Individual and group level analysis

張葶葶
國立政治大學心理系
國立政治大學心腦學中心
台灣心智科學腦造影中心

Overview

- How to analyze fMRI data

  - General Linear Modeling (GLM)

  - Individual and group level

  - Multiple comparison correction

- A quick overview of using SPM to implement

  individual and group level analysis

## Inferential Statistics
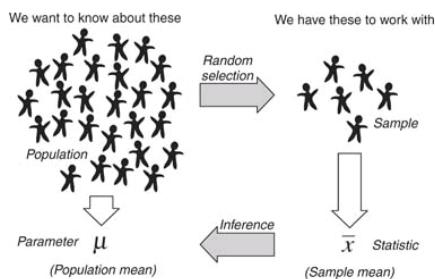
Hypothesis testing

- H_0 : condition 1 = condition 2
- H_1: condition1 ≠condition 2

What does this mean in fMRI data?
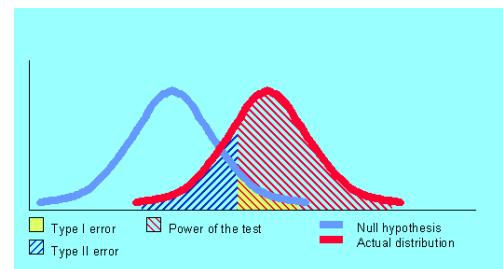
### Consider a very simple fMRI experiment

One session

Passive word listening

versus rest

7 cycles of rest and listening

Stimulus function
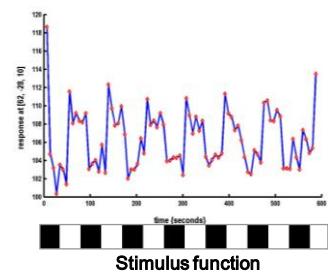
Question: Is there a change in the BOLD response between listening and rest?

Modelling the measured data

Why? Make inferences about effects of interest (listening > rest is real?)

How?
1. Decompose data into effects (contrast map) and error (sample errors etc.)
2. Form statistic (t map) using estimates of effects and error

stimulus function

data → linear model → effects estimate / error estimate → statistic

Each voxel is analyzed separately.

Each voxel presents a time-series data.
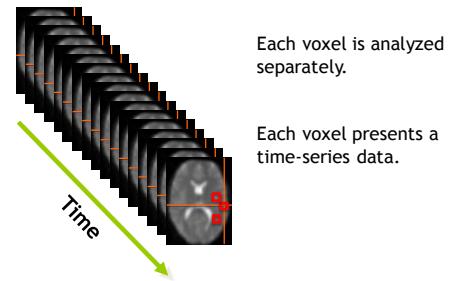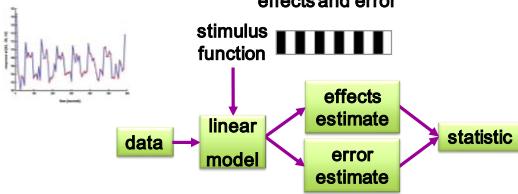
Time

Voxel-wise time series analysis

Time

single voxel time series

Model Specification
Parameter estimation
Hypothesis
Statistic

BOLD signal

SPM
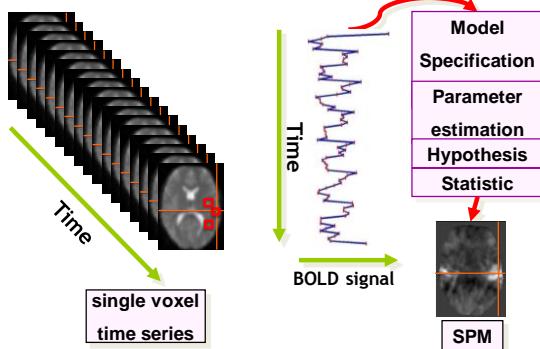
General Linear Model

$$y = x_1 b_1 + e$$

Measured Data   Design Model   Amplitude (solve for)   Noise

Cf. Boynton et al., 1996

Model specification: Single voxel regression model

Time

BOLD signal x_1 e

$= \beta_1$ +   error

$$y = x_1 b_1 + e$$

Time   BOLD signal x_1 $= \beta_1$

Time   BOLD signal x_1 $= \beta_1$
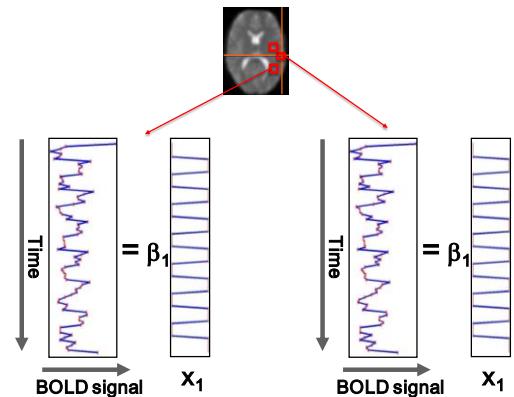
Parameter estimation

$y = X\beta + e$

Objective: estimate parameters to minimize $\sum_{t=1}^{N} e_t^2$

## Consider this example
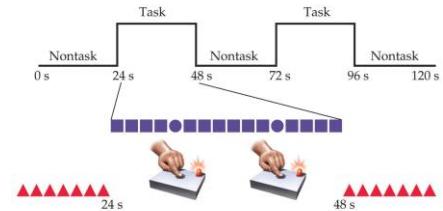
Design Matrix
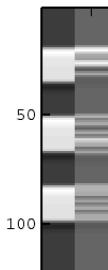
The 1st regressor is block
The 2nd regressor is trial type

Task and Nuisance Regressors
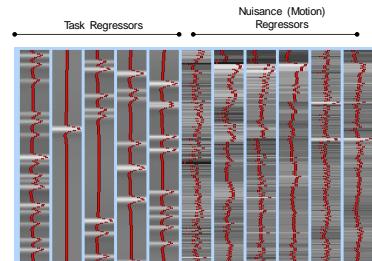
There is one problem of this model.

BOLD responses have a delayed and dispersed form.


HRF
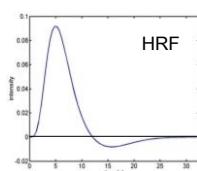
Solution: Convolution model

$$f \otimes g(t) = \int_0^t f(\tau)g(t-\tau)d\tau$$
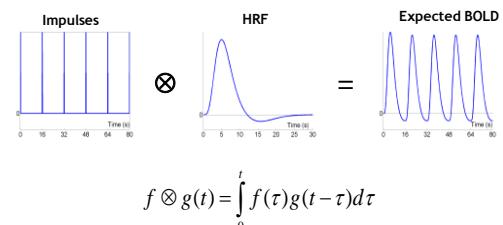
expected BOLD response
= input function ⊗ impulse response function (HRF)

Convolution model of the BOLD response

Convolve stimulus function with a canonical hemodynamic response function (HRF):

⊗ HRF

$$f \otimes g(t) = \int_0^t f(\tau) g(t-\tau) d\tau$$


Assumptions of GLM

- Same design matrix throughout the brain

- Homoscedastic vs. heteroscedastic

- All voxels represent independent statistical test
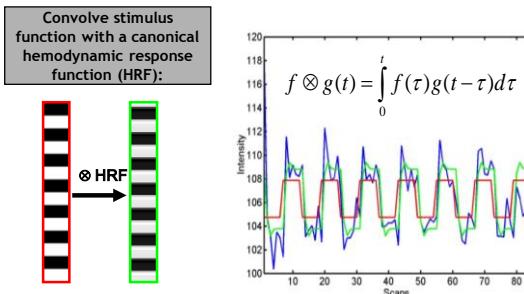
In the SPM interface

The following images are created each time an analysis is performed

- beta.img: images of estimated regression coefficients (parameter estimate).

- con.img: contrast values between two beta images.

- spmT.img: T-value of the contrast image.

Single subject results

Not as pretty as the data you often seen in fMRI papers.

Typically you will need ~20 subjects to obtain meaningful results.

Group Level Analysis

Fixed Effects

- Fixed-effects Model
  - Assumes that effect is constant ("fixed") in the population
  - Uses data from all subjects to construct statistical test
  - Allows inference to subject sample

Random Effects

- Random-effects Model
  - Assumes that effect varies across the population
  - Accounts for inter-subject variance in analyses
  - Allows inferences to population from which subjects are drawn
  - Especially important for group comparisons
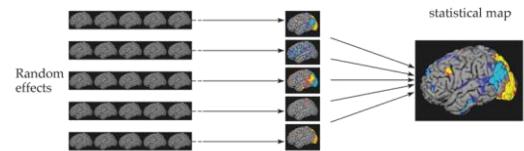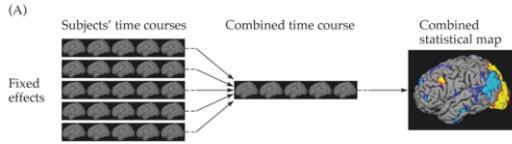
Subject 1

For voxel v in the brain
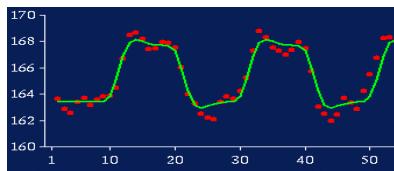
Effect size, c ~ 4

Subject 3

For voxel v in the brain
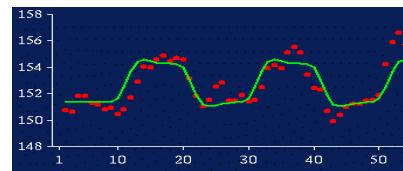
Effect size, c ~ 2

Subject 12

For voxel v in the brain
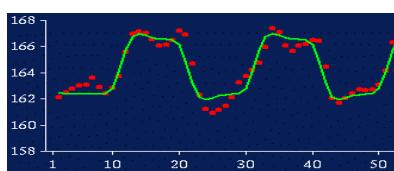
Effect size, c ~ 4

Random Effects Analysis

For group of N=12 subjects effect sizes are

c= [3, 4, 2, 1, 1, 2, 3, 3, 3, 2, 4, 4]

Group effect (mean), m=2.67
Between subject variability (stand dev), s_b =1.07

This is called a Random Effects Analysis (RFX) because we are comparing the group effect to the between-subject variability.

Subject 1

For voxel v in the brain
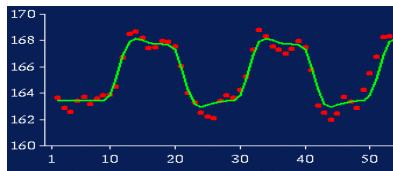
Within subject variability, s_w~0.9

Subject 3

For voxel v in the brain
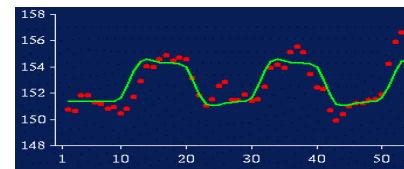
Within subject variability, s_w~1.5

Subject 12

For voxel v in the brain
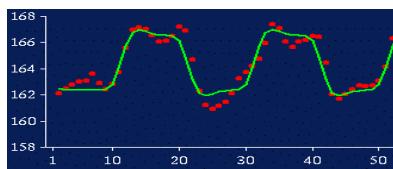
Within subject variability, s_w~1.1

Fixed Effects Analysis

Time series are effectively concatenated – as though we had one subject with N=50x12=600 scans.

s_w = [0.9, 1.2, 1.5, 0.5, 0.4, 0.7, 0.8, 2.1, 1.8, 0.8, 0.7, 1.1]

**Mean effect, m=2.67
Average within subject variability (stand dev), s_w =1.04**

Fixed-effects Analysis in SPM
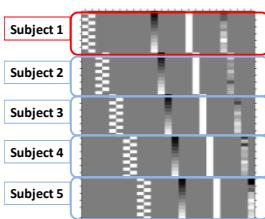
Multisubject 1ˢᵗ level :
5 subjects x 1 run each

Fixed-effects
- each subjects entered as separate sessions
- create contrast across all subjects
  c = [1 -1 1 -1 1 -1 1 -1 1 -1]
- perform one sample t-test

## Random-effects Analysis in SPM
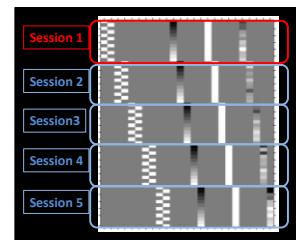
Random-effects
- 1ˢᵗ level design per subject
- generate contrast image per subject (con.*img)
- images MUST have same dimensions & voxel sizes
- con*.img for each subject entered in 2ⁿᵈ level analysis
- perform stats test at 2ⁿᵈ level

6

2nd Level Analysis

FIRST LEVEL (per person) | SECOND LEVEL
Data Design Contrast — Group analysis

$$t = \frac{c^T \hat{a}}{\sqrt{Var(c^T \hat{a})}}$$

SPM(t)

One-sample t-test @ 2nd level
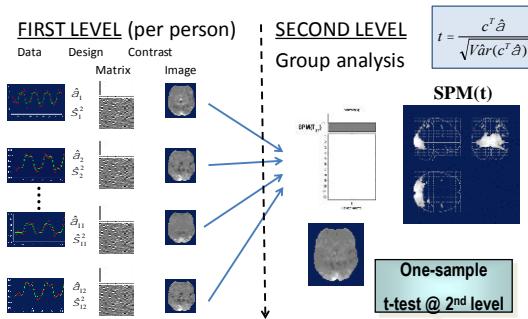
What statistics does SPM do?
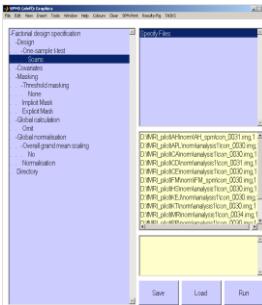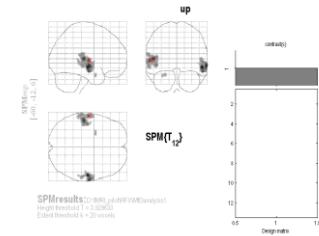
Specify 2nd level: One-sample t-test
Simplest example.
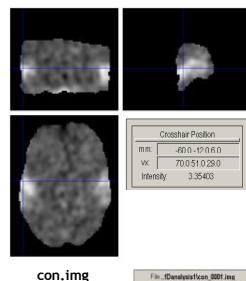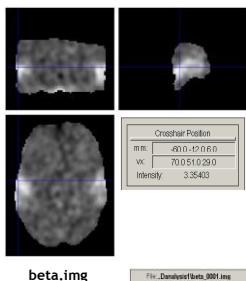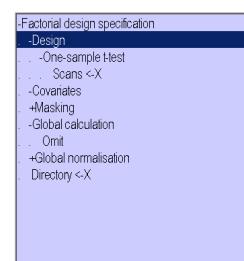
beta.img   con.img

Other tests

One sample t-test

with a covariate added.

Test correlations between task specific activations and some other measure

(age, performance, etc.).

Vectors added here.

Full factorial

Statistical thresholding

If we set the threshold too low, we will accuse innocent people (high rate of false alarms).

If threshold set too high, we fail to detect dopers (high rate of misses).
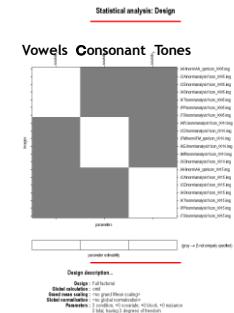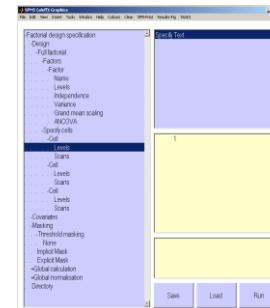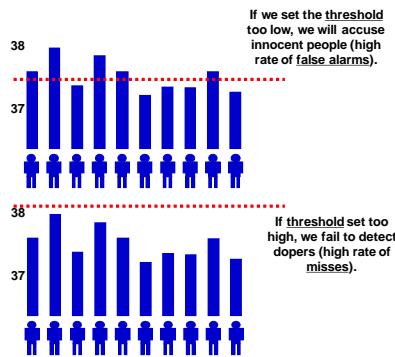
Statistical thresholding

- We need to choose a threshold that balances the benefits of finding effects with the cost of making false alarms.
- α is our statistical threshold: it measures our chance of Type I error.
  - A 5% alpha level (Z>1.64) means only 1/20 chance of false alarm (p < 0.05).
  - A 1% alpha level (Z>2.3) means only 1/100 chance of false alarm (p< 0.01).
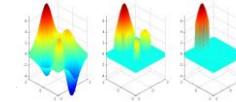
Fewer peaks survive as we apply a more stringent threshold.
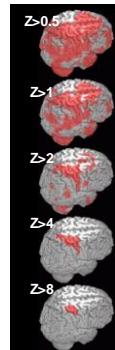
The Problem of Multiple Comparisons
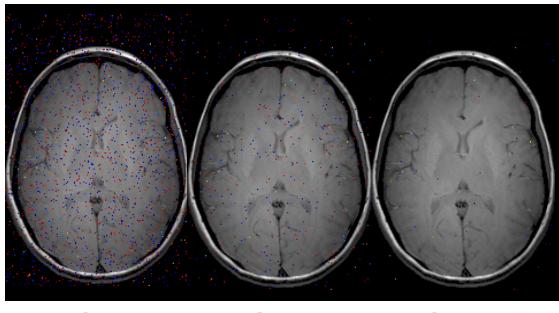
P < 0.05   P < 0.01   P < 0.001

Options for Multiple Comparisons

- Statistical Correction
  - Family-Wise Error Rate (FEW)
  - False Discovery Rate (FDR)
  - Random Field Theory (RFT)
- Cluster Analyses
- ROI Approaches

Bonferroni Correction

- Very severe correction
  - Results in very strict significance values
  - Typical brain may have up to ~30,000 functional voxels
    - Alpha .1, Corrected alpha ~ 0.000003
- Benefits
  - Controls for FWE.
- Problem
  - Very conservative = very little chance of detecting real effects

False Discovery Rate

- Controls the expected proportion of false positive values among suprathreshold values
 - Genovese, Lazar, and Nichols (2002, *NeuroImage*)

- Algorithm
  - $p1 <= p2 <= p3\ldots\ldots<= pV$
  - $Pi <= q/V$
  - E.g. q=.1 means control voxel does not exceed 10 out of 100 voxels

- Advantage
  - Less stringent

Small volume Comparison

- Only test a small proportion of voxels.
- Should only be done before analyses, based on strong a priori hypotheses.
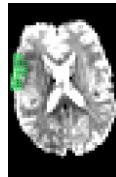
Random field theory

- Estimate the number of independent test

- Algorithm
  - $R = x * y * z / V^3$, v = smooth voxel size

- Recommendation: Use a combination of *voxel* and *cluster* correction methods
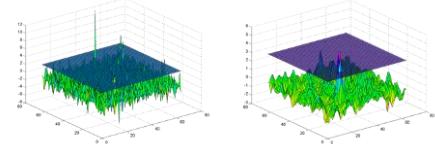
Cluster Analyses

- Adopting a minimum size of a cluster of active voxels to be labeled as significant

- Assumptions
  - Assumption I: Areas of true fMRI activity will typically extend over multiple voxels
  - Assumption II: The probability of observing an activation of a given voxel extent can be calculated

Two approaches of fMRI data analysis

A. Whole volume statistical approach

- **Requires no prior hypotheses about areas involved**
- **Includes entire brain**
- **Can lose spatial resolution with intersubject averaging**
- **Can produce meaningless "laundry lists of areas" that are difficult to interpret**
- **Depends highly on statistics and threshold selected**

B. Region of interest (ROI) approach

- **Gives you more statistical power because you do not have to correct for the number of comparisons**
- **Hypothesis-driven**
- **ROI is not smeared due to intersubject averaging**
- **Easy to analyze and interpret**
- **Neglects other areas which may play a fundamental role**

Anatomical ROI
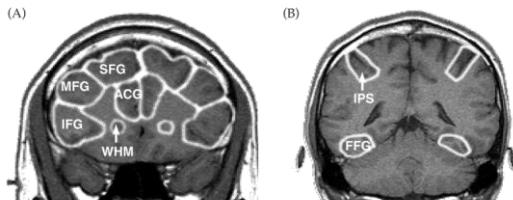
Fuctional ROI

- ROIs that were activated by a particular stimulus
- How to select
  - Functional localizer
  - Previous studies
  - meta-analysis
- Problem
  - Selection bias

Alternatives to voxelwise analysis

- Conventional fMRI statistics compute one statistical comparison per voxel.
  - Advantage: can discover effects anywhere in brain.
  - Disadvantage: low statistical power due to multiple comparisons.
- Small Volume Comparison: Only test a small proportion of voxels.
- Region of Interest: Pool data across anatomical region for single statistical test.

Example: how many comparisons on this slice?
- Voxelwise: 1600
- SVC: 57
- ROI: 1
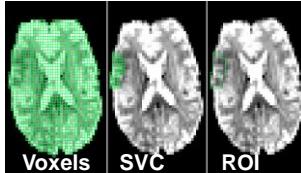
Group level analysis

- Many different ways of conducting group-level analysis
- Choice depends primarily on:
  1. Initial study design.
  2. Research questions
  3. Parsimonious models vs. more complex ones.