An Introduction to Group Analysis in fMRI

Alexis Roche, PhD

CEA, SHFJ, Orsay, France
roche@shfj.cea.fr
www.madic.org

Plan

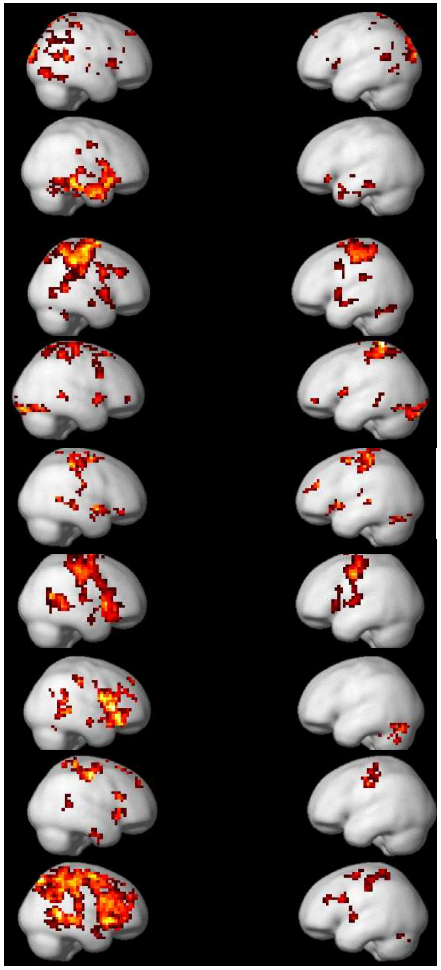
I. Introduction

III. The spatial normalization problem

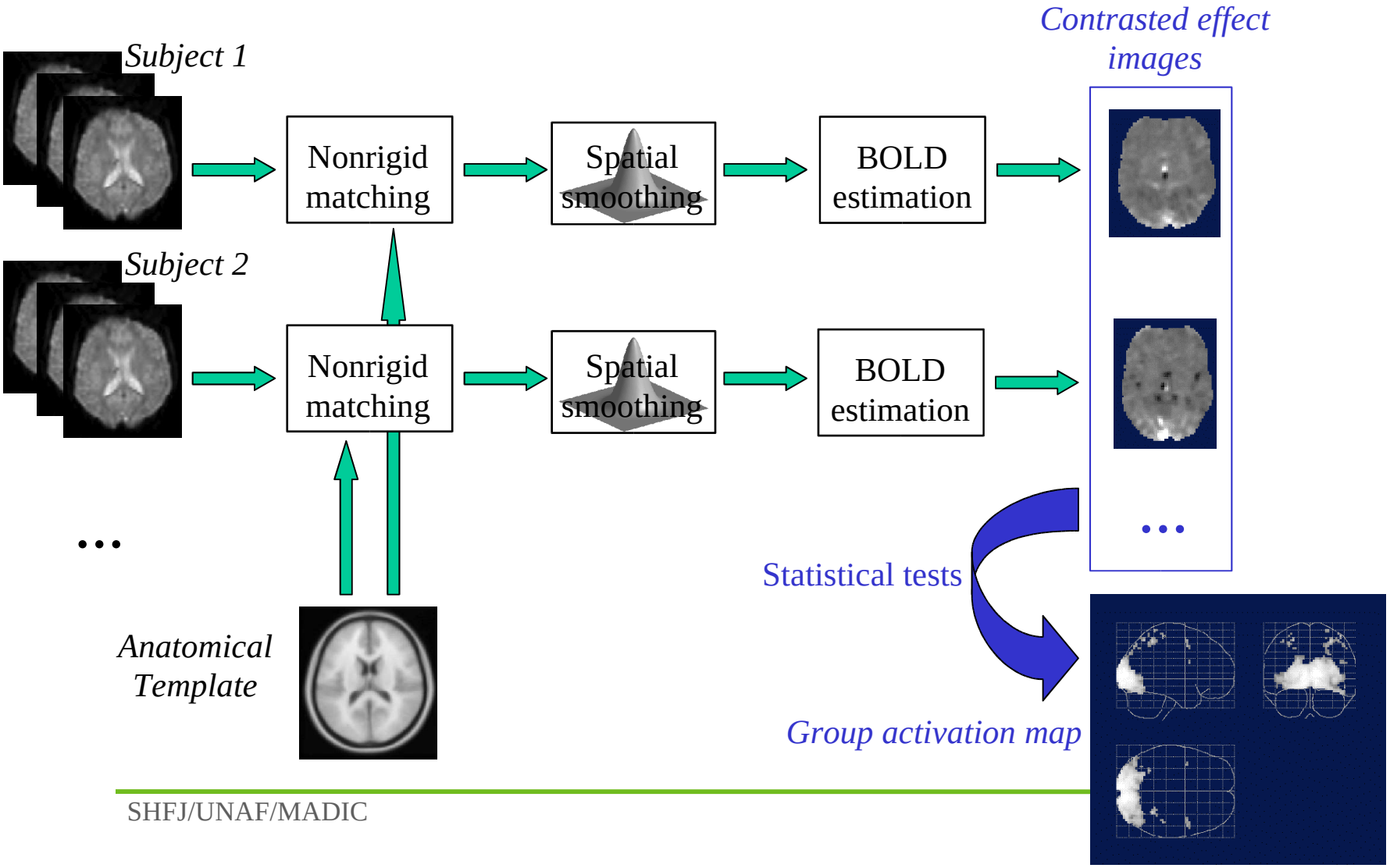
IV. The group inference problem

Variability of functional areas

- Intra-subject findings do not easily generalize to a population
- Inferring consistent activation patterns leads to questions of
  - Localization
  - Statistical significance



Classical (SPM) approach



Issues

- Successful group analysis relies on solving two problems
  - The spatial normalization problem
  - The inference problem

- To date, only partial solutions exist for each of those

Plan

I. Introduction

II. The spatial normalization problem

- Matching macro anatomy
  - Intensity-based registration
  - Sulcal matching
- Matching micro anatomy
- Matching functional activity

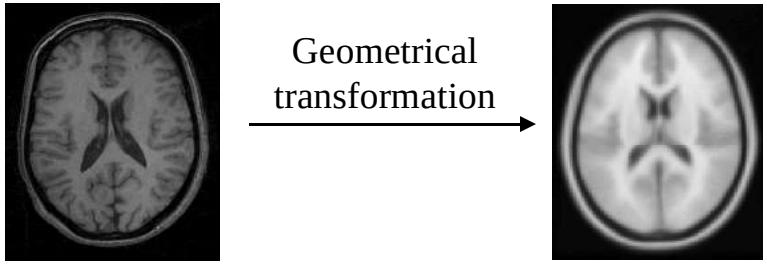
III. The group inference problem

Spatial normalization

- How to define homologous regions between two brains?
  - Match macro anatomy?
  - Match micro anatomy (cytoarchitecture)?
  - Match functional activity?

Matching macro anatomy: optical flow

- Classical solution: intensity-based matching between a structural image (usually T1) and a template (usually MNI)
  - "Optical flow" or "Morphing": doesn't use explicit landmarks
  - Amounts to maximizing the overlap between corresponding tissue classes (GM, WM, CSF)

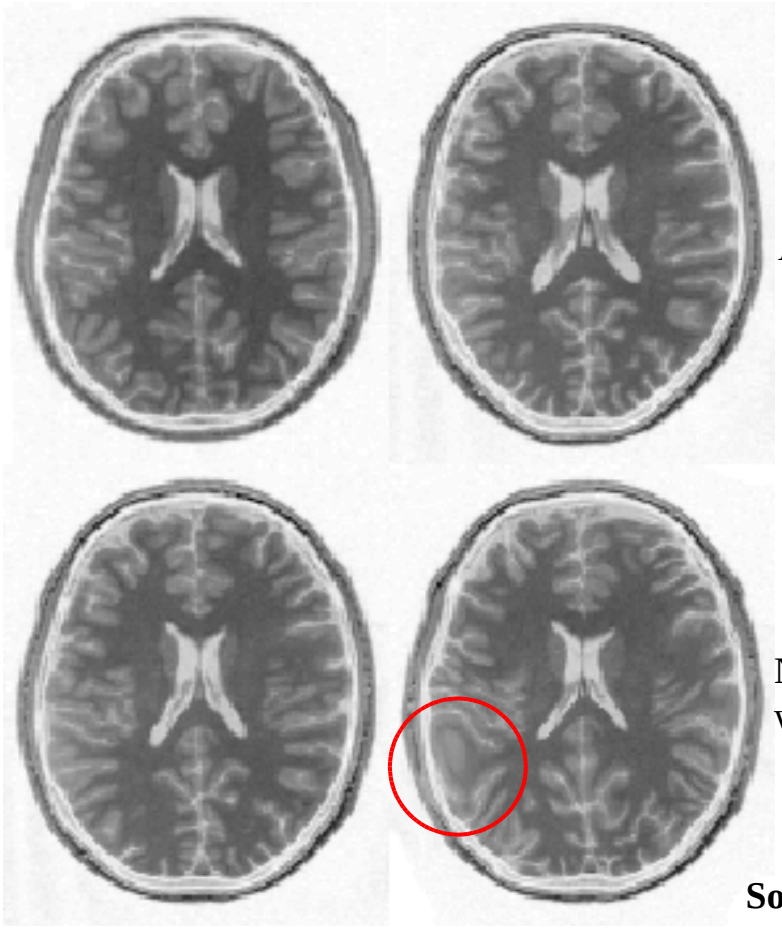


- Requires topology-preserving constraints on the transformation
  - Affine, Parametric (RBF, DCT basis), Regularized free-form, Diffeomorphic constraints…

Optical flow

- Importance of transformation regularity

Template image



Affine registration

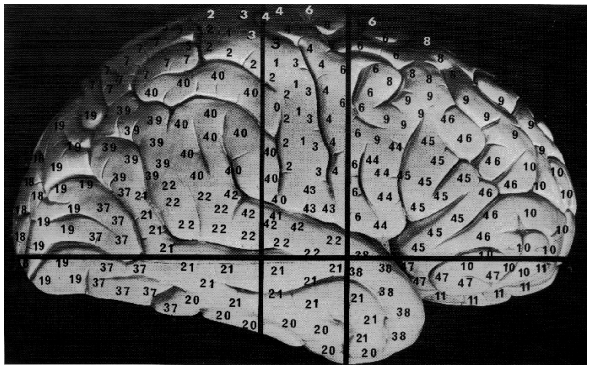
Nonrigid registration using regularization

Nonrigid registration without regularization

Source: John Ashburner

Optical flow

- Many registration packages are available on the web (SPM, FSL, AIR, ANIMAL, Image Registration Toolkit, …)
- Allows for positioning in the Talairach coordinate system



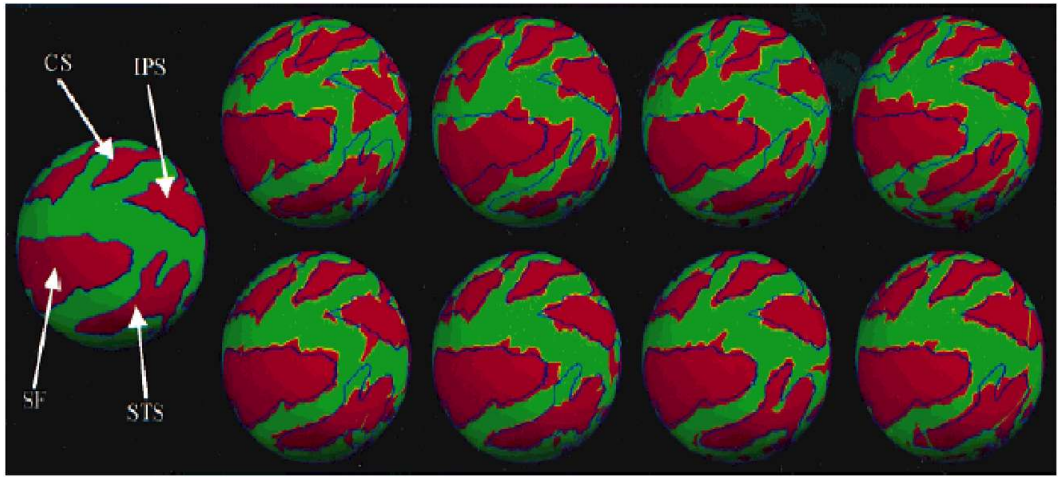
- Relies on the correspondence of function and gross anatomy
- OK to match deep brain structures, but poor matches between cortical surfaces (Hellier et al, 2001)

Sulcal vs. intensity-based matching

- Can we do any better while still relying on macro anatomy?
- Evidence from cytoarchitecture indicates that stable sulci may reflect functional subdivisions
- Explicit sulcal matching may therefore be more accurate than intensity-based matching

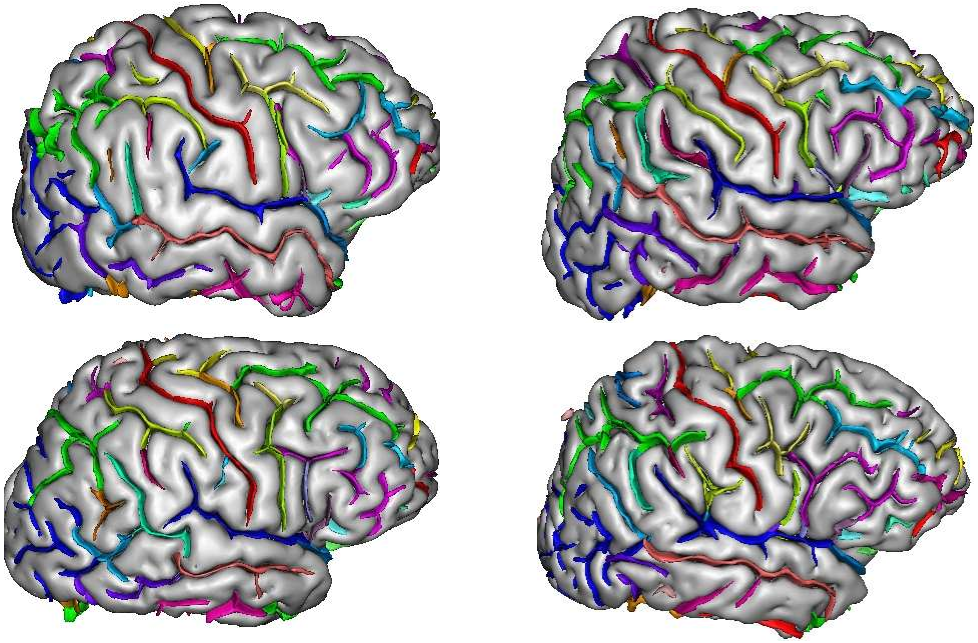
Sulcal matching

- (Fischl et al, 1999)
  - Transform the cortical surface into a sphere (cortex inflation)
  - Assign to each spherical coordinate a gray value encoding for local convexity
  - Perform 2D intensity-based registration between spherical images



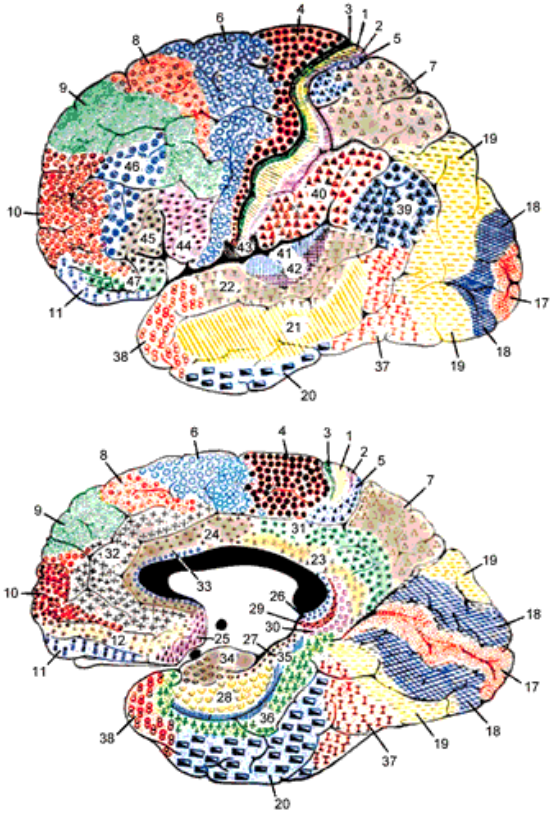
Sulcal matching

- (Cachier et al, 2001; Rivière et al, 2002)
  - Automatic extraction and labeling of the cortical sulci
  - Mixed feature-based and intensity-based nonrigid registration



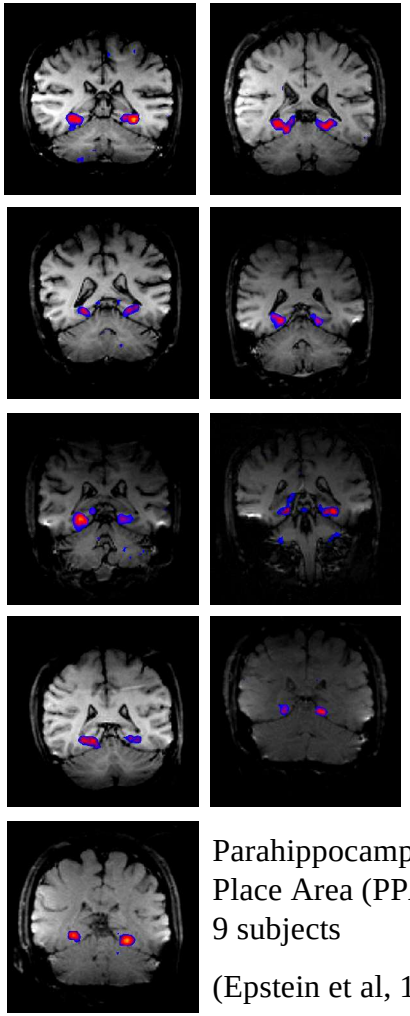
Matching micro anatomy

- Cytoarchitecture reveals functionally homogeneous brain regions (Brodmann areas)
- Despite high variability in location and shape, those are natural functional landmarks
- Not (yet) available in vivo
  - Perhaps with high resolution Diffusion Tensor Imaging…



Matching functional activity

- Idea:
  - Identify functional regions from fMRI itself using a preliminary “localizer” protocol
  - Restrict the analysis of subsequent experiments to those regions
- Useful when hypothesizing activity within a particular brain area
- Activation delineation may require user interaction
  - *MarsBAR* SPM toolbox



Parahippocampal Place Area (PPA) in 9 subjects

(Epstein et al, 1999)

Spatial normalization: summary

- Current “whole-brain” normalization techniques are prone to unquantifiable inaccuracy
  - Spatial mismatch results in blurring the average activation map
  - The residual spatial variability is often accounted for by spatially smoothing each subject’s fMRI data
- Sulcal matching and functional matching as emerging alternatives
- The best normalization scheme depends on the questions asked to the data

Plan

I. Introduction

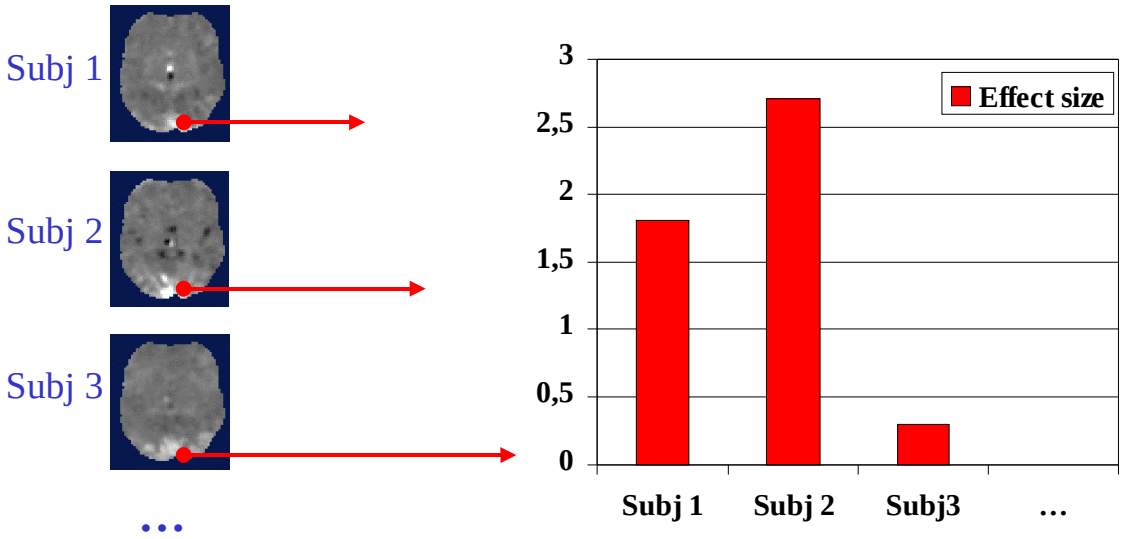
II. The spatial normalization problem

III. The group inference problem

- Fixed-effect vs. random-effect analysis
- Random-effect models
  - Simplistic model
  - Multilevel models
- Robust random-effect analysis
  - Inter-subject distance measures
  - Nonparametric tests

Context

- After spatial normalization, each voxel in the reference space is associated with a list of estimated effects



- Samples are generally small (a tenth of subjects)

Fixed effect vs. Random effect

- Do we want to make inferences about:
  - The *particular group* we sampled?
  - The *population* they were sampled from?

- The first approach refers to fixed-effect analysis (FFX)
  - “I can see this effect in this cohort” (Tom Nichols)

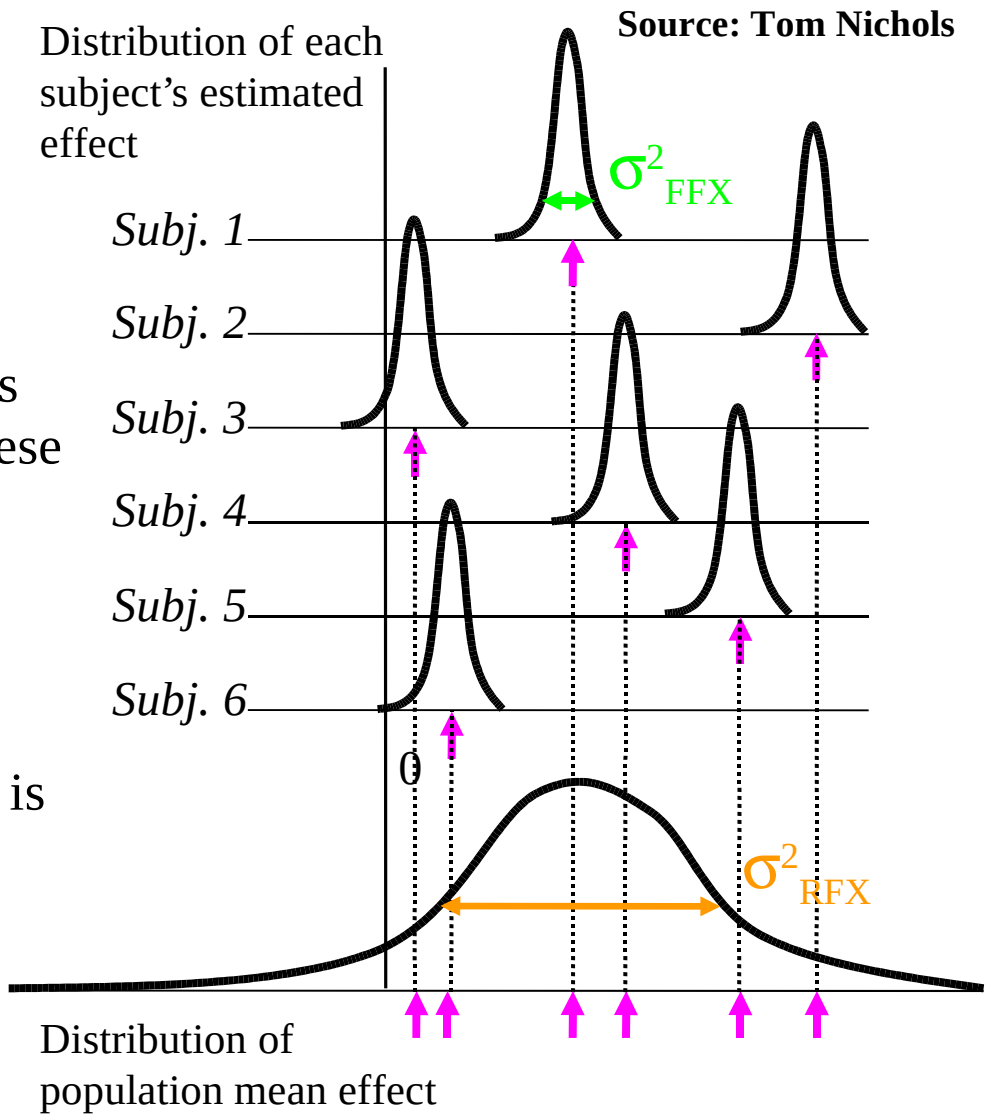
- The second approach refers to random-effect analysis (RFX)
  - “Would another cohort be sampled, I would see the same effect”

Fixed effect vs. Random effect

- Observed effects are affected by two distinct variability sources
  - *Intra-subject variability*: there is some uncertainty on each subject's estimated effect due to "noise" in the fMRI data
  - *Inter-subject variability*: different subjects may have intrinsically different effects
- FFX is concerned with *intra*-subject variability only, while RFX is concerned with both

Example

- FFX
  - The *empirical mean* effect is significantly positive (all these subjects are!)
- RFX
  - The *population mean* effect is not so significantly positive



Simple RFX model (SPM'99)

- Assume:
  - (A1) The distribution of each subject's estimated effect is normal with constant variance
  - (A2) The population is normal
  - Amounts to assuming all estimated effects are drawn from *the same normal distribution*

- The *t*-test is then a valid procedure to assess significance of the population mean

$$t = \frac{\bar{\beta}}{\bar{\sigma}_{RFX} / \sqrt{n}}, \quad p(t|H_0) = T_{n-1}$$

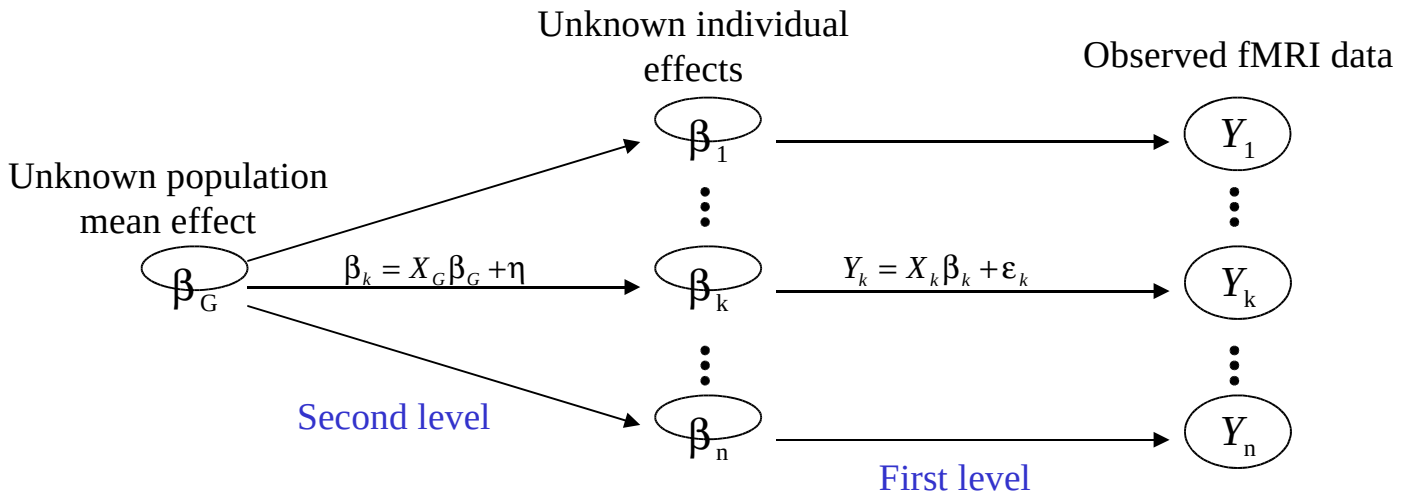
Simple RFX model (SPM'99)

- Issues:
  - Assumption (A1) cannot be controlled (σ_{FFX} may not be constant across subjects)
  - Population normality (A2) is questionable

- This RFX approach tends to be over-conservative
  - Multilevel models try to overcome the first assumption
  - Nonparametric techniques try to overcome both assumptions

Multilevel modeling for RFX

- Integrated generative model for both intra-subject (first-level) and inter-subject (second-level) analyses



- General inference problem (in Bayesian terms): compute the posterior pdf of the population effect, $p(\beta_G|Y_1, Y_2,\ldots, Y_n)$

Multilevel modeling for RFX

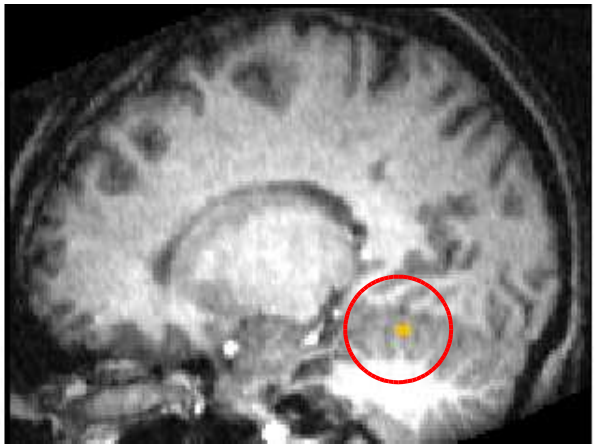
- No analytical solution even when all random processes involved are assumed Gaussian
  - Mainly because first-level variance components are unknown
- Analytical approximation by running first- and second-level analyses sequentially (Beckmann et al, 2003)
- More accurate solvers are iterative techniques
  - Expectation-Maximization algorithm (Friston et al, 2002)
  - Stochastic sampling (Woolrich et al, 2004)

Simple RFX vs. multilevel RFX

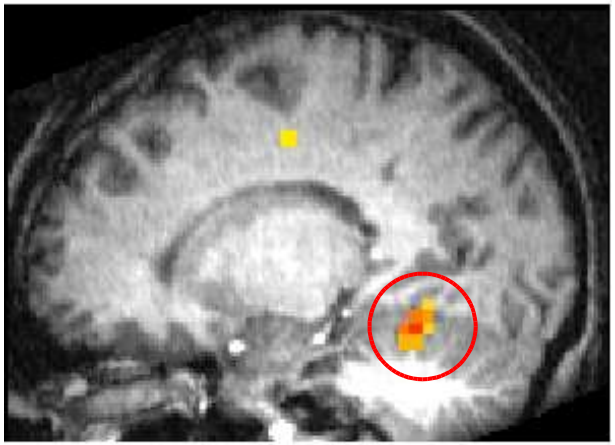
- Simple RFX is typically more conservative than multilevel RFX

  - Motor contrast: “left click vs. right click”
  - 21 subjects
  - Uncorrected threshold ($P=10^{-3}$)

Activation found in the ipsi-lateral cerebellum



Simple RFX



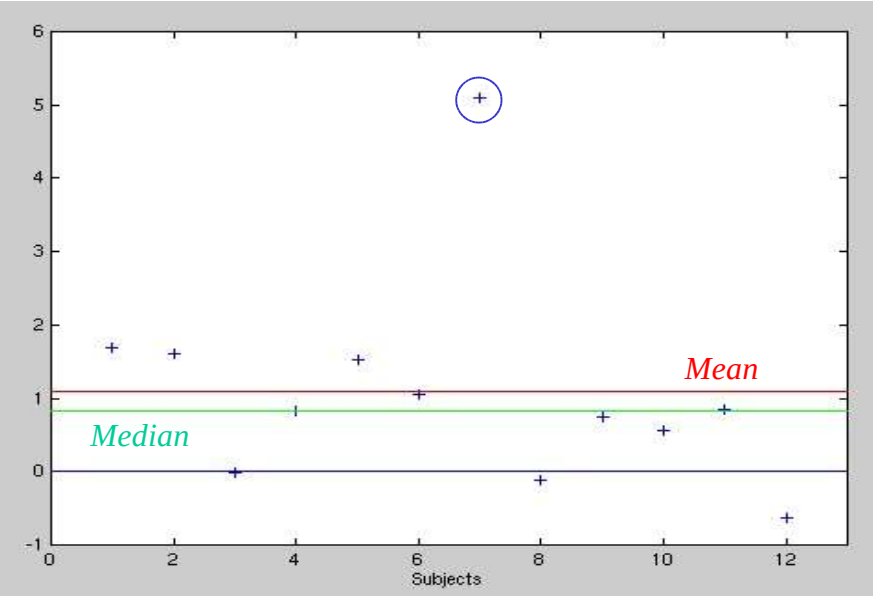
Multilevel RFX

Source: Bertrand Thirion

Robustness issues in RFX

- Estimated effects often reveal some “outliers”, possibly due to:
  - Poor first-level estimation: noisy data, subject’s motion, …
  - Spatial normalization errors
  - Intrinsically heterogeneous population

Estimated effects in one voxel



Inter-subject distances

- Outliers may also be appreciated by assessing how different subjects are in terms of their global activation patterns

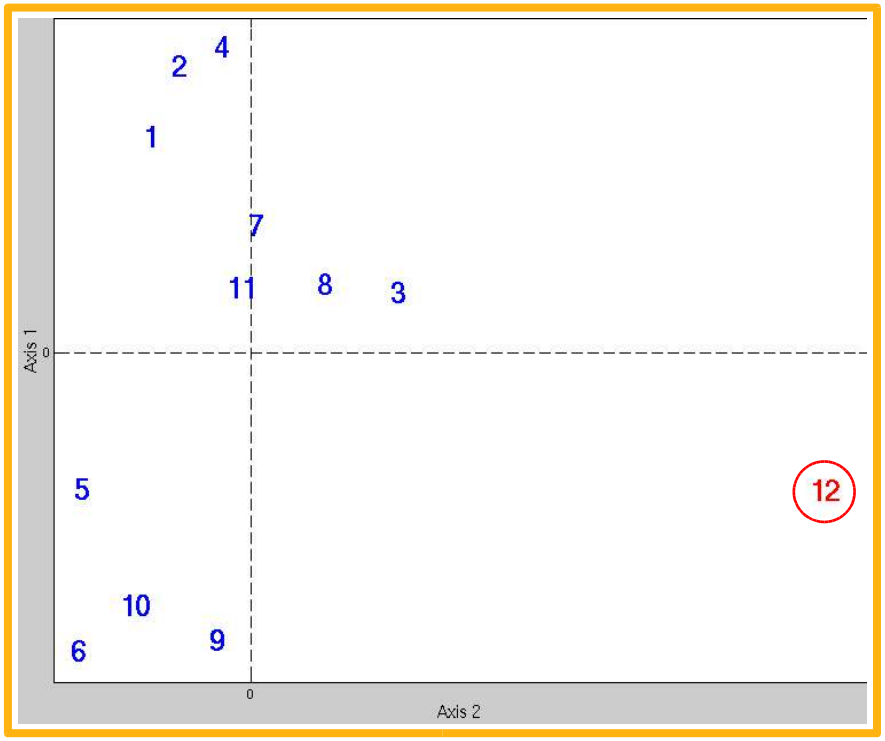
- Distance measure using the RV-coefficient between two subjects' effect images (Kherif et al, 2003)

$$RV(\beta_i, \beta_j) = \frac{\text{trace}(\beta_i \beta_j^t)}{\sqrt{\text{trace}(\beta_i \beta_i^t) \times \text{trace}(\beta_j \beta_j^t)}}$$

- Implemented in the *DISTANCE* SPM toolbox

Inter-subject distances

- Distance matrix visualized in a plan using multi-dimensional scaling
- Outliers diagnosis (Cook test)
- Outlier detection is a frequent situation in fMRI group studies (Mériaux et al, *in prep*.)



Source: Sébastien Mériaux

Nonparametric tests: motivations

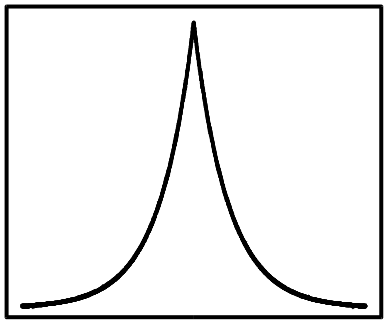
- Both simple and multilevel RFX models rely on *homogeneity*
  - Assume the estimated effects are normally distributed
  - Multilevel RFX only accounts for possibly different first-level variances
- Deviations from normality (“outliers”) cause inexact P-values
  - Increased false negative risk *and* false positive risk
- It is not statistically valid to force normality by rejecting outliers *a posteriori*

Nonparametric tests: motivations

- We lack models of population effect distribution
  - Experimental model selection would necessitate fMRI studies with hundreds of subjects…
- Nonparametric tests rely on vague distributional assumptions
  - Generally less sensitive than parametric tests, but more robust
  - In other words: “worse if parametric assumptions hold, better otherwise”

The Laplace test

- Assume a Laplacian distribution form



- Derive the associated likelihood ratio statistic

$$t = \frac{\sum_i |\beta_i|}{\sum_i |\beta_i - med\{\beta\}|}$$

- Null distribution tabulated from Monte Carlo simulations

Simple nonparametric tests

- The sign test
  - Statistic: $t=n_+/n$ (the proportion of positive effects)
  - The null distribution is known (Binomial law) if the population distribution is assumed symmetric
  - No assumption required if first-level variances are zero

- The Wilcoxon signed rank test
  - Statistic: $t = \sum_i \text{rank}(|\beta_i|) \times \text{sign}(\beta_i)$
  - Assumes population symmetry *and* constant first-level variance

Permutation tests

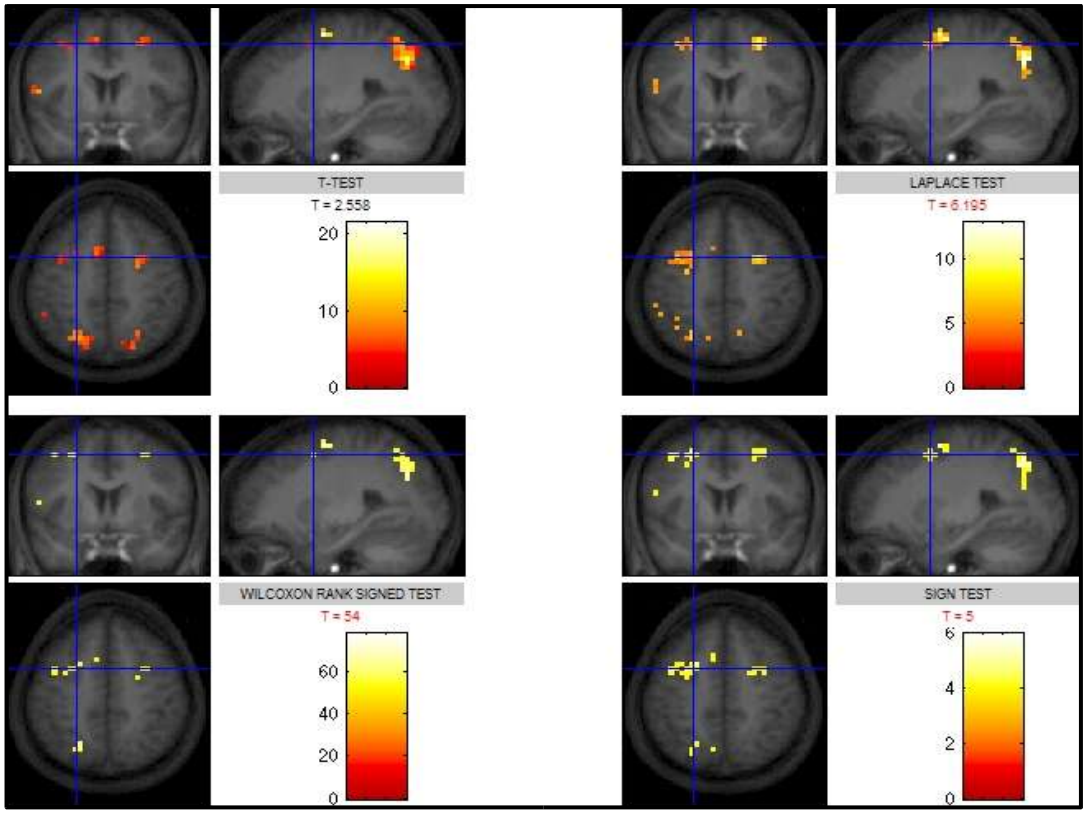
- General principle:
  - Consider *any* statistic
  - Under some *exchangeability* assumption, the null distribution may be approximated by permutations from the data

- Application to RFX in fMRI (Holmes et al, 1996; Nichols et al, 2001)
  - Use the Student statistic
  - Assuming population symmetry implies exchangeability of signs
  - Asymptotically equivalent to the *t*-test if data is truly normal i.i.d.
  - Easily accounts for the multiple comparison problem
  - Implemented in the *SnPM* toolbox

Robust tests: comparison

- Linguistic contrast: “french vs. english”
- 12 subjects
- Uncorrected threshold ($P=10^{-3}$)

t-test



Laplace test

Wilcoxon test

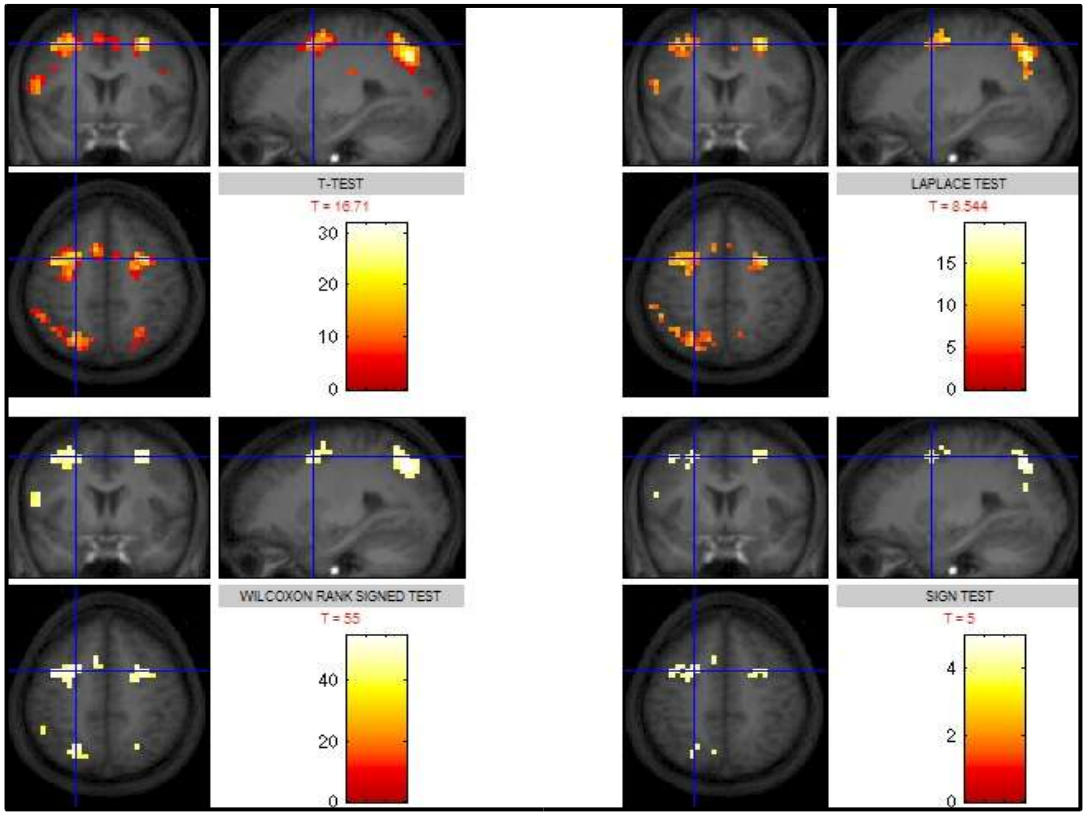
Sign test

Robust tests: comparison

- Linguistic contrast: “french vs. english”
- 12 subjects
- Uncorrected threshold ($P=10^{-3}$)

Two subjects removed

t-test



Laplace test

Wilcoxon test

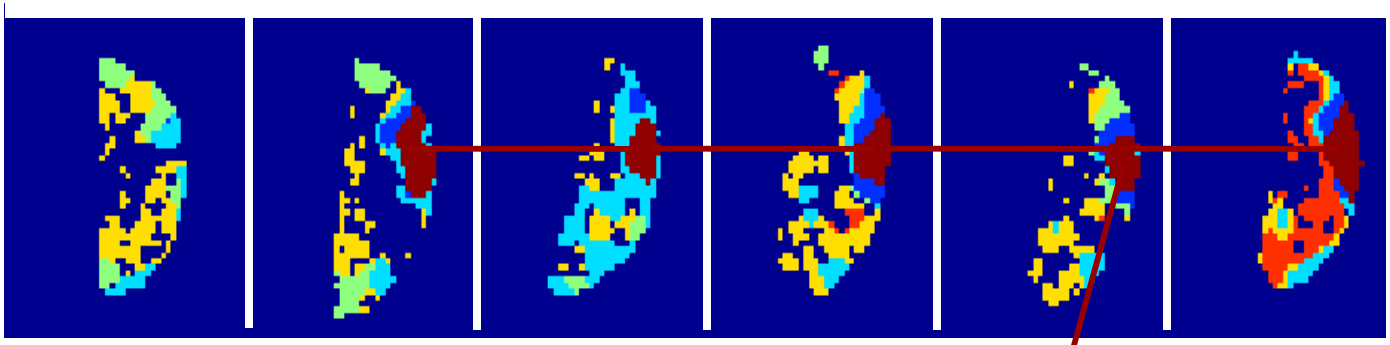
Sign test

Group inference: summary

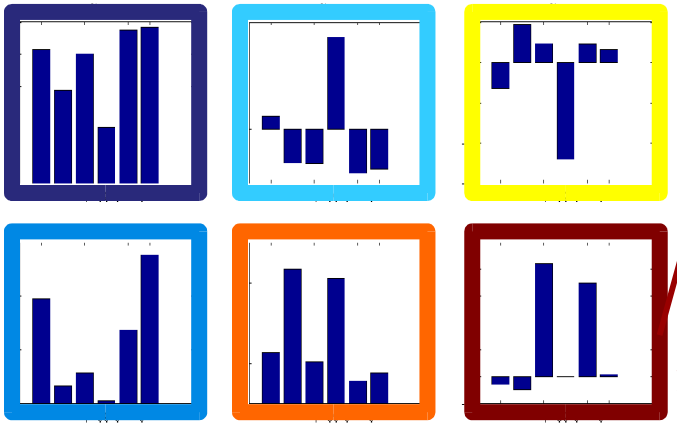
- RFX involves a two-level analysis (unlike FFX)
- Current parametric inference techniques are questionable as they rely on normality assumptions
- Nonparametric techniques are valuable alternatives
  - They may however lack sensitivity
  - Current techniques are probably suboptimal w.r.t. the treatment of first-level variability

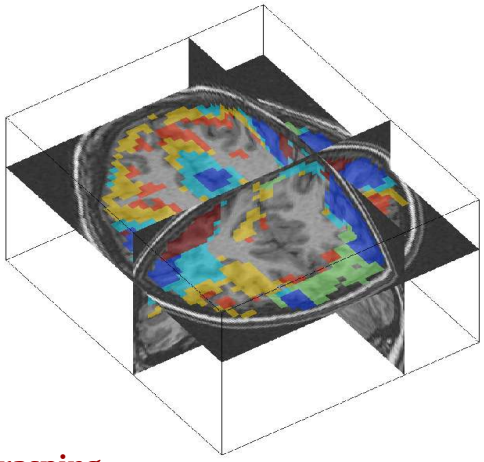
The future: simultaneous spatial and functional clustering?

(Flandin et al, 2003; Thirion et al, 2004)



- "Parietal protocol" (Simon et al, 2002)
- 6 subjects
- 6 contrasts of interest





Pointing and grasping