# Detecting a gene effect in pharmacokinetic models: comparison of different methods

**Julie Bertrand, Emmanuelle Comets, France Mentré**
INSERM U738, Paris 7 University, Bichat Hospital, Paris, France.

## Objectives

Genetic factors constitute part of the interindividual variability in pharmacokinetics (PK). The impact of genetic polymorphisms on pharmacokinetics is often analyzed using a non-compartmental approach but this requires extensive pharmacokinetic sampling and brings limited information, whereas modeling approaches provide deeper insight in the pharmacokinetics and the underlying processes. With non-linear mixed effects models, several methods can be used for the inclusion of genetic covariates during model building. In this work we place ourselves in the framework of a design devised to show the influence of a single nucleotide polymorphism (SNP) on the bioavailability of a drug, and we evaluate by simulation the statistical properties of strategies using non-linear mixed effects models.

## Simulation study

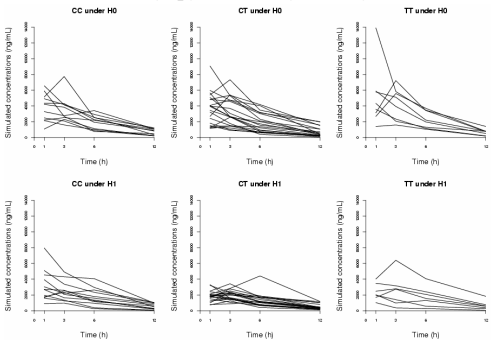
### Models

- Statistical model:
  - $f$ is a classic PK model with one compartment, first order absorption and elimination, at steady state
    $$y_{i,j} = f(t_{i,j}, \theta_i) + \varepsilon_{i,j}$$
  - parameters $\theta_i = \{ka_i, ke_i, V_i/F_i\}$ with $\theta_i$ defined by fixed effects vector $\mu$ and random effects vector $b_i$:
    $$\theta_i = \mu \times e^{bi}$$
  - residual error $\varepsilon_{i,j}$ normally distributed with 0 mean and variance $\sigma_{i,j}^2$
- Model of the genetic polymorphism effect:
  - SNP (C>T) leading to 3 genotypes: CC, CT, TT
  - effect on bioavailability trough $V/F$
    $$V_i/F_i = V/F \times \beta(G_i) \times e^{bi}$$
    - $G_i$ is the genotype for subject i
    - $\beta(G_i)=\{1,\beta_1,\beta_2\}$ for $G_i$=CC, CT or TT, respectively

### Simulation settings

- Based on COPHAR2-ANRS111 clinical trial, where indinavir concentrations were collected at 1, 3, 6 and 12 hours after two weeks of treatment
  - N = 40 (an average of 9 TT)
- Simulation of two exons combination effect
  $$V_i/F_i = V/F \times \beta(G_{1i}) \times \delta(G_{2i}) \times e^{bi}$$
  - polymorphism distribution and effect inspired from literature on exon 26 and 21 of MDR1[1]
- 1000 data sets simulated under H0 ($M_0$)
  - evaluation of type I error
- 1000 data sets simulated under H1 ($M_{CCvsCTvsTT}$)
  - $\beta(G_{1i})=\{1,\beta_1=1.2,\beta_2=1.6\}$, $\delta(G_{2i})=\{1,\delta_1=1.2, \delta_2=1.3\}$
  - evaluation of power and corrected power (with the 5th percentile computed under H0 as threshold)

### Drug simulated concentrations under H0 (top) and H1 (bottom)




## Testing a gene effect

### Methods

- Three estimation methods
  - FO and FOCE in NONMEM
  - SAEM in Monolix using EM and MCMC approaches[2]
- Test based on an ANOVA
  - the empirical Bayes estimates (EBE) of the individual PK parameters from the model with no covariate ($M_0$) are compared between the 3 genotypes using ANOVA
- Wald test
  - Wald tests of the estimates of $\beta_1$ and $\beta_2$ from the model with the covariate in 3 classes ($M_{CCvsCTvsTT}$)
  - three tests: $\{\beta_1=1\}$, $\{\beta_2=1\}$ and $\{\beta_1-\beta_2=0\}$ using estimation errors (SE) of estimates
  - the global test is significant if at least one of the tests is significant with alpha=0.05/3
- Likelihood ratio test (LRT)
  - the models $M_0$ and $M_{CCvsCTvsTT}$ are compared using the LRT with a $\chi^2$ with 2 degrees of freedom

### Type one error of the tests

| Test | Algorithm | Number of data sets | Type I error (%) |
|---|---|---|---|
| ANOVA | FO | 991 | 21.6 |
| | FOCE | 987 | 5.6 |
| | SAEM | 1000 | 5.3 |
| LRT | FO | 989 | 46.9 |
| | FOCE | 965 | 7.9 |
| | SAEM | 1000 | 5.8 |
| Wald | FO | 976 | 20.5 |
| | FOCE | 928 | 9.3 |
| | SAEM | 1000 | 8.1 |

- FO algorithm shows bad performances for the 3 tests.
- For ANOVA, both FOCE and SAEM have a type one error close to 5%.
- For the LRT, FOCE shows a slight significant increase
- FOCE and SAEM obtain a significantly elevated type I error for the Wald test

### Power of the tests

| Test | Algorithm | Number of data sets | Power (%) | Corrected power (%) |
|---|---|---|---|---|
| ANOVA | FOCE | 970 | 71.2 | 69.0 |
| | SAEM | 1000 | 71.2 | 70.1 |
| LRT | FOCE | 949 | 78.7 | 70.9 |
| | SAEM | 1000 | 77.6 | 73.7 |
| Wald | FOCE | 914 | 55.5 | 31.2 |
| | SAEM | 1000 | 81.7 | 72.7 |

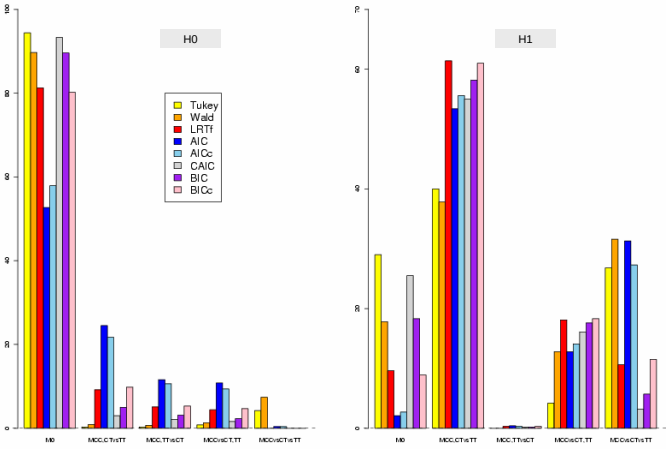
- Due to its results in term of type I error, FO power estimates are not represented on the table
- Using FOCE or SAEM the three strategies have a corrected power around 70%, except for the Wald test for FOCE
- Only SAEM achieve convergence on all data sets for all models

## Strategies for model building

### Methods

- Models
  - $M_0$: no gene effect
  - $M_{CCvsCTvsTT}$: gene effect in 3 classes
  - $M_{CC,CTvsTT}$, $M_{CC,TTvsCT}$, and $M_{CCvsCT,TT}$: three intermediate models with the covariate in two classes
- Selection based on tests
  - Selection based on Tukey tests after ANOVA on the EBE from model $M_0$
    - $M_0$ is selected, if none of the 3 Tukey tests is significant
    - $M_{CCvsCTvsTT}$ is selected if the three tests are significant
    - intermediate models are selected depending on which tests are significant
  - Selection on Wald tests on the estimates of the genotype effects from $M_{CCvsCTvsTT}$
    - tests as described previously
    - model selection similar to that using EBE
  - Forward selection using the LRT
- Selection based on criterion
  - Several criterion are studied, the model with the minimal criterion is chosen
    - $AIC = -2L + 2P$
    - $AICc = AIC + \frac{2P(P+1)}{ntot-P-1}$
    - $CAIC = -2L + P(\ln ntot+1)$
    - $BIC = -2L + P \ln ntot$
    - $BICc = -2L + P \ln N$
    - where $L$: model loglikelihood; $P$: total number of population model parameters, $N$: sample size, $ntot$: total number of observations

### Percentage of data sets where each model was selected




- Under H0, the AIC and AICc show poor results whereas other strategies choose the correct model more often
- Under H1, with all methods the model $M_{CC,CTvsTT}$ is more often selected than the true one $M_{CCvsCTvsTT}$.
- The selection result strongly depends on the strategy and/or the criteria.

## Discussion

- FOCE ran into convergence problems in up to 9% of the data sets tested, while SAEM provided estimates for all models
- With FO, false covariate inclusion was very important for all tests
- With a realistic design, ANOVA based on EBE and LRT maintained a 5% type I error using SAEM.
- Once corrected with the result under H0, the power was similar for the 3 strategies for FOCE and SAEM, except for the Wald test with FOCE, where correlation between estimates and their estimation error leads to lack of power
- Under H0, AIC and AICc show poor selection capacity
- Under H1, performances to detect the good model where somewhat disappointing, but the design is rather small (40 patients)
- Further studies are required to provide recommendations in model selection strategies

(1) Marzolini C, Paus E, Buclin T, Kim R B. Polymorphisms in Human MDR1 (P-glycoprotein): Recent advances and clinical relevance. *Clinical Pharmacology and Therapeutics* 2003; **75**:13-33.
(2) Kuhn and Lavielle. Maximum likelihood estimation in nonlinear mixed effects model. *Computational Statistics and Data Analysis* 2005; **49**:1020-1038.