What’s the best way to pre-process your data?

“To resample or not resample”

Included: a detour into cluster-failure land.

The short answer

There is a short, simple and almost always correct answer to this question.

The short answer

There is a short, simple and almost always correct answer to this question.

Unfortunately it is also a very unhelpful answer.

The short answer

There is a short, simple and almost always correct answer to this question.

Unfortunately it is also a very unhelpful answer.

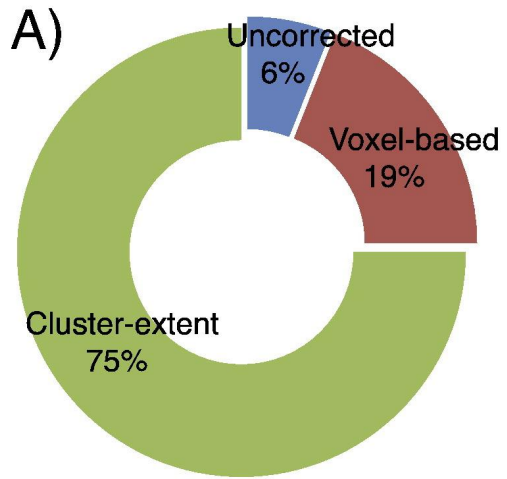
IT DEPENDS.

Caveats

- Not a systematic review on the issue.

Caveats

- Not a systematic review on the issue.
- This concerns (mostly) cluster-based univariate analysis.



Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations

Choong-Wan Woo, Anjali Krishnan, Tor D. Wager *

1500 papers and included 814 studies

https://doi.org/10.1016/j.neuroimage.2013.12.058

Caveats

- Not a systematic review on the issue.
- This concerns (mostly) cluster-based univariate analysis.
- Voxel based univariate analysis is not the issue.
  - But remember to correct for multiple comparison.
- How do those choices affect MVPA analysis ?
  - Anyone knows papers on that shows something more systematic than “YMMV depending on smoothing”?

Reminders

Preprocessing helps increase signal to noise ratio.

One sample t-test : $t = \frac{Z}{s} = \frac{\bar{X} - \mu}{\hat{\sigma}/\sqrt{n}}$

Removing noise: ↓ s ; ↑ t

Like with any methods: balance false positive and false negative results.

How to answer the question

A) Take one dataset and run it through a “multiverse analysis”.

A Study of Analysis Parameters That Influence the Sensitivity of Event-Related fMRI Analyses

Joseph B. Hopfinger,* Christian Büchel,† Andrew P. Holmes,‡ and Karl J. Friston†

doi:10.1006/nimg.2000.0549

How to answer the question

A) Take one dataset and run it through a “multiverse analysis”.

36 analyses were performed on a single fMRI data-set (one subject).

Varying parameters along four axes:

1. **resampled voxel size (2 or 3 mm^3 voxels)**
2. spatial smoothing (FWHM: 6, 10, 14 mm)
3. temporal smoothing (FWHM: 4, 8 seconds)
4. the set of basis functions (HRF, HRF + temp derivatives, HRF + temp + disp derivatives)

16 ROIs

How to answer the question

A) Take one dataset and run it through a “multiverse analysis”.

36 analyses were performed on a single fMRI data-set (one subject).

Varying parameters along four axes:

1. **resampled voxel size (2 or 3 mm^3 voxels)**
2. spatial smoothing (FWHM: 6, 10, 14 mm)
3. temporal smoothing (FWHM: 4, 8 seconds)
4. the set of basis functions (HRF, HRF + temp derivatives, HRF + temp + disp derivatives)

16 ROIs

Sensitivity (i.e., the probability of detecting an activation given it exists) was assessed in terms of:
- Z scores (as metric for uncorrected P values)
- the negative log of the expected Euler characteristic (as metric for corrected P values).

How to answer the question

A) Take one dataset and run it through a “multiverse analysis”.

“*Smaller voxel sizes ensure that the data conform to a good lattice representation and therefore reduce the likelihood that any of the assumptions required by the application of Gaussian field theory are violated.*”

How to answer the question

A) Take one dataset and run it through a “multiverse analysis”.

“*Smaller voxel sizes ensure that the data conform to a good lattice representation and therefore reduce the likelihood that any of the assumptions required by the application of Gaussian field theory are violated.*”

“*In terms of resampled voxel size, the smaller resampled voxel size did produce a significantly greater sensitivity, although the absolute measures of sensitivity changed only a small amount.* ***The increase insensitivity with 2-mm voxels is probably a result of the slight implicit smoothing in the sinc interpolation.***”

TABLE 4

Summary of Average Scores

	Z Scores	−ln (Euler)
Resampled voxel dimensions		
$2 \times 2 \times 2$ mm^3	5.26	6.59
$3 \times 3 \times 3$ mm^3	5.20	6.36
Spatial smoothing		
6 mm	5.35	5.74
10 mm	5.32	7.01
14 mm	5.02	6.67
Temporal smoothing		
4 s	5.60	8.47
8 s	4.86	4.48
Set of basis functions		
HRF (alone)	5.31	7.33
HRF and Temporal derivative	5.39	7.16
HRF and Temp and Dispersion deriv	4.99	5.17

How to answer the question

A) Take one dataset and run it through a “multiverse analysis”.

```
edit spm_defaults

Defaults.normalise.write.vox = [2 2 2];
```

TABLE 4

Summary of Average Scores

	Z Scores	−ln (Euler)
Resampled voxel dimensions		
2 × 2 × 2 mm³	5.26	6.59
3 × 3 × 3 mm³	5.20	6.36
Spatial smoothing		
6 mm	5.35	5.74
10 mm	5.32	7.01
14 mm	5.02	6.67
Temporal smoothing		
4 s	5.60	8.47
8 s	4.86	4.48
Set of basis functions		
HRF (alone)	5.31	7.33
HRF and Temporal derivative	5.39	7.16
HRF and Temp and Dispersion deriv	4.99	5.17

How to answer the question

B) Simulate data and analyse it in different ways

How to answer the question

B) Simulate data and analyse it in different ways

“Sensitivity (i.e., the probability of detecting an activation **given** it exists)”

“*We employed an activation paradigm that evoked responses in multiple cortical and subcortical areas that are **assumed** to represent canonical cortical and subcortical activations.*”

But what’s the “**ground truth**”?

How to answer the question

B) Simulate data and analyse it in different ways

```
%% False positive rate simulation

number_simulation = 10000;
number_observation = 10000;

% mean and standard deviation for our "noise" distribution
signal = 0;   % < ----- This is the "ground 'truth'"
std = 1;

% generate data that follows N(M = signal, STD = std)
data = randn(number_observation, number_simulation) * std + signal;

% Test how often a t-test gives us a false positive results
H = ttest(data, M, 'alpha', 0.05, 'tail', 'both');

% Compute number of positive results
sum(H)/number_simulation

ans = 0.0529
```

How to answer the question

B) Simulate data and analyse it in different ways

If the assumptions of your “forward model” (the model you use to generate your data) do not “match” the real data generation process then your simulations could be misleading.

https://www.discovermagazine.com/mind/fmri-and-false-positives-a-basic-flaw

“Specifically, the problem is that the [fMRI] software assumes that the spatial autocorrelation function has a Gaussian shape but in fact it has 'long tails', with more long-range correlations than expected.”

How to answer the question

C) Cluster F… ailure approach

- Use real data with no signal,
- Fit different “dummy models”
- See how often they come up with false positive result

- Ground truth → there is no signal
- Realistic “noise”

How to answer the question

C) Cluster F… ailure approach

Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates

Anders Eklund[a,b,c,1], Thomas E. Nichols[d,e], and Hans Knutsson[a,c]

- Take resting state data (no “activation” signal, but realistic noise)
- Fit GLMs “pretending” the subjects did a task
- See how often those analysis give positive results at the voxel / cluster level.

doi:10.1073/pnas.1602413113

How to answer the question

C) Cluster F… ailure approach

- Several large datasets
- Type of dummy designs
  - Block: B1, B2
  - Event related: E1, E2
- Level of spatial smoothing
  - 4, 6, 8, 10 mm FWHM
- “Softwares”
  - SPM
  - FSL
  - AFNI
  - Non-parametric (permutation test)

How to answer the question

C) Cluster F… ailure approach

- Looked at
  - Voxel based inference
  - Cluster based inference
- Different levels of cluster-defining-threshold (CDT)
  - p = 0.01
  - p = 0.001

CDT also known as:

- Peak level threshold
- Voxel level threshold
- Height threshold
- Cluster forming threshold

Can also be expressed as a T or F value.

How to answer the question

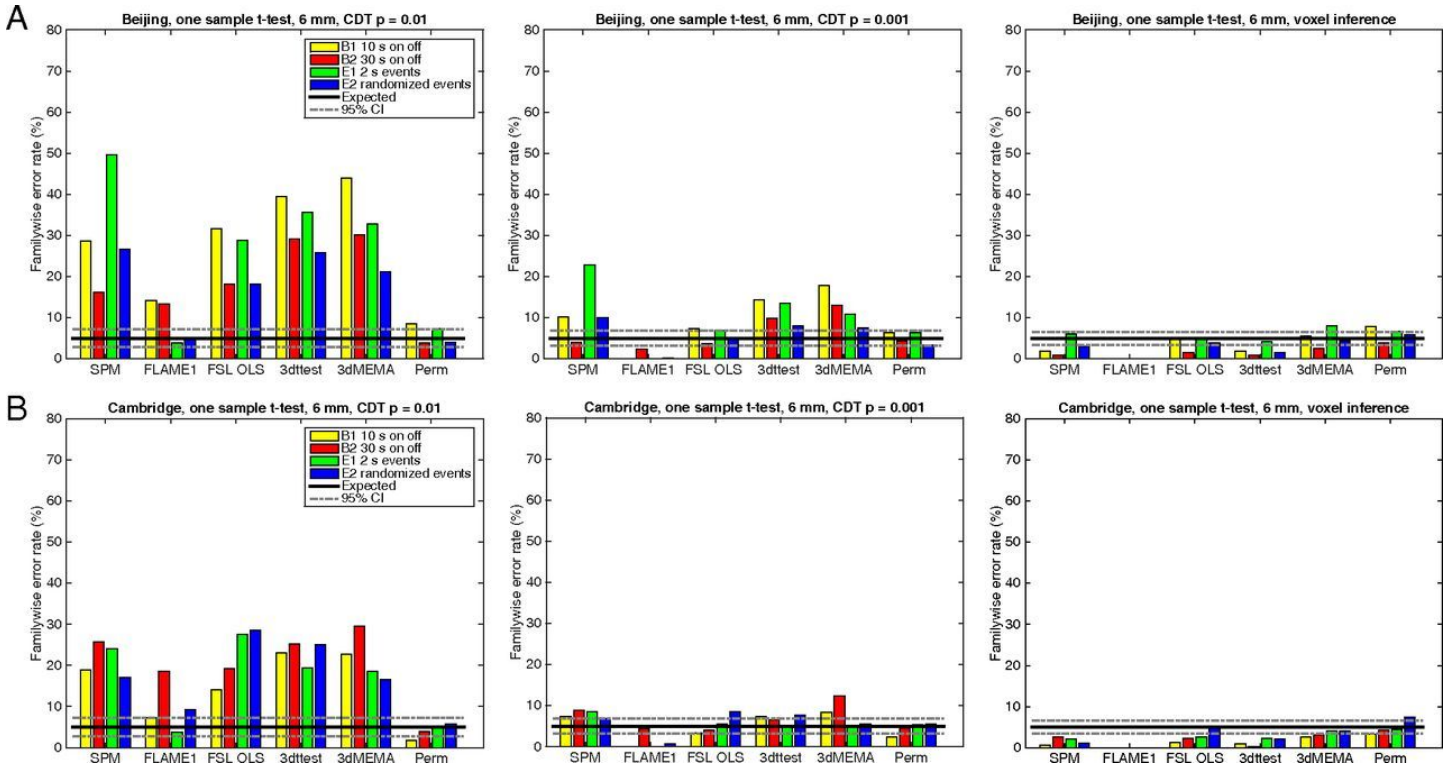
C) Cluster F… ailure approach

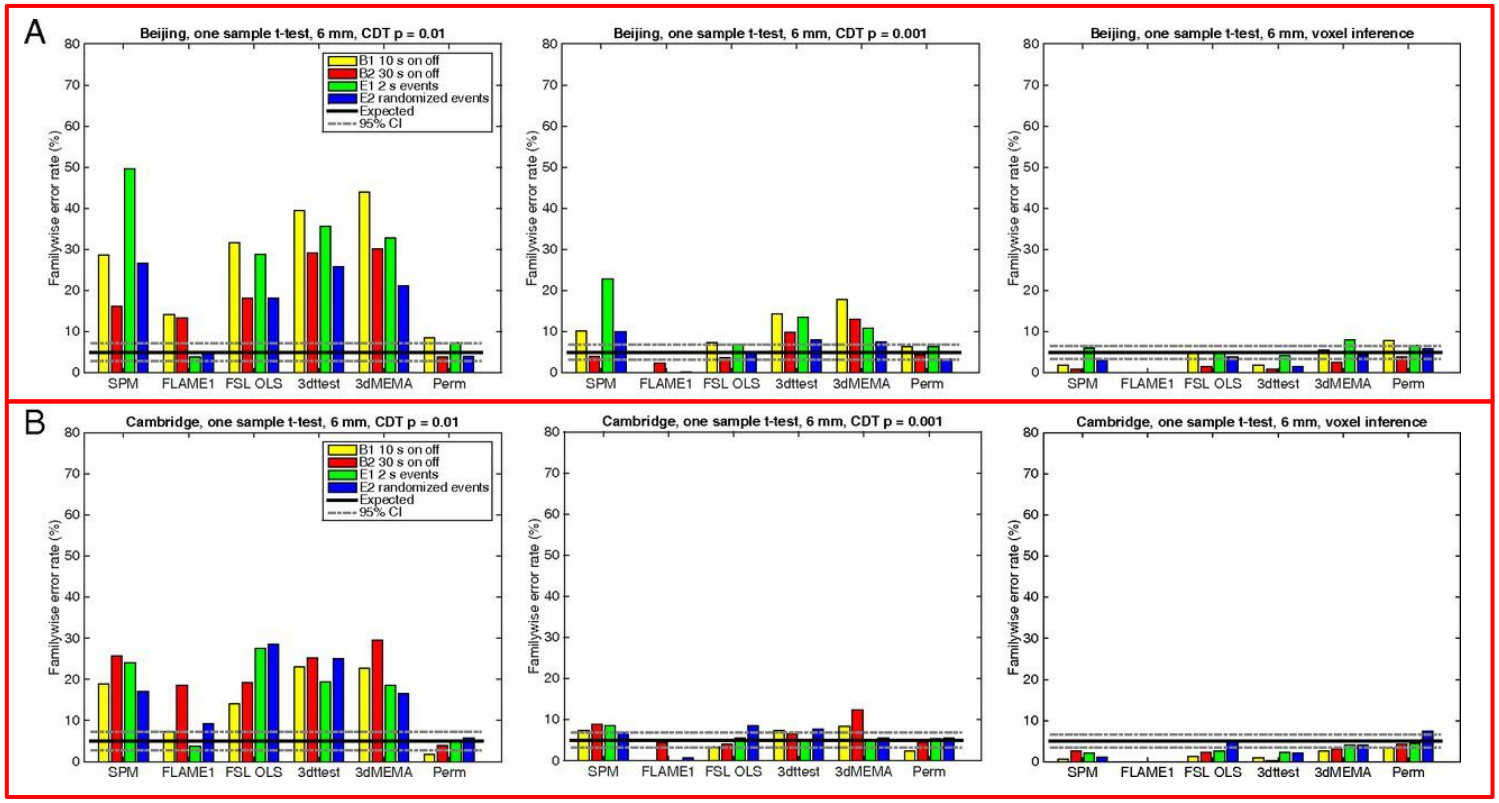
- Looked at
  - Voxel based inference
  - Cluster based inference
- Different levels of cluster-defining-threshold (CDT)
  - p = 0.01
  - p = 0.001
- SPM “demo”

CDT also known as:

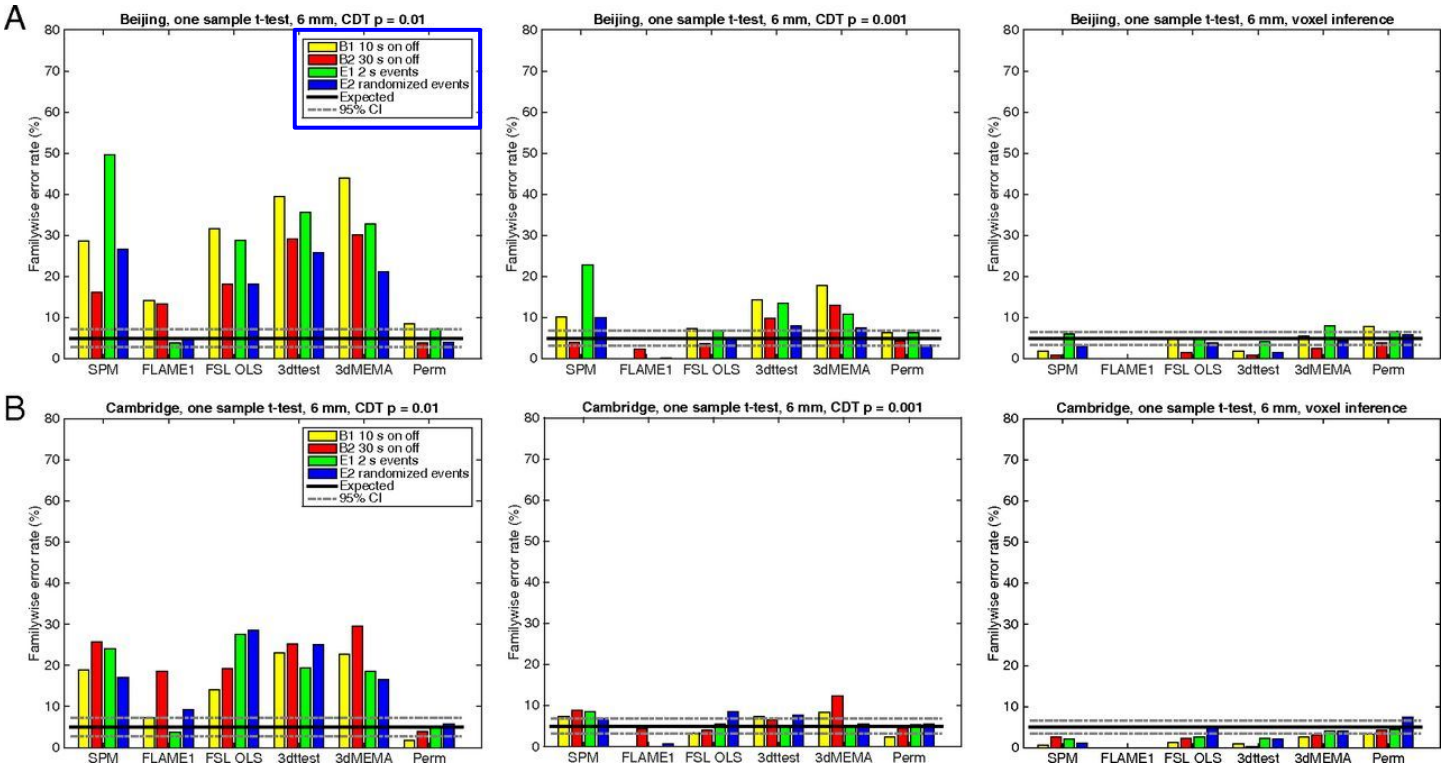
- Peak level threshold
- Voxel level threshold
- Height threshold
- Cluster forming threshold

Can also be expressed as a T or F value.

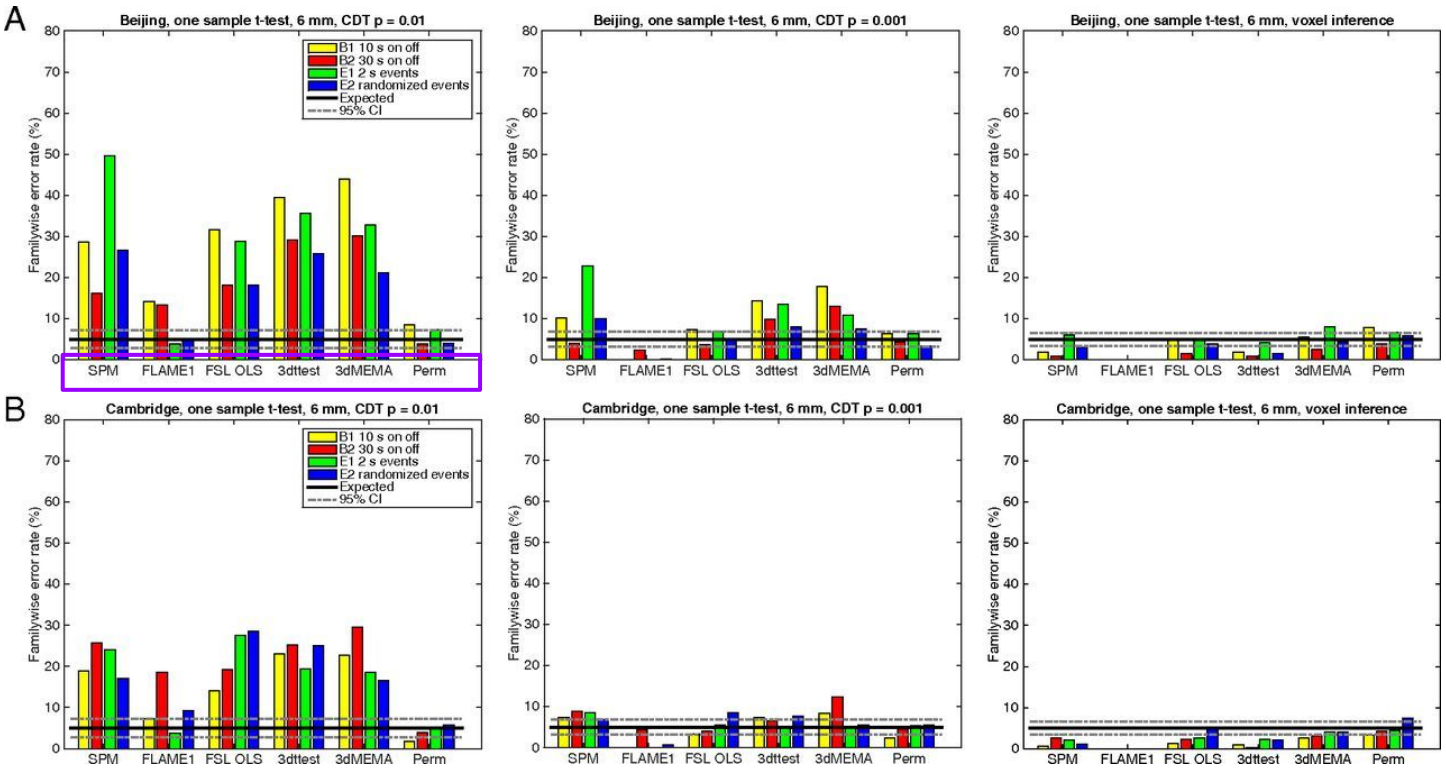

A
Beijing, one sample t-test, 6 mm, CDT p = 0.01
Beijing, one sample t-test, 6 mm, CDT p = 0.001
Beijing, one sample t-test, 6 mm, voxel inference
B
Cambridge, one sample t-test, 6 mm, CDT p = 0.01
Cambridge, one sample t-test, 6 mm, CDT p = 0.001
Cambridge, one sample t-test, 6 mm, voxel inference
B1 10 s on off
B2 30 s on off
E1 2 s events
E2 randomized events
Expected
95% CI
Familywise error rate (%)
SPM
FLAME1
FSL OLS
3dttest
3dMEMA
Perm



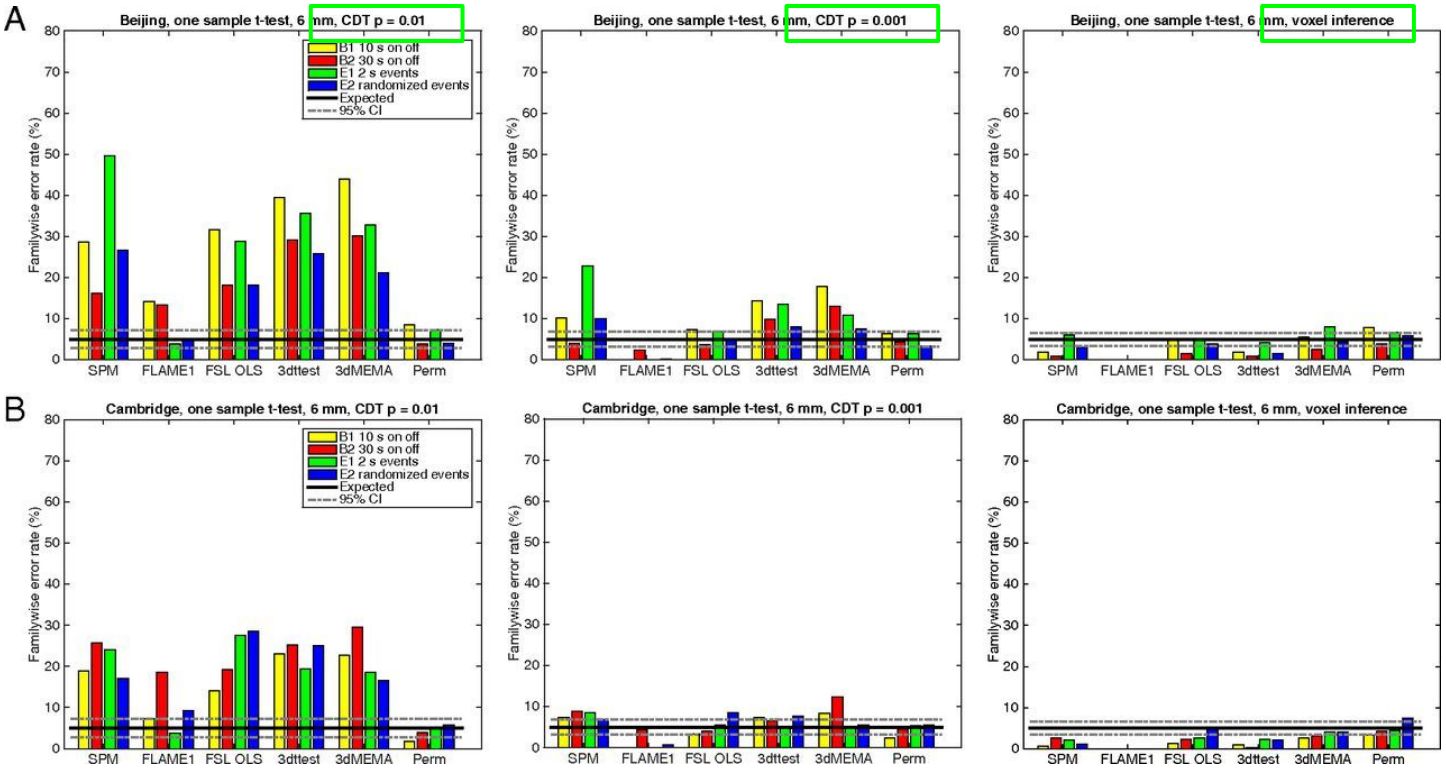
- dataset
- design
- software
- inference



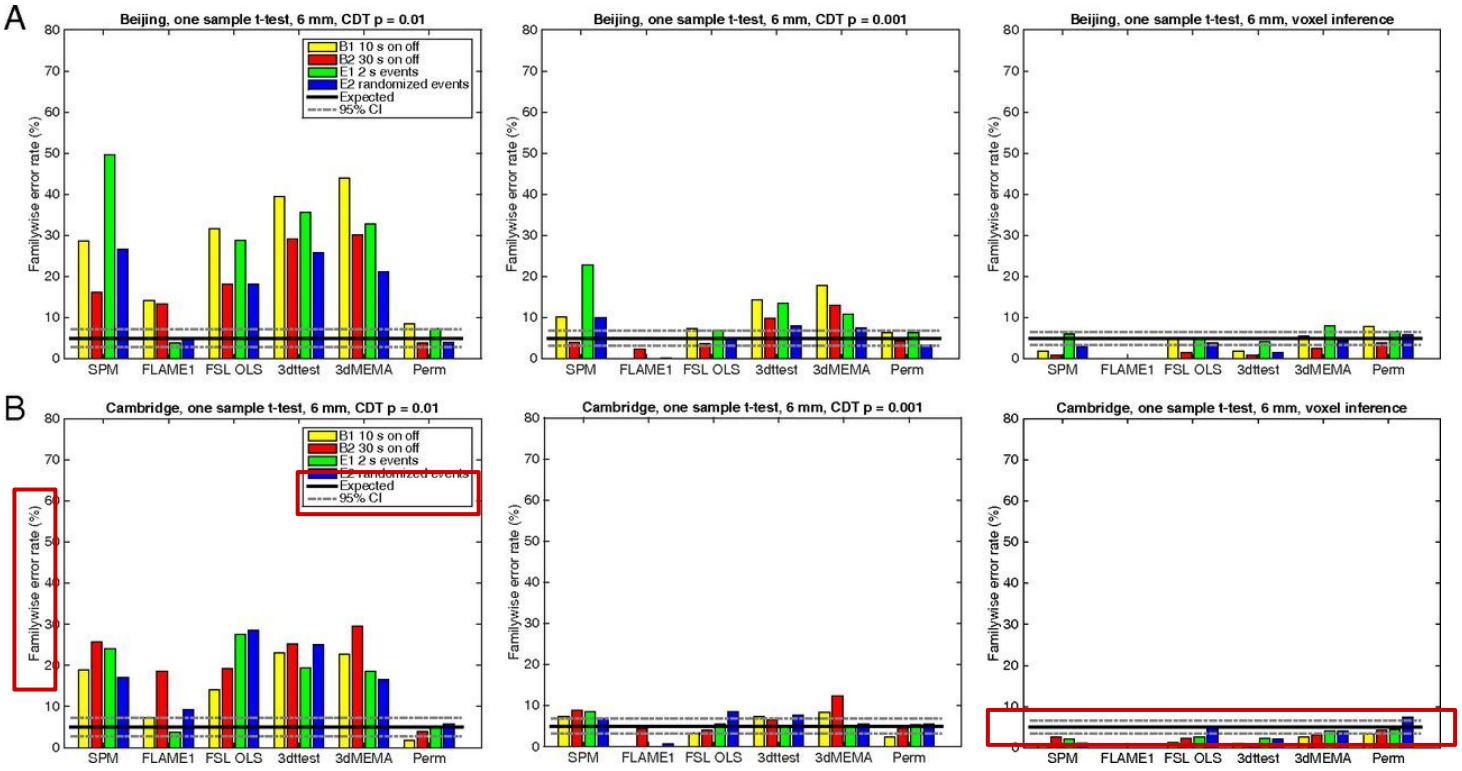
- dataset
- design
- software
- inference



- dataset
- design
- software
- inference

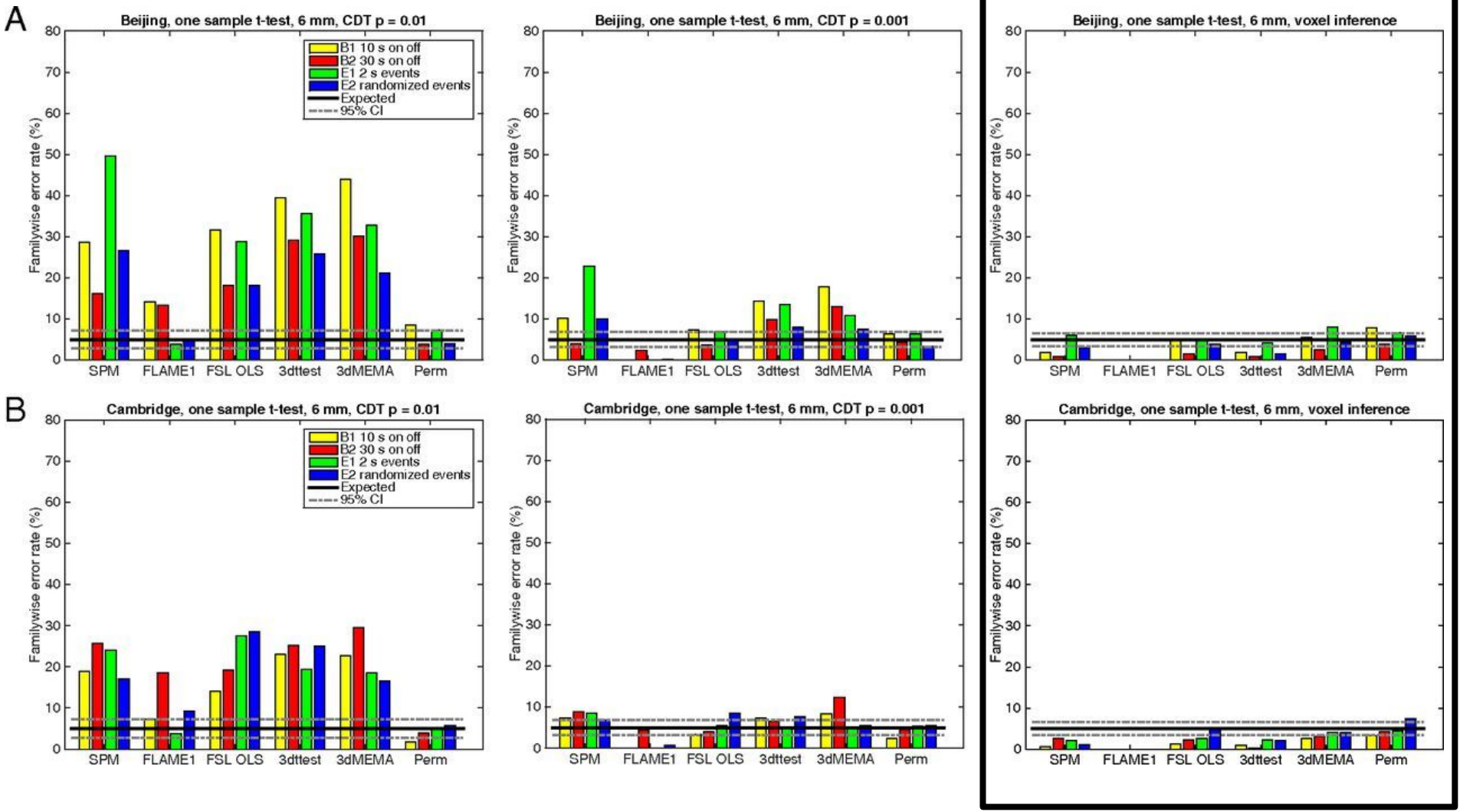


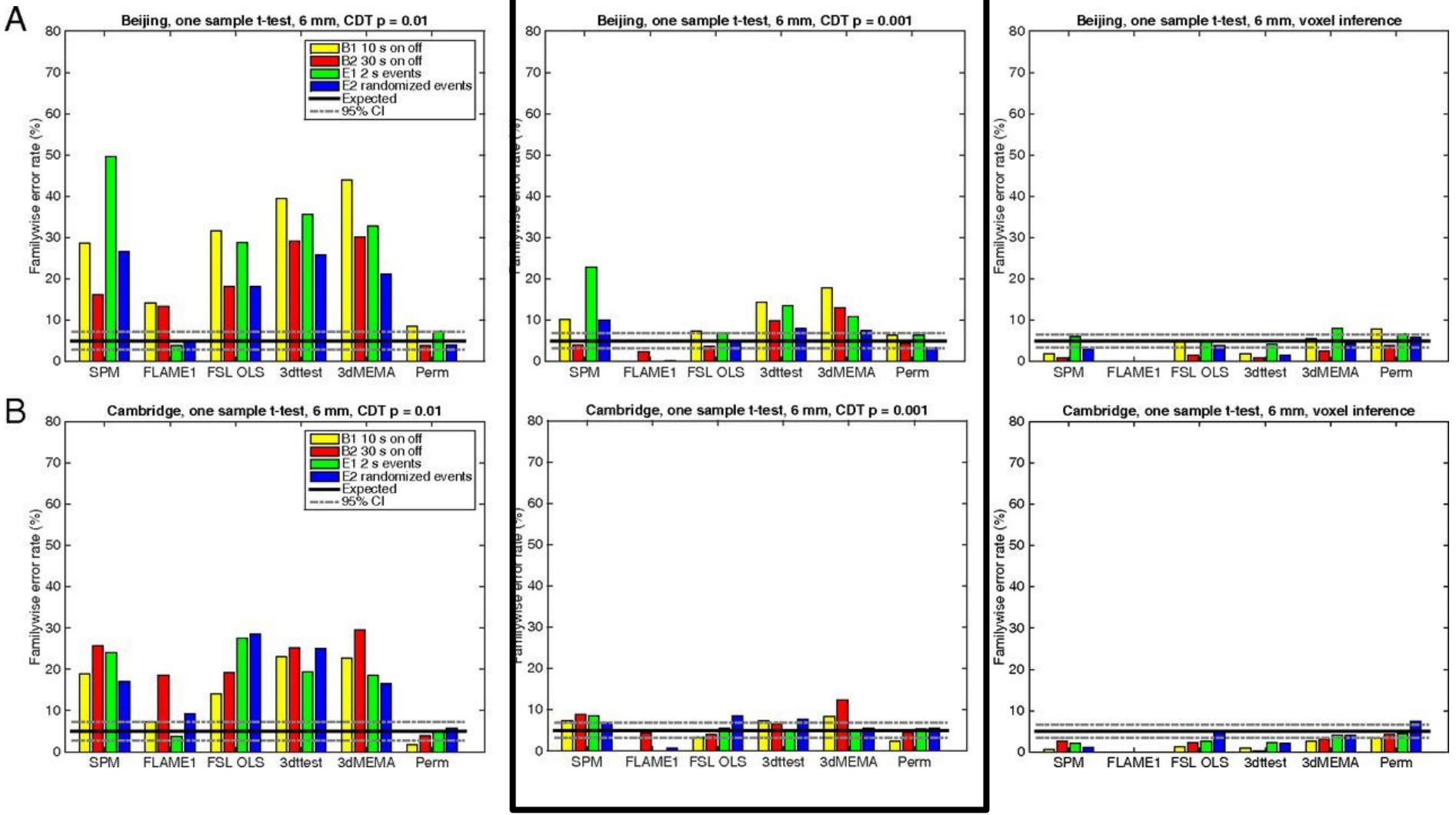
- dataset
- design
- software
- inference

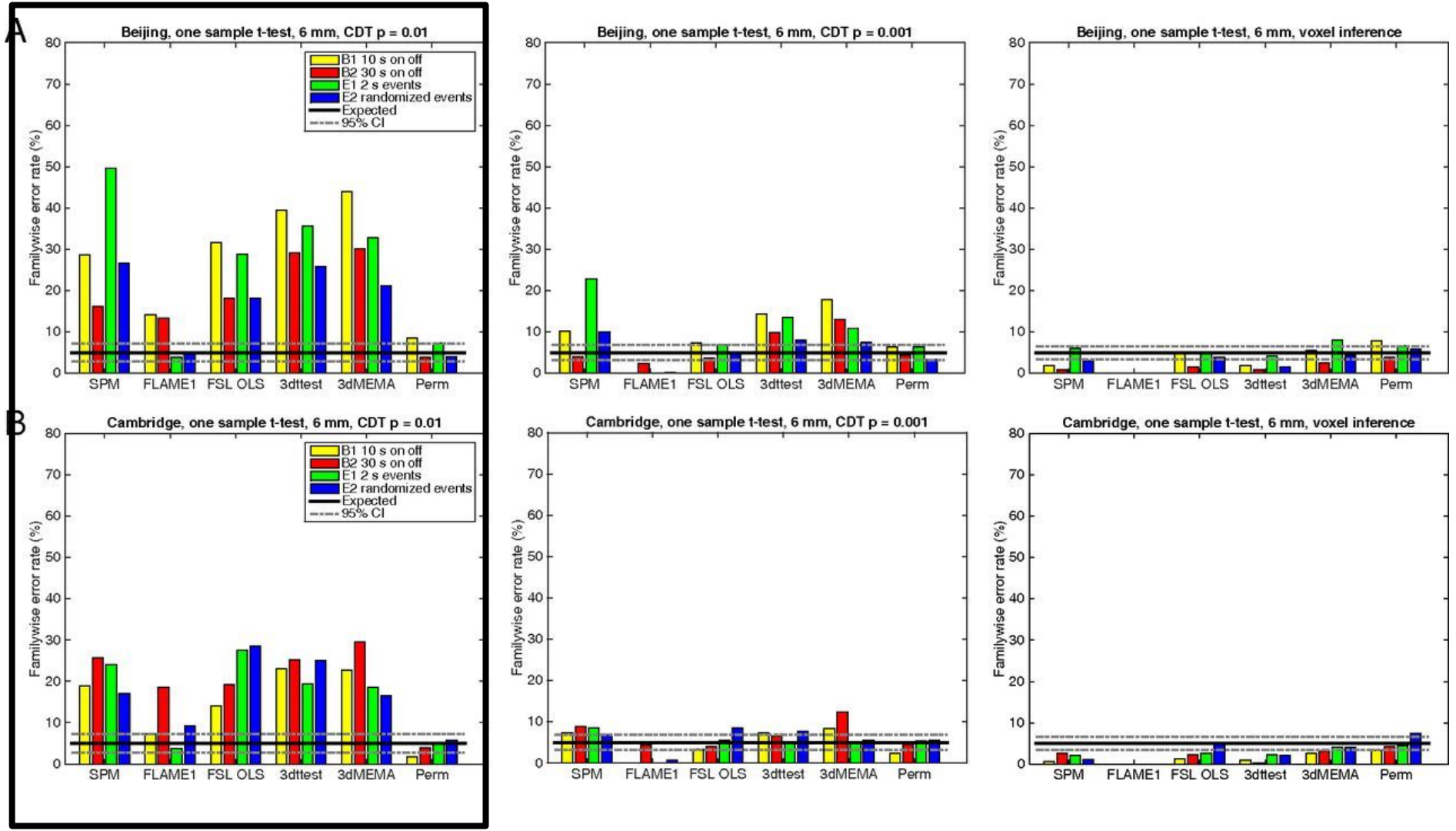


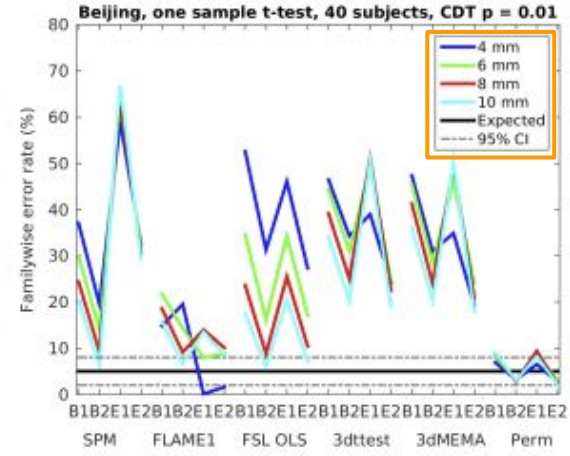
- dataset
- design
- software
- inference

- FWE

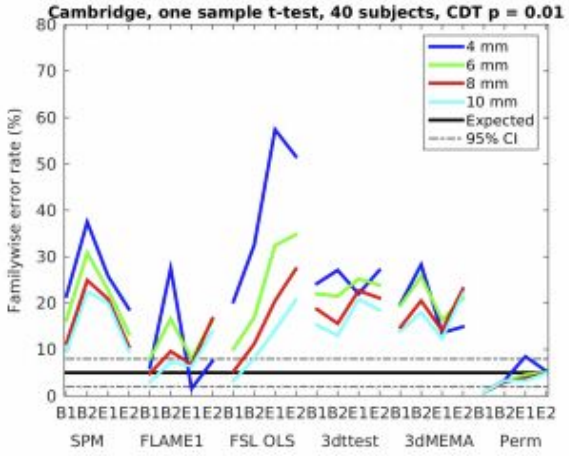

A
Beijing, one sample t-test, 6 mm, CDT p = 0.01
Beijing, one sample t-test, 6 mm, CDT p = 0.001
Beijing, one sample t-test, 6 mm, voxel inference
B
Cambridge, one sample t-test, 6 mm, CDT p = 0.01
Cambridge, one sample t-test, 6 mm, CDT p = 0.001
Cambridge, one sample t-test, 6 mm, voxel inference
B1 10 s on off
B2 30 s on off
E1 2 s events
E2 randomized events
Expected
95% CI
Familywise error rate (%)
SPM
FLAME1
FSL OLS
3dttest
3dMEMA
Perm


A
Beijing, one sample t-test, 6 mm, CDT p = 0.01
Beijing, one sample t-test, 6 mm, CDT p = 0.001
Beijing, one sample t-test, 6 mm, voxel inference
B
Cambridge, one sample t-test, 6 mm, CDT p = 0.01
Cambridge, one sample t-test, 6 mm, CDT p = 0.001
Cambridge, one sample t-test, 6 mm, voxel inference
B1 10 s on off
B2 30 s on off
E1 2 s events
E2 randomized events
Expected
95% CI
Familywise error rate (%)
SPM
FLAME1
FSL OLS
3dttest
3dMEMA
Perm

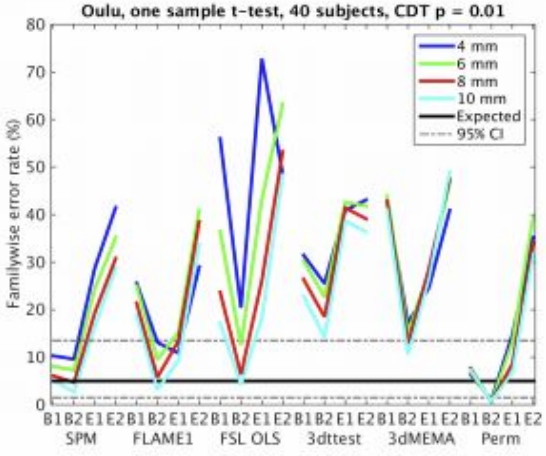

A
Beijing, one sample t-test, 6 mm, CDT p = 0.01
Beijing, one sample t-test, 6 mm, CDT p = 0.001
Beijing, one sample t-test, 6 mm, voxel inference
B
Cambridge, one sample t-test, 6 mm, CDT p = 0.01
Cambridge, one sample t-test, 6 mm, CDT p = 0.001
Cambridge, one sample t-test, 6 mm, voxel inference
B1 10 s on off
B2 30 s on off
E1 2 s events
E2 randomized events
Expected
95% CI
Familywise error rate (%)
SPM
FLAME1
FSL OLS
3dttest
3dMEMA
Perm



(a)

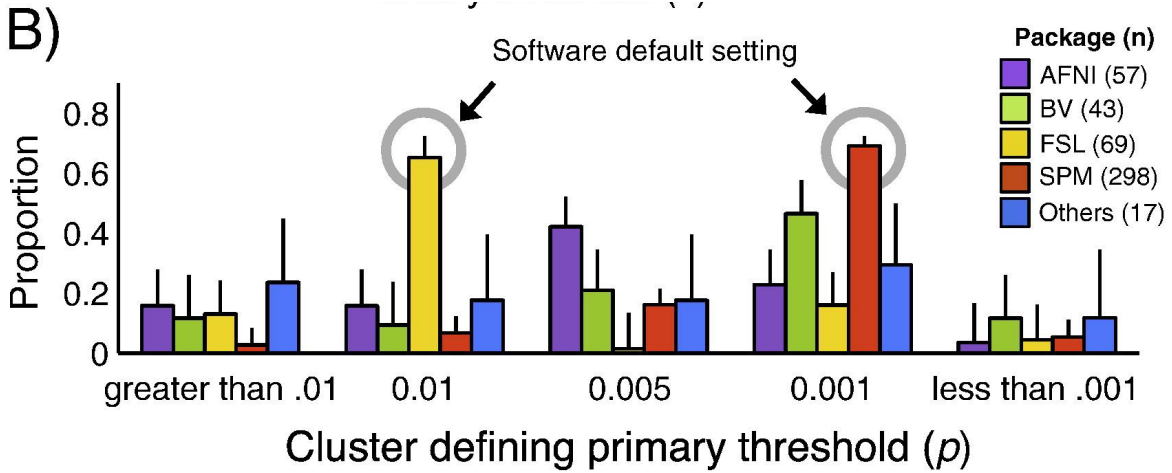


(b)



(c)

- dataset
- design
- software
- inference
- smoothing



Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations

Choong-Wan Woo, Anjali Krishnan, Tor D. Wager *

https://doi.org/10.1016/j.neuroimage.2013.12.058

How to answer the question

C) Cluster F… ailure approach - “spin off”

Analysis of Family-Wise Error Rates in Statistical Parametric Mapping Using Random Field Theory

Guillaume Flandin* and Karl J. Friston

doi:10.1002/hbm.23839

How to answer the question

C) Cluster F… ailure approach - “spin off”

Commentary: Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates

Karsten Mueller[1], Jöran Lepsien[1], Harald E. Möller[1] and Gabriele Lohmann[2,3]*

doi:10.3389/fnhum.2017.00345

How to answer the question

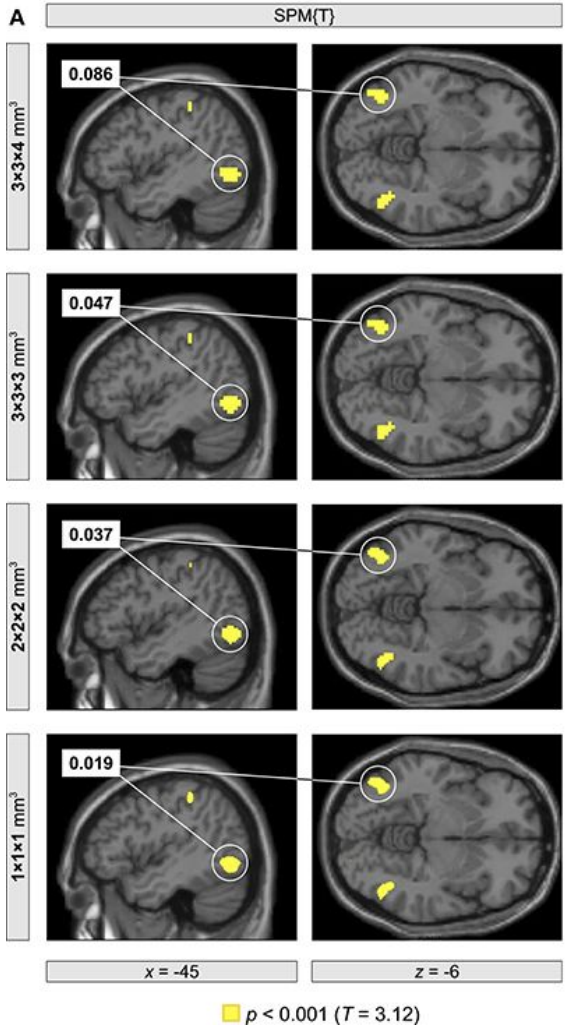
C) Cluster F… ailure approach - “spin off”

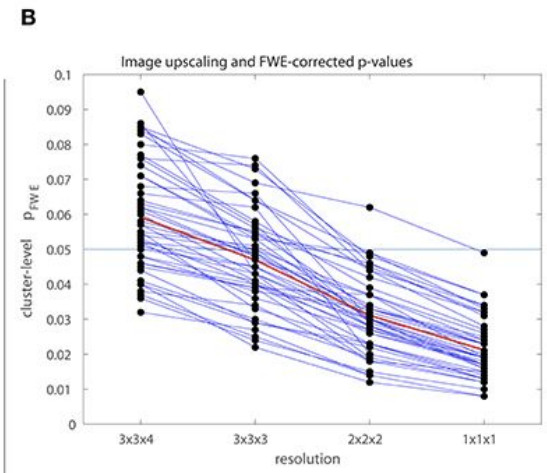
“*Eklund et al. (2016) used the default setting of 2×2×2 mm^3.*

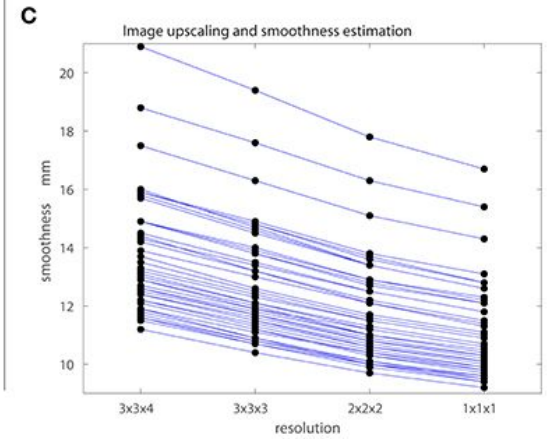
In response to Eklund's paper, Flandin and Friston (2016) used a different setting of this parameter, namely 3×3×3 mm^3. Together with a more stringent initial cluster-forming threshold, they did not observe inflated false positive rates.

However, the 2×2×2 mm^2 setting is the default in two major software packages (SPM, FSL), and in previous work, Friston and colleagues Hopfinger et al. (2000) stated that resampling to 2×2×2 mm^3 renders the analysis “more sensitive.”

In other words, at present it is unclear what a valid setting for this parameter should be.”



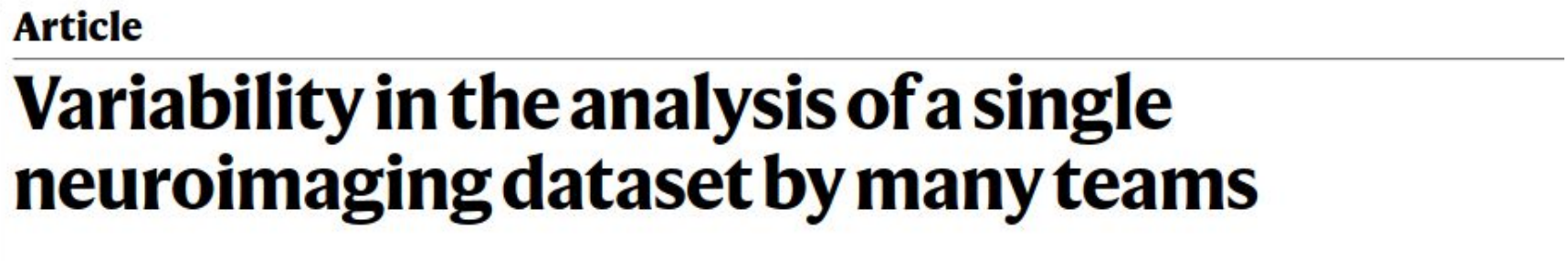




"*In other words, it appears that there is a systematic dependence of the false positive rate on the resampling parameter with* ***smaller voxel sizes leading to smaller FWE-corrected p-values and hence more false positives.***"

How to answer the question

D) One dataset - Many analysts approach



"*The strongest factor was spatial smoothness;* ***higher estimated smoothness of the unthresholded statistical maps was associated with a greater likelihood of significant outcomes****.*

The ***relevant smoothness arose from analytical steps beyond explicit smoothing****.*

No significant effect was detected for the use of standardized preprocessed data versus custom preprocessing pipelines *or for the modelling of head motion parameters.*"

In practice for CPP_SPM

In the spatial preprocessing workflow:

- During normalization to MNI.
  - By default, functional images resolution is not changed.
  - Override possible by setting `opt.funcVoxelDims` to the desired resolution.
  - Anatomical images are resampled at 1 mm (why ?).
  - Tissue probability maps downsampled at resolution of functional images (mostly to help with potential with creation of tissue-based mask and also QA)

- When no normalization is requested (`opt.space = 'individual'` instead of the default `opt.space = 'MNI'`)
  - Functional images resolution is not changed. This cannot be overridden.