Modèles non linéaires à effets mixtes, pénalités structurées et algorithme SAEM

E Ollier[1],A Samson[2], X Delavenne[1], V Viallon[3]

(1) Groupe de Recherche sur la Thrombose, EA3065, Université Jean Monnet, Saint-Etienne, France
(2) Université Grenoble-Alpes, Laboratoire Jean Kuntzmann, UMR CNRS 5224, Grenoble, France
(3) Université de Lyon, F-69622, Lyon, France

Group Comparison with Non Linear Mixed Effect Model (NLMEM)

Non Linear mixed effects model (NLMEM) for groups of subjects ($g = 1, ..., G$) :

$$y_{i,j}^g = f(t_{i,j}^g, \phi_i^g) + h(t_{i,j}^g, \phi_i^g)\epsilon_{i,j}^g \text{ with } \epsilon_{i,j}^g \sim \mathcal{N}(0,1)$$

$$\phi_i^g = A_i^g \begin{pmatrix} \mu^g \\ \beta^g \end{pmatrix} + \omega_i^g \text{ with } \omega_i^g \sim \mathcal{N}(0, \Omega^g)$$

Ω^g diagonal

$h = af + b$

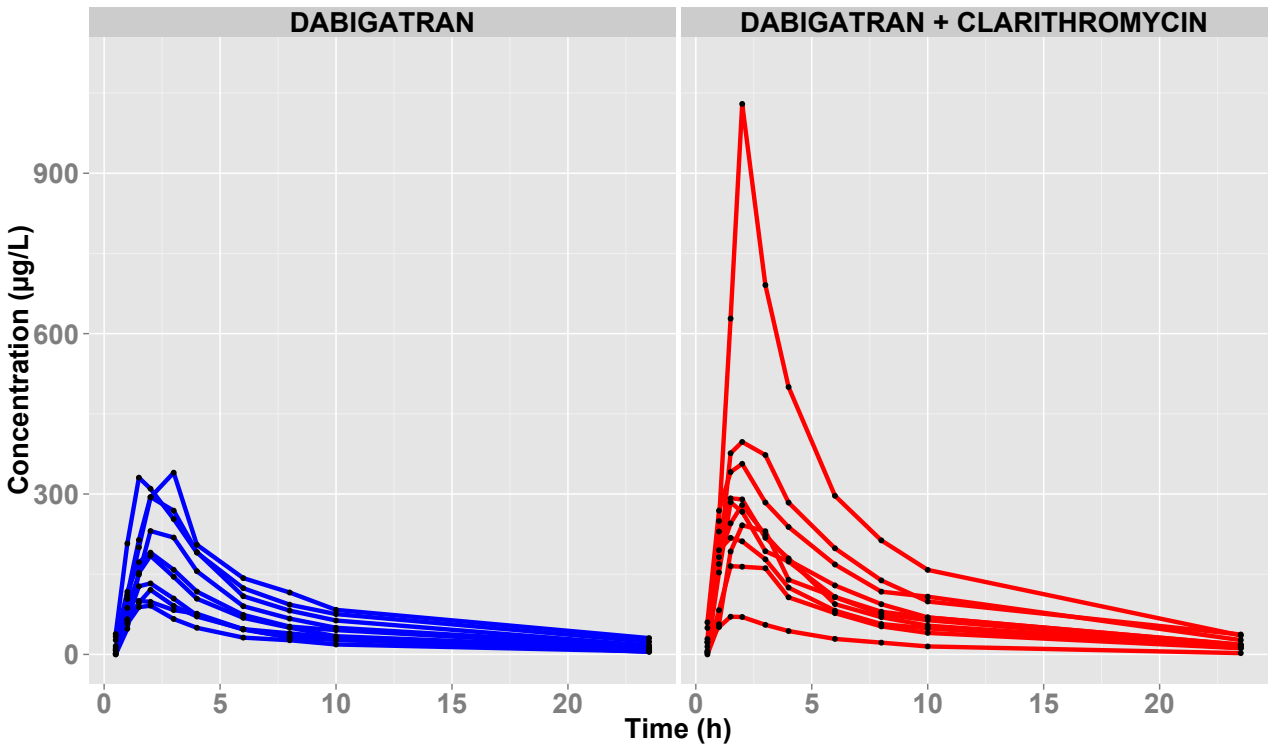
Hypothesis

$\boldsymbol{\mu^{g1} - \mu^{g2}}$, $\boldsymbol{\Omega^{g1} - \Omega^{g2}}$, $\boldsymbol{\beta^{g1} - \beta^{g2}}$ **and** $\boldsymbol{\beta}$ **are sparse**

Goal

Detect significant difference between group parameters

Dabigatran PK with or without Clarithromycin[1]



Is there a drug-drug interaction ?

Parameters influenced by clarithromycin ?

[1] Delavenne et al., 2013

Model Selection methods

- **Stepwise methods**:
  - Forward, Backward,...
  - Based on Likelihood Ratio Tests (LRT) or BIC

- **Greedy approach**:
  - All the possible model are evaluated

    $\hookrightarrow$Difficult with large number of parameter

- **Penalized likelihood**:
  - Preselect relevant models
  - example: the LASSO [2] (covariates selection)

[2]Bertrand J, Balding DJ, 2013

Joint Modelling with the Fused LASSO Penalty

When $G = 2$:

$$\underset{\theta^1,\theta^2}{\text{ArgMin}} \sum_{g=1}^{2} -2LL(y^g, \theta^g) + \boldsymbol{\lambda}\|\theta^1 - \theta^2\|_1$$

Estimate a vector $\theta^1 - \theta^2$ that is sparse

λ tunes the level of sparsity :

- $\lambda = 0 \Rightarrow$ between group differences are estimated for each parameters
- high levels of λ values $\Rightarrow \theta^1 = \theta^2$

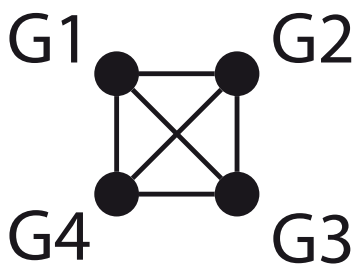
Joint Modelling with the Fused LASSO Penalty

When $G > 2$:

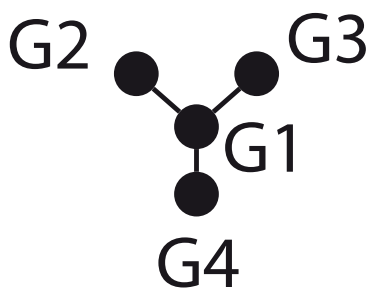
$$\underset{\theta^1,\ldots,\theta^G}{\text{ArgMin}} \sum_{g=1}^{G} -2LL(y^g, \theta^g) + \lambda \sum_{(g_1,g_2)\in\mathcal{E}} \|\theta^{g_1} - \theta^{g_2}\|_1$$

where $\mathcal{E}$ is the edges set of a graph.

Clique Graph



Star Graph



Penalty used in this work

- **Fixed effects**

$$P_F(\mu^1, ..., \mu^G) = \sum_{(g_1, g_2) \in \mathcal{E}} \|\mu^{g_1} - \mu^{g_2}\|_1$$

- **Random effects variances**

$$P_V(\Omega^{1^{-1}}, ...\Omega^{G^{-1}}) = \sum_{(g_1, g_2) \in \mathcal{E}} \|\Omega^{g_1^{-1}} - \Omega^{g_2^{-1}}\|_1$$

- **Covariates effects**

$$P_C(\beta^1, ..., \beta^G) = \sum_{(g_1, g_2) \in \mathcal{E}} \|\beta^{g_1} - \beta^{g_2}\|_1 + \sum_{g=1}^{G} \|\beta^g\|_1$$

Some notations

- $\theta = (\theta^1, ..., \theta^G)$
- $\theta^g = (\mu^g, \beta^g, \Omega^g, a, b)$
- $\mu = (\mu^1, ..., \mu^G)$
- $\beta = (\beta^1, ..., \beta^G)$

SAEM algorithm for joint modeling (iteration k)

- **E step**:
  1. **Simulation**:
     - simulation of ϕ^g under $p(\phi|y^g;\theta_k^g)$ by MCMC
  2. **Stochastic Approximation**: Stochastic approximation of $\mathbb{E}[logp(y^g,\phi^g,\theta^g)|y^g,\theta_k^g]$:
     - $Q_{k+1}^g(\theta) = \gamma_k log\ p(y^g,\phi^g,\theta_k^g) + (1-\gamma_k)Q_k(\theta)$[1]

     with $(\gamma_k\ )$ sequence of decreasing step sizes

- **M step**: actualization of θ:

$$(\theta_{k+1}^1,...,\theta_{k+1}^G) = \underset{\theta^1,...,\theta^G}{\text{ArgMax}} \sum_{g=1}^{G} Q_{k+1}^g(\theta^g)$$

Introducing a penalty within SAEM

Classical SAEM algorithm (M-step):

$$(\theta_{k+1}^1, ..., \theta_{k+1}^G) = \underset{\theta^1,...,\theta^G}{\text{ArgMax}} \sum_{g=1}^{G} Q_{k+1}^g(\theta^g)$$

Penalized SAEM algorithm (M-step):

$$(\theta_{k+1}^1, ..., \theta_{k+1}^G) = \underset{\theta^1,...,\theta^G}{\text{ArgMax}} \sum_{g=1}^{G} Q_{k+1}^g(\theta^g) - P(\theta^1, ..., \theta^G)$$

Penalized version of SAEM algorithm (iteration k)

- **E step**:

UNCHANGED

- **M step**: actualization of θ_k^g:

$$(\theta_{k+1}^1, ..., \theta_{k+1}^G) = \underset{\theta^1,...,\theta^G}{\text{ArgMax}} \begin{cases} \sum_{g=1}^G Q_{k+1}^g(\theta^g) \\ -\lambda_F P_F(\mu^1, ..., \mu^G) \\ -\lambda_C P_C(\beta^1, ..., \beta^G) \\ -\lambda_V P_V(\Omega^{1^{-1}}, ...\Omega^{G^{-1}}) \end{cases}$$

M step in details

1. Fixed effects update[3]:

$$(\mu_{k+1}, \beta_{k+1}) = \underset{\mu,\beta}{\text{ArgMax}} \begin{cases} \sum_{g=1}^{G} Q_k^g(\mu^g, \Omega_k^g, a_k, b_k) \\ -\lambda_F P_F(\mu^1, ..., \mu^G) \\ -\lambda_C P_C(\beta^1, ..., \beta^G) \end{cases}$$

2. Random effects variance update[4]:

$$(\Omega_{k+1}^1, ..., \Omega_{k+1}^G) = \underset{\Omega}{\text{ArgMax}} \begin{cases} \sum_{g=1}^{G} Q_k^g(\mu_{k+1}^g, \Omega^g, a_k, b_k) \\ -\lambda_V P_V(\Omega^{1^{-1}}, ...\Omega^{G^{-1}}) \end{cases}$$

3. Error model's parameters update: same as in SAEM because they are not penalized.

[3] Boyd et al., 2011
[4] Danaher et al., 2011

BIC Selection of Optimal λ_F and λ_V Values

1. Preselect relevant models by solving the fused LASSO problem on a grid of λ_F and λ_V values:

$$(\ (\lambda_F^m, \lambda_V^m) \in \mathbb{R}^2, \ m \in \{1, ..., M\})$$

2. The optimal model (among the preselected) is the one that minimize the BIC [5]:

$$BIC(\lambda_F, \lambda_V) = \sum_{g=1}^{G} -2LL(y^g, \tilde{\theta}^g_{\lambda_F, \lambda_V}) + log(N) \times df_{\lambda_F, \lambda_V}$$

- N: total number of patients
- $\tilde{\theta}^g_{\lambda_F, \lambda_V}$: unpenalized reestimation of the selected model
- df_λ: number of distinct coefficients (degree of freedom)

[5] Delattre et al., 2014

Exemple: a Simulated Data Set with 3 Groups

- 100 patients per group
- Model : first order absorption, one compartment

$$C(t) = \frac{D}{V}\frac{k_a}{k_a - Cl}(e^{-\frac{Cl}{V}t} - e^{-k_a t})$$

- Fixed effects parameters were set to :

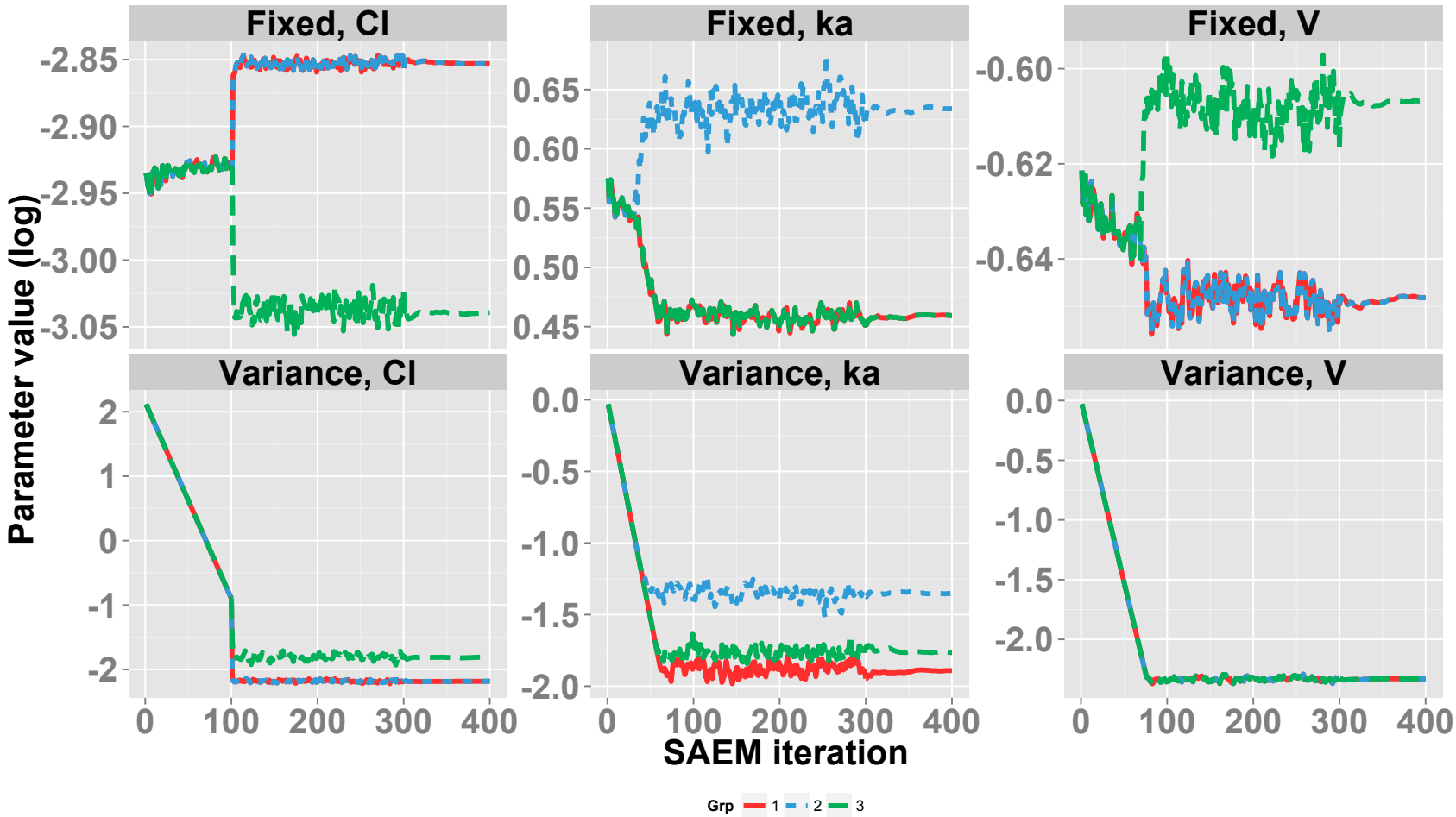
$$\mu^1_{Cl} = 0.0597 \quad \mu^1_V = 0.4819 \quad \mu^1_{k_a} = 1.47$$
$$\mu^2_{Cl} = 0.0597 \quad \mu^2_V = 0.4819 \quad \mu^2_{k_a} = 2.18$$
$$\mu^3_{Cl} = 0.0420 \quad \mu^3_V = 0.5784 \quad \mu^3_{k_a} = 1.47$$

- Random effects variances were set to :

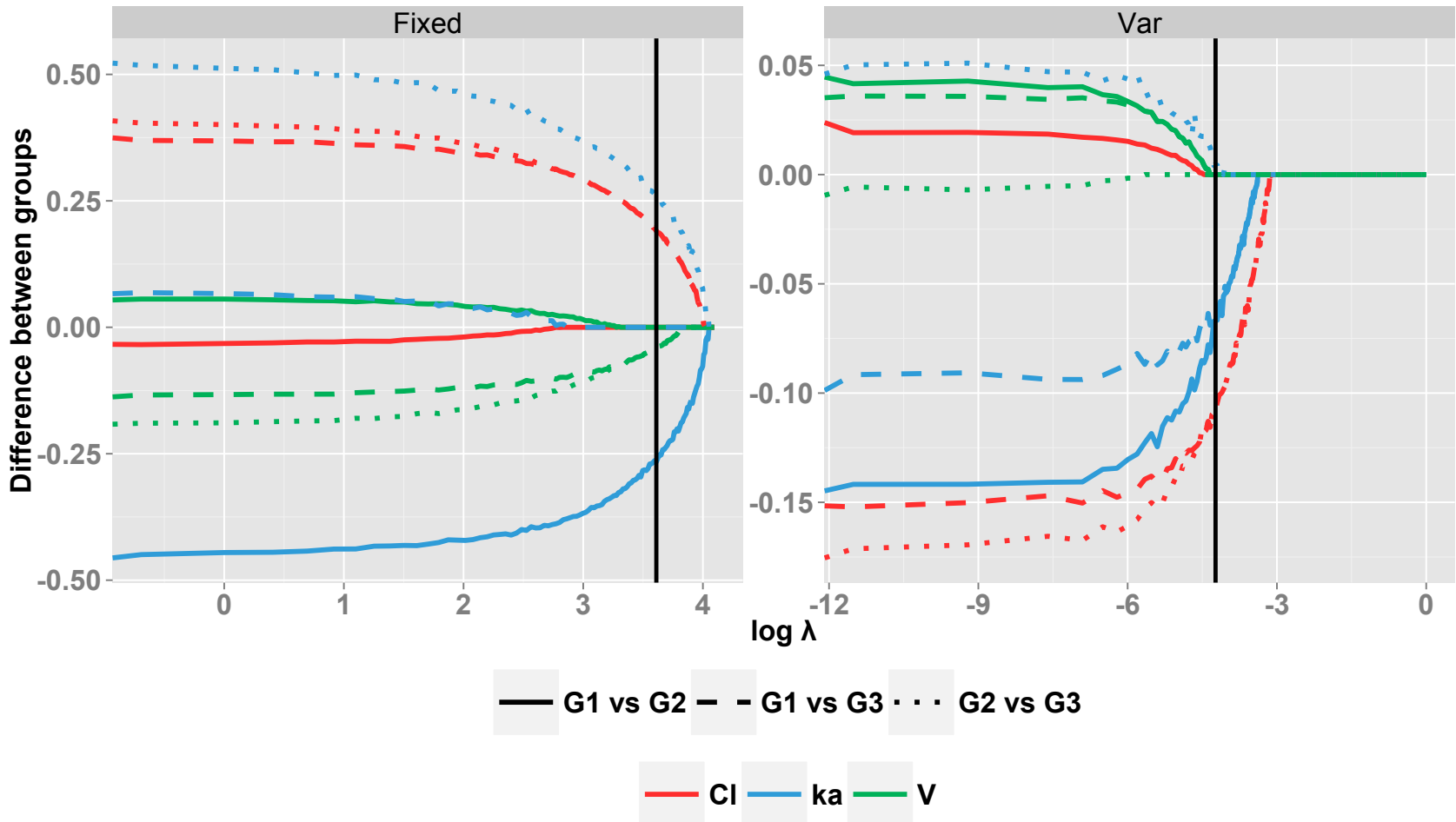
$$\Omega^1_{Cl} = 0.1 \quad \Omega^1_V = 0.1 \quad \Omega^1_{k_a} = 0.1$$
$$\Omega^2_{Cl} = 0.1 \quad \Omega^2_V = 0.1 \quad \Omega^2_{k_a} = 0.25$$
$$\Omega^3_{Cl} = 0.2 \quad \Omega^3_V = 0.1 \quad \Omega^3_{k_a} = 0.2$$

- Penalty structure: clique graph

SAEM Estimates



Regularization path



Model Selection Performance: Simulation Set up

- Model : first order absorption, one compartment
- 2 groups with 25, 50 or 100 patients per group
- Fixed effects parameters were set to :

$$\mu_{Cl}^1 = 0.0597 \quad \mu_V^1 = 0.4819 \quad \mu_{k_a}^1 = 1.47$$
$$\mu_{Cl}^2 = 0.0420 \quad \mu_V^2 = 0.5784 \quad \mu_{k_a}^2 = 1.47$$

- Random effects variances were set to :

$$\Omega_{Cl}^1 = 0.1 \quad \Omega_V^1 = 0.1 \quad \Omega_{k_a}^1 = 0.1$$
$$\Omega_{Cl}^2 = 0.21 \quad \Omega_V^1 = 0.1 \quad \Omega_{k_a}^2 = 0.21$$

- 50 simulations for each group size

Model Selection Performance

	Fixed effects		
	25	50	100
Forward + BIC	30%	76%	68%
Fused + BIC	44%	74%	76%

	Variances		
	25	50	100
Forward + BIC	10%	30%	52%
Fused + BIC	38%	58%	78%

	Both		
	25	50	100
Forward + BIC	6%	24%	46%
Fused + BIC	14%	40%	60%

Table: Proportion of correctly selected model on 50 simulations for the fixed effects model, the variances model and the whole model

Dabigatran PK with or without Clarithromycin

Inverse Gaussian absorption model:

$$IG(t) = Dose \times F \times \sqrt{\frac{MAT}{2\pi CV^2 t^3}} \times e^{\frac{-(t-MAT)^2}{2CV^2 MAT t}}$$

$$\frac{dD_c}{dt} = IG(t) - D_c(\frac{Q}{V_c} + \frac{Cl}{V_c}) + \frac{Q}{V_p} D_p$$

$$\frac{dD_p}{dt} = \frac{Q}{V_c} D_c - \frac{Q}{V_p} D_p$$

- Individual parameters: Log-Normal distribution
- Random effects on all the parameters

Dabigatran PK with or without Clarithromycin

First step:

- No covariates
- Effect of clarithromycin tested on 8 parameters:
  - Absorption:

$$\mu_F,\ \mu_{MAT},\ \mu_{CV} \text{ and } \Omega_F,\ \Omega_{MAT},\ \Omega_{CV}$$

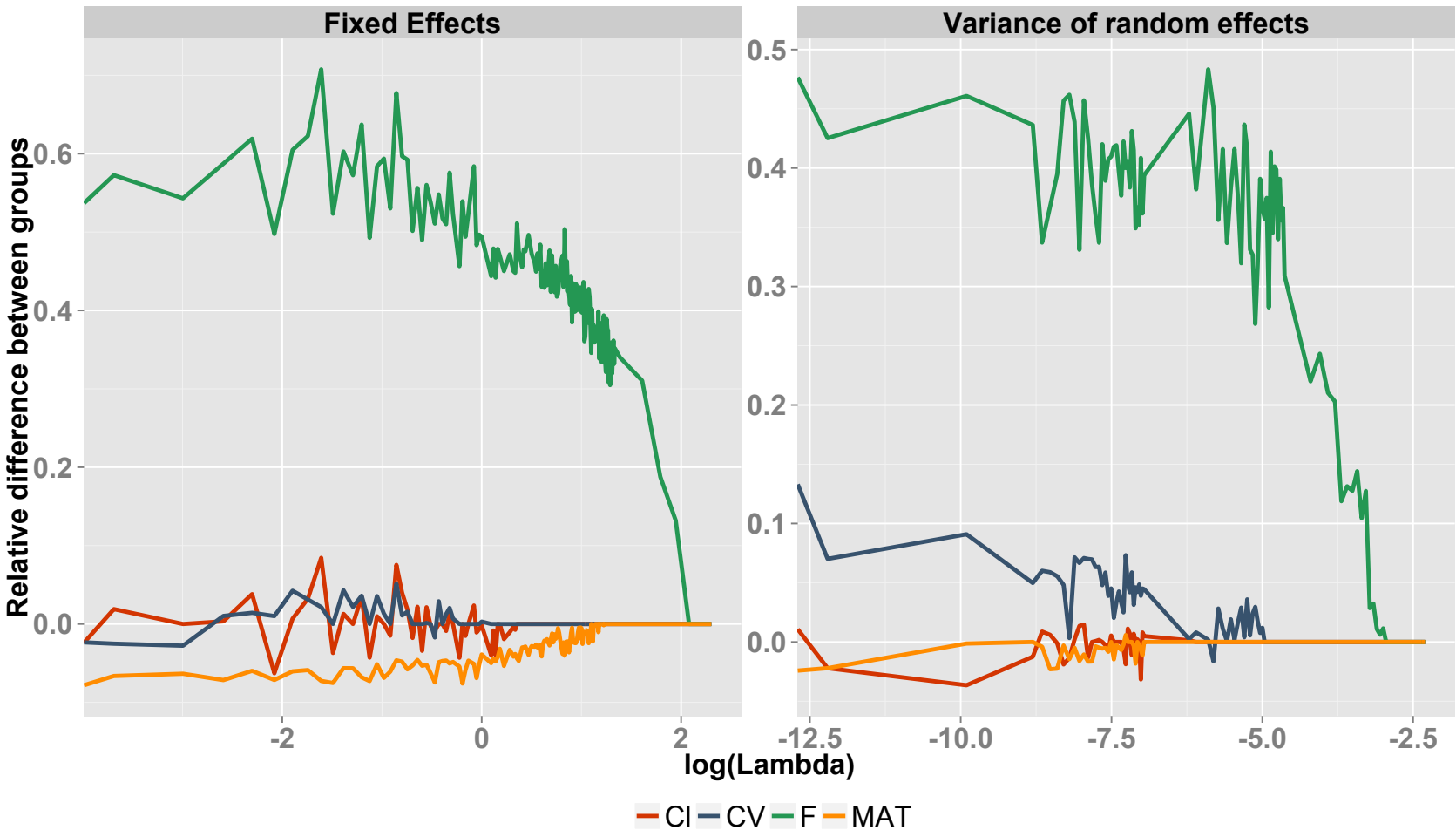
  - Elimination:

$$\mu_{Cl} \text{ and } \Omega_{Cl}$$

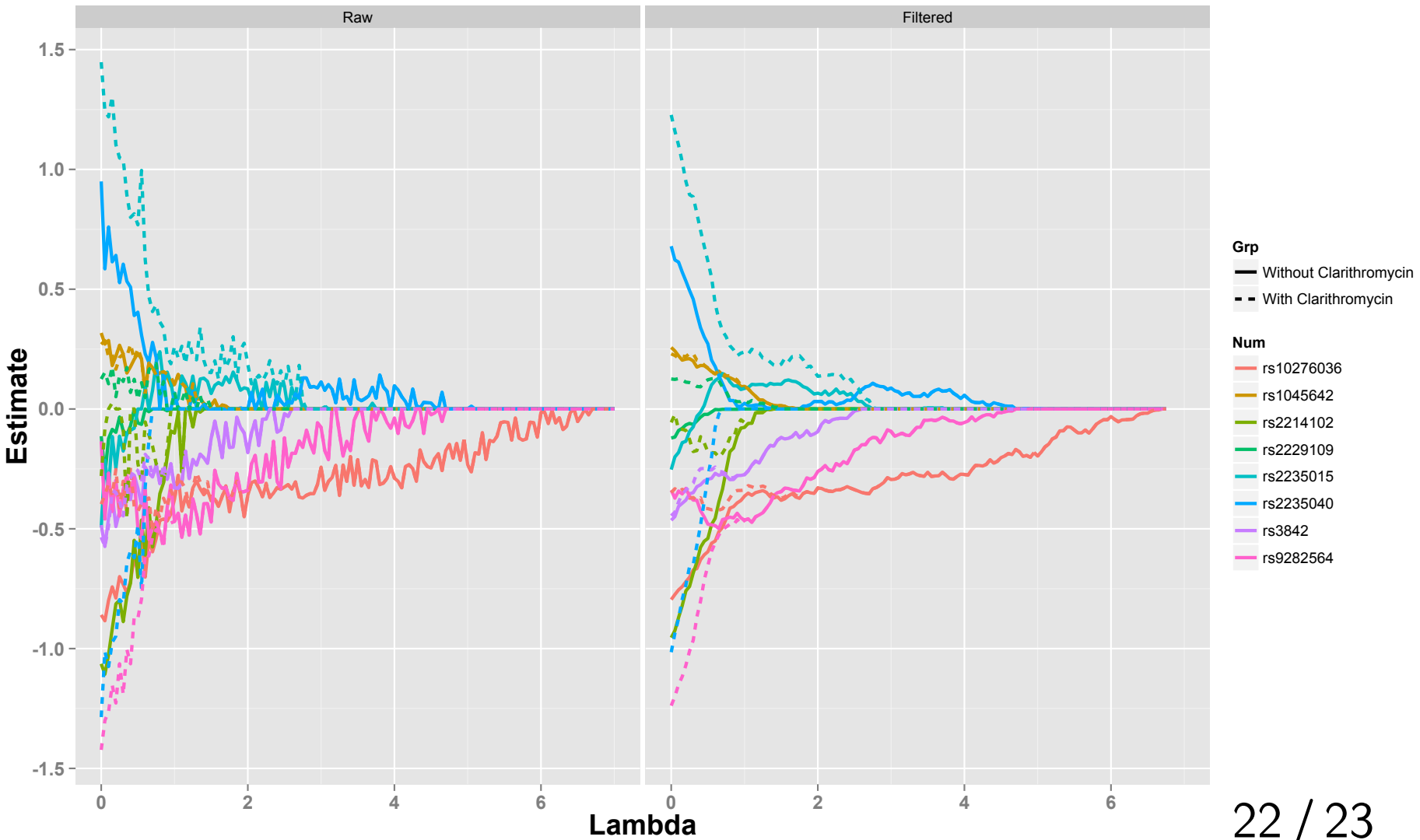
Second step:

- 8 SNP of the P-gp included as covariates on F
- The model selected in first step was used for fixed and variance parameters

Dabigatran PK with or without Clarithromycin: Results



Dabigatran PK with or without Clarithromycin: Covariates Analysis



Conclusion

- Extension of the SAEM algorithm to take into account a fused LASSO penalty on fixed, covariates and variances parameters
- Fused LASSO penalty is well suited to do group comparison
- Open question:

  Does this algorithm maximizes the penalized likelihood ?
- Perspectives:
  - Variance selection
  - Work in high dimensional settings
  - Take into account nested random effects (inter-occasion variability)