Extension of the covariate clinical relevance determination in population pharmacokinetic analyses to FREM, LASSO & GA approaches

Morgane PHILIPP[1,2], Sylvie RETOUT[2,3], Simon BUATOIS[3], France MENTRE[1]

[1] Université Paris Cité, INSERM, IAME, UMR 1137, Paris, France
[2] Institut Roche, Boulogne-Billancourt, France
[3] Roche Pharma Research and Early Development, Pharmaceutical Sciences, Roche Innovation Center Basel, Switzerland

CONTEXT

- Covariate clinical relevance (CCR) assessment is essential for patient dose adjustment and is a key step according to health authorities.
- Forest plots are used to illustrate parameter changes for covariate values relative to typical ones.
- CCR is inferred from the position of covariate effect ratios (CER) with their 90% confidence interval (CI90) relative to the reference area/line.
- Our previous work [1] compared the accuracy of CCR assessment of Stepwise Covariate Model (SCM), SCM+ [2] & Full Model (FM) [3]; all approaches provided satisfactory results close to those of the reference model (estimations using the simulated covariate models).
- Here, 3 alternative approaches are investigated: full random effect modelling (FREM) [4], least absolute shrinkage and selection operator (LASSO) [5] & a genetic algorithm (GA) [6]; PsN algorithms were used for FREM & LASSO; the algorithm developed by Ronchi et al. [6] was used for the GA.

OBJECTIVE

To compare the results of CCR determination by FREM, LASSO & GA with the ones by FM, SCM+ and the reference model in a proof of concept study

METHODS

Simulation study

PK model [1,7]
One compartment model (V/F)
First order absorption (ka)
Linear elimination (CL/F)

Design [1,7]
N = 383
4 phase III studies n = {1, ..., 26} &
1 phase I/II study n = {21, ..., 71}

Error model [1,7]
Combined (a and b)
→ with a the additive (fixed) and b the proportional term

Covariate models [1,7]
Body weight (BW), age (AGE), albumin (ALB), black (BLK) race

Base model

$$CL/F_i = \mu_{CL/F} \times \left(\frac{BW_i}{70}\right)^{\beta_{CL/F,BW}} \times e^{\eta_{CL/F,i}}$$

$$V/F_i = \mu_{V/F} \times \left(\frac{BW_i}{70}\right)^{\beta_{V/F,BW}} \times e^{\eta_{V/F,i}}$$

$$KA_i = \mu_{KA} \times e^{\eta_{KA,i}}$$

Covariate model

$$CL/F_i = \mu_{CL/F} \times \left(\frac{BW_i}{70}\right)^{\beta_{CL/F,BW}} \times \left(\frac{AGE_i}{30}\right)^{\beta_{CL/F,AGE}} \times \left(\frac{ALB_i}{45}\right)^{\beta_{CL/F,ALB}} \times e^{\eta_{CL/F,i}}$$

$$V/F_i = \mu_{V/F} \times \left(\frac{BW_i}{70}\right)^{\beta_{V/F,BW}} \times \left(\frac{AGE_i}{30}\right)^{\beta_{V/F,AGE}} \times (1 + \beta_{V/F,BLK} \times 1_{RACE_i=BLK}) \times e^{\eta_{V/F,i}}$$

$$KA_i = \mu_{KA} \times e^{\eta_{KA,i}}$$

→ with μ the fixed effects and $\eta_i \sim N(0,\Omega)$ the between subject random-effects of individual i, Ω being the variance-covariance matrix, with $\beta_{par,cov}$ the effect of a covariate (cov) on a parameter (par)

Data simulation
1 dataset simulated with the Base model & with the Covariate model

Parameter estimation
Software: NONMEM 7.4
Estimation algorithm: FOCEi
SE: derived from the covariance matrix computed as $R^{-1}SR^{-1}$
(R: hessian matrix, S: cross-product gradient matrix)

Reference model fitting
Software: PsN 5.3.2
Base model & Covariate model

Covariate model building

Parsimonious covariate set (14 relations par,cov, 18 $\beta_{par,cov}$) & Saturated covariate set (21 relations par,cov, 27 $\beta_{par,cov}$)

Continuous covariates (med [min - max])	CL/F	V/F	KA	Categorical covariates category (N [%])	CL/F	V/F	KA
Body Weight (BW, kg), 69.1 [9.50–156]	B: $\beta_{CL/F,BW}$ = 0.939 C: $\beta_{CL/F,BW}$ = 0.801, Y, Y	B: $\beta_{CL/F,BW}$ = 1.066 C: $\beta_{CL/F,BW}$ = 0.867, Y, Y	N, Y	**Status** Non-inhibitor 195 [50%] FVIII inhibitor (INH) 194 [50%]	Y, Y	Y, Y	Y, Y
Age (AGE, years), 30.0 [1.22 - 77.00]	C: $\beta_{CL/F,AGE}$ = 0.127, Y, Y	Y, Y	Y, Y				
Albumine (ALB, g/L), 45.0 [33.0–56.6]	C: $\beta_{ALB,BW}$ = -0.918, Y, Y	Y, Y	N, Y	**Race** White, 244 [63%] Black (BLK), 31 [8%] Asian (ASN), 89 [23%] Other (OTH), 25 [6%]	Y, Y	C: $\beta_{V/F,RACE=BLK}$=-0.212, Y, Y	N, Y
Aspartate aminotransferase (AST, U/L), 23.0 [11.0–91.0]	Y, Y	N, Y	N, Y				
Bilirubin (BILI, µmol/L), 9.0 [0.33–46.0]	Y, Y	N, Y	N, Y				

→ B,C: covariate effect simulated for the Based & the Covariate model, respectively
→ Y/N, Y/N: covariate effect tested (yes, Y) or not (No, N) on the parameter considering the Parsimonious or the Saturated covariate set

FM
Software: PsN 5.3.2
Retries = 5

SCM+
Software: PsN 5.3.2
P-cutoff = 0.05
P-forward = 0.05
P-backward = 0.01

FREM
Software: PsN 5.3.2

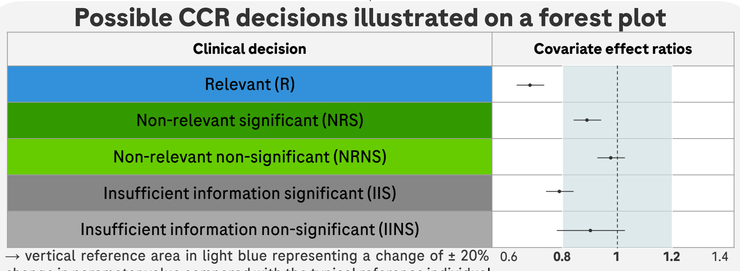
LASSO
Software: PsN 5.3.2
Stop_t = 10
→ fit the selected covariate model with the right relation shape

GA
Software: R 4.0.5 & Perl 5.30.2
Selection pressure = 0.75
Crossover probability = 0.7
Mutation probability = 0.025
Chromosomes preserved by elitism = 10%
GA 150: 150 generations, 50 chromosomes
GA 50: 50 generations, 30 chromosomes

Results evaluation
Covariates selection, CCR assessment using forest plot, OFV = -2LL & BICc = -2LL + P_R x ln(N) + P_F x ln(n_{tot}) values, runtime
→ with LL the log likelihood, P_R the number of µ, $\beta_{par,cov}$ and dim(Ω), P_F the number of error model parameters = 1, N the number of patients and n_{tot} the total number of observations

Forest plot generation
Software: R 4.1.2
Package: PMXForest 1.2.6 (https://github.com/pharmetheus/PMXForest),
PMXFrem 1.2.2 (https://github.com/pharmetheus/PMXFrem),
Number of samples: 1000

Possible CCR decisions illustrated on a forest plot



→ vertical reference area in light blue representing a change of ± 20% change in parameter value compared with the typical reference individual

CONCLUSION

- FREM, LASSO & GA gave satisfactory results for CCR assessment in line with those found with FM & SCM+.
- FREM & FM gave very comparable results, making both methods very suitable for CCR assessment.
- GA150 selected more parsimonious models than GA50 & LASSO and resulted in better BIC; however the number of models generated is substantial and requires a high run time; if computing resources are not an issue, GA150 may be preferable.
- Between full modelling methods and covariate selection methods, the former have the benefits to get a CCR evaluation of all relations while the latter provide a parsimonious model suitable for prediction; the best choice depend on one's objective.
- Further evaluations are necessary to enhance the robustness of those results by using simulation frameworks with increased complexity in model structures, expanded covariate sets, and a larger number of simulated datasets.

References
[1] Philipp et al. *J. Pharmacokinet. Pharmacodyn.* 2024; [2] Svensson et al. *CPT. Pharmacomet. Syst. Pharmacol.* 2022; [3] Xu et al. *Br. J. Clin. Pharmacol.* 2018; [4] Yngman et al. *CPT. Pharmacomet. Syst. Pharmacol.* 2022; [5] Ribbing et al. *J. Pharmacokinet. Pharmacodyn.* 2007; [6] Ronchi et al. *Pharmacokinet. Pharmacodyn.* 2023; [7] Retout et al. *Clin. Pharmacokinet.* 2020;

RESULTS

Covariates selection & CCR assessment

	Simulation case: Base model							Simulation case: Covariate model						
	Ref model	FM	FREM	SCM+	LASSO	GA 150	GA 50	Ref model	FM	FREM	SCM+	LASSO	GA 150	GA 50
Parsimonious covariate set														
CL/F,BW	R	R	/	R	R	R	R	R	R	/	R	R	R	R
CL/F, AGE		NRS	/		NRS			IIS	IIS	/	IIS	IIS	IIS	IIS
CL/F, ALB		NRNS	/					NRS	NRS	/	NRS	NRS	NRS	NRS
CL/F, AST		NRNS	/				NRNS		NRNS	/		NRNS	NRNS	NRNS
CL/F, BILI		NRNS	/						NRNS	/				
CL/F, RACE		NRS	/		NRNS				NRS	/		NRS		IIS
CL/F, INH		NRNS	/						NRNS	/		NRNS		
V/F, BW	R	R	/	R	R	R	R	R	R	/	R	R	R	R
V/F, AGE		NRNS	/					IIS	IIS	/	IIS	IIS	IIS	IIS
V/F, ALB		NRNS	/						NRNS	/				
V/F, RACE		NRS	/		NRS			IIS	IIS	/			IIS	IIS
V/F, INH		NRNS	/		NRS				NRS	/		NRNS	NRS	NRS
KA, AGE		IIS	/	IIS		IIS	IIS		IIS	/		IIS	IIS	
KA, INH		IINS	/				NRS		IINS	/				
Saturated covariate set														
CL/F, BW	R	R	R	R	R	R	R	R	R	R	R	R	R	R
CL/F, AGE		NRNS	NRNS		NRS			IIS	IIS	IIS	IIS	IIS	IIS	IIS
CL/F, ALB		NRNS	NRNS					NRS	NRS	NRS	NRS	NRS	NRS	NRS
CL/F, AST		NRNS	NRNS						NRNS	NRNS		NRNS		
CL/F, BILI		NRNS	NRNS						NRNS	NRNS				
CL/F, RACE		NRNS	NRNS		NRNS				NRS	NRNS		NRNS		NRNS
CL/F, INH		NRNS	NRNS						NRNS	NRNS				
V/F, BW	R	R	R	R	R	R	R	R	R	R	R	R	R	R
V/F, AGE		NRNS	NRNS					IIS	IIS	IIS	IIS	IIS	IIS	IIS
V/F, ALB		NRNS	NRNS						NRNS	NRNS				NRNS
V/F, AST		NRNS	NRNS						NRNS	NRNS				NRNS
V/F, BILI		IIS	IIS	NRS	IIS	IIS	IIS		NRNS	NRNS		NRS		IIS
V/F, RACE		IIS	IIS		IIS		IIS	IIS	IIS	IIS	IIS	IIS	IIS	IIS
V/F, INH		IIS	IIS	NRS	IIS	IIS	IIS		IIS	IIS		NRS	NRS	
KA, BW		IINS	IINS						IINS	IINS	NSEL		IINS	
KA, AGE		IINS	IINS	IIS		IIS	IIS		IINS	IINS				IIS
KA, ALB		IINS	NRNS				NRNS		NRNS	IINS				
KA, AST		IINS	IINS						IIS	IINS				IINS
KA, BILI		NRNS	NRNS						NRNS	NRNS				
KA, RACE		IINS	IINS						IINS	IINS				
KA, INH		IINS	IINS						IINS	IINS				

R | NRS | NRNS | IIS | IINS | NSEL | NIM | / | NP

NSEL: not selected, NIM: not in the model, NP: not performed

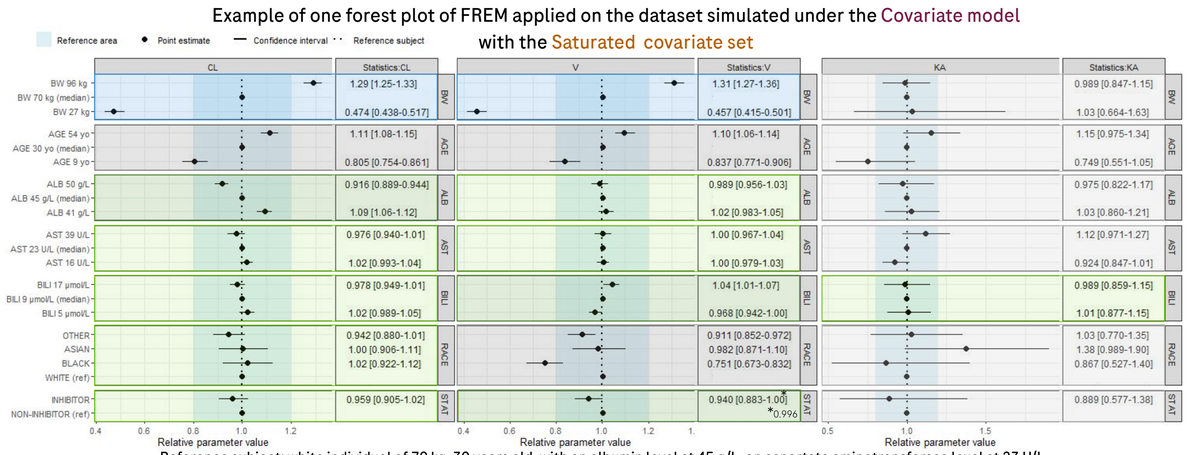
→ **Covariate selection**
- Covariate selection approaches (i.e. SCM+, LASSO & GA) selected all simulated relations
- Few additional relationships were retained, with a larger number using the saturated covariate set and with SCM+ being the more restrictive in selection

Number of additional relations retained

Simulation case: Base model				Simulation case: Covariate model			
SCM+	LASSO	GA 150	GA 50	SCM+	LASSO	GA 150	GA 50
Parsimonious covariate set							
1	4	1	3	0	5	2	3
Saturated covariate set							
1	5	3	5	1	4	2	6

→ **CCR assessment**
- $\beta_{par,cov} \neq 0$: all covariates modelling methods led to consistent CCR assessment and in line with those found with the reference model
  → BW on CL/F & V/F: **R**; AGE on CL/F & V/F & RACE on V/F: **IIS**; ALB on V/F: **NRS**
- $\beta_{par,cov} = 0$: none of covariates modelling methods conclude to the R of one of these relations
  - Full modelling approaches (i.e. FM & FREM): NR or with II to conclude
  - Covariate selection approaches: mainly NSEL and the few spurious relations retained NR or with II to conclude

Example of one forest plot of FREM applied on the dataset simulated under the Covariate model with the Saturated covariate set



Reference subject: white individual of 70 kg, 30 years old, with an albumin level at 45 g/L, an aspartate aminotransferase level at 23 U/L, a bilirubin level at 9 µmol/L and without FVIII inhibitor

OFV, BICc & runtime

	Simulation case: Base model							Simulation case: Covariate model						
	Ref model	FM	FREM*	SCM+	LASSO	GA 150	GA 50	Ref model	FM	FREM*	SCM+	LASSO	GA 150	GA 50
Parsimonious covariate set														
OFV	24 135	**24 110**	/	24 126	24 121	24 126	24126	24 056	**24 038**	/	24 056	24 039	24 044	24050
P_R	8	24	/	9	16	9	11	14	24	/	14	21	16	19
BICc	24 191	24 261	/	**24 188**	24 225	**24 188**	24 200	**24 148**	24 189	/	**24 148**	24 172	**24 148**	24 172
Runtime (h)	0.04	**0.44**	/	0.66	3.31	5.99	2.15	0.04	**0.44**	/	0.81	5.38	9.89	2.32
Saturated covariate set														
OFV	24 135	**24 103**	19 786*	24 126	24 119	24121	24114	24 056	**24 032**	19 715*	24 048	24 046	24 044	24042
P_R	8	33	90	9	17	11	15	14	33	90	15	20	16	22
BICc	24 191	24 308	20 330*	**24 188**	24 229	24 195	24 212	24 148	24 237	20 259*	**24 146**	24 173	24 148	24 181
Runtime (h)	0.04	**0.44**	2.05	1.55	4.51	5.99	2.09	0.04	**0.44**	2.18	2.09	6.33	10.73	3.71

*FREM OFV & BICc are not comparable with other methods due to differences regarding the dataset (covariates treated as observations), the way of coding covariate effects (as random-effects) and the Ω matrix (omega block for FREM vs diagonal matrix for the other methods)

→ **OFV & BICc**
- FM → best OFV in all cases, but a worse BIC due to the very large number of $\beta_{par,cov}$ to be estimated
- FREM → OFV and BICc not comparable with other methods, see *
- SCM+ → best BICc in all cases, with values lower than or equal to the ones of ref model
- GA 50 → results very close to the ones of SCM+ in terms of BICc (from 0 to 7 points difference)
- LASSO & GA 150 → slightly higher BICc (9 to 38 points compared to ref model) due to less parsimonious models selected

→ **Runtime**
- FM showed the best runtime; only SCM+ can compete with the runtime of the full modelling methods
- GA 50 & 150: high runtime due to the large number of models to run (max of 1500/7500 for GA50/150)
- LASSO: runtime depends mainly on the cross-validation number of iterations (10 models ran at each iteration)