Second level analysis of fMRI

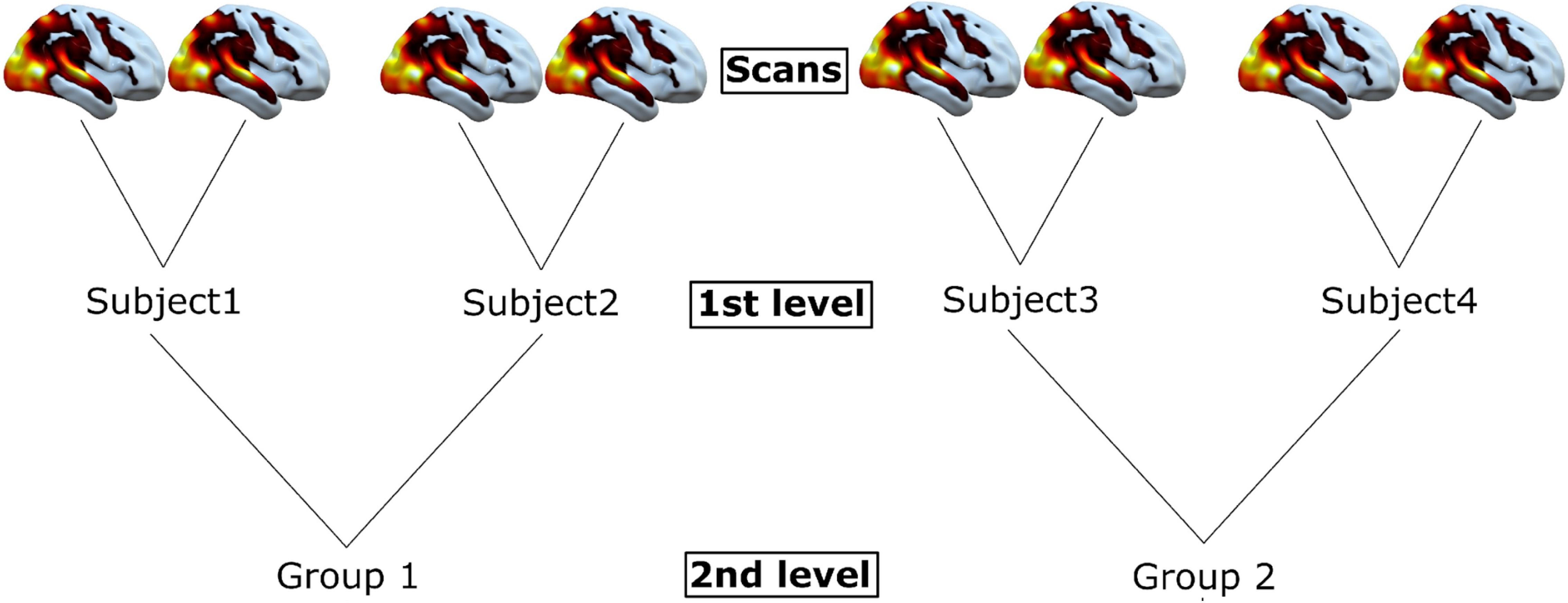
Severi Santavirta

Turku PET Centre

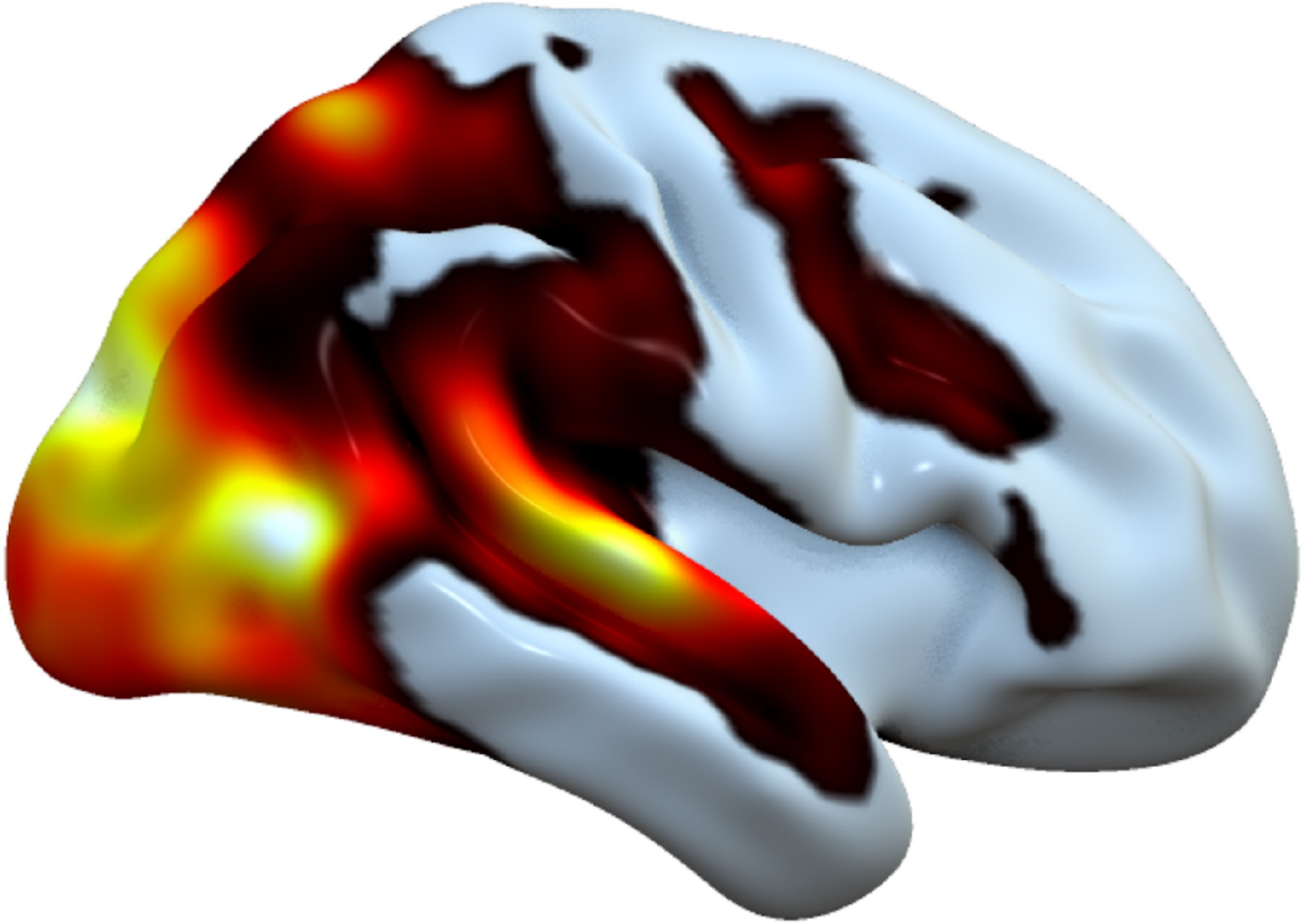
Topics

1. **Introduction to the second level analysis**
2. Theoretical framework of the group analysis
3. Second level models
4. Multiple comparisons problem
5. Non-parametric tests

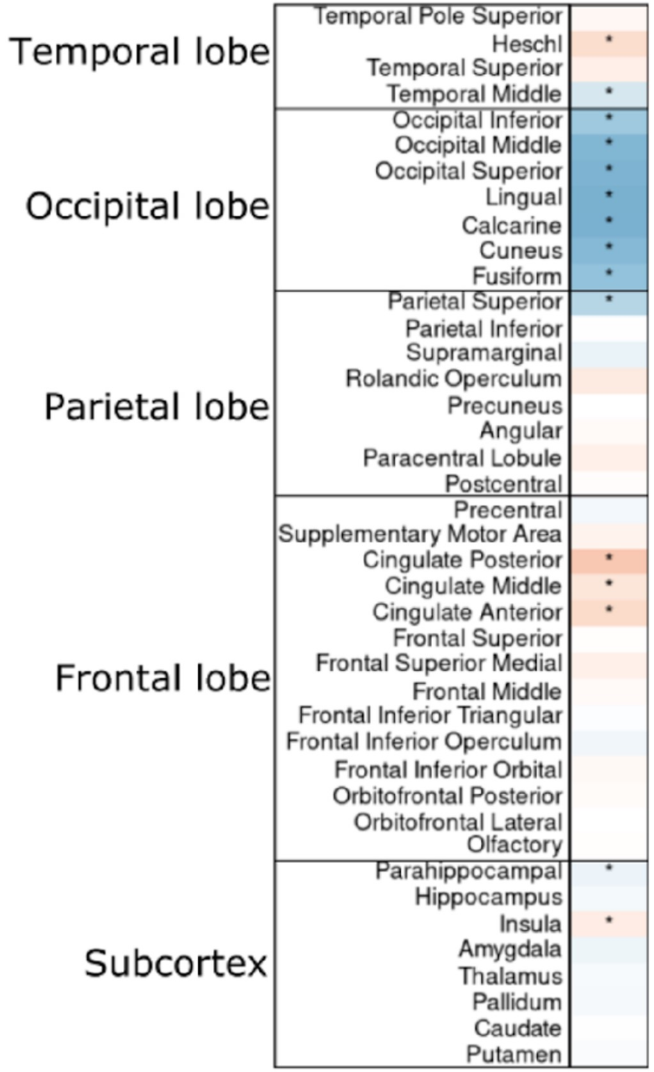
fMRI data are hierarchical



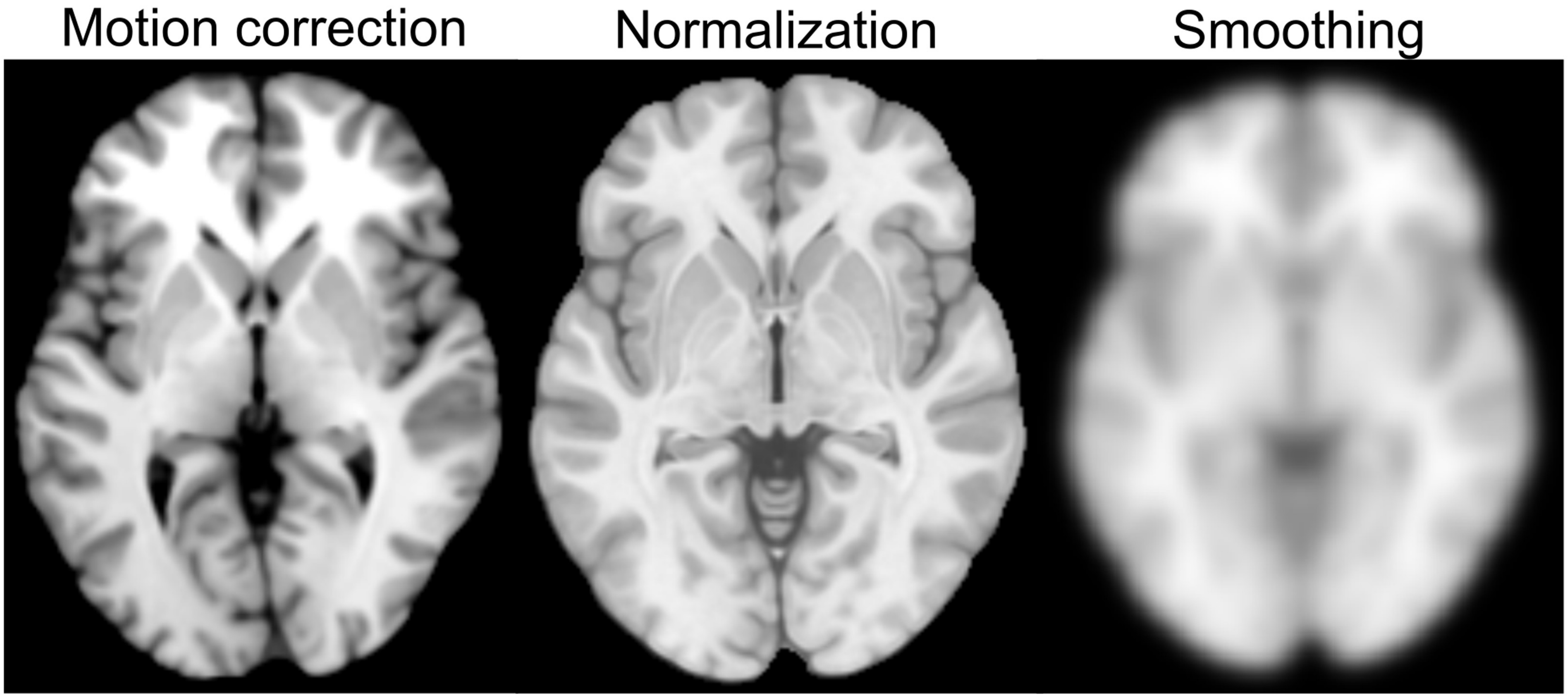
Whole brain analysis
Voxel-level analysis
Massive univariate analysis



Region-of-interest (ROI) analysis



Preprocessing



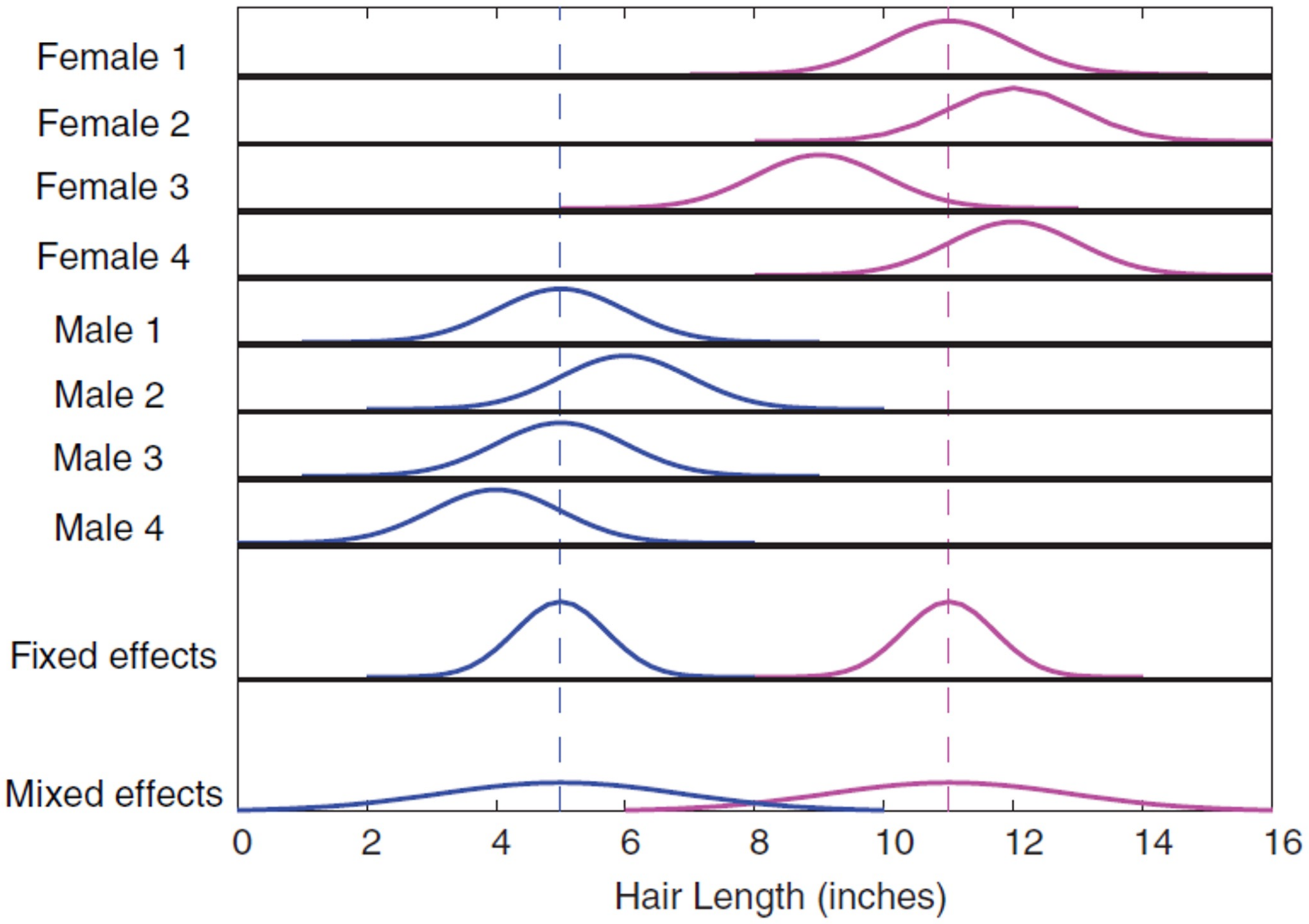
Topics

1. Introduction to the second level analysis
2. **Theoretical framework of the group analysis**
3. Second level models
4. Multiple comparisons problem
5. Non-parametric tests

Sources of variation in fMRI (or hair length)

Var^W = within-subject variance

Var^B = between-subject variance



(Poldrack, Nichols, & Mumford, 2011)

Fixed effects model		Mixed effects model
Var^W	Variance used in the analysis	$Var^W + Var^B$
Describe study sample only	Results	Generalize to population
Combine repeated measures within subjects	Application	Group analysis of fMRI

Mixed effects model mathematically

- Within subject variance estimation (1st level)

$$Y = \beta X + \varepsilon, \qquad \varepsilon \sim N(0, \sigma) \tag{1}$$

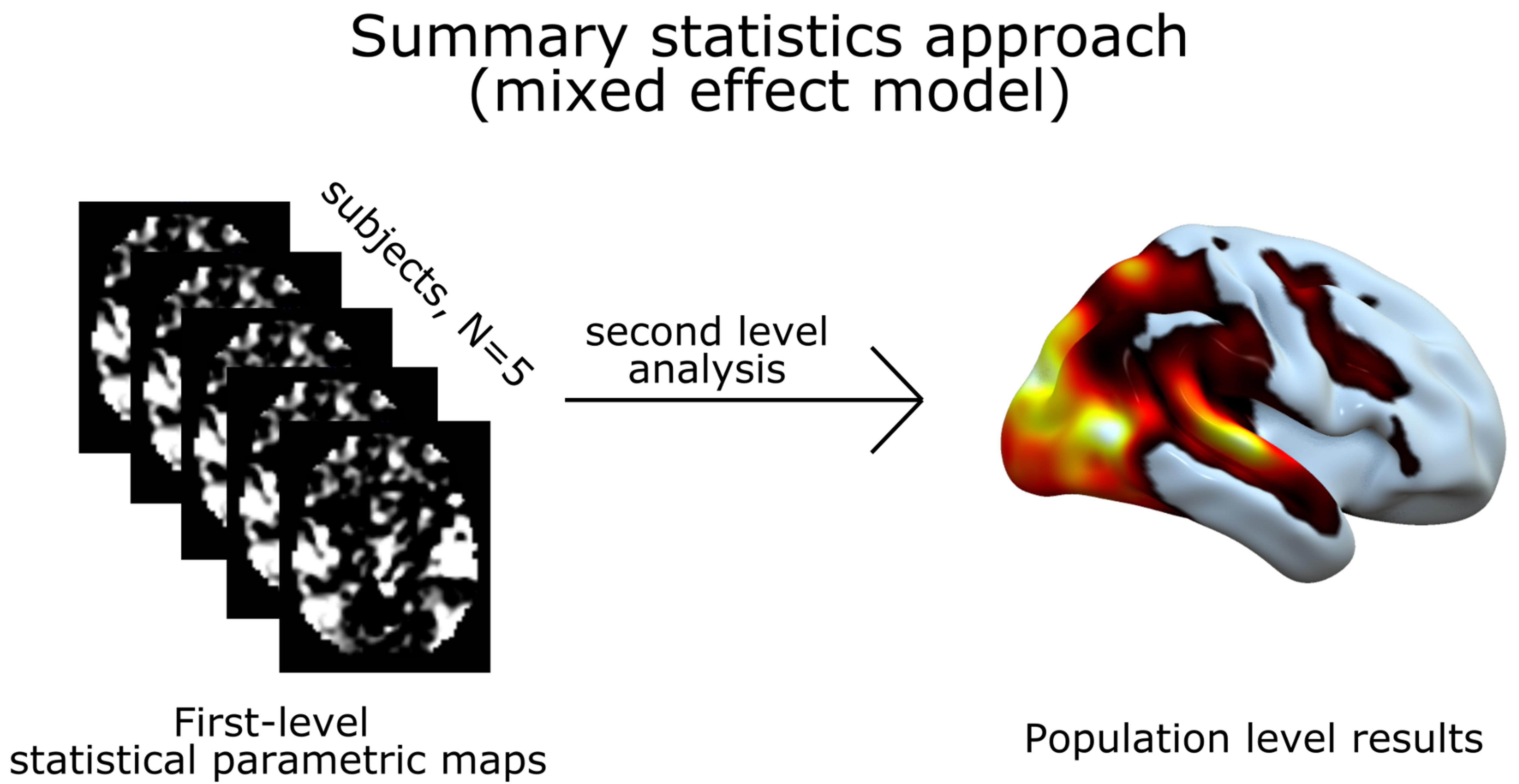
- Between subject variance estimation (2nd level)

$$\beta = \beta_b X_b + \varepsilon_b \qquad \varepsilon_b \sim N(0, \sigma_b) \tag{2}$$

- Full mixed effects model

$$Y = (\beta_b X_b + \varepsilon_b) X + \varepsilon \tag{3}$$

Computationally demanding to estimate!

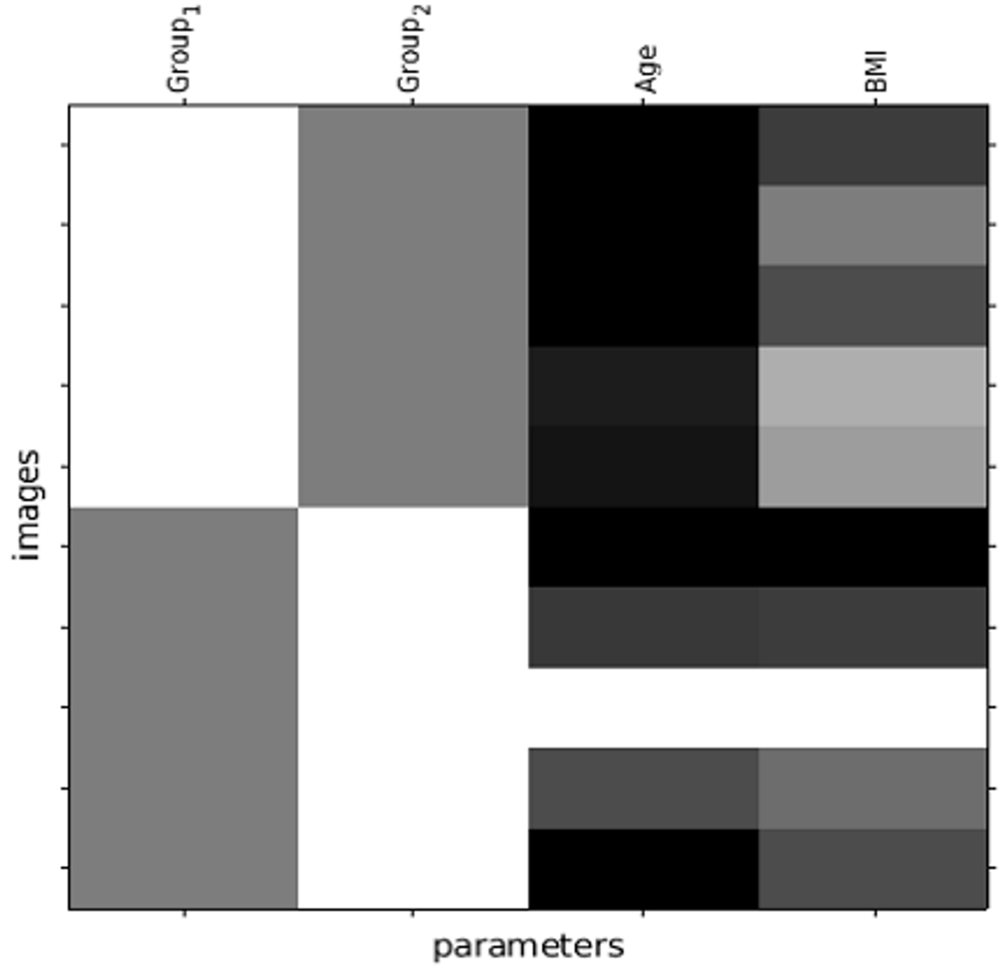

Summary statistics approach
(mixed effect model)
subjects, N=5
second level
analysis
First-level
statistical parametric maps
Population level results

Topics

- Introduction to the second level analysis
- Theoretical framework of the group analysis
- **Second level models**
- Multiple comparisons problem
- Non-parametric tests

Second level design matrix in SPM

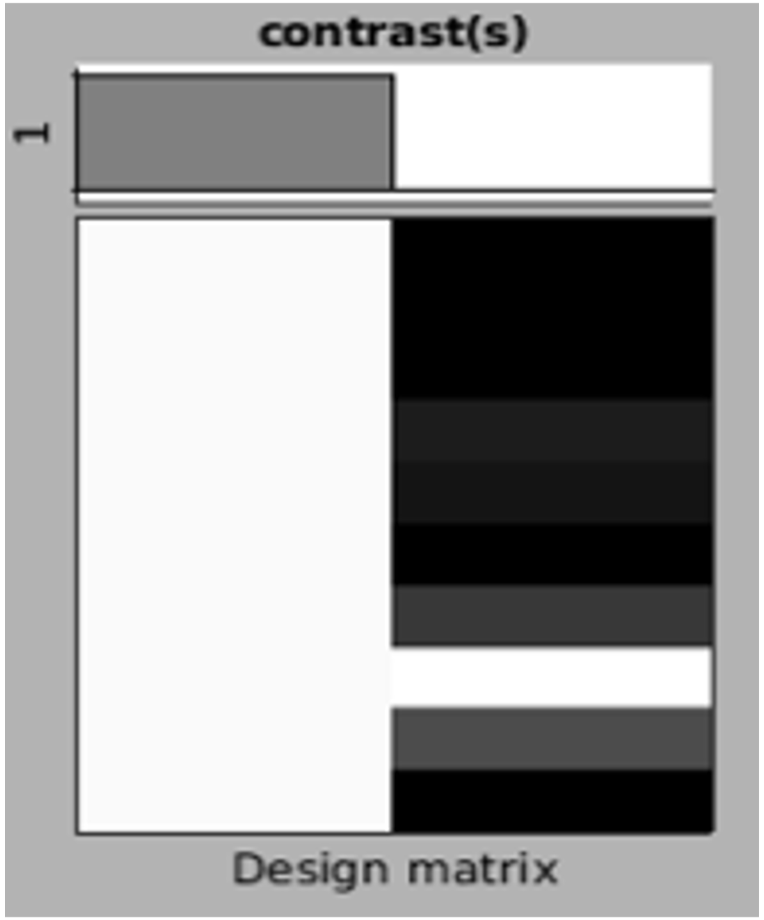
- Dependent variable:
  - β values from 1st level analysis
- Rows
  - subjects
- Columns:
  - Variables describing between subject differences
  - Effect of each column will be estimated



One sample T-test

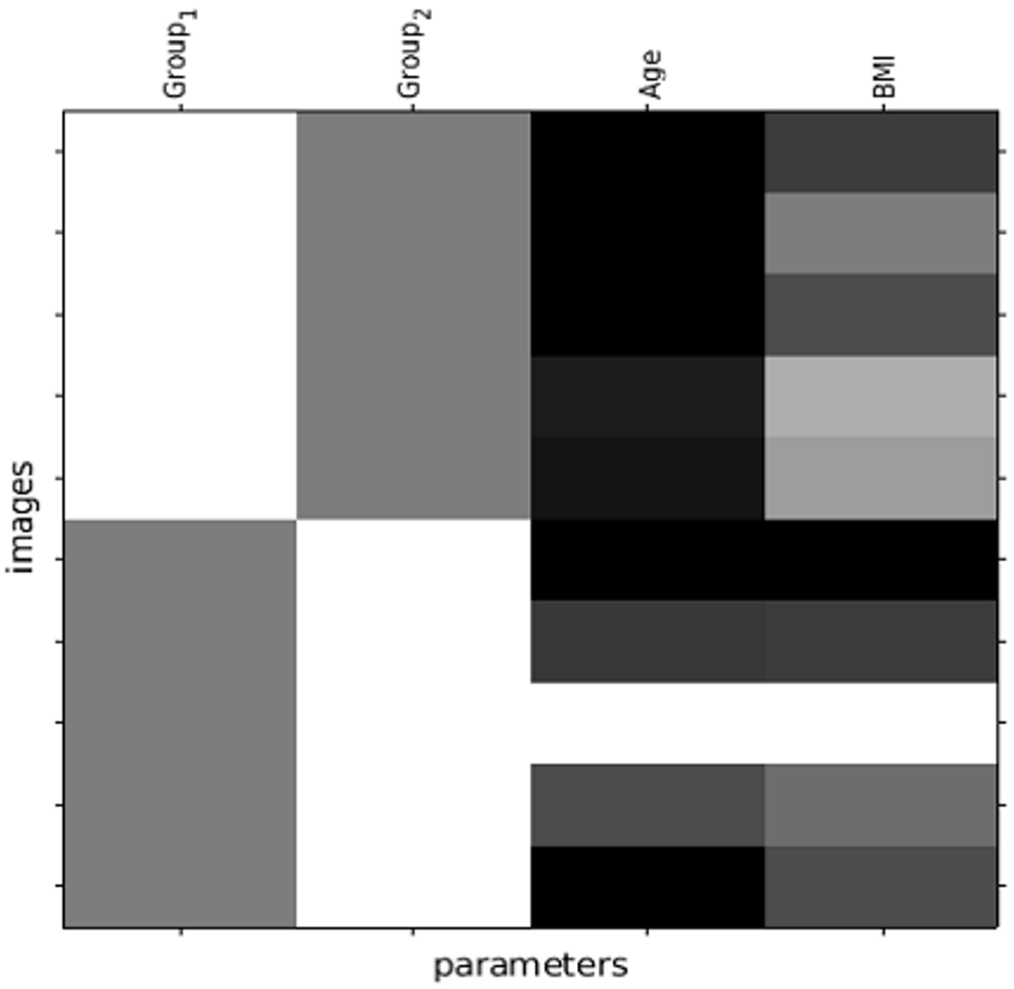
$\beta_{contrast} = \beta_{main} + \beta_{covariate} X_{covariate}$

- H: Brain activity is associated with the study condition somewhere in the brain
- Contrast (specified in the 1st level)
  - Main effect
  - Condition1 – Condition2
- Covariates of interest
  - Positive association [0 1] = $\beta_{covariate}$
  - Negative association [0 -1] = $-\beta_{covariate}$
- Main effect [1 0] = β_{main}



Two sample T-test

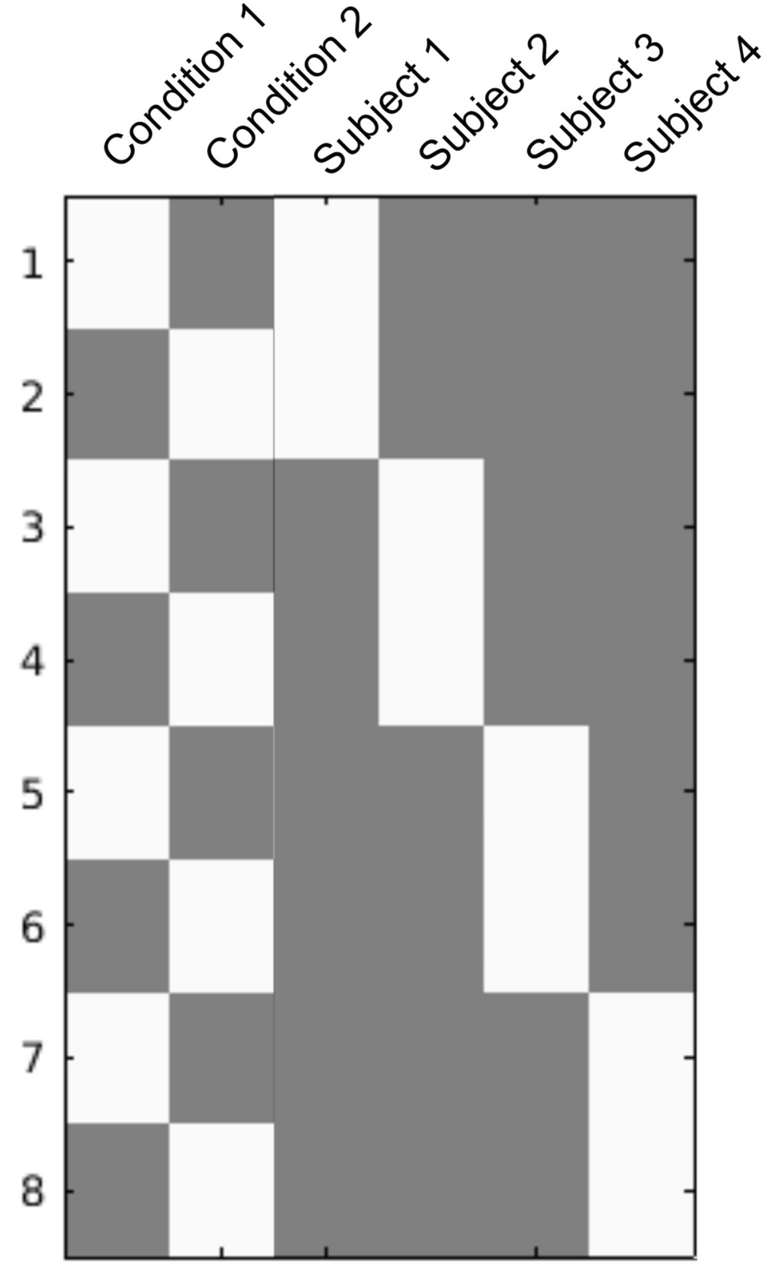
- H: Brain activity is different between two groups of subjects
- Group1 > Group2 with nuisance covariates
  - [1 -1 0 0] = β_{gr1} - β_{gr2}
  - "Whether females have increased brain response for the condition than males when age and BMI are controlled for"



$\beta_{contrast} = \beta_{gr1}X_{gr1} + \beta_{gr2}X_{gr2} + \beta_{age}X_{age} + \beta_{bmi}X_{bmi}$

Paired T-test

- H: There is a difference between conditions in the brain response
- 4 subjects, 2 scans per subject
- Condition
  - Cross-sectional
    - Baseline (1) vs. Stimulus (2)
  - Longitudinal
    - Baseline (1) vs. After treatment (2)
- Condition 2 > Condition 1
  - [-1 1 0 0 0 0 0] = $-\beta_{c1} + \beta_{c2}$
  - “Brain response associated with happy faces is higher after treatment"

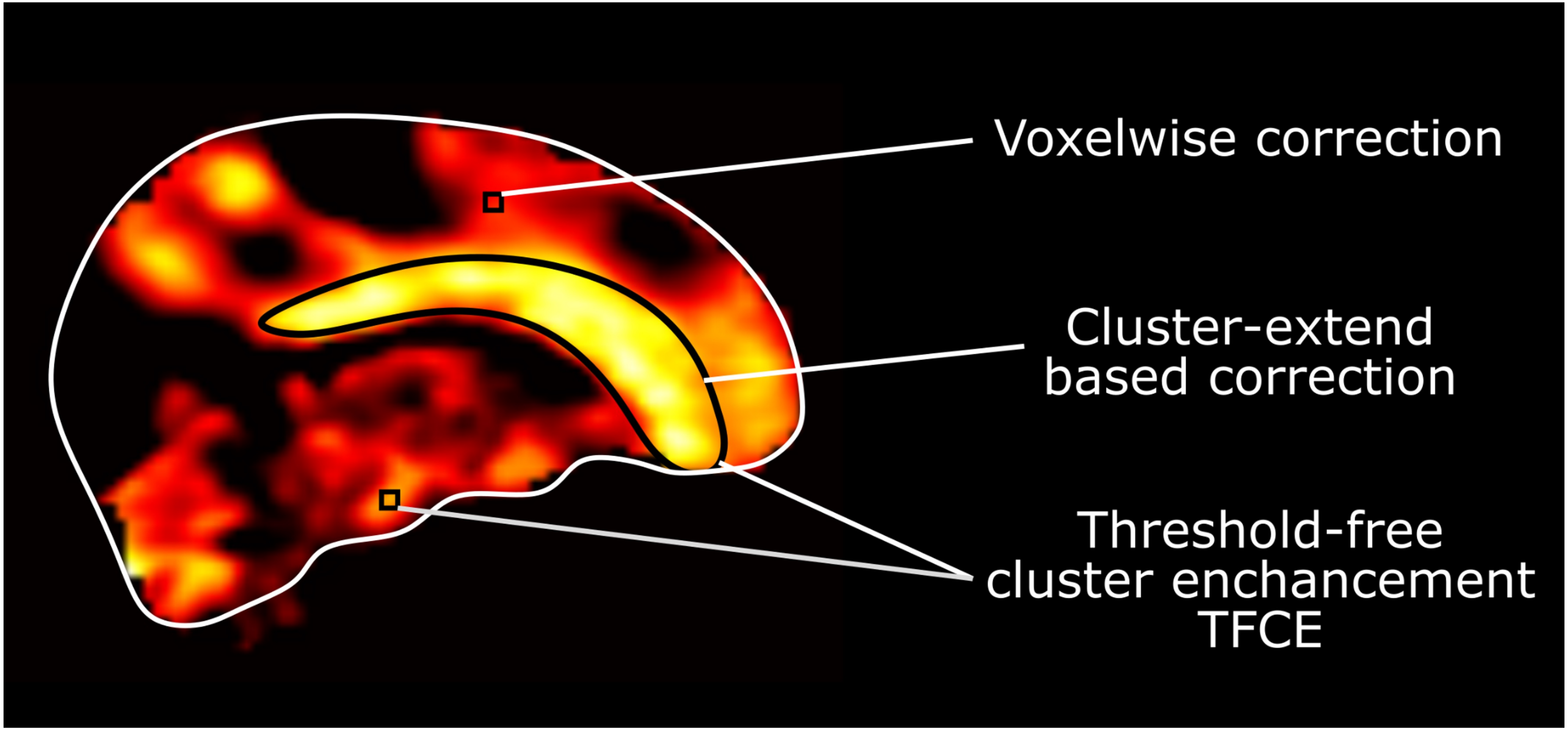


$\beta_{contrast} = \beta_{c1}X_{c1}+\beta_{c2}X_{c2}+\beta_{s1}X_{s1}+\beta_{s2}X_{s2}+\beta_{s2}X_{s2}+\beta_{s2}X_{s2}$

Topics

- Introduction to the second level analysis
- Theoretical framework of the group analysis
- Second level models
- **Multiple comparisons problem**
- Non-parametric tests

Multiple comparisons correction methods



(Review: Lindquist & Mejia, 2015)

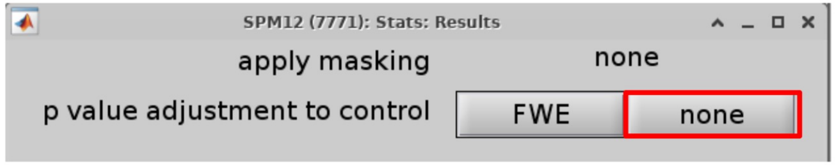
Voxelwise multiple comparisons correction

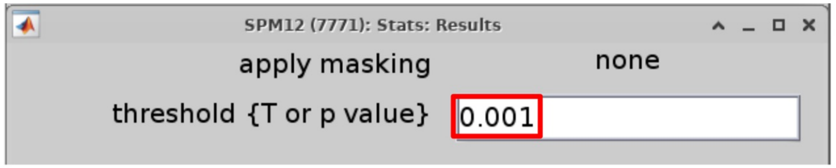
- Family-wise error rate (FWER) (Lindquist & Mejia, 2015)
    - Probability of making one or more false positives
    - Bonferroni correction
        - "There is a 5% probability of making **at least one** false positive finding"
        - 0.05 / number of tests = corrected p-value threshold
- False discovery rate (Benjamini & Hochberg, 1995)
    - "**No more than 5%** of our findings are false positives"

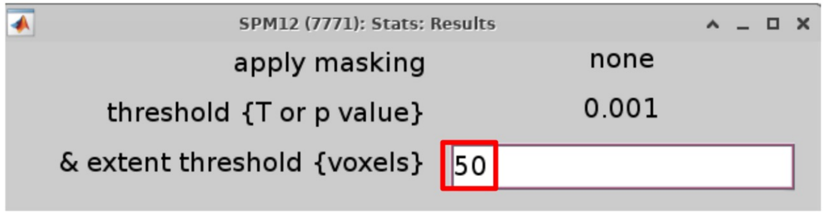
Cluster-extend based correction (Lindquist & Mejia, 2015)

- Accounts for the spatial dependency between voxels
- "What is the probability to observe a activating cluster of this size under the null hypothesis of no activation"
- Three-step procedure
  1. Choose primary voxel-level threshold
     e.g. $p < 0.001$
  2. Choose minimum size of the cluster
     e.g. 50
  3. Control for FWER on a cluster level
- The approach may be problematic

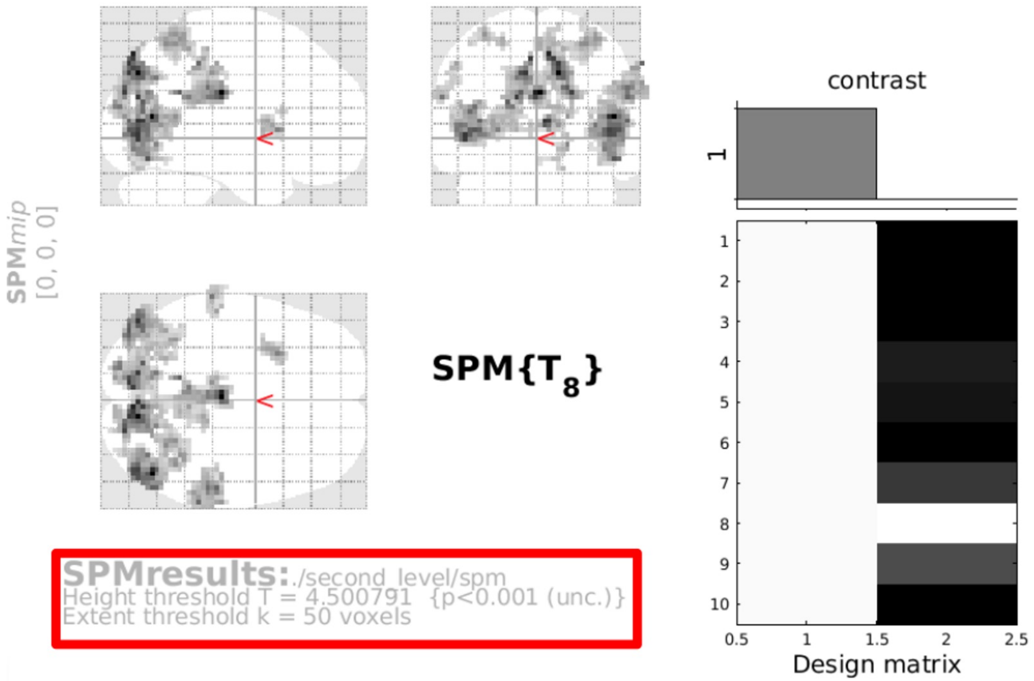
1


2


3


4


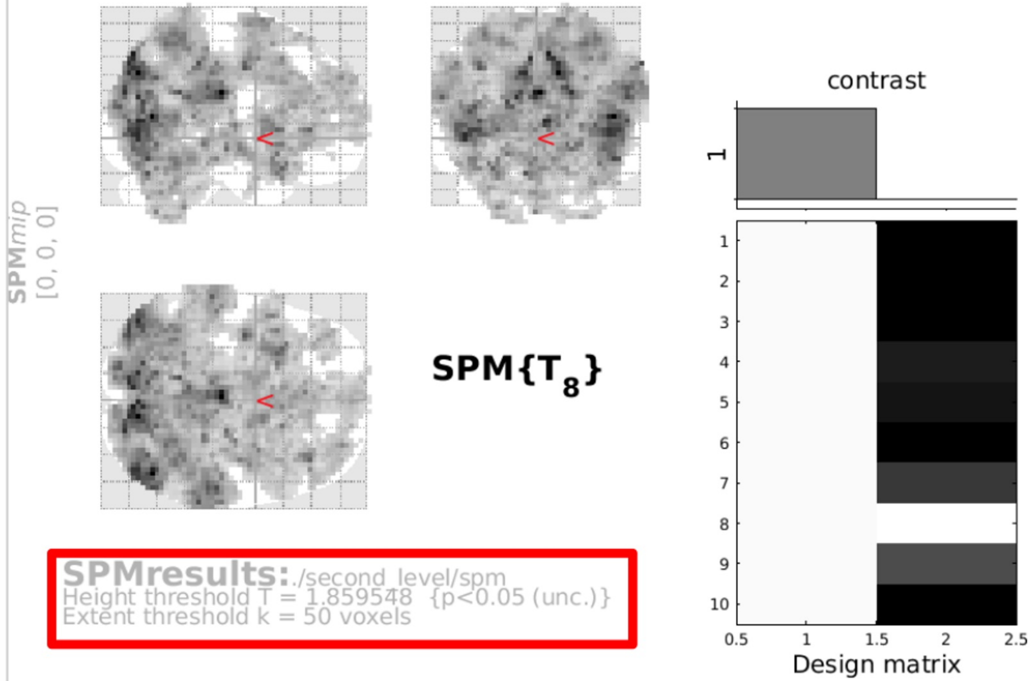
Main effect



Statistics: ***p-values adjusted for search volume***

set-level p	set-level c	cluster-level $p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	peak-level $p_{FWE-corr}$	$q_{FDR-corr}$	T	(Z_E)	p_{uncorr}	mm	mm	mm
0.000	**10**	**0.000**	**0.000**	**453**	**0.000**	**0.014**	**0.507**	**14.69**	**5.04**	**0.000**	**48**	**-69**	**12**
						0.099	0.507	11.33	4.65	0.000	45	-78	3
						0.188	0.581	10.40	4.52	0.000	48	-72	-15
		0.000	**0.000**	**166**	**0.000**	**0.016**	**0.507**	**14.39**	**5.01**	**0.000**	**0**	**-24**	**27**
						0.242	0.581	10.06	4.46	0.000	-12	-51	27
						0.593	0.581	8.91	4.27	0.000	0	-33	24
		0.000	**0.000**	**1011**	**0.000**	**0.019**	**0.507**	**14.10**	**4.98**	**0.000**	**12**	**-78**	**39**
						0.040	0.507	12.77	4.83	0.000	6	-81	51
						0.061	0.507	12.06	4.75	0.000	3	-78	9
		0.000	**0.000**	**202**	**0.000**	**0.296**	**0.581**	**9.79**	**4.42**	**0.000**	**66**	**-39**	**45**
						0.296	0.581	9.79	4.42	0.000	57	-27	24
						0.425	0.581	9.32	4.34	0.000	57	-33	30
		0.000	**0.000**	**57**	**0.000**	**0.536**	**0.581**	**9.03**	**4.29**	**0.000**	**-63**	**-30**	**24**
						0.999	0.606	7.34	3.94	0.000	-54	-27	45
						1.000	0.824	5.45	3.43	0.000	-57	-33	33
		0.000	**0.000**	**68**	**0.000**	**0.764**	**0.581**	**8.61**	**4.21**	**0.000**	**-27**	**12**	**3**
						1.000	0.641	6.70	3.79	0.000	-30	3	3
						1.000	0.719	6.05	3.61	0.000	-21	15	15
		0.000	**0.000**	**55**	**0.000**	**0.831**	**0.591**	**8.51**	**4.19**	**0.000**	**18**	**-66**	**-6**
						1.000	0.694	6.24	3.66	0.000	18	-60	0
						1.000	0.705	6.20	3.65	0.000	12	-72	-9
		0.000	**0.000**	**130**	**0.000**	**0.991**	**0.606**	**8.17**	**4.12**	**0.000**	**30**	**-60**	**57**

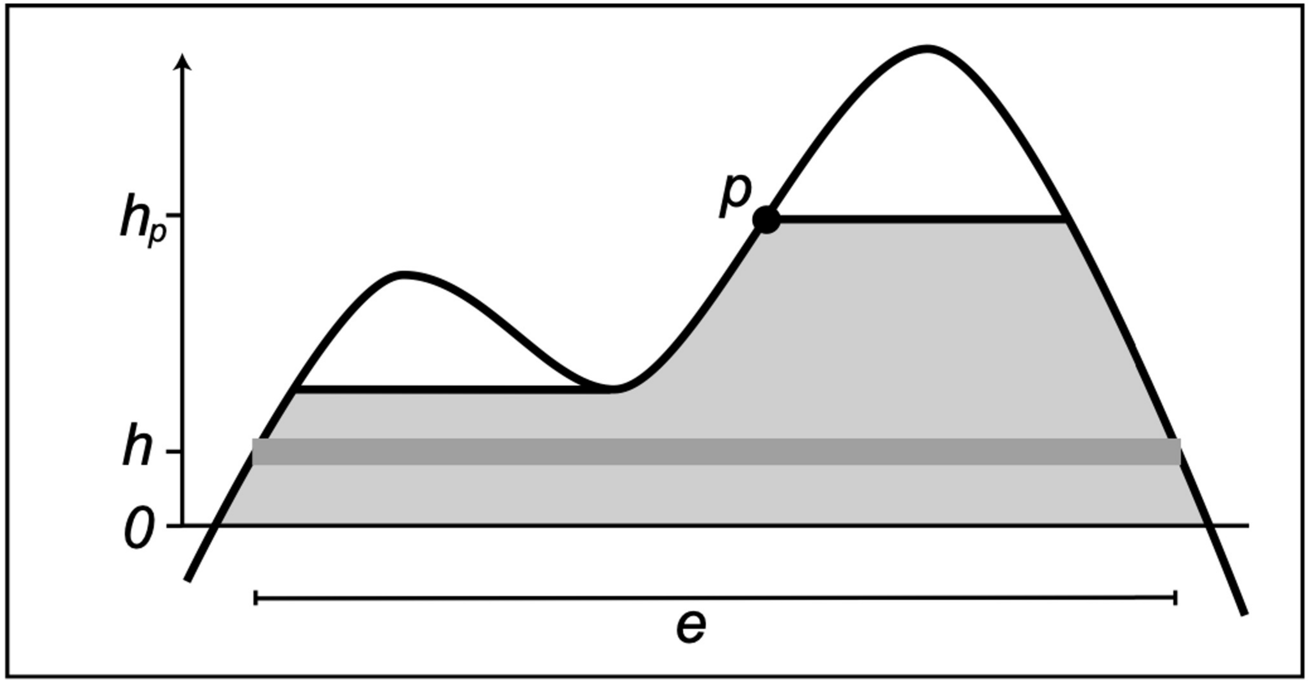
Main effect

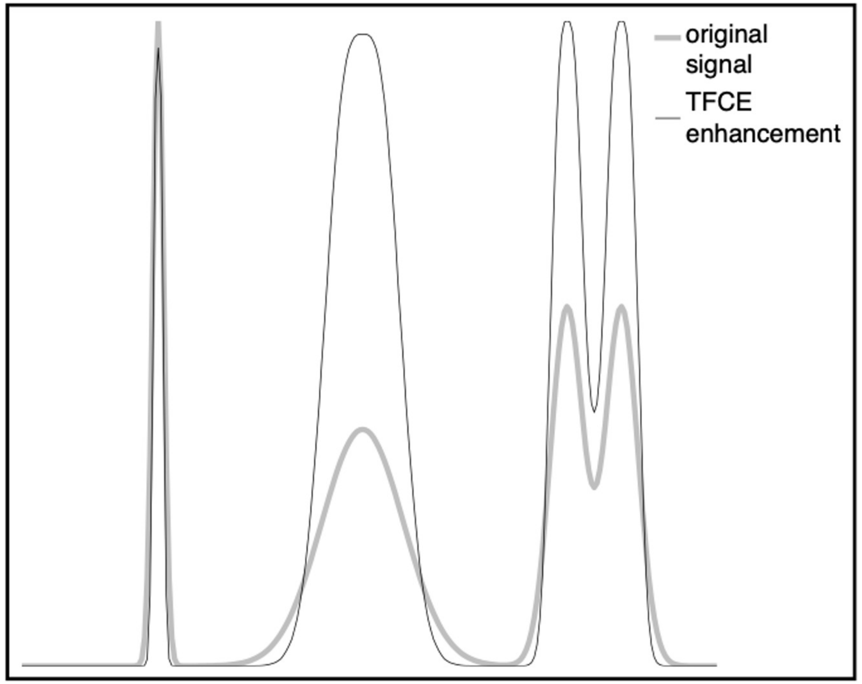


Statistics: ***p-values adjusted for search volume***

set-level p	set-level c	cluster-level $p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	peak-level $p_{FWE-corr}$	$q_{FDR-corr}$	T	(Z_E)	p_{uncorr}	mm	mm	mm
1.000	**3**	**0.000**	**0.000**	**18190**	**0.000**	**0.014**	**0.309**	**14.69**	**5.04**	**0.000**	**48**	**-69**	**12**
						0.016	0.309	14.39	5.01	0.000	0	-24	27
						0.019	0.309	14.10	4.98	0.000	12	-78	39
		0.003	**0.000**	**668**	**0.000**	**0.536**	**0.355**	**9.03**	**4.29**	**0.000**	**-63**	**-30**	**24**
						0.999	0.370	7.34	3.94	0.000	-54	-27	45
						1.000	0.503	5.45	3.43	0.000	-57	-33	33
		0.893	**0.253**	**163**	**0.013**	**1.000**	**0.490**	**5.60**	**3.48**	**0.000**	**-36**	**39**	**24**
						1.000	0.503	5.34	3.39	0.000	-39	45	33
						1.000	0.658	4.22	2.98	0.001	-24	42	33

Threshold-free cluster enchancement (TFCE)

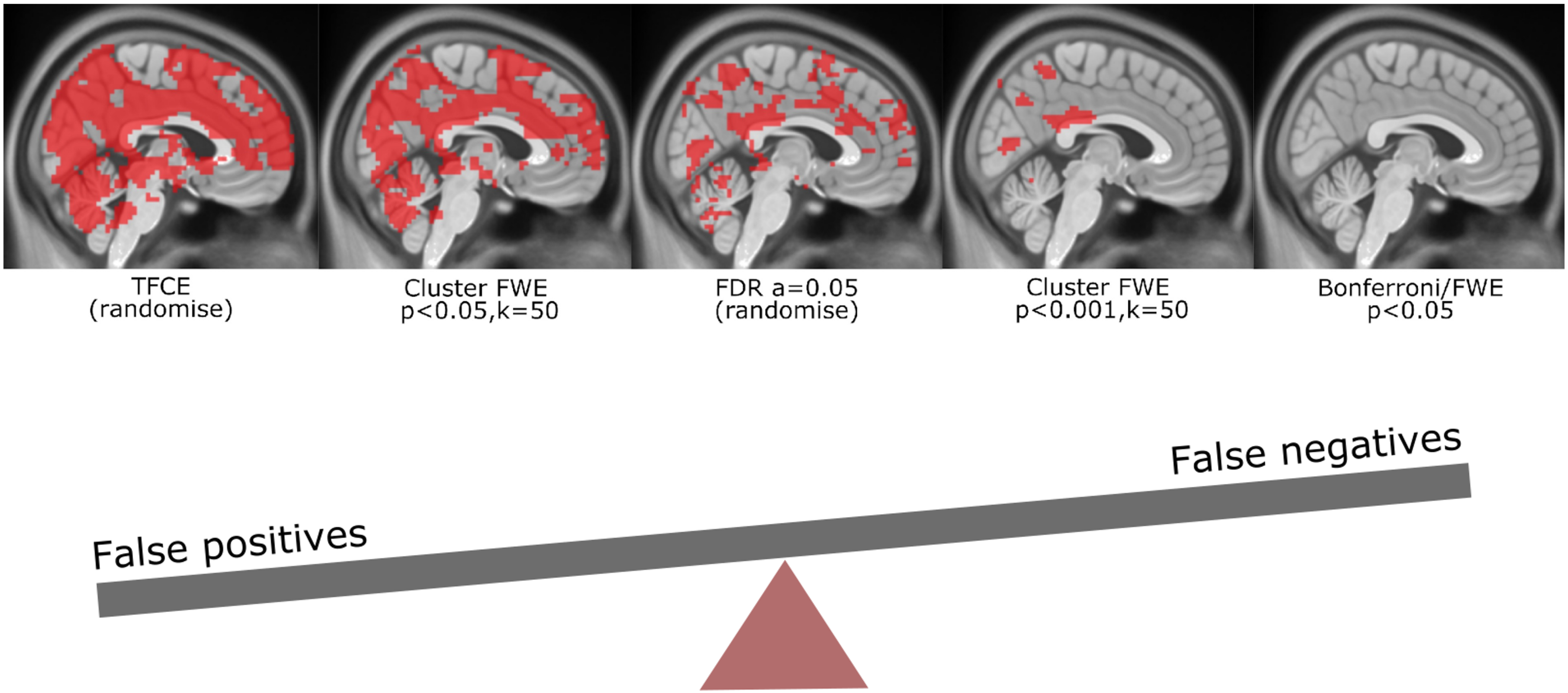




TFCE = h_p (voxelwise t-value) * e (amount of supporting voxels)

- The voxelwise significance is adjusted by the amount supporting voxels
- Significance of each voxel is assessed with permutations and then corrected for multiple comparisons

(Smith & Nichols, 2009)



Conservative correction inflates effect size estimates

Topics

- Introduction to the second level analysis
- Theoretical framework of the group analysis
- Second level models
- Multiple comparisons problem
- **Non-parametric tests**

Non-parametric tests

- Motivation for non-parametric tests in group analyses (Eklund, Nichols, & Knutsson, 2016)
  1. Parametric tests may produce false positive findings in group level analyses after multiple comparisons corrections
  2. Voxelwise multiple comparisons methods may produce too conservative findings and cluster-based methods false positives
  3. Non-parametric tests have been shown to correct better for multiple comparisons.
- Tools for non-parametric tests
  - SnPM (Doc: https://warwick.ac.uk/fac/sci/statistics/staff/academic-research/nichols/software/snpm)
  - FSL Randomise (Winkler, Ridgway, Webster, Smith & Nichols, 2014)
    - One and two sample (unpaired/paired) T-tests, repeated measures anova
    - Easy to output statistical result maps with various different multiple comparisons methods
      - Includes TFCE method
    - Doc: https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Randomise

References

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Statistical Methodology, 57*(1), 289-300. doi:DOI 10.1111/j.2517-6161.1995.tb02031.x

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci U S A, 113*(28), 7900-7905. doi:10.1073/pnas.1602413113

Lindquist, M. A., & Mejia, A. (2015). Zen and the art of multiple comparisons. *Psychosom Med, 77*(2), 114-125. doi:10.1097/PSY.0000000000000148

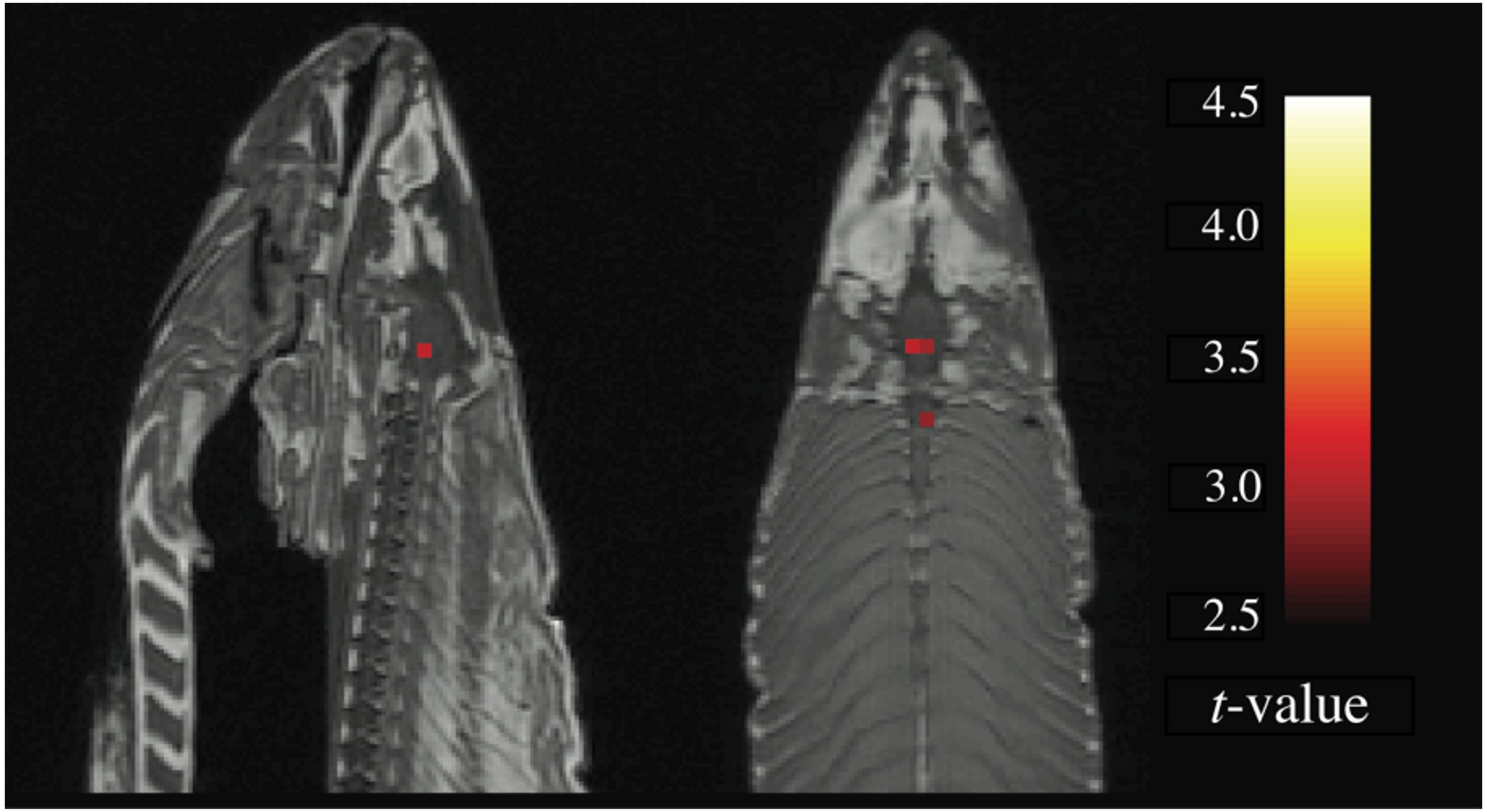
Poldrack, R. A., Nichols, T., & Mumford, J. (2011). Handbook of Functional MRI Data Analysis. *Cambridge: Cambridge University Press*. doi:10.1017/cbo9780511895029

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage, 44*(1), 83-98. doi:10.1016/j.neuroimage.2008.03.061

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *Neuroimage, 92*, 381-397. doi:10.1016/j.neuroimage.2014.01.060

Bennett, B. Miller, & Wolford.(2011). Neural Correlates of Interspecies Perspective Taking in the Post-Mortem Atlantic Salmon: An Argument For Proper Multiple Comparisons Correction. Journal of Serendipitous and Unexpected Results, 1(1), 1-5.

Thank you!



(Bennet, 2011)