Analysis of longitudinal imaging data with OLS & Sandwich Estimator standard errors

Bryan Guillaume

GSK / University of Liège / University of Warwick

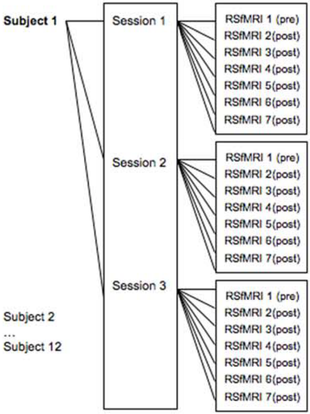
FMRIB Centre - 07 Mar 2012

Supervisors: Thomas Nichols (Warwick University) and Christophe Phillips (Liège University)

Outline

Example of longitudinal studies in neuroimaging
Example 1

Study design:



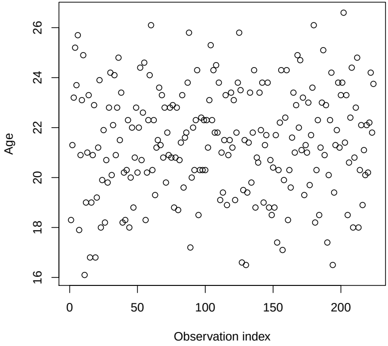
Effect of drugs (morphine and alcohol) versus placebo over time on Resting State Networks in the brain (Khalili-Mahani et al, 2011)

- 12 subjects
- 21 scans/subject!!!
- Balanced design

Example of longitudinal studies in neuroimaging
Example 2

fMRI study of longitudinal changes in a population of adolescents at risk for alcohol abuse
(Heitzeg et al, 2010)

- 86 subjects
- 2 groups
- 1, 2, 3 or 4 scans/subjects (missing data)
- Total of 224 scans
- Very unbalanced design (no common time points for scans)



Why is it challenging to model longitudinal data in neuroimaging ?

Longitudinal modeling is a standard biostatistical problem and standard solutions exist:

- Gold standard: Linear Mixed Effects (LME) model
  - Iterative method $\rightarrow$ generally slow and may fail to converge
    - E.g., 12 subjects, 8 visits, Toeplitz, LME with unstructured intra-visit correlation fails to converge 95 % of the time.
    - E.g., 12 subjects, 8 visits, CS, LME with random int. and random slope fails to converge 2 % of the time.
- LME model with a random intercept per subject
  - May be slow (iterative method) and only valid with Compound Symmetric (CS) intra-visit correlation structure
- Naive-OLS (N-OLS) model which include subject indicator variables as covariates
  - Fast, but only valid with CS intra-visit correlation structure

Why is it challenging to model longitudinal data in neuroimaging ?

Longitudinal modeling is a standard biostatistical problem and standard solutions exist:

- Gold standard: Linear Mixed Effects (LME) model
  - Iterative method → generally slow and may fail to converge
    - E.g., 12 subjects, 8 visits, Toeplitz, LME with unstructured intra-visit correlation fails to converge 95 % of the time.
    - E.g., 12 subjects, 8 visits, CS, LME with random int. and random slope fails to converge 2 % of the time.
- LME model with a random intercept per subject
  - May be slow (iterative method) and only valid with Compound Symmetric (CS) intra-visit correlation structure
- Naive-OLS (N-OLS) model which include subject indicator variables as covariates
  - Fast, but only valid with CS intra-visit correlation structure

Why is it challenging to model longitudinal data in neuroimaging ?

Longitudinal modeling is a standard biostatistical problem and standard solutions exist:

- Gold standard: Linear Mixed Effects (LME) model
  - Iterative method $\rightarrow$ generally slow and may fail to converge
    - E.g., 12 subjects, 8 visits, Toeplitz, LME with unstructured intra-visit correlation fails to converge 95 % of the time.
    - E.g., 12 subjects, 8 visits, CS, LME with random int. and random slope fails to converge 2 % of the time.
- LME model with a random intercept per subject
  - May be slow (iterative method) and only valid with Compound Symmetric (CS) intra-visit correlation structure
- Naive-OLS (N-OLS) model which include subject indicator variables as covariates
  - Fast, but only valid with CS intra-visit correlation structure

Why is it challenging to model longitudinal data in neuroimaging ?

Longitudinal modeling is a standard biostatistical problem and standard solutions exist:

- Gold standard: Linear Mixed Effects (LME) model
  - Iterative method $\rightarrow$ generally slow and may fail to converge
    - E.g., 12 subjects, 8 visits, Toeplitz, LME with unstructured intra-visit correlation fails to converge 95 % of the time.
    - E.g., 12 subjects, 8 visits, CS, LME with random int. and random slope fails to converge 2 % of the time.
- LME model with a random intercept per subject
  - May be slow (iterative method) and only valid with Compound Symmetric (CS) intra-visit correlation structure
- Naive-OLS (N-OLS) model which include subject indicator variables as covariates
  - Fast, but only valid with CS intra-visit correlation structure

Outline

The Sandwich Estimator (SwE) method

- Use of a simple OLS model (without subject indicator variables)
- The fixed effects parameters β are estimated by

$$\hat{\beta}_{OLS} = \left(\sum_{i=1}^{M} X_i' X_i\right)^{-1} \sum_{i=1}^{M} X_i' y_i$$

- The fixed effects parameters covariance $\mathrm{var}(\hat{\beta}_{OLS})$ are estimated by

$$\mathrm{SwE} = \underbrace{\left(\sum_{i=1}^{M} X_i' X_i\right)^{-1}}_{\text{Bread}} \underbrace{\left(\sum_{i=1}^{M} X_i' \hat{V}_i X_i\right)}_{\text{Meat}} \underbrace{\left(\sum_{i=1}^{M} X_i' X_i\right)^{-1}}_{\text{Bread}}$$

Property of the Sandwich Estimator (SwE)

$$\text{SwE} = \left(\sum_{i=1}^{M} X_i' X_i\right)^{-1} \left(\sum_{i=1}^{M} X_i' \hat{V}_i X_i\right) \left(\sum_{i=1}^{M} X_i' X_i\right)^{-1}$$

If V_i are consistently estimated, the SwE tends **asymptotically** (Large samples assumption) towards the true variance $\text{var}(\hat{\beta}_{OLS})$. (Eicker, 1963; Eicker, 1967; Huber, 1967; White, 1980)

The Heterogeneous HC0 SwE

In practice, V_i is generally estimated from the residuals $r_i = y_i - X_i\hat{\beta}$ by

$$\hat{V}_i = r_i r_i^{'}$$

and the SwE becomes

$$\text{Het. HC0 SwE} = \left(\sum_{i=1}^{M} X_i' X_i\right)^{-1} \left(\sum_{i=1}^{M} X_i' r_i r_i^{'} X_i\right) \left(\sum_{i=1}^{M} X_i' X_i\right)^{-1}$$

Simulations: setup

- Monte Carlo Gaussian null simulation (10,000 realizations)
- For each realization,
  1. Generation of longitudinal Gaussian null data (no effect) with a CS or a Toeplitz intra-visit correlation structure:

Compound Symmetric

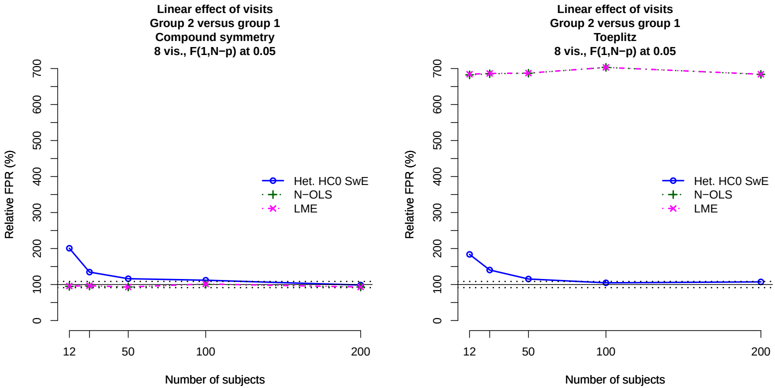
$$\begin{pmatrix} 1 & 0.8 & 0.8 & 0.8 & 0.8 \\ 0.8 & 1 & 0.8 & 0.8 & 0.8 \\ 0.8 & 0.8 & 1 & 0.8 & 0.8 \\ 0.8 & 0.8 & 0.8 & 1 & 0.8 \\ 0.8 & 0.8 & 0.8 & 0.8 & 1 \end{pmatrix}$$

Toeplitz

$$\begin{pmatrix} 1 & 0.8 & 0.6 & 0.4 & 0.2 \\ 0.8 & 1 & 0.8 & 0.6 & 0.4 \\ 0.6 & 0.8 & 1 & 0.8 & 0.6 \\ 0.4 & 0.6 & 0.8 & 1 & 0.8 \\ 0.2 & 0.4 & 0.6 & 0.8 & 1 \end{pmatrix}$$

  2. Statistical test (F-test at α) on the parameters of interest using each different methods (N-OLS, LME and SWE) and recording if the method detects a (False Positive) effect
- For each method, rel. FPR$= \frac{\text{Number of False Positive}}{10{,}000\alpha}$

Simulations: LME vs N-OLS vs Het. HC0 SwE



Outline

Bias adjustments: the Het. HC2 SwE

- In an OLS model, we have

$$(I - H)\text{var}(y)(I - H) = \text{var}(r)$$

where $H = X(X'X)^{-1}X'$

- Under independent homoscedastic errors,

$$\begin{aligned}(I - H)\sigma^2 &= \text{var}(r)\\ (1 - h_{ik})\sigma^2 &= \text{var}(r_{ik})\\ \sigma^2 &= \text{var}\left(\frac{r_{ik}}{\sqrt{1 - h_{ik}}}\right)\end{aligned}$$

- This suggests to estimate V_i by

$$\hat{V}_i = r_i^* r_i^{*'} \text{ where } r_{ik}^* = \frac{r_{ik}}{\sqrt{1 - h_{ik}}}$$

Bias adjustments: the Het. HC2 SwE

Using in the SwE

$$\hat{V}_i = r_i^* r_i^{*'} \text{ where } r_{ik}^* = \frac{r_{ik}}{\sqrt{1 - h_{ik}}}$$

We obtain

$$\text{Het. HC2 SwE} = \left(\sum_{i=1}^{M} X_i' X_i\right)^{-1} \left(\sum_{i=1}^{M} X_i' r_i^* r_i^{*'} X_i\right) \left(\sum_{i=1}^{M} X_i' X_i\right)^{-1}$$

Homogeneous SwE

In the standard SwE, each V_i is normally estimated from only the residuals of subject i. It is reasonable to assume a common covariance matrix V_0 for all the subjects and then, we have

$$\hat{V}_{0kk'} = \frac{1}{N_{kk'}} \sum_{i=1}^{N_{kk'}} r_{ik} r_{ik'}$$

$\hat{V}_{0kk'}$: element of $\hat{V}_0$ corresponding to the visits k and k'
$N_{kk'}$: number of subjects with both visits k and k'
r_{ik}: residual corresponding to subject i and visit k
$r_{ik'}$: residual corresponding to subject i and visit k'

$$\hat{V}_i = f(\hat{V}_0)$$

Null distribution of the test statistics with the SwE

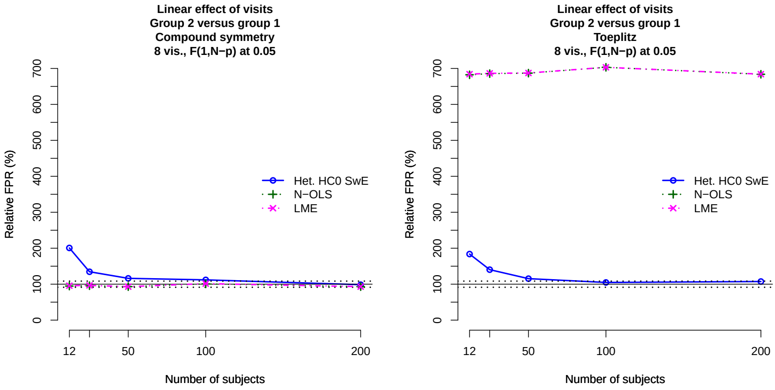
- $H_0 : L\hat{\beta} = 0, H_1 : L\hat{\beta} \neq 0$
  L: contrast matrix of rank q
- Using multivariate statistics theory and assuming a balanced design, we can derive the test statistic

$$\frac{M - p_B - q + 1}{(M - p_B)q}(L\hat{\beta})'(L\mathrm{SwE}L')^{-1}(L\hat{\beta}) \sim F(q, M - p_B - q + 1)$$

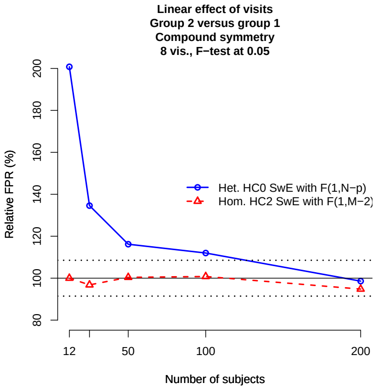
- q=1, the test becomes

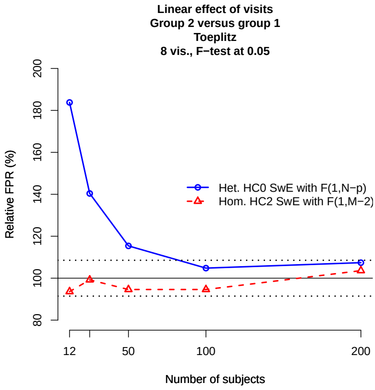
$$(L\hat{\beta})'(L\mathrm{SwE}L')^{-1}(L\hat{\beta}) \sim F(1, M - p_B) \neq F(1, N - p)$$

Simulations: LME vs N-OLS vs unadjusted SwE



Simulations: unadjusted SwE vs adjusted SwE

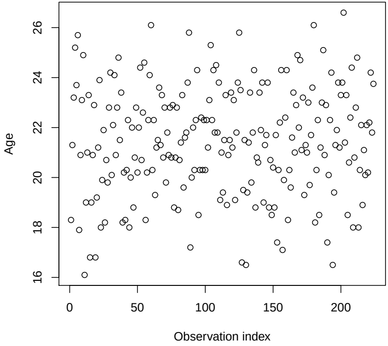




Simulation with real design
Example 2

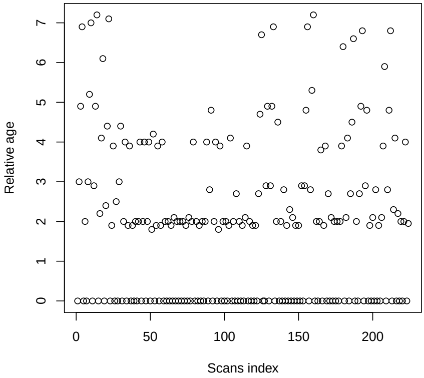
fMRI study of longitudinal changes in a population of adolescents at risk for alcohol abuse

- 86 subjects
- 2 groups
- 1, 2, 3 or 4 scans/subjects (missing data)
- Total of 224 scans
- Very unbalanced design (no common time points for scans)

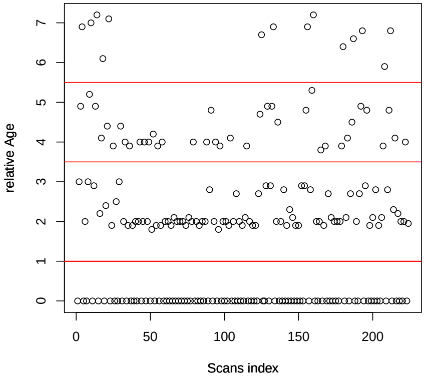


Simulation with real design

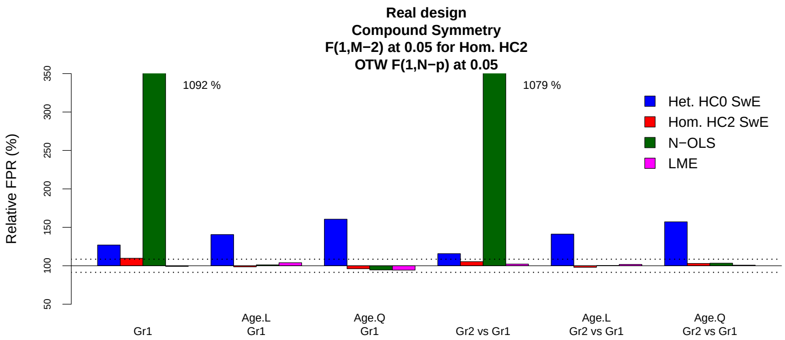
Example 2



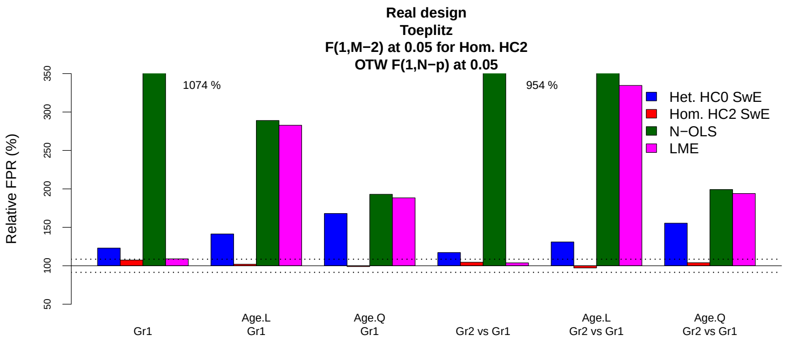
Simulation with real design
Example 2



Real design


Real design
Compound Symmetry
F(1,M−2) at 0.05 for Hom. HC2
OTW F(1,N−p) at 0.05
Relative FPR (%)
50
100
150
200
250
300
350
1092 %
1079 %
Het. HC0 SwE
Hom. HC2 SwE
N−OLS
LME
Gr1
Age.L
Gr1
Age.Q
Gr1
Gr2 vs Gr1
Age.L
Gr2 vs Gr1
Age.Q
Gr2 vs Gr1

Real design


Real design
Toeplitz
F(1,M−2) at 0.05 for Hom. HC2
OTW F(1,N−p) at 0.05
Relative FPR (%)
50
100
150
200
250
300
350
1074 %
954 %
Het. HC0 SwE
Hom. HC2 SwE
N−OLS
LME
Gr1
Age.L
Gr1
Age.Q
Gr1
Gr2 vs Gr1
Age.L
Gr2 vs Gr1
Age.Q
Gr2 vs Gr1

Outline

Remarks about the SwE method

- Power of the SwE method generally lower than the power of the LME method
  - Power loss not significant with a high number of subject (e.g., 86 subjects)
  - Power loss may be significant with a low number of subject and a low significance level α
    - Solution: spatial regularization of the SwE
- Test statistic with an unbalanced design and a low number of subject
  - Estimation of the effective degrees of freedom of the test needed

Summary

- Longitudinal standard methods not really appropriate to neuroimaging data:
  - Convergence issues with LME
  - N-OLS & LME with random intercepts: issues when CS does not hold
- The SwE method
  - Accurate in a large range of settings
  - Easy to specify
  - No iteration needed
    - Quite fast
    - No convergence issues
  - Can accommodate pure between covariates
  - But, careful in small samples:
    - Adjustment essential
    - If low significance level, spatial regul. needed for power
    - If unbalanced design, effective dof estimation needed

Thanks for your attention!