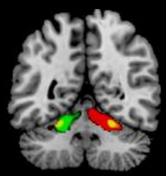
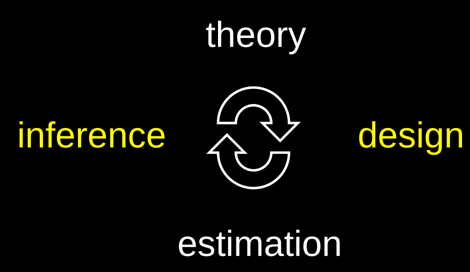
SPM8 for Basic and Clinical Investigators

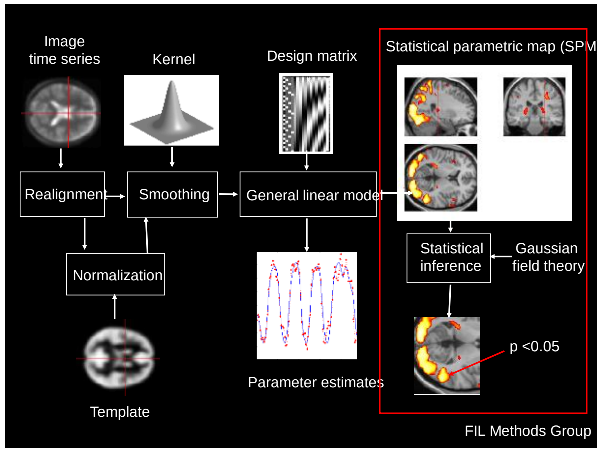
Inference



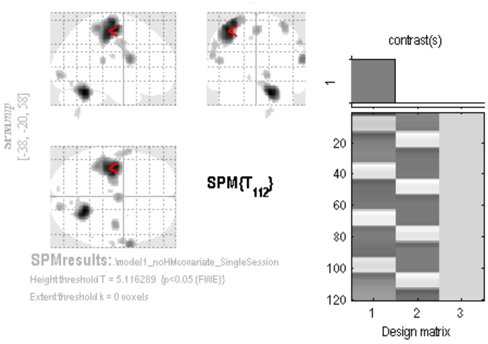


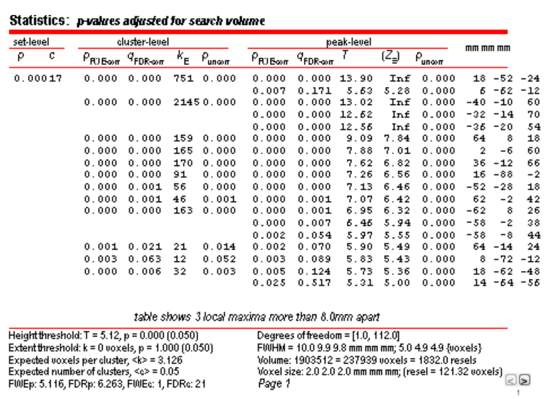
Inferential Statistics

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



Right Hand





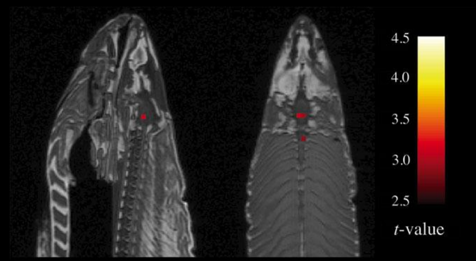
Statistics: ***p-values adjusted for search volume***

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	17	0.000	0.000	751	0.000	0.000	0.000	13.90	Inf	0.000	18	-52	-24
						0.007	0.171	5.63	5.28	0.000	6	-62	-12
		0.000	0.000	2145	0.000	0.000	0.000	13.02	Inf	0.000	-40	-10	60
						0.000	0.000	12.62	Inf	0.000	-32	-14	70
						0.000	0.000	12.56	Inf	0.000	-36	-20	54
		0.000	0.000	159	0.000	0.000	0.000	9.09	7.84	0.000	64	8	18
		0.000	0.000	165	0.000	0.000	0.000	7.88	7.01	0.000	2	-6	60
		0.000	0.000	170	0.000	0.000	0.000	7.62	6.82	0.000	36	-12	66
		0.000	0.000	91	0.000	0.000	0.000	7.26	6.56	0.000	16	-88	-2
		0.000	0.001	56	0.000	0.000	0.000	7.13	6.46	0.000	-52	-28	18
		0.000	0.001	46	0.001	0.000	0.001	7.07	6.42	0.000	62	-2	42
		0.000	0.000	163	0.000	0.000	0.001	6.95	6.32	0.000	-62	8	26
						0.000	0.007	6.46	5.94	0.000	-58	-2	38
						0.002	0.054	5.97	5.55	0.000	-58	-8	44
		0.001	0.021	21	0.014	0.002	0.070	5.90	5.49	0.000	64	-14	24
		0.003	0.063	12	0.052	0.003	0.089	5.83	5.43	0.000	8	-72	-12
		0.000	0.006	32	0.003	0.005	0.124	5.73	5.36	0.000	18	-62	-48
						0.025	0.517	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)
Extent threshold: k = 0 voxels, p = 1.000 (0.050)
Expected voxels per cluster, <k> = 3.126
Expected number of clusters, <c> = 0.05
FWEp: 5.116, FDRp: 6.263, FWEc: 1, FDRc: 21

Degrees of freedom = [1.0, 112.0]
FWHM = 10.0 9.9 9.8 mm mm mm; 5.0 4.9 4.9 {voxels}
Volume: 1903512 = 237939 voxels = 1832.0 resels
Voxel size: 2.0 2.0 2.0 mm mm mm; (resel = 121.32 voxels)
Page 1



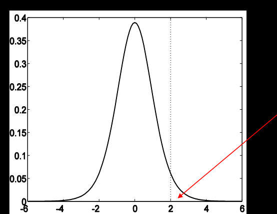
A *t-contrast was used to test for regions with significant BOLD signal change* during the photo condition compared to rest. The parameters for this comparison were *t(131) > 3.15, p(uncorrected) < 0.001, 3 voxel extent* threshold.

Bennett et al., OHBM (2009)

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



u=2

t-distribution

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

False Positives and Negatives

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

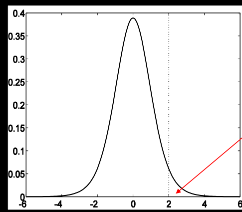
Type I and II Errors

	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 - β = power
Do not reject Null	1 - α	β

Inference at a single voxel



u=2

t-distribution

NULL hypothesis
H: activation is zero

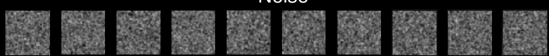
a = p(*t*>u|H)

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

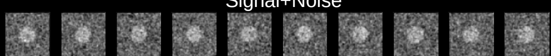
Inference for Images

Noise



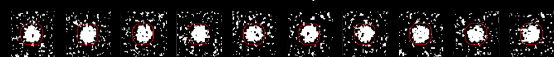
Signal

Signal+Noise



Penny

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



11.3% 11.3% 12.5% 10.8% 11.5% 10.0% 10.7% 11.2% 10.2% 9.5%

Percentage of Null Pixels that are False Positives

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 □ □ = {1,…,K}
  - $H^□$ = H^1 □ H^2 … □ H^k □ H^K

FAMILY-WISE NULL HYPOTHESIS:
Activation is zero everywhere

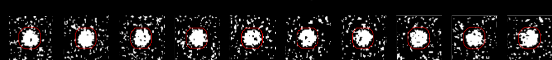
If we reject a voxel null hypothesis at *any* voxel, we reject the family-wise Null hypothesis

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

FWE

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

$$P_{corrected} = TESTS * P_{uncorrected}$$

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 α=0.05 we need v=0.0005 !

Therefore, to ensure a particular FWE we choose

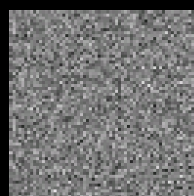
$$v = \alpha / N$$

A Bonferroni correction is appropriate for independent tests

A correction for multiple comparisons

Bonferroni correction

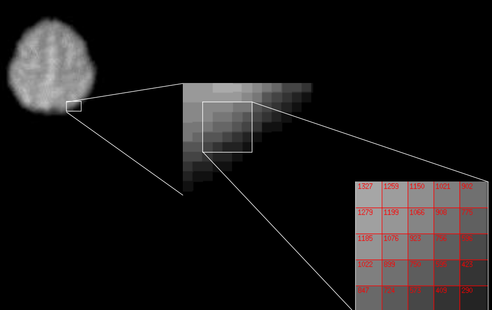
Independent Voxels



Spatially Correlated Voxels

Bonferroni is too conservative for brain images

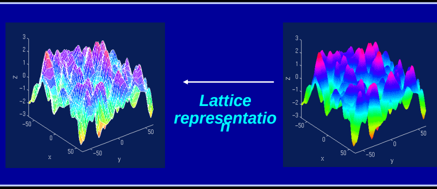
Image Smoothness



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

Random Field Theory

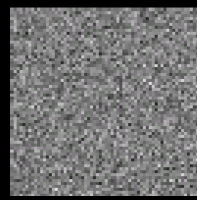
- Consider a statistic image as a lattice representation of a continuous random field
- Use results from continuous random field theory



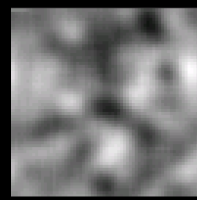
FIL Methods Group

RESEL correction

Independent Voxels



Spatially Correlated Voxels



RESEL correction is more appropriate for brain images

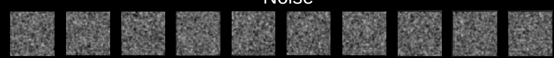
Resolution Elements (RESELS)

RESELS= V / $FWHM^3$

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

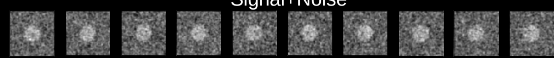
False Discovery Rate

Noise



Signal

Signal+Noise



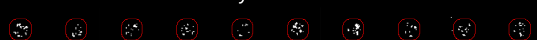
Penny

Control of Per Comparison Rate at 10%

11.3% 11.3% 12.5% 10.8% 11.5% 10.0% 10.7% 11.2% 10.2% 9.5%

Percentage of Null Pixels that are False Positives

Control of Familywise Error Rate at 10%



Occurrence of Familywise Error FWE

Control of False Discovery Rate at 10%

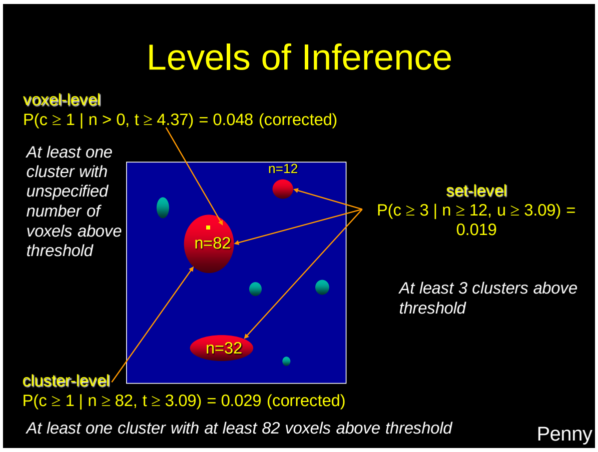
6.7% 10.4% 14.9% 9.3% 16.2% 13.8% 14.0% 10.5% 12.2% 8.7%

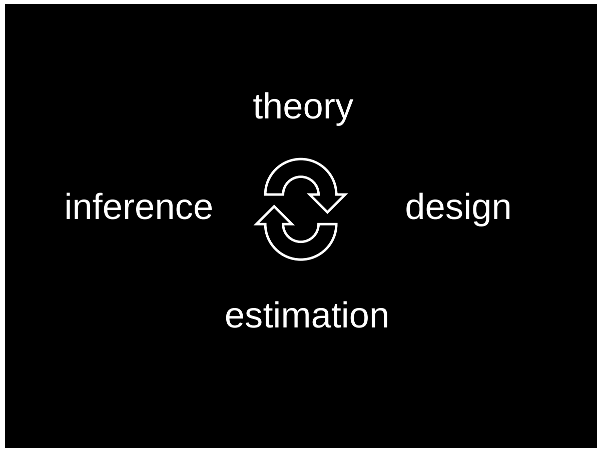
Percentage of Activated Pixels that are False Positives

Penny

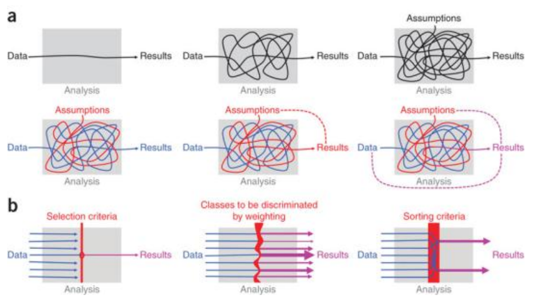
Critical Threshold Determination

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

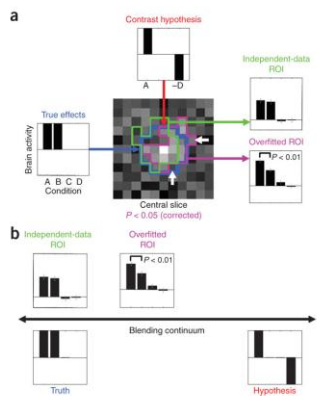




Selection Bias in Functional Neuroimaging



Kriegeskorte et al., Nature Neuroscience (2009)



Kriegeskorte et al., Nature Neuroscience (2009)