

Logiciel statistique d'analyse des images cérébrales : SPM

Su Ruan

Professeure à l'Université de Rouen

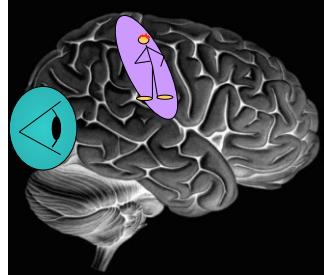
Introduction

- L'identification d'une région particulière du cerveau avec une fonction spécifique est un thème central dans les neurosciences
- Deux principes fondamentaux de l'organisation fonctionnelle
 - La spécialisation (ségrégation) fonctionnelle
 - Une région corticale est spécialisée pour certains aspects du traitement perceptif ou moteur, et que cette spécialisation est anatomiquement distincte dans le cortex.
 - L'intégration fonctionnelle
 - Ensembles de régions (nœuds) distribuées sur le cerveau qui interagissent pour réalisation d'une tâche: interaction entre nœuds.
 - Relations entre elles
 - L'infrastructure corticale peut faire impliquer ensemble de nombreuses régions spécialisées dont l'union est coordonnée par l'intégration fonctionnelle entre elles.
 - La spécialisation fonctionnelle n'a de sens que dans le contexte de l'intégration fonctionnelle et vice versa.

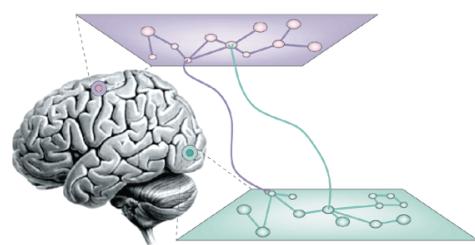


Principles of organisation: complementary approaches

Functional Specialisation

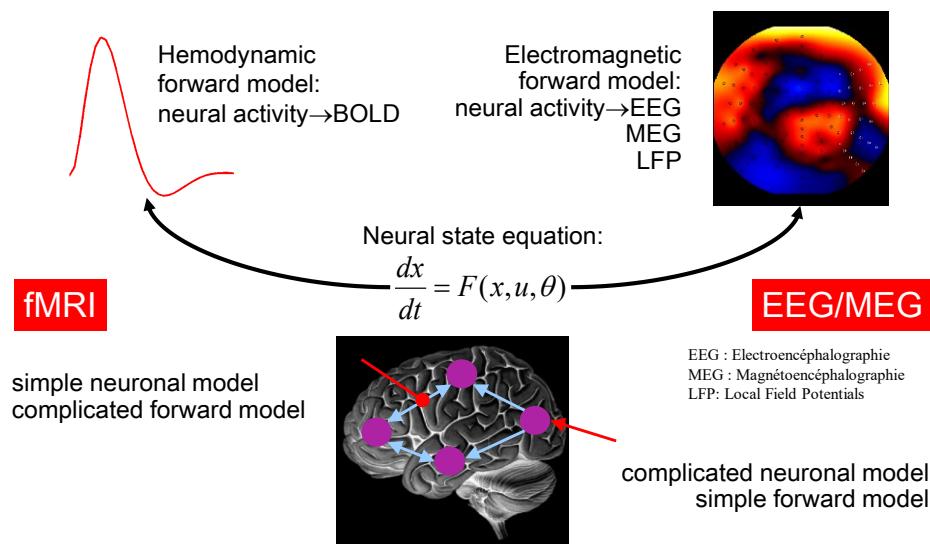


Functional Integration



3

Dynamic Causal Modelling (DCM)

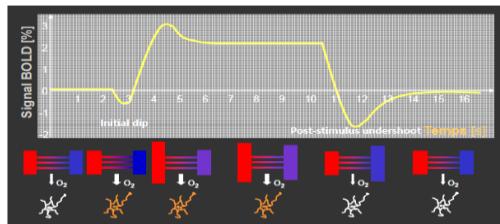


IRM fonctionnelle

➤ Activité neuronale

- Réponse métabolique
- Réponse hémodynamique
 - > variation de la susceptibilité magnétique du sang:
Signal BOLD (Blood-Oxygen-Level Dependent)

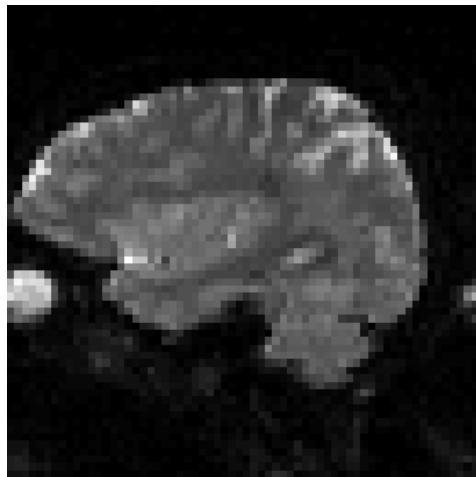
➤ Décours temporel d'un signal BOLD



Le rapport signal / bruit
est très faible.

Image de [A. Krainik, J. Wamking](#)

fMRI time-series movie

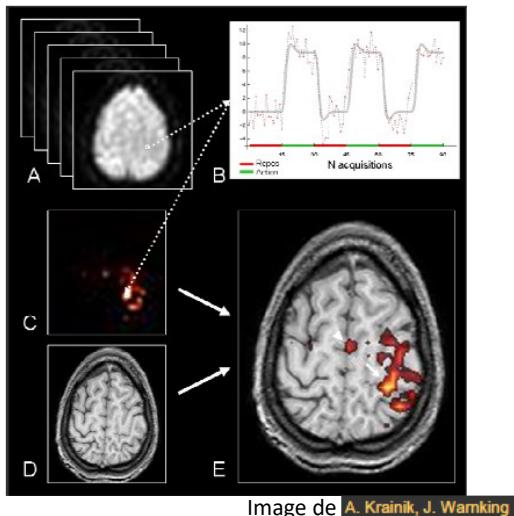
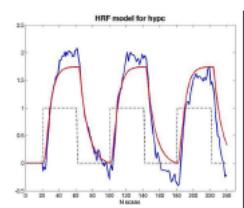


IRM fonctionnelle

- Paradigme :

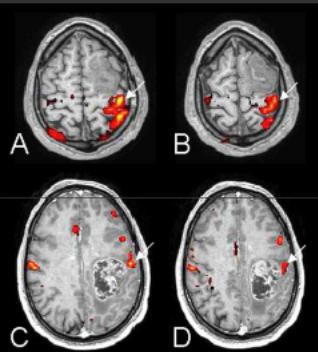
- deux états

- repos et activé
- activé 1 et activé 2

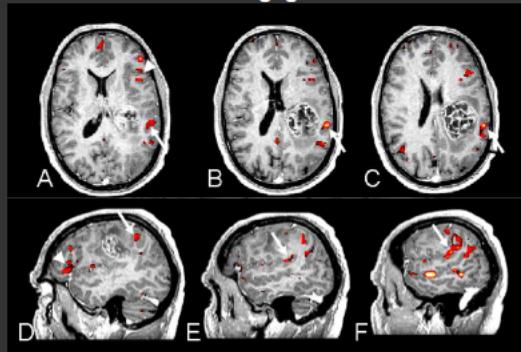


Cartographie des principales zones fonctionnelles

Motricité



Langage



Validation

MEG Stippich, Neuroreport 1998, ESPO Lehéricy, J Neurosurg 2000

WADA Binder, Neurology 1996; Hertz-Pannier, Neurology 1997, Lésions Krainik, Neurology 2001, 2003, 2004,

Image de A. Krainik, J. Wamking

SPM: Statistical Parametric Mapping

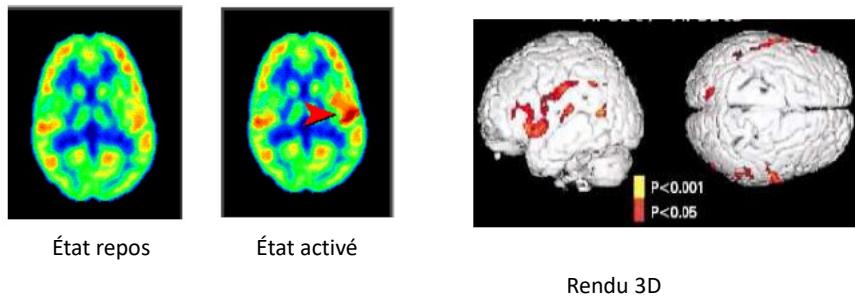
- Statistic
 - Les valeurs représentées dans les images paramétriques sont des **valeurs de statistique** (ou de valeurs de p)
 - Chaque valeur des images paramétriques est fonction de la probabilité qu'a le pixel considéré à suivre (ou à ne pas suivre) le modèle considéré
- Parametric
 - La méthode suppose **un modèle paramétrique des cinétiques** suivies par les différents pixels, c'est-à-dire qu'elle suppose que ces cinétiques suivent une fonction caractérisée par certains paramètres
 - Il faut donc avoir une idée a priori assez forte concernant **la cinétique** suivie par les régions d'intérêt
- Mapping
 - On obtient des **images de paramètres** (paramètre = valeur du test statistique), donc une cartographie

SPM: Statistical Parametric Mapping

- Développé par le département Wellcome de neurologie Cognitive de l'institut de Neurologie de Londres
- L'approche la plus répandue à la caractérisation de l'anatomie fonctionnelle et les changements liés à la maladie
 - Voxel-based analyses
- Identification des réponses fonctionnelles du cerveau, basée **sur l'analyse des séquences de données** d'imagerie cérébrale: IRMf, PET, SPECT, EEG, MEG

SPM: Statistical Parametric Mapping

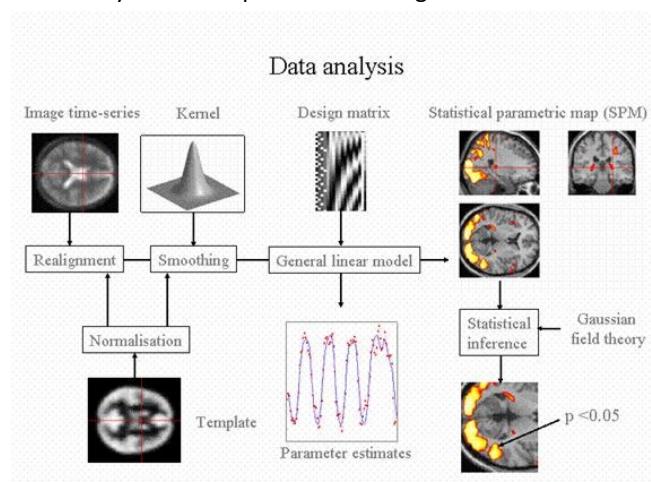
Exemple: une cartographie



Images extraites du cours d'Irène Buvat

SPM: Statistical Parametric Mapping

– But: Analyses des séquences des images cérébrales



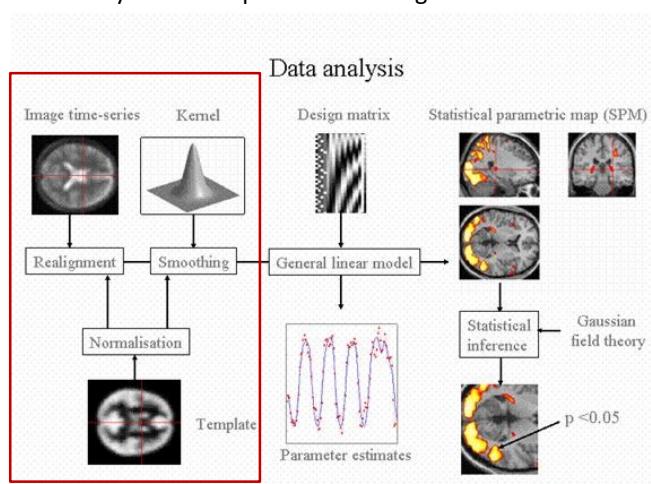
SPM: Statistical Parametric Mapping

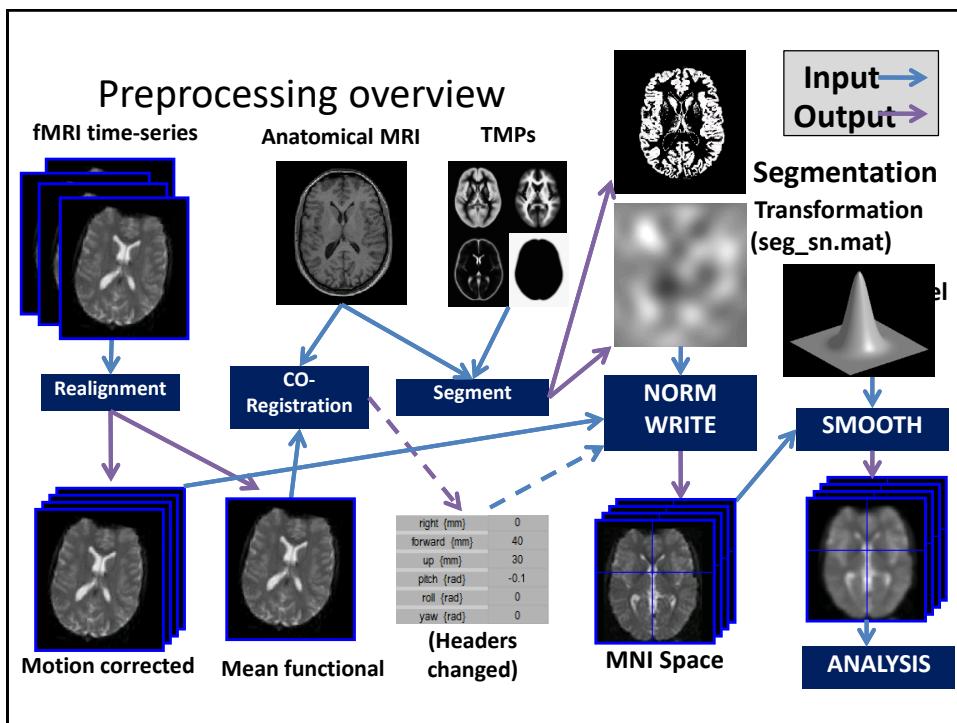
– Fonctions

- Images are realigned, spatially normalised into a standard space, and smoothed.
- Parametric statistical models are assumed at each voxel, using the General Linear Model GLM to describe the data in terms of experimental and confounding effects, and residual variability.
- For fMRI the GLM is used in combination with a temporal convolution model.
- Classical statistical inference is used to test hypotheses that are expressed in terms of GLM parameters. This uses an image whose voxel values are statistics, a *Statistic Image*, or *Statistical Parametric Map* (SPM{t}, SPM{Z}, SPM{F})

SPM: Statistical Parametric Mapping

– But: Analyses des séquences des images cérébrales





Realignment spatial et normalisation

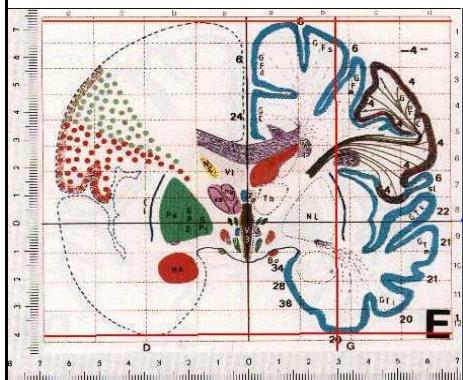
- Realignment spatial
 - Recalage des images
 - Une série des images
 - Images multimodales
- Normalisation
 - Les images sont recalées, puis normalisées spatialement pour être représentées sur un espace anatomique standard (MNI: Montreal Neurological Institute).

Realignment spatial et normalisation

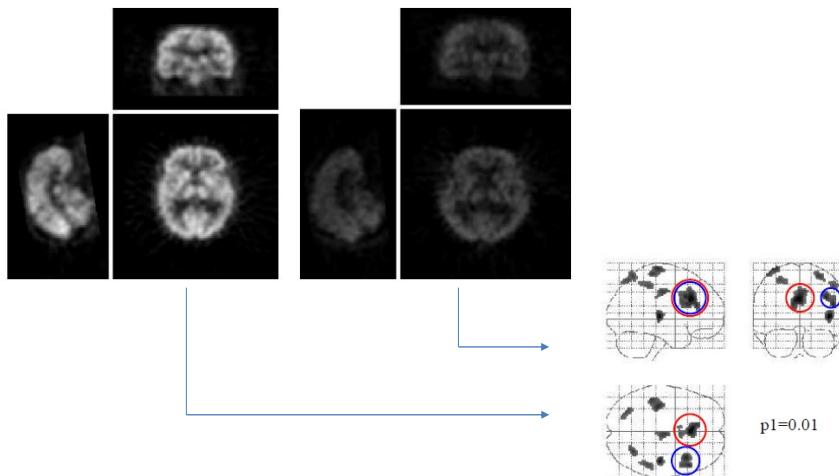
- Inter-subject averaging
 - Increase sensitivity with more subjects
 - Fixed-effects analysis
 - Extrapolate findings to the population as a whole
 - Mixed-effects analysis
- Make results from different studies comparable by aligning them to standard space

Standard spaces

The Talairach Atlas



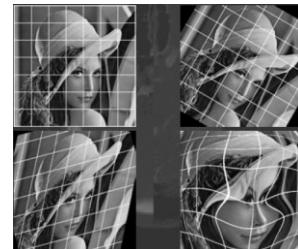
Realignment spatial et normalisation



Realignment spatial et normalisation

– Method of spatial Realignment

- Rigide transformation on 3D
 - 6 parameters (rotation and translation)
- Principe:
 - estimating the 6 parameters of an affine 'rigid-body' transformation that minimize the [sum of squared] differences between each successive scan and a reference scan
 - applying the transformation by re-sampling the data using tri-linear, sinc or spline interpolation.
- Adjusting for movement related effects in fMRI
 - in fMRI, even after perfect realignment, movement-related signals can still persist.
 - The movement-related signal is firstly estimated and then simply subtracted from the original data.



Realignment spatial et normalisation

- movement related effects in fMRI

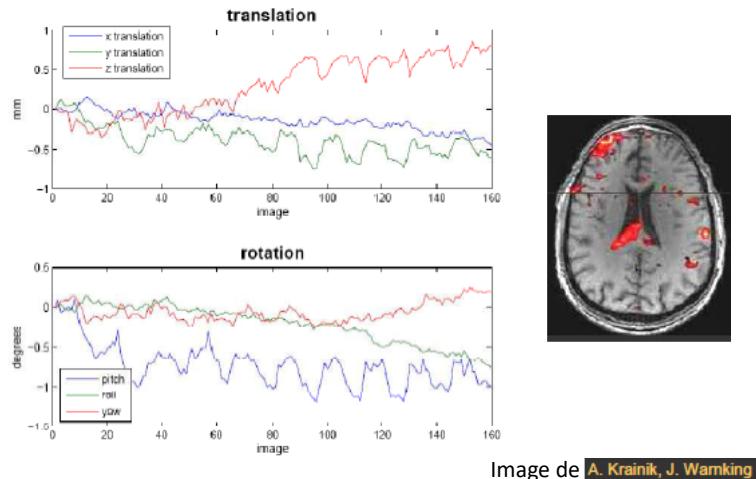
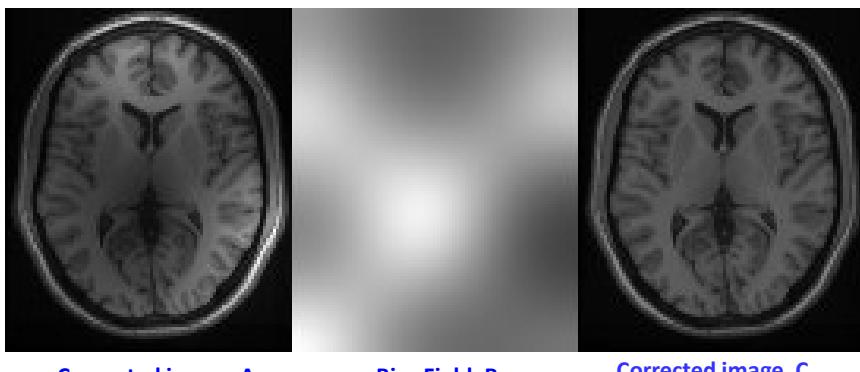


Image de [A. Krainik, J. Womking](#)

Inhomogeneity correction

- A multiplicative bias field can be modelled as a linear combination of basis functions: $C = A + B$



Realignment spatial et normalisation

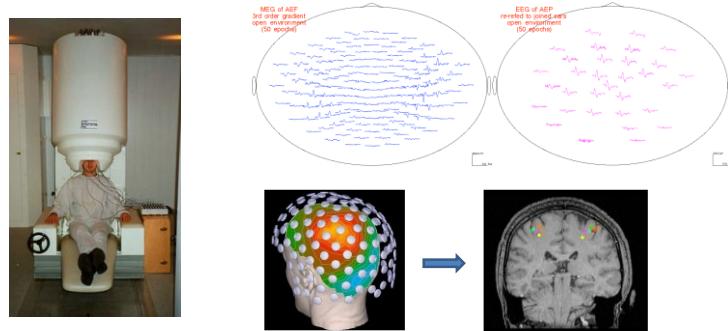
– Method of Normalisation

- A mean image of the series, or some other co-registered (e.g. a T₁-weighted) image, is used to estimate some warping parameters that map it onto a template that already conforms to some standard anatomical space
- A special consideration is the spatial normalization of brains that have gross anatomical pathology.
 - two sorts (i) quantitative changes in the amount of a particular tissue compartment (e.g. cortical atrophy) or (ii) qualitative changes in anatomy involving the insertion or deletion of normal tissue compartments (e.g. ischemic tissue in stroke or cortical dysplasia).

Realignment spatial et normalisation

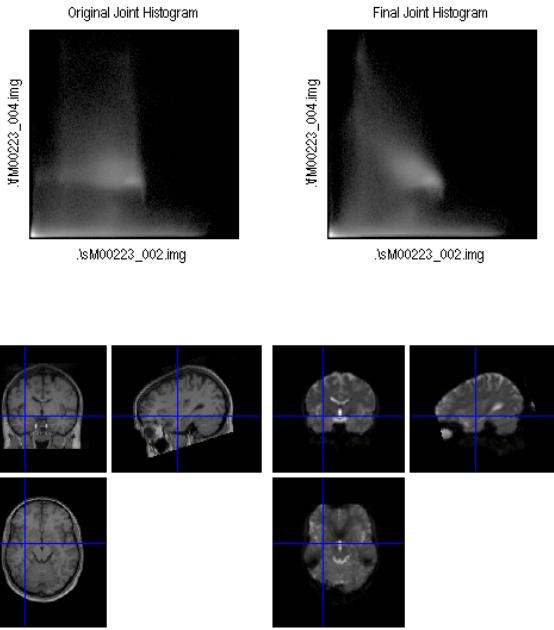
– Co-registration of functional and anatomical data

- IRM - IRMf- EEG (Electroencéphalographie) – MEG (Magnétoencéphalographie)

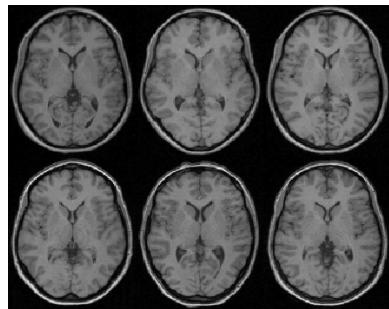


Coregistration (NMI)

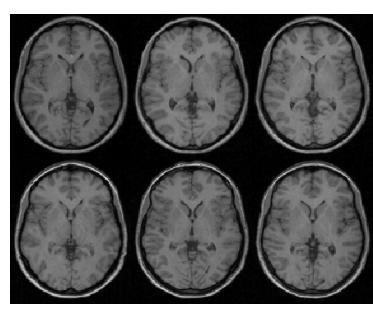
- Intermodal coreg.
 - Can't simply use intensity difference
 - Quantify how well one image predicts the other = how much shared info
 - Info from joint probability distribution.
 - Estimated from joint histogram



Realignment spatial et **normalisation**

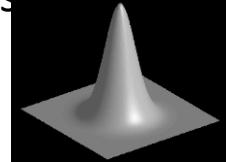


Affine registration



Non-linear registration

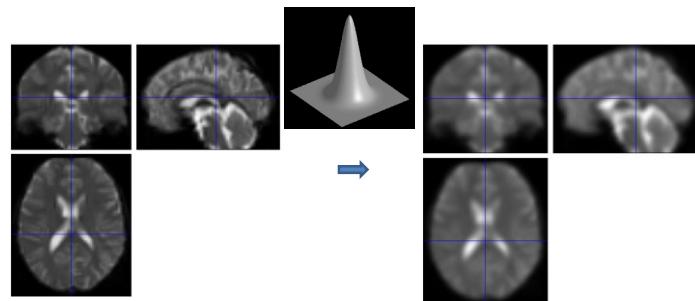
Realignment spatial et normalisation



- Why would we deliberately blur the data?
 - Improves spatial overlap by blurring over minor anatomical differences and registration errors
 - Averaging neighbouring voxels suppresses noise
 - Increases sensitivity to effects of similar scale to kernel (matched filter theorem)
 - Makes data more normally distributed (central limit theorem)
 - Reduces the effective number of multiple comparisons
- How is it implemented?
 - Convolution with a 3D Gaussian kernel, of specified full-width at half-maximum (FWHM) in mm

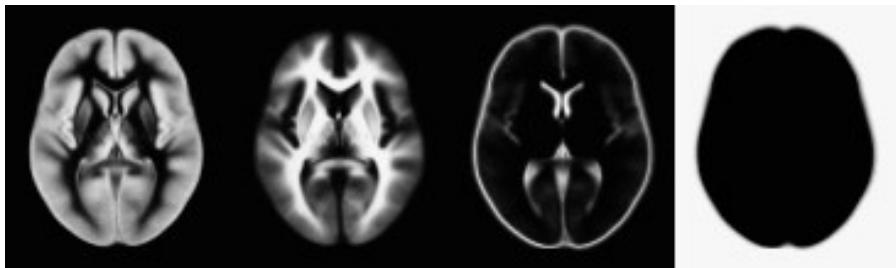
Realignment spatial et normalisation

- Spatial smoothing



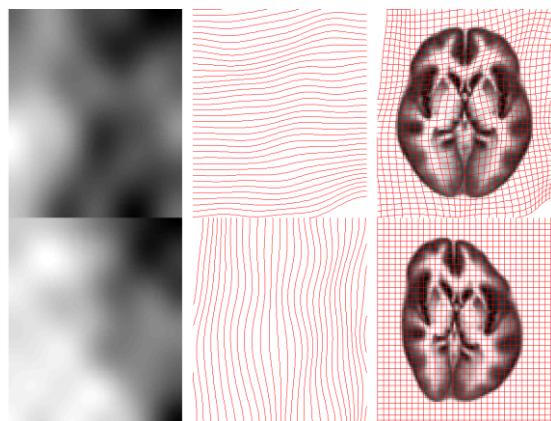
Tissue Probability Maps

- Tissue probability maps (TPMs) are used as the prior, instead of the proportion of voxels in each class



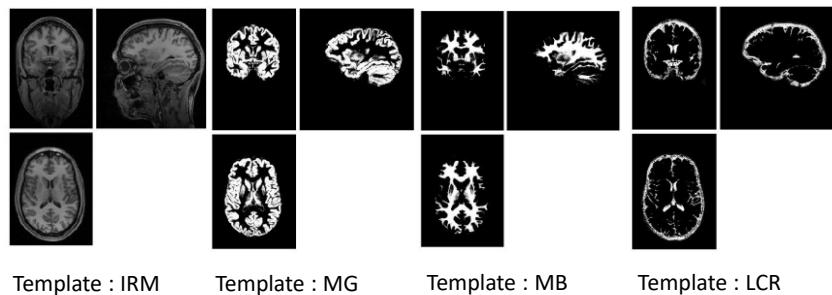
Deforming the Tissue Probability Maps

- * Tissue probability images are warped to match the subject
- * The inverse transform warps to the TPMs



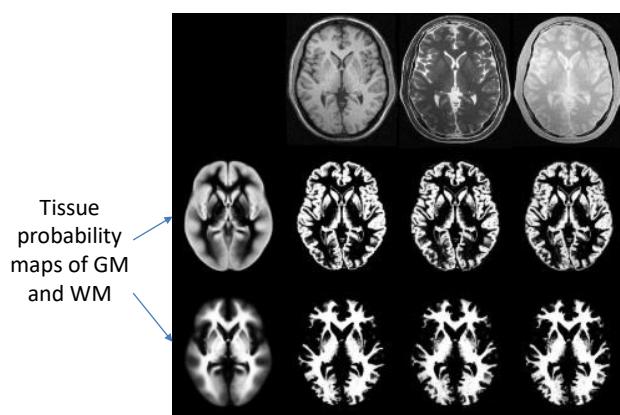
Segmentation

- Recalage entre le template et l'image à segmenter
 - Espace de l'image

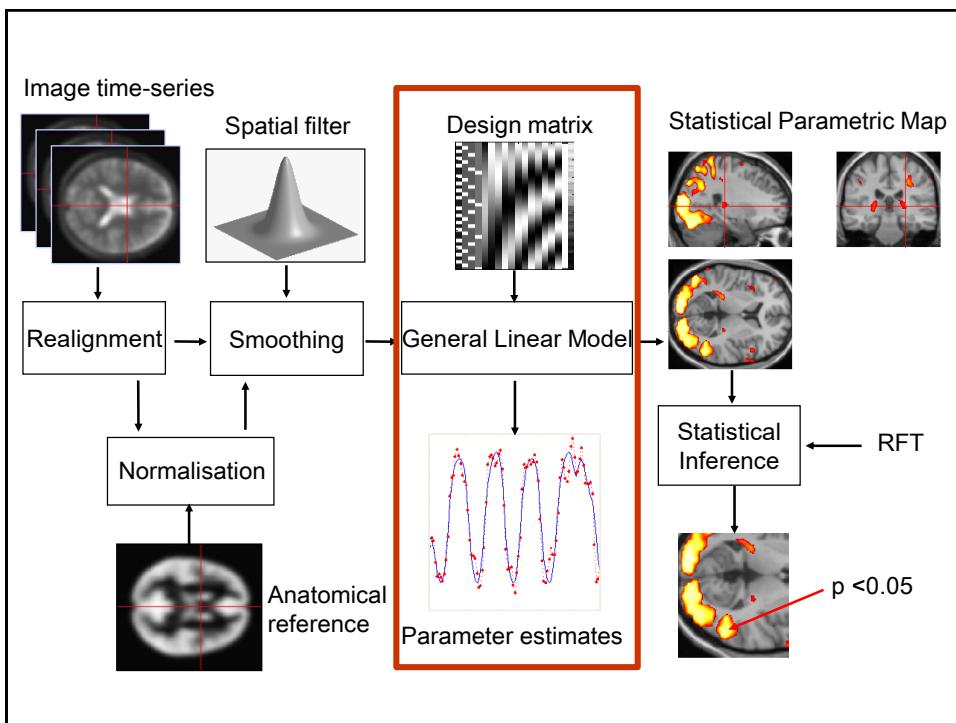


Segmentation

- Recalage entre le template et l'image à segmenter
 - Espace normalisé

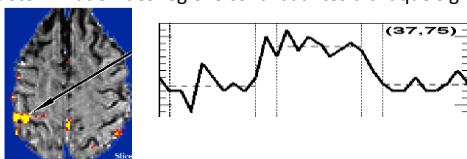


Cocosco, Kollokian, Kwan & Evans. "BrainWeb: Online Interface to a 3D MRI Simulated Brain Database". NeuroImage 5(4):S425 (1997)



General Linear Model (GLM)

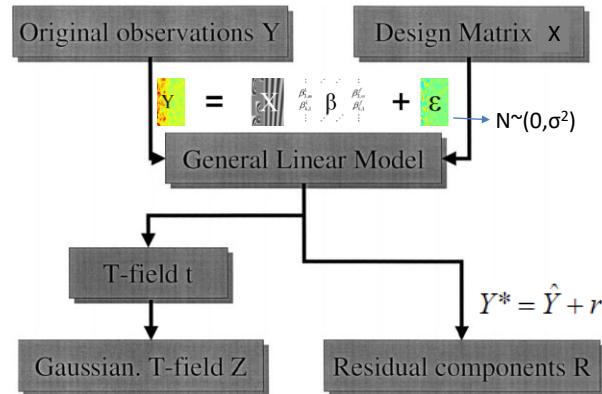
- Modèle *a priori* de la réponse neurophysiologique : modèle linéaire général
 - ⇒ détermination d'une matrice décrivant les signaux attendus en réponse au *paradigme* fonctionnel mis en œuvre (matrice explicative)
- Ajustement du modèle
 - ⇒ détermination des régions contribuantes à chaque signal réponse attendu



- Inférence statistique régionale possible
 - Hypothèse testée H_0 : région X significativement activée par le stimulus ?

General Linear Model (GLM)

– Schéma général

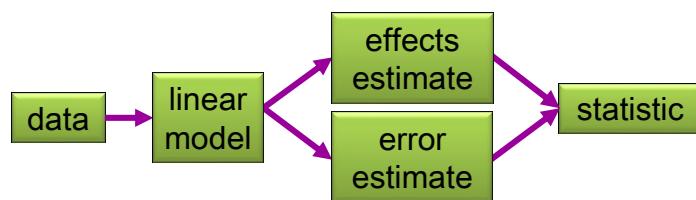


Modelling the measured data

Why? Make inferences about effects of interest

How?

1. Decompose data into effects and error
2. Form statistic using estimates of effects and error



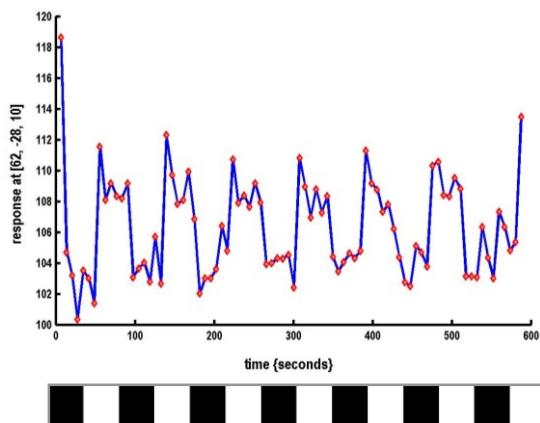
A very simple fMRI experiment

One session

Passive word
listening
versus rest

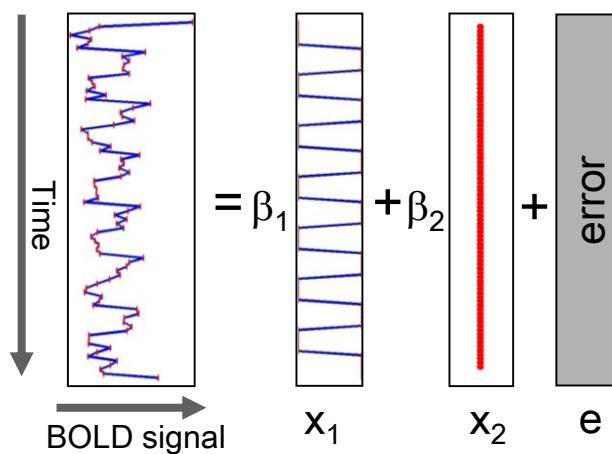
7 cycles of
rest and listening

Blocks of 6 scans
with 7 sec TR



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

Single voxel regression model



$$y = x_1\beta_1 + x_2\beta_2 + e$$

Mass-univariate analysis: voxel-wise GLM

$y = X\beta + e$

$e \sim N(0, \sigma^2 I)$

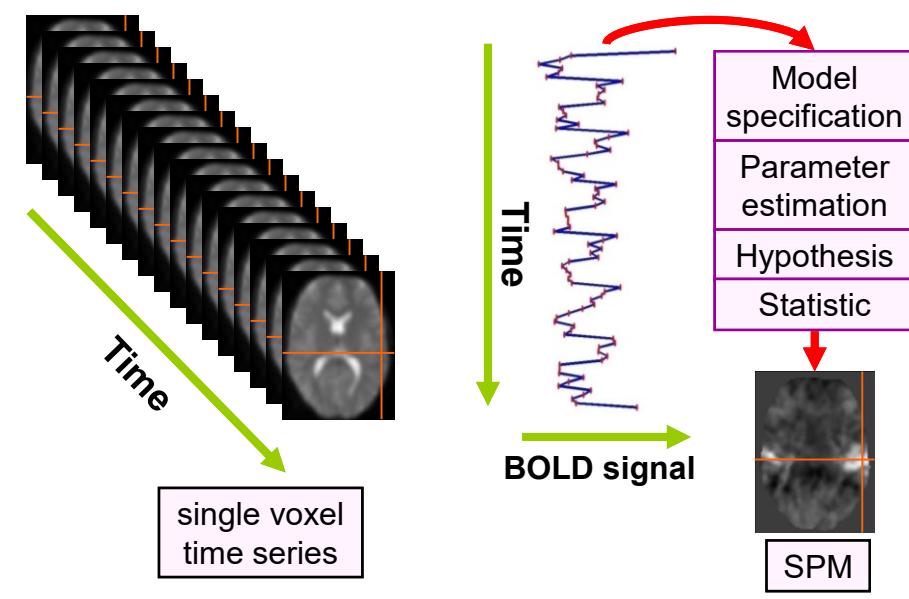
Model is specified by

1. Design matrix X
2. Assumptions about e

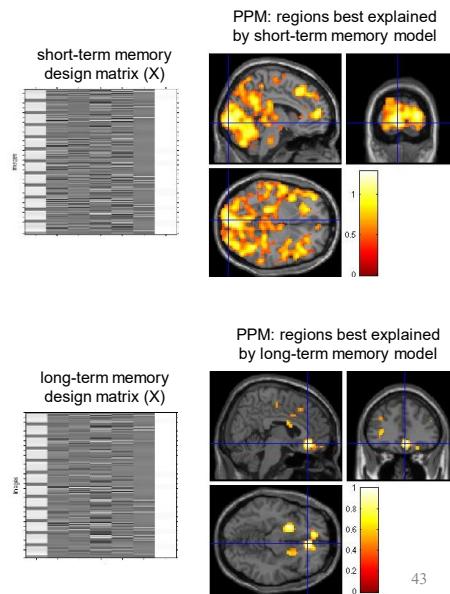
N : number of scans
 p : number of regressors

The design matrix embodies all available knowledge about experimentally controlled factors and potential confounds.

Voxel-wise time series analysis



Design matrix

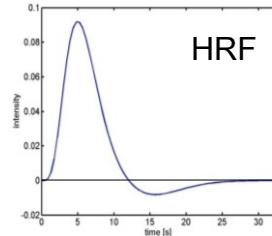


IMPROVING THE MODEL

What are the problems of this model?

1. BOLD responses have a delayed and dispersed form.

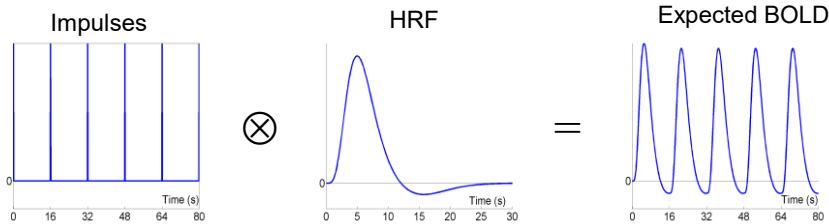
HRF: Hemodynamic Response Function



2. The BOLD signal includes substantial amounts of low-frequency noise (eg due to scanner drift).
3. Due to breathing, heartbeat & unmodeled neuronal activity, the errors are serially correlated. This violates the assumptions of the noise model in the GLM

Problem 1: Shape of BOLD response

Solution: Convolution model

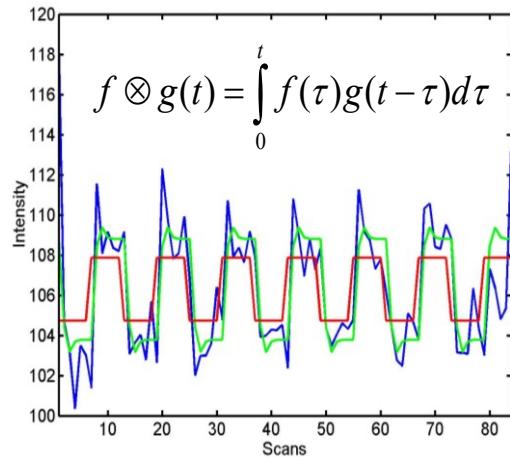
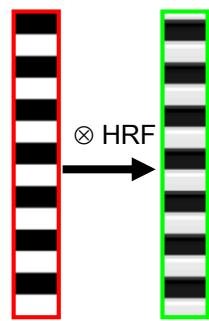


$$f \otimes g(t) = \int_0^t f(\tau)g(t-\tau)d\tau$$

expected BOLD response
= input function \otimes impulse response function (HRF)

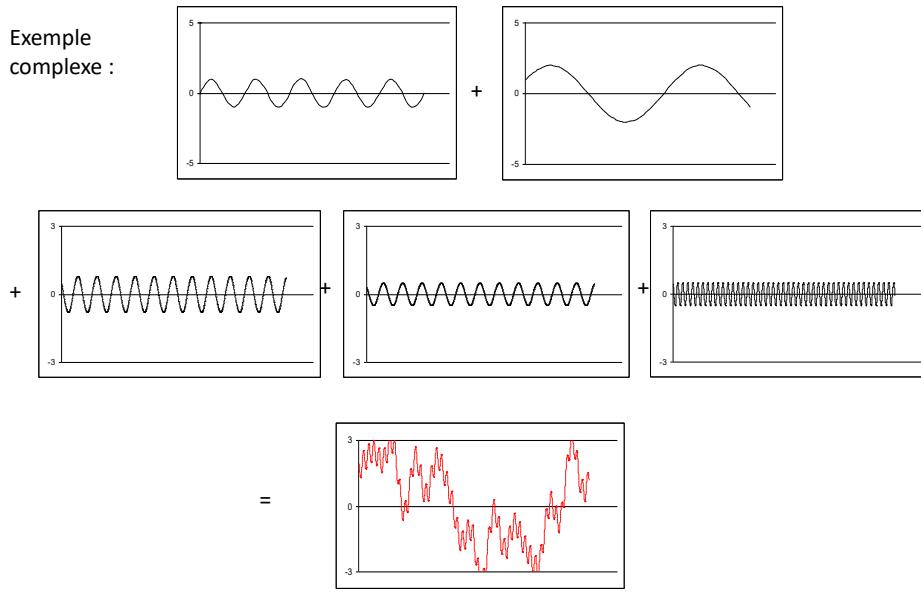
Convolution model of the BOLD response

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



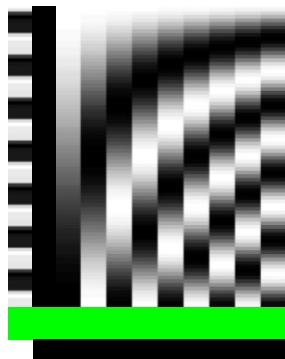
Transformée de Fourier

Exemple complexe :

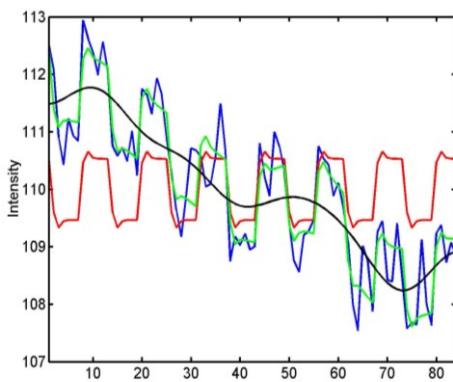


Problem 2: Low-frequency noise

Solution: High pass filtering

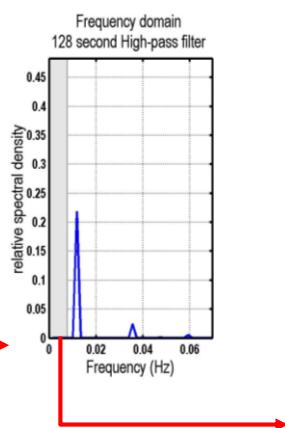


discrete cosine
transform (DCT)
set

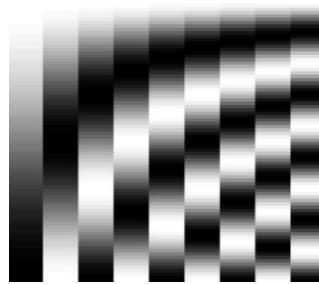


- blue = data
- black = mean + low-frequency drift
- green = predicted response, taking into account low-frequency drift
- red = predicted response, NOT taking into account low-frequency drift

High pass filtering



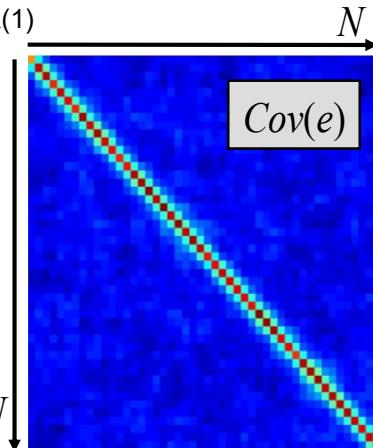
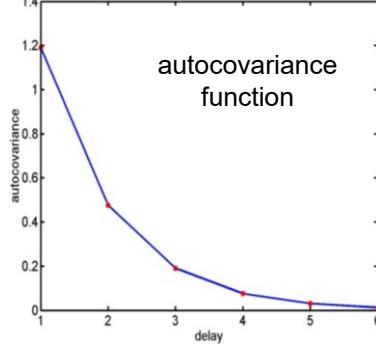
discrete cosine
transform (DCT)
set



Problem 3: Serial correlations

$$e_t = ae_{t-1} + \varepsilon_t \text{ with } \varepsilon_t \sim N(0, \sigma^2)$$

1st order autoregressive process: AR(1)



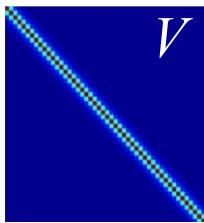
Multiple covariance components

$$e_i \sim N(0, C_i)$$

enhanced noise model at voxel i

$$\begin{aligned} C_i &= \sigma_i^2 V \\ V &= \sum \lambda_j Q_j \end{aligned}$$

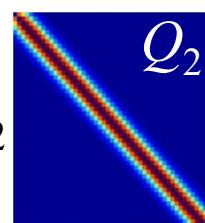
error covariance components Q
and hyperparameters λ



$$= \lambda_1$$



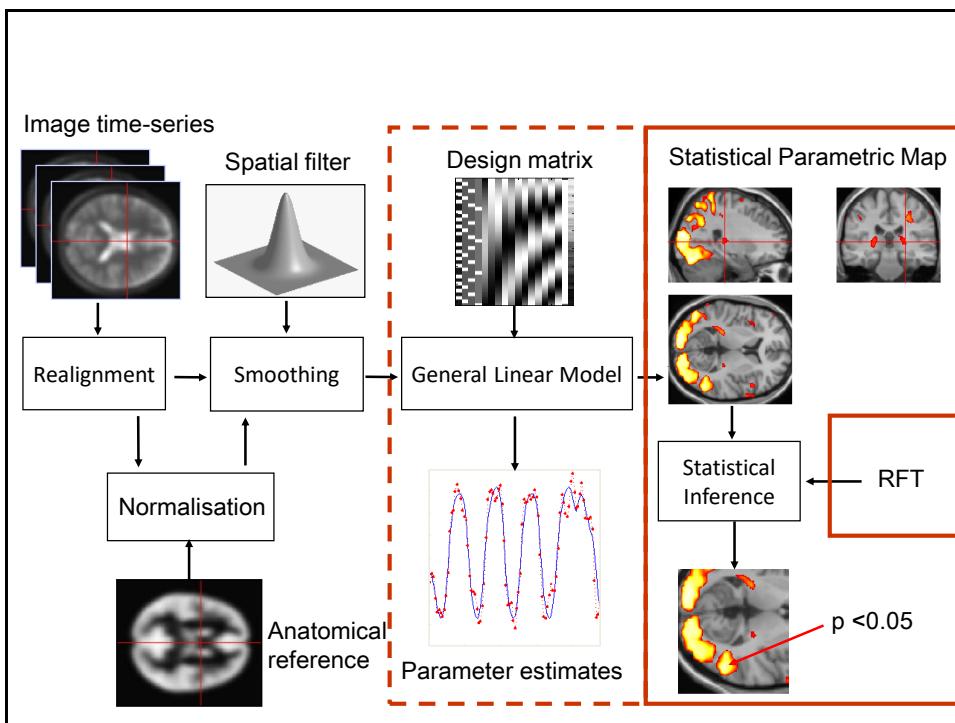
$$+ \lambda_2$$



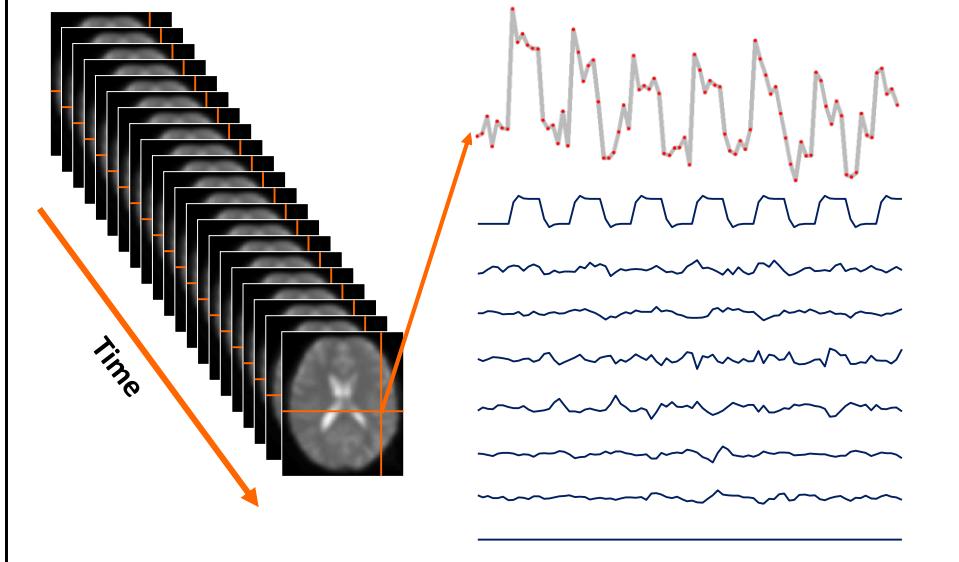
Estimation of hyperparameters λ with ReML (Restricted Maximum Likelihood).

General Linear Model (improved version)

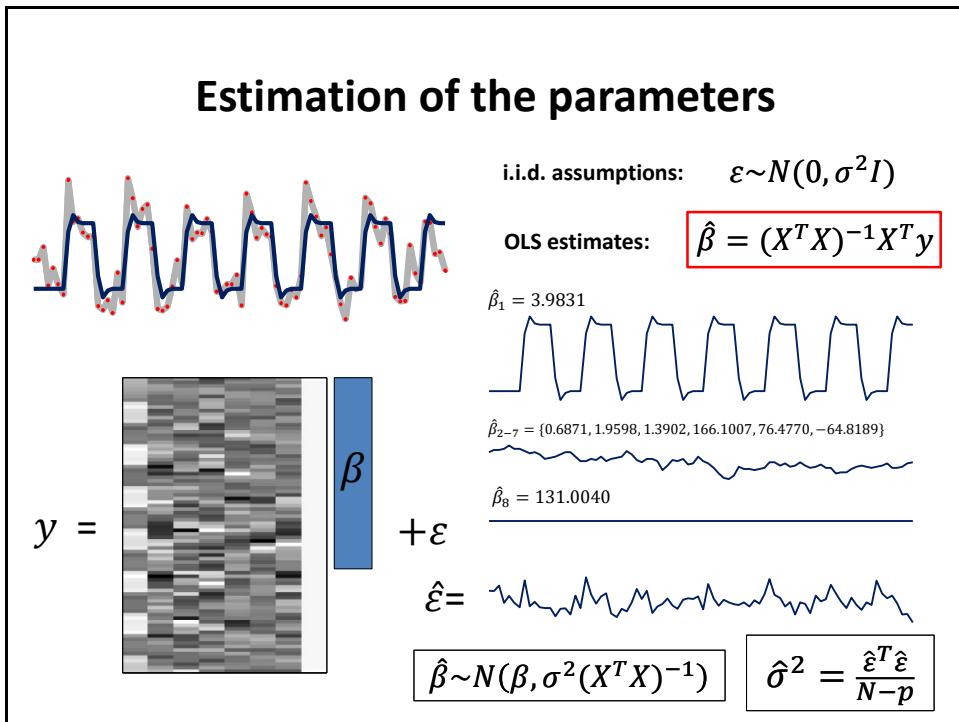
1. A general linear model of the data is used
2. The model is combined with the **Hemodynamic Response Function (HRF)**, **high-pass filtered** and **serial correlations corrected**
3. The model is applied to every voxel, producing beta images.



A mass-univariate approach

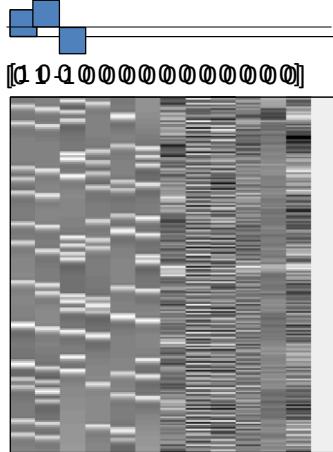


Estimation of the parameters



Contrasts

□ A contrast selects a specific effect of interest.



⇒ A contrast c is a vector of length p .

⇒ $c^T \beta$ is a linear combination of regression coefficients β .

$$c = [1 \ 0 \ 0 \ 0 \ \dots]^T$$

$$\begin{aligned} c^T \beta &= \mathbf{1} \times \beta_1 + \mathbf{0} \times \beta_2 + \mathbf{0} \times \beta_3 + \mathbf{0} \times \beta_4 + \dots \\ &= \boldsymbol{\beta}_1 \end{aligned}$$

$$c = [0 \ 1 \ -1 \ 0 \ \dots]^T$$

$$\begin{aligned} c^T \beta &= \mathbf{0} \times \beta_1 + \mathbf{1} \times \beta_2 + \mathbf{-1} \times \beta_3 + \mathbf{0} \times \beta_4 + \dots \\ &= \boldsymbol{\beta}_2 - \boldsymbol{\beta}_3 \end{aligned}$$

$$c^T \hat{\beta} \sim N(c^T \beta, \sigma^2 c^T (X^T X)^{-1} c)$$

Hypothesis Testing

To test an hypothesis, we construct “test statistics”.

- **Null Hypothesis H_0**

Typically what we want to disprove (no effect).

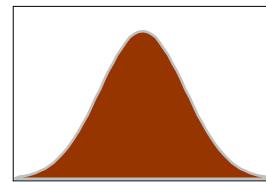
⇒ The Alternative Hypothesis H_A expresses outcome of interest.

- **Test Statistic T**

The test statistic summarises evidence about H_0 .

Typically, test statistic is small in magnitude when the hypothesis H_0 is true and large when false.

⇒ We need to know the distribution of T under the null hypothesis.



Hypothesis Testing

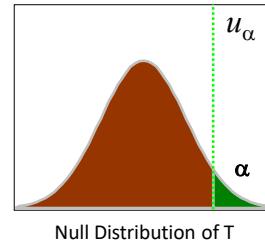
□ **Significance level α :**

Acceptable *false positive rate* α .

⇒ threshold u_α

Threshold u_α controls the false positive rate

$$\alpha = p(T > u_\alpha \mid H_0)$$



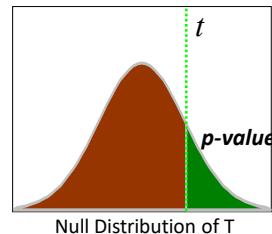
□ **Conclusion about the hypothesis:**

We reject the null hypothesis in favour of the alternative hypothesis if $t > u_\alpha$

□ **p-value:**

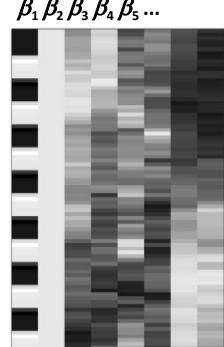
A *p-value* summarises evidence against H_0 .

This is the chance of observing value more extreme than t under the null hypothesis.



T-test - one dimensional contrasts – SPM{t}

$$c^T = \begin{matrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{matrix}$$



Question: box-car amplitude > 0 ?

$$= \\ \beta_1 = c^T \beta > 0 ?$$

Null hypothesis:

$$H_0: c^T \beta = 0$$

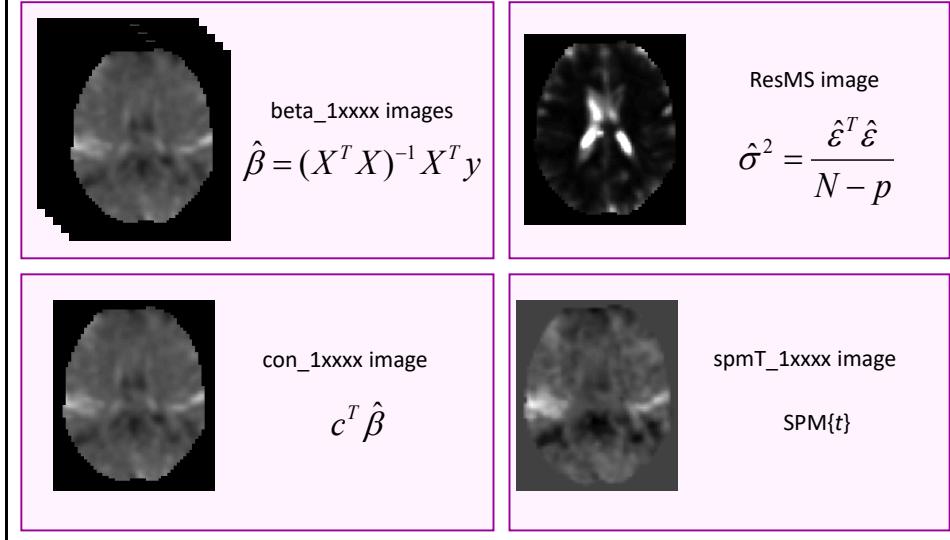
Test statistic:

$$T = \frac{\text{contrast of estimated parameters}}{\sqrt{\text{variance estimate}}}$$

$$T = \frac{c^T \hat{\beta}}{\sqrt{\text{var}(c^T \hat{\beta})}} = \frac{c^T \hat{\beta}}{\sqrt{\hat{\sigma}^2 c^T (X^T X)^{-1} c}} \sim t_{N-p}$$

T-contrast in SPM

□ For a given contrast c :



T-test: a simple example

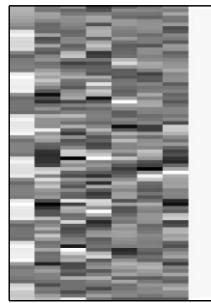
□ Passive word listening versus rest

$$c^T = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$$

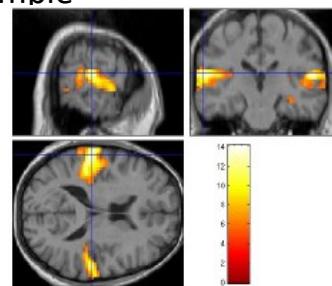
Q: activation during listening ?



Null hypothesis: $\beta_1 = 0$



$$t = \frac{c^T \hat{\beta}}{\sqrt{\text{var}(c^T \hat{\beta})}}$$



SPM results:

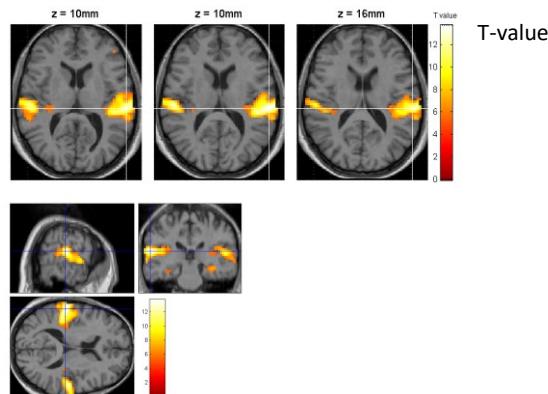
Height threshold $T = 3.2057$ ($p < 0.001$)

voxel-level mm mm mm

T	(Z)	p uncorrected	mm	mm	mm
13.94	Inf	0.000	-63	-27	15
12.04	Inf	0.000	-48	-33	12
11.82	Inf	0.000	-66	-21	6
13.72	Inf	0.000	57	-21	12
12.29	Inf	0.000	63	-12	-3
9.89	7.83	0.000	57	-39	6
7.39	6.36	0.000	36	-30	-15
6.84	5.99	0.000	51	0	48
6.36	5.65	0.000	-63	-54	-3
6.19	5.53	0.000	-30	-33	-18
5.96	5.36	0.000	36	-27	9
5.84	5.27	0.000	-45	42	9
5.44	4.97	0.000	48	27	24
5.32	4.87	0.000	36	-27	42

Images paramétriques et tests statistiques

- Seuillage



T-test: summary

- T -test is a *signal-to-noise* measure (ratio of estimate to standard deviation of estimate).

- Alternative hypothesis:

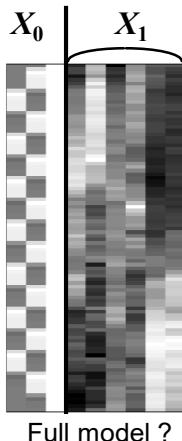
$$H_0: c^T \beta = 0 \quad \text{vs} \quad H_A: c^T \beta > 0$$

- T -contrasts are simple combinations of the betas; the T-statistic does not depend on the scaling of the regressors or the scaling of the contrast.

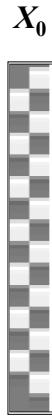
F-test - the extra-sum-of-squares principle

- Model comparison:

Null Hypothesis H_0 : True model is X_0 (reduced model)



$$\rightarrow \text{RSS} \sum \hat{\epsilon}_{full}^2$$



$$\rightarrow \text{RSS}_0 \sum \hat{\epsilon}_{reduced}^2$$

Test statistic: ratio of explained variability and unexplained variability (error)

$$F \propto \frac{RSS_0 - RSS}{RSS}$$

$$F \propto \frac{ESS}{RSS} \sim F_{v_1, v_2}$$

$$v_1 = \text{rank}(X) - \text{rank}(X_0)$$

$$v_2 = N - \text{rank}(X)$$

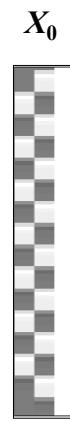
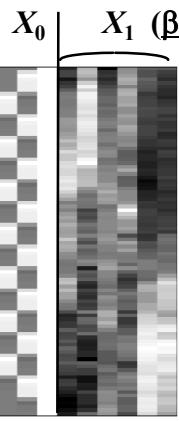
F-test - multidimensional contrasts – SPM{F}

- Tests multiple linear hypotheses:

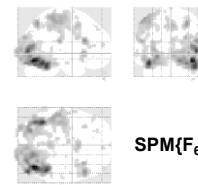
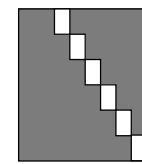
H_0 : True model is X_0

H_0 : $\beta_4 = \beta_5 = \dots = \beta_9 = 0$

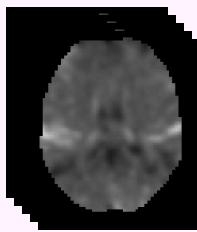
test H_0 : $c^T \beta = 0$?



$$c^T = \begin{matrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{matrix}$$

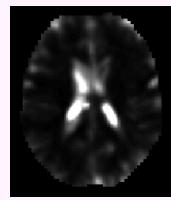


F-contrast in SPM



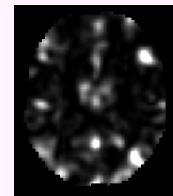
beta_00111 images

$$\hat{\beta} = (X^T X)^{-1} X^T y$$



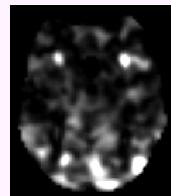
ResMS image

$$\hat{\sigma}^2 = \frac{\hat{\epsilon}^T \hat{\epsilon}}{N - p}$$



ess_00111 images

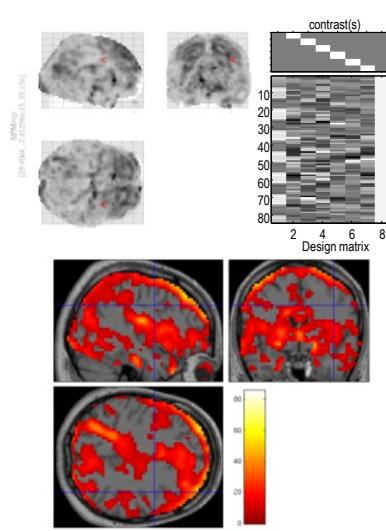
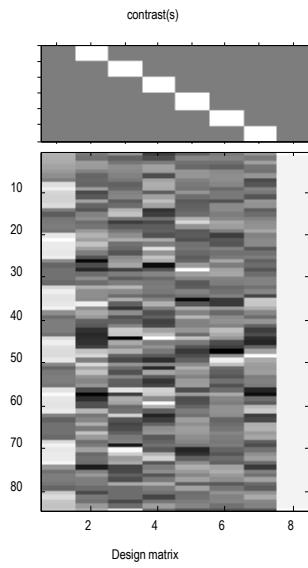
$$(RSS_{\theta} - RSS)$$



spmF_00111 images

$$SPM\{F\}$$

F-test example: movement related effects



F-test: summary

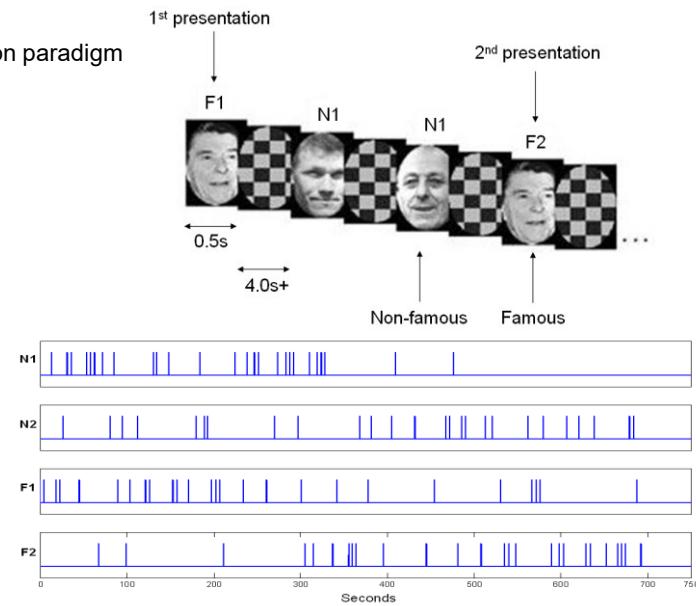
- F-tests can be viewed as testing for the additional variance explained by a larger model wrt a simpler (*nested*) model \Rightarrow **model comparison**.
- ❑ F tests a weighted **sum of squares** of one or several combinations of the regression coefficients β .
- ❑ In practice, we don't have to explicitly separate X into $[X_1 X_2]$ thanks to **multidimensional contrasts**.
- ❑ Hypotheses:

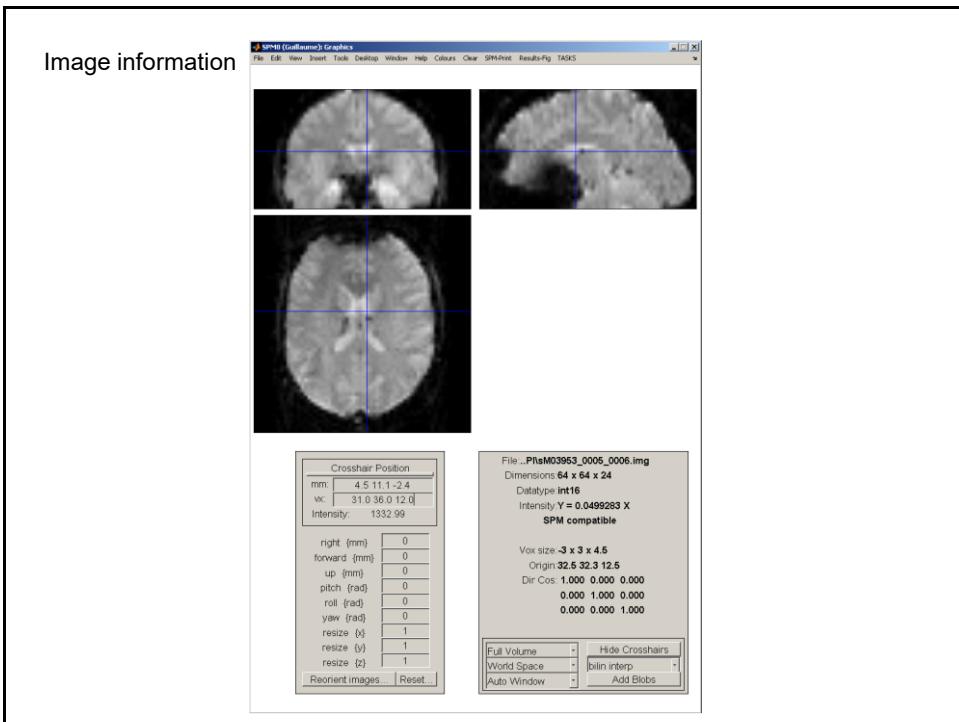
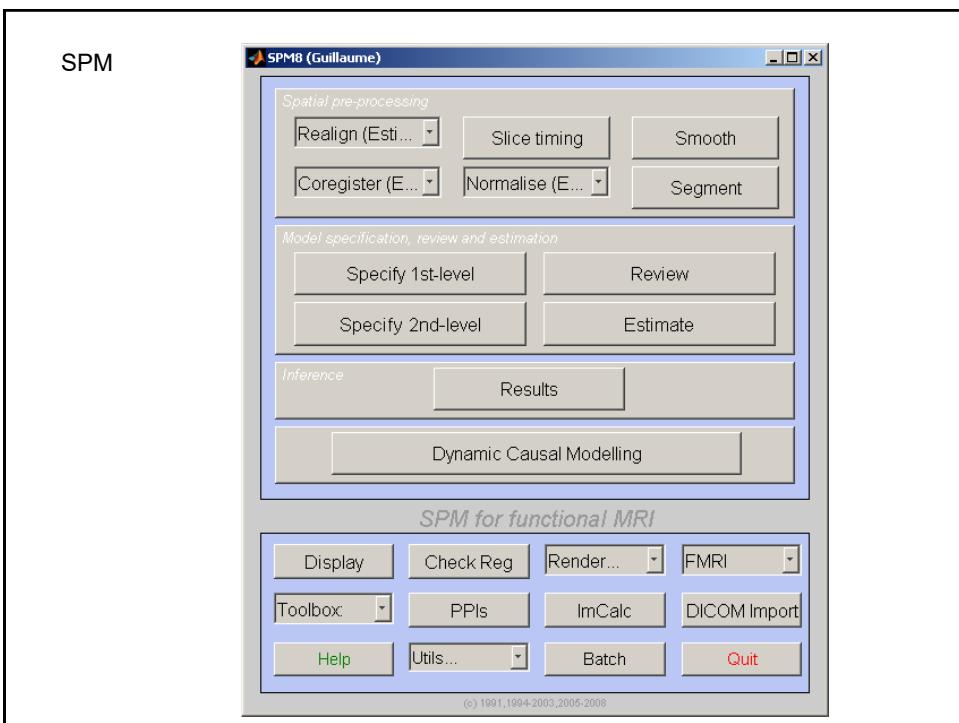
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

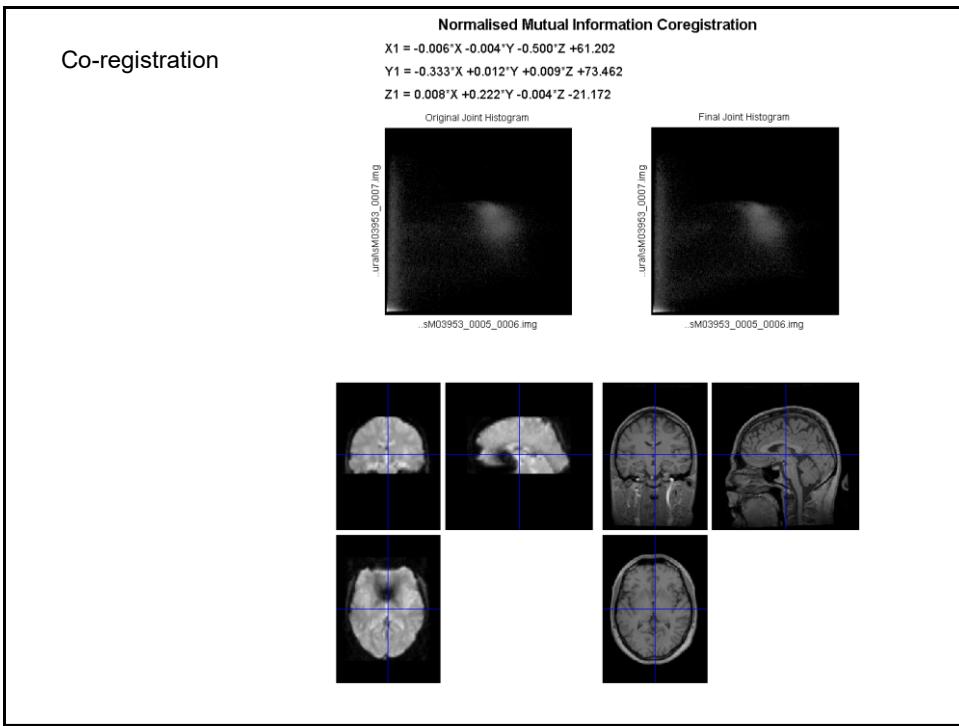
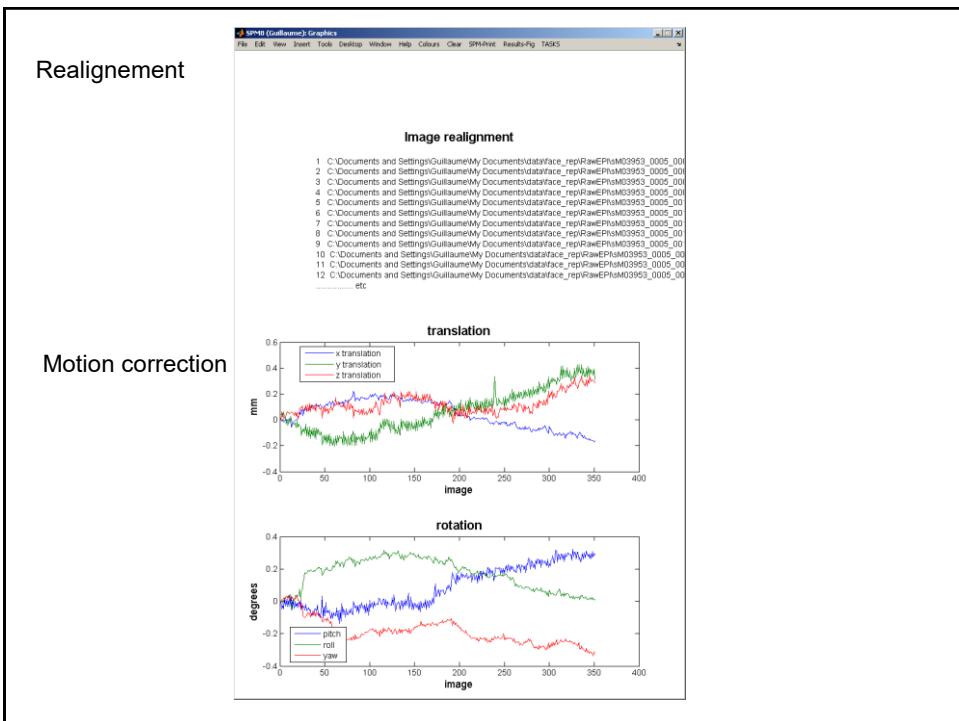
Null Hypothesis $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$

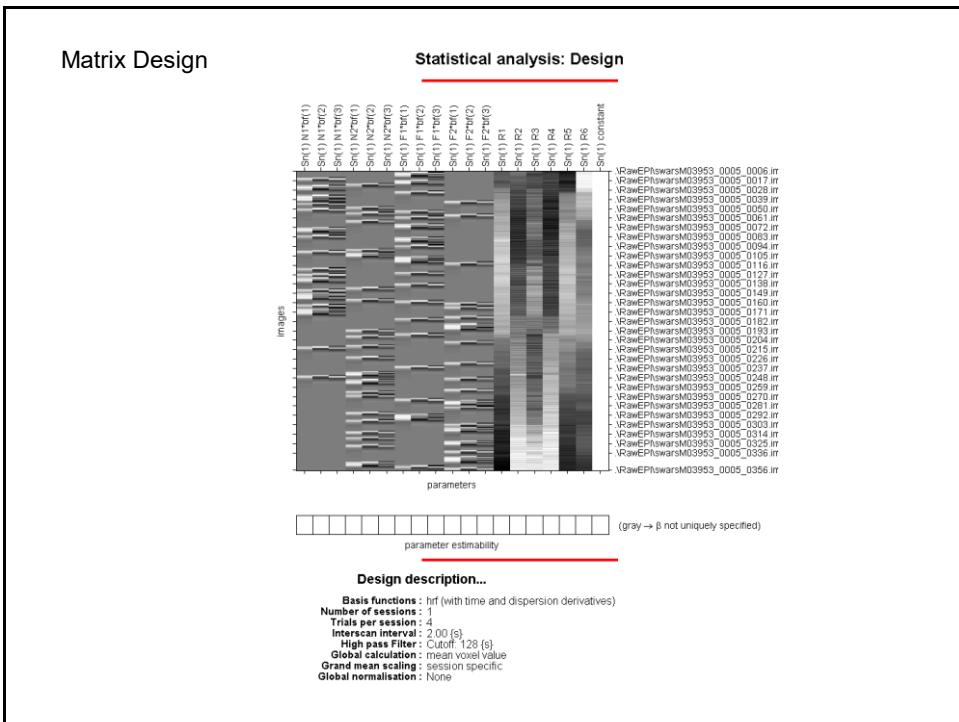
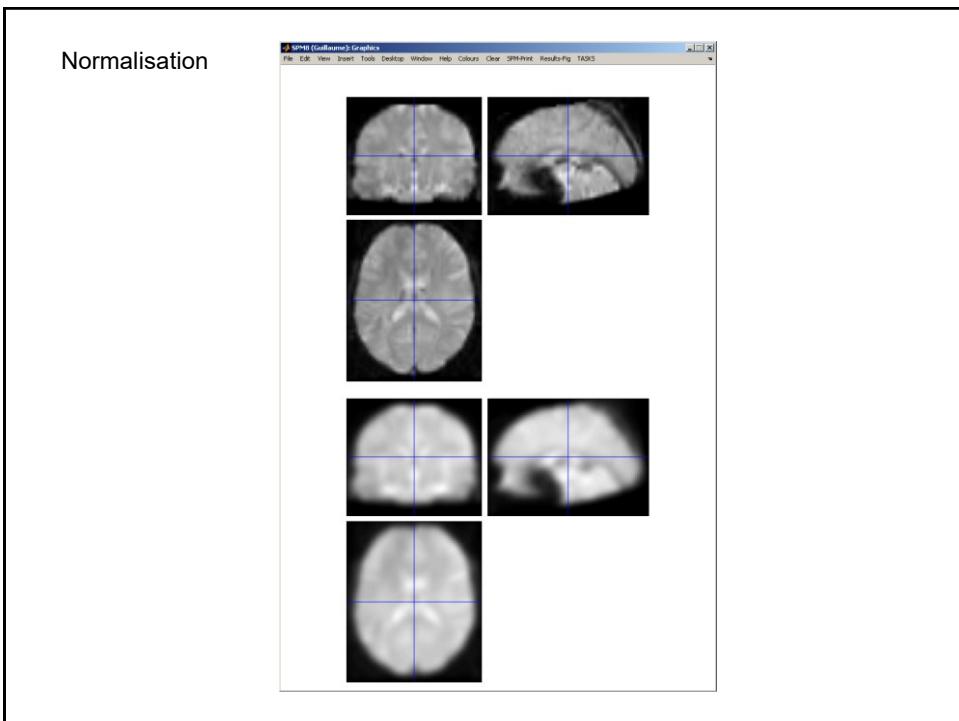
Alternative Hypothesis $H_A : \text{at least one } \beta_k \neq 0$
- ❑ In testing uni-dimensional contrast with an F-test, for example $\beta_1 - \beta_2$, the result will be the same as testing $\beta_2 - \beta_1$. It will be exactly the square of the t-test, testing for both positive and negative effects.

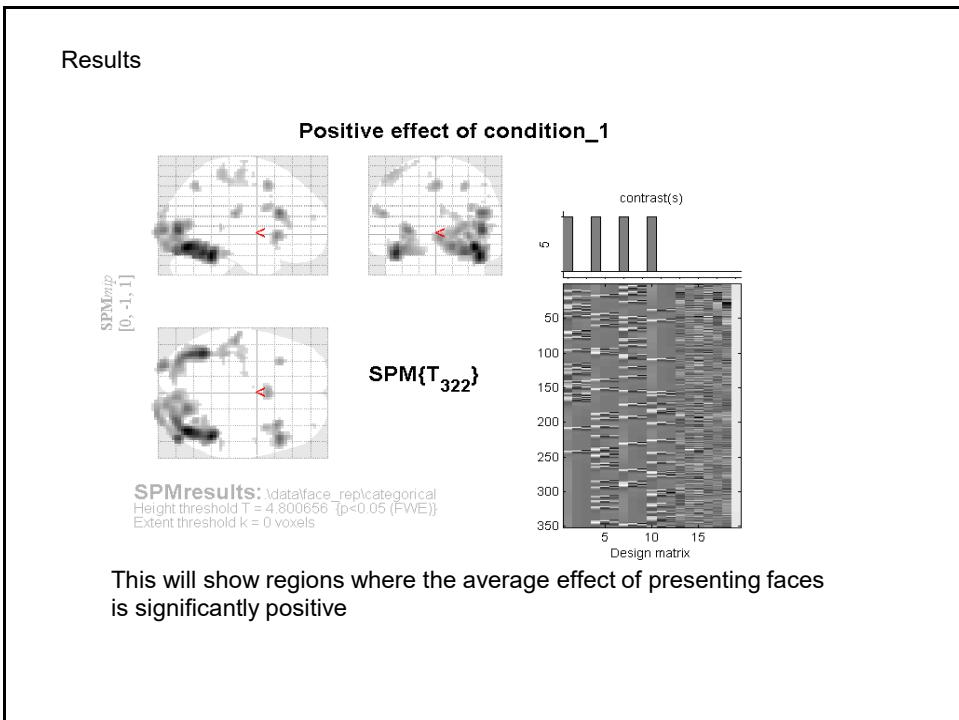
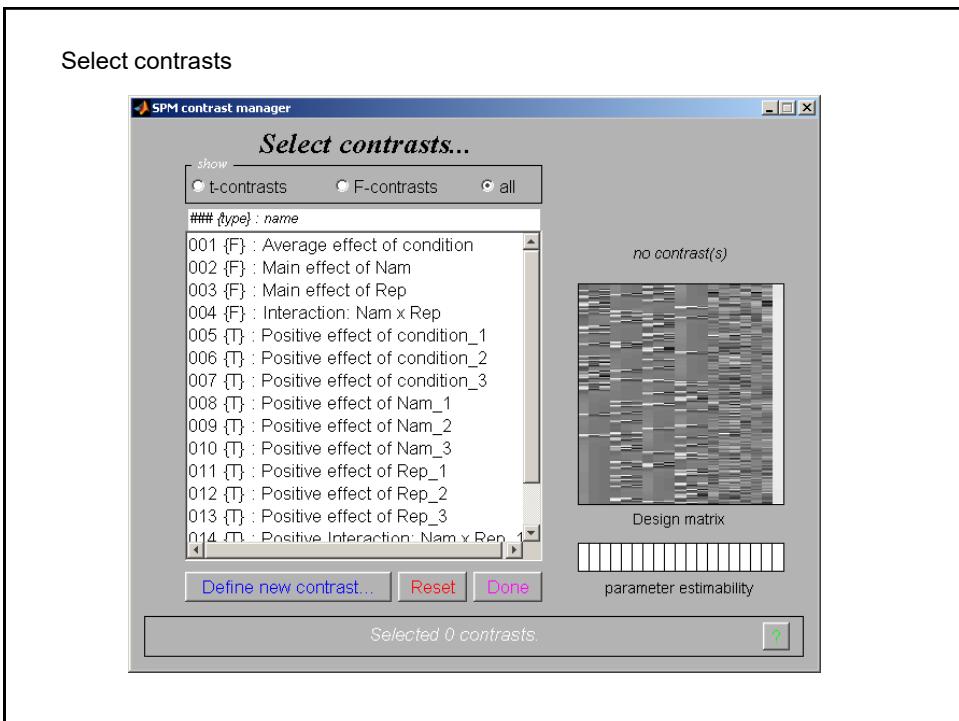
Exemple:
Face repetition paradigm











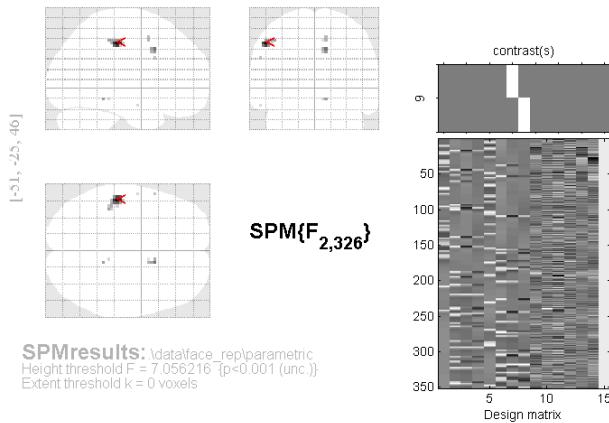
Results

Statistics: p-values adjusted for search volume

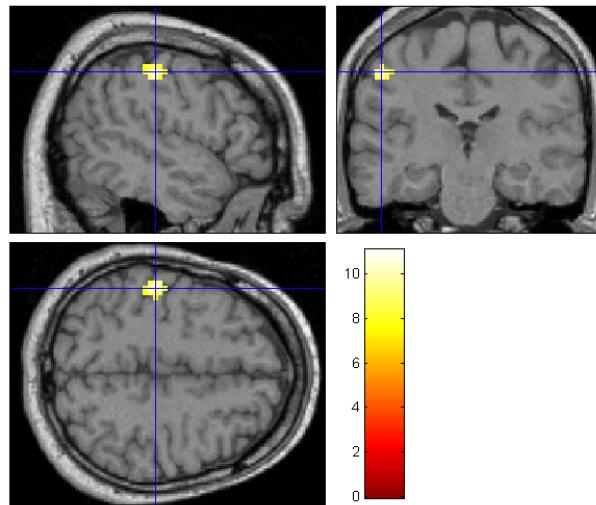
set-level	<i>p</i>	<i>c</i>	cluster-level				peak-level				mm mm mm mm			
			<i>p</i> _{FWE-corr}	<i>q</i> _{FDR-corr}	<i>k</i> _E	<i>p</i> _{uncorr}	<i>p</i> _{FWE-corr}	<i>q</i> _{FDR-corr}	<i>T</i>	(<i>Z</i>)	<i>p</i> _{uncorr}	mm	mm	mm
0.000	16	0.000	0.000	1338	0.000	0.000	0.000	14.45	Inf	0.000	39 -70 -14			
			0.000	0.000	14.04	Inf	0.000	45 -46 -23						
			0.000	0.000	11.25	Inf	0.000	48 -79 4						
			0.000	0.000	11.13	Inf	0.000	-39 -67 -20						
			0.000	0.000	8.95	Inf	0.000	-33 -79 -14						
			0.000	0.000	9.95	Inf	0.000	48 23 22						
			0.000	0.000	6.89	6.65	0.000	36 11 28						
			0.000	0.000	8.96	Inf	0.000	-27 -94 4						
			0.000	0.000	8.55	Inf	0.000	33 20 -2						
			0.001	0.017	5.76	5.61	0.000	54 17 1						
			0.004	0.082	5.41	5.29	0.000	51 20 -14						
			0.000	0.000	7.94	7.58	0.000	-54 -22 22						
			0.000	0.000	7.71	7.38	0.000	0 11 52						
			0.000	0.000	7.16	6.89	0.000	-33 23 -2						
			0.000	0.001	7.02	6.76	0.000	30 -61 52						
			0.000	0.000	6.61	6.39	0.000	-45 -34 61						
			0.000	0.002	6.24	6.06	0.000	-39 -16 64						
			0.000	0.003	6.14	5.96	0.000	-51 -28 55						

Results

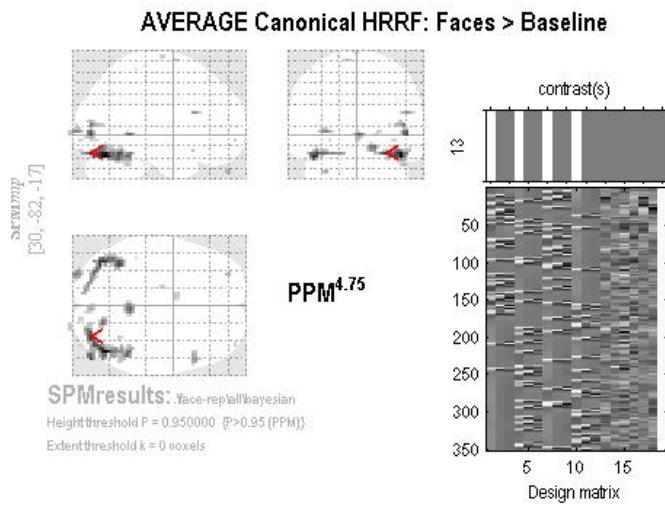
Famous Lag (masked [incl.] by Positive effect of condition_1 at p=0.05)



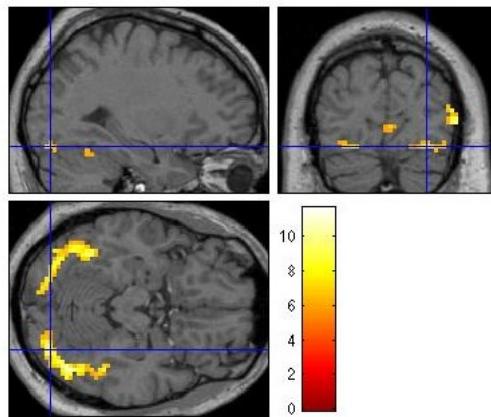
Results



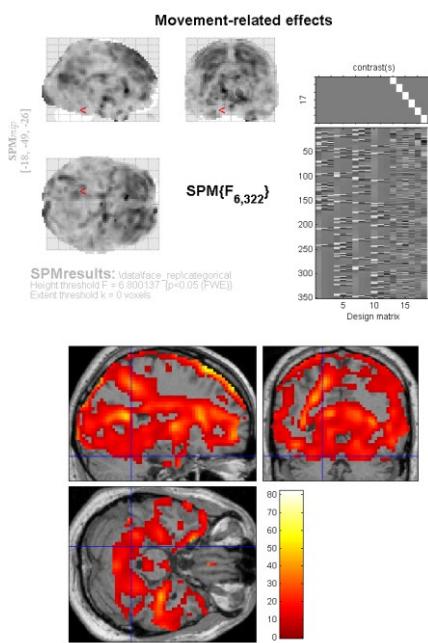
Results



Results



Results



Référence

- [http://www.fil.ion.ucl.ac.uk/spm/course/
slides14-may/](http://www.fil.ion.ucl.ac.uk/spm/course/slides14-may/)