding Structural Changes of the | 27 |
| sial temporal lobe epilepsy | 48 |
| e network suppression and poorer medication | 6 |
| eta-Analysis of Network Labels | 8 |
| eralization across states and actions: | 49 |
| prain activation in a social-stress paradigm | 17 |









Kantonen et al (2020 Neurolmage)

Lowered mu-opioid receptor availability in subclinical depression and anxiety





y = -2



Nummenmaa et al (2020 Neuropsychopharmacology)

Z - 1 Right Left







Sun et al (2021 J Neurosci)



Sun et al (2021 J Neurosci)

Common problems with data integration

- Variable imaging equipment
- Standarization of data ulletacquisition
- Metadata description



- Processing pipelines
- Comparability of conditions ullet
- Specificity of effects

- GE Advance
- HR+
- HRRT
- GE Discovery VCT PET/CT 0
- GE Discovery 690 PET/CT 0

Malen et al (submitted)

Solution 3: More is more in the first place

а



Human Connectome Project



UK Biobank



Physical and size Physical cardiac genera /le alcoho tobaccc assav and 3000 /sical ifesty. ensi fest É bone de

Cognitive





Comparison of the approaches

| | Specificity | Price | Computational demands |
|---------------------------------|--------------------|--------------------|--------------------------|
| Meta-analysis | Low-medium | Low | Low |
| Retrospective reuse | Medium | Medium-high | Moderate |
| Dedicated large- scale study | High | High-stratospheric | High |

