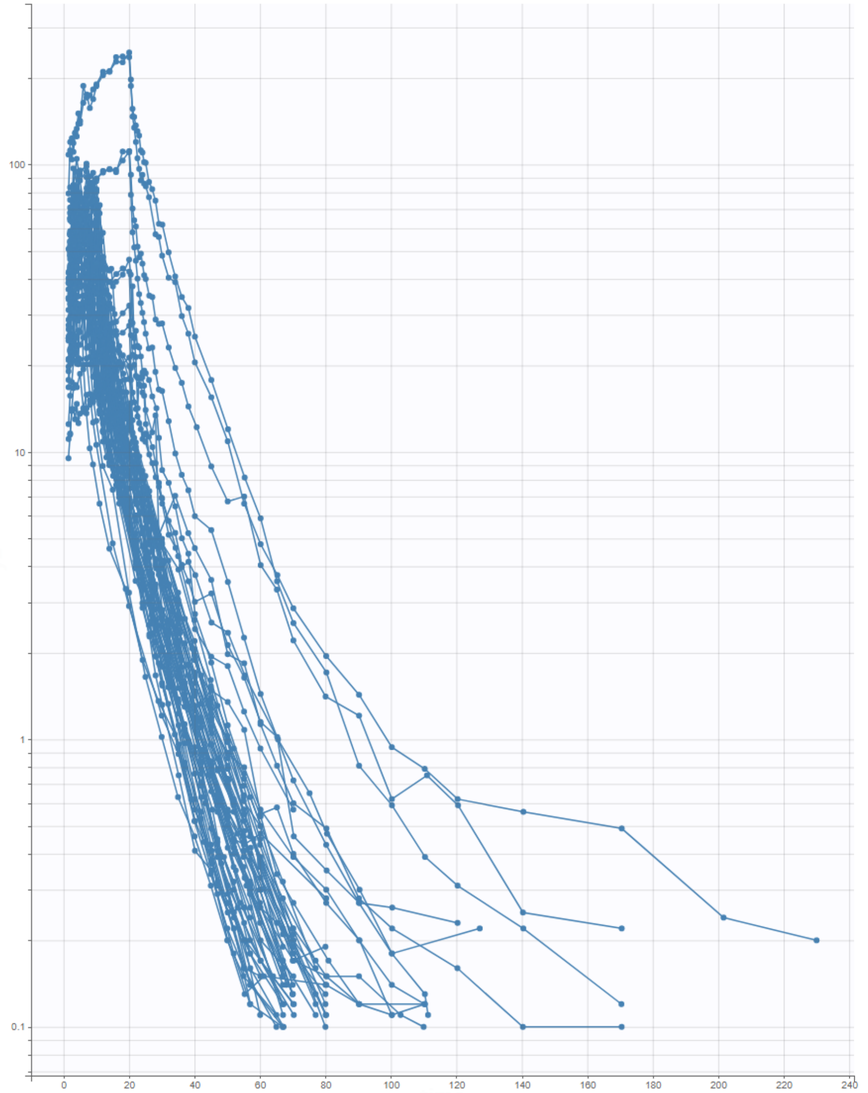
30th PAGEmeeting
Ljubljana • Slovenia
28 June – 1 July 2022
PAGE
Population Approach Group Europe

Automatic construction of nonlinear mixed-effects models with SAMBA

Marc Lavielle & Mélanie Prague

Inria
Inserm
ÉCOLE POLYTECHNIQUE
IP PARIS

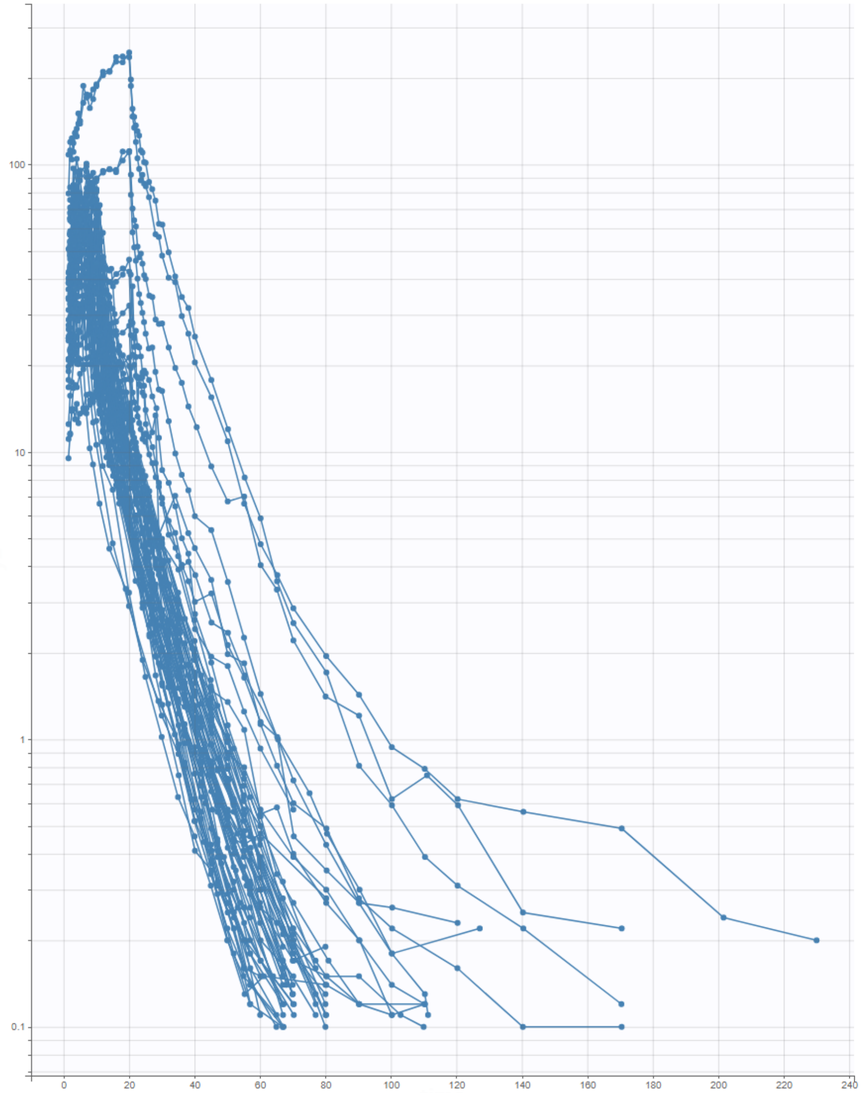
Building a pharmacometric model is a complex process that requires in-depth expertise and can be very time consuming.



Building a pharmacometric model is a complex process that requires in-depth expertise and can be very time consuming.

Questions are many:

- what is the “best” structural model ? *(1, 2, 3cpt? linear elimination?)*



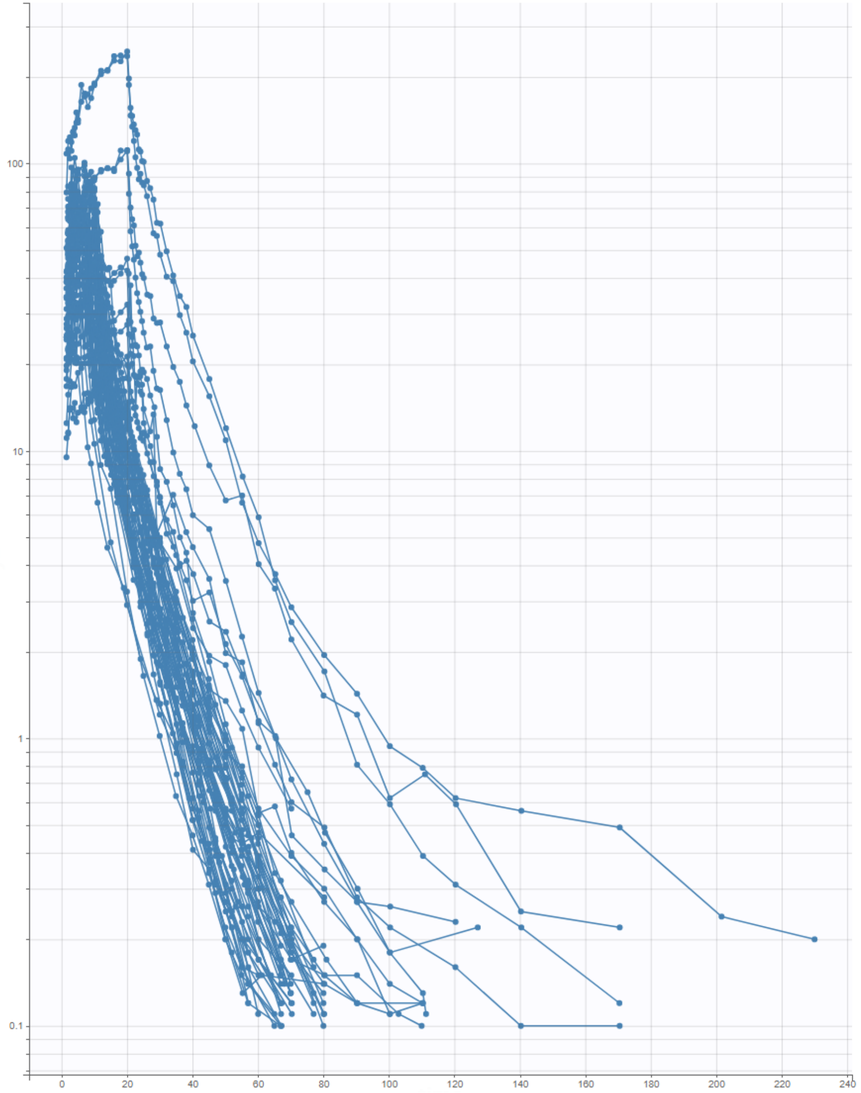
Building a pharmacometric model is a complex process that requires in-depth expertise and can be very time consuming.

Questions are many:

- what is the “best” structural model ? *(1, 2, 3cpt? linear elimination?)*

- what is the “best” statistical model ?

  - covariate model *(is V related to WEIGHT?)*
  - correlation model *(are η_V and η_{CL} correlated?)*
  - variance model *(does k_{12} show inter individual variability?)*
  - residual error model *(constant variance?)*



Standard approach for selecting the *"best"* model among all the possible models:

Minimize a global criterion

Penalized likelihood: $-2\,LL(\mathcal{M}; y) + pen(\mathcal{M})$

Standard approach for selecting the “best” model among “all” the available models:

Minimize a global criterion

Penalized likelihood: $-2\,LL(\mathcal{M}; y) + pen(\mathcal{M})$

AIC $-2\,LL(\mathcal{M}; y) + 2 \times nb.param(\mathcal{M})$

BIC $-2\,LL(\mathcal{M}; y) + \log(N) \times nb.param(\mathcal{M})$

BICc $-2\,LL(\mathcal{M}; y) + \log(N) \times nb_1(\mathcal{M}) + \log(n_{obs}) \times nb_2(\mathcal{M})$

$\widehat{\mathcal{M}}$ minimizes - 2 LL($\mathcal{M}$; y) + pen($\mathcal{M}$)

Two “orthogonal” challenges:

1. **Statistical challenge:** choose a “good” criterion
   => *able to select a model with good predictive performances*

$\widehat{\mathcal{M}}$ minimizes - 2 LL($\mathcal{M}$; y) + pen($\mathcal{M}$)

Two “orthogonal” challenges:

1. **Statistical challenge:** choose a “good” criterion => **BICc**
   => *able to select a model with good predictive performances*

$$\widehat{\mathcal{M}} \text{ minimizes } -2\,LL(\mathcal{M}\,;\,y) + pen(\mathcal{M})$$

Two “orthogonal” challenges:

1. **Statistical challenge:** choose a “good” criterion **=> BICc**
   => *able to select a model with good predictive performances*

2. **Computational challenge:** minimize the chosen criterion
   => *able to find a good solution in a decent time*
   (generally impossible to fit and compare all possible models)

$$\widehat{\mathcal{M}} \text{ minimizes } -2\,LL(\mathcal{M}\,;\,y) + pen(\mathcal{M})$$

Two “orthogonal” challenges:

1. **Statistical challenge:** choose a “good” criterion **=> BICc**
   => *able to select a model with good predictive performances*

2. **Computational challenge:** minimize the chosen criterion **=> SAMBA**
   => *able to find a good solution in a decent time*
   (generally impossible to fit and compare all possible models)

SAMBA

Stochastic Approximation for Model Building Algorithm

- Choose an initial model $\mathcal{M}_o$

- At iteration *k*

  - **Simulation step**: Generate a sequence of *simulated* individual parameters $\psi^{(k)} = (\psi_1^{(k)}, \ldots, \psi_N^{(k)})$

$$\psi_i^{(k)} \sim \mathrm{p}(\psi_i | y; \mathcal{M}_{k-1})$$

  - **Maximization step**: Select a new model $\mathcal{M}_k$

$$\mathcal{M}_k = \arg\min_{\mathcal{M} \in \mathbb{M}} \left\{ -2\mathcal{LL}(\mathcal{M}\ ;\ y, \psi^{(k)}) + \mathrm{pen}(\mathcal{M}) \right\}$$

Example 1: the covariate model

2 available covariates WT and AGE
4 possible covariate models for the volume *V*:

$$\log(V_i) = \log(V_{\text{pop}}) + \eta_i \qquad V_i = V_{\text{pop}} e^{\eta_i}$$

$$\log(V_i) = \log(V_{\text{pop}}) + \beta_W \log(\text{WT}_\text{i}/70) + \eta_i \qquad V_i = V_{\text{pop}} (\text{WT}_\text{i}/70)^{\beta_W} e^{\eta_i}$$

$$\log(V_i) = \log(V_{\text{pop}}) + \beta_A \log(\text{AGE}_\text{i}/40) + \eta_i \qquad V_i = V_{\text{pop}} (\text{AGE}_\text{i}/40)^{\beta_A} e^{\eta_i}$$

$$\log(V_i) = \log(V_{\text{pop}}) + \beta_W \log(\text{WT}_\text{i}/70) + \beta_A \log(\text{AGE}_\text{i}/40) + \eta_i \qquad V_i = V_{\text{pop}} (\text{WT}_\text{i}/70)^{\beta_W} (\text{AGE}_\text{i}/40)^{\beta_A} e^{\eta_i}$$

Example 1: the covariate model

2 available covariates WT and AGE
4 possible covariate models for the volume *V*:

$$\log(V_i) = \log(V_{\text{pop}}) + \eta_i \qquad V_i = V_{\text{pop}} e^{\eta_i}$$
$$\log(V_i) = \log(V_{\text{pop}}) + \beta_W \log(\text{WT}_\text{i}/70) + \eta_i \qquad V_i = V_{\text{pop}} (\text{WT}_\text{i}/70)^{\beta_W} e^{\eta_i}$$
$$\log(V_i) = \log(V_{\text{pop}}) + \beta_A \log(\text{AGE}_\text{i}/40) + \eta_i \qquad V_i = V_{\text{pop}} (\text{AGE}_\text{i}/40)^{\beta_A} e^{\eta_i}$$
$$\log(V_i) = \log(V_{\text{pop}}) + \beta_W \log(\text{WT}_\text{i}/70) + \beta_A \log(\text{AGE}_\text{i}/40) + \eta_i \qquad V_i = V_{\text{pop}} (\text{WT}_\text{i}/70)^{\beta_W} (\text{AGE}_\text{i}/40)^{\beta_A} e^{\eta_i}$$

1. Fit these 4 linear models using the simulated volumes $V_1^{(k)}$, $V_2^{(k)}$, … , $V_N^{(k)}$
2. Select the covariate model with the smallest BICc

Example 2: the correlation model

3 random effects: η_{ka} , η_V , η_{CL}
5 possible correlation models:

- No correlation
- Correlation between η_{ka} and η_V
- Correlation between η_{ka} and η_{CL}
- Correlation between η_V and η_{CL}
- Correlation between η_{ka} , η_V and η_V

Example 2: the correlation model

3 random effects: η_{ka} , η_V , η_{CL}
5 possible correlation models:

- No correlation
- Correlation between η_{ka} and η_V
- Correlation between η_{ka} and η_{CL}
- Correlation between η_V and η_{CL}
- Correlation between η_{ka} , η_V and η_V

1. Fit these 5 Gaussian models using the simulated random effects $\eta_1^{(k)}$, … , $\eta_N^{(k)}$

2. Select the correlation model with the smallest BICc

Comparison with other existing procedures for covariate model building

10 representative pharmacometrics datasets

Comparison with SCM and COSSAC in terms of:

- total number of runs for reaching the solution
- value of the criterion (BICc)

		SCM		COSSAC		SAMBA		Δ BICc	
Dataset	Characteristics	#Runs[2]	Final Model[1]	#Runs[2]	Final Model[1]	#Runs[2]	Final Model[1]	SAMBA-SCM	SAMBA-COSSAC
Warfarin Linear PK	32 ind. - 247 obs. 4 param. - 3 cov. 4 re - 1 outcome	44	logtWt - V, Cl logtAge - C	4	Identical	2	Identical	0	0
Remifentanil Linear PK	65 ind. - 1992 obs. 6 param. - 6 cov. 4 re - 1 outcome	295	logLBM - $V1$ logAGE - Cl,$Q2$,$\mathbf{Q3}$,$V2$,$V3$ logBSA - Cl logHT - $\mathbf{V2}$	13	logLBM - $V1$,$\mathbf{V2}$ logAGE - Cl,$Q2$,$V2$,$V3$ logBSA - Cl SEX - $\mathbf{V3}$	4	logLBM - $V1$ logAGE - Cl,$Q2$,$\mathbf{Q3}$,$V2$,$V3$ logBSA - Cl SEX - $\mathbf{V2}$	0.8	0.5
Theophylline Linear PK	12 ind. - 20 obs. 3 param. - 2 cov. 4 re - 1 outcome	12	logtWEIGHT - $\mathbf{ka}$	4	Identical	2	Identical	0	0
Quinidine Sparse PK	136 ind. - 361 obs. 3 param. - 2 cov. 3 re - 1 outcome	22	none	11	Identical	1	Identical	0	0
Tobramycin Sparse PK	97 ind. - 322 obs. 3 param. - 2 cov. 2 re - 1 outcome	22	logCLCR - Cl logWT - $\mathbf{V}$	6	logCLCR - Cl logWT - $\mathbf{V}$	2	logCLCR - Cl logWT - $\mathbf{Cl}$	4.2	4.2
Theophylline Ext. Rel. Linear PK	18 ind. - 362 obs. 7 param. - 3 cov. 7 re - 1 outcome	98	logWT - $\mathbf{Tlag1}$, $\mathbf{V}$	8	logWT - $\mathbf{Tlag1}$ logAGE - $\mathbf{ka2}$	6	logWT - $\mathbf{F}$, $\mathbf{V}$ logAGE - $\mathbf{F}$ logHT - $\mathbf{ka1}$, $\mathbf{ka2}$, $\mathbf{Tlag1}$, $\mathbf{diffTlag2}$	-12	-21
Warfarin PK/PD Joint	32 ind. - 247+232 obs. 8 param. - 3 cov. 8 re - 2 outcomes	92	logWT - Cl	10	logWT - Cl	2	logWT -Cl,$\mathbf{V}$ logAGE - $\mathbf{Cl}$, $R0$	-1.4	-1.4
Cholesterol Disease Progression	200 ind. - 1044 obs. 2 param. - 2 cov. 2 re - 1 outcome	12	logAGE - $Chol0$,$\mathbf{slope}$ SEX - $\mathbf{slope}$	5	logAGE - $Chol0$,$\mathbf{slope}$ SEX - $\mathbf{slope}$	2	logAGE - $Chol0$	13.5	13.5
Alzheimer Sparse PK	896 ind. - 3707 obs. 2 param. - 7 cov. 2 re - 1 outcome	73	APOE - $alpha$,$p0$ logAGE - $p0$,$\mathbf{alpha}$ logBMI - $\mathbf{alpha}$ logWT - $p0$	8	APOE - $alpha$,$p0$ logAGE - $p0$,$\mathbf{alpha}$ logBMI - $\mathbf{alpha}$ logWT - $p0$	2	APOE - $alpha$,$p0$ logAGE - $p0$ logWT - $p0$	6	1.5
Tranexamic PK	166 ind. - 817 obs. 4 param. - 10 cov. 4 re - 1 outcome	298	GROUP - Cl,$V2$ logBMI - Cl logCOCK - Cl logLBW - Q logWeight - V2	12	Identical	2	Identical	0	0

- SAMBA is implemented in Monolix and Rsmlx (R speaks Monolix)

- Some other features for model building implemented in Rsmlx 5.0.1:
  - automatic selection of the PK model (from a PK model library)
  - automatic selection of the variance model
  - introduction of prior information about the statistical model

Monte-Carlo experiment

Number of replicates: $M = 100$

Number of individuals: $N = 100$

Observation times: $0.1, 0.25, 0.5, 1, 2, 4, 6, 8, 12, 16, 20, 24h$

PK model: 2 compartments model for oral administration $(ka, Cl, V1, Q, V2)$

Population parameters: $ka_{\text{pop}} = 0.6$, $Cl_{\text{pop}} = 5$, $V1_{\text{pop}} = 6$, $Q_{\text{pop}} = 20$, $V2_{\text{pop}} = 10$.

Variance-covariance of the random effects: $\omega_{ka} = \omega_{V1} = \omega_{Cl} = 0.3$; $\omega_Q = \omega_{V2} = 0$; $\rho_{Cl,V1} = 0.7$

Residual error model: proportional model $b = 0.15$

10 covariates: $X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}$ *i.i.d.* $\mathcal{N}(0,1)$

Covariate model:

$$\begin{aligned}
\log(ka_i) &= \log(ka_{\text{pop}}) + \eta_{i,ka} \\
\log(Cl_i) &= \log(Cl_{\text{pop}}) + \beta_{Cl,X3} X_{i,3} + \eta_{i,Cl} \\
\log(V1_i) &= \log(V1_{\text{pop}}) + \beta_{V1,X1} X_{i,1} + \beta_{V1,X2} X_{i,2} + \eta_{i,V1} \\
\log(Q_i) &= \log(Q_{\text{pop}}) + \beta_{Q,X1} X_{i,1} \\
\log(V2_i) &= \log(V2_{\text{pop}})
\end{aligned}$$

$\beta_{Cl,X3} = 0.3$, $\beta_{V1,X1} = 0.2$, $\beta_{V1,X2} = 0.3$, $\beta_{Q,X1} = 0.5$,

Monte-Carlo experiment

Number of replicates: $M = 100$

Number of individuals: $N = 100$

Observation times: $0.1, 0.25, 0.5, 1, 2, 4, 6, 8, 12, 16, 20, 24h$

PK model: 2 compartments model for oral administration $(ka, Cl, V1, Q, V2)$

Population parameters: $ka_{\text{pop}} = 0.6$, $Cl_{\text{pop}} = 5$, $V1_{\text{pop}} = 6$, $Q_{\text{pop}} = 20$, $V2_{\text{pop}} = 10$.

Variance-covariance of the random effects: $\omega_{ka} = \omega_{V1} = \omega_{Cl} = 0.3$; $\omega_Q = \omega_{V2} = 0$; $\rho_{Cl,V1} = 0.7$

Residual error model: proportional model $b = 0.15$

10 covariates: $X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}$ *i.i.d.* $\mathcal{N}(0,1)$

2^{50} possible covariate models!

Covariate model:

$$
\begin{aligned}
\log(ka_i) &= \log(ka_{\text{pop}}) + \eta_{i,ka} \\
\log(Cl_i) &= \log(Cl_{\text{pop}}) + \beta_{Cl,X3} X_{i,3} + \eta_{i,Cl} \\
\log(V1_i) &= \log(V1_{\text{pop}}) + \beta_{V1,X1} X_{i,1} + \beta_{V1,X2} X_{i,2} + \eta_{i,V1} \\
\log(Q_i) &= \log(Q_{\text{pop}}) + \beta_{Q,X1} X_{i,1} \\
\log(V2_i) &= \log(V2_{\text{pop}})
\end{aligned}
$$

$\beta_{Cl,X3} = 0.3$, $\beta_{V1,X1} = 0.2$, $\beta_{V1,X2} = 0.3$, $\beta_{Q,X1} = 0.5$,

id	time	y	amount	X01	X02	X03	X04	X05	X06	X07	X08	X09	X10
1	0	.	1000	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	0.1	12.121	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	0.25	25.856	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	0.5	39.487	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	1	34.788	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	2	37.137	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	4	22.598	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	6	8.8408	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	8	4.7898	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	12	1.2996	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	16	0.32332	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	20	0.07816	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	24	0.02103	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177

id	time	y	amount	X01	X02	X03	X04	X05	X06	X07	X08	X09	X10
1	0	.	1000	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	0.1	12.121	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	0.25	25.856	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	0.5	39.487	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	1	34.788	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	2	37.137	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	4	22.598	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	6	8.8408	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	8	4.7898	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	12	1.2996	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	16	0.32332	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	20	0.07816	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	24	0.02103	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177

```
# data information
d <- list(dataFile="data001.csv",
          headerTypes = c("id","time" ,"observation", "amount",
                          "contcov", "contcov", "contcov", "contcov", "contcov",
                          "contcov", "contcov", "contcov", "contcov", "contcov"),
          administration = "oral")
```

id	time	y	amount	X01	X02	X03	X04	X05	X06	X07	X08	X09	X10
1	0	.	1000	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	0.1	12.121	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	0.25	25.856	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	0.5	39.487	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	1	34.788	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	2	37.137	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	4	22.598	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	6	8.8408	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	8	4.7898	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	12	1.2996	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	16	0.32332	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	20	0.07816	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177
1	24	0.02103	.	-0.502	-0.333	0.028	-1.855	-1.285	-1.446	1.934	0.507	-0.316	0.177

```
# data information
d <- list(dataFile="data001.csv",
          headerTypes = c("id","time" ,"observation", "amount",
                          "contcov", "contcov", "contcov", "contcov", "contcov",
                          "contcov", "contcov", "contcov", "contcov", "contcov"),
          administration = "oral")

# build the structural model and the complete statistical model
r <- pkbuild(data=d    , stat="TRUE")
```

The “true” PK model is selected using BICc (or AIC or BIC…)

	model	OFV	AIC	BIC	BICc
1	lib:oral1_2cpt_kaClV1QV2.txt	4002.784	4026.784	4058.046	4075.440
2	lib:oral1_2cpt_TlagkaClV1QV2.txt	4002.053	4030.053	4066.526	4086.405
3	lib:oral1_3cpt_kaClV1Q2V2Q3V3.txt	4006.107	4038.107	4079.790	4102.154
4	lib:oral1_1cpt_kaVCl.txt	4551.987	4567.987	4588.829	4601.253
5	lib:oral1_1cpt_TlagkaVCl.txt	4552.544	4572.544	4598.595	4613.505
6	lib:oral0_1cpt_Tk0VCl.txt	5127.031	5143.031	5163.872	5176.297
7	lib:oral0_1cpt_TlagTk0VCl.txt	5125.823	5145.823	5171.875	5186.784

Computing time = 3’

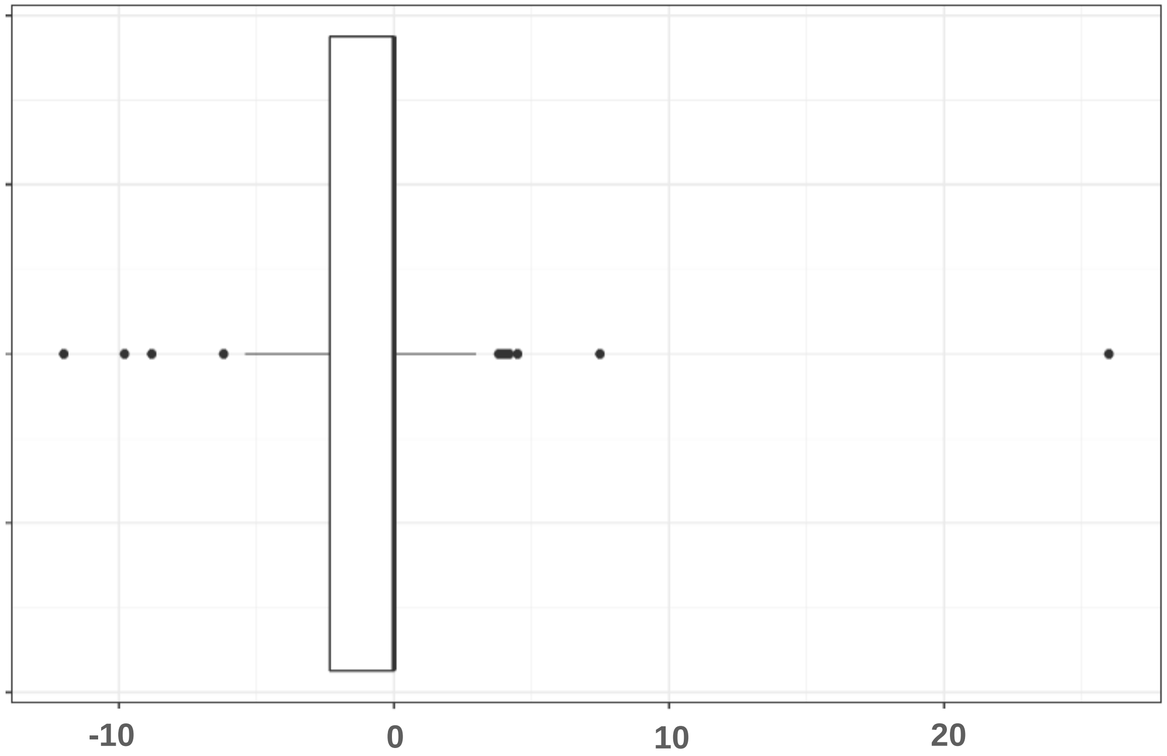
The “true” statistical model is also selected:

```
Variance model:
Parameters without variability: Q V2
Parameters with variability   : ka Cl V1

Covariate model:
   X01 X02 X03 X04 X05 X06 X07 X08 X09 X10
ka   0   0   0   0   0   0   0   0   0   0
Cl   0   0   1   0   0   0   0   0   0   0
V1   1   1   0   0   0   0   0   0   0   0
Q    1   0   0   0   0   0   0   0   0   0
V2   0   0   0   0   0   0   0   0   0   0

Correlation model:
[[1]]
[1] "Cl" "V1"


Residual error model:
             y
"proportional"
```

Computing time = 5’

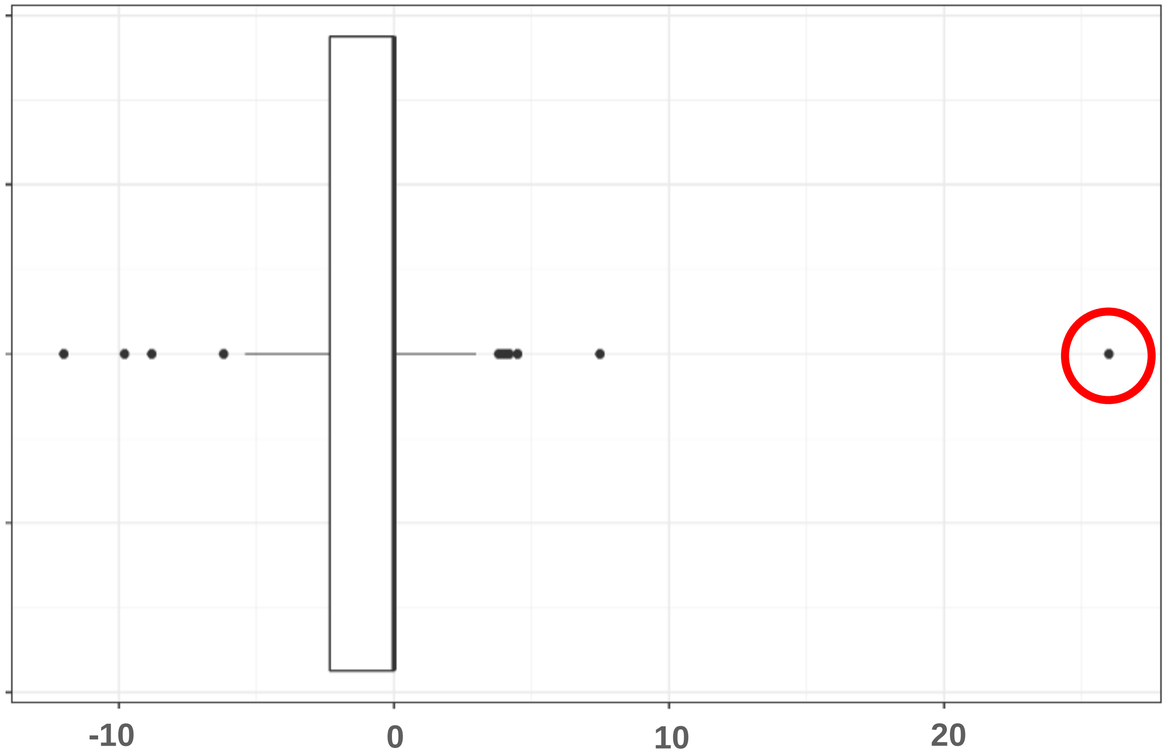
Monte-Carlo results

Difference $\widehat{BICc}$ - BICc*



Monte-Carlo results

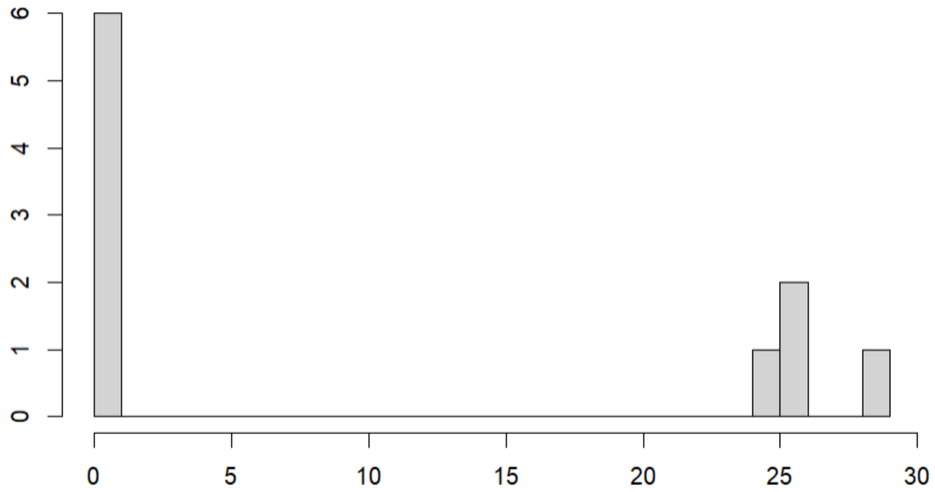
Difference $\widehat{BICc}$ - BICc*



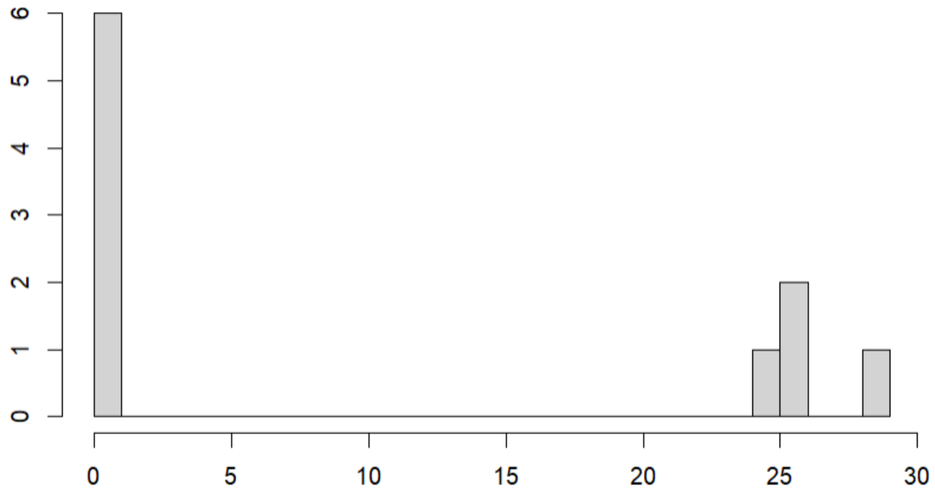
BICc is not correctly minimized for one of the simulated data:

$\widehat{BICc}$ - BICc* = 26

10 runs using 10 different seeds
(different sequences of random numbers)



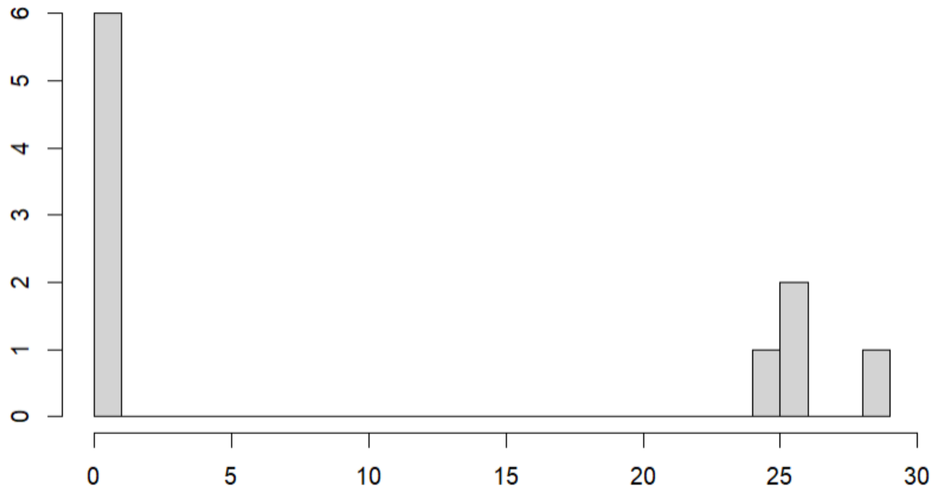
10 runs using 10 different seeds
(different sequences of random numbers)



We don't pretend that SAMBA ***always*** converge to the optimal solution…

But we pretend that SAMBA converges (very) quickly to a (very) good solution very a high probability.

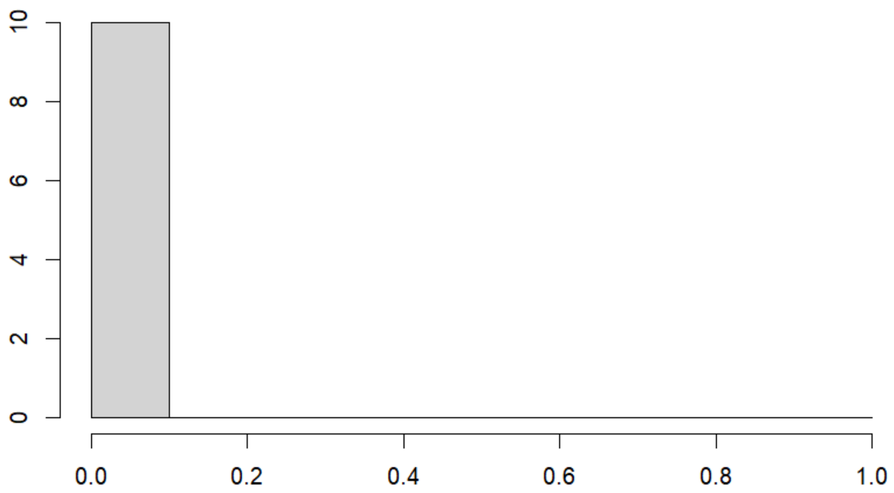
10 runs using 10 different seeds
(different sequences of random numbers)



We don't pretend that SAMBA ***always*** converge to the optimal solution…

But we pretend that SAMBA converges (very) quickly to a (very) good solution very a high probability.

10 runs using a different initial model



Using different initializations allows to improve the results by avoiding to stay stuck in a bad minimum

Some final remarks:

- Under “general conditions” SAMBA Converges to a (possibly local) minimum of the observed criteria

$$-2\,\mathrm{LL}(\mathcal{M}\,;\,y) \;+\; \mathrm{pen}(\mathcal{M})$$

- SAMBA (and SAEM) are algorithms that “learn from their mistakes”. Indeed, fitting a “wrong” model always provides some information about a better model.

- SAMBA does not select a model: it is nothing more than an optimization tool designed for optimizing a criteria selected by the modeler.

- To obtain a biologically meaningful model, a priori information can be introduced (using either a prior distribution or an appropriate penalization).

- By minimizing a penalized criterion, SAMBA maximizes a posterior probability (MAP estimation).
  An upcoming development will consist in estimating this posterior distribution by sampling it (full Bayesian approach).

Thank you!

- M. Prague, M. Lavielle, *SAMBA: a Novel Method for Fast Automatic Model Building in Nonlinear Mixed-Effects Models*, CPT: Pharmacometrics and Systems Pharmacology, 2021,
- M. Lavielle, *Some EM-type algorithms for incomplete data model building*, preprint, 2022, [hal:hal-03512130].

Theoretical property of SAMBA

At iteration k,

- **Simulation step**: Generate a sequence of *simulated* individual parameters $\psi^{(k)} = (\psi_1^{(k)}, \ldots, \psi_N^{(k)})$

$$\psi_i^{(k)} \sim \mathrm{p}(\psi_i | y; \mathcal{M}_{k-1})$$

- **Maximization step**: Select a new model $\mathcal{M}_k$

$$\mathcal{M}_k = \arg \min_{\mathcal{M} \in \mathbb{M}} \left\{ -2 \sum_{j=1}^{k} \gamma_{k-j} \mathcal{LL}(\mathcal{M} \, ; \, y, \psi^{(j)}) + \mathrm{pen}(\mathcal{M}) \right\}$$

Theoretical property of SAMBA

At iteration k,

- **Simulation step**: Generate a sequence of *simulated* individual parameters $\psi^{(k)} = (\psi_1^{(k)}, \ldots, \psi_N^{(k)})$

$$\psi_i^{(k)} \sim \mathrm{p}(\psi_i | y; \mathcal{M}_{k-1})$$

- **Maximization step**: Select a new model $\mathcal{M}_k$

$$\mathcal{M}_k = \arg \min_{\mathcal{M} \in \mathbb{M}} \left\{ -2 \sum_{j=1}^{k} \gamma_{k-j} \mathcal{LL}(\mathcal{M} \,;\, y, \psi^{(j)}) + \mathrm{pen}(\mathcal{M}) \right\}$$

Proposition *Under "general conditions", SAMBA converges to a (possibly local) minimum of the observed criterion*

$$-2\mathcal{LL}(\mathcal{M} \,;\, y) + \mathrm{pen}(\mathcal{M})$$

Relationship between prior distribution $p(\mathcal{M})$ and penalization pen($\mathcal{M}$)

Maximizing the posterior distribution $p(y \mid \mathcal{M})$

$\Leftrightarrow$

Minimizing the penalized criteria $-2\log(p(y \mid \mathcal{M})) - 2\log(p(\mathcal{M}))$

Relationship between prior distribution $p(\mathcal{M})$ and penalization pen($\mathcal{M}$)

Maximizing the posterior distribution $p(y \mid \mathcal{M})$

$\Leftrightarrow$

Minimizing the penalized criteria $-2\log(p(y \mid \mathcal{M})) - 2\log(p(\mathcal{M}))$

BIC (or BICc) for the covariate model: $\text{pen}(\mathcal{M}) = \log(N) \times \#\beta\text{'s}$

$\Leftrightarrow$

Prior distribution for the covariate model: $p(\beta_{jk} \neq 0) = \frac{1}{1+\sqrt{N}}$

Some of the existing methods for covariate model building:

- **SCM** (Stepwise Covariate Modelling)

Jonsson, E. N. & Karlsson, M. O. Automated covariate model building within NONMEM. *Pharm. Res.* 15, 1463–1468 (1998).

- **COSSAC** (COnditional Sampling use for Stepwise Approach based on Correlation tests)

Ayral G, Si Abdallah J-F, Magnard C, Chauvin J. A novel method based on unbiased correlations tests for covariate selection in nonlinear mixed-effects models–the COSSAC approach. *CPT Pharmacomentrics Sys Pharmacol*.10(4):318-329 (2021).

- **EBEs regression using GAM** (Generalized Additive Model)

Mandema, J. W., Verotta, D. & Sheiner, L. B. Building population pharmacokinetic-pharmacodynamic models. I. Models for covariate effects. *J. Pharmacokinet. Biopharm.* 20, 511–528 (1992).