Post-hoc Inference

Modern Inference

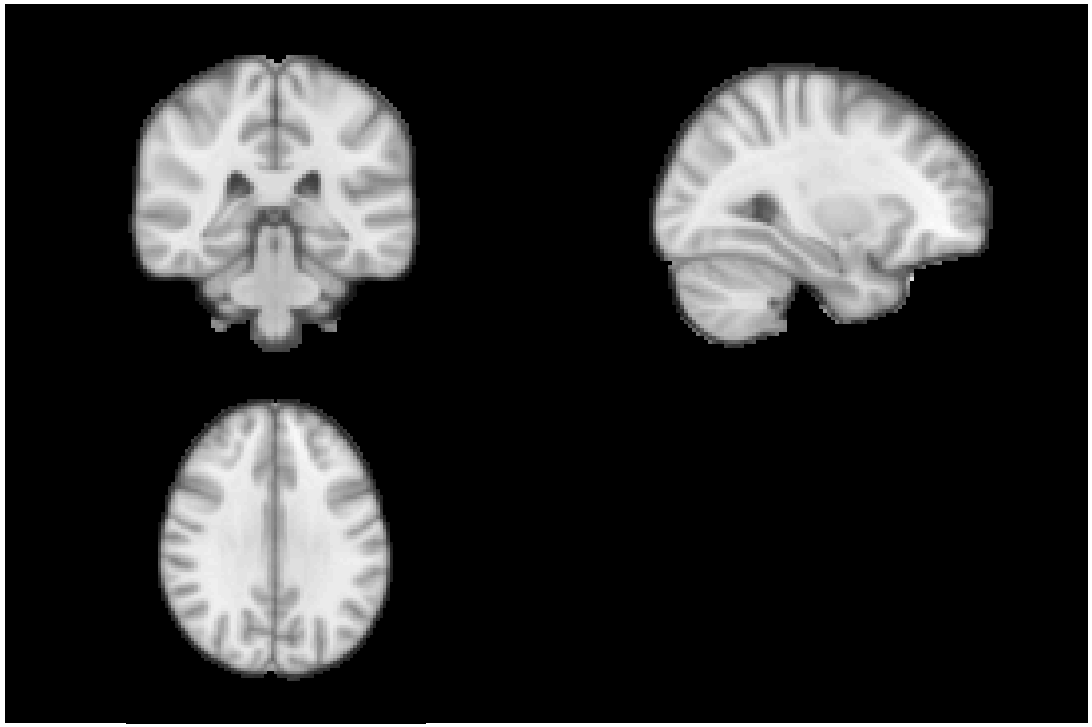
Aldo Solari

Go/No-go data

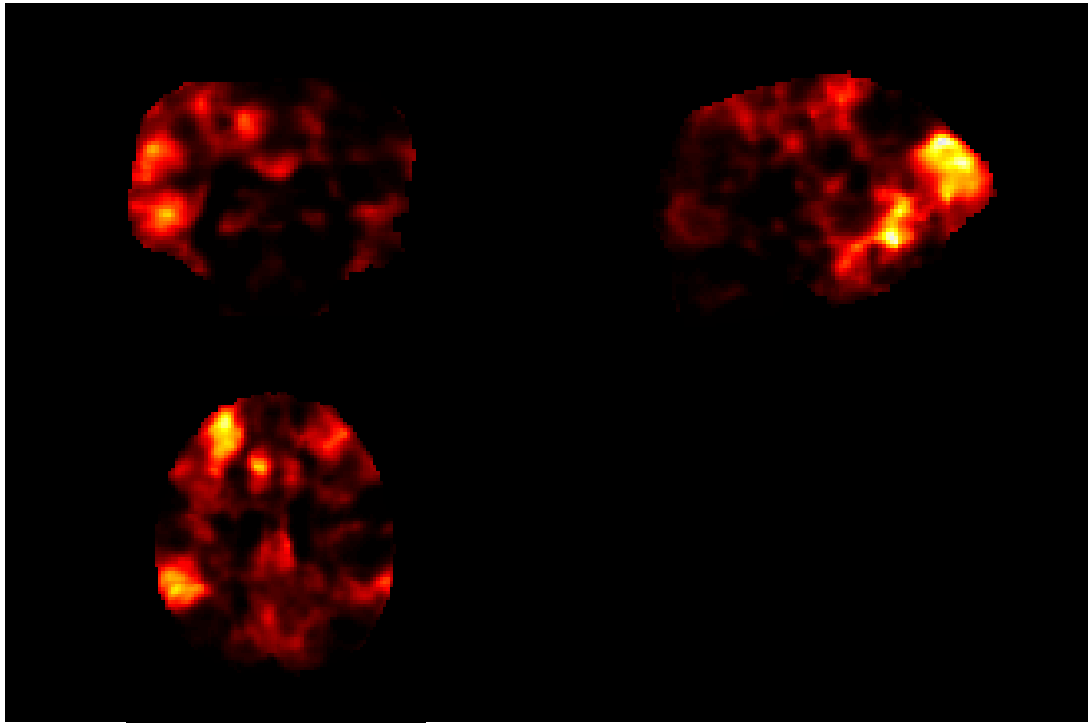
Lee et al. (2018)

$n = 34$ subjects

$m = 225212$ variables (voxels)



Brain



Statistical Parametric Map

Statistical Parametric Map

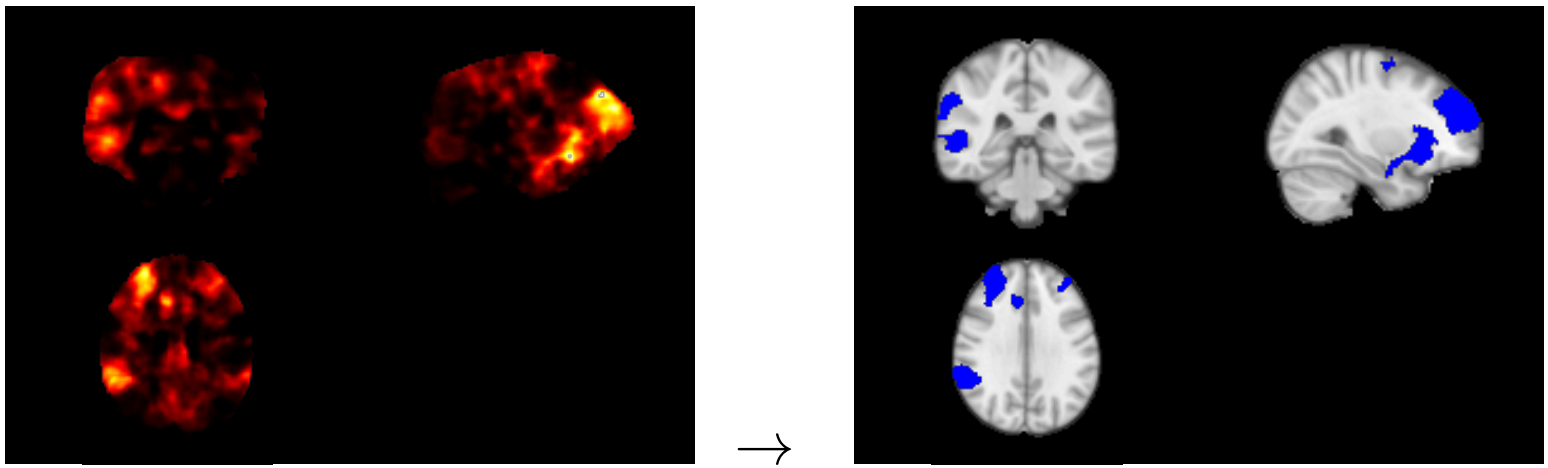
t-statistic/p-value testing 'no activation' at each location (voxel)

Aim

Find brain *regions* of activations

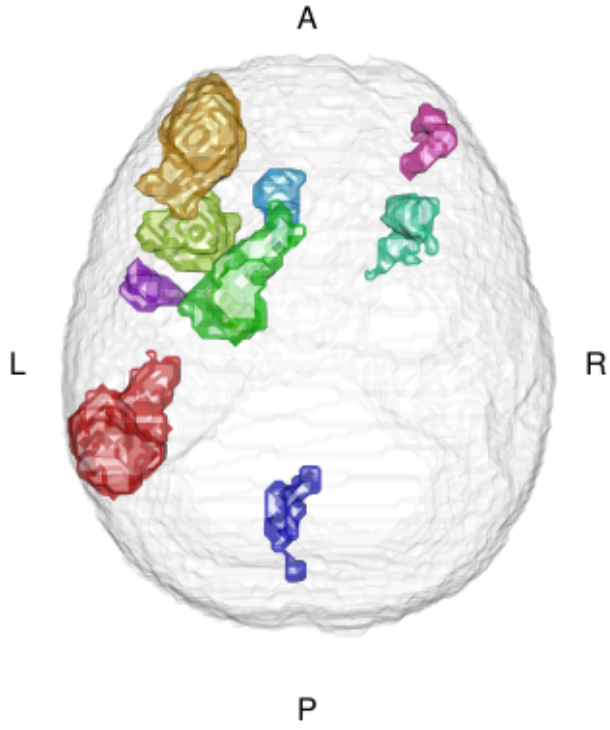
Aggregation

Micro-inferences (voxels) can be aggregated to larger-scale inferences (regions)



Data-driven selection

Clusters



Post-selection inference

Problem

How to assess the significance of selected clusters?

Clusters are both (i) selected and (ii) tested with the same data

Solution

Correcting overoptimism in inference due to data-driven selection

Assume a random sample of size n from m-variate Gaussian

$$N_m(\mu, \Sigma)$$

with $\mu = (\mu_1, \ldots, \mu_m)^\mathsf{T}$ and $\mathrm{diag}(\Sigma) = (\sigma_1^2, \ldots, \sigma_m^2)^\mathsf{T}$

Voxel null hypotheses

$$H_i : \mu_i = 0 \qquad i = 1, \ldots, m$$

Test statistic

$$T_i = \sqrt{n}\,\frac{\hat{\mu}_i}{\hat{\sigma}_i} \sim t_{n-1, \sqrt{n}\frac{\mu_i}{\sigma_i}} \qquad i = 1, \ldots, m$$

Notation

Brain

$B = \{1, \ldots, m\}$ collection of m voxels

Effects

$T = \{i \in B : \mu_i = 0\}$ voxels with no effect
$F = \{i \in B : \mu_i \neq 0\}$ voxels with effect

Selection

$S \subseteq B$

$\phi(S) = |S \cap F|$ number of effects in the selection

Counting the number of effects

Simultaneous lower bounds for the number of effects

$$\mathrm{P}(\underline{\phi}_{\alpha}(S) \leq \phi(S) \ \forall \ S) \geq 1 - \alpha$$

These bounds are valid for *all* S, e.g. clusters (data-driven selections), anatomical regions (knowledge-driven selections), etc.

Reference

Goeman and Solari (2011)
Multiple Testing for Exploratory Research
Statistical Science

Other approaches

- If the selection S is data-driven (e.g. clusters), then S is a random variable
- This implies that the global null $H_S : \bigcap H_i$ and the hypotheses $\{H_i, i \in S\}$ are random. How to deal with random hypotheses?
- FWER control over *all* the hypotheses $\{H_i, i \in B\}$ implies

$$R_\alpha \subseteq F \Rightarrow S \cap R_\alpha \subseteq S \cap F \quad \forall S$$

  with probability $\geq 1 - \alpha$
- FDR control does not translate to subsets $S \subseteq R_\alpha$:

$$\mathbb{E}\Big(\frac{|R_\alpha \cap T|}{|R_\alpha|}\Big) \leq \alpha \nRightarrow \mathbb{E}\Big(\frac{|S \cap T|}{|S|}\Big) \leq \alpha$$

Lower bounds

cluster	*size*	# *effects*	% *effects*
C_1	2191	≥ 624	≥ 29 %
C_2	1835	≥ 847	≥ 46 %
C_3	1400	≥ 454	≥ 32 %
C_4	698	≥ 0	≥ 0 %
C_5	421	≥ 25	≥ 6 %
C_6	304	≥ 33	≥ 11 %
C_7	245	≥ 0	≥ 0 %
C_8	232	≥ 0	≥ 0 %
C_9	187	≥ 0	≥ 0 %

Bonferroni

cluster	*size*	# *effects*	% *effects*
C_1	2191	≥ 7	≥ 0.3 %
C_2	1835	≥ 86	≥ 4 %
C_3	1400	≥ 82	≥ 6 %
C_4	698	≥ 0	≥ 0 %
C_5	421	≥ 0	≥ 0 %
C_6	304	≥ 0	≥ 0 %
C_7	245	≥ 0	≥ 0 %
C_8	232	≥ 0	≥ 0 %
C_9	187	≥ 0	≥ 0 %

Estimate [bound]

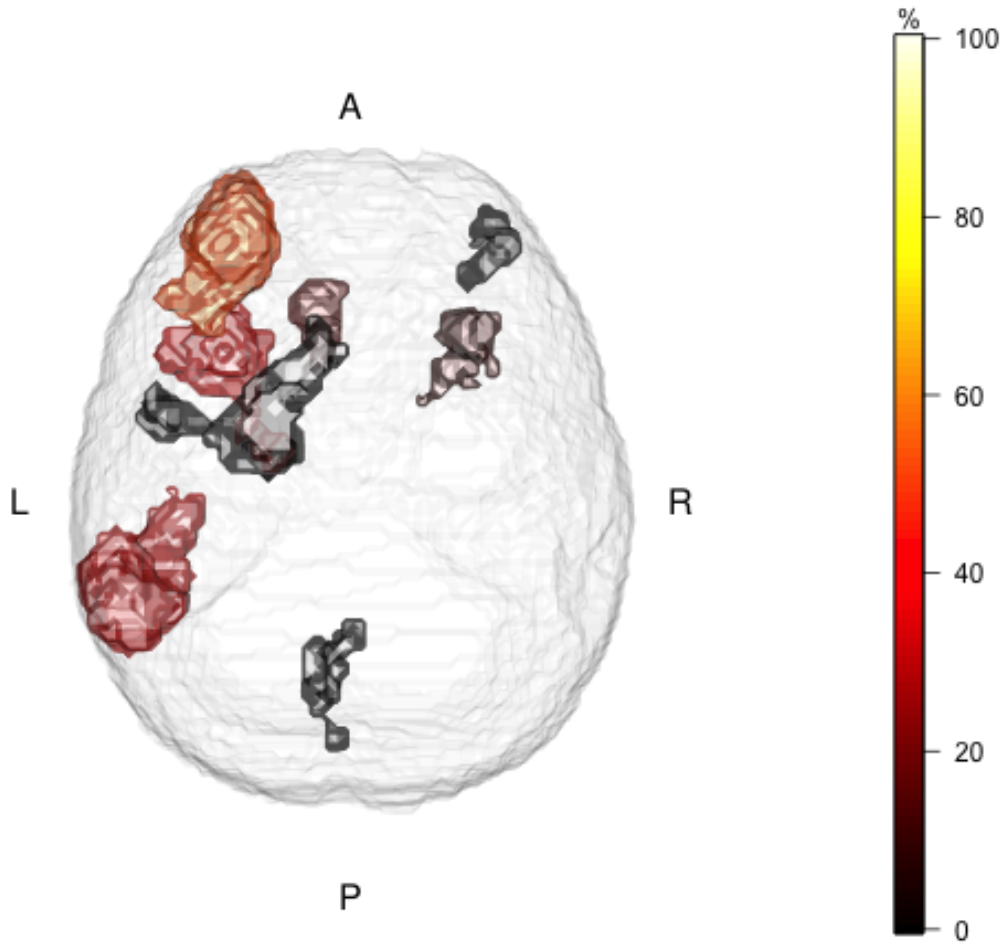
$$\hat{\phi}(S)\ [\underline{\phi}(S)\,] = \underline{\phi}_{50\%}(S)\ [\underline{\phi}_{95\%}(S)]$$

cluster	*estimate [bound]*	
C_1	88 %	[29 %]
C_2	86 %	[46 %]
C_3	81 %	[32 %]
C_4	62 %	[0 %]
C_5	42 %	[6 %]
C_6	49 %	[11 %]
C_7	0 %	[0 %]
C_8	20 %	[0 %]
C_9	1 %	[0 %]

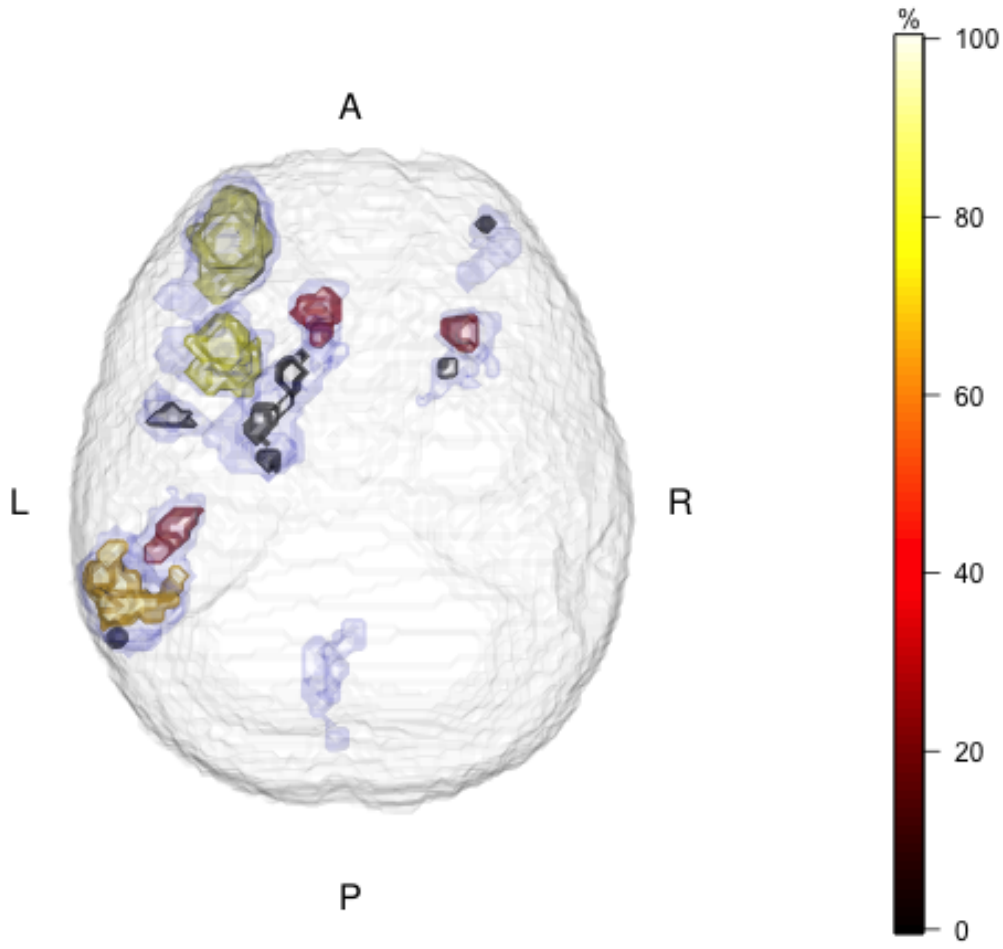
Interactive inference

- The user looks at the data and select interesting $S_1, S_2, \ldots$
- The method informs the user about $\underline{\phi}_\alpha(S_1), \underline{\phi}_\alpha(S_2), \ldots$
- The user may consider different $S'_1, S'_2, \ldots$ based on results, domain knowledge, etc.
- The method informs the user about $\underline{\phi}_\alpha(S'_1), \underline{\phi}_\alpha(S'_2), \ldots$
- Etc.

T > 3.2



T > 4



Sub-clusters

cluster		*threshold*	*size*	*# effects*	*% effects*
C_1		$T > 3.2$	2191	624	29 %
	1	$T > 4$	405	267	66 %
	2	$T > 4$	133	31	23 %
	3	$T > 4$	6	0	0 %
C_2		$T > 3.2$	1835	847	46 %
	1	$T > 4$	963	826	86 %
C_3		$T > 3.2$	1400	454	32 %
	1	$T > 4$	583	449	77 %
	2	$T > 4$	4	0	0 %
	3	$T > 4$	1	0	0 %
$\vdots$					

Anatomical regions