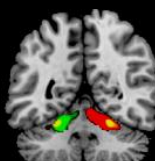
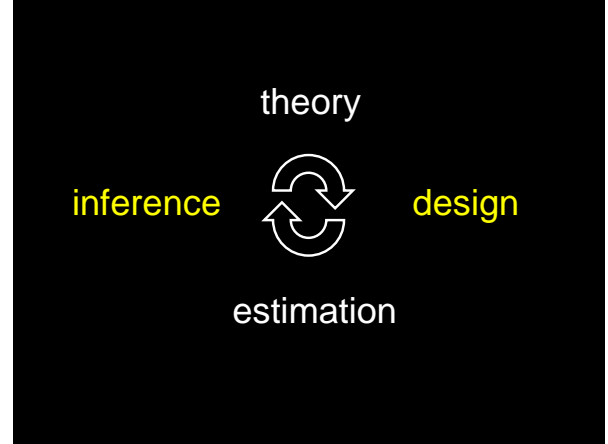


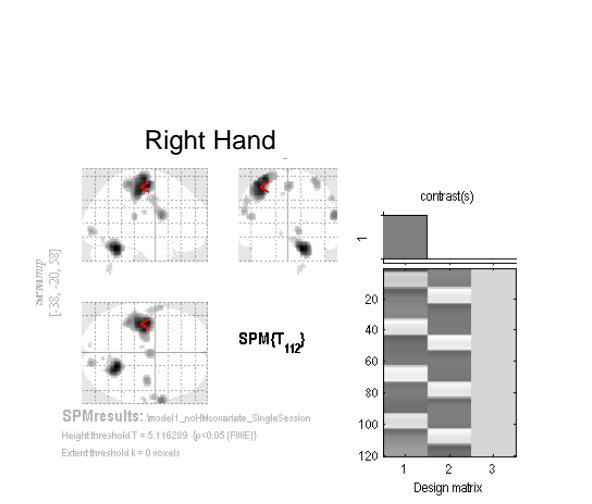
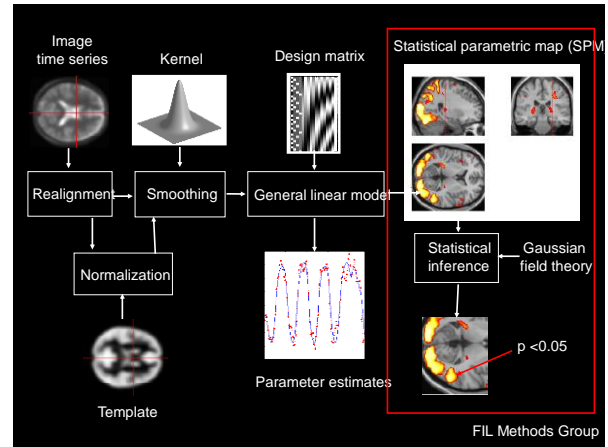
SPM8 for Basic and Clinical Investigators

Inference

Inferential Statistics

- Infer something about the population from which a sample is drawn
- Hypothesis testing
- Significance - p-value

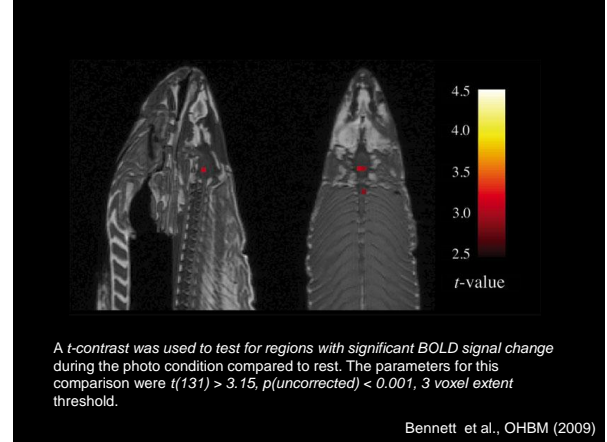


Statistics: p-values adjusted for search volume

set-level	cluster-level				peak-level				mm mm mm				
	p	c	$\rho_{\text{ROI}}^{\text{Bonf}}$	$\rho_{\text{ROI}}^{\text{FDR}}$	k_E	ρ_{unort}	$\rho_{\text{ROI}}^{\text{Bonf}}$	$\rho_{\text{ROI}}^{\text{FDR}}$		T			
0.00017	0.000	0.000	0.000	751	0.000	0.000	0.000	13.90	Inf	0.000	18	-52	-24
	0.000	0.000	0.000	2145	0.000	0.000	0.000	13.02	Inf	0.000	6	-62	-12
								12.62	Inf	0.000	-32	-14	70
								12.56	Inf	0.000	-26	-20	54
				159	0.000	0.000	0.000	9.09	7.84	0.000	64	8	15
				165	0.000	0.000	0.000	7.88	7.01	0.000	2	-6	60
				170	0.000	0.000	0.000	7.62	6.82	0.000	36	-12	66
				91	0.000	0.000	0.000	7.26	6.56	0.000	16	-88	-2
				56	0.000	0.000	0.000	7.15	6.45	0.000	-52	-28	15
				46	0.001	0.001	0.000	7.07	6.42	0.000	62	-2	42
				163	0.000	0.001	0.000	6.95	6.32	0.000	-62	8	26
								6.46	5.94	0.000	-58	-2	38
								5.97	5.55	0.000	-58	-8	44
				21	0.014	0.002	0.070	5.90	5.49	0.000	64	-14	24
				12	0.052	0.003	0.089	5.83	5.43	0.000	8	-72	-12
				32	0.003	0.005	0.124	5.73	5.36	0.000	18	-62	-48
								5.31	5.00	0.000	14	-64	-56

table shows 3 local maxima more than 8.0mm apart

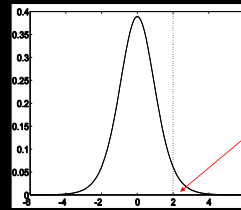
Height threshold: T = 5.12, p = 0.000 (0.050) Degrees of freedom = [10, 112.0]
Extent threshold: k = 0 voxels, p = 1.000 (0.050) FWHM = 10.0 9.9 9.9 mm mm mm, 5.0 4.9 4.9 (voxels)
Expedited voxels per cluster, $\rho_c = 3.126$ Volume: 190510 = 237929 voxels = 1832.0 resels
Expedited number of clusters, $\rho_c > 0.05$ Voxel size: 2.0 2.0 2.0 mm mm mm, (resel = 121.32 voxels)
FWEp: 5.116, FDRp: 6.263, FWEc: 1, FDRc: 21 Page 1



Critical Threshold

t-statistic	Z-score	P-value
1.00	1.64	.05
2.00	2.33	.01
3.00	3.11	.001

Inference at a single voxel



NULL hypothesis
H: activation is zero

$$\alpha = p(t > u | H)$$

p-value: probability of getting a value of *t* at least as extreme as *u*. If α is small we reject the null hypothesis.

Penny

False Positives and Negatives

	Null is True	Null is False
Reject Null	False Positive	Correct Decision
Do not reject Null	Correct Decision	False Negative

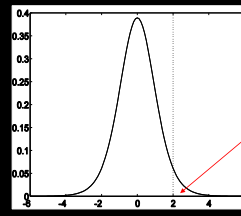
Type I and II Errors

	Null is True	Null is False
Reject Null	Type I Error	Correct Decision
Do not reject Null	Correct Decision	Type II Error

Alpha and Beta

	Null is True	Null is False
Reject Null	α	$1 - \beta = \text{power}$
Do not reject Null	$1 - \alpha$	β

Inference at a single voxel



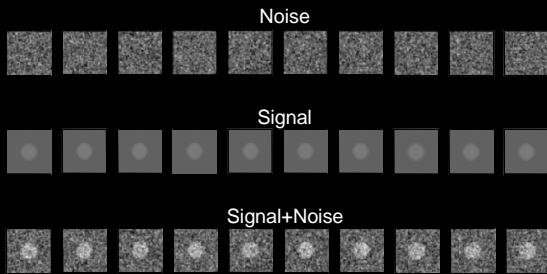
NULL hypothesis
H: activation is zero

$$a = p(t > u | H)$$

p-value: probability of getting a value of t at least as extreme as u. If a is small we reject the null hypothesis.

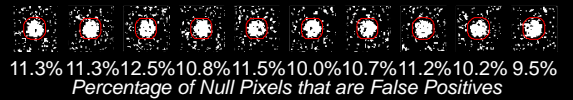
Penny

Inference for Images



Penny

Use of 'uncorrected' p-value, $a=0.1$



Using an 'uncorrected' p-value of 0.1 will lead us to conclude on average that 10% of voxels are active when they are not.

To correct for this we can define a null hypothesis for images of statistics.

Penny

Family-wise Null Hypothesis

- Family of hypotheses
 - H^k $k \in \{1, \dots, K\}$
 - $H^0 = H^1 \cap H^2 \dots \cap H^k \cap H^K$

FAMILY-WISE NULL HYPOTHESIS:
Activation is zero everywhere

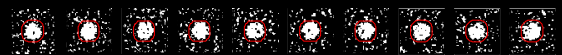
If we reject a voxel null hypothesis at any voxel, we reject the family-wise Null hypothesis

A FP anywhere gives a Family Wise Error (FWE)

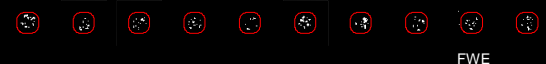
Family-wise error rate = 'corrected' p-value

Penny

Use of 'uncorrected' p-value, $a=0.1$



Use of 'corrected' p-value, $a=0.1$



Penny

Critical Threshold Determination

- no adjustment for multiple comparisons
- Family-Wise Error (FWE)
 - Bonferroni correction
 - resolvable element correction (RESELS)
 - brain volume correction
 - small volume correction (SVC)
- False Discovery Rate (FDR)

Bonferroni Correction

$$P_{\text{corrected}} = \text{TESTS} * \frac{P_{\text{uncorrected}}}{N}$$

The Bonferroni correction

Given a family of N independent voxels and a voxel-wise error rate v - the Family-Wise Error rate (FWE) or 'corrected' error rate is

$$\alpha = 1 - (1-v)^N$$

$$\sim Nv$$

If $v=0.05$ then over 100 voxels we'll get 5 voxel-wise type I errors. But we'll get a much higher α . To ensure $\alpha=0.05$ we need $v=0.0005!$

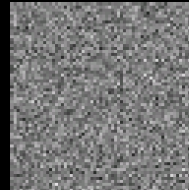
Therefore, to ensure a particular FWE we choose $v = \alpha / N$

A Bonferroni correction is appropriate for independent tests

A correction for multiple comparisons

Bonferroni correction

Independent Voxels

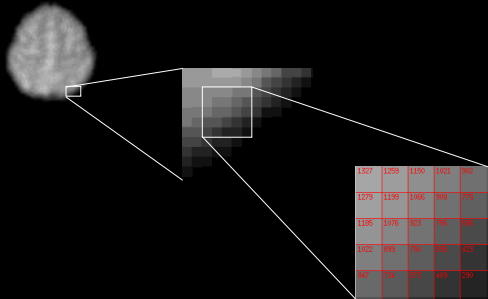


Spatially Correlated Voxels



Bonferroni is too conservative for brain images

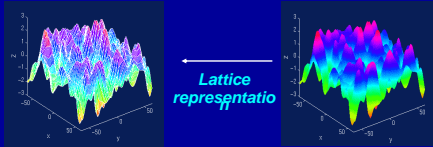
Image Smoothness



A star-forming region, N11B, in the Large Magellanic Cloud

Random Field Theory

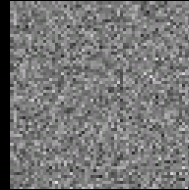
- Consider a statistic image as a lattice representation of a continuous random field
- Use results from continuous random field theory



FIL Methods Group

RESEL correction

Independent Voxels



Spatially Correlated Voxels



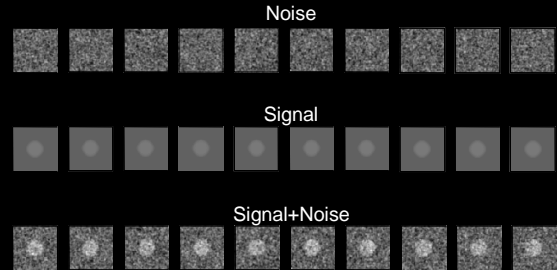
RESEL correction is more appropriate for brain images

Resolution Elements (RESELS)

$$RESELS = V / FWHM^3$$

- V - volume of the search region
- FWHM – the effective full-width at half maximum of the Gaussian smoothing kernel

False Discovery Rate

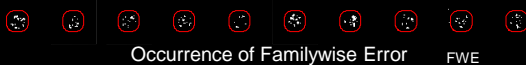


Penny

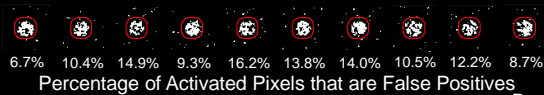
Control of Per Comparison Rate at 10%



Control of Familywise Error Rate at 10%



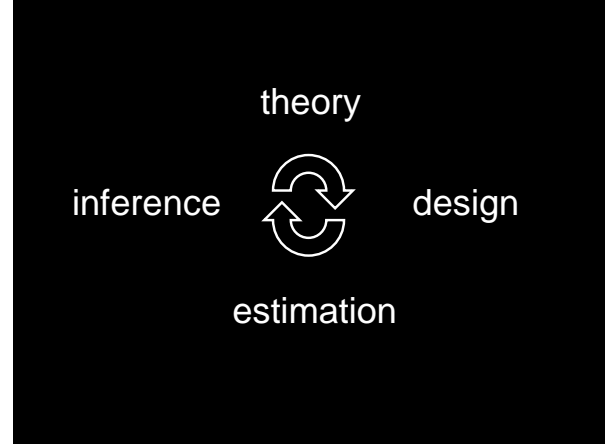
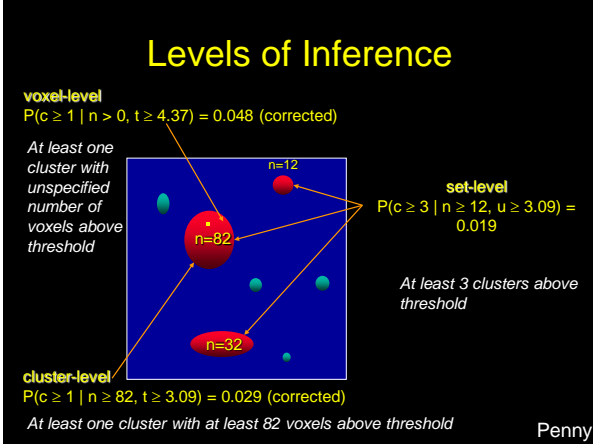
Control of False Discovery Rate at 10%



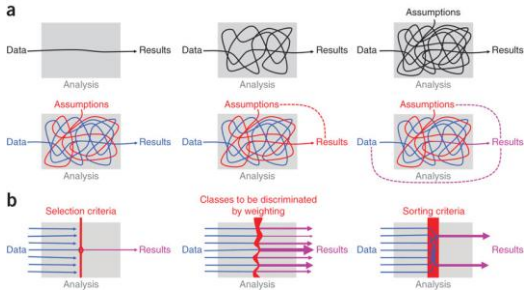
Penny

Critical Threshold Determination

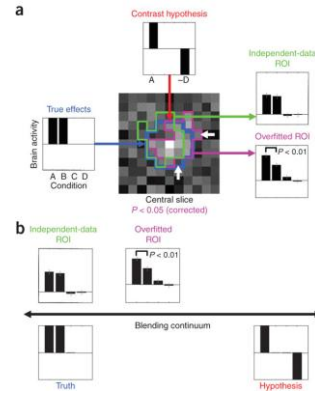
- no correction for multiple comparisons
- Bonferroni correction
 $p(\text{corrected}) = p(\text{uncorrected}) * \text{number of tests}$
- resolvable element correction (RESELS)
 $p(\text{corrected}) = p(\text{uncorrected}) * \text{resolvable elements}$
- False discovery rate



Selection Bias in Functional Neuroimaging



Kriegeskorte et al., Nature Neuroscience (2009)



Kriegeskorte et al., Nature Neuroscience (2009)