

BRISBANE
OHBM 2025 JUNE
24-28

Cluster analysis revisited

Wouter Weeda

'Classical' cluster-extent analysis

LABAR ET AL.

700

a. Working memory

b. Conjunction: Brain regions engaged in both working memory and spatial attention

c. Brain regions selective for working memory

d. Brain regions selective for spatial attention

FIG. 3. Functional MRI results. (a) Dorsal and medial views of the activation patterns in the working memory and spatial attention tasks. (b) Conjunction analysis: brain regions equally engaged in verbal working memory and covert spatial attention. (c) Subtraction analysis: brain regions selectively engaged in verbal working memory. (d) Subtraction analysis: brain regions selectively engaged in covert spatial attention. Abbreviations: Chl, cerebellum; Ex, extrastriate cortex; FEF, frontal eye field; IFG, inferior frontal gyrus; IPS, intraparietal sulcus; MFG, middle frontal gyrus; OT, occipitotemporal junction; Pre, precentral sulcus; SMA, supplementary motor area; Th, thalamus.

Journal, A. Muts, H.A. Alho et al.

NeuroImage 265 (2023) 119812

319

C.L. Harenski, S. Hamann / NeuroImage 30 (2006) 313–324

a) **b)**

Fig. 3. Activity in (a) posterior STS (BA 39) and (b) posterior cingulate (BA 31) ($P < 0.001$, uncorrected) during the viewing of moral relative to non-moral pictures. Activations are displayed on representative sagittal structural images at the level $x = -51$ and $x = -15$, respectively, in MNI coordinates.

viewed both moral and non-moral pictures. This region has been activated by a variety of tasks involving emotional appraisal, including viewing emotional pictures (Lane et al., 1999; Gusnard and Raichle, 2001) and emotional autobiographical recall (Reiman, 1997). MPFC activation has also been associated with theory of mind tasks (Fryd, 2001), and tasks involving self-referential judgments (Kelley et al., 2002). Taken together, these studies implicate MPFC in the processing of emotional stimuli and the judgments (Kelley et al., 2002). Although MPFC activation has been reported during moral but not non-moral pictures in previous studies (Greene et al., 2001; Moll et al., 2002a,b; Heekeren et al., 2003, 2005), the current findings suggest that this region is also recruited by non-moral stimuli that are closely matched to moral stimuli on degree of social and emotional content. Thus, it seems more parsimonious to interpret activation in this region while processing morally salient picture stimuli as reflecting the more general processing of social and emotional elements and thus activate this region strongly. Whether this conclusion applies to increased MPFC activation that has been reported in studies using other moral tasks (Greene et al., 2001, 2004; Heekeren et al., 2003), however, remains an open question.

Although we observed similar patterns of activation during passive viewing of moral and non-moral pictures, a direct contrast between both conditions revealed greater activation during the moral condition in two regions: posterior STS (BA 39) and the posterior cingulate (BA 31). The finding of STS activation in the moral viewing condition only replicates the Moll et al. (2002b) study that also used a moral vs. non-moral picture viewing task.

Responses to viewing moral and non-moral pictures

Activation was observed in the amygdala when subjects were passively viewing both moral and non-moral pictures, consistent with previous studies showing that the amygdala is strongly responsive to emotional arousal (Hamann, 2001) and expected given that emotional content was equivalent across both types of pictures. MPFC activation was also observed while subjects

Developmental Cognitive Neuroscience 31 (2018) 20–34

non-targets in comparison to calm non-targets ($t(33) = -3.610$, $p = .001$; Fig. 2a). The number of miss trials differed marginally between emotions ($p = .065$). Post-hoc tests showed that more misses were found between the groups ($p = .000$) in the number of miss trials. Subsequent analyses of reaction times showed a similar main effect of emotion on hit trials ($F(2, 64) = 4.863$, $p = .011$), due to faster responses to angry ($t(33) = 2.392$, $p = .023$) and happy ($t(33) = 3.128$, $p = .004$) compared to calm trials. Reaction times for due to faster responses to happy faces than angry faces ($t(29) = -2.071$, $p = .047$; Fig. 2b). No differences were found when comparing the emotional to calm stimuli, suggesting that the heightened false alarm rate in the emotional conditions was not due to a speed-accuracy trade-off. No differences were found between risk-taking groups for any of the behavioural or reaction time indices.

3.2. Imaging results

3.2.1. Whole brain analysis

Initial whole brain analyses (FWE cluster corrected, $Z > 3.1$, $p < .05$) confirmed that regions previously implicated in go/no-go and emotional task performance were also activated in our sample. For the no-go > go contrast, clusters of activation were found in the temporal, dorso- and ventrolateral prefrontal cortex, parietal cortex and the emotional (happy > angry) > calm contrast, one cluster of activations in Table 3. Fig. 3 shows an overview of the a priori ROIs (blue) and the actual activation values in our sample (red, FWE clusterwise $p < .05$). For the emotional > calm contrast no overlap was found. For the no-go > go contrast overlap between a priori ROIs was concentrated on the amygdala.

3.2.2. Region of interest analysis

Further examination, targeting the a priori ROIs revealed differences in the magnitude of activation between go and no-go trials in the right ventral striatum ($F(2, 160) = 9.593$, $p = .002$), and in the right amygdala at trend level ($F(1, 160) = 6.383$, $p = .012$). In these regions activation was greater during no-go trials, which required higher levels of cognitive control than during go trials.

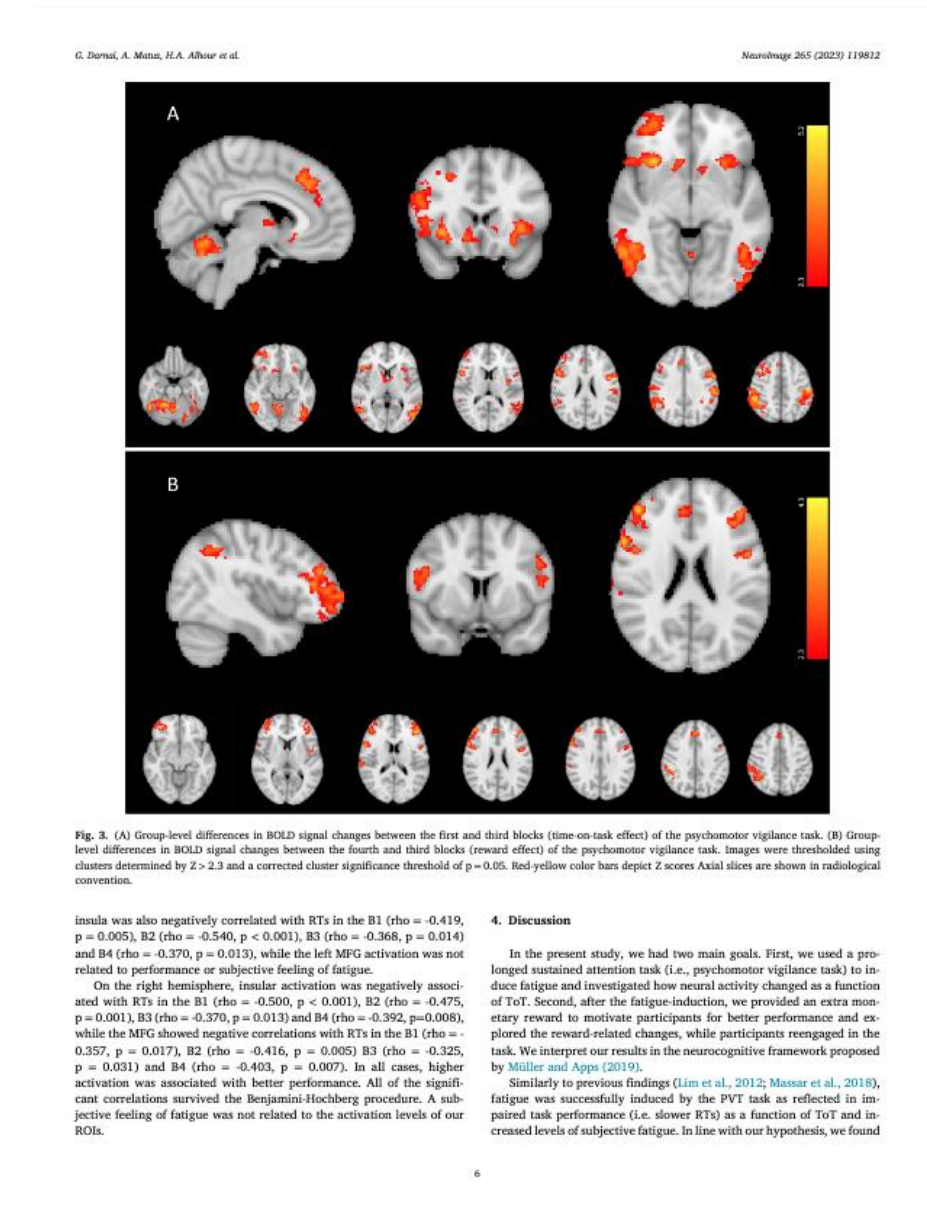
Analyses also showed a bilateral interaction effect of Trial Type and Emotion in the left ventral striatum ($F(2, 160) = 3.367$, $p = .037$) and right ventral striatum ($F(2, 160) = 3.536$, $p = .031$). While it did not meet the Bonferroni adjusted p-value, it did fall within uncorrected thresholds. We tentatively report the results here as Bonferroni corrected (Feldman et al., 2011, p. 117), similar effects were found bilaterally, and the findings are in line with previous research demonstrating that the ventral striatum is known to play an important role in reward processing (e.g. Delgado, 2007). Post hoc t-tests showed this was due to greater activation during happy no-go than happy go trials in right/left

Fig. 8. VBM-style analysis showing reduction of grey-matter in adolescent-onset schizophrenic patients, compared with controls. 3 different thresholding methods were compared: TFCE, cluster-based and voxel-based. Thresholding for all methods was at $P < 0.05$, corrected for multiple comparisons across space using permutation testing.

Fig. 3. Whole brain activation for the no-go > go contrast (FWE cluster corrected, $Z > 3.1$, $p < .05$) showing activation in the right temporal cortex, the right dorso- and ventrolateral prefrontal cortex, right parietal cortex and the right dorso-ventral cingulate cortex. Coordinates are in MNI space.

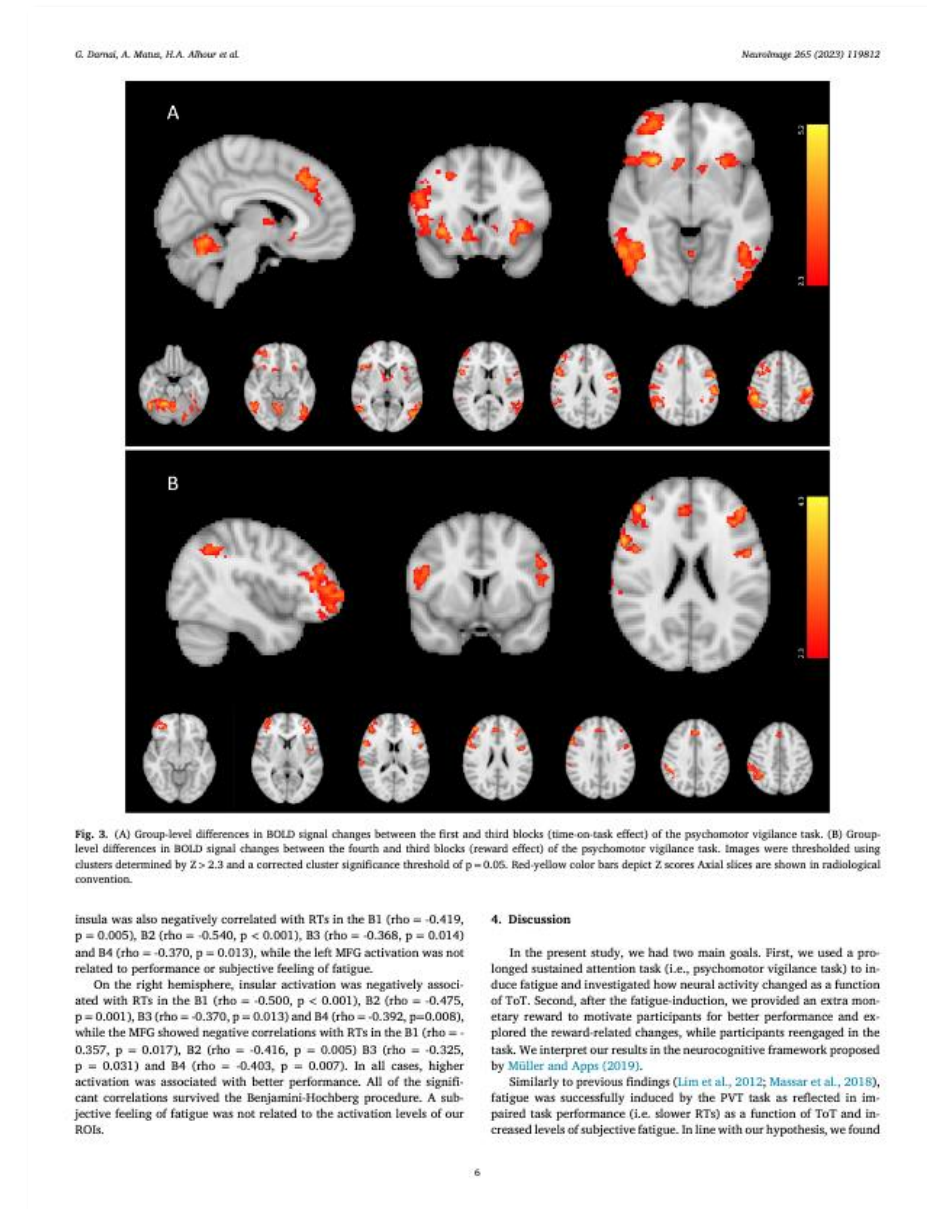
‘Classical’ cluster-extent analysis

- For most functional MRI studies measured signal comes from distinct locations in the brain called voxels: a 3-dimensional grid of 3x3x3 mm cubes.
- Inference in functional MRI is done on each location (voxel) separately.
- The maps that you often see are the outcomes of this inference (usually in the form of a z or t-statistic indicating significance).



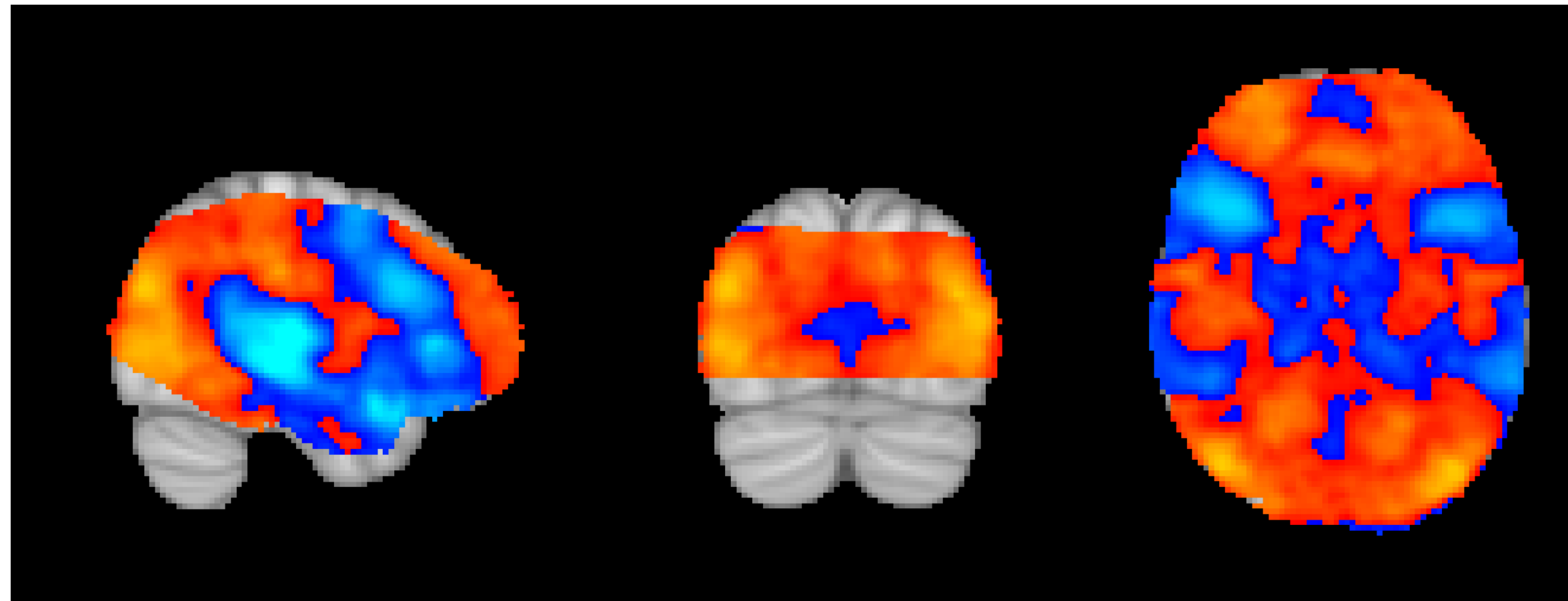
‘Classical’ cluster-extent analysis

- The goal of fMRI inference is to decide for each voxel whether it is active or not (using a hypothesis test).
- For each test we allow a little uncertainty of whether our decision is the right one.
- When doing multiple tests, the chances of making a wrong decision somewhere in our ‘family’ of tests increases dramatically.
- The family-wise error rate (FWER) of our family of tests is what we want ‘controlled’.



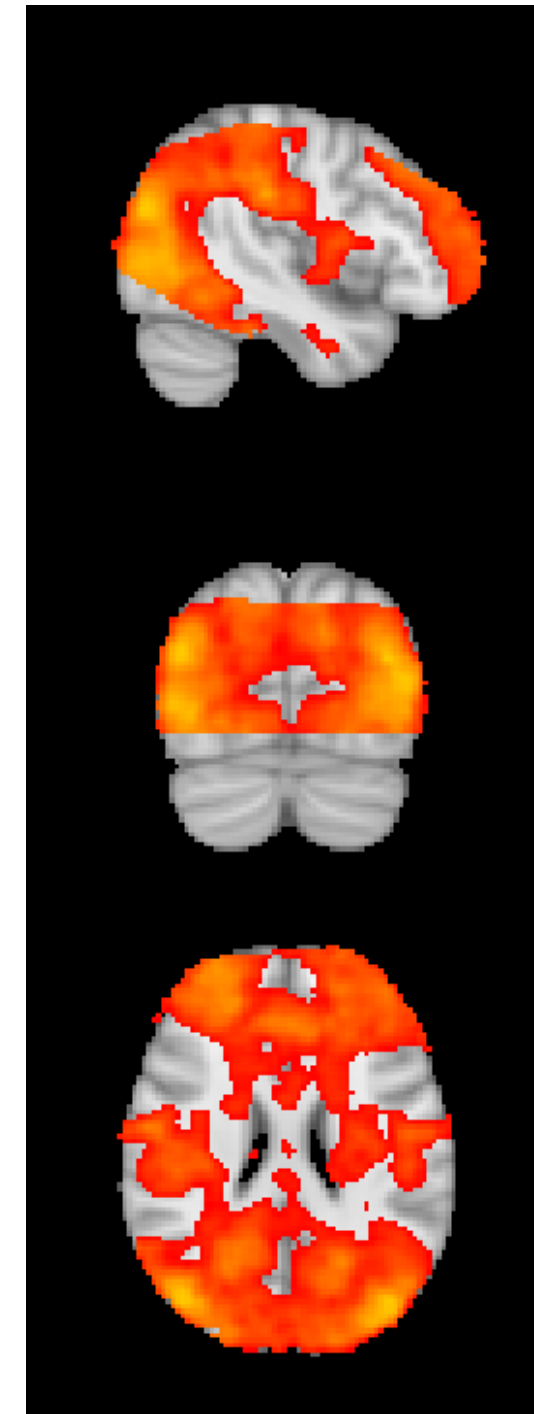
‘Classical’ cluster-extent analysis

Study on vocal and non-vocal sounds, Pernet et al., 2015



‘Classical’ cluster-extent analysis

- In total 166.407 in-mask voxels.
- Focus only on positive values for now.
- Z-statistics indicate whether a voxel is more active in the *non-vocal* condition than in the *vocal* condition.
 - H_0 = not active (z-value = 0)
 - H_1 = active (z-value > 0)

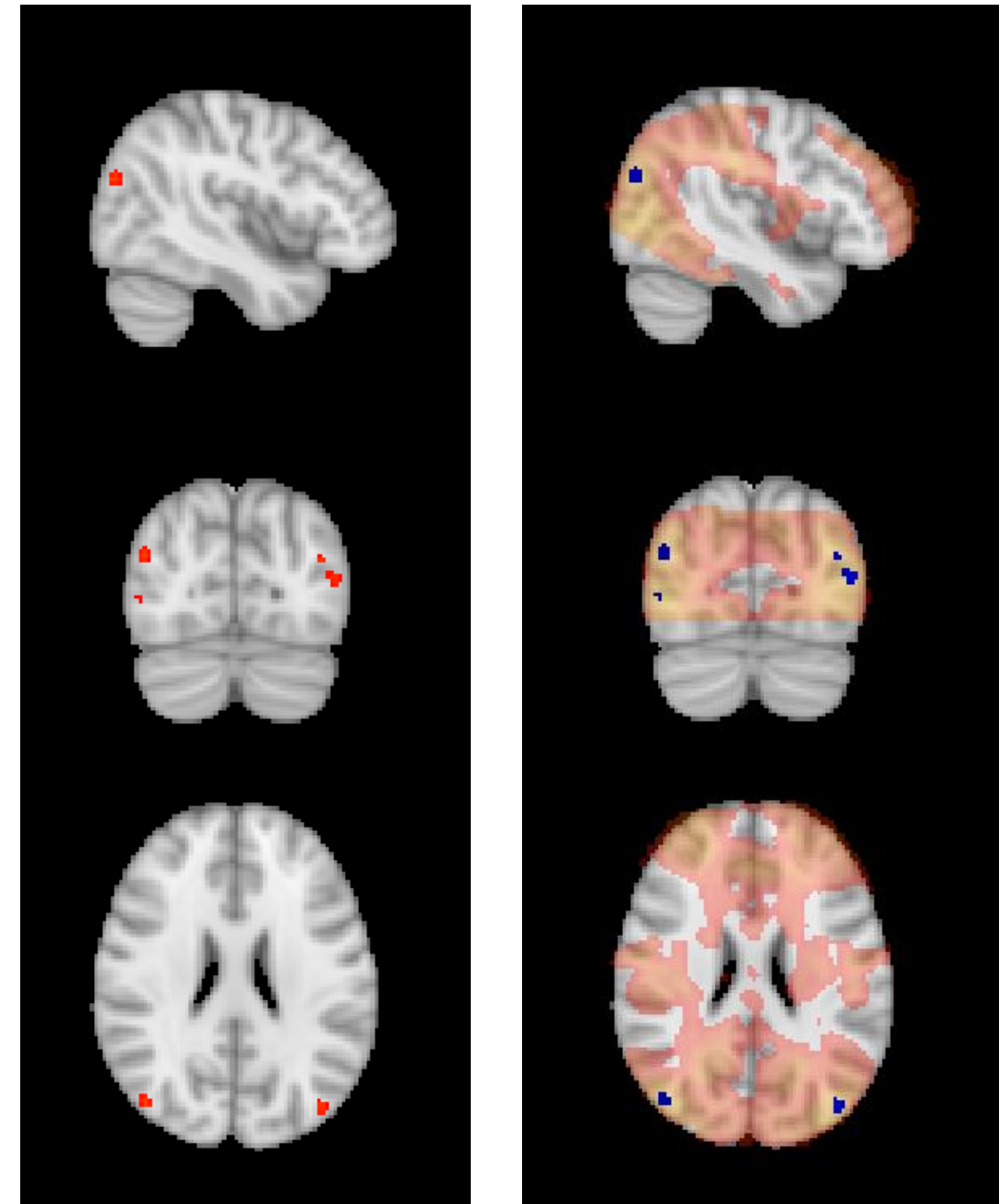


‘Classical’ cluster-extent analysis

- Controls the FWER over all voxels in the brain (mask). Family = all voxels.
- Easiest method to control the FWER is Bonferroni correction.
- Calculated by setting the per-voxel α to be $\alpha / \# \text{voxels}$

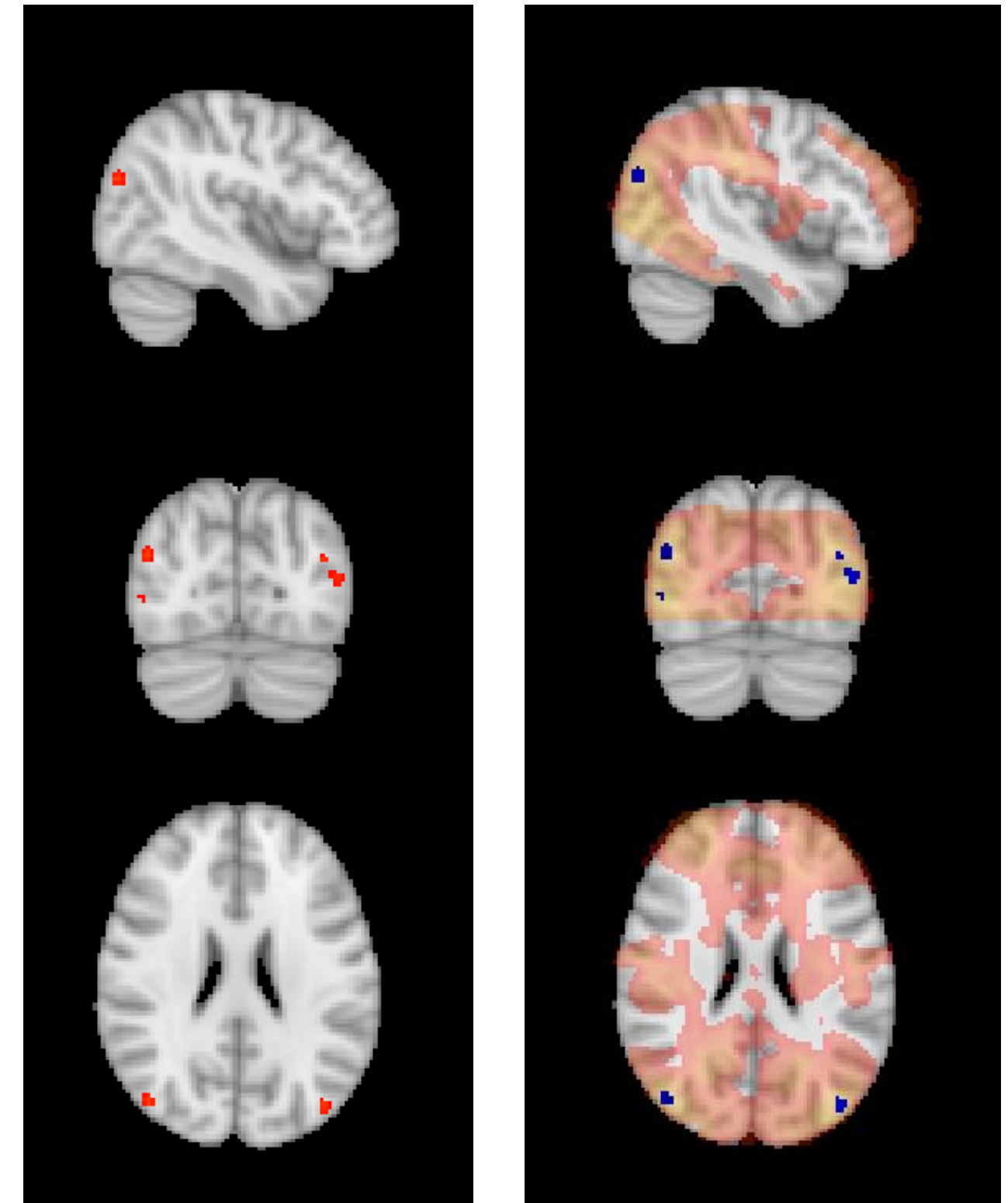
$$.05 / 166407 = .0000003$$

- Usually not very powerful.



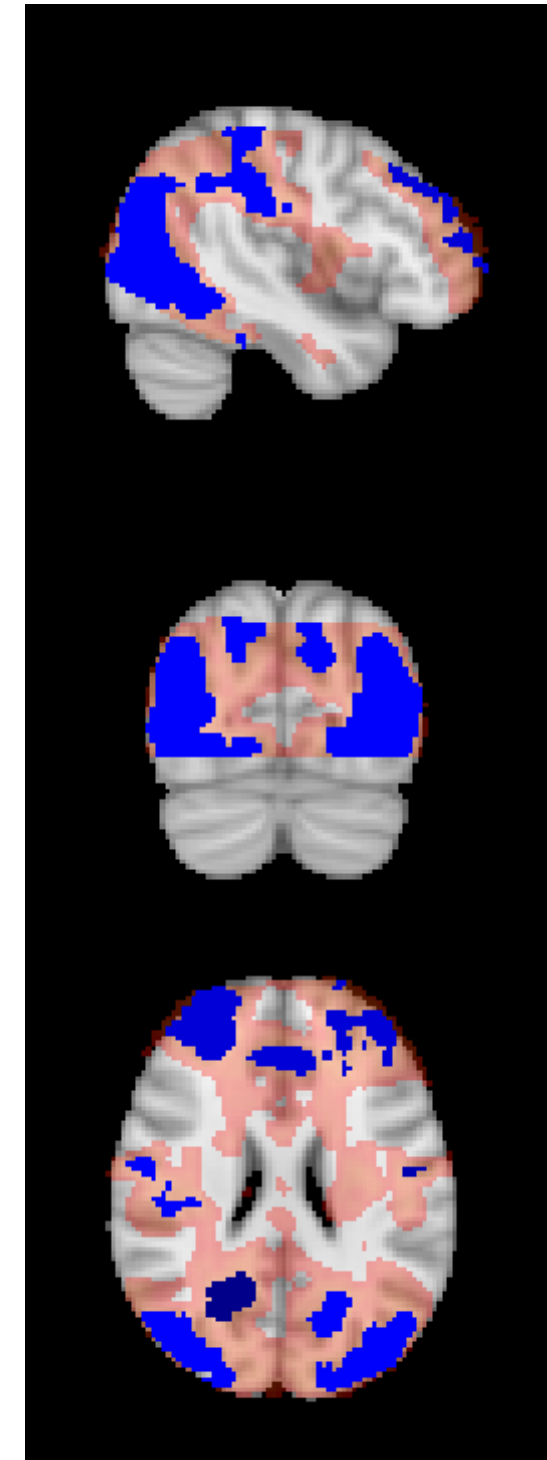
‘Classical’ cluster-extent analysis

- But...
- Since our family is all voxels, we know exactly where the activation is!
- In other words: we have high spatial specificity.
- (because the chance of any of these voxels being a false-positive $< 5\%$)



‘Classical’ cluster-extent analysis

- Usually, we are not interested in single-voxel activity per se. A more natural unit is a ‘cluster’ of voxels (which we will name ‘blob’).
- A cluster or blob is defined as a contiguous/connected set of voxels.
- We control the number of false-positive blobs (our family in FWER is thus all possible blobs, not all voxels).

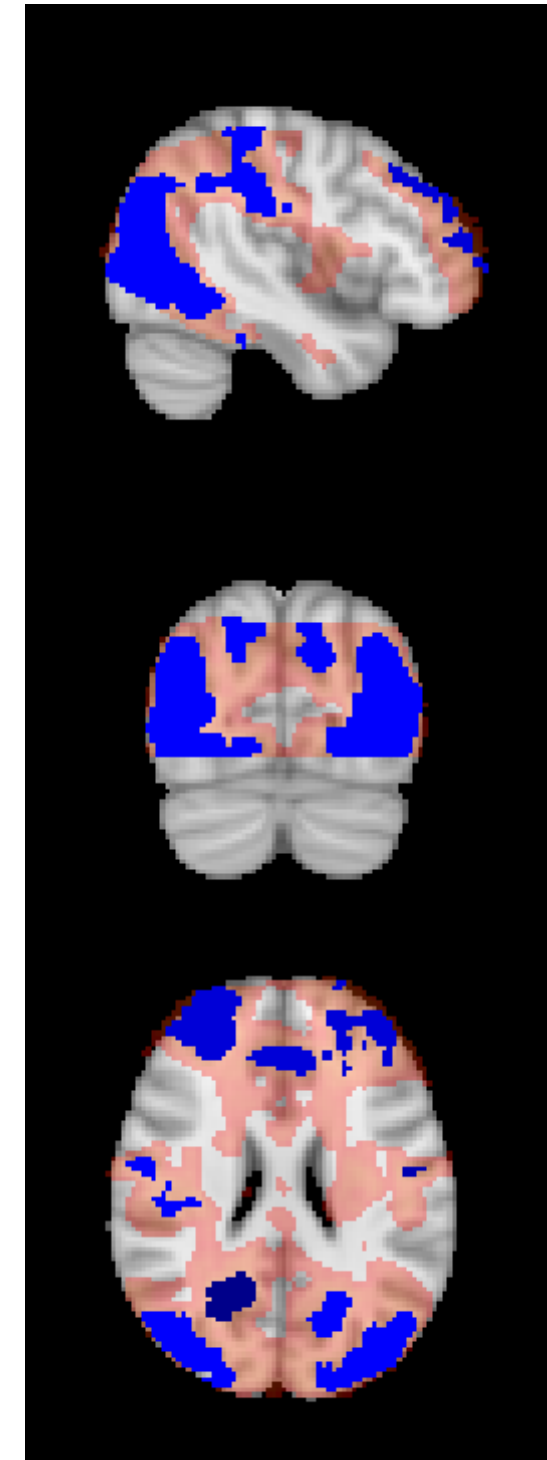


‘Classical’ cluster-extent analysis

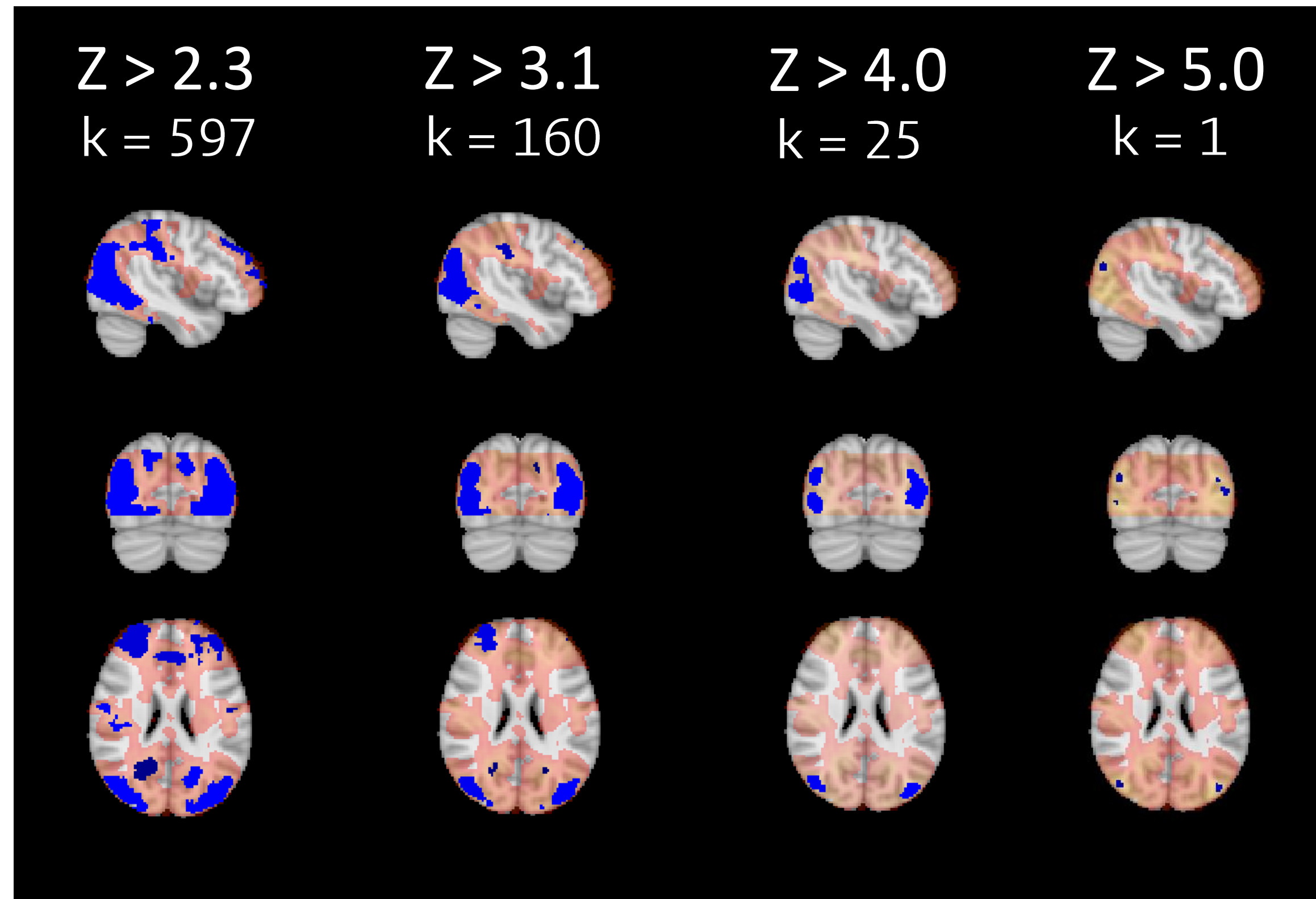
- In practice, using a two-step approach:
 1. Choose a ‘cluster-forming’ threshold z and estimate the size of all contiguous clusters above this threshold.

Determine the minimum cluster size k that occurs by chance under the null (95%) given the smoothness of the data and the chosen threshold z (e.g., using RFT or permutations)

2. Check which clusters are larger than k (all clusters that are larger are significant).

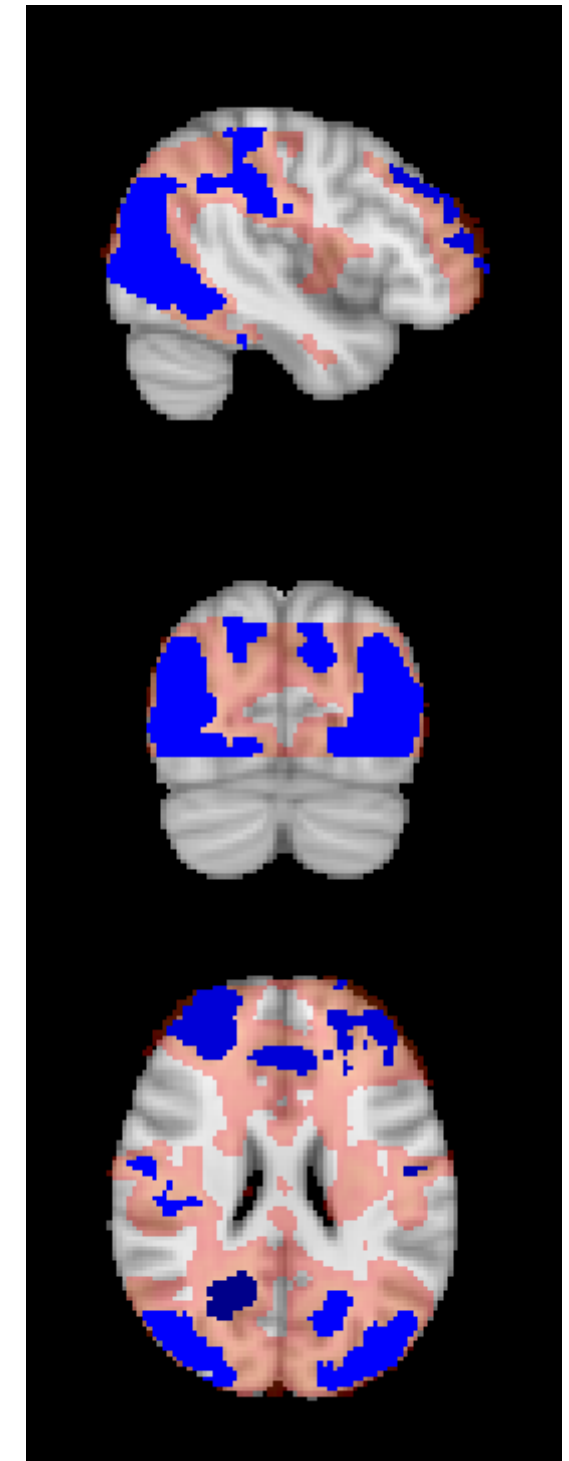


‘Classical’ cluster-extent analysis



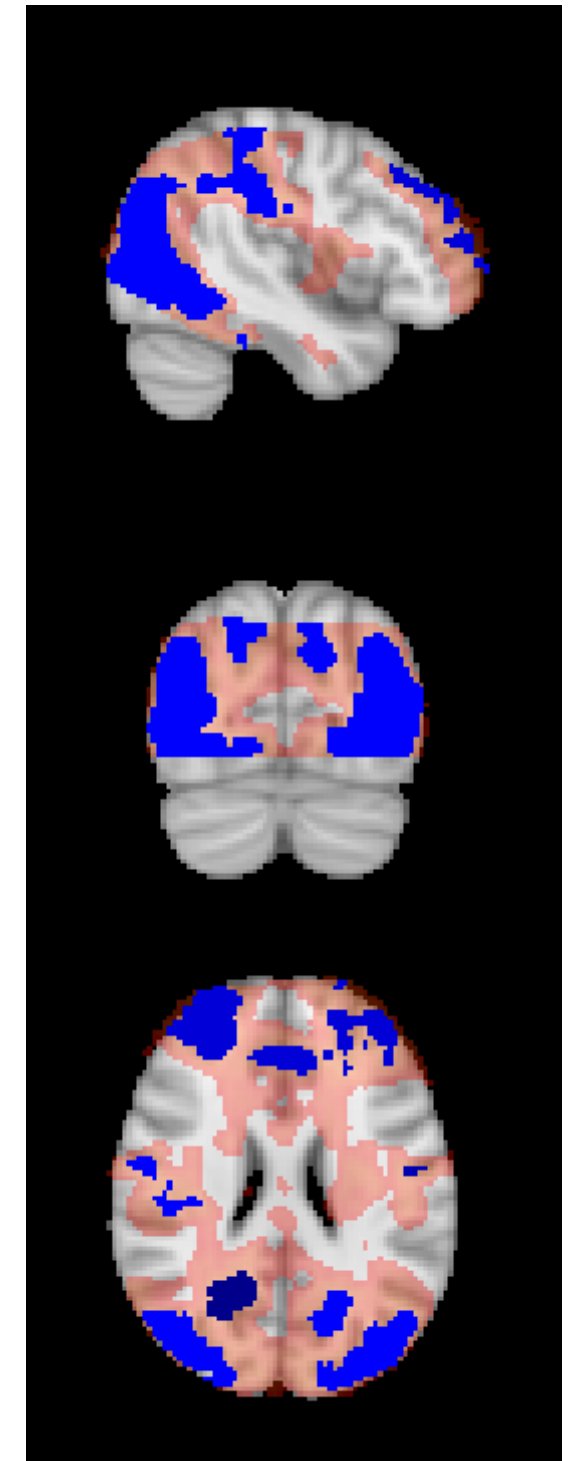
‘Classical’ cluster-extent analysis

- More powerful than voxel-wise approaches, but... more powerful in detecting activation, not in localizing it.
- Because of hypotheses being on the ‘cluster’ level:
 - Non-significant when cluster-extent is smaller than k
 - Significant when cluster-extent is larger than k
- No information about voxels within a cluster (clusters are large enough or not).

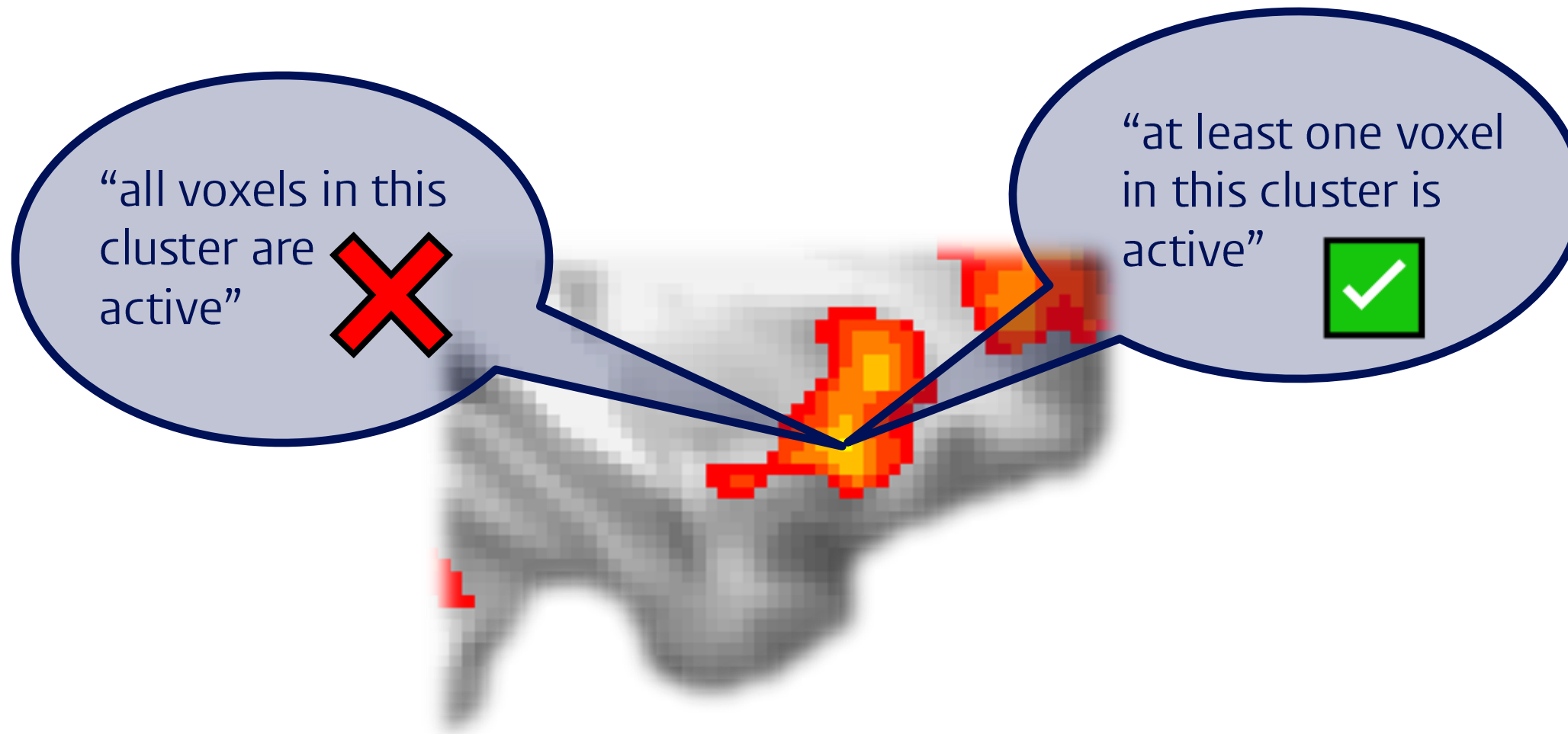


‘Classical’ cluster-extent analysis

- Formal way of stating this:
 H_0 = no activation within a cluster
 H_1 = at least one voxel active within a cluster
- So, the larger the cluster found, the less we know about activation within a cluster.
- This is called the Spatial specificity paradox (Woo et al., 2014, Lindquist & Mejia, 2015).



Spatial Specificity Paradox

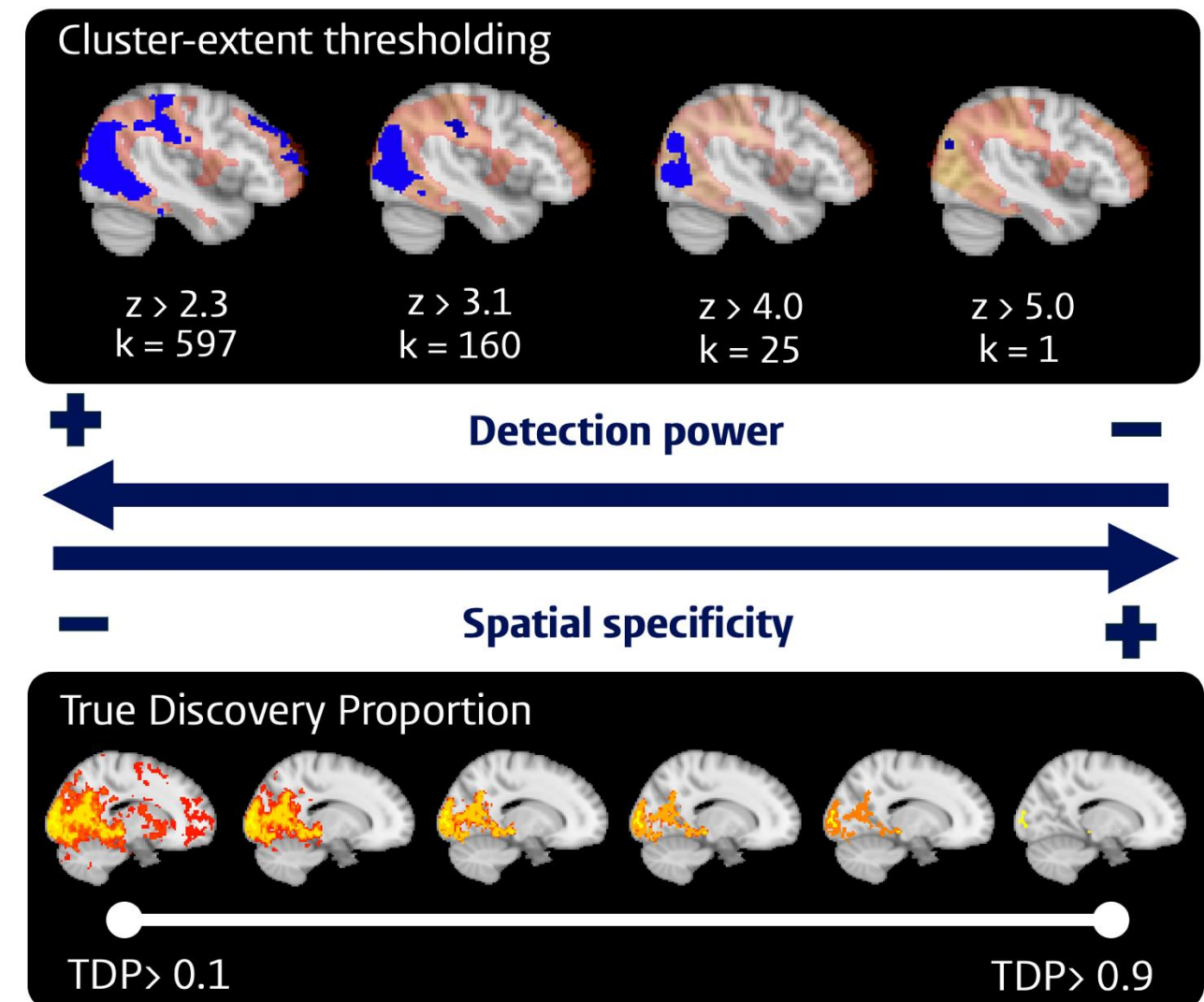


H_0 = no activation within a cluster

H_1 = at least one voxel active within a cluster

Spatial Specificity Paradox

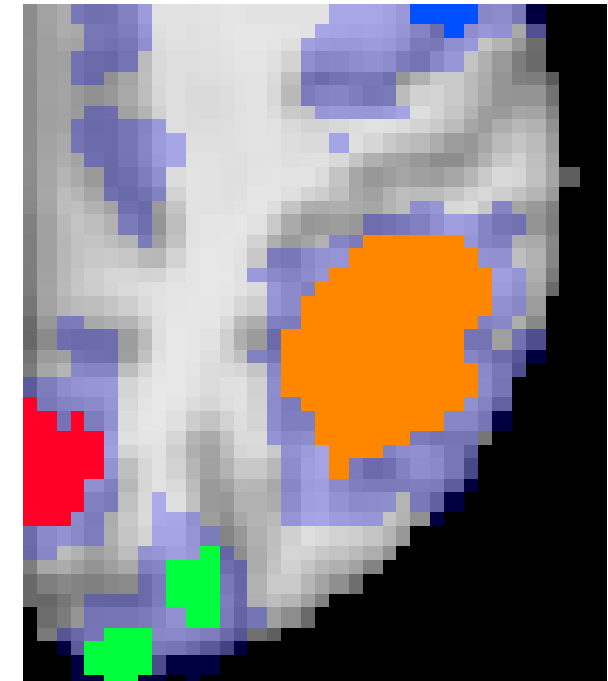
- The statement “there is at least one voxel active” could mean that all voxels in a cluster are active. Or it could be one, we just don’t know.
- Intuition is that there is usually more than one voxel active. But cluster-extent statistics don’t allow us to test that.
- We can use TDP based methods to give us an in-depth analysis of what clusters are made of.



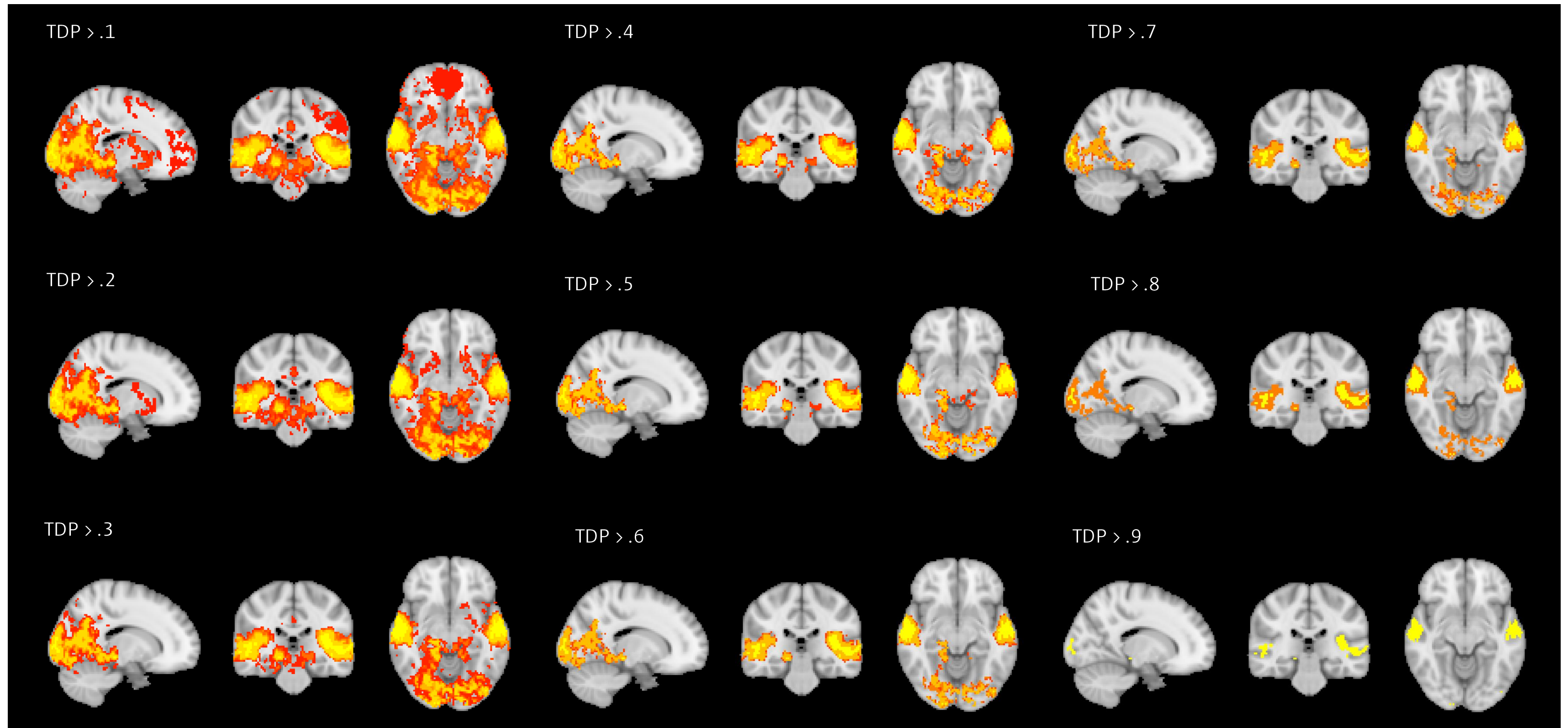
True Discovery Proportion (TDP) based methods

Rosenblatt, Finos, Weeda, Solari & Goeman (2018)

- Multiple testing correction method based on closed testing.
- True Discovery Proportion (TDP) based methods allow us to estimate the number of truly active voxels within a cluster, for all possible clusters, as many times a researcher wants, with full FWER control.
- We can also estimate clusters with at least a certain TDP. For example, “what is the largest cluster that contains at least 60% active voxels ($\text{TDP} > .6$)?”

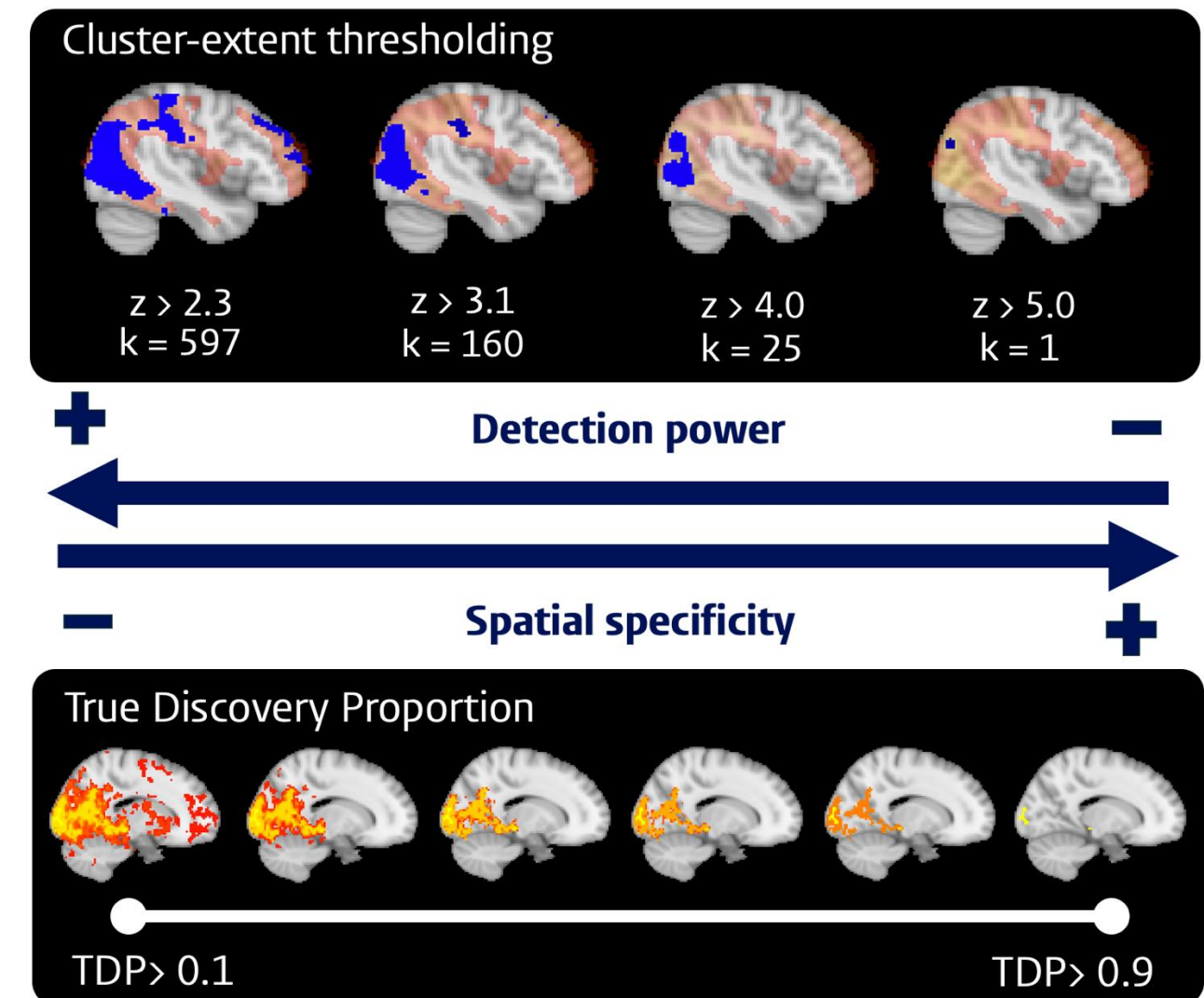


True Discovery Proportion (TDP) based methods



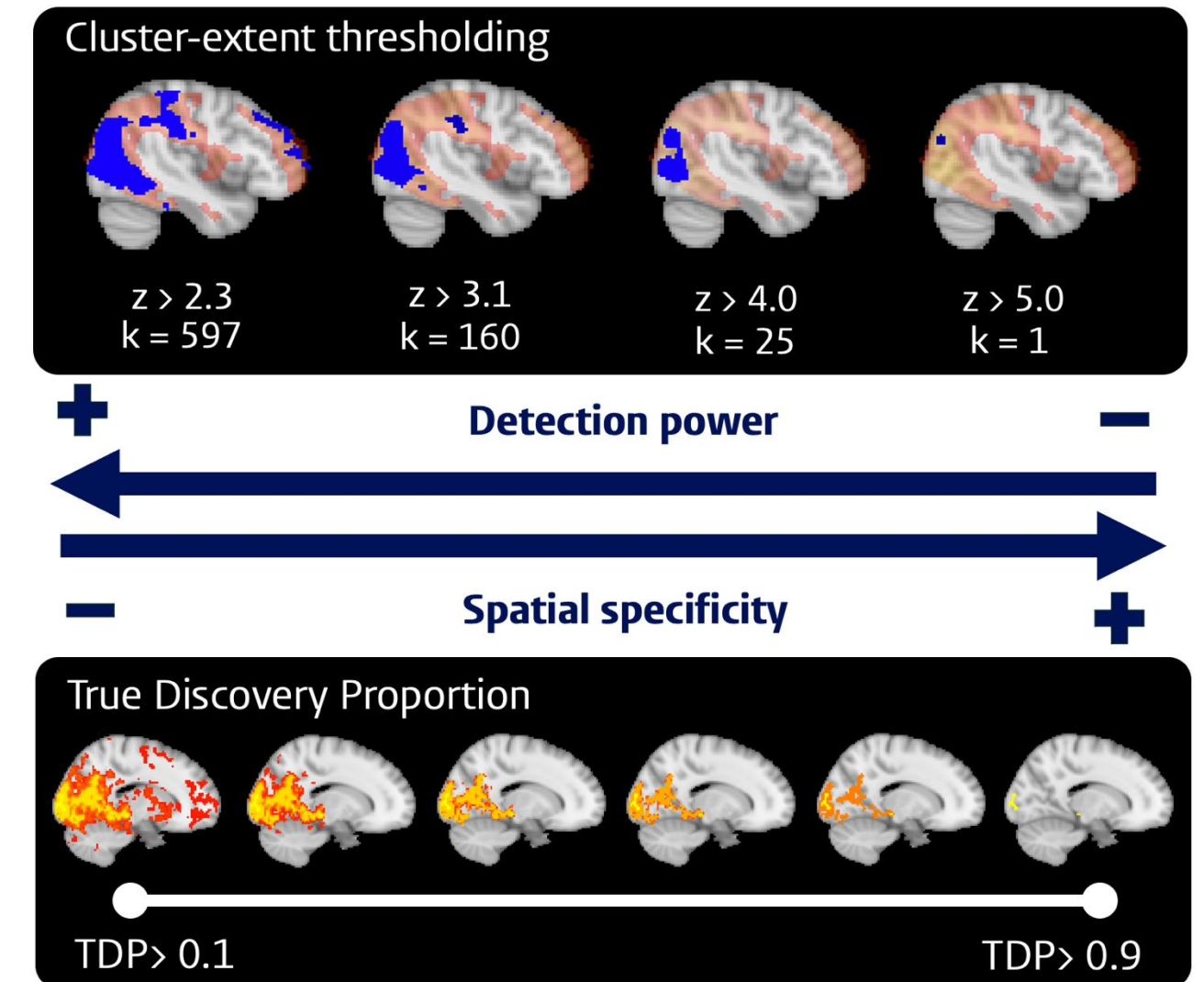
True Discovery Proportion (TDP) based methods

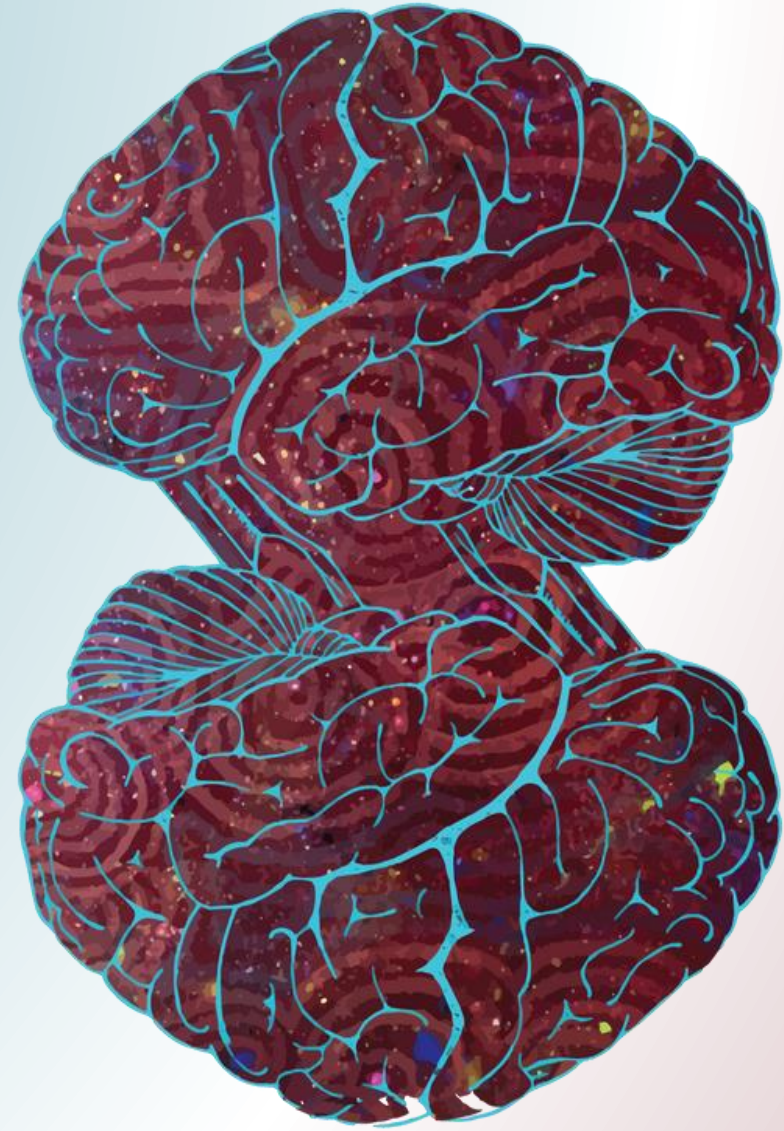
- Both methods are a trade-off between detection and localization, but:
- While cluster extent inference cannot go beyond the cluster-level
- TDP based methods can quantify this trade-off explicitly.



True Discovery Proportion (TDP) based methods

- TDP-based methods can mitigate the spatial shortcomings of cluster-extent thresholding by:
- Quantifying the amount of spatial uncertainty/specificity.
- Allowing researchers to drill-down into clusters until they are happy with a certain TDP (post-hoc).
- Setting a desired TDP level and estimating all clusters with at least this TDP (i.e., choosing a spatial specificity level).





BRISBANE

OHBM 2025 JUNE 24-28