# Overview:

***generalisability of inferences from multi-subject functional mapping experiments***

- **Most currently employed statistical assessments (parametric & non-parametric) only account for the error variability from scan-to-scan – *fixed-effects* analyses**
  - **only assesses the mean (average) effect for these subjects**
  - **not generalisable to the population(s) from which subjects were drawn – "case study" inference**
  - **unsatisfactory for group comparisons**

    **a group difference may simply reflect small differences between the particular subjects studied, rather than a systematic difference between the populations from which the groups were drawn**

- **To extend inference to the population(s) from which the subjects are drawn, analysis must account for sampling subjects from the population – a *random-effects* analysis**
  - **model the variability in response from subject to subject**
  - **account for both between-subject & error variance**

  - **most designs are balanced and subject-separable**

  ∴ **appropriate *hierarchical random effects models* can be implemented via a multi-level approach**

# Theory: *fixed effects*

**(illustrated for parametric General Linear Model / multiple regression)**

**Consider a 2-condition *n*-subject *f*MRI experiment…**

- **General Linear Model [1] for the timecourses (at a voxel)**

$$Y_{ij} = \gamma_i + \alpha_i f(j) + \ldots + \varepsilon_{ij}$$

**where…**

- $Y_{ij}$ – **measured response on scan *i* of subject *j***
- $f(\bullet)$ – **reference function – such as a box-car**
- $\varepsilon_{ij}$ – **residual errors**
  - **usually** $\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$

**with parameters…**

- $\alpha_i$ – **magnitude of the activation for subject *i***
- $\gamma_i$ – **additive subject effects** $(i = 1,\ldots, n)$

- **Current analyses assume parameters are *fixed effects***
  - **assess significance of mean activation of these subjects –** $\overline{\alpha}_\bullet$
    - **using** $\overline{\hat{\alpha}}_\bullet \sim N(\overline{\alpha}_\bullet, \sigma^2_\varepsilon / nw)$

      $\overline{\hat{\alpha}}_\bullet$ **is the mean measured activation**

      ***w*** **is a weight computed from the design matrix, accounting for the model and temporal auto-correlation in the *f*MRI time series [2]**
  - **only error component of variance (**$\sigma^2_\varepsilon$**) is considered**

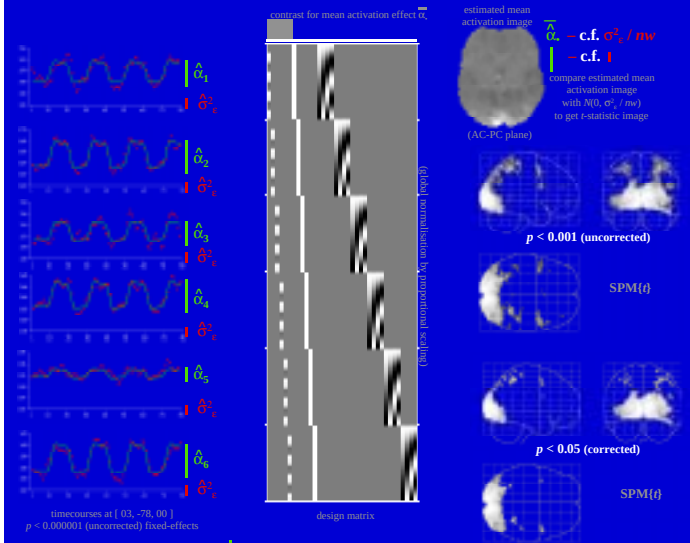
  ∴ **only mean activation for these subjects is assessed**

  - **this is effectively a case study**
  - **results can not be generalised to population**

# *Example:* *fixed effects analysis*

- Six subject visual activation:
  - Epoch *f*MRI : BABABABABABABABA : 10 scans/epoch
    - B → "baseline" – fixation point
    - A → "active" – starfield simulation
  - smoothed box-car reference waveform
  - "high-pass" filter of discrete cosine basis functions [3]

RT=3.2s ⇒ 32s / epoch ⇒ cut off 128s




- mean activation | compared with error variance ↓
  - between-subject variability not taken into account
  - strong evidence of average activation for these subjects

# *Theory:* *random effects*

…to extend inference to population

- subjects are chosen at random from population
- so $\alpha_i$'s are *random effects*
  - $\alpha_i \sim N(\alpha, \sigma^2_\alpha)$
  - $\alpha$ is population mean activation about which we wish to infer
  - $\sigma^2_\alpha$ is *between-subject* variability
- account for error ($\sigma^2_\varepsilon$) and between-subject ($\sigma^2_\alpha$) *components* of variance when assessing observed average activation
- test population mean activation $\alpha$ using:

$\hat{\bar{\alpha}} \sim N(\alpha, \sigma^2_\alpha / n + \sigma^2_\varepsilon / nw)$ (assuming balanced design)

…a *hierarchical model*

- within-subject (1st) level:

$Y_{ij} = \gamma_i + \alpha_i f(j) + \ldots + \varepsilon_{ij}$ – $\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$

- between-subject (2nd) level:

$\alpha_i = \alpha + \varepsilon_i$ – $\varepsilon_i \sim N(0, \sigma^2_\alpha)$

a model on the parameters of the within-subject model

- In general such *multi-level* models are difficult to assess
  - requiring iterative algorithms & special statistics
- However: In functional neuroimaging…
  - designs are balanced (or only slightly imbalanced)
  - models are separable into individual subject models
    - all parameters are independent from subject to subject & can therefore be estimated separately
    - individual subject models all (nearly) identical (in form)

→ two-stage approach

# *Theory: two-stage approach*

1 **Fit level-one** (within-subject) **level models:**

- $Y_{ij} = \gamma_i + \alpha_i f(j) + \ldots + \varepsilon_{ij}$
- **for each subject** *i*

→**giving subject activation estimates**

- $\hat{\alpha}_i \sim N(\alpha_i, \sigma^2_\varepsilon / w)$

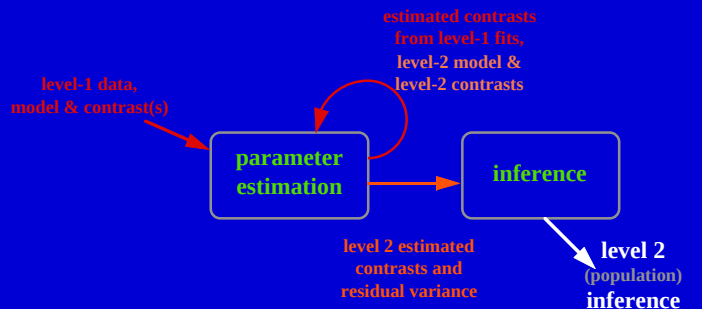
2 **Plug estimates into second level** (between-subject) **model:**

- $\hat{\alpha}_i = \alpha + \varepsilon_i'$
- **a one-sample *t*-test model**
- **residuals** $\varepsilon_i' \sim N(0, \sigma^2)$
- **where** $\sigma^2 = \sigma^2_\alpha + \sigma^2_\varepsilon / w$

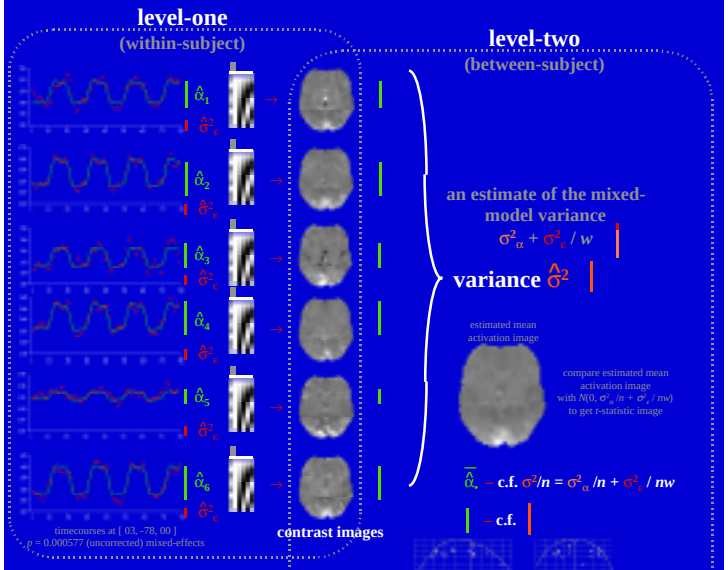
⇒ $\overline{\hat{\alpha}}_\bullet \sim N(\alpha, \sigma^2_\alpha / n + \sigma^2_\varepsilon / nw)$

—**as required for population inference**

—**a one sample *t*-test on the subject activation estimates**




# *Example: random effects analysis*




- **mean activation compared with level-two residual variance**
- **both error and between-subject components of variance accounted for**
- **inference extends to population**
- **much lower significance**

**attempting to infer about the population from a sample of six!**

## *Implementation: contrast images*

- **In general: Consider general single-subject model:**
  - $\underline{Y} = X\,\underline{\beta} + \underline{\varepsilon}$
    - **in matrix form:**
      - data vector $\underline{Y}$
      - design matrix $X$
      - parameter vector $\underline{\beta}$
      - residual errors $\underline{\varepsilon}$
  - **contrast** $\underline{c}^T\underline{\beta}$
  - **estimated by** $\underline{c}^T\underline{\hat{\beta}}$

- **General two-stage approach**
  - **write contrast image** $\underline{c}^T\underline{\hat{\beta}}$ **for each subject**
  - **assess contrast images across subjects in second level model**

- **Computation of contrast images**
  - $\underline{\hat{\beta}} = (X^TX)^{-1}X^T\underline{Y}$
  - $\Rightarrow \underline{c}^T\underline{\hat{\beta}} = \underline{c}^T(X^TX)^{-1}X^T\underline{Y}$
  - **a weighted sum of the data**
  - **no need to fit first-level model**

## *Discussion*

- **General inter subject-level modelling**
  - **via more elaborate second level models**
  - **group comparisons:**
    - **2nd level consists of a two-sample *t*-test**
  - **correlation of activation with a subject score**
    - **2nd level model is regression of subject activations on score**
- **Between-subject variability $\sigma^2_\alpha$ is key**
  - **usually much greater than error variance – $\sigma^2_\alpha \gg \sigma^2_\varepsilon$**
    - **especially for cognitive tasks**
  - **$\sigma^2_\alpha$ dominates random-effects analysis variance term**
    - $\sigma^2_\alpha / n + \sigma^2_\varepsilon / nw$
  - **high chance of *fixed-effects* analysis (ignoring $\sigma^2_\alpha$) producing significant results not representative of the population**
- **Number of subjects ($n$) is crucial**
  - ***n*-1 degrees of freedom**
  - **number of scans per subject not so important**
    - **more subjects & less scans per subject**
- **Unbalanced experiments (non-identical within-subject designs)**
  - **error variance weighting (*w*) not common to all subjects**
  - **but between-subject variability » error variability ($\sigma^2_\alpha \gg \sigma^2_\varepsilon$)**
    - ***w* appears with error variability as $\sigma^2_\varepsilon / nw$ in random-effects analysis variance term ($\sigma^2_\alpha / n + \sigma^2_\varepsilon / nw$)**
  - **$\Rightarrow$ changes in *w* have little effect**
  - **slight imbalances have negligible effect on analysis**
  - **can still use simple two-stage framework**

# Conclusions

***Random effects analyses are required for generalisation of study results to population(s) from which subjects drawn***

- **Particularly relevant for**
  - **group comparisons**
    - **where population inference is sought**
  - **multi-subject *f*MRI**
    - **where error variability is small**
      (relative to between-subject variability)
  - **cognitive paradigms**
    - **where between-subject variability is great**

- **Two-stage approach:**
  - **collapse data for each subject into a single image parameterising the effect of interest** (within-subject modelling)
  - **assess these images across subjects using a simple between-subject model**
  - **extends to multiple levels**
    (when designs are separable and balanced at all levels)
- **Multi-stage approach is**
  - **intuitive**
  - **practicable with existing software**
  - **flexible**
  - **extensible**

***Random effects analyses are standard fare in other disciplines — and are beginning to be demanded by discerning journals***

# References


1 Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ (1995)
**"Statistical Parametric Maps in Functional Imaging: A General Linear Approach"**
*Human Brain Mapping* 2:189–210

2 Frackowiak RSJ, Friston KJ, Frith CD, Dolan RJ, Mazziotta JC (1997)
***Human Brain Function***
Academic Press, London

3 Holmes AP, Josephs O, Büchel C, Friston KJ (1997)
**"Statistical Modelling of Low-Frequency Confounds in fMRI"**
*Neuroimage* 5(4):S480

- Searle SR, Casella G, McCulloch CE (1992)
***Variance Components***
John Wiley & Son, New York

- Everitt BS (1995)
**"The analysis of repeated measures: a practical review with examples"**
The Statistician 1:113–135


- **"Random effects kit for SPM96"**
  - `http://www.fil.ion.ucl.ac.uk/spm/spm96.html#RFX96`
  - **with discussion and references**