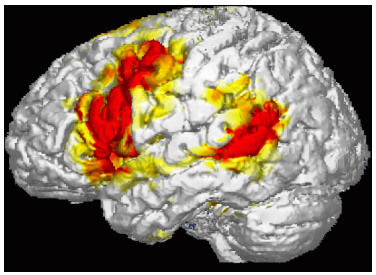
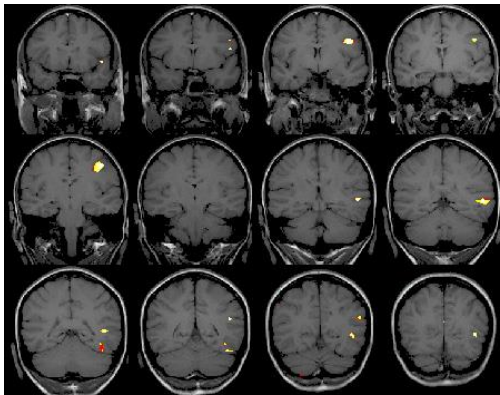
Functional MR Imaging: Experimental Designs and Analyses

Michelle Hampson

Lecture 9/11/2018



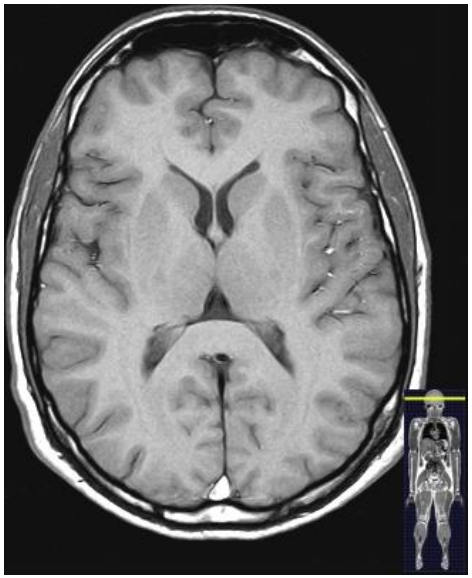


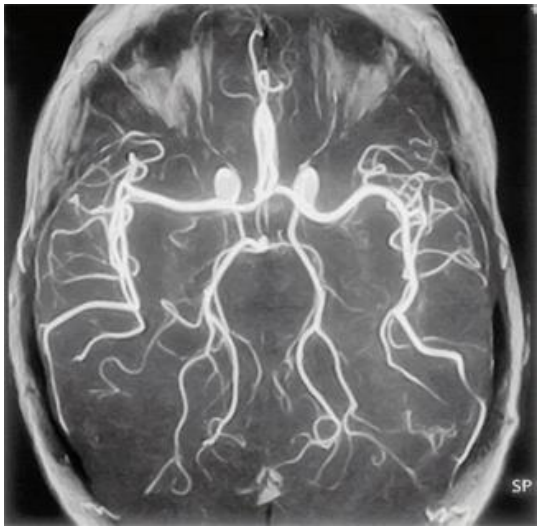
High resolution human imaging

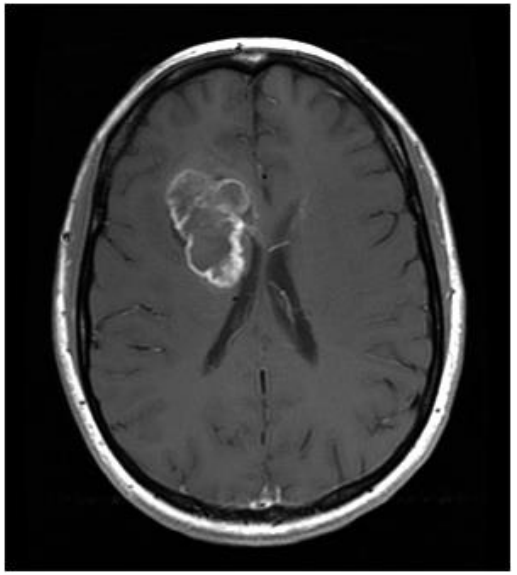
CT scans, x-rays, PET

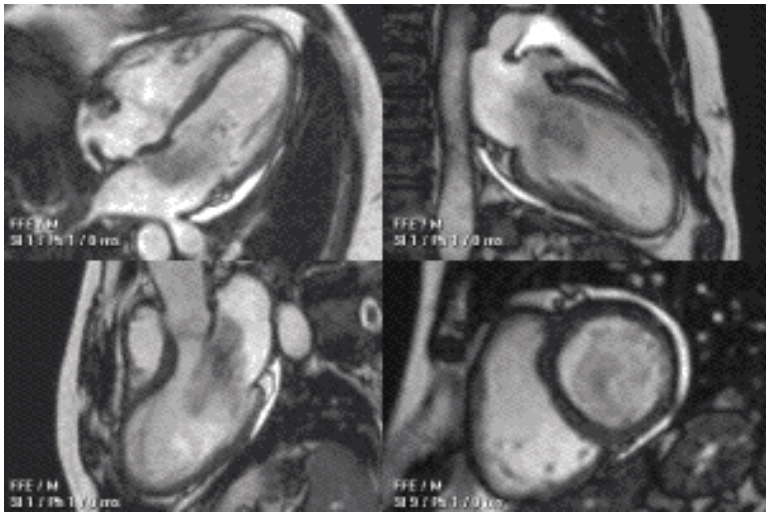
MRI

Beautiful images!









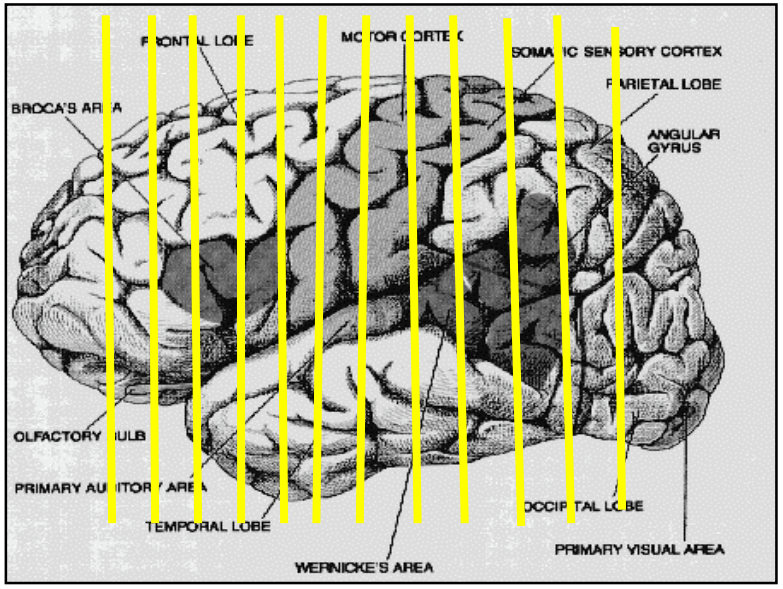
Using MR to study the human brain: what kind of information can we get?

- Structural Imaging: anatomical information
- Diffusion Tensor Imaging: By examining whether water diffuses more easily in one direction than another, can gain information about alignment of fibre pathways in the brain
- Spectroscopy: information regarding chemical composition in the brain
- **Functional Imaging (fMRI):** By examining changes in blood oxygenation level over time, can gain information regarding the changing patterns of neural activity

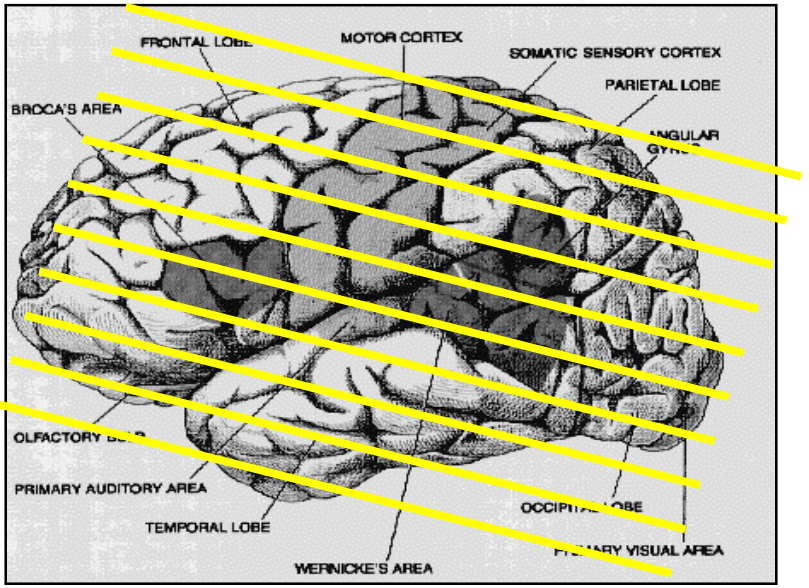
Basics of fMRI

Things you need to understand before you beginning designing/analyzing fMRI studies

Data acquired in slices



Coronal



Axial-oblique

Slice acquisition of functional data

Repetition time (TR) is the time it takes to collect a single volume

Slices are not obtained simultaneously
In our default sequences, they are spaced evenly across the TR
e.g. TR=1.5s for 10 slices there will be 150ms between slices
so the last slice is obtained almost a whole TR after the first

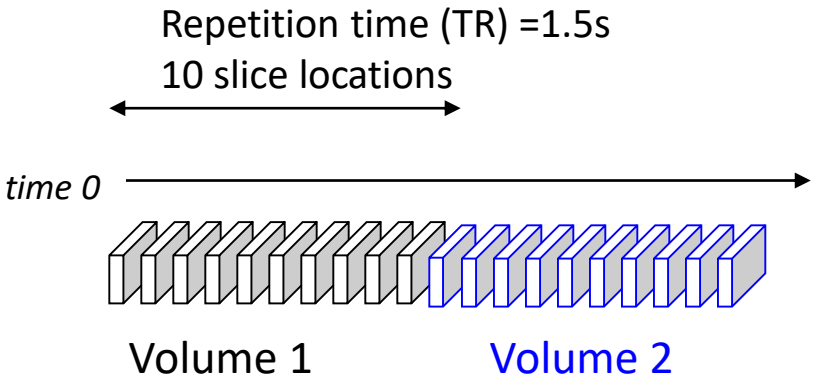
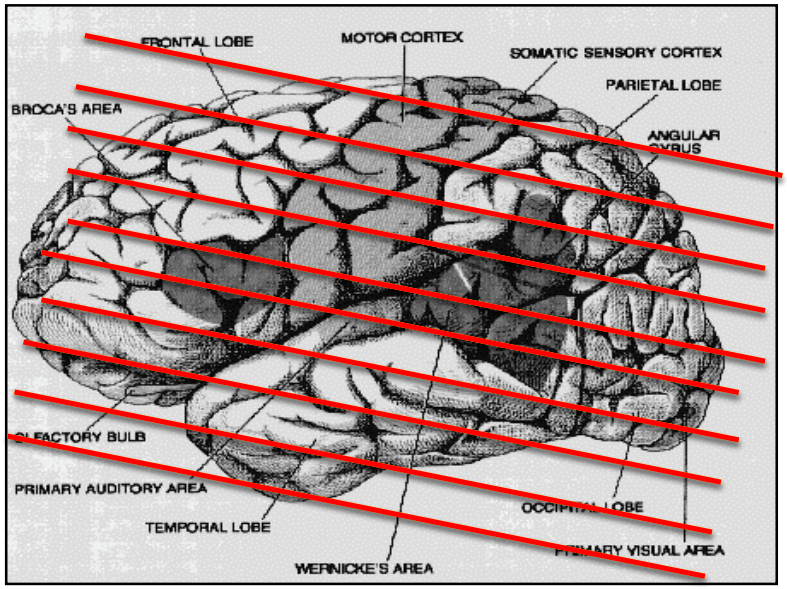
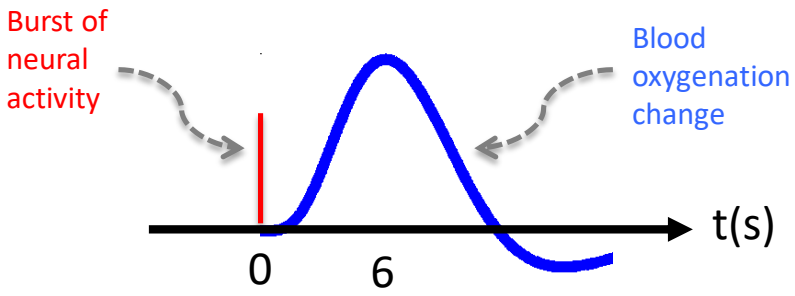


Image continuously
Typically use interleaved slice acquisition order is as follows:
[1, 3, 5, 7, 9, 2, 4, 6, 8, 10] x 200
which then repeats approx. 200 times



Functional Magnetic Resonance Imaging (fMRI)



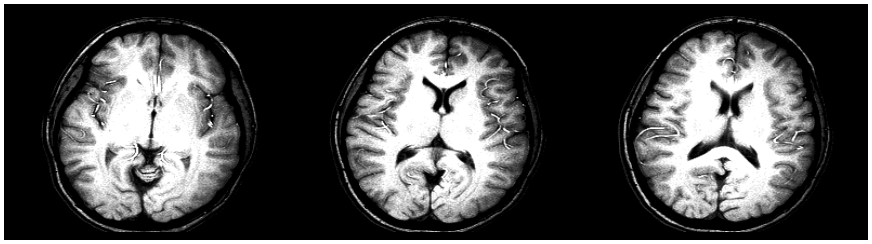
Neural activity and blood flow are tightly coupled throughout the brain.

<u>BOLD imaging (Blood Oxygenation Level Dependent)</u>

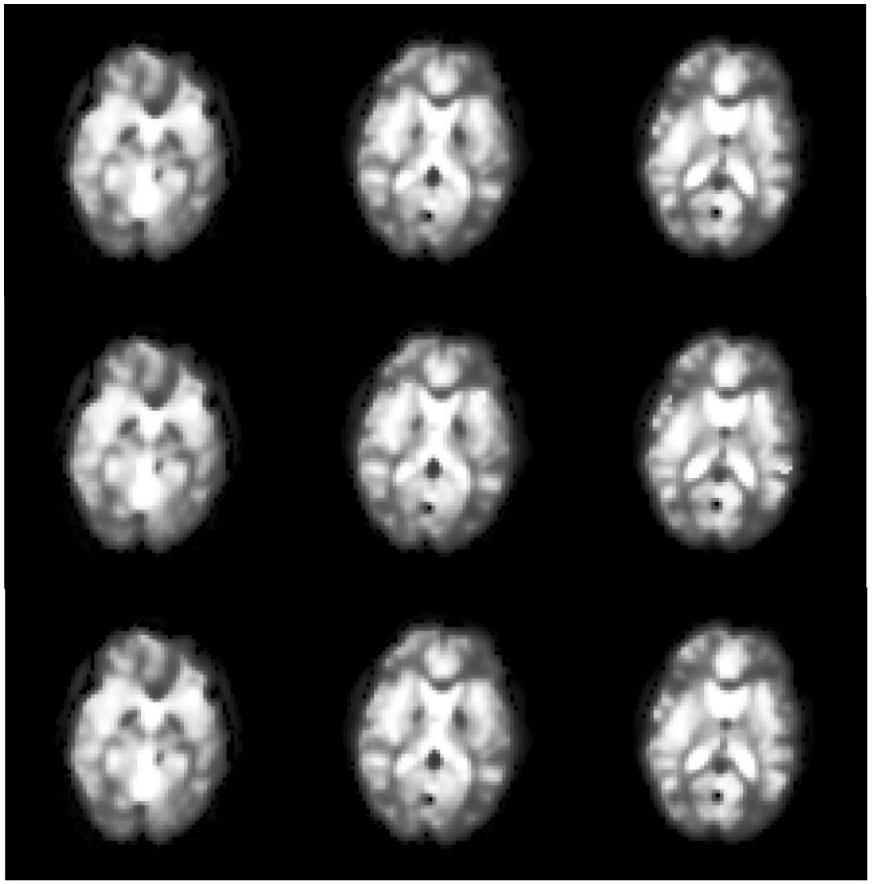
We can measure blood oxygenation fluctuations and infer neural activity changes

Functional magnetic resonance imaging (fMRI)

1. Structural data



2. Functional data



Changing over time as a result of neural activity

Time 1

Time 2

Time 3

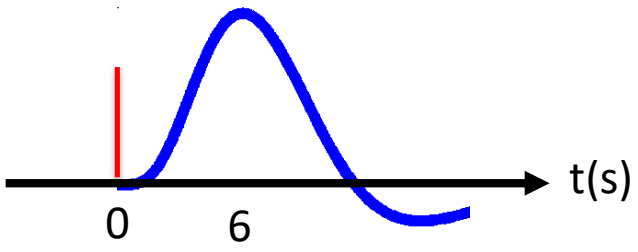
⋮

BOLD imaging

1. Everything is relative

- Absolute value of signal meaningless
- Activation studies always comparing to a baseline

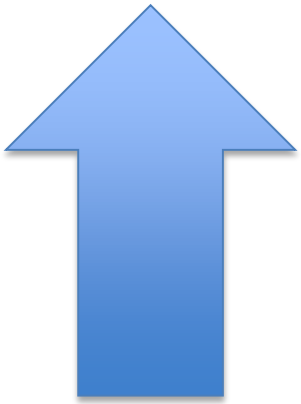
2. Temporal resolution limited to seconds



Functional MRI Data Analyses

Second level => Group Analyses

- Do healthy people have a certain pattern of brain activity?
- Do patients have a different pattern?
- Are brain patterns correlated with personal variables?



First level => Individual Subject Analyses

- What brain areas activate during a task or event?
- How are brain areas synchronized?

Individual subject analyses

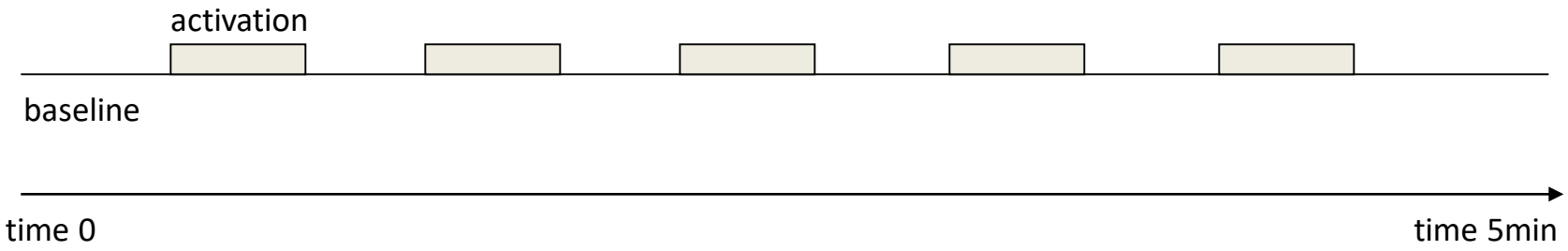
- Activation
  - block design: e.g., brain areas activated during a task
  - event-related: e.g., brain areas activated before, during and after a given event
- Functional connectivity
  - how different brain areas are synchronized with each other

Activation studies

What brain areas are involved in that mental task or event?

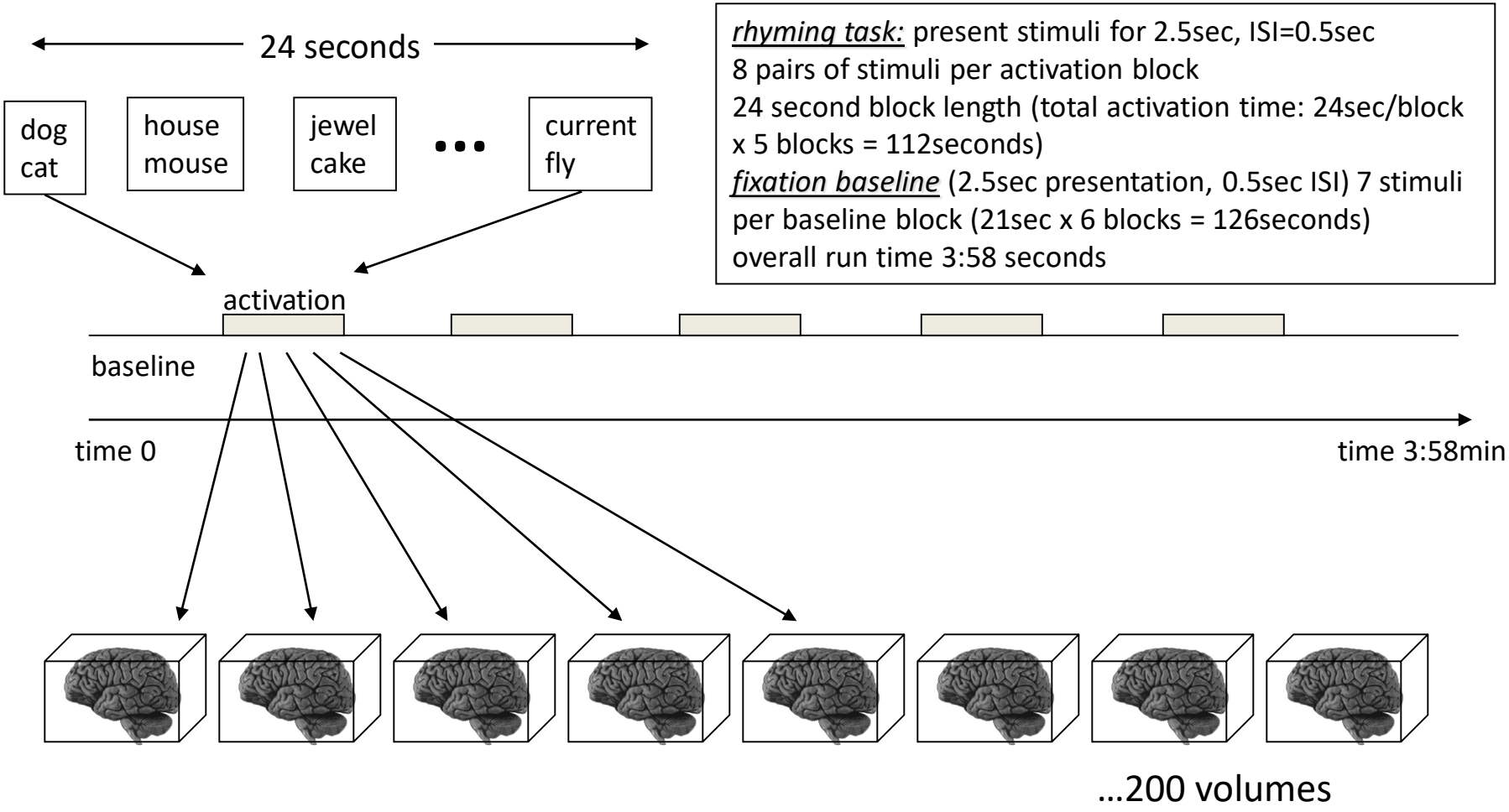
Block Design

- each run contains 4 or 5 activation blocks to be compared with 5 or 6 control blocks
- each block lasts 15-40 seconds (around 16s optimal SNR)
- images are acquired continuously throughout complete run



Block Design

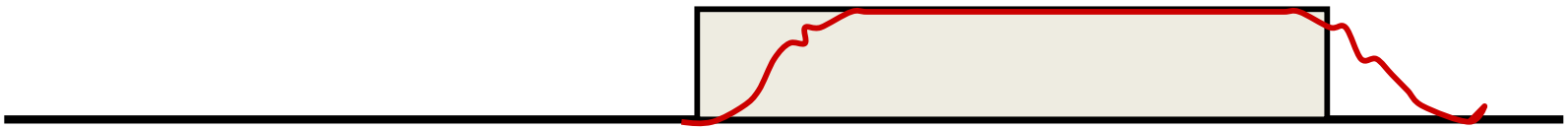
- multiple stimuli typically presented within a block
- multiple images per slice collected within a block



rhyming task: present stimuli for 2.5sec, ISI=0.5sec
8 pairs of stimuli per activation block
24 second block length (total activation time: 24sec/block x 5 blocks = 112seconds)
fixation baseline (2.5sec presentation, 0.5sec ISI) 7 stimuli per baseline block (21sec x 6 blocks = 126seconds)
overall run time 3:58 seconds

Discarding Images

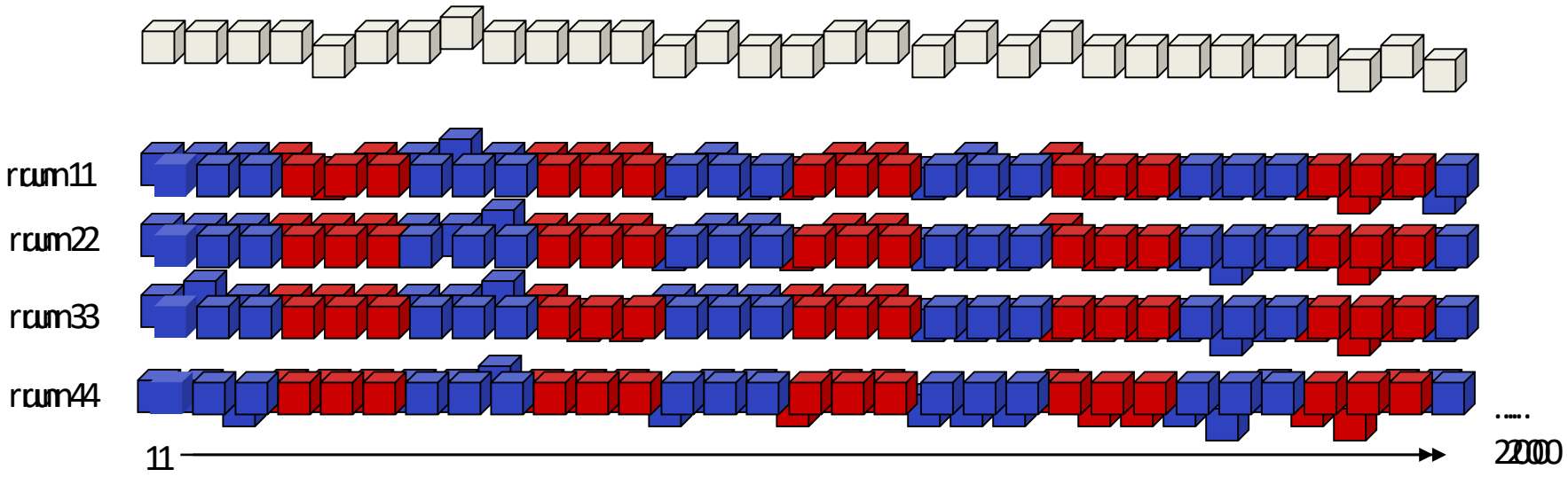
- skip first 2 acquisition images to allow magnetization to achieve steady-state
- recall blood flow response is delayed and slow - therefore skip a few images at each transition between blocks



1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23...

Baseline images for analysis	Activation images for analysis
3 to 12, 24 to ...	15 - 21, ...

Motion correction

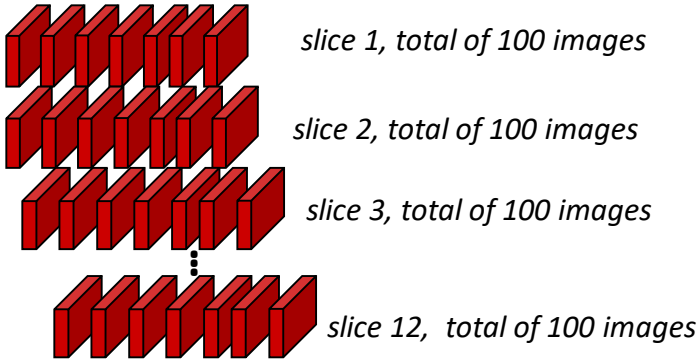


- realign within a run (and between runs: to compensate for movement)

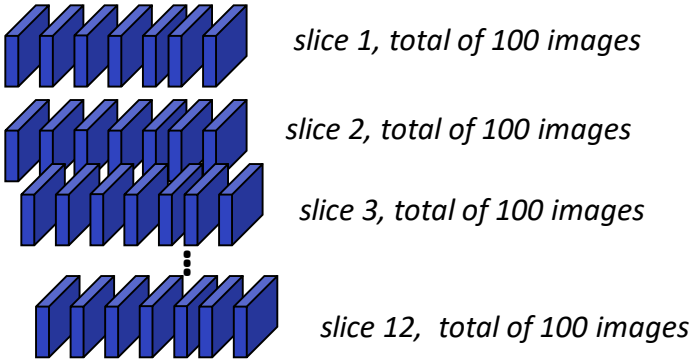
- ONLY PARTIALLY CORRECTS FOR MOTION

Separate Images into Activation/Baseline Groups

activation images



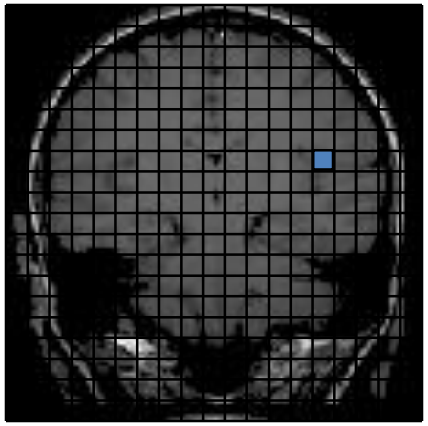
baseline images



ImageAnalysis

For each voxel in a slice: A = average value in activation condition, B = average value in baseline

percent signal change = ((A-B)/B)*100



Compute percent signal change for

- *voxel*

Repeat for all other voxels.

Strengths of Block Design Studies

- simple!
- powerful

Limitations of Block Design Studies

- cannot examine temporal response to individual stimuli
- Sometimes you are interested in spontaneous event that you are not controlling (e.g. what happens in the brain when a schizophrenic patient has a hallucination?)

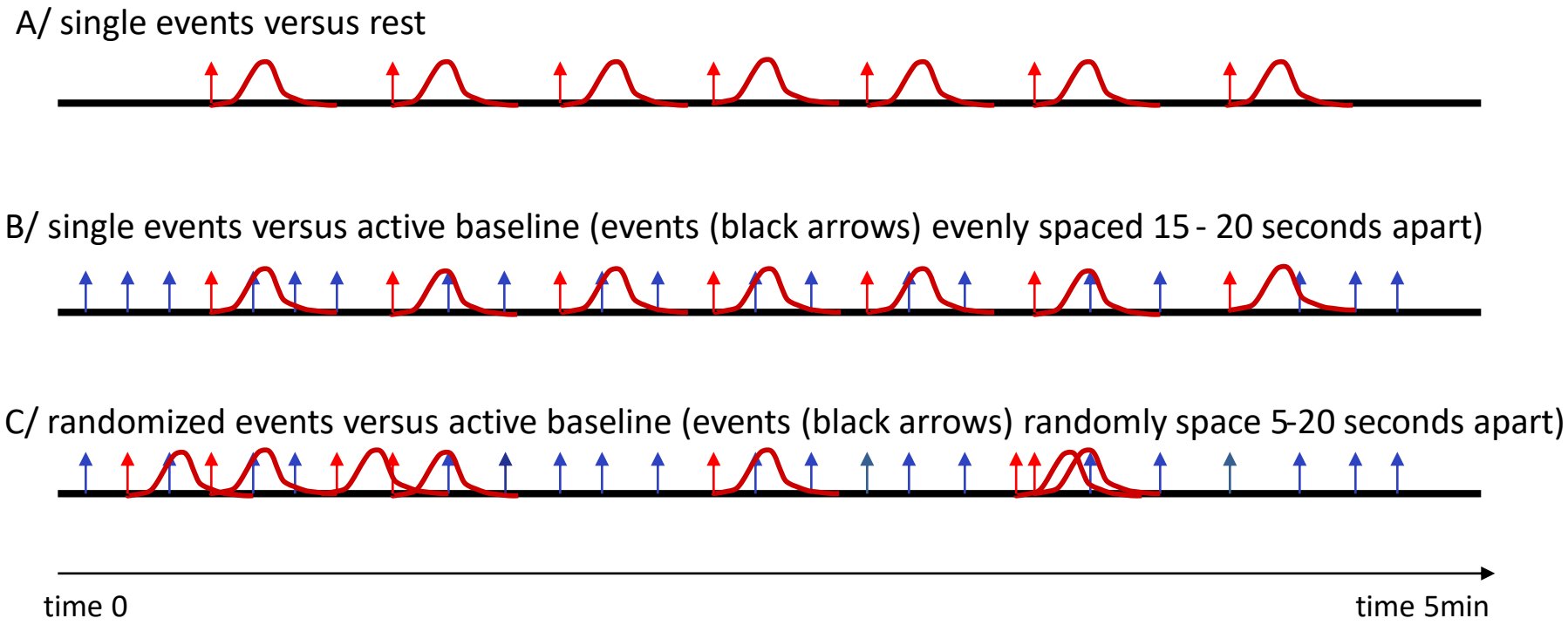
Event-related Experimental Designs

- allow response to individual stimuli to be examined
- allow examination of temporal pattern of response
- Allows examination of brain patterns associated with spontaneously occuring events

Two approaches to spacing stimuli :

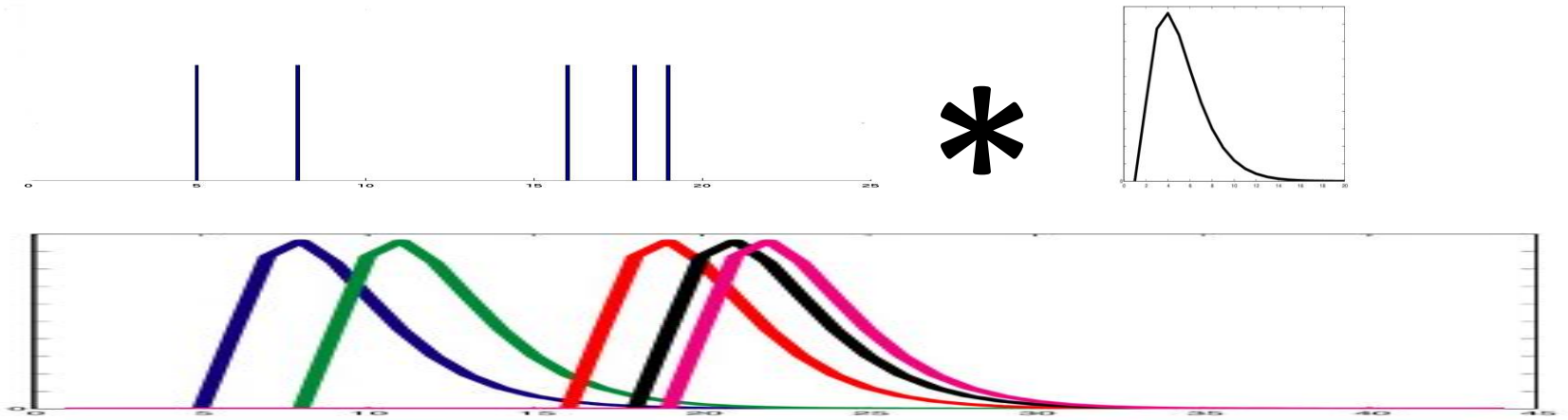
1. stimuli are separated by 15-20seconds in order to measure complete HRF of each event
   - not very many events in a single run
2. stimuli randomly spaced, often close together (“rapid event-related”)
   + Allows timecourse of response to be measured
   + allows study of spontaneously occurring event
   + compacts more events into a run
   - must assume BOLD response to series of events is sum of BOLD responses to individual events

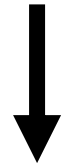
Event-Related Experimental Design

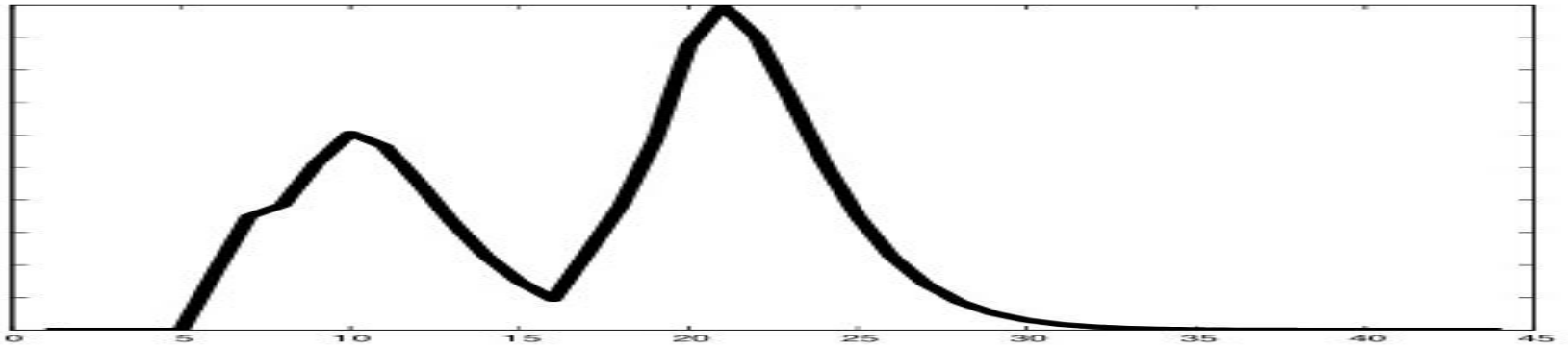


- In A and B HRF's do not overlap
- In C the HRF's overlap – must assume response to all stimuli is linear sum of responses to the individual stimuli

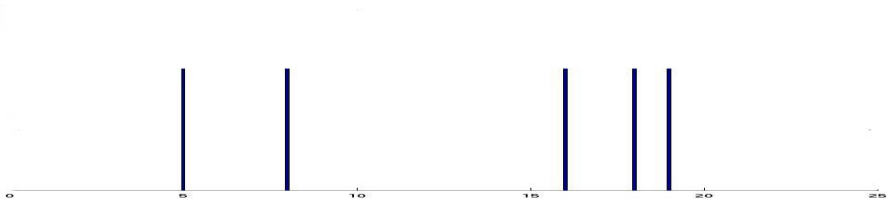
Jittered event-related study

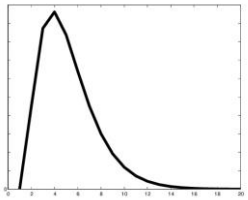




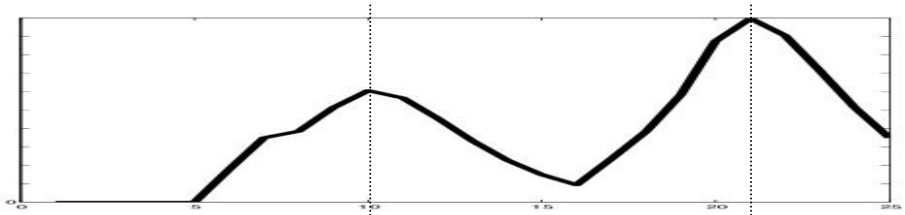


Typical event-related study

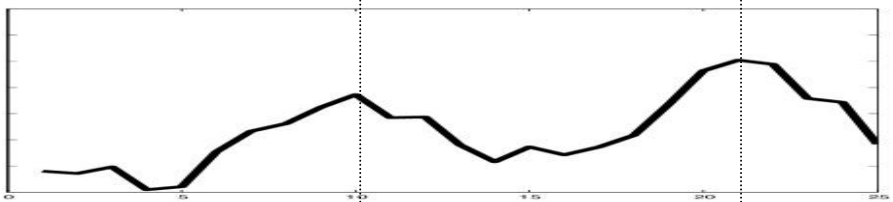




Reference timecourse

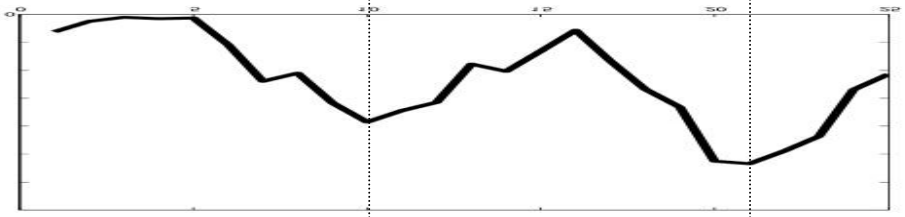


Voxel 1 (activated)



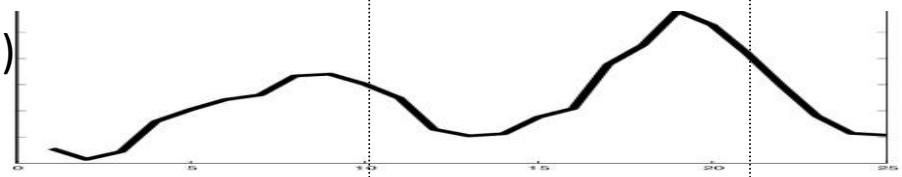
r = .98

Voxel 2 (deactivated)



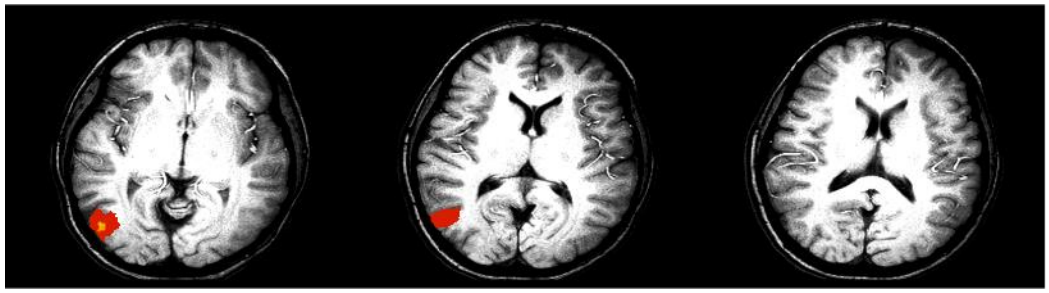
r = -.98

Voxel 3 (active before events)

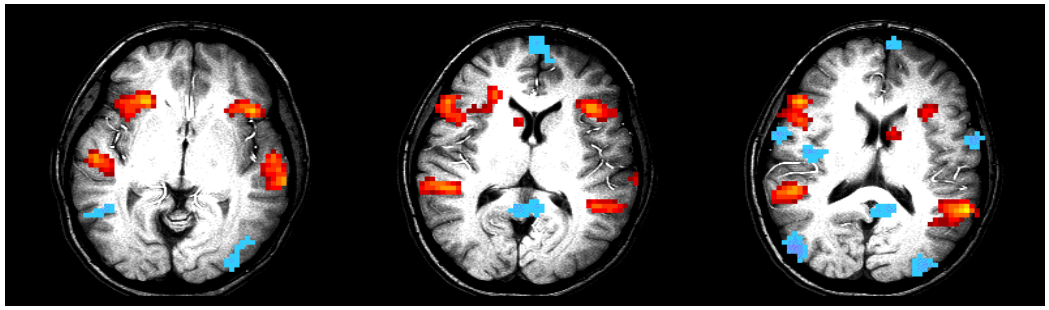


r = .6 (.98)

Results of event-related analysis:



Brain region activated prior to the hallucination

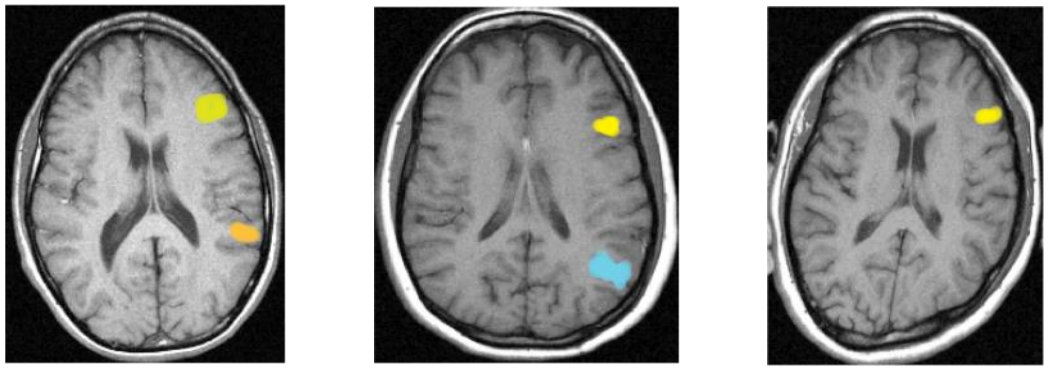


Brain regions activated/deactivated at the time of the hallucination

- But how do these brain areas interact?
- ➔ Functional/effective connectivity research

Individual subject analyses

- Activation
  - block design: e.g., brain areas activated during a task
  - event-related: e.g., brain areas activated before, during and after a given event
- Functional connectivity
  - how different brain areas are synchronized with each other

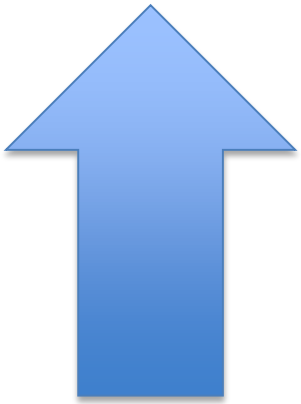


Individual subject result maps

Functional MRI Data Analyses

Second level => Group Analyses

- Do healthy people have a certain pattern of brain activity?
- Do patients have a different pattern?
- Are brain patterns correlated with personal variables?



First level => Individual Subject Analyses

- What brain areas activate during a task or event?
- How are brain areas synchronized?

Group level analyses

1. register the data from all subjects to a common space
2. For each voxel in the common space, do statistics across subjects
3. Correct for multiple comparisons

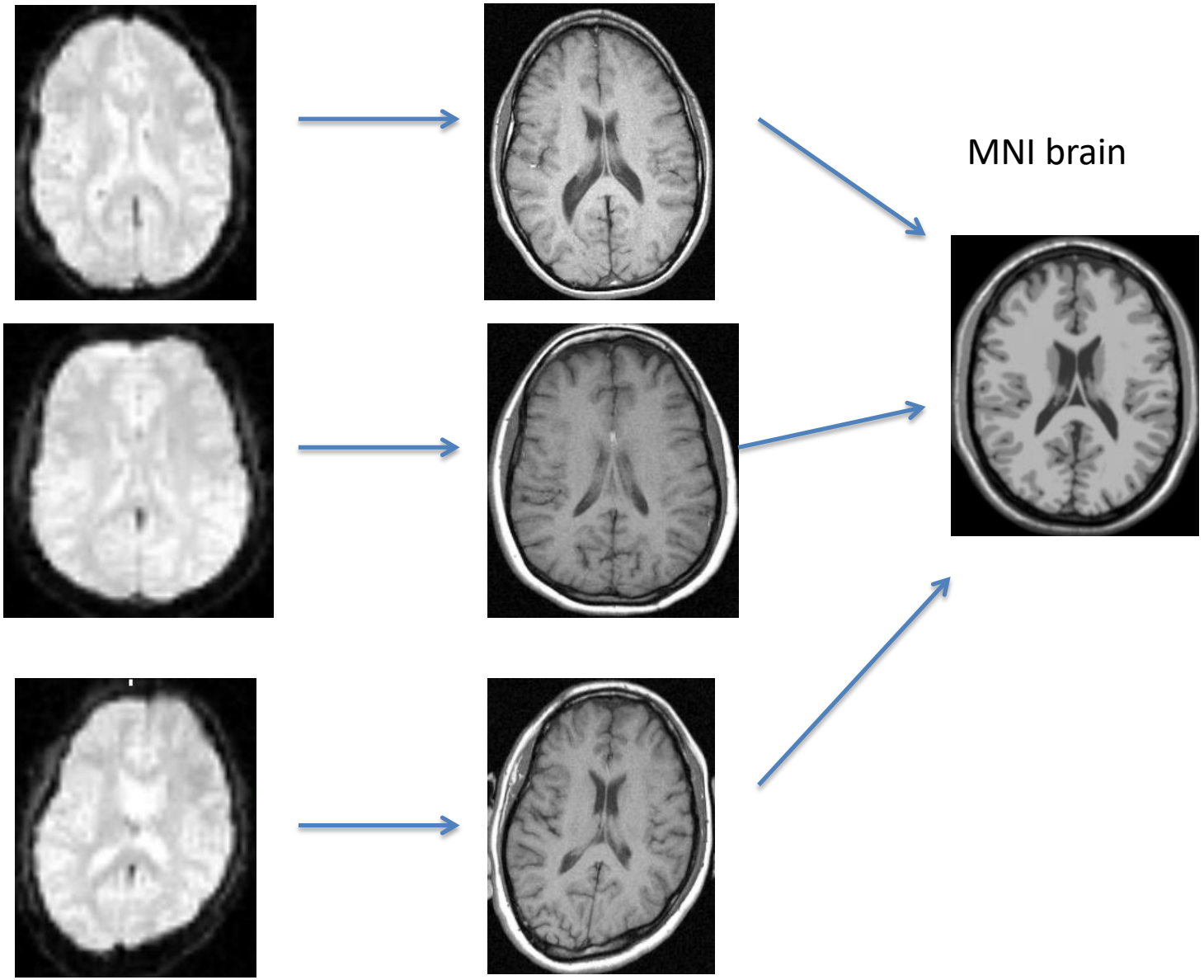
Registering data

Subject 1

Subject 2

Subject 2

MNI brain



Voxel-wise statistics

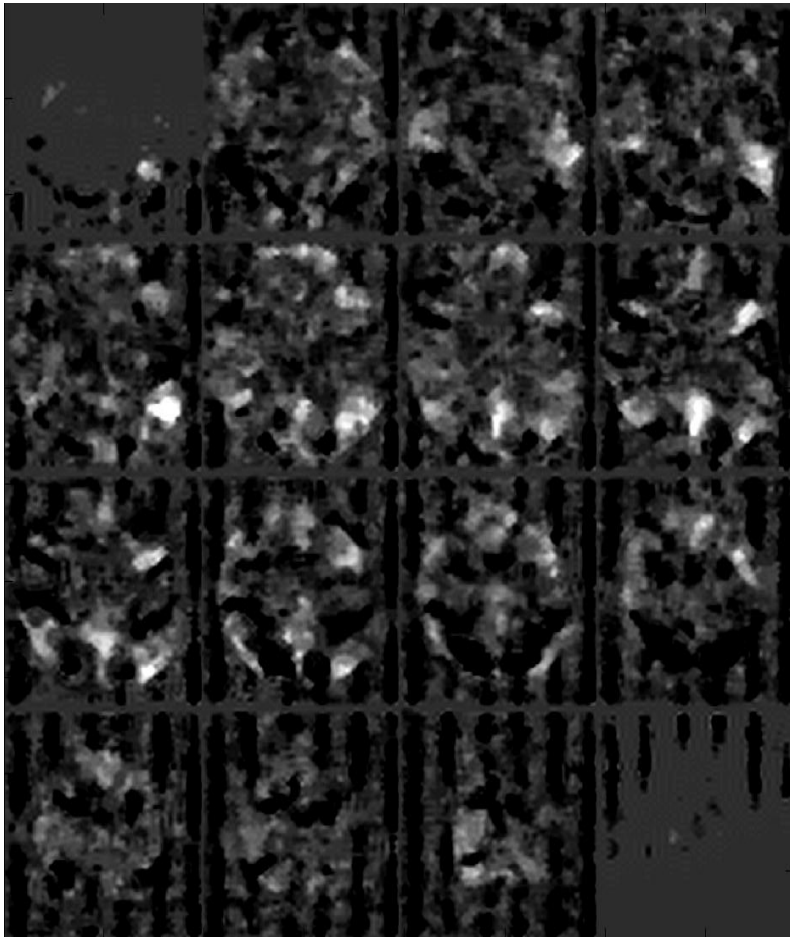
Once the functional data from each subject has been transformed into the common space, the same voxel should correspond to the same part of the brain across subjects.

Perform voxel-wise statistics:

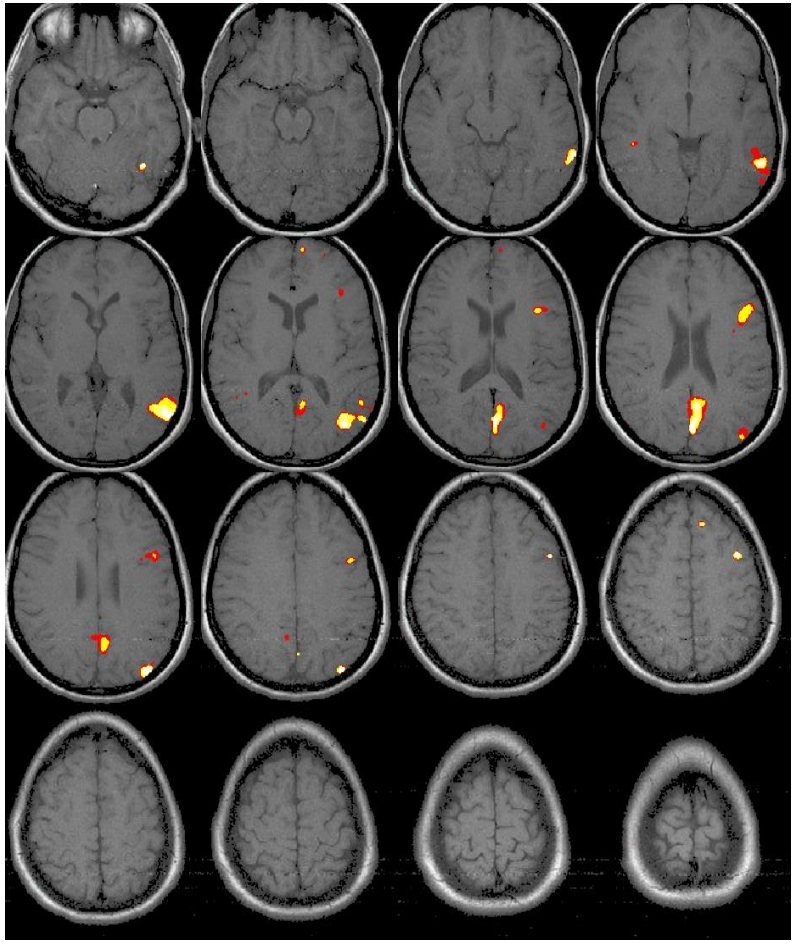
*2 sample t-test for each voxel to compare across groups

*voxel-wise correlation with behavioral measure

Results



- T-maps



- t-maps, thresholded and overlaid on anatomic scans

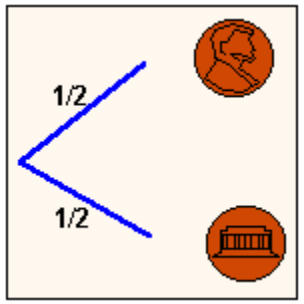
The problem of multiple comparisons

Classical Hypothesis Testing

- Assume the Null Hypothesis, H_0
- Compute test statistic, e.g. t-test = 2 with 48 d.f.
- Convert test statistic to p-value – probability of getting a t-value that large if there were no real effect
- If p-value is very low, *reject* H_0

Slide c/o Dr. Joe Maisog, Georgetown University

Flip 5 Heads in a Row?



P(5 heads in a row | fair coin) = $0.5^5 \cong$

0.03

Or 1 in 32 probability

Slide c/o Dr. Joe Maisog, Georgetown University

Null Hypothesis: Coin Toss

- H_0: “This is a fair coin.”
- If we flip 5 heads in a row, we’d strongly consider rejecting H_0

Norman Rockwell, “The Coin Toss”

Slide c/o Dr. Joe Maisog, Georgetown University

Football Stadium of Coin Flippers

- 70,000 People

Assuming all coins are fair:

Expect ~2000 people to flip 5 heads in a row (1/32 of the crowd).

Slide c/o Dr. Joe Maisog, Georgetown University

The problem of multiple comparisons

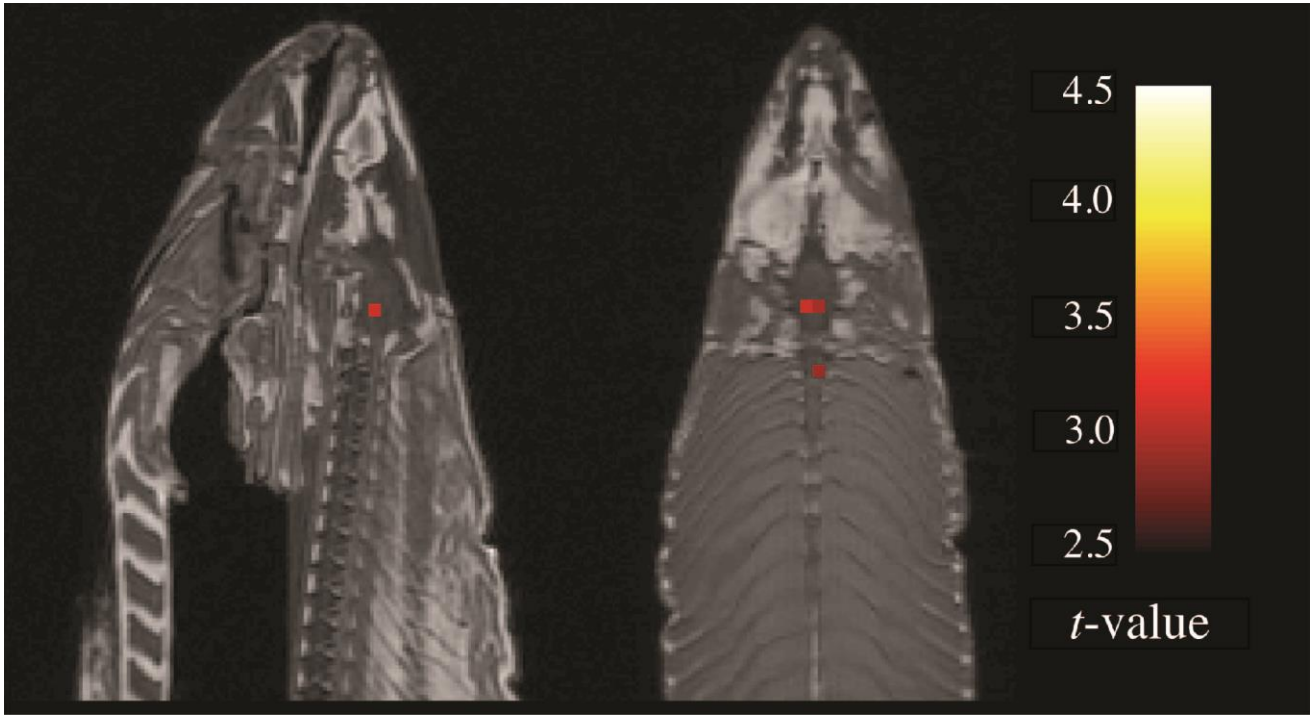
- Approximately 50,000 voxels in the brain
- If you do a t-test at each voxel and threshold at $p<0.05$ then by chance expect 5 in every 100 voxels will give false positives

- In data set with no real effect, 2500 false positives expected by chance!

MUST CORRECT FOR MULTIPLE COMPARISONS

Journal of Serendipitous and Unexpected Results

Neural Correlates of Interspecies Perspective Taking in the Post-Mortem Atlantic Salmon: An Argument For Proper Multiple Comparisons Correction



Activation in a dead fish!

Methods for correcting for multiple comparisons

1. Bonferroni correction

2. Cluster correction

3. Gaussian Random Field Theory

4. False Discovery Rate (FDR) correction

Bonferroni correction:

Ensure that your probability of getting a false positive is less than 5% by requiring significance level of $p<0.05*(1/num\_voxels) \sim 10\text{-}6$

Pros

- Yes, this will reduce your chance of getting a false positive to 5%

Cons

- Your chance of seeing real effects will be also be extremely slim!

=> Outrageous Type II error (will not find any of the real effects)

- Almost never used in neuroimaging – too stringent

Cluster correction:

- Neural activity tends to occur across regions of cortex (not in tiny areas)
- take advantage of this to discern real activations!
- false positives randomly distributed, less likely to be clustered than real activations.
- Simulations determine the chance of finding a cluster of a certain size if the data were just random noise – must take into account the smoothness of the noise in the data

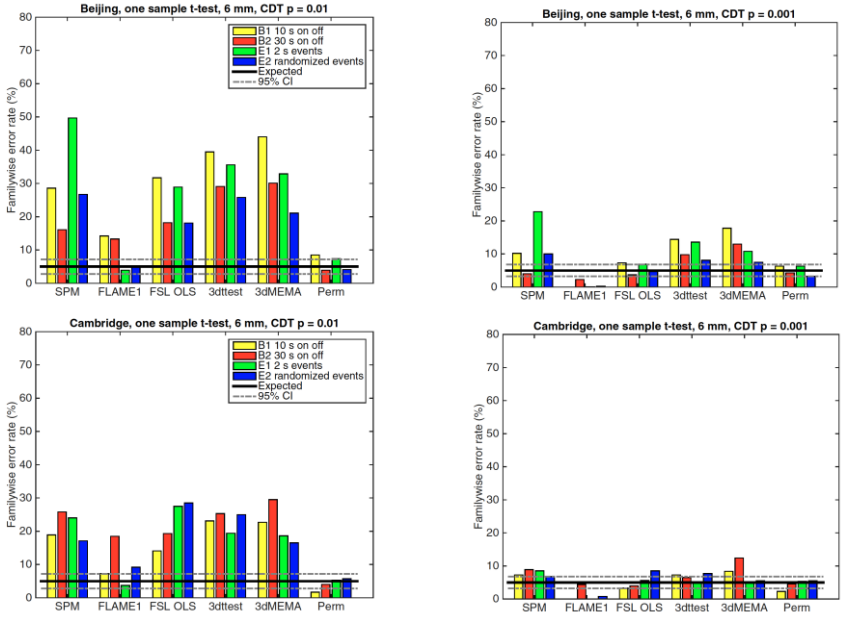
- Limitation: doesn't allow you to find small activations.

Gaussian random field theory:

Similar idea

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]



- Strict voxel wise threshold MUCH better performance, but not perfect
- Permutation based cluster correction good
- New software designed to better model distributions also now available
- AFNI bug identified and corrected

False Discovery Rate (FDR)

Family-wise error correction (e.g., Bonferroni, cluster correction) ensures your chance of getting ANY error is less than 5%

In contrast, FDR allows some of your findings to be false positives, but limits the false positive voxels to be, on average, 5% of the voxels above threshold.

In other words, almost all of what you show is correct, but a little bit is wrong (who knows which bit that is)

A totally different approach: ROI analyses

But what if I was only interested in two brain areas? Why do I have to correct for all the voxels all over the whole brain?

You don't!!

Multiple comparison correction is only necessary for exploratory, whole-brain analyses. If you have **a-priori** hypotheses about specific regions, you can get much more power using ROI analyses.

ROI analyses

Only need to correct for the number of ROIs you had a-priori hypotheses about – MUCH, much more power.

The issue of whether you really had a-priori hypotheses is key.

**always a temptation to claim that what you see in the uncorrected whole brain map was an a-prior hypothesis.

Summary: Methods for correcting for mult comparisons

1. **Bonferroni correction:** require significance level of $p<0.05*(1/num\_voxels) \sim 10\text{-}6$
- Outrageous Type II error (will not find any of the real effects)
- Almost never used in neuroimaging – too stringent

2. False Discovery Rate (FDR) correction: instead of setting threshold so that expectation of any false positive is less than 5%, set threshold so that 5% of all the voxels surviving the threshold are false positives.

3. Cluster correction:
- false positives randomly distributed, less likely to be clustered than real activations.
- Simulations determine the chance of finding a cluster of a certain size
- Limitation: doesn't allow you to find small activations.

4. Gaussian Random Field Theory:
- Smoothness affects chance of finding statistical patterns/clusters
- Limitation: extra smoothing required (loss of resolution)
- can only be used for statistics (r,f, t) where random fields have been mapped

OR, avoiding making so many mult comparisons!

Example synopsis

Michelle Hampson

Association of Marijuana Use with Blunted Nucleus Accumbens Response to Reward Anticipation (Martz et al)

Goal: To determine if marijuana use affects subsequent nucleus accumbens (NAcc) activation during anticipation of reward.

Modality: fMRI

Drug: Marijuana

Signal being measured: BOLD / T2*-weighted

Finding: Marijuana use associated with decreased activation in the NAcc in reward anticipation phase of monetary incentive task years later.

What was computed from imaging data of each subject: NAcc activation during reward anticipation (relative to neutral anticipation) in a monetary incentive task

How did they handle multiple comparisons: ROI analysis employed, so no whole brain correction for multiple comparisons required in imaging analysis. Did their hypothesis about the ROI involve laterality differences? Did they plan to treat activations to large and small gains the same? Number of associations in longitudinal analysis not corrected for.

Experimental design: Longitudinal study with event-related activation paradigm used for the imaging portion.

Assumptions: Monetary reward is assumed to be non-drug related.

Errors in design or interpretation: They interpret their results as arising from a causal relationship between marijuana use and later blunted NAcc activation. However, there could be a third variable driving both (e.g. low SES, chronic anxiety, or sleep deprivation could cause people to take more marijuana and also to develop blunted NAcc activity over time). Did not control for risk level, anxiety, SES, etc.

Biases in design or analyses: Study sample predominately white and high risk (and limited to young adults).

What questions do you have? What don't you understand? They interpret the findings as a causal relationship of marijuana use on NAcc activation, but do not directly discuss the fact that the blunting takes years to show up after the marijuana use. They suggest heightened NAcc could be associated with increased risk of drug use in Introduction, but then suggest in conclusions that the blunted NAcc activation could increase risk of future drug use.

Next logical question to ask/experiment to do:
Do those with blunted NAcc responses go on to develop addiction?

Commented [HM1]: Skip tracer question for fMRI studies

Commented [MH2]: Describe what kind of analysis was used at the first level