

Case Study: auditory data

The Auditory dataset was collected by Pernet et al. (2015), and generously shared via the OpenfMRI initiative at <https://openfmri.org/>. It consists of 218 subjects passively listening to vocal (i.e. speech) and non-vocal sounds. We used an original sample with a typical fMRI sample size of 33 subjects.

A type of inference which deserves particular attention is *cluster-based inference*. It is now the most common type of inference, being the default option in several popular software suites. Cluster inference can be seen as inference at a *data-driven scale*. This is because the size of the clusters is not selected a priori, but rather determined by the data used for inference.

The fact that clusters are both defined and tested with the same data introduces a statistical circularity challenge typically solved using a *random field theory* (RFT) approach, which permits both FWER control on clusters.

Standard analysis

Given a Z -score map, we defined clusters of interest using pre-specified cluster-forming Z -threshold and minimal cluster size.

- The pre-specified cluster-forming Z -threshold is 3.2
- The cluster size threshold 118, computed by using RFT at $\alpha = 5\%$

Cluster-based analysis showed activity in 6 regions of interest commonly found in auditory studies.

Cluster	Threshold	Size	Max statistic	MNI coordinates	RFT p-value
Right Heschl's/STG/IFG/PT	$Z > 3.2$	6907	7.83	58 -14 2	0.0000
Left Heschl's/STG/PT	$Z > 3.2$	4607	7.51	-60 -22 4	0.0000
Left IFG	$Z > 3.2$	385	4.54	-58 14 30	0.0005
Right precentral gyrus	$Z > 3.2$	249	4.88	52 2 52	0.0050
Left amygdala	$Z > 3.2$	168	4.59	-18 -8 -10	0.0249

We observed activity bilaterally in the superior temporal gyrus (STG), planum temporale (PT), Heschl's gyrus (HG), inferior frontal gyrus (IFG), and amygdala, and activity in the right precentral gyrus.

Import data:

```
# import Z-map
require(RNifti)
Zmap = readNifti("zmap.nii.gz")
# mask
mask = readNifti("mask.nii.gz")
mask = (mask >= 1)
Zmap[!mask] = 0
Zmap
```

FALSE Image array of mode "double" (6.9 Mb)
FALSE - 91 x 109 x 91 voxels
FALSE - 2 x 2 x 2 mm per voxel

To see the sizes of the 14 clusters of contiguous voxels with $Z > 3.2$:

```
clusters = readNifti("id_clusters.nii.gz")
clusters = clusters[mask==1]
table(clusters)
```

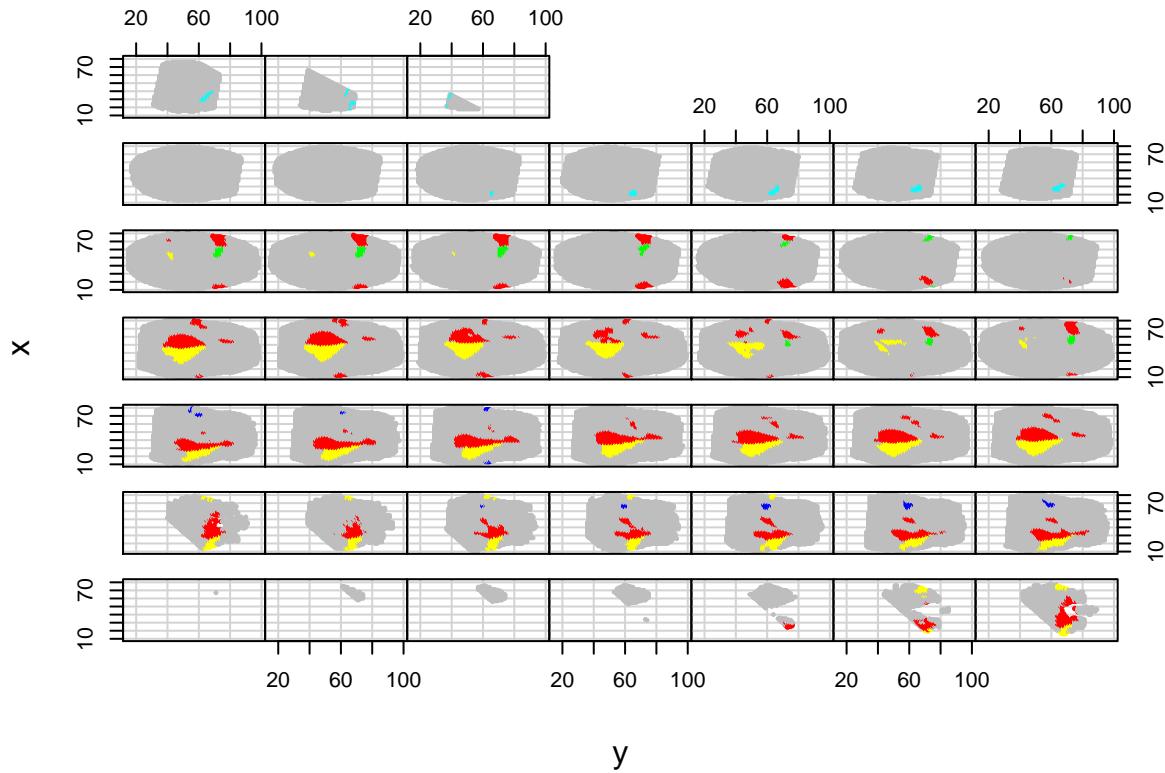
FALSE clusters

	0	1	2	3	4	5	6	7	8	9
FALSE	133273	1	1	2	7	9	11	13	18	23
FALSE	10	11	12	13	14	15	16	17	18	
FALSE	28	30	32	108	168	249	385	4607	6907	

To see the location of 5 clusters with size > 118:

```
coord=which(Zmap!=0, arr.ind = TRUE)
x = coord[,1]
y = coord[,2]
z = coord[,3]
cols7 = clusters
cols7[cols7 <= 13] = 0
cols7[cols7==0]<- "gray"
for (i in 14:18) cols7[cols7==i]<-rainbow(6)[18-i+1]
coplot( x ~ y | as.factor(z), col=cols7, show.given=F, cex=.3, pch=19)
```

Given : as.factor(z)



FDP confidence

Cluster inference also suffers from low spatial resolution, which is demonstrated by the following paradox. Since discovering a cluster means that “there exists at least one voxel with an evoked response in the cluster”, and not that “all the voxels in the cluster have an evoked response”, it follows that *the larger the detected cluster, the less information we have on the location of the activation*. Moreover, cluster-based inference gives no information on the extent of the activation within the cluster.

The matter of low spatial resolution can be remedied by a FDP confidence approach to quantify the activation

within each cluster.

<i>Cluster</i>	<i>Threshold</i>	<i>Size</i>	<i># active</i>	<i>% active</i>
Right Heschl's/STG/IFG/PT	$Z > 3.2$	6907	5179	74.9 %
Left Heschl's/STG/PT	$Z > 3.2$	4607	3409	73.9 %
Left IFG	$Z > 3.2$	385	0	0 %
Right precentral gyrus	$Z > 3.2$	249	15	6.0 %
Left amygdala	$Z > 3.2$	168	0	0 %

The activity in the right hemisphere covered one large cluster (6907 voxels), with a proportion of true discoveries (PTD) of 74.9% (with exception of the precentral gyrus; 249 voxels, PTD = 6.0%). In the left hemisphere these same areas were divided amongst three regions: HG/STG/PT (4607 voxels, PTD = 73.9%), IFG (385 voxels, PTD = 0%), and the amygdala (168 voxels, PTD = 0%).

```
# Z to p
pmap = readNifti("pmap.nii.gz")

library(hommel)
hom = hommel(pmap[mask==1])

res = matrix(NA, nrow=5,ncol=4)
colnames(res)=c("cluster","size","discoveries","PTD")
for (i in 14:18){
  ix = which(clusters==i)
  res[i-13,1] = i
  res[i-13,2] = length(ix)
  res[i-13,3] = discoveries(hom, ix = ix)
  res[i-13,4] = tdp(hom, ix=ix)
}
res

##      cluster size discoveries      PTD
## [1,]      14   168          0 0.00000000
## [2,]      15   249         15 0.06024096
## [3,]      16   385          0 0.00000000
## [4,]      17  4607        3409 0.73996093
## [5,]      18  6907        5179 0.74981902
```