Nonparametric Inference with SnPM

Thomas Nichols, Ph.D.
Department of Statistics
Warwick Manufacturing Group
University of Warwick

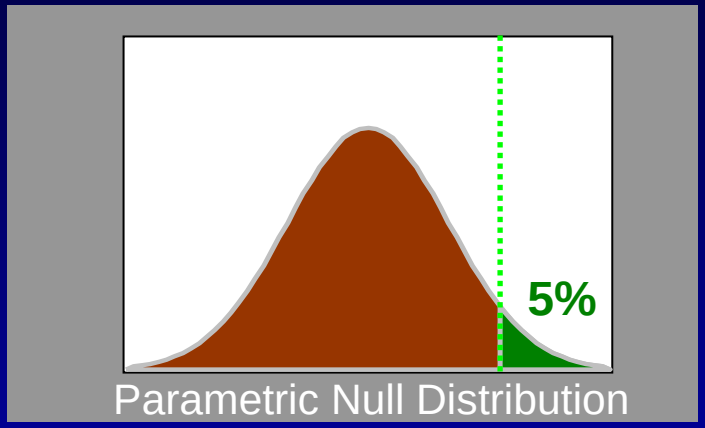
Zurich SPM Course
February 16, 2012

Outline

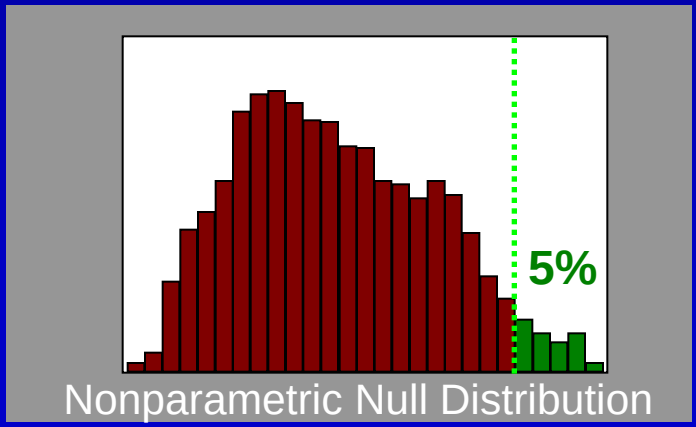
- Permutation test – one voxel
- Permutation test – Controlling FWE
- SnPM with small groups
- SnPM stability

Nonparametric Inference

- Parametric methods
  - Assume distribution of statistic under null hypothesis
  - Needed to find P-values, u_α



- Nonparametric methods
  - Use *data* to find distribution of statistic under null hypothesis
  - Any statistic!



Permutation Test
Toy Example

- Data from V1 voxel in visual stim. experiment

  A: Active, flashing checkerboard B: Baseline, fixation

  6 blocks, ABABAB Just consider block averages...

A	B	A	B	A	B
103.00	90.48	99.93	87.83	99.76	96.06

- Null hypothesis H_o
  - No experimental effect, A & B labels arbitrary
- Statistic
  - Mean difference

Permutation Test Toy Example

- Under H_o
  - Consider all equivalent relabelings

AAABBB	ABABAB	BAAABB	BABBAA
AABABB	ABABBA	BAABAB	BBAAAB
AABBAB	ABBAAB	BAABBA	BBAABA
AABBBA	ABBABA	BABAAB	BBABAA
ABAABB	ABBBAA	BABABA	BBBAAA

Permutation Test
Toy Example

- Under H_o
  - Consider all equivalent relabelings
  - Compute all possible statistic values

AAABBB 4.82	ABABAB 9.45	BAAABB -1.48	BABBAA -6.86
AABABB -3.25	ABABBA 6.97	BAABAB 1.10	BBAAAB 3.15
AABBAB -0.67	ABBAAB 1.38	BAABBA -1.38	BBAABA 0.67
AABBBA -3.15	ABBABA -1.10	BABAAB -6.97	BBABAA 3.25
ABAABB 6.86	ABBBAA 1.48	BABABA -9.45	BBBAAA -4.82

Permutation Test Toy Example

- Under H_o
  - Consider all equivalent relabelings
  - Compute all possible statistic values
  - Find 95%ile of permutation distribution

AAABBB 4.82	ABABAB 9.45	BAAABB -1.48	BABBAA -6.86
AABABB -3.25	ABABBA 6.97	BAABAB 1.10	BBAAAB 3.15
AABBAB -0.67	ABBAAB 1.38	BAABBA -1.38	BBAABA 0.67
AABBBA -3.15	ABBABA -1.10	BABAAB -6.97	BBABAA 3.25
ABAABB 6.86	ABBBAA 1.48	BABABA -9.45	BBBAAA -4.82

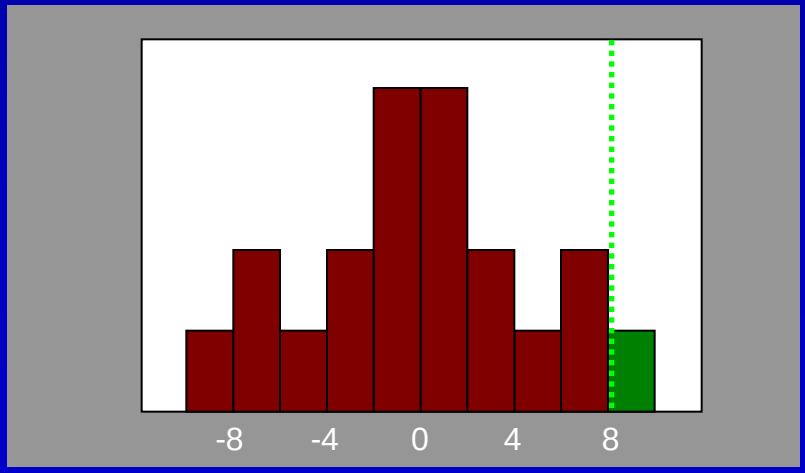
Permutation Test Toy Example

- Under H_o
  - Consider all equivalent relabelings
  - Compute all possible statistic values
  - Find 95%ile of permutation distribution

AAABBB 4.82	ABABAB 9.45	BAAABB -1.48	BABBAA -6.86
AABABB -3.25	ABABBA 6.97	BAABAB 1.10	BBAAAB 3.15
AABBAB -0.67	ABBAAB 1.38	BAABBA -1.38	BBAABA 0.67
AABBBA -3.15	ABBABA -1.10	BABAAB -6.97	BBABAA 3.25
ABAABB 6.86	ABBBAA 1.48	BABABA -9.45	BBBAAA -4.82

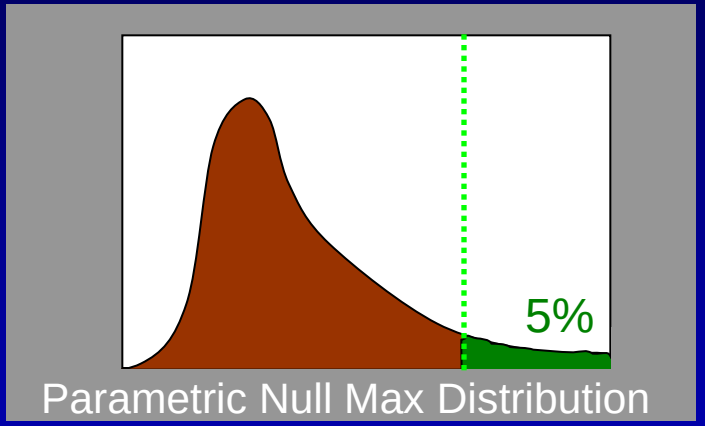
Permutation Test Toy Example

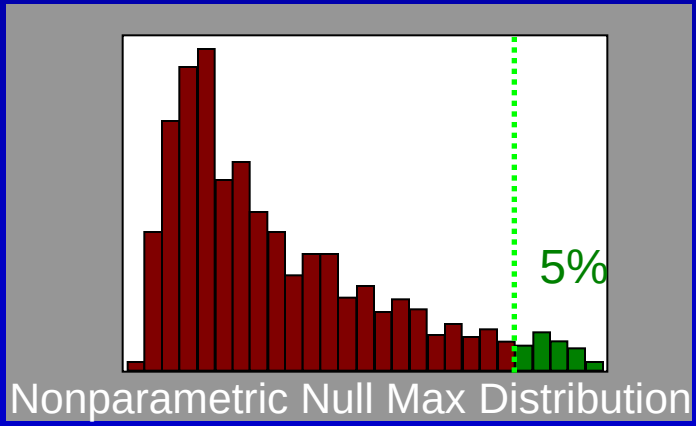
- Under H_o
  - Consider all equivalent relabelings
  - Compute all possible statistic values
  - Find 95%ile of permutation distribution



Controlling FWE: Permutation Test

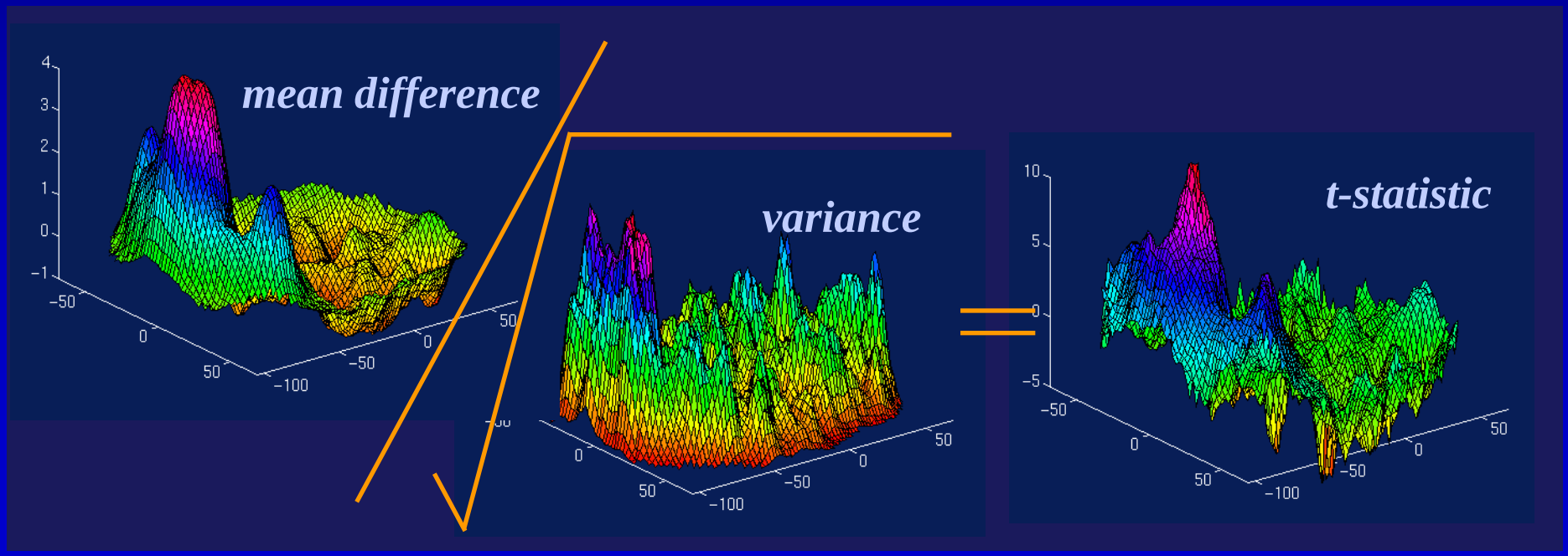
- Parametric methods
  - Assume distribution of *max* statistic under null hypothesis
- Nonparametric methods
  - Use *data* to find distribution of *max* statistic under null hypothesis
  - Again, any max statistic!





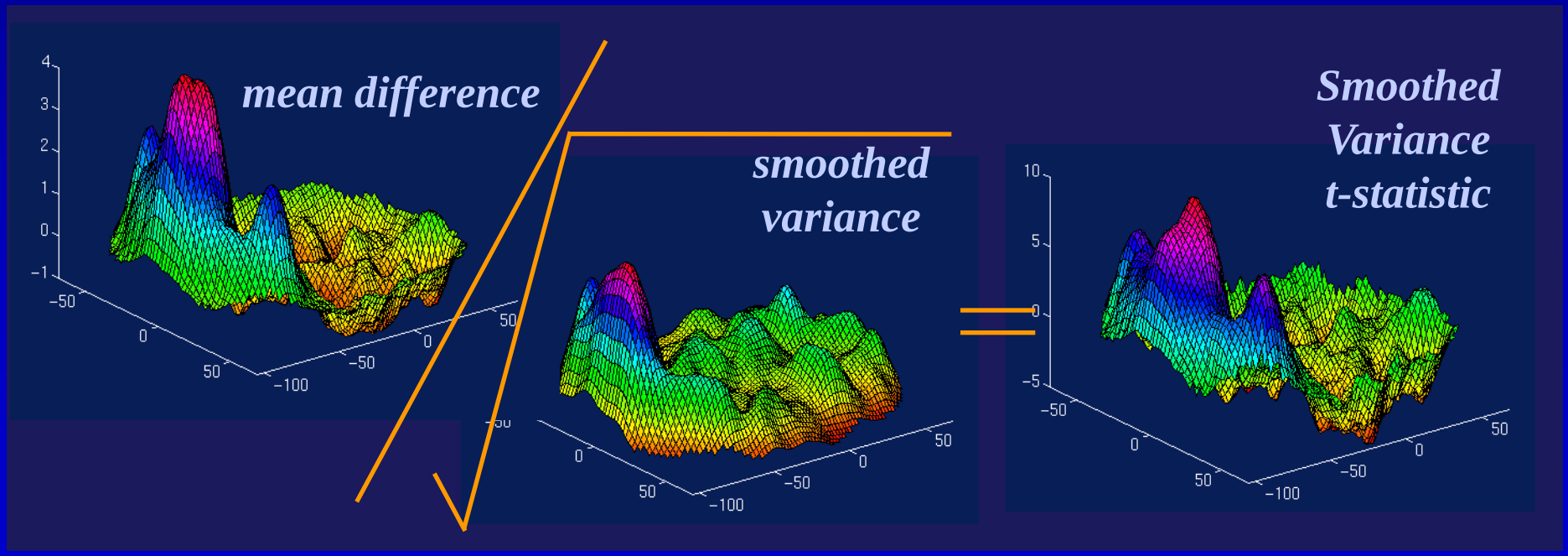
Permutation Test Smoothed Variance *t*

- Collect max distribution
  - To find threshold that controls FWE
- Consider smoothed variance *t* statistic



Permutation Test Smoothed Variance *t*

- Collect max distribution
  - To find threshold that controls FWE
- Consider smoothed variance *t* statistic



Permutation Test Strengths

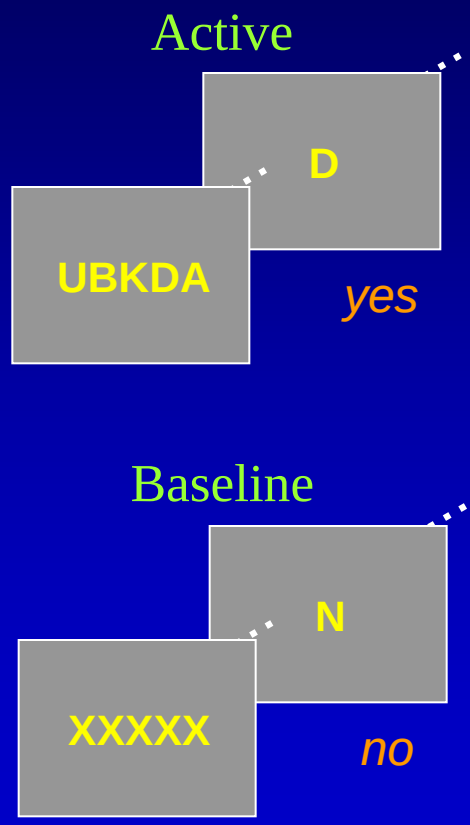
- Requires only assumption of exchangeability
  - Under Ho, distribution unperturbed by permutation
  - Allows us to build permutation distribution
- Subjects are exchangeable
  - Under Ho, each subject's A/B labels can be flipped
- fMRI scans not exchangeable under Ho
  - Due to temporal autocorrelation

Permutation Test Limitations

- Computational Intensity
  - Analysis repeated for each relabeling
  - Not so bad on modern hardware
    - No analysis discussed below took more than 3 hours
- Implementation Generality
  - Each experimental design type needs unique code to generate permutations
    - Not so bad for population inference with t-tests

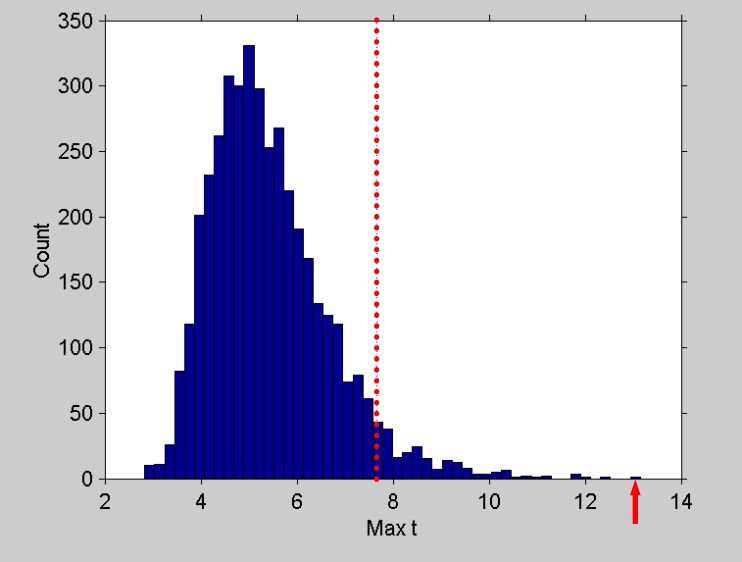
Permutation Test Example

- fMRI Study of Working Memory
  - 12 subjects, block design Marshuetz et al (2000)
  - Item Recognition
    - Active:View five letters, 2s pause, view probe letter, respond
    - Baseline: View XXXXX, 2s pause, view Y or N, respond
- Second Level RFX
  - Difference image, A-B constructed for each subject
  - One sample, smoothed variance *t* test

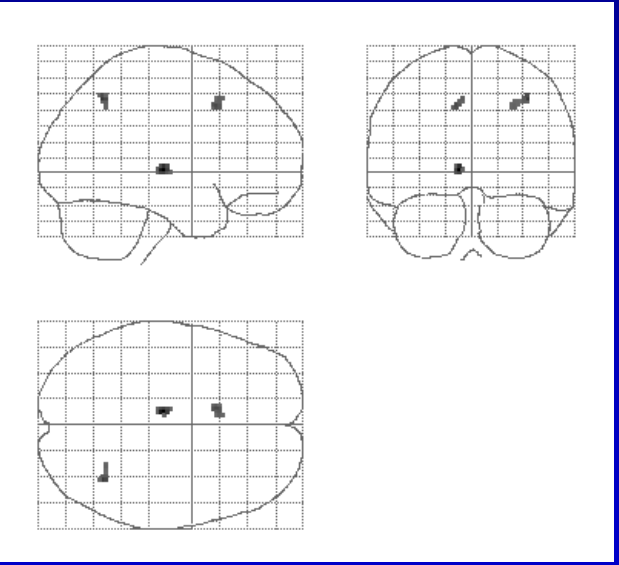


Permutation Test Example

- Permute!
  - 2^{12} = 4,096 ways to flip 12 A/B labels
  - For each, note maximum of *t* image



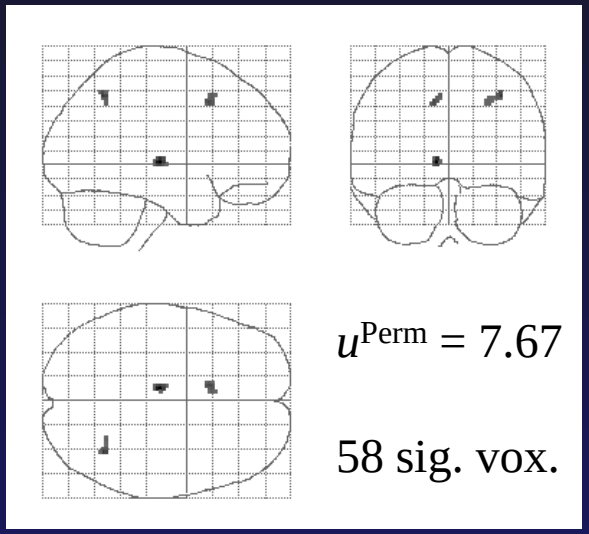
Permutation Distribution
Maximum *t*



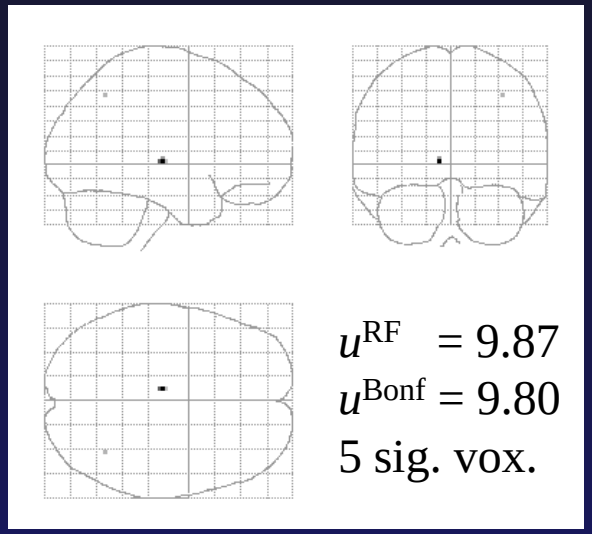
Maximum Intensity Projection
Thresholded *t*

Permutation Test Example

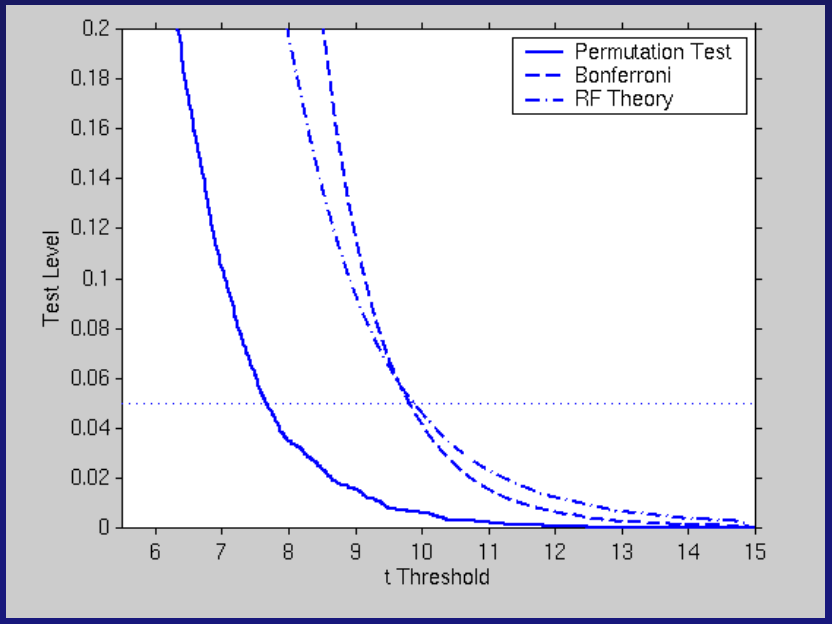
- Compare with Bonferroni
  - $\alpha = 0.05/110{,}776$
- Compare with parametric RFT
  - 110,776 2×2×2mm voxels
  - 5.1×5.8×6.9mm FWHM smoothness
  - 462.9 RESELs



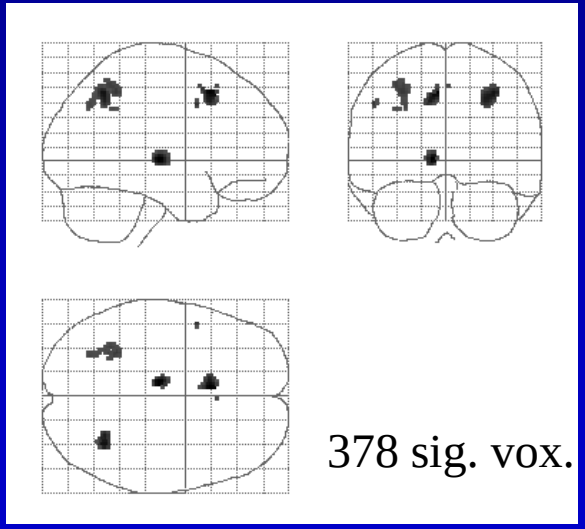
t_{11} Statistic, Nonparametric Threshold



t_{11} Statistic, RF & Bonf. Threshold



Test Level vs. t_{11} Threshold



Smoothed Variance *t* Statistic, Nonparametric Threshold

Does this Generalize? RFT vs Bonf. vs Perm.

	df	*t* Threshold (0.05 Corrected) RF	Bonf	Perm
Verbal Fluency	4	4701.32	42.59	10.14
Location Switching	9	11.17	9.07	5.83
Task Switching	9	10.79	10.35	5.10
Faces: Main Effect	11	10.43	9.07	7.92
Faces: Interaction	11	10.70	9.07	8.26
Item Recognition	11	9.87	9.80	7.67
Visual Motion	11	11.07	8.92	8.40
Emotional Pictures	12	8.48	8.41	7.15
Pain: Warning	22	5.93	6.05	4.99
Pain: Anticipation	22	5.87	6.05	5.05

RFT vs Bonf. vs Perm.

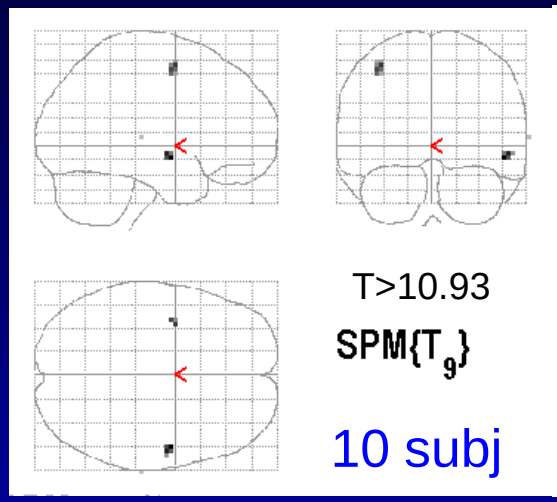
		No. Significant Voxels (0.05 Corrected)			
			t		SmVar *t*
	df	RF	Bonf	Perm	Perm
Verbal Fluency	4	0	0	0	0
Location Switching	9	0	0	158	354
Task Switching	9	4	6	2241	3447
Faces: Main Effect	11	127	371	917	4088
Faces: Interaction	11	0	0	0	0
Item Recognition	11	5	5	58	378
Visual Motion	11	626	1260	1480	4064
Emotional Pictures	12	0	0	0	7
Pain: Warning	22	127	116	221	347
Pain: Anticipation	22	74	55	182	402

Using SnPM to Assess Reliability with Small Groups

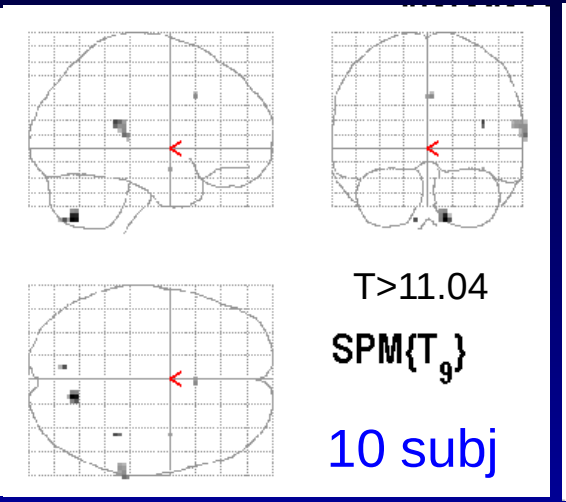
- Consider n=50 group study
  - Event-related Odd-Ball paradigm, Kiehl, et al.
- Analyze all 50
  - Analyze with SPM and SnPM, find FWE thresh.
- Randomly partition into 5 groups 10
  - Analyze each with SPM & SnPM, find FWE thresh
- Compare reliability of small groups with full
  - With and without variance smoothing

Skip

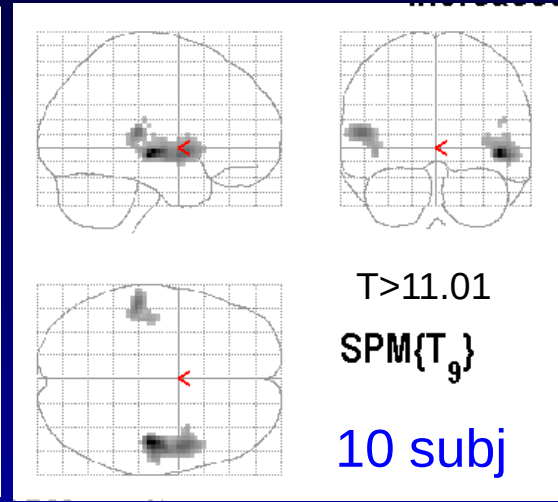
SPM t_{11}: 5 groups of 10 vs all 50
5% FWE Threshold



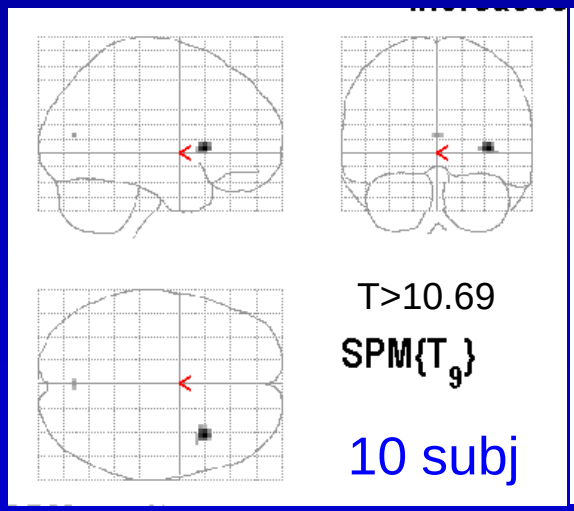
2 8 11 15 18 35 41 43 44 50



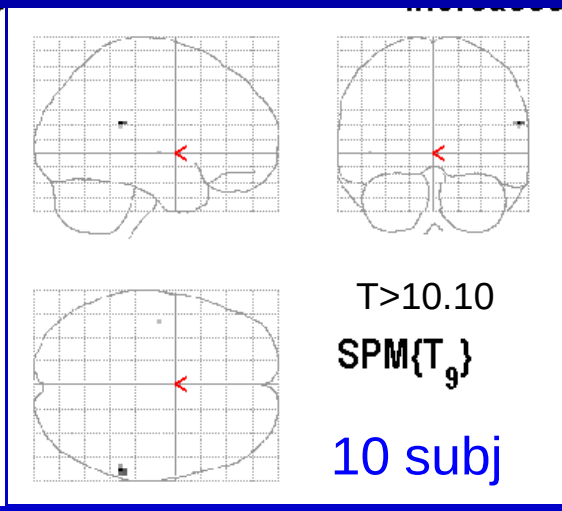
1 3 20 23 24 27 28 32 34 40



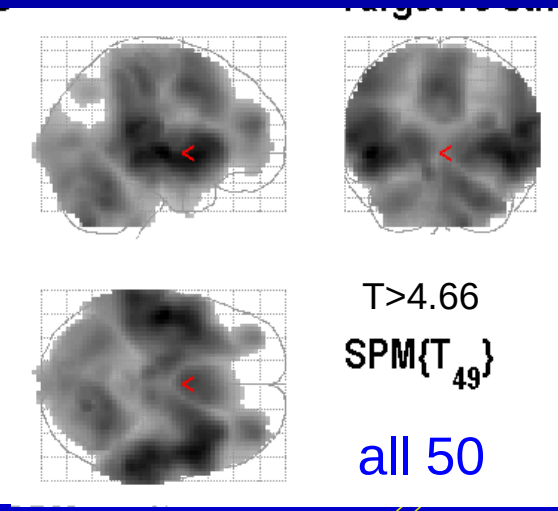
9 13 14 16 19 21 25 29 30 45



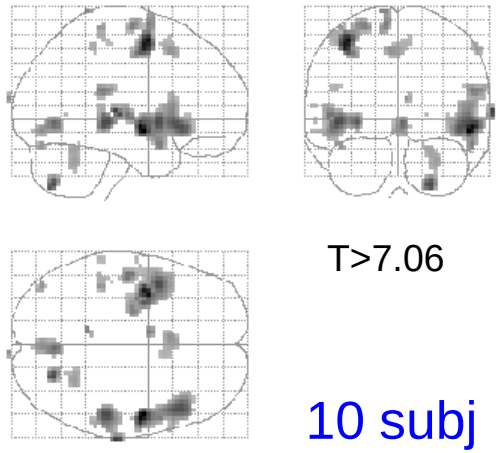
4 5 10 22 31 33 36 39 42 47



6 7 12 17 26 37 38 46 48 49



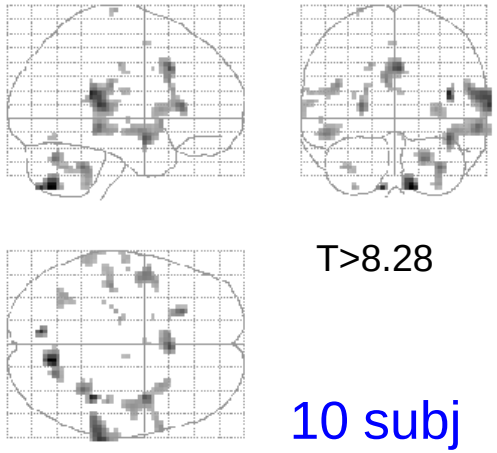
SnPM *t*: 5 groups of 10 vs. all 50
5% FWE Threshold



T>7.06

10 subj

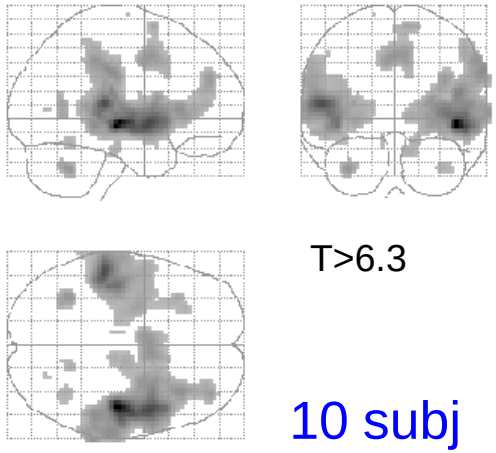
2 8 11 15 18 35 41 43 44 50



T>8.28

10 subj

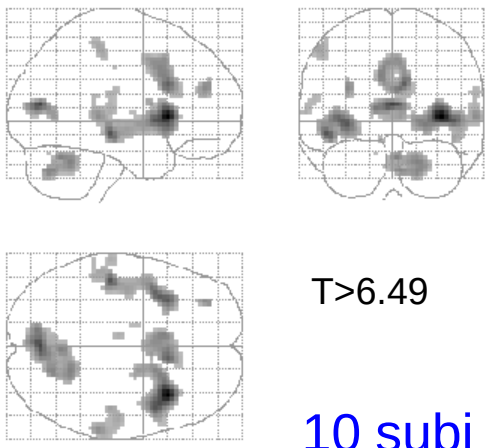
1 3 20 23 24 27 28 32 34 40



T>6.3

10 subj

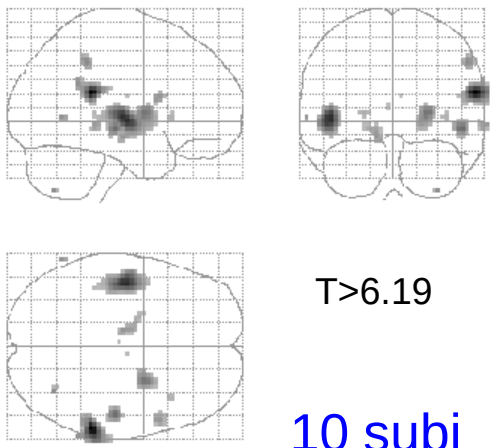
9 13 14 16 19 21 25 29 30 45



T>6.49

10 subj

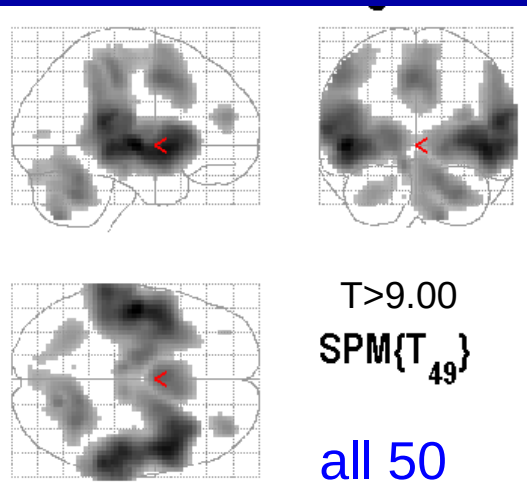
4 5 10 22 31 33 36 39 42 47



T>6.19

10 subj

6 7 12 17 26 37 38 46 48 49



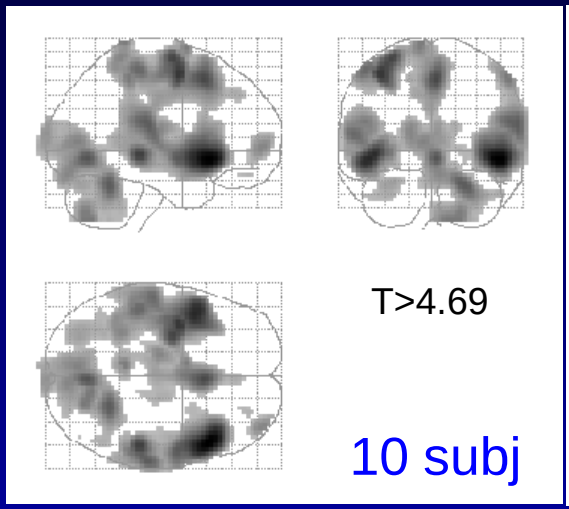
T>9.00

SPM{T_{49}}

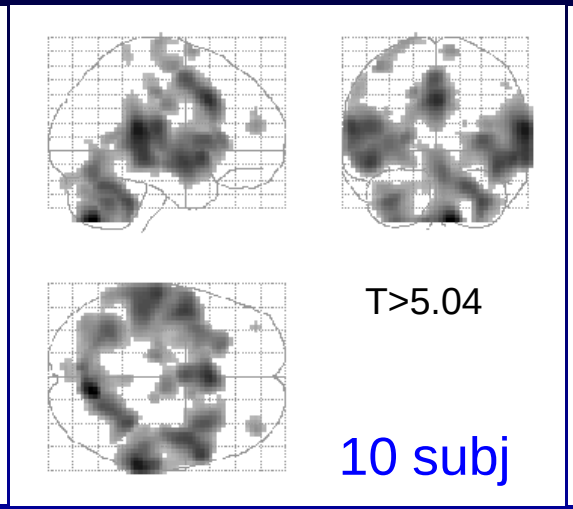
all 50

Arbitrary thresh of 9.0

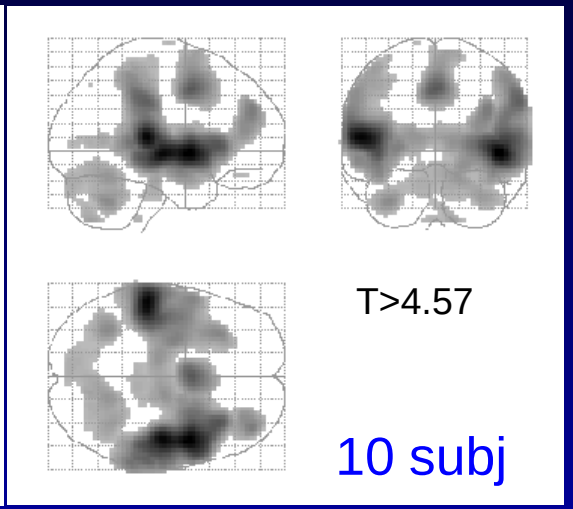
SnPM SmVar *t*: 5 groups of 10 vs. all 50
5% FWE Threshold



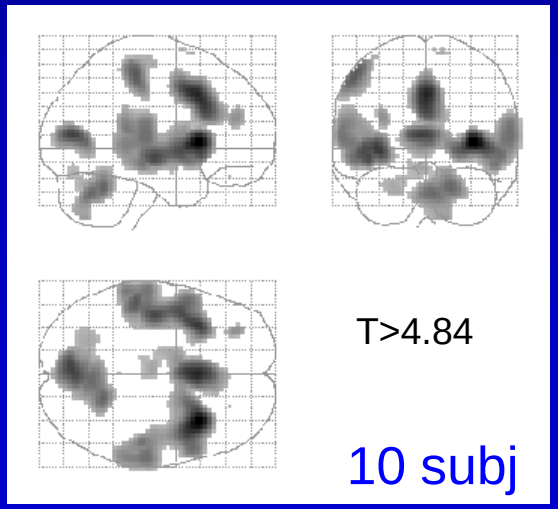
2 8 11 15 18 35 41 43 44 50



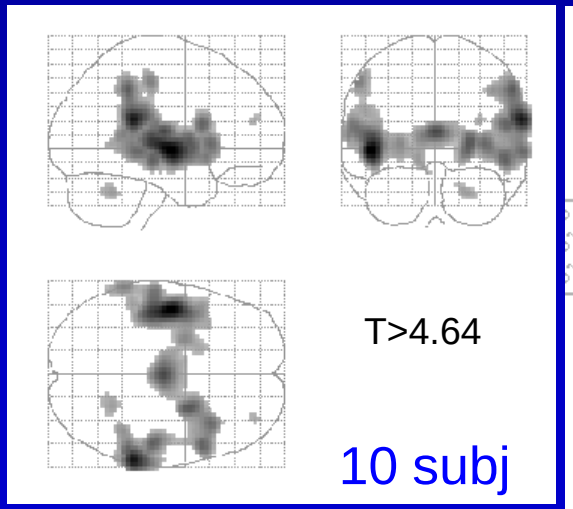
1 3 20 23 24 27 28 32 34 40



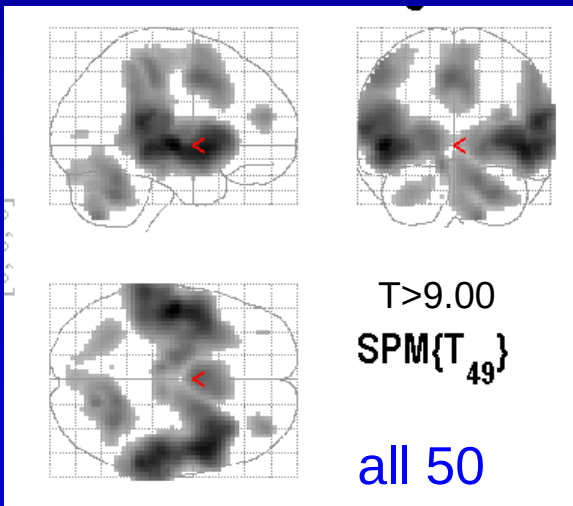
9 13 14 16 19 21 25 29 30 45



4 5 10 22 31 33 36 39 42 47



6 7 12 17 26 37 38 46 48 49



Arbitrary thresh of 9.0

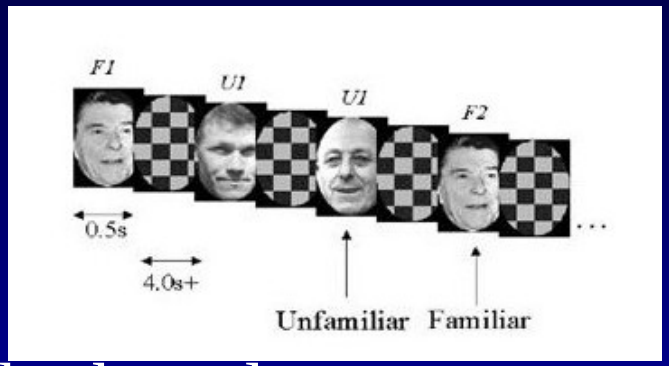
Conclusions

- Nonparametric Permutation
  - Good when Normality is question
  - Good with tiny group inference & variance smoothing
- Come to practical for more!

Using SnPM

- Choose right ‘Plug in’
  - Depends on your design
- Cluster size inference?
  - Yes: Commit to cluster-forming threshold now
  - Yes: Don’t commit, collect *huge* SnPM_ST file
- Approximate Test
  - Are you in a hurry?
  - Or are more than 10,000 perms possible?
  - Then ‘Yes’ approximate, and choose 1k-10k perms
- Which image? T or P?
  - T: Traditional voxel-wise or cluster-wise FWE
  - P: Uncorrected voxel-wise or FDR

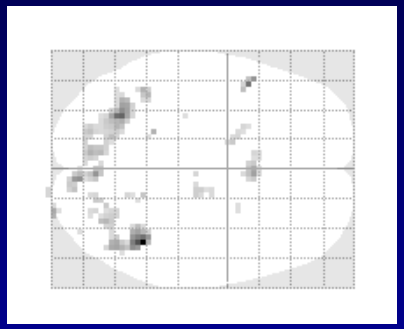
Henson et al. Faces Data



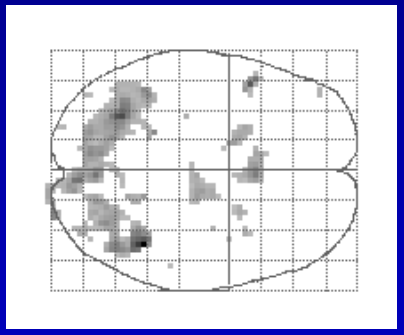
- Famous-vs-Nonfamous faces
  - Chapter 30 of SPM manual
  - Main effect, Any Faces – Checkerboard
  - 12 subjects
    - 'cons_can' Canonical HRF only
    - 'cons_informed' Canonical + Temp Deriv + Disp Deriv
- Will compare SnPM to SPM
  - For 1-sample t-test (cons_can) & Repeated measures ANOVA (cons_informed)

Voxel-Wise Results
Canonical HRF t test

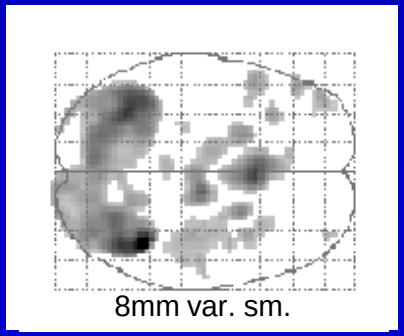
- SPM
  - u_{FWE} = 9.071, 371 voxels



- SnPM
  - u_{FWE} = 7.925, 917 voxels

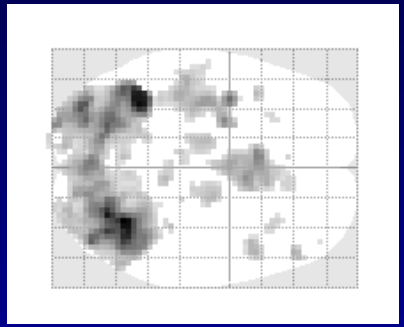


- SnPM w/ Var Smoothing
  - (u_{FWE} not comparable) 3575 voxels w/ 6mm
  - 3483 voxels w/ 4mm



Voxel-Wise Results
3-Basis F test

- SPM
  - u_{FWE} : 3733 voxels



- SnPM
  - u_{FWE} : 5105 voxels