Aalto University
School of Science

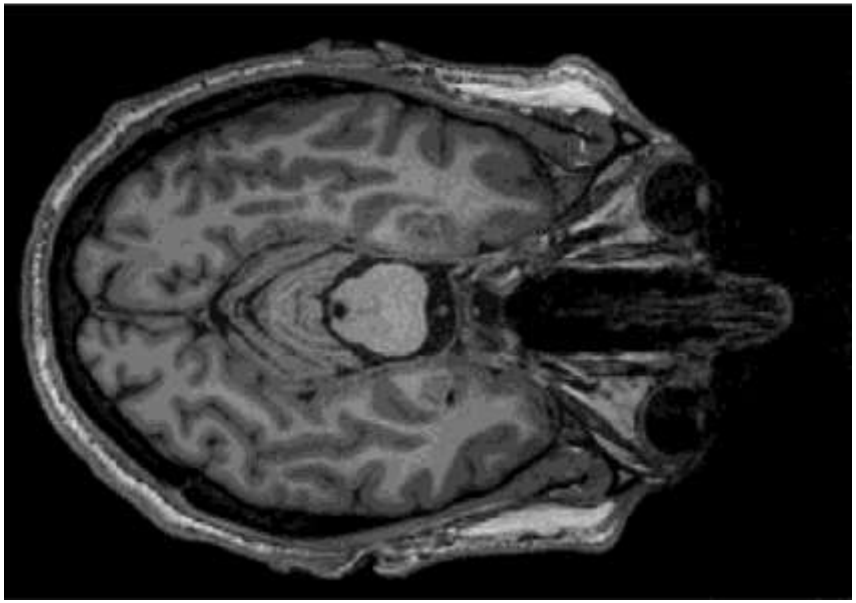
General Linear Model for random-effects fMRI

Linda Henriksson

Department of Neuroscience and Biomedical Engineering,
Aalto University, Finland

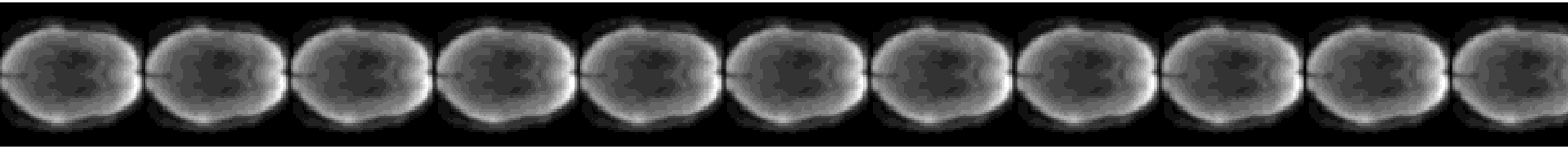
fMRI data

MRI



3T MR scanner @ Otaniemi (ani.aalto.fi)

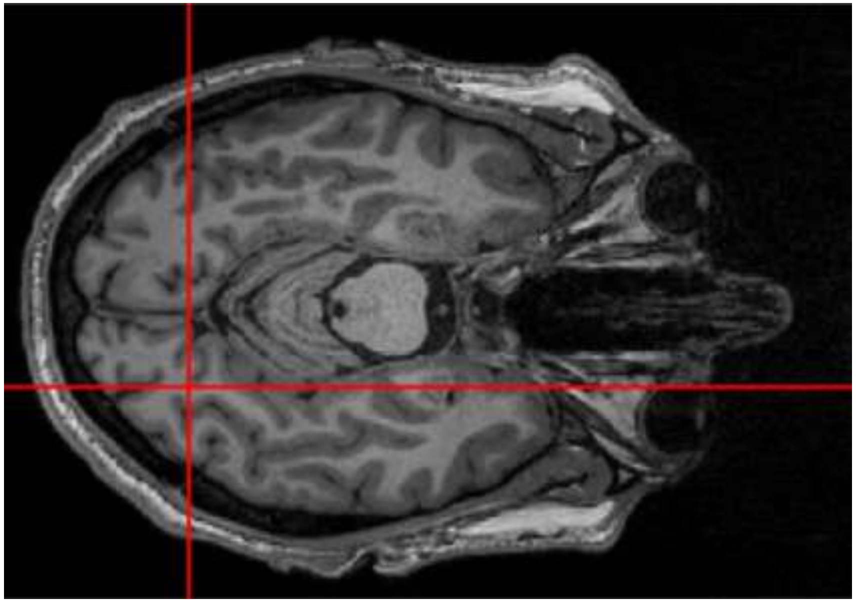
fMRI

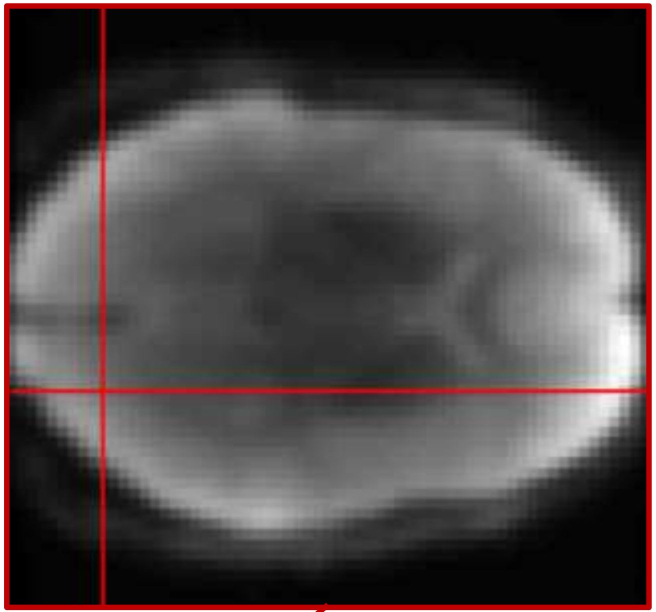


time

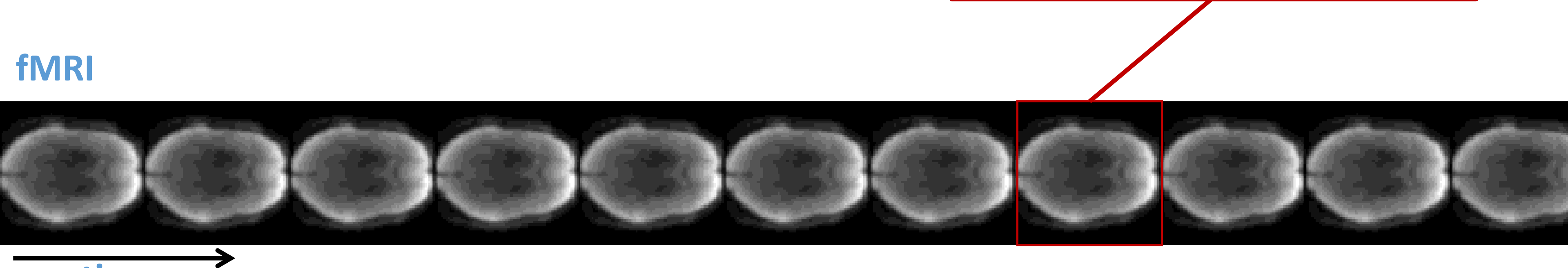
fMRI data

MRI





fMRI

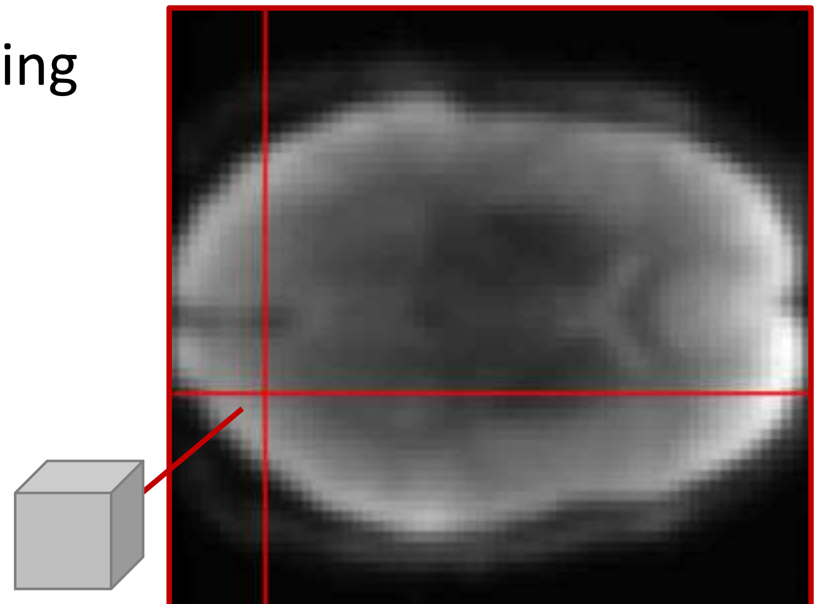


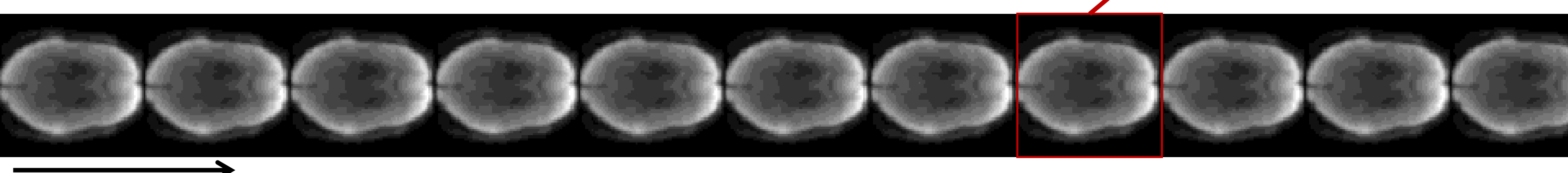
time

fMRI data

- fMRI is a technique for measuring and mapping brain activity
- fMRI data are time series of image volumes
- Typical dimensions:
  - 64 x 64 voxels within a slice
  - about 30 slices
  - 150 – 300 volume images (time points)

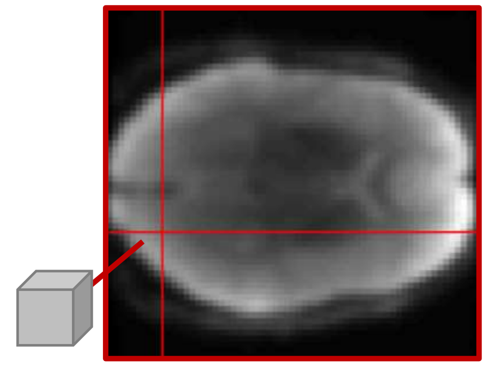
122 880 voxels!

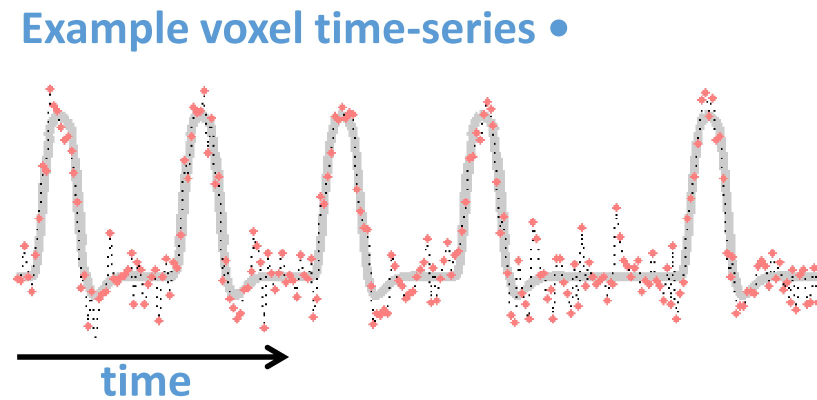


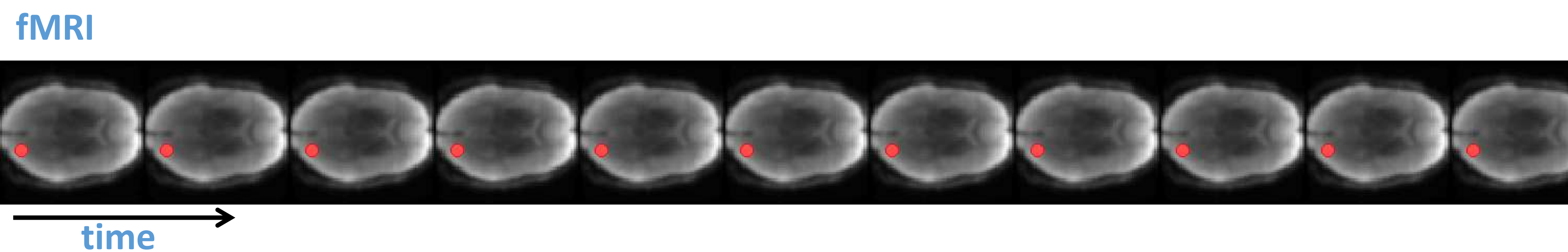


fMRI data

- fMRI is a technique for measuring and mapping brain activity
- fMRI data are time series of image volumes
- Voxel-by-voxel time-series analysis







fMRI data

- fMRI is an **indirect** measure of neural activity.
- Absolute **magnitude** of the fMRI response depends on
  - voxel size, proportion of gray matter in a voxel, local vascular density, number of voxels in a cluster, physiological variability in signal strength, field strength,...
- fMRI is a **relative** measure of activity: typically you need a control/baseline condition to compare with.

Typical questions for an fMRI experiment

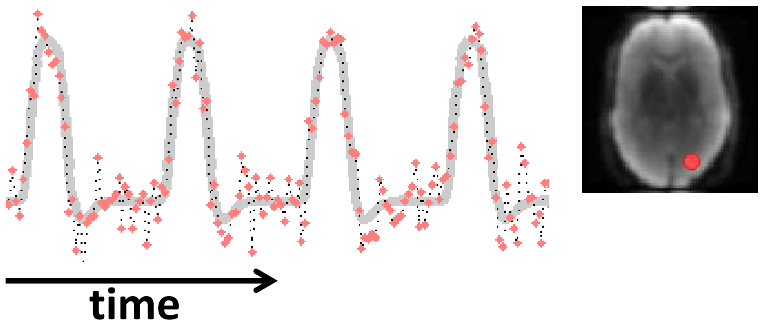
- Which brain regions are activated during a specific task? (functional **localization**)
- Is the response to task A larger than the response to task B **in a particular brain region**?
- Spatial **maps**
- …

“**Where**” questions

Preprocessing of fMRI data

- Typical preprocessing steps:
  - Data format conversion
  - Slice timing correction
  - Movement correction
  - Distortion correction
  - Spatial smoothing
  - Spatial normalization

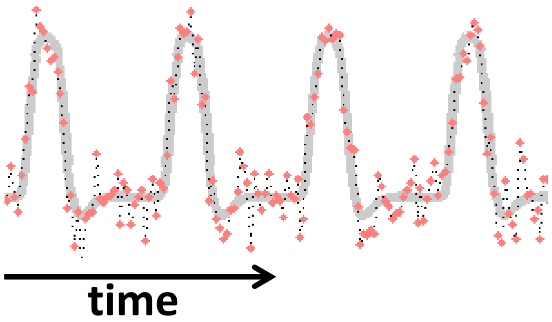
Example voxel time-series

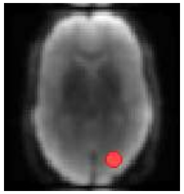


Preprocessing of fMRI data

- Typical preprocessing steps:
  - Data format conversion
  - Slice timing correction
  - **Movement correction**
  - *(Distortion correction)*
  - **Spatial smoothing to increase SNR**
  - **Spatial normalization to common brain atlas**
- Main goal of preprocessing is to **reduce non-task-related (uninteresting) variability in the data**

Example voxel time-series

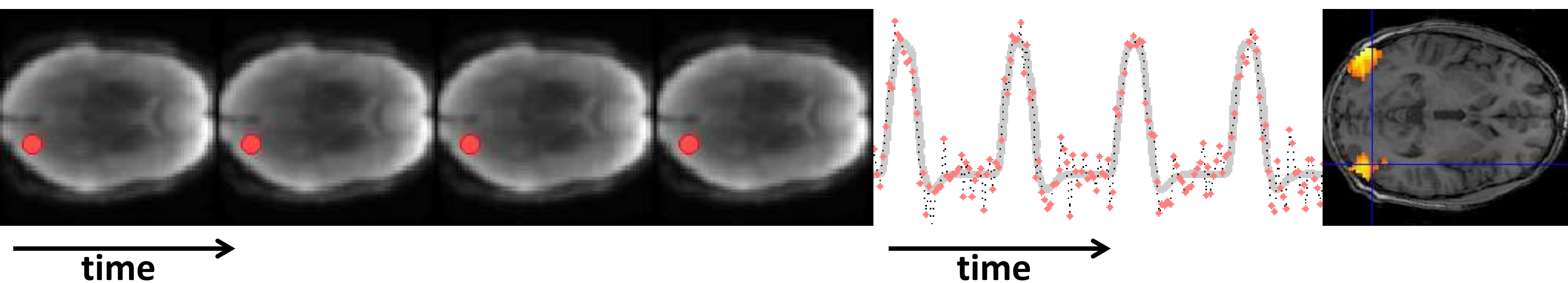




Statistical analysis of fMRI data

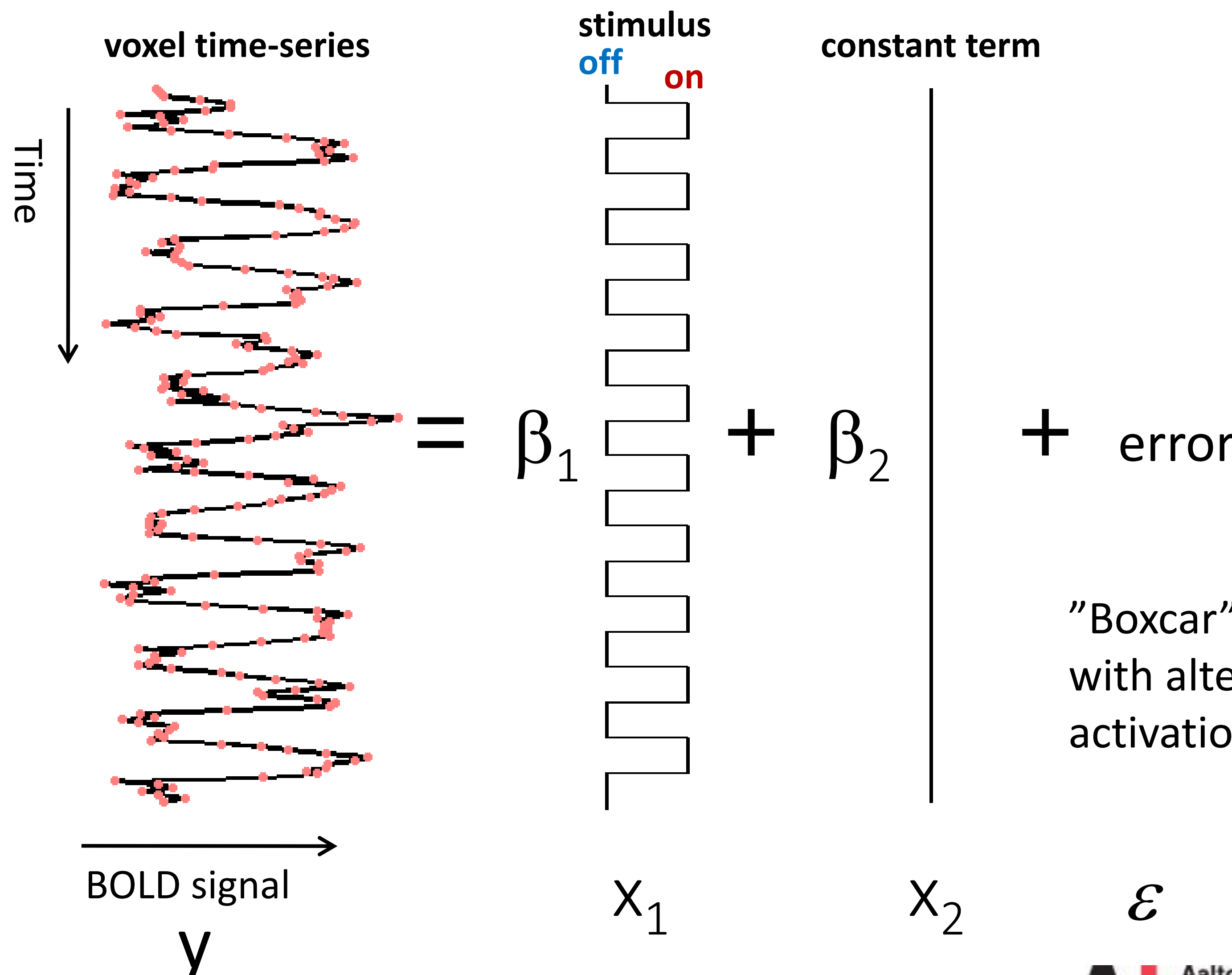
- Typical question: **Which brain regions (voxels) are activated** by the stimulus or task?
- Standard approach*:
  1. Construct a model of predicted brain activity
  2. Fit the model to data
  3. Perform statistical tests

  *repeat for each voxel



General linear model (GLM)

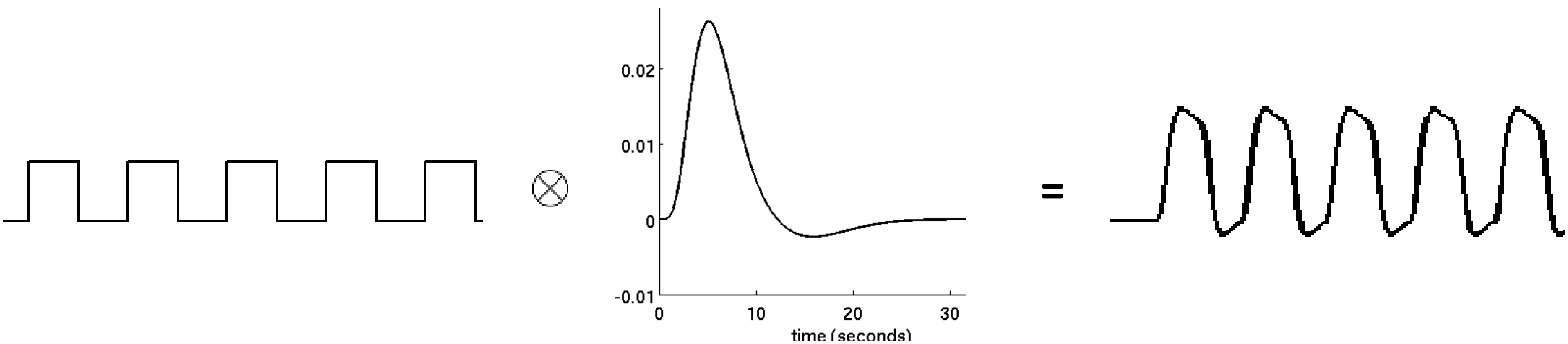
$$y = X\beta + \varepsilon$$



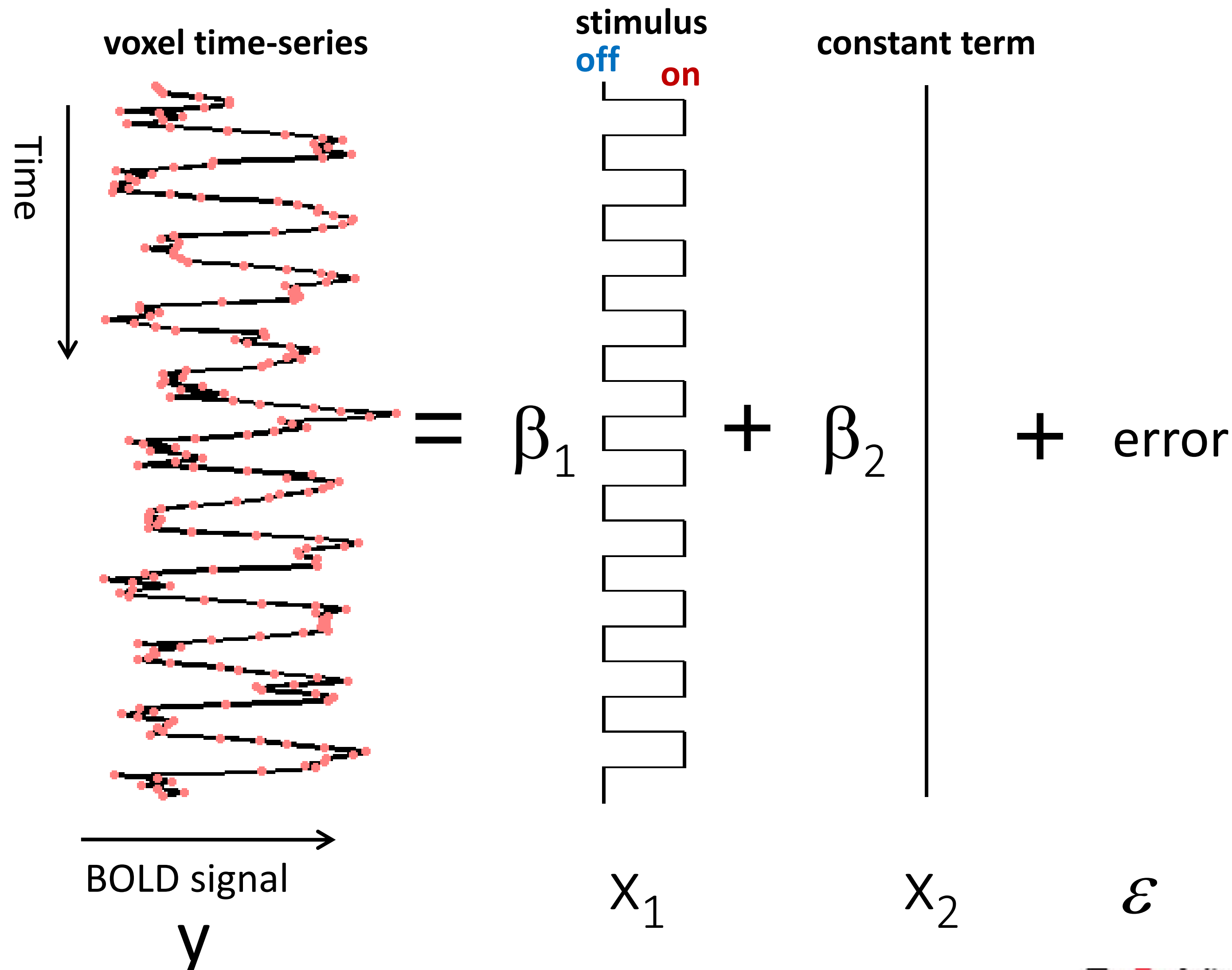
"Boxcar" model with alternating activation & rest

a better model

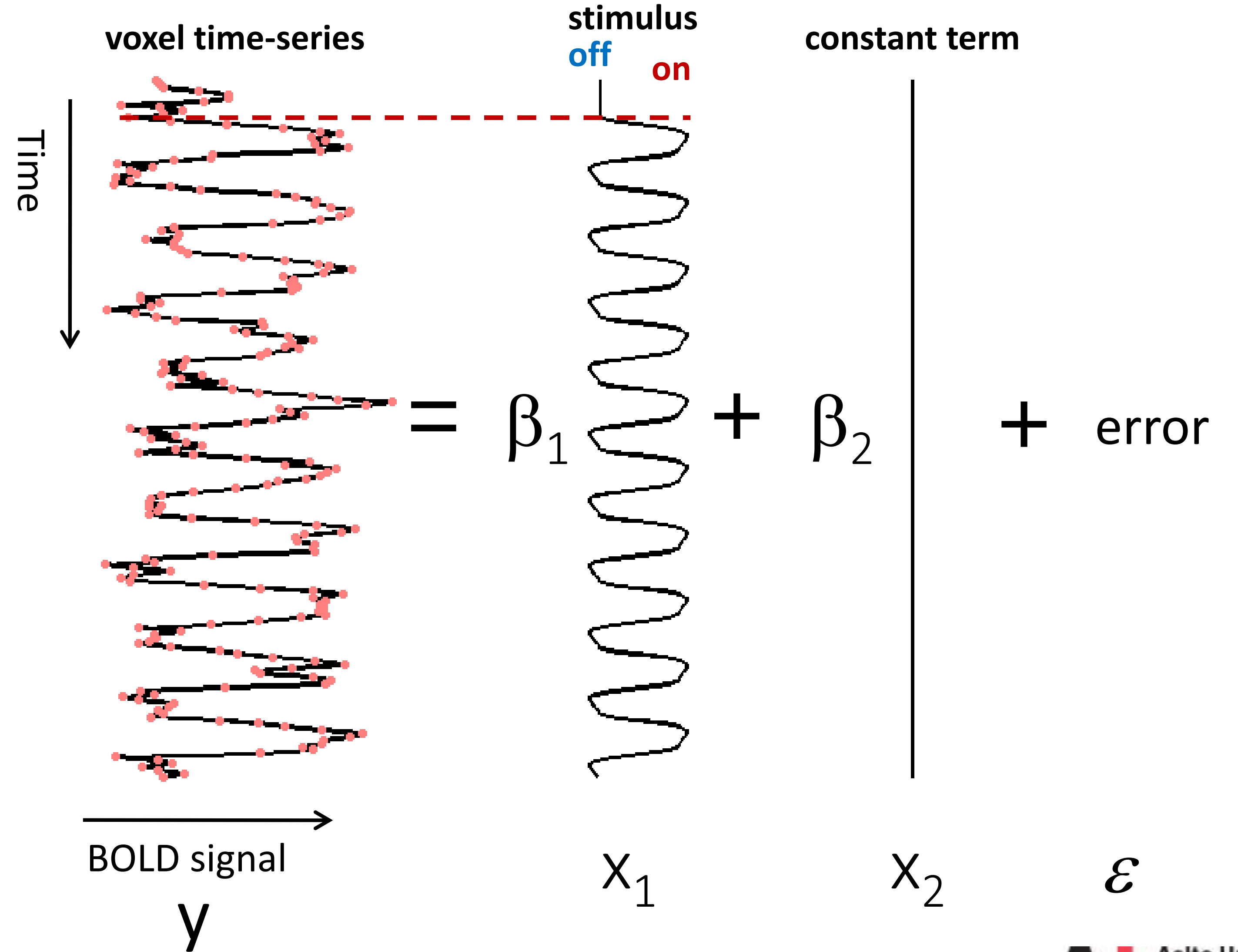
- Take into account the **shape of the BOLD response**: convolve the stimulus timing vector with a model of the hemodynamic response function (hrf)



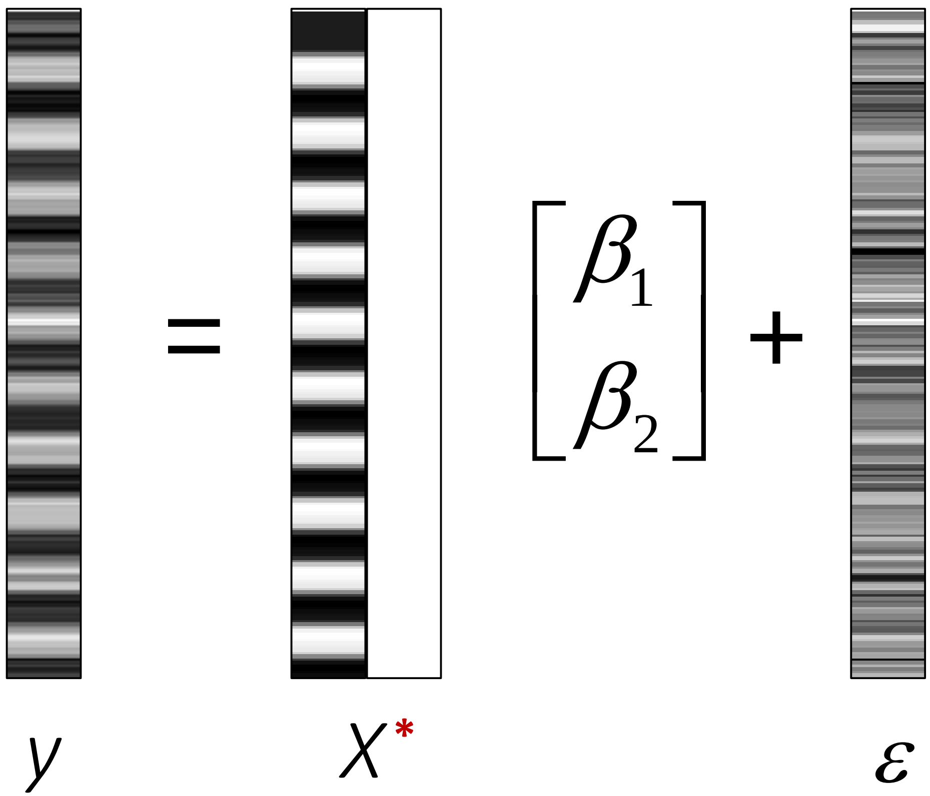
$$y = X\beta + \varepsilon$$



$$y = X\beta + \varepsilon$$

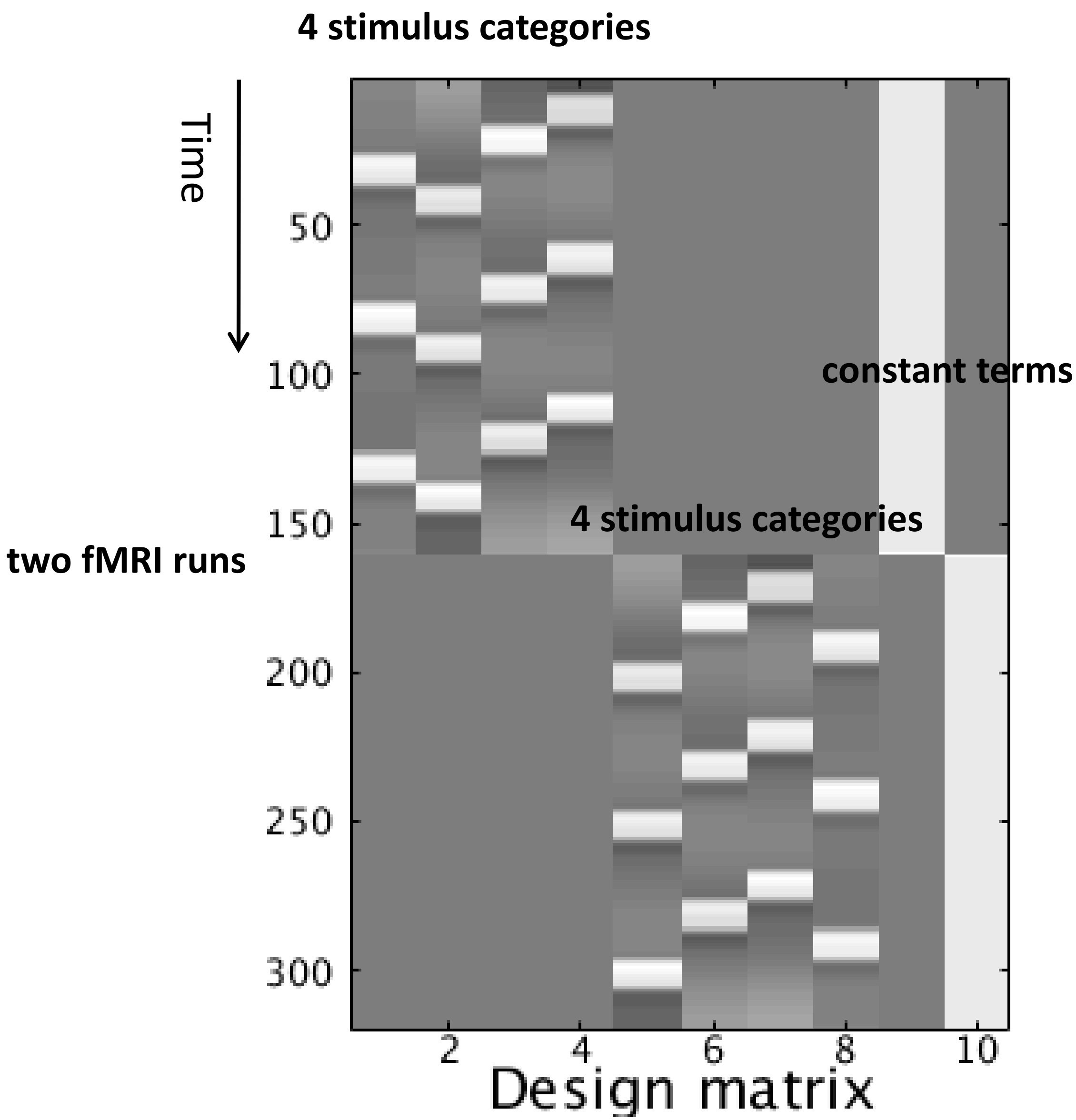


$$y = X\beta + \varepsilon$$



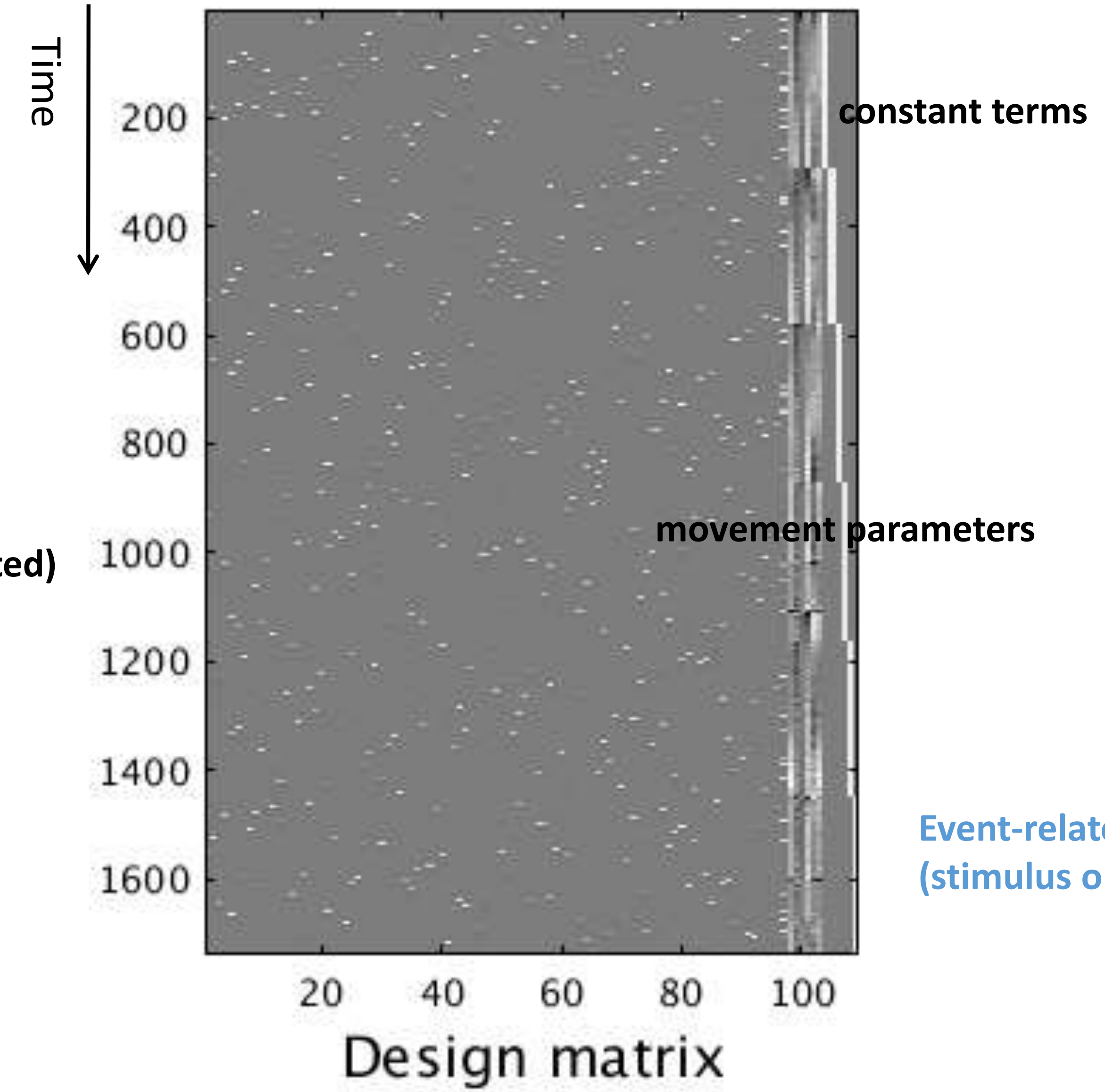
***this is a basic design matrix**—should model (add columns for) all known effects-of-interest and nuisance variables**

→ **Find parameters β** that best explain the data



Blocked fMRI design
(20s ’stimulus on’ periods)

96 stimulus types

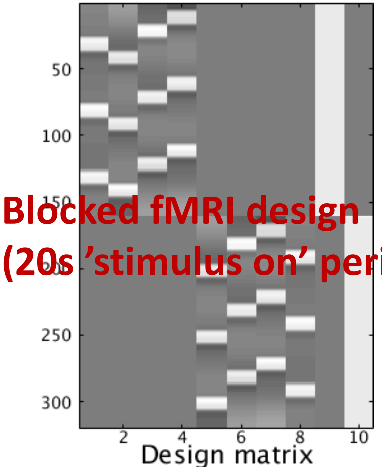


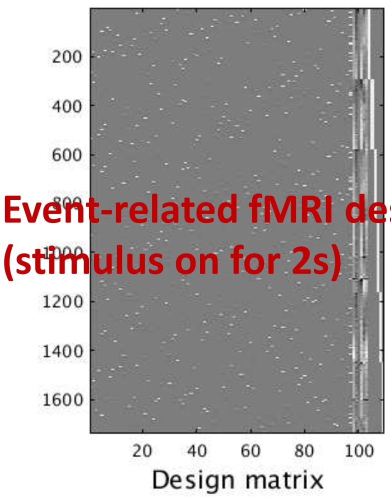
6 fMRI runs (concatenated)

Event-related fMRI design (stimulus on for 2s)

Experimental design? Design efficiency?

- Blocked designs
  - Multiple repetitions of stimuli of the same category (or task) shown in “blocks”
  - Good detection power
- Event-related designs
  - More stimulus types
  - Transient activity
  - Good estimation power
- For details on design efficiency, see http://imaging.mrc-cbu.cam.ac.uk/imaging/DesignEfficiency

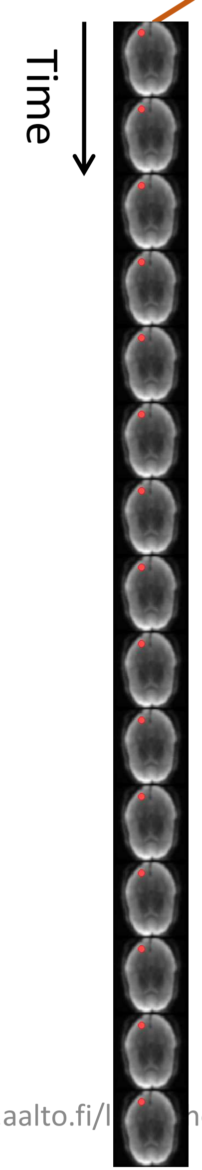


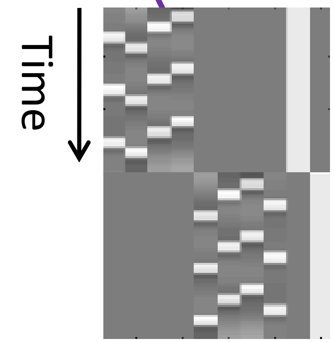


General linear model

$$y = X\beta + \varepsilon$$

- $y = X\beta + \varepsilon$
- GLM aims to explain the variation in the measured fMRI time-course in terms of a linear combination of predictors (columns in the design matrix).
- Find parameters β that best explain the data by minimizing the sum of the squared error values ($\sum \varepsilon^2$).





Design matrix

fMRI data
(voxel-by-voxel time-series analysis)

Statistical inference

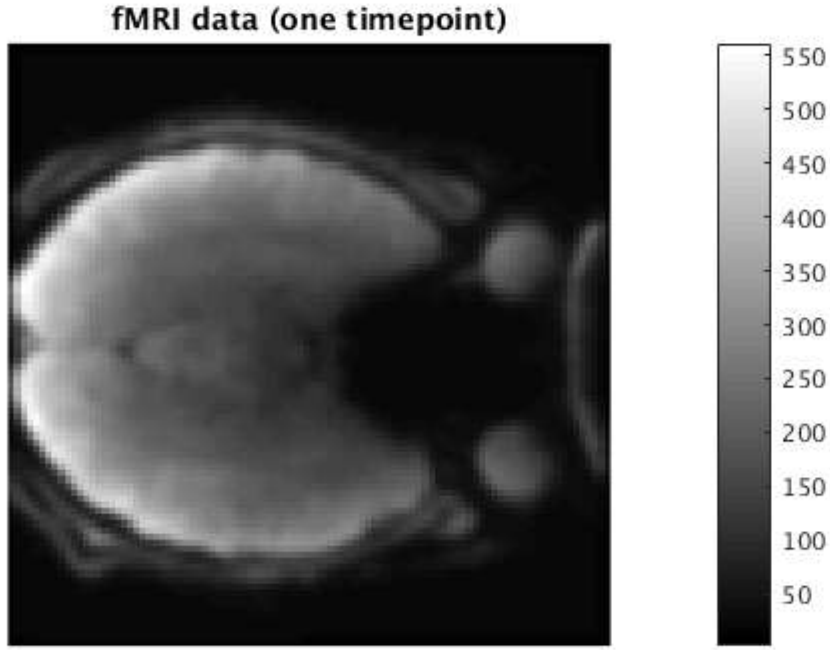
- Voxel-by-voxel hypothesis testing: **Does my model explain variance in the data?**
- Specify contrast (hypothesis), c, a linear combination of the estimated parameters (e.g., [1 0])
- Calculate, e.g., the T-statistic for the contrast **separately for each voxel**

$$T = \frac{c^T \hat{\beta}}{\mathrm{std}(c^T \hat{\beta})}$$

From fMRI data images to an activation map

$$y = X\beta + \varepsilon$$

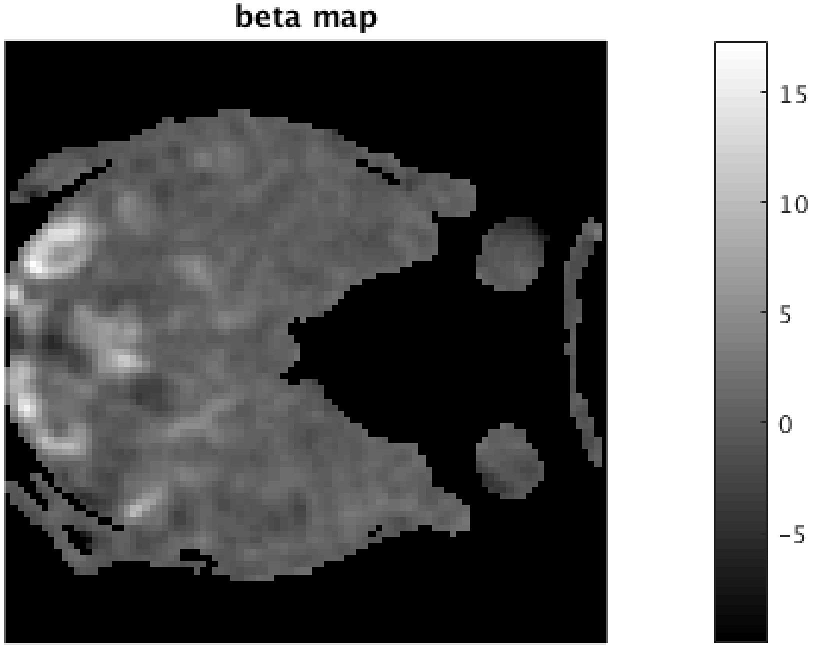
$$T = \frac{c^T \hat{\beta}}{\mathrm{std}(c^T \hat{\beta})}$$



From fMRI data images to an activation map

$$y = X\beta + \varepsilon$$

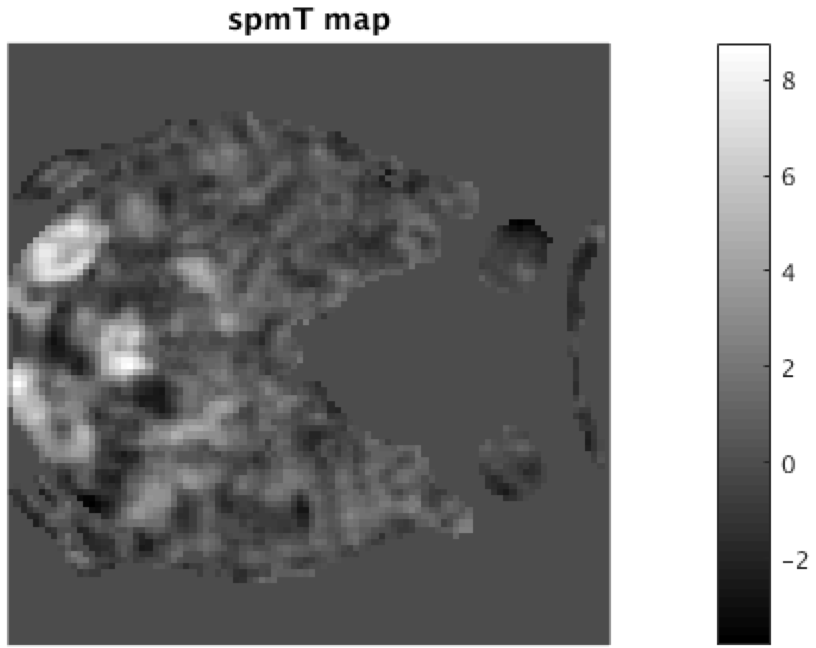
$$T = \frac{c^T \hat{\beta}}{\mathrm{std}(c^T \hat{\beta})}$$



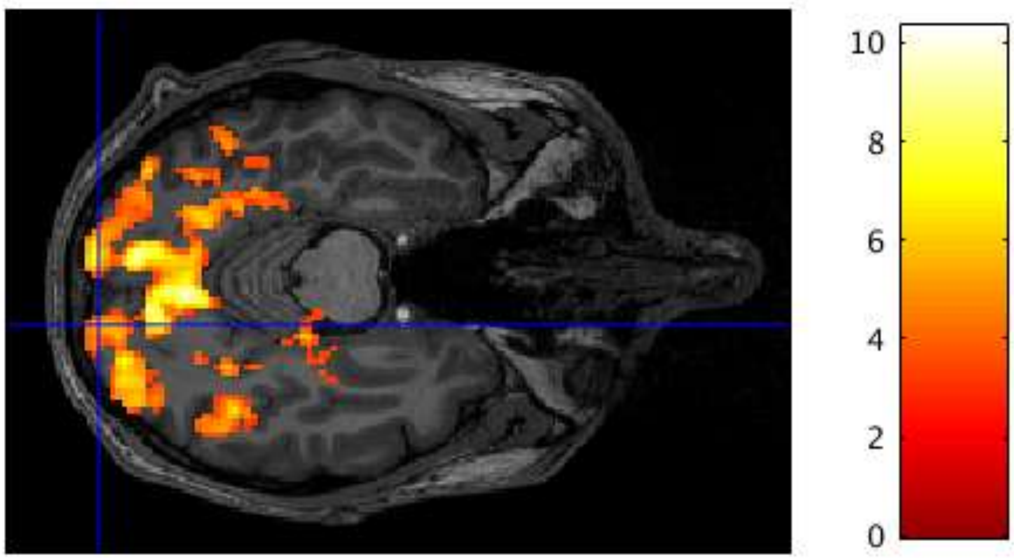
From fMRI data images to an activation map

$$y = X\beta + \varepsilon$$

$$T = \frac{c^T \hat{\beta}}{\mathrm{std}(c^T \hat{\beta})}$$



From fMRI data images to an activation map



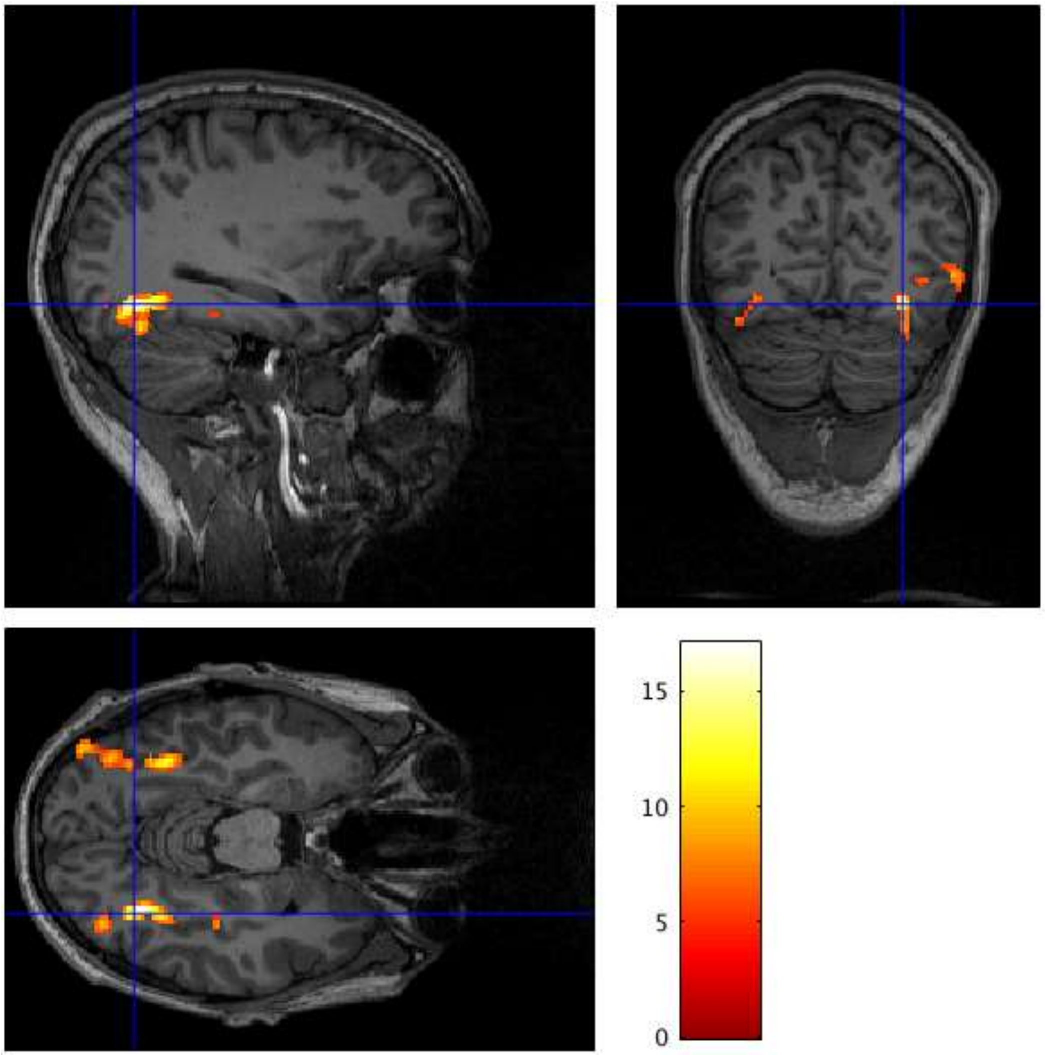
Statistical parametric map (T map) thresholded and overlaid on an anatomical MR image

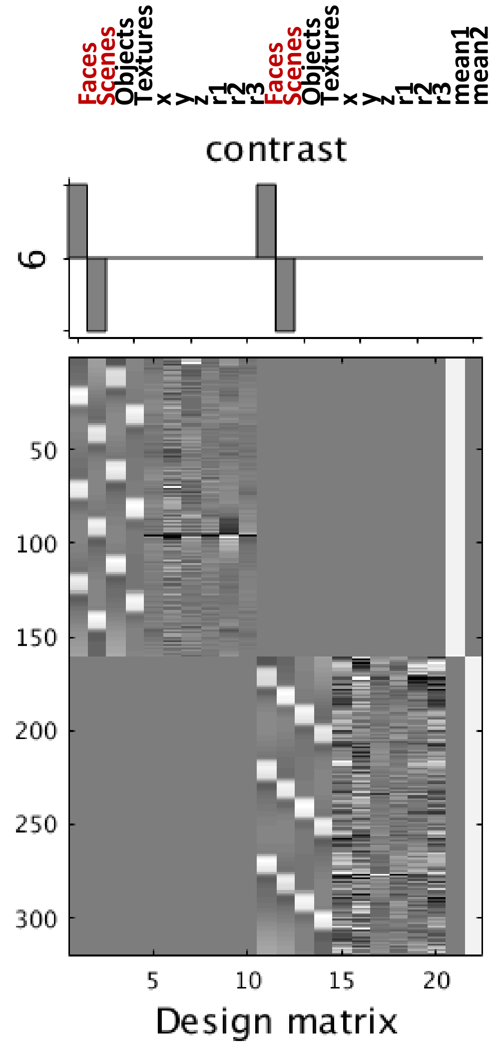
Aalto University
School of Science

Constructing contrasts

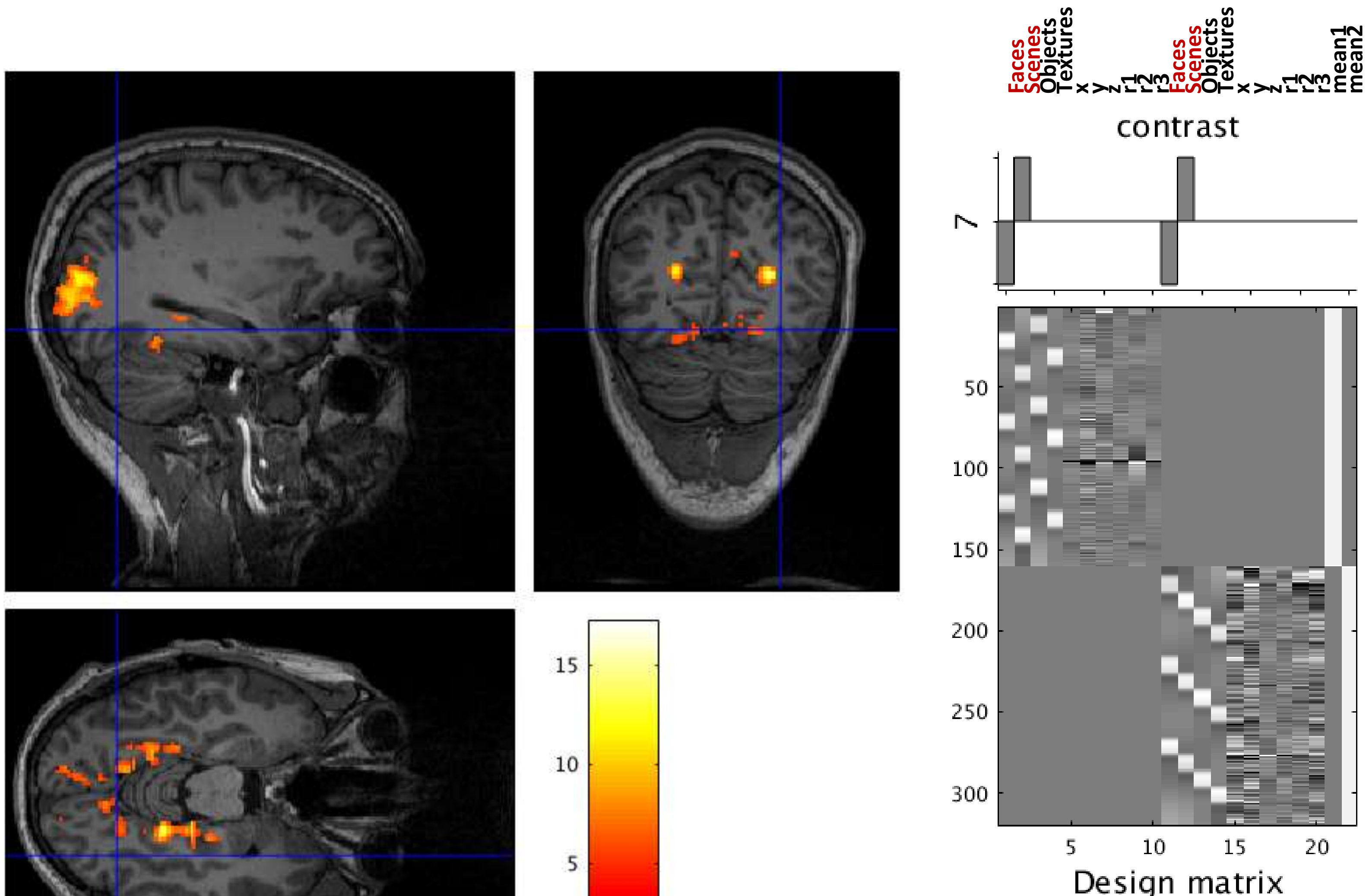
- Are there distinct regions that are specialized for a particular function (*e.g.*, perception of faces)?
- Before neuroimaging: focal lesions → specific perceptual problems
- Why? e.g., behavioral relevance of specific stimulus categories
- Always a network of brain regions
- In practice: **Where in the brain stimulus X evokes a larger response than stimulus Y**?

Faces > scenes

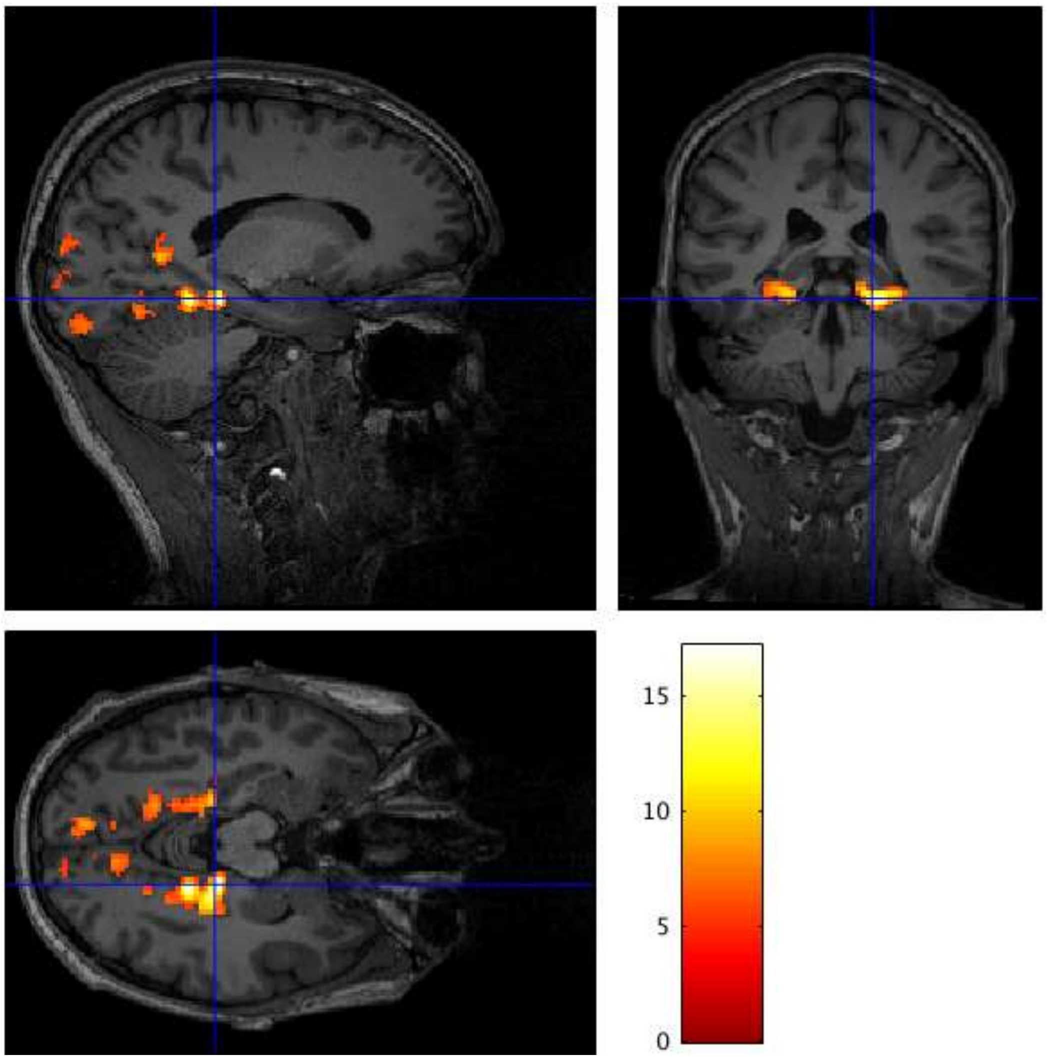


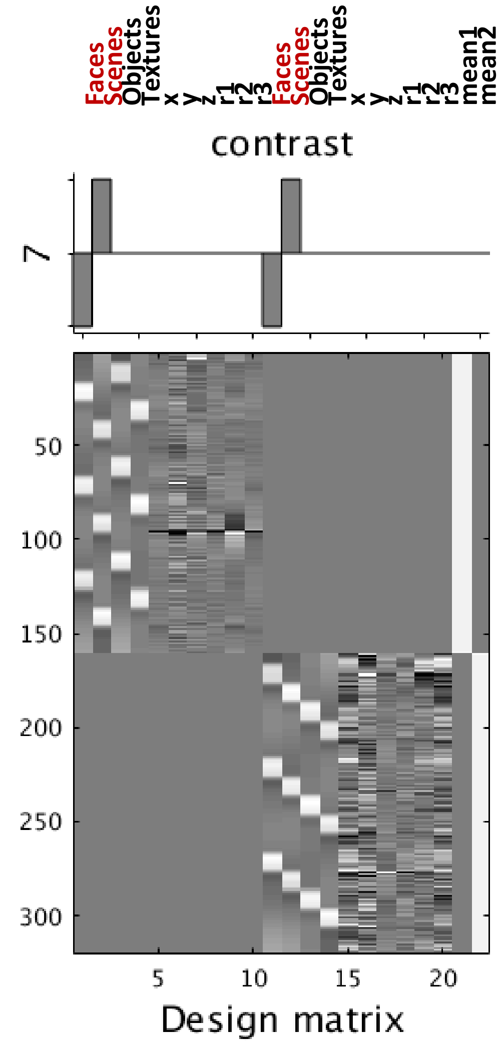


Scenes > faces

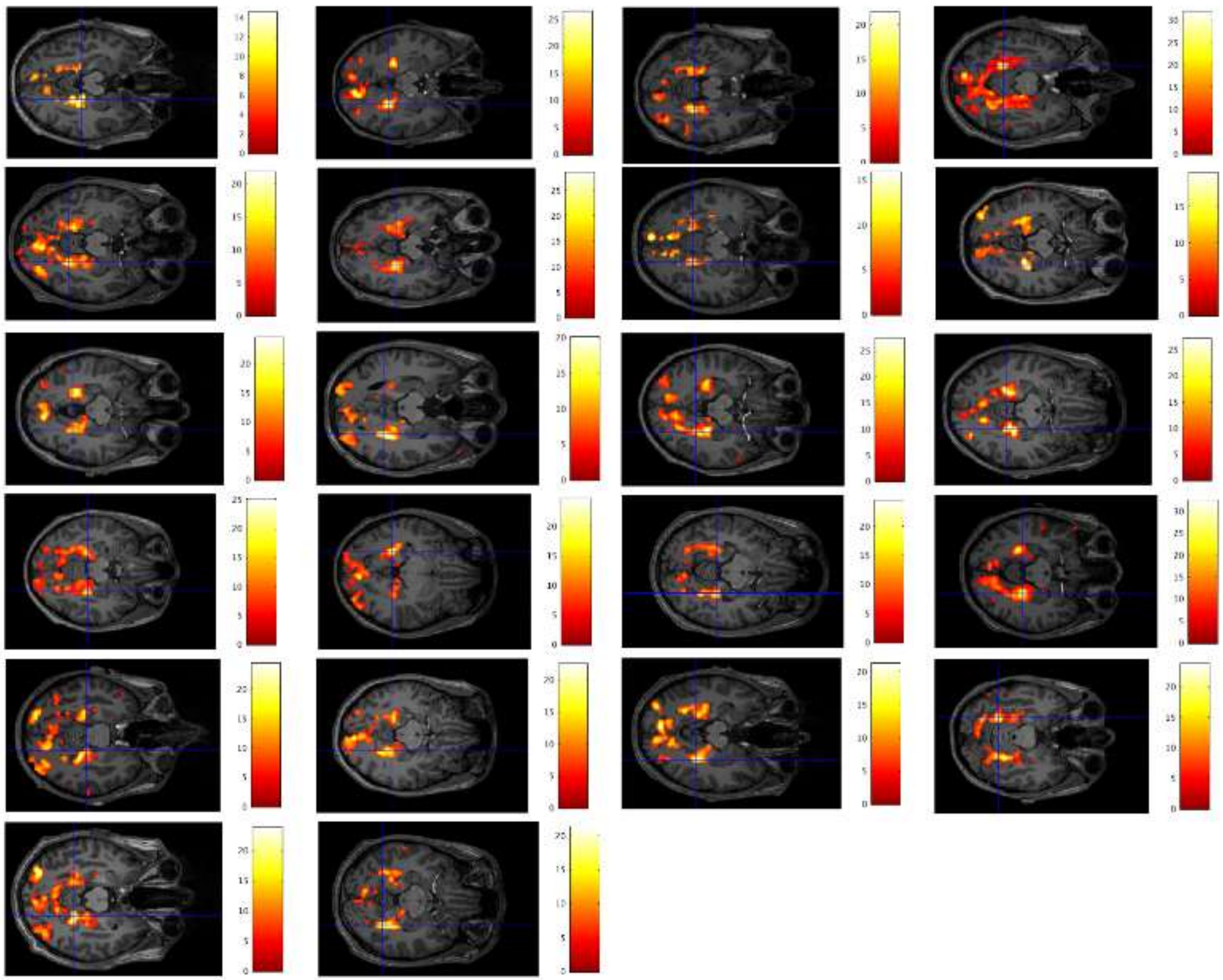


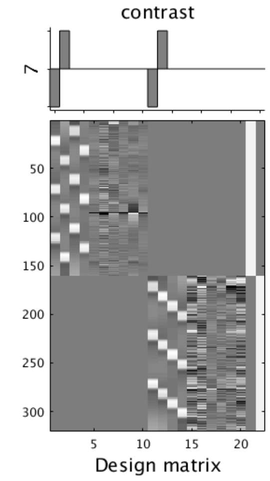
Scenes > faces





22 individuals (c: scenes > faces)



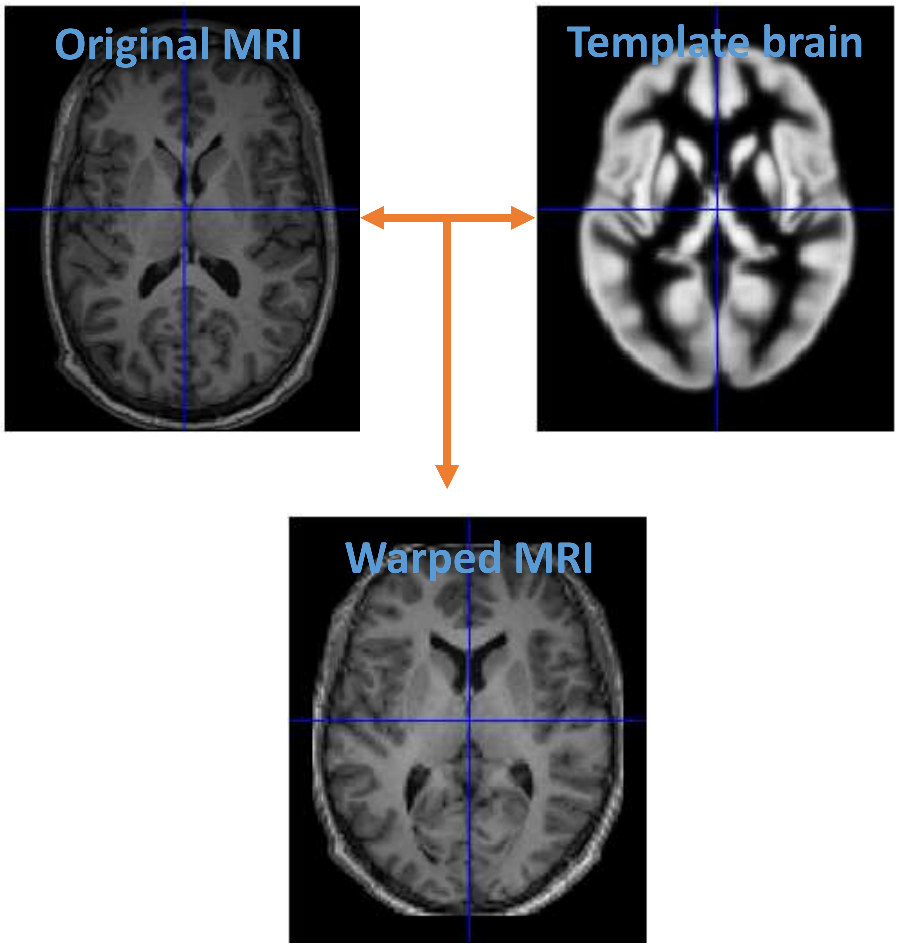


Group analysis?

- How to generalize findings from a sample of subjects to the population (from which the subjects have been drawn)?
- Whole-brain group analysis
  - Spatial normalization: match brains across individuals
    - Warp each individual brain data to a common space (Talairach; MNI)
    - Same voxel ≈ same location in the brain across individuals
  - Power of statistical analysis depends on the quality of normalization
    - Smoothing of functional data increases SNR and overlap of active brain regions across individuals (but aso **spreads activations** across sulci, **increses partial-volume effects** and **reduces spatial resolution**)

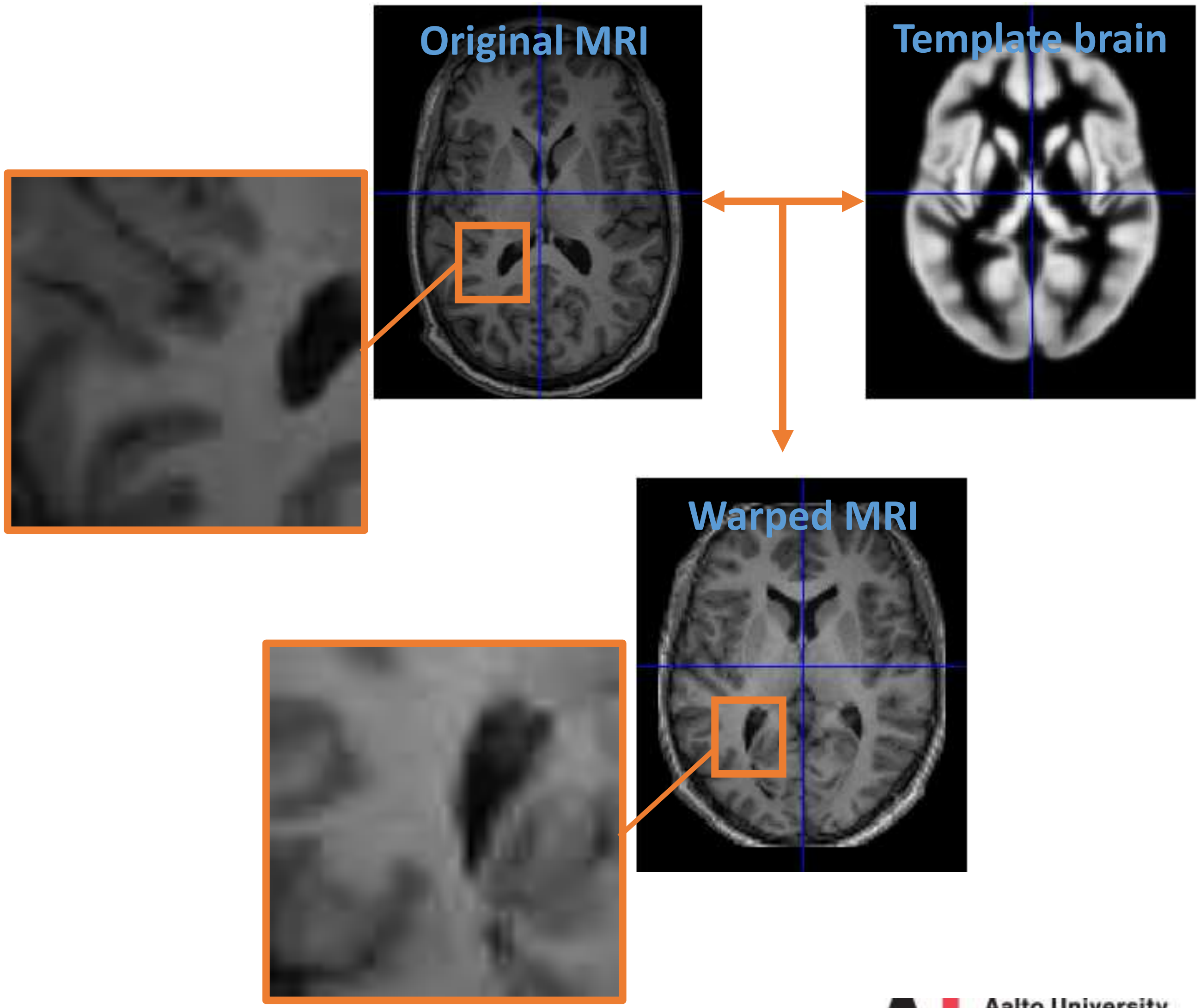
Spatial normalization in volume space

- Spatial normalization: match brains across individuals
  - Warp each individual brain data to a common space (Talairach; MNI)
  - Same voxel ≈ same location in the brain across individuals



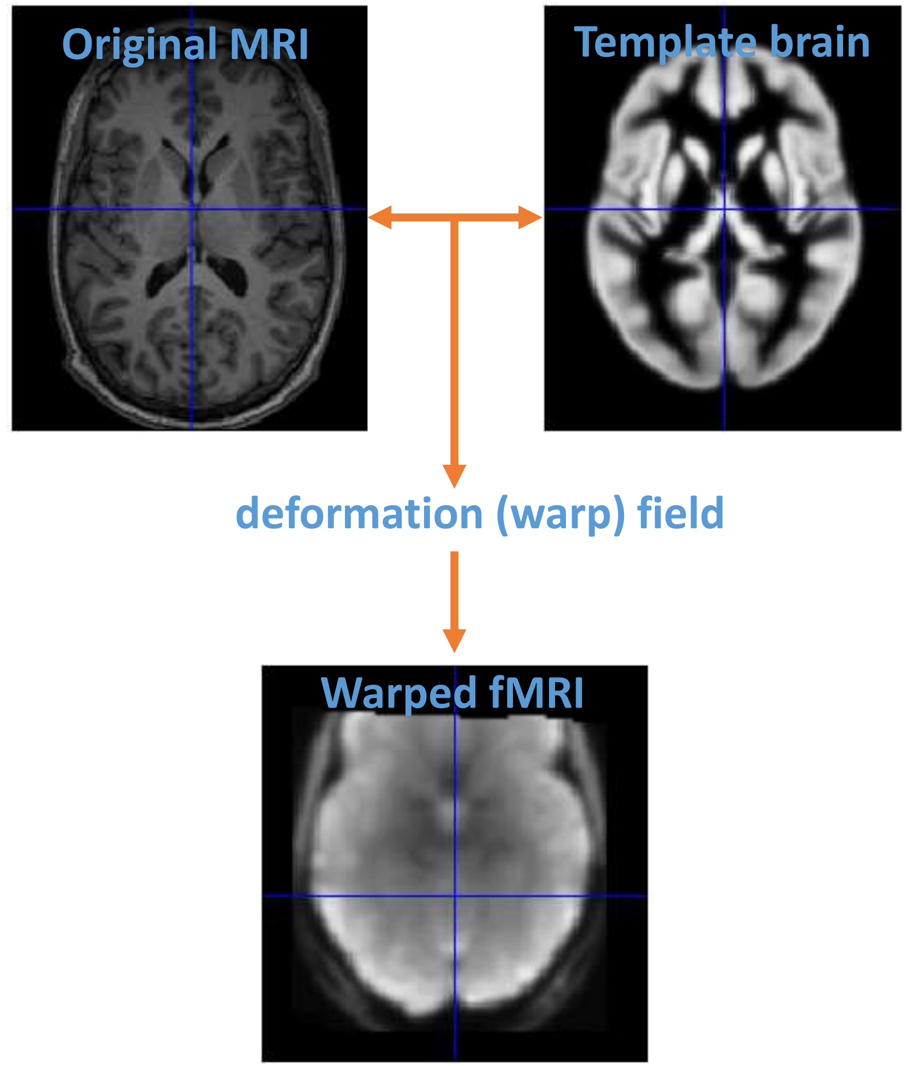
+ deformation (warp) field

Spatial normalization in volume space



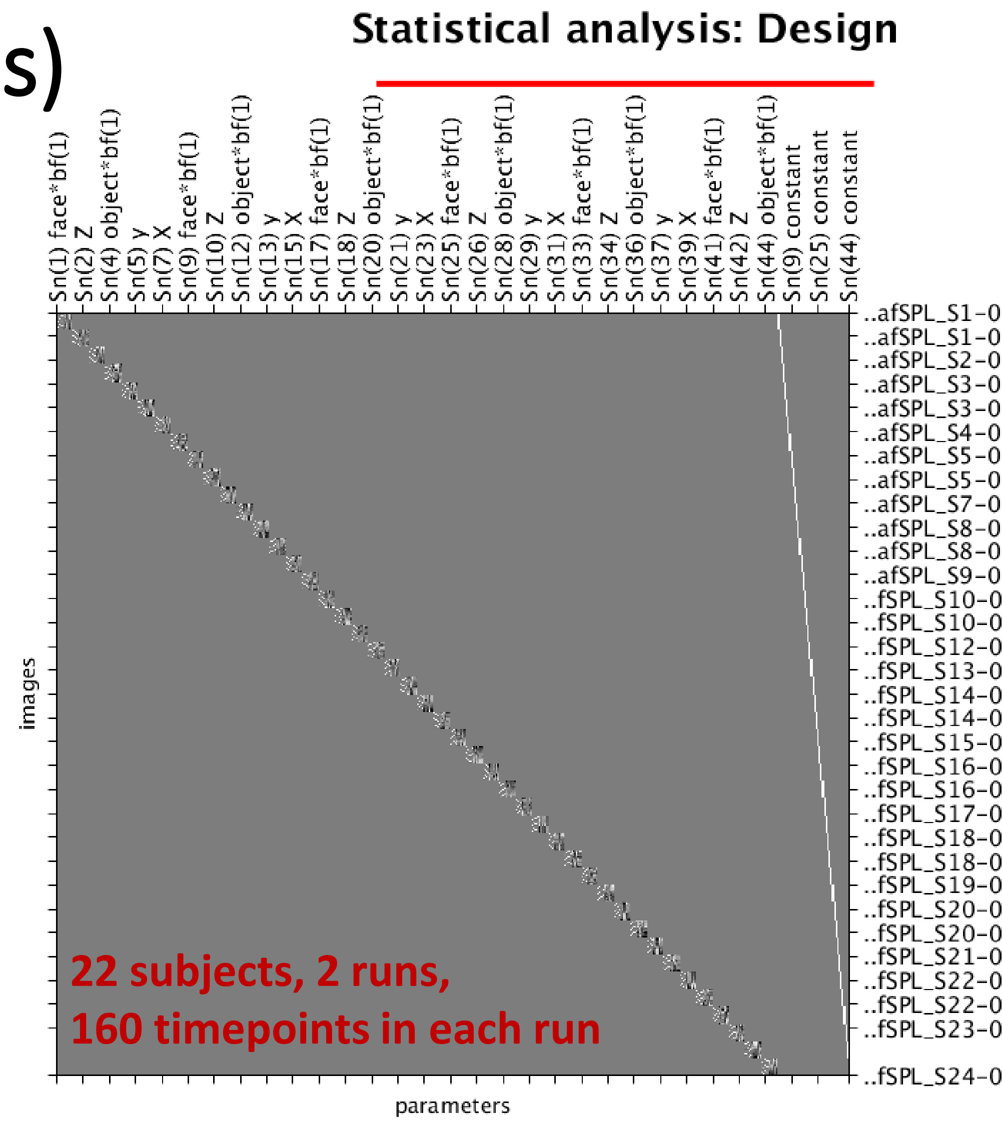
Spatial normalization in volume space

- Deformation (warp) field can be applied to any other data co-registered with the original MRI image (*e.g.*, functional data)



Concatenate all data (fixed-effects analysis)

- Data analyzed as originated from a single subject of a very long experiment
- High statistical power

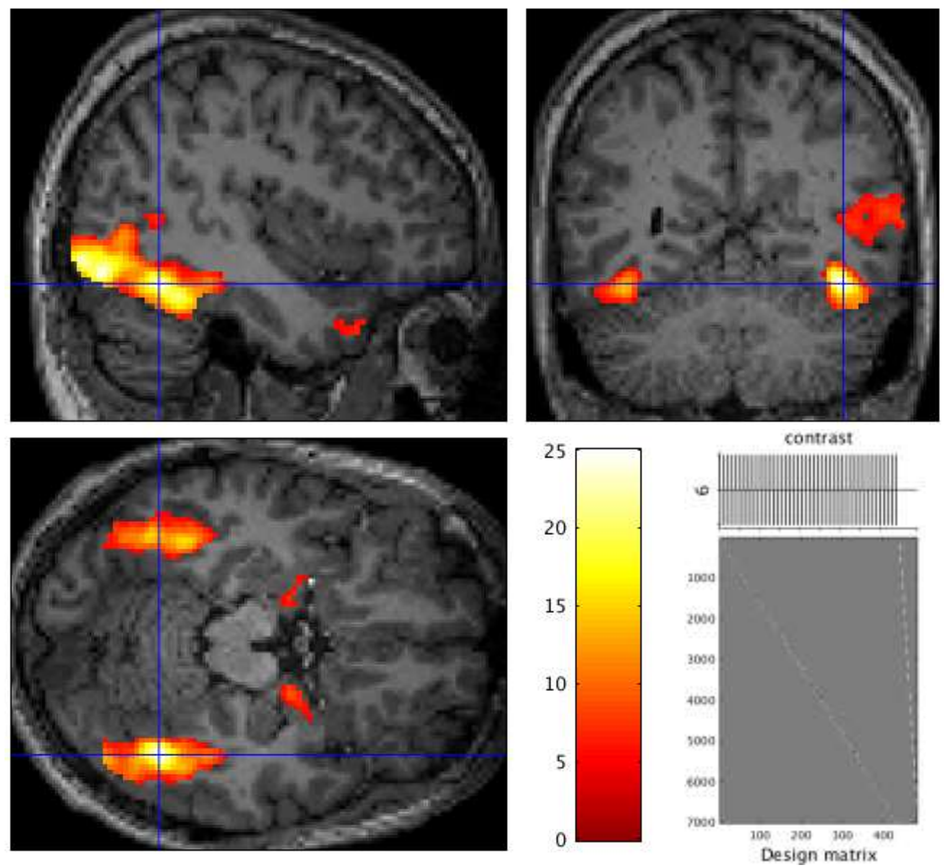


Concatenate all data (fixed-effects analysis)

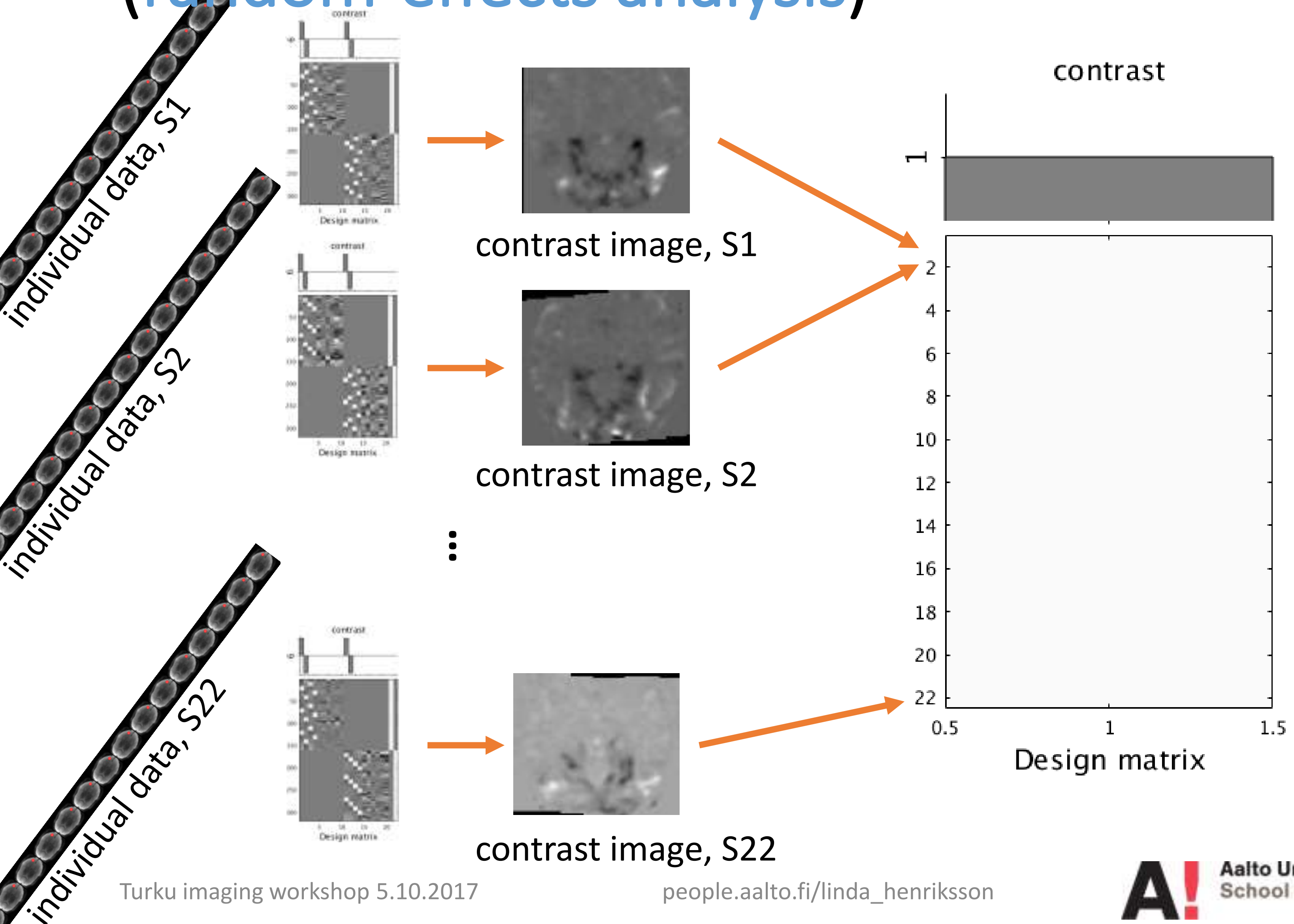
- Data analyzed as originated from a single subject of a very long experiment
- High statistical power
- Results **cannot be generalized to population**!

→Need to consider that subjects constitute a randomly drawn sample from a large population

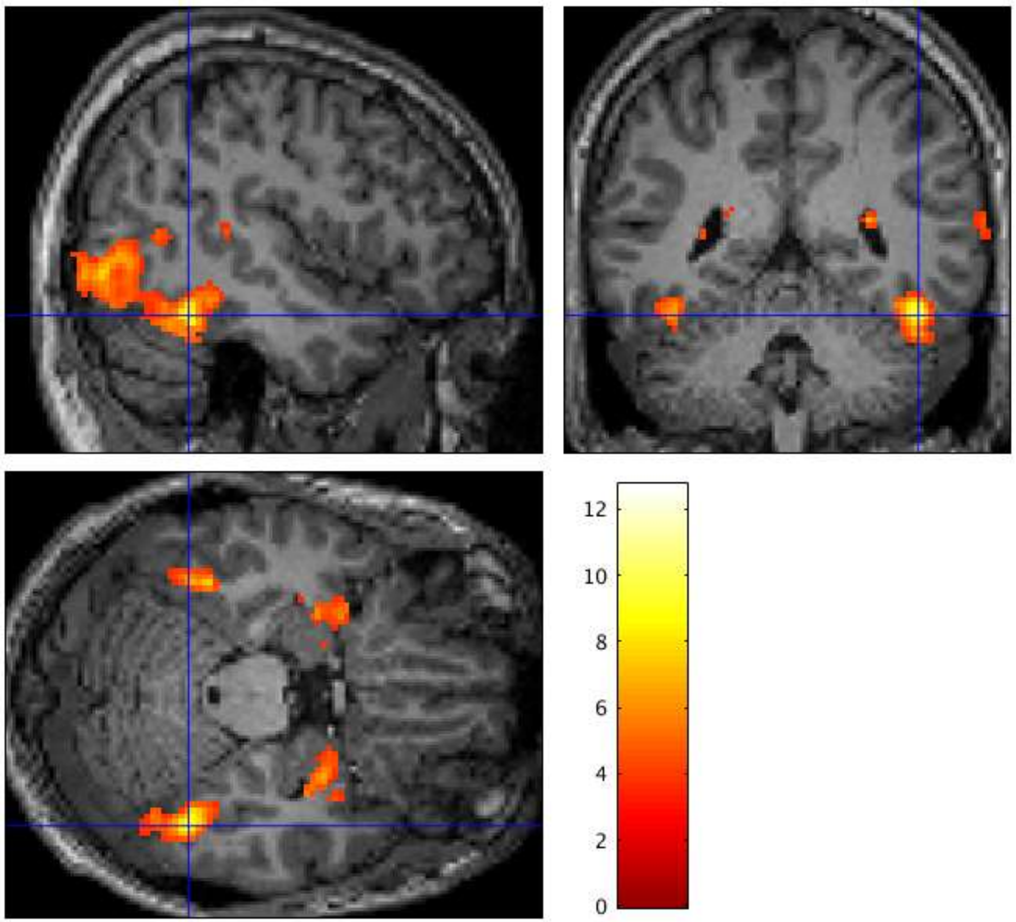
→**Random-effects analysis**

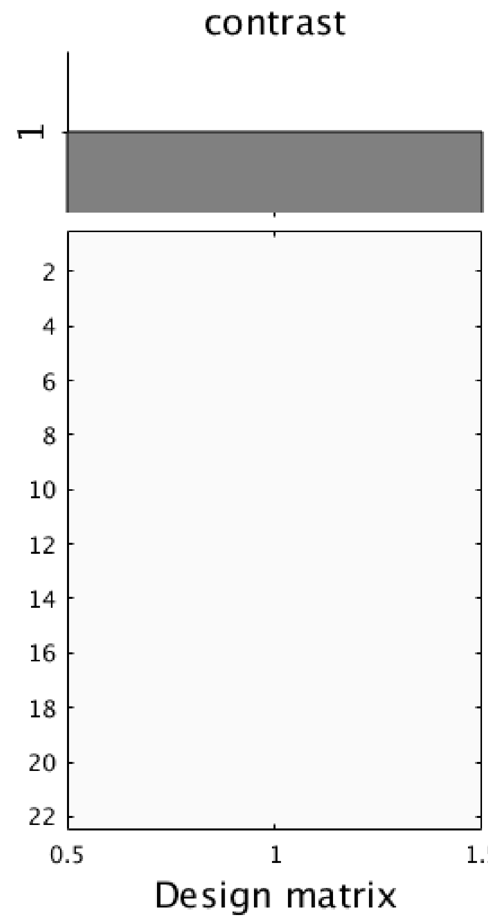


Account for between-subject variance (random-effects analysis)

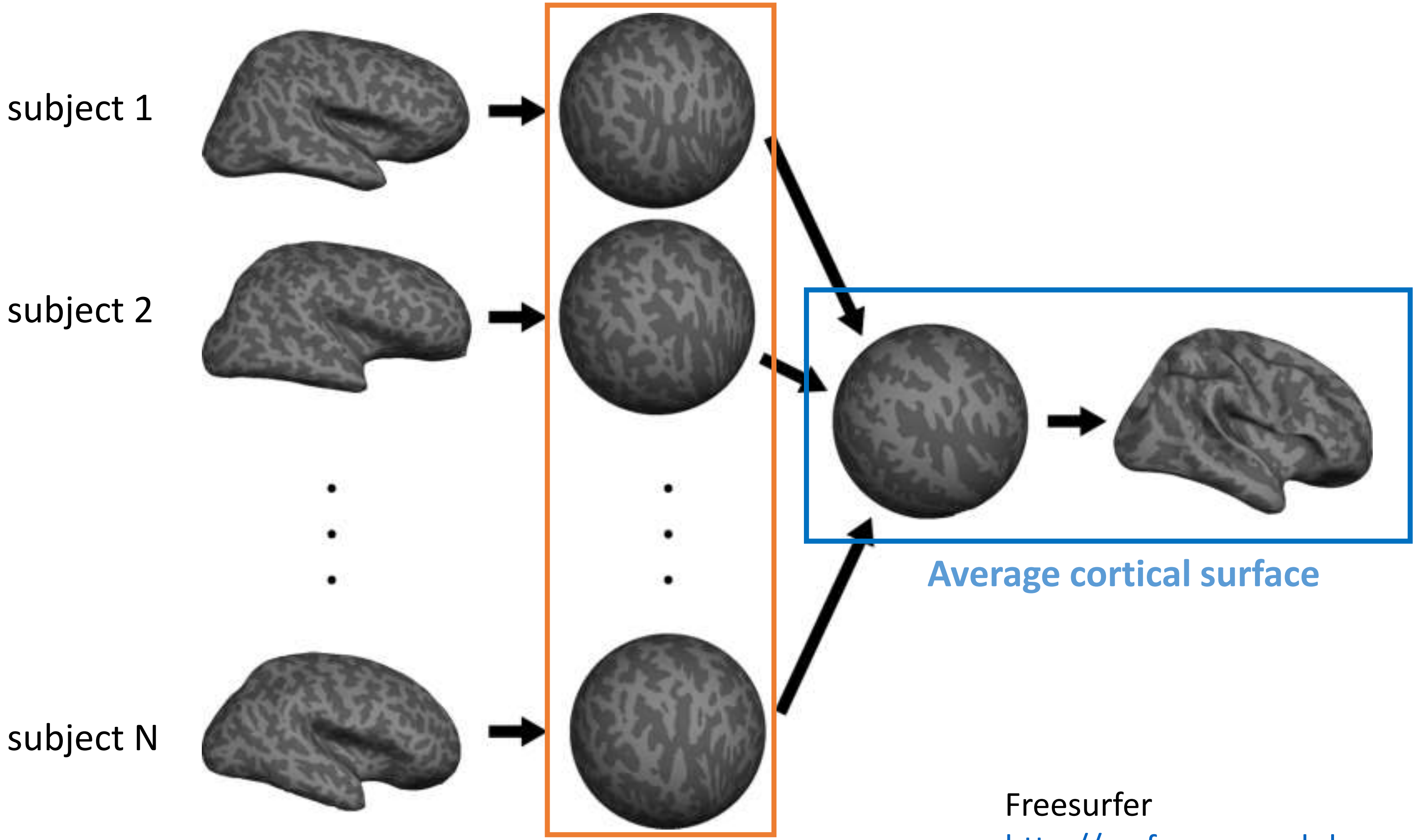


Account for between-subject variance (random-effects analysis)





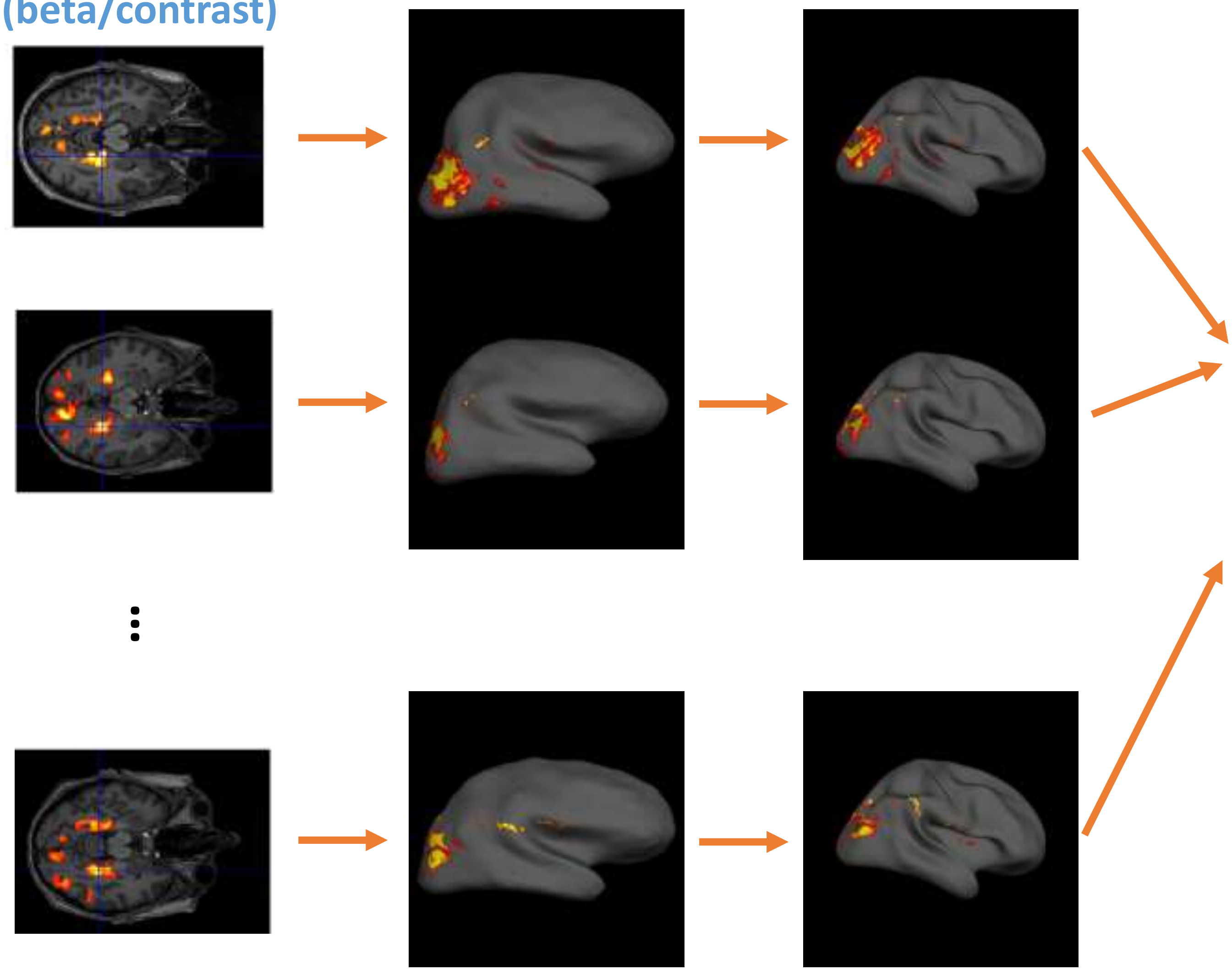
Spatial normalization in surface space

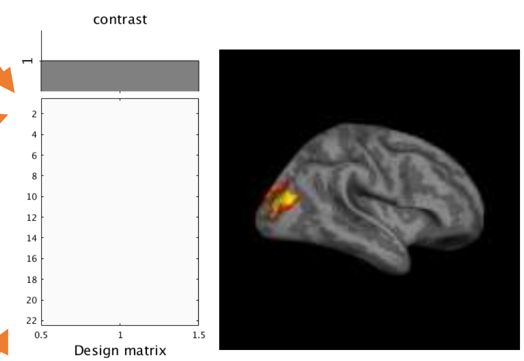


Freesurfer
http://surfer.nmr.mgh.harvard.edu

Surface-based group analysis

Volume data (beta/contrast) **Surface data** **...on average surface**





Freesurfer:
mris_preproc
mri_glmfit

Multiple comparisons correction

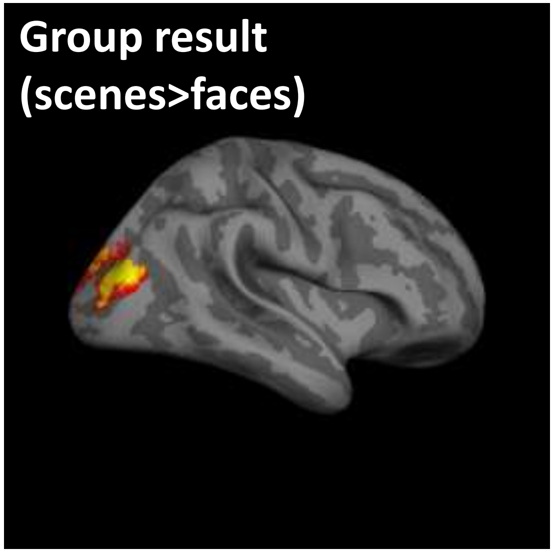
- GLM is applied independently to a huge number of voxels (>100 000) = "**massively univariate**" approach
- At 5% chance level, we might label **5000 voxels** "significant" due to chance
- Different approaches to **correct for the multiple tests**: Bonferroni correction, Gaussian random field theory, false discovery rate approach, permutation methods,…

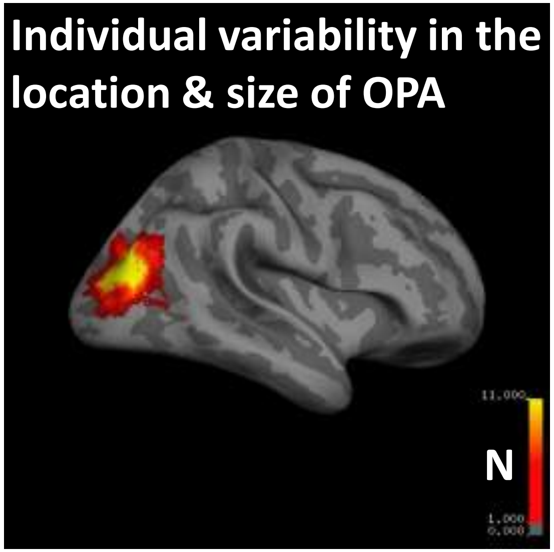
Multiple comparisons correction

- fMRI data has significant **spatial correlations** (neighboring voxels exhibit similar behavior); for example, Bonferroni is typically too conservative for single subject data.
- You can also control for (reduce the number of) the multiple tests by **masking** or **region-of-interest** analysis.
- Importance of **meta-analysis** and **replication studies** to identify consistent results across studies.
- For more details, see, for example, Lindquist et al 2015: Zen and the Art of Multiple Comparisons.

Whole-brain group analysis

- Correspondence problems:
  - Individuals differ both in the global and in the more fine-grained cortical folding patterns
  - Relationship between brain **functions and anatomical structures** across subjects?





Whole-brain group analysis

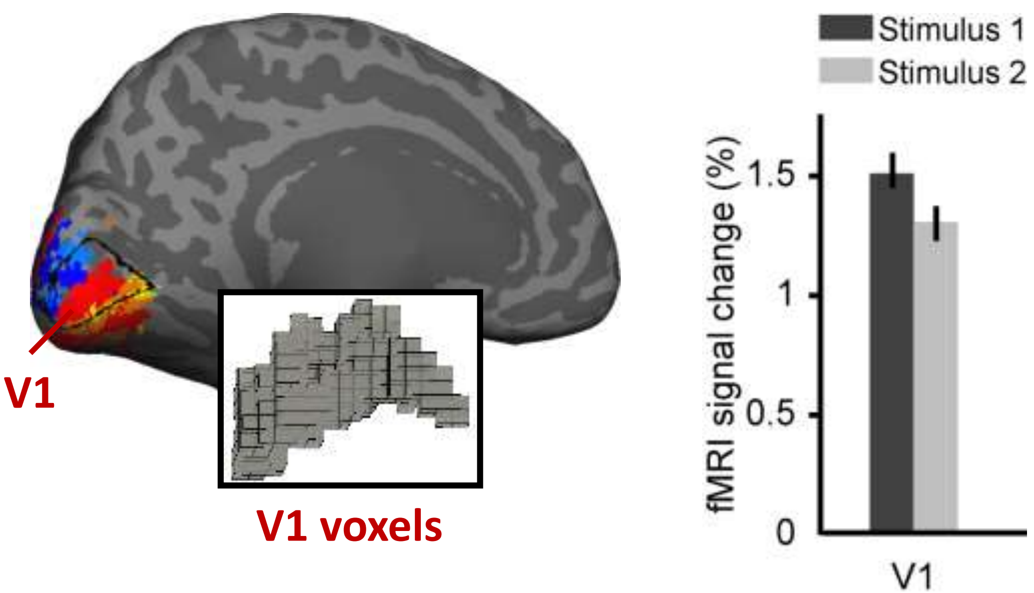
- Correspondence problems:
    - Individuals differ both in the global and in the more fine-grained cortical folding patterns
    - Relationship between brain **functions and anatomical structures** across subjects?

→ **Region-of-interest** analysis

- Region(s)-of-interest can be defined based on anatomical landmarks or functional criteria (separate functional localizer scan)

Region-of-interest analysis

- Region of Interest (ROI) analysis = select a cluster of voxels or brain region ***a priori*** when investigating a region for effects



Why region-of-interest analysis?

- To keep the data in the individuals' space (no need to normalize all data to a common space)
- To **explore** your data
  - Average response time-course within a ROI
- To **limit the number of statistical tests** (>100 000 vs 1 – 500)
  - Average activation-level within a ROI
  - Control for multiple tests within the ROI only
- To **investigate** the function of a region **in detail**

Risks of region-of-interest analysis

- You are looking at the “wrong” region
- The effect is not specific to the region
- Avoid “**double dipping**”, “**circular analysis**”
    - Always use independent data to select the ROIs and ask the research question
    - For more details, see:
        - Kriegeskorte et al 2009: Circular analysis in systems neuroscience – the dangers of double dipping.
        - Kriegeskorte et al 2010: Everything You Never Wanted to Know about Circular Analysis, but Were Afraid to Ask

Summary

- fMRI data are time-series of image volumes
- **Voxel-by-voxel** time-series analysis
- **Preprocessing** to reduce non-task/stimulus related variance in the data
- Spatial **normalization** to a common space necessary when applying whole-brain group analysis

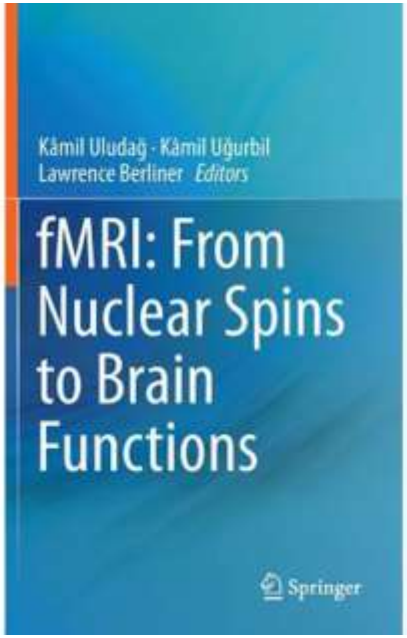
Summary

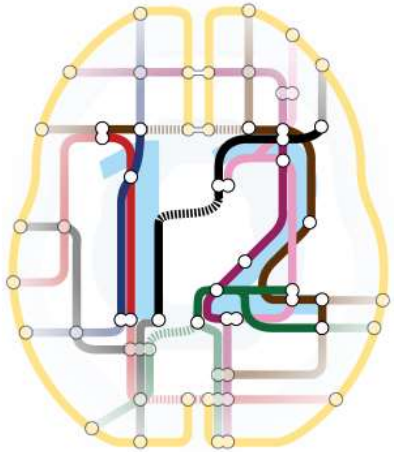
- Statistical analysis…
  1) …aims at **localizing** the brain regions that show an increased (decreased) response in response to the stimulus or task, and
  2) **quantifies** the likelihood that an observer effect can be explained by noise fluctuations.
- **General linear model**: explain the variation in the measured fMRI time-course using a linear combination of predictors.
  1) **Find** optimal **beta** weights (parameter estimates) for each predictor.
  2) **Test** the **significance** of the beta weights (or difference between betas) using, for example, t statistics.
- **Massively univariate** analysis = statistical tests performed independently for each voxel → statistical (parametric) map → **multiple comparison correction**

Summary

- Random-effects analysis asses the variability of observed effects between subjects (cf. fixed-effect analysis).
- Volume-based vs. surface-based group analysis
- **Spatial correspondence problem** in group-analysis: differences in anatomy, validity of structure-function relationship.
- **Region-of-interest analysis** (a priori defined, based on independent data, anatomical and/or functional criteria)

References



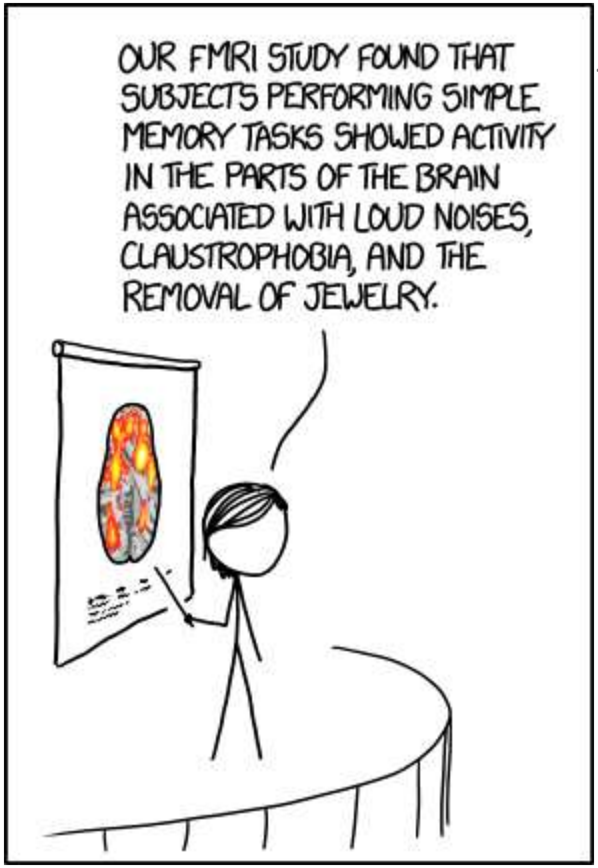


SPM12:
http://www.fil.ion.ucl.ac.uk/spm/software/spm12/

Freesurfer:
https://surfer.nmr.mgh.harvard.edu

Thank you for your attention! Questions?

contact: linda.henriksson@aalto.fi



http://xkcd.com/1453/

PS. Pay attention to careful planning of your experiment!