Multiple testing

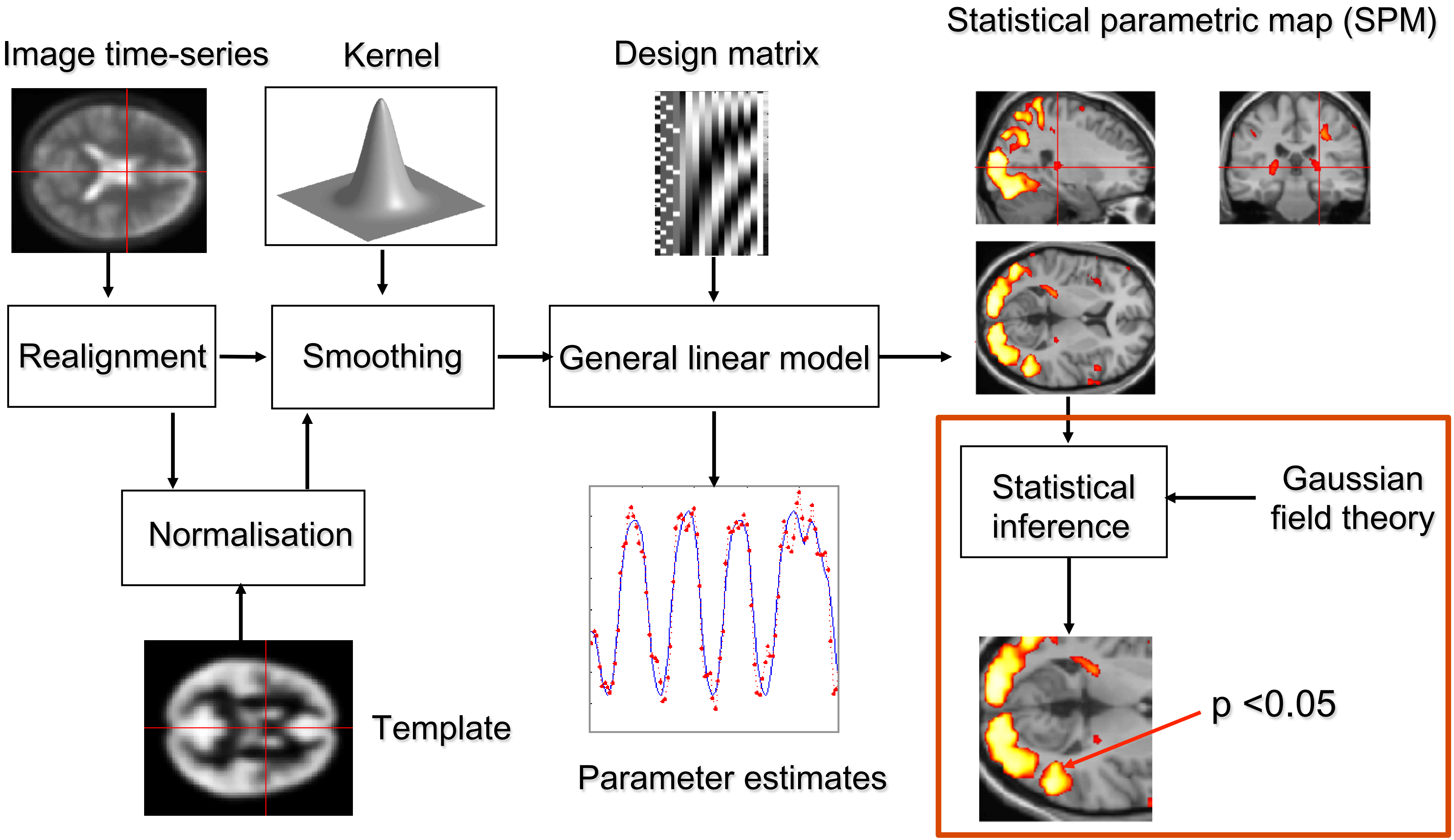
Justin Chumbley

Laboratory for Social and Neural Systems Research
University of Zurich

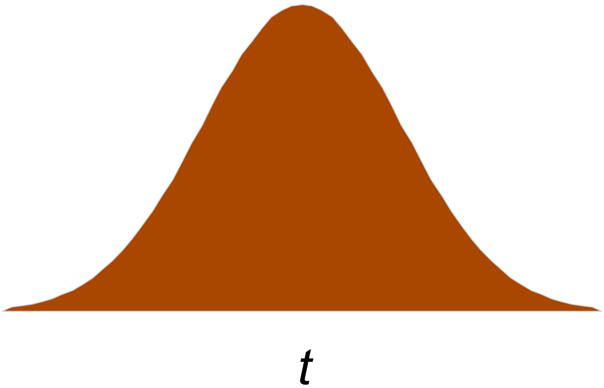
With many thanks for slides & images to:

FIL Methods group

Overview of SPM – Random field theory

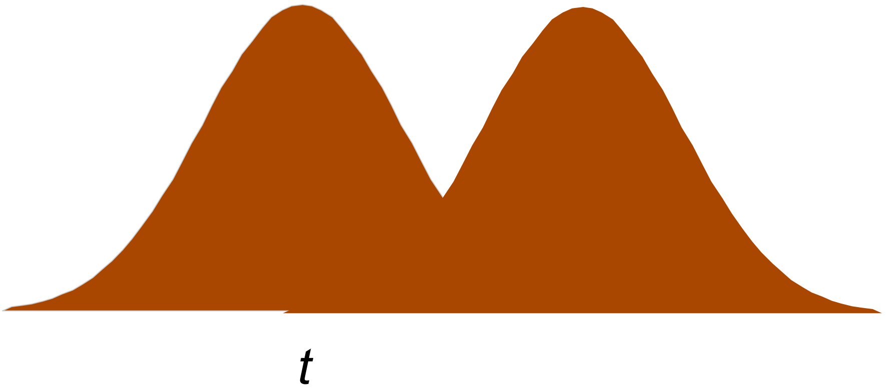


Error at a single voxel



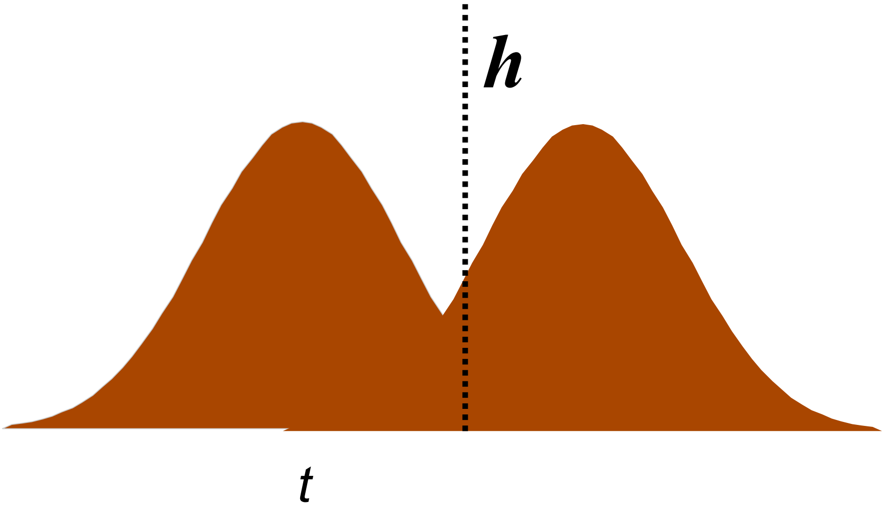
$$t = \frac{\textit{contrast}\ \text{of estimated parameters}}{\sqrt{\text{variance estimate}}}$$

Error at a single voxel



$$t = \frac{\textit{contrast} \text{ of estimated parameters}}{\sqrt{\text{variance estimate}}}$$

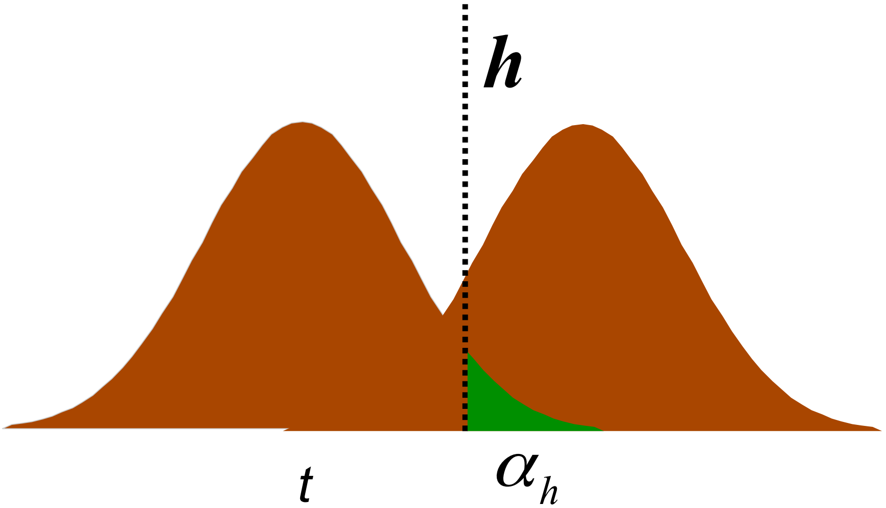
Error at a single voxel



Decision:
H_0 , H_1: zero/non-zero activation

$$t = \frac{\textbf{\textit{contrast} of estimated parameters}}{\sqrt{\textbf{variance estimate}}}$$

Error at a single voxel

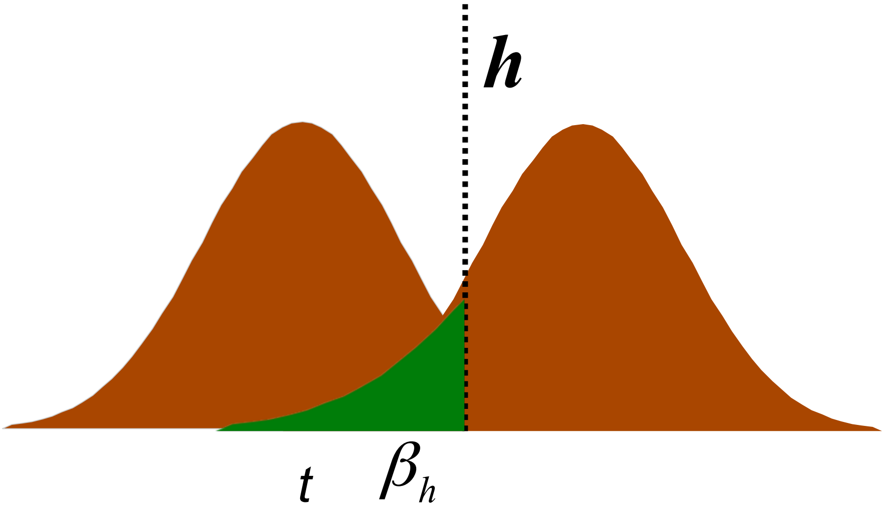


Decision:

H_0 , H_1: zero/non-zero activation

$$t = \frac{\textbf{\textit{contrast} of estimated parameters}}{\sqrt{\textbf{variance estimate}}}$$

Error at a single voxel

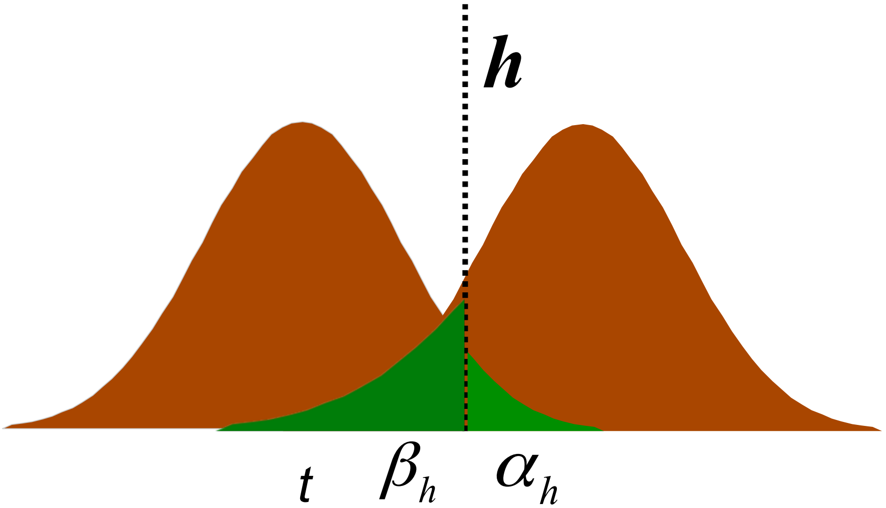


Decision:

H_0 , H_1: zero/non-zero activation

$$t = \frac{\textbf{\textit{contrast} of estimated parameters}}{\sqrt{\textbf{variance estimate}}}$$

Error at a single voxel



Decision:

H_0 , H_1: zero/non-zero activation

Decision rule (threshold) *h*,
determines related error rates α_h , β_h

Convention: Penalize complexity

Choose *h* to give acceptable α_h under H_0

$$t = \frac{\textit{contrast}\ \text{of estimated parameters}}{\sqrt{\text{variance estimate}}}$$

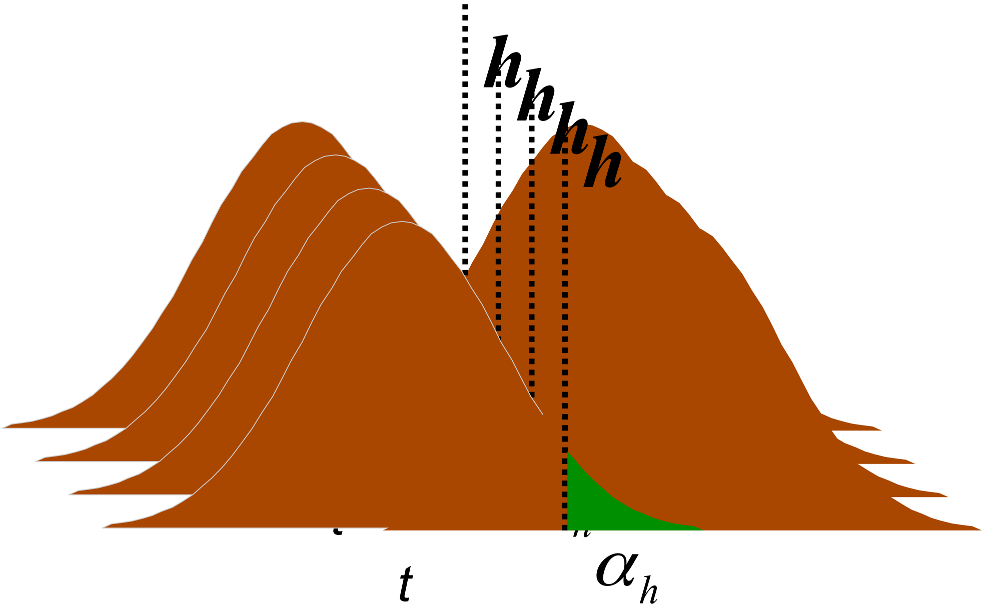
Types of error

Decision \ Reality	H_0	H_1
H_1	False positive (FP) α_h	True positive (TP)
H_0	True negative (TN)	False negative (FN) β_h

specificity: 1- α_h
= TN / (TN + FP)
= proportion of actual negatives which are correctly identified

sensitivity (power): 1- β_h
= TP / (TP + FN)
= proportion of actual positives which are correctly identified

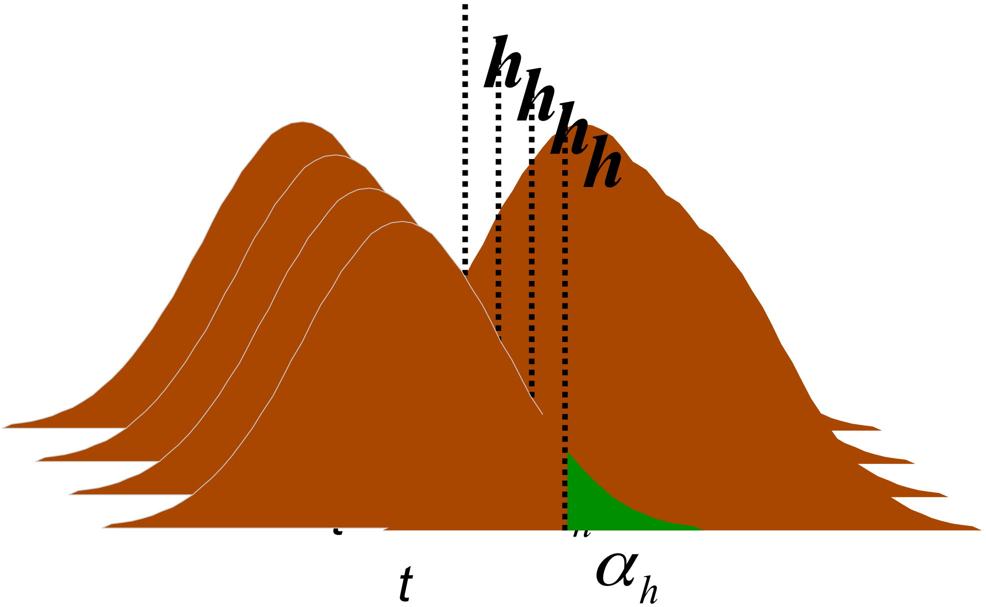
Multiple tests



What is the problem?

$$t = \frac{\textit{contrast} \text{ of estimated parameters}}{\sqrt{\text{variance estimate}}}$$

Multiple tests



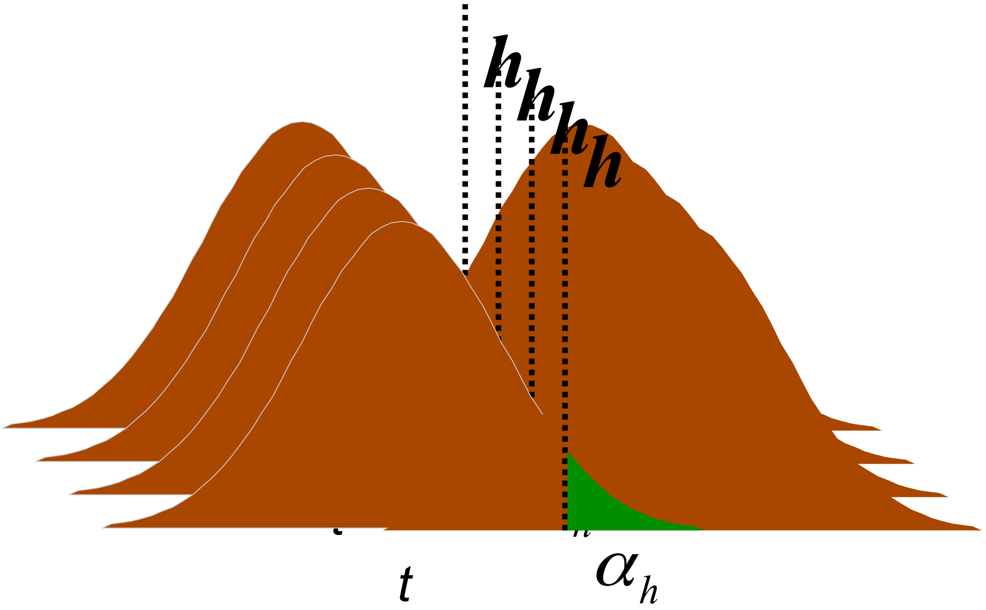
Penalize each independent opportunity for error.

$$p(\ 1 \text{ or more } FP\ ) = FWER_h$$

$$E\left(\frac{FP}{\text{All positives}}\right) = FDR$$

$$t = \frac{\textit{contrast} \text{ of estimated parameters}}{\sqrt{\text{variance estimate}}}$$

Multiple tests



Bonferonni

$$FWER_h \leq N\alpha_h$$

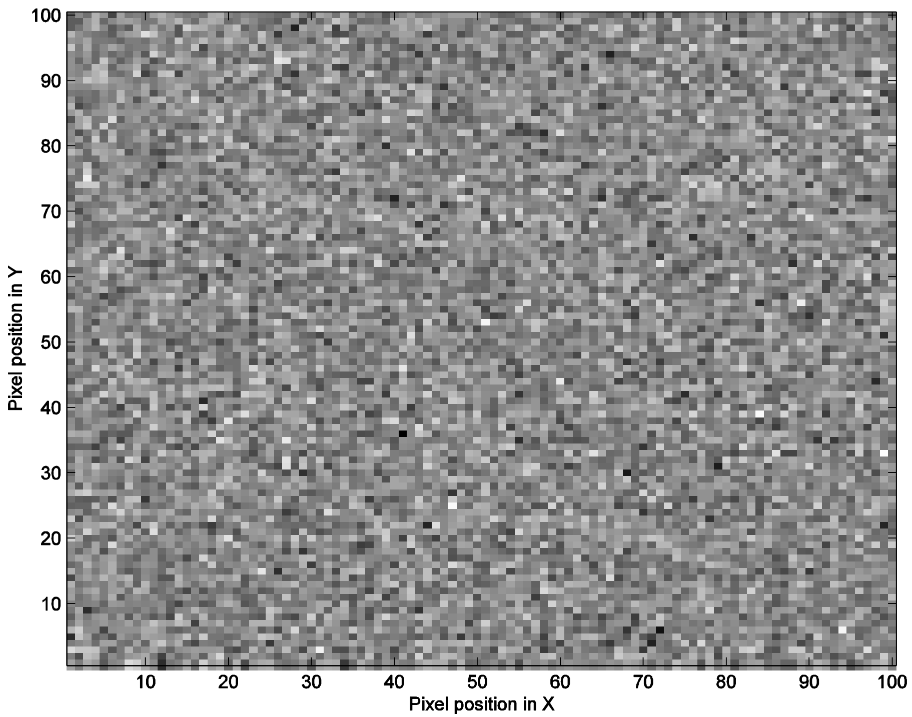
$$FWER_h / N \leq \alpha_h$$

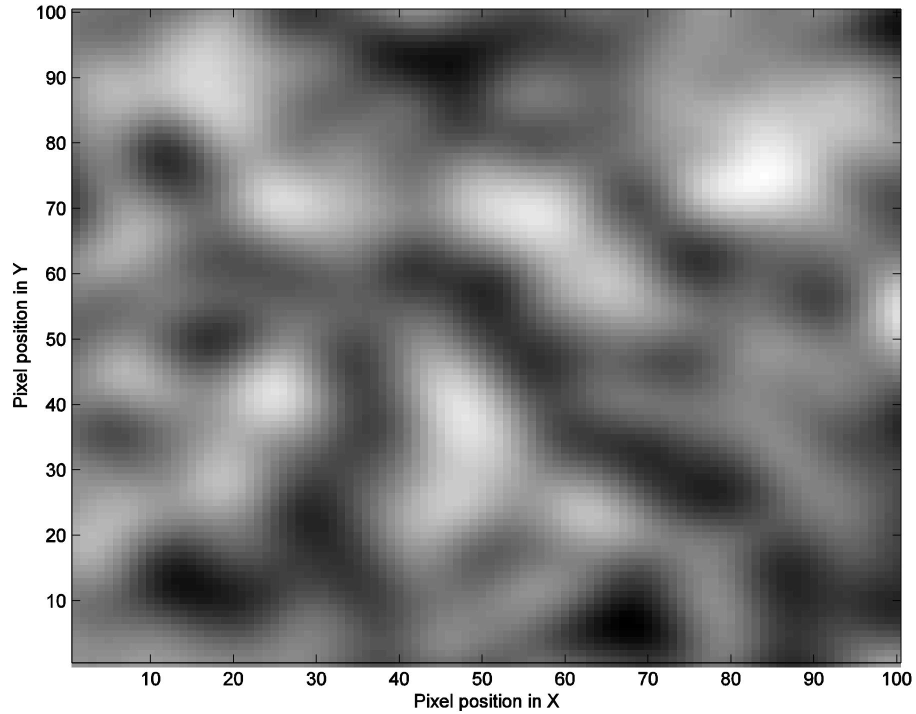
$t = \dfrac{\textit{contrast} \text{ of estimated parameters}}{\sqrt{\text{variance estimate}}}$

Convention: Choose *h* to limit $FWER_h$ assuming family-wise H_0

Issues

1. Voxels or regions
2. Bonferroni too harsh (insensitive)
    - Unnecessary penalty for sampling resolution (#voxels/volume)
    - Unnecessary penalty for independence





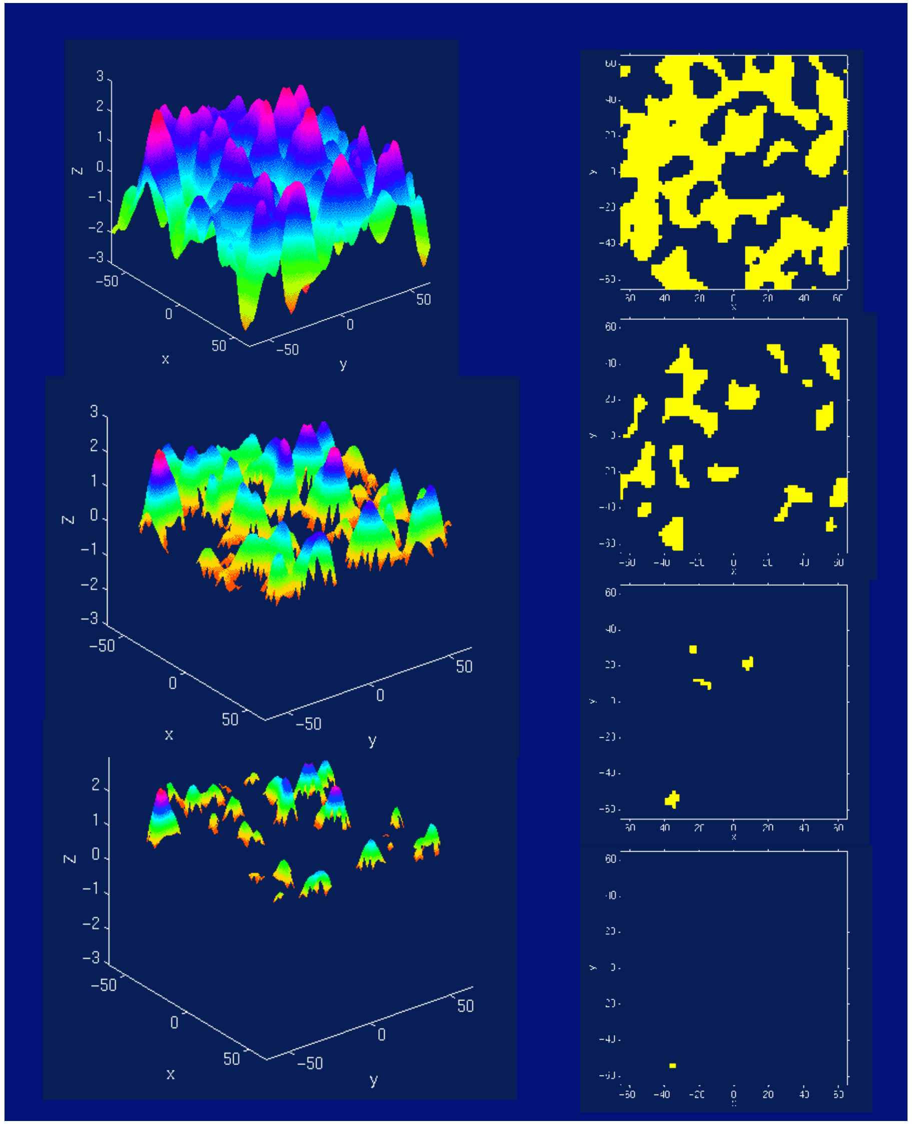
- intrinsic smoothness
  - MRI signals are aquired in k-space (Fourier space); after projection on anatomical space, signals have continuous support
  - diffusion of vasodilatory molecules has extended spatial support
- extrinsic smoothness
  - resampling during preprocessing
  - matched filter theorem
    → deliberate additional smoothing to increase SNR
  - Robustness to between-subject anatomical differences

Acknowledge/estimate dependence
Detect effects in smooth landscape, not voxels

1. Apply high threshold: identify improbably high peaks

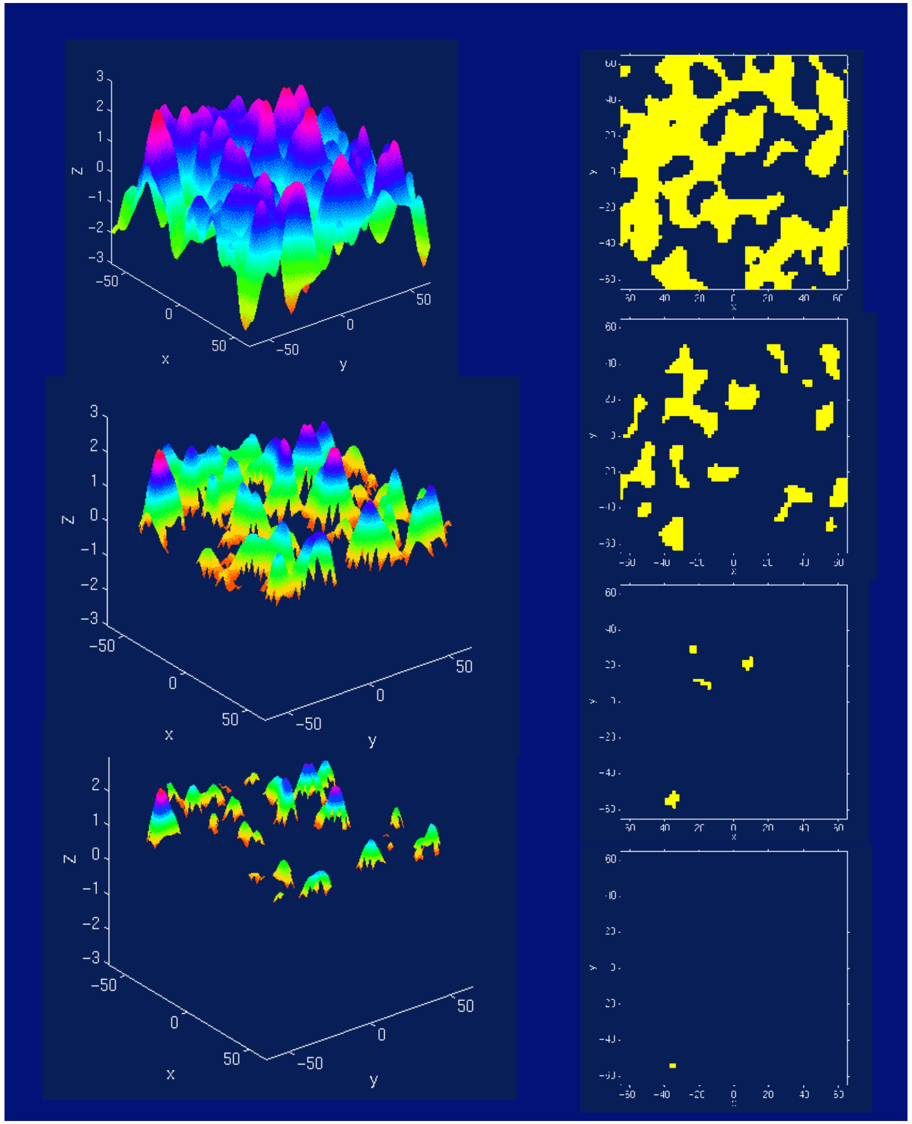
2. Apply lower threshold: identify improbably broad peaks

3. Total number of regions?



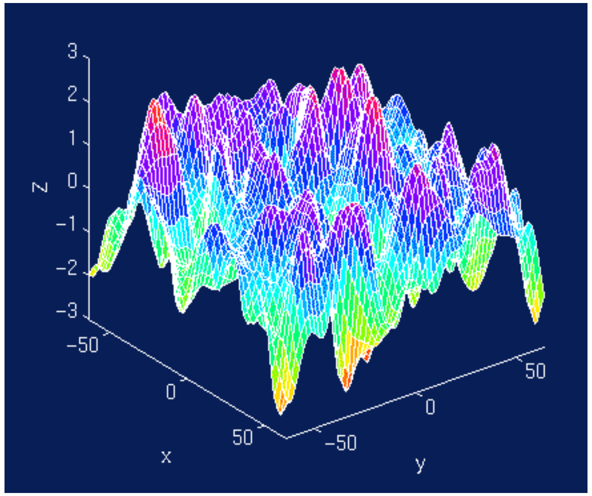
Null distribution?

1. Simulate null experiments
2. Model null experiments

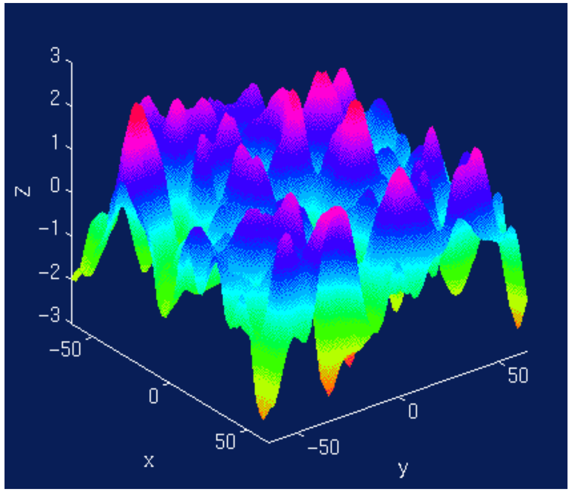


Use continuous random field theory

- image ≈ discretised continuous random field



Discretisation ("lattice approximation")



Smoothness quantified: resolution elements ('resels')

- similar, but not identical to # independent observations
- computed from spatial derivatives of the residuals

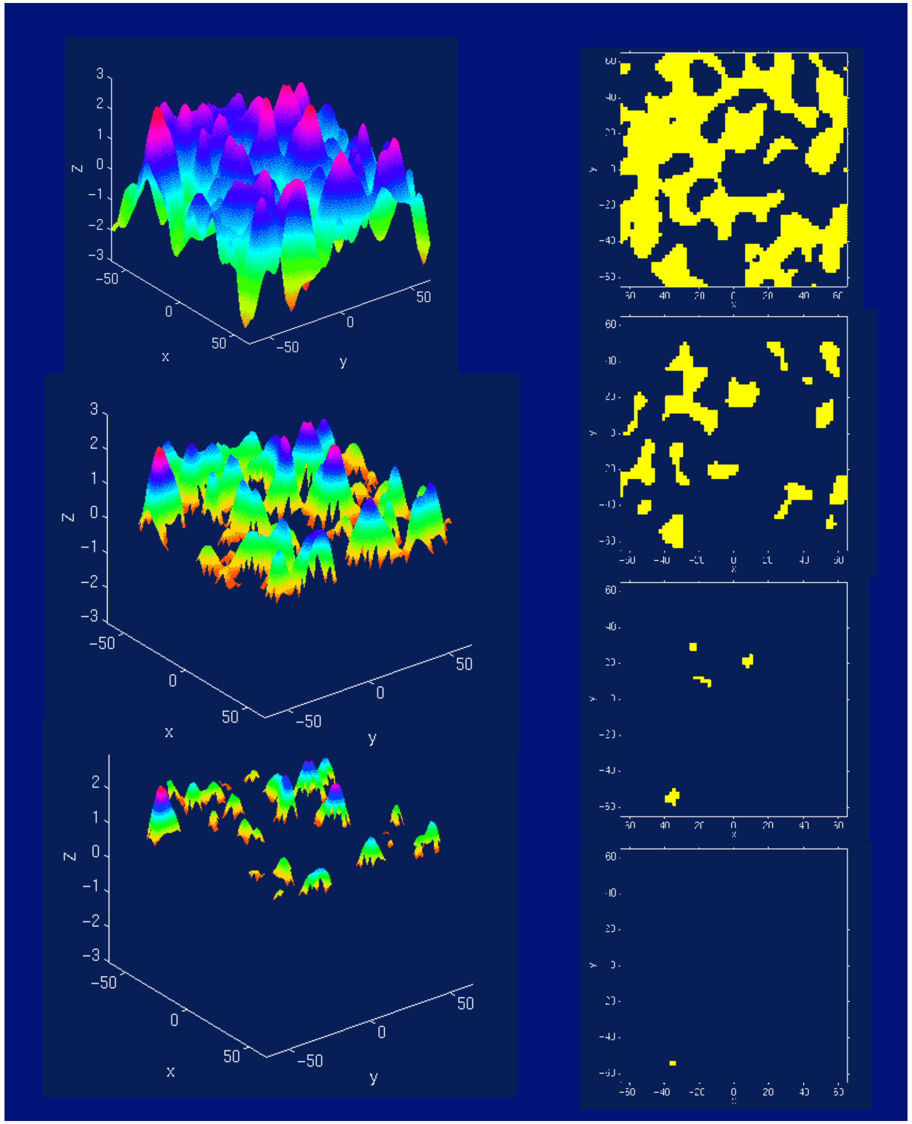
Euler characteristic

– threshold an image at high *h*

blobs = N_h

FWER $\approx$ E [N_h]

= p (blob)



Unified Formula

- General form for expected Euler characteristic
  - χ^2, F, & t fields

$$E[N_h(\Omega)] = \sum_d R_d(\Omega)\, \rho_d(h)$$

Small volumes: Anatomical atlas, ‘functional localisers’, orthogonal contrasts, volume around previously reported coordinates…

$R_d(\Omega)$: *d*-dimensional Minkowski functional of Ω

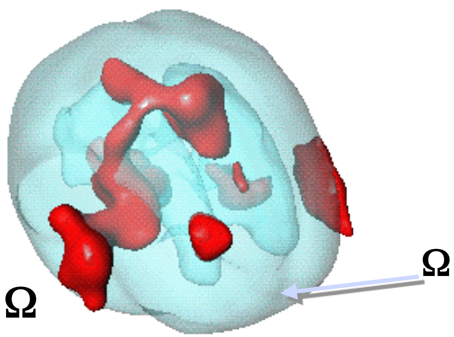
– function of dimension, space **Ω and smoothness:**

$R_0(\Omega)$ = $N(\Omega)$ Euler characteristic of Ω

$R_1(\Omega)$ = resel diameter

$R_2(\Omega)$ = resel surface area

$R_3(\Omega)$ = resel volume



$\rho_d(\Omega)$: *d*-dimensional EC density of $Z(\underline{x})$

– function of dimension and threshold, specific for RF type:

E.g. Gaussian RF:

$$\rho_0(h) = 1 - \Phi(h)$$

$$\rho_1(h) = (4 \ln 2)^{1/2} \exp(-h^2/2) / (2\pi)$$

$$\rho_2(h) = (4 \ln 2) \exp(-h^2/2) / (2\pi)^{3/2}$$

$$\rho_3(h) = (4 \ln 2)^{3/2} (h^2 - 1) \exp(-h^2/2) / (2\pi)^2$$

$$\rho_4(h) = (4 \ln 2)^2 (h^3 - 3h) \exp(-h^2/2) / (2\pi)^{5/2}$$

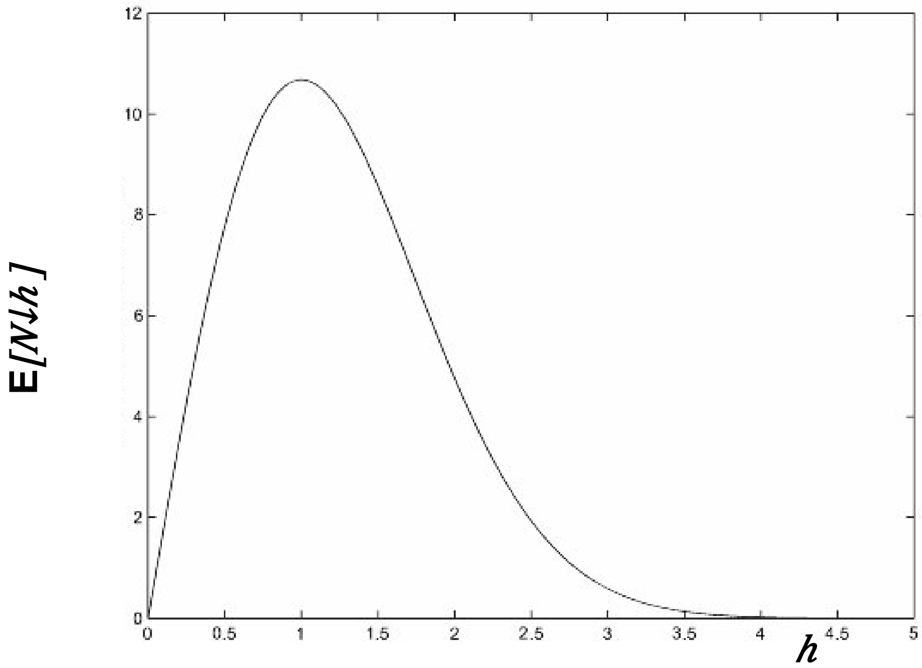
Euler characteristic (EC) for 2D images

$$\mathrm{E}\left[N_h\right] = R(4\log 2)(2\pi)^{-3/2}\, h \exp(-0.5h^2)$$

R = number of resels
h = threshold

Set *h* such that $E[N\downarrow h] = 0.05$

Example: For 100 resels, $E[N\downarrow h]$ = 0.049 for a Z threshold of 3.8. That is, the probability of getting one or more blobs where Z is greater than 3.8, is 0.049.



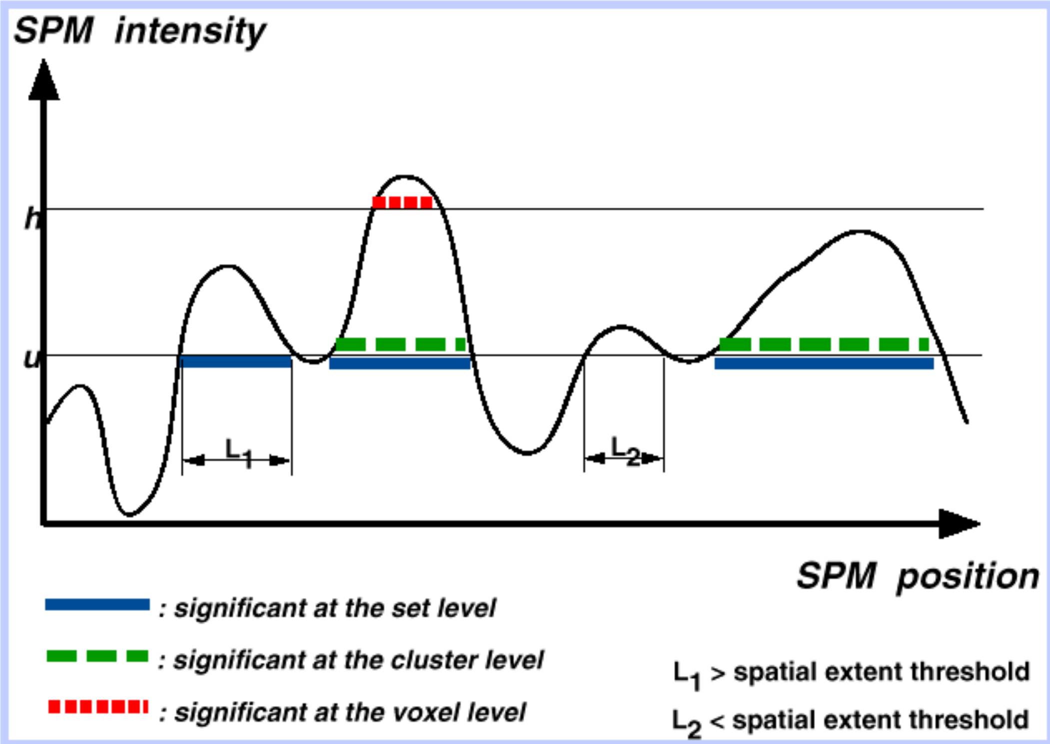
Spatial extent: similar

Voxel, cluster and set level tests

e

u

h



Statistics: *p-values adjusted for search volume*

set-level		cluster-level				peak-level					mm mm mm
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	
0.000	**16**	0.000	**0.000**	**138**	**0.000**	**0.000**	**0.000**	**11.04**	**7.64**	**0.000**	-34 **-70** -28
						0.000	0.009	7.31	5.90	0.000	-44 -74 -24
		0.000	**0.000**	**452**	**0.000**	**0.000**	**0.000**	**9.82**	**7.14**	**0.000**	**6 16 40**
		0.000	**0.000**	300	**0.000**	**0.000**	**0.000**	**9.14**	**6.84**	**0.000**	**44 16 0**
						0.041	0.833	5.29	4.64	0.000	38 12 16
		0.000	**0.000**	**173**	**0.000**	**0.000**	**0.009**	**7.39**	**5.95**	**0.000**	**44 -58 -28**
						0.000	0.009	7.35	5.93	0.000	52 -58 -20
						0.002	0.087	6.42	5.38	0.000	50 -66 -24
		0.000	**0.000**	**112**	**0.000**	**0.000**	**0.025**	**6.93**	**5.69**	**0.000**	**-2 -66 -24**
						0.012	0.418	5.73	4.94	0.000	4 -78 -24
						0.014	0.472	5.65	4.89	0.000	2 -86 -28
		0.013	**0.374**	**3**	**0.257**	**0.010**	**0.406**	**5.77**	**4.97**	**0.000**	-52 **20 4**
		0.000	**0.019**	**20**	**0.008**	**0.011**	**0.406**	**5.76**	**4.96**	**0.000**	**10 -10 8**
		0.008	**0.263**	**5**	**0.148**	**0.016**	**0.472**	**5.63**	**4.87**	**0.000**	**-8 -16 12**
		0.000	**0.012**	**24**	**0.004**	**0.016**	**0.472**	**5.61**	**4.86**	**0.000**	**44 4 28**
						0.035	0.736	5.34	4.68	0.000	46 6 20
		0.006	**0.231**	**6**	**0.116**	**0.018**	**0.472**	**5.59**	**4.84**	**0.000**	**-6 -48 -16**
		0.026	**0.520**	**1**	**0.520**	**0.021**	**0.538**	**5.52**	**4.80**	**0.000**	**-6 -54** -16
		0.026	**0.520**	**1**	**0.520**	**0.030**	**0.713**	**5.40**	**4.72**	**0.000**	**6 -84** -28

table shows 3 local maxima more than 8.0mm apart

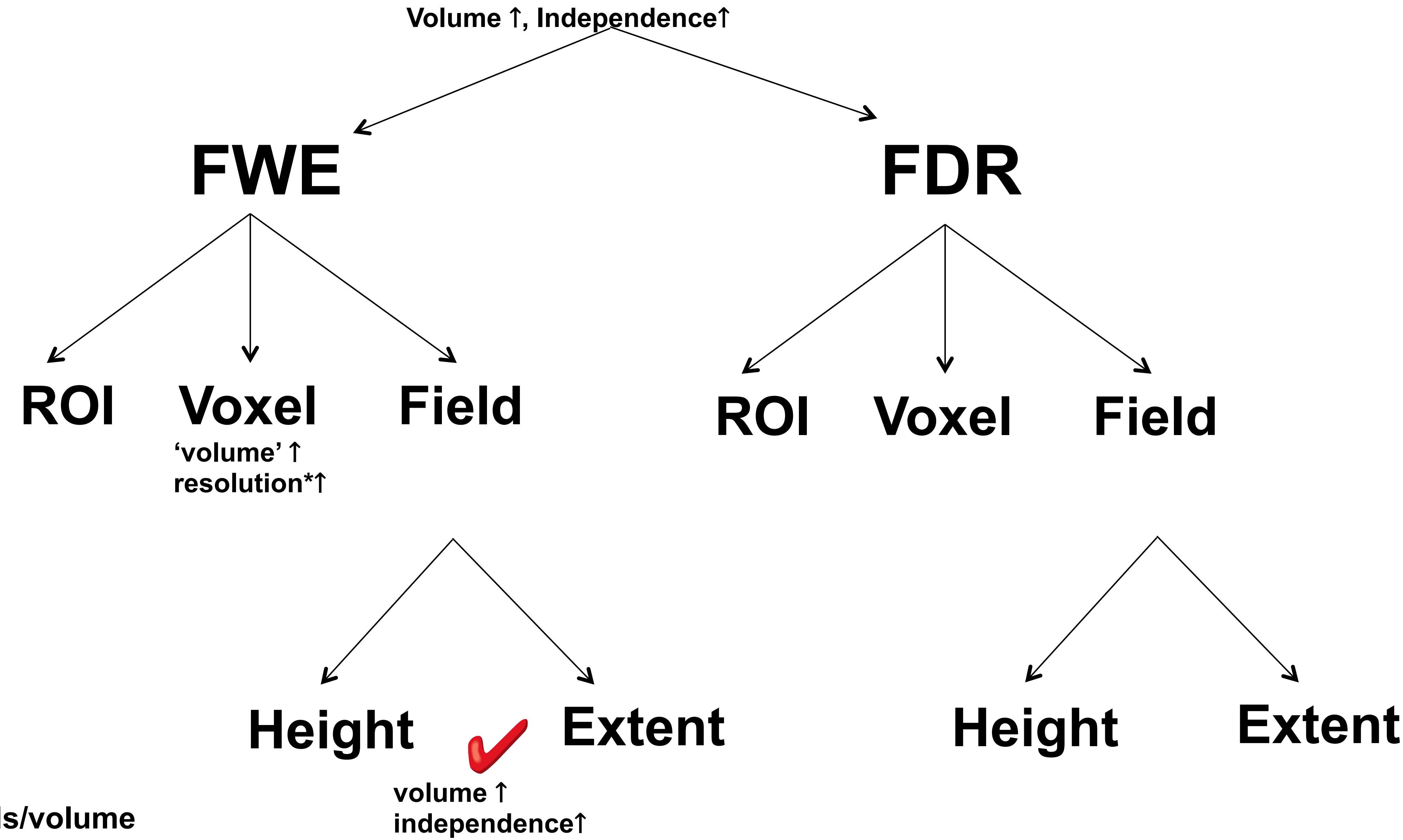
Height threshold: T = 5.21, p = 0.000 (0.050)
Extent threshold: k = 0 voxels, p = 1.000 (0.050)
Expected voxels per cluster, <k> = 2.519
Expected number of clusters, <c> = 0.05
FWEp: 5.213, FDRp: 6.702, FWEc: 1, FDRc: 20

Degrees of freedom = [1.0, 45.0]
FWHM = 9.8 10.6 15.6 mm mm mm; 4.9 5.3 3.9 {voxels}
Volume: 880432 = 55027 voxels = 472.2 resels
Voxel size: 2.0 2.0 4.0 mm mm mm; (resel = 102.26 voxels)
Page 1

Detect an effect of *unknown* extent & location

There is a multiple testing problem ('voxel' or 'blob' perspective).

More corrections needed as ..



*voxels/volume

Further reading

- Friston KJ, Frith CD, Liddle PF, Frackowiak RS. Comparing functional (PET) images: the assessment of significant change. J Cereb Blood Flow Metab. 1991 Jul;11(4):690-9.
- Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage. 2002 Apr;15(4):870-8.
- Worsley KJ Marrett S Neelin P Vandal AC Friston KJ Evans AC. A unified statistical approach for determining significant signals in images of cerebral activation. Human Brain Mapping 1996;4:58-73.