

**ARTICLE**

A novel method based on unbiased correlations tests for covariate selection in nonlinear mixed effects models: The COSSAC approach

Géraldine Ayrat | Jean-François Si Abdallah | Claude Magnard | Jonathan Chauvin

Lixoft, Antony, France

Correspondence

Géraldine Ayrat, Lixoft, 8 rue de la renaissance, Bât D, 92160 Antony, France.

Email: geraldine.ayrat@lixoft.com

Abstract

Building a covariate model is a crucial task in population pharmacokinetics and pharmacodynamics in order to understand the determinants of the interindividual variability. Identifying a good covariate model usually requires many runs. Several procedures have been proposed in the past to automatize this task. The most commonly used is Stepwise Covariate Modeling (SCM). Here, we present a novel stepwise method based on statistical tests between individual parameters sampled from their conditional distribution and the covariates. This strategy, called the COnditional Sampling use for Stepwise Approach based on Correlation tests (COSSAC), makes use of the information contained in the current model to choose which parameter-covariate relationship to try next. This strategy greatly reduces the number of covariate models tested, while retaining on its search path the models improving the log-likelihood (LL). In this article, we detail the COSSAC method and its implementation in Monolix, and evaluate its performance. The performance was assessed by comparing COSSAC to the traditional SCM method on 17 representative data sets. For the large majority of cases (15 out of 17), the final covariate model is identical (11 cases) or very similar (4 cases with LL differences less than 3.84) with both procedures. Yet, COSSAC requires between 2 to 20 times fewer runs than SCM. This represents a decisive speed up, especially for models that take long to run and would not be tractable using the SCM method.

Study Highlights**WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?**

Current covariate search methods are very costly in computational time, difficult to implement, or rely on subjective preselection of covariates.

WHAT QUESTION DID THIS STUDY ADDRESS?

What are the principles of the COnditional Sampling use for Stepwise Approach based on Correlation tests (COSSAC) covariate search method and how does it compare to the standard Stepwise Covariate Modeling (SCM) method?

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 Lixoft. *CPT: Pharmacometrics & Systems Pharmacology* published by Wiley Periodicals LLC on behalf of American Society for Clinical Pharmacology and Therapeutics

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

The COSSAC method yields identical or very similar final covariate models as SCM but requires far fewer iterations.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?

The tremendous speedup of the COSSAC method allows to apply an automated and systematic procedure to computationally intensive models which would not be manageable with SCM.

INTRODUCTION

The population pharmacokinetic/pharmacodynamic (PK/PD) approach focuses on the characterization of the typical shape of the observations and of its interindividual variability. This variability can often be partly related to patient characteristics, called covariates. Identifying and quantifying the relationships between covariates and the model parameters allows to better understand the drug behavior and improve the accuracy of model predictions.

Typical data sets can have tens of covariates and PK/PD models a couple of parameters, which makes it impossible to test all possible covariate models. Thus, several procedures have been developed over the years to explore the space of covariate models.^{1,2}

The most common approach is Stepwise Covariate Modeling (SCM).³ SCM is a stepwise procedure, with a forward selection followed by backward elimination. In the forward selection, each possible covariate addition is tested in turn, the one improving the likelihood the most is kept, and the addition of a second covariate is tested, etc. Although the simplicity of the method is appealing, it is computationally expensive as it is prone to combinatorial explosion.

To reduce the number of tested models, several methods have been proposed⁴⁻⁷ (Appendix S1). Among those, we find the generalized additive model (GAM) method⁸ particularly interesting. It uses the empirical Bayes estimates (EBEs) obtained in the base run to identify a candidate covariate model. A linear regression is built iteratively between the EBEs and the covariates. This linear regression does not require running the population model and is computationally fast. The candidate model is then implemented in the population PK/PD model and run. The main drawback of GAM is that it relies on the EBEs. The informativeness of the EBEs is poor when the individual data are sparse, because the EBEs shrink toward the population value.⁹ In this situation, inferences based on EBEs can be misleading.

To circumvent the drawbacks of the EBEs, the use of random samples from the conditional distribution of each individual has been suggested.¹⁰ For each individual, a sample represents a plausible value for the individual parameters, although not the most probable one. These samples are spread over the entire marginal distribution of individual parameters and do not suffer from shrinkage. They can be used to

perform unbiased statistical tests for which the type I error is correctly controlled.¹⁰

Our novel procedure called COnditional Sampling use for Stepwise Approach based on Correlation tests (COSSAC) exploits these reliable statistical tests to identify correlations between individual parameters and covariates. This information is used to choose which parameter-covariate relationship to add in the model, in a stepwise manner, with an alternance of forward and backward steps. In this article, we present the COSSAC procedure and compare it to the standard SCM procedure on a representative set of examples. COSSAC is implemented in Monolix, together with the SCM, covSAMBA-COSSAC,¹¹ and SAMBA¹² approaches.

METHODS

The COSSAC procedure relies on the correlations between the covariates and the individual parameters sampled from the conditional distributions. We first explain how these correlations are calculated, then describe the COSSAC procedure step-by-step and finally explicit its implementation in Monolix.

Correlation tests between covariates and individual parameters

Conditional distribution sampling

Correlation tests between the conditional expectation of the parameters and the covariates are unbiased. Thus, we estimate the conditional expectation by a Monte Carlo approach (Appendix S2). This is done by sampling several random effects (called replicates) from each individual conditional distribution and averaging them.

Forward step

For continuous covariates, a Pearson correlation test is performed between the replicates averaged random effects and the covariates. The test statistic is compared to a t-distribution with $N-2$ degrees of freedom with N the number of individuals. For

categorical covariates, a one-way analysis of variance (ANOVA) is performed on the averaged random effects. When the covariate has more than two categories, the ANOVA procedure tests if at least one is significantly different from the others. These tests are performed for each pair of random effects and covariates, leading to one p value for each. It indicates the strength of the correlation between the covariate and a parameter random effects.

Backward step

Once covariates are included in the model, we would like to detect covariates bringing redundant information, which are not needed in the model. For this, pairwise correlation tests, as presented above, are not appropriate (Appendix S1). Instead, we perform a linear regression between the covariates and the parameters and test whether the estimated β coefficients are significantly different from zero. A t -test is used, comparing the statistic $\frac{\hat{\beta}}{se(\hat{\beta})}$ to a t -distribution with $N - n_C - 1$ degrees of freedom, with n_C the number of coefficients and $se(\hat{\beta})$ the estimated standard error obtained by least squares estimation during the regression. The calculations are very fast and yield one beta coefficient and p value for each relationship with a continuous covariate, and one beta and p value per category (apart from the reference category) for categorical covariates. As above, samples from the conditional distribution are used and averaged over replicates. An example is provided in the Appendix S1.

The COSSAC procedure

COSSAC is an iterative stepwise procedure. It alternates between deletions of covariates (backward) and additions of

covariates (forward). The choice of the parameter-covariate relationship to test for addition or removal is based on the p values of the correlation tests. For covariate additions, we first try relationships with small p values indicating significant correlations between the covariates and the random effects. Covariates with p values above a threshold, which have very weak or no correlation with the random effects, are not tested at all. For covariate removals, we try to remove relationships with large p values, where the contribution of the relationship is uncertain. After model estimation, the addition or removal of a covariate is accepted according to a likelihood-based criterion, for instance, the likelihood ratio test (LRT). If addition or removal is accepted, this constitutes the new current run from which we will try to add and remove further covariates.

Below, we detail each step of the COSSAC procedure, which is also depicted in Figure 1. To ease the reading, we use the default threshold values and acceptance criterion of Monolix but these values can be modified. The run of a model corresponds to the estimation of the population parameters, the estimation of the log-likelihood ($-2LL$ abbreviated as LL), the sampling of individual parameters from the conditional distributions, and the calculation of the correlation tests. An iteration is generally composed of several runs and a new iteration starts once a run is accepted.

At each iteration:

For iteration 1, start with the backward step. For the subsequent iterations, start with the step opposite to the previous one.

Backward step:

- Correlation test p values for parameter-covariates relationships already included in the model, which are above 0.01, are ranked from largest to smallest.

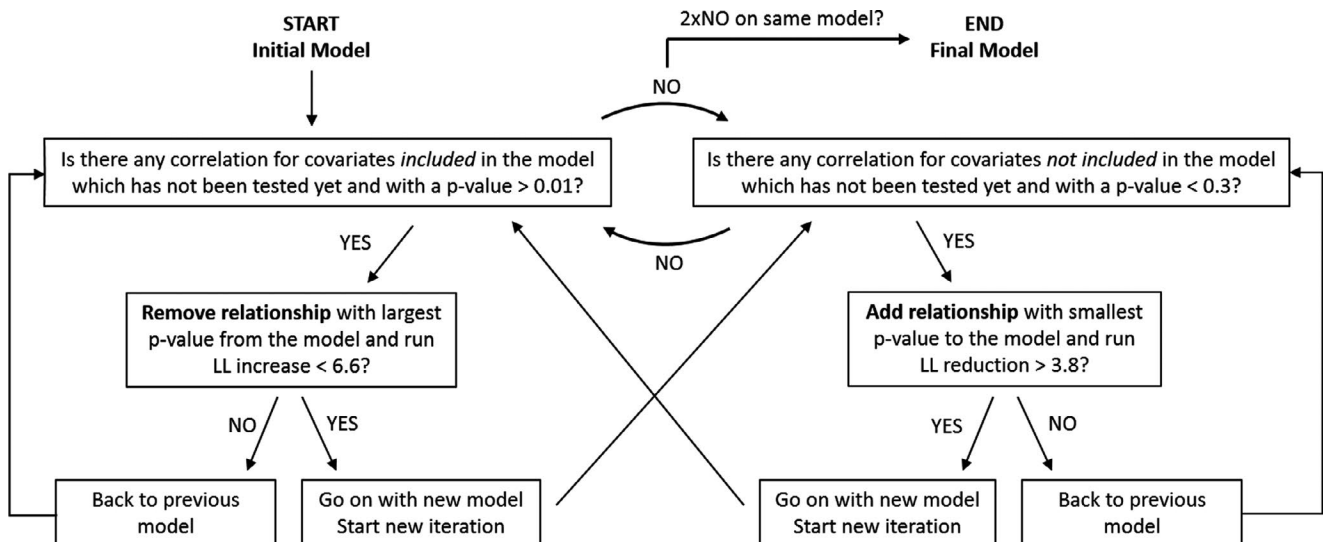


FIGURE 1 Scheme of the COnditional Sampling use for Stepwise Approach based on Correlation tests (COSSAC) covariate model building procedure. LL stands for -2 times the log-likelihood

- For each possible relationship starting with the one with the largest p value:
 - Remove the relationship and run the corresponding model.
 - If the increase of the LL is smaller than 6.635 (corresponding to a p value of the LRT of 0.01), accept the removal of the relationship, and start a new iteration with this run.
 - Otherwise, discard the run and test the next relationship in p value order.
 - If no relationship with correlation test p value above 0.01 remains to be tested, move to the forward step. If all forward and backward candidates have been tested, stop.

Forward step:

- Correlation test p values for parameter-covariate relationships not yet in the model, which are below 0.3, are ranked from smallest to largest.
- For each possible relationship starting with the one with the smallest p value:
 - Add the relationship and run the corresponding model.
 - If the reduction of the LL is larger than 3.841 (corresponding to a p value of the LRT smaller than 0.05), accept the addition of the relationship, and start a new iteration with this run.
 - Otherwise, discard the run and test the next relationship in p value order.
 - If no relationship with correlation test p value below 0.3 remains to be tested, move to backward step. If all forward and backward candidates have been tested, stop.

When the addition or removal of a covariate leads to a model which has already run, the model is not re-run and the likelihood value is read directly.

In the ensemble of runs of a COSSAC procedure, most of the runs come from a forward step. Most forward candidates are accepted, except at the end of the COSSAC procedure where remaining weaker correlations are also tested. The alternance of forward and backward steps avoids being trapped in a suboptimal path.^{1,13,14}

Implementation in Monolix

We here describe the implementation of COSSAC in Monolix, which can be launched from the GUI or via the `lixoftConnectors` R package. Note that the implementation in the R package `Rsmxlx` is a less efficient beta version of the procedure. The sampling from the conditional distribution is

described in the Appendix S1, as well as additional implementation details.

In the “model building” tab of Monolix (version 2019 and above), the SCM and COSSAC procedures are available to automatically build the covariate model. The settings include:

- *Method*: COSSAC or SCM (among others).
- *Parameter-covariate relationships to test*: The user can select the covariates and the parameters to test. In addition, each parameter-covariate relationship can also be set as “to test,” “always included,” or “always excluded.”
- *Criterion for acceptance*: LRT or corrected Bayesian Information Criteria (BICc, asymptotically consistent for nonlinear mixed effect models¹⁵).
- *Thresholds for LRT*: Thresholds on the LRT p values, one for the forward step (default 0.05, corresponding to a 3.841 points difference in the LL) and one for the backward step (default 0.01, corresponding to a 6.635 points difference).
- *Thresholds for correlations*: Thresholds on the correlation test p values, one for the forward step (default 0.3) and one for backward (default 0.01).
- *Method to calculate the likelihood*: Via linearization (default) or importance sampling.

For parameters without random effects, it is not possible to calculate the correlations between covariates and random effects. Thus, it is not possible to apply the COSSAC procedure. In Monolix, parameters without random effects are tested using SCM, once the COSSAC procedure on the parameters with random effects has finished.

COSSAC versus SCM

In COSSAC, relationships which are the most likely to improve the likelihood (as indicated by the correlation tests) are tested first. In addition, as soon as a relationship improves the likelihood sufficiently (according to the criterion) it is accepted, and additional relationships are tested on top. On the opposite, with SCM, all possible relationships are tested and the one improving the likelihood the most is taken as the basis for the next addition. Furthermore, SCM proceeds with forward steps first followed by backward steps, whereas COSSAC switches between forward and backward frequently.

RESULTS

Step-by-step example

To illustrate the functioning of the COSSAC algorithm, we first present the procedure step-by-step on a small data set of

warfarin PK with few covariates.¹⁶ For comparison, the SCM procedure is also run on the same data set.

The data set contains plasma concentration measurements for 32 individuals after a single oral dose of 1.5 mg/kg body weight. There are 6 to 13 observations per individual, 247 in total. Three covariates have been recorded: age, sex, and weight. A one-compartment model with delayed first-order absorption and linear elimination (parameters lag time of absorption [Tlag], absorption rate constant [ka], volume [V], and clearance [Cl]) properly captures the data. All parameters have random effects. This constitutes the base run, the starting point for the covariate model building.

The continuous covariates age and weight (WT) are log-transformed and centered. Only the log-transformed counterparts, which correspond to power-law relationships (Appendix S2), are tested. All covariates (logAge, logWT, and sex) are tested on all parameters (Tlag, ka, V, and Cl). The default Monolix settings are used, except that the LRT threshold of the backward steps is set to 0.05 (same as forward) to mirror the SCM implementation in PsN¹⁷ (Perl modules to aid Nonmem usage). Files to reproduce this example are provided (Appendix S2).

COSSAC procedure

All information used during the procedure is depicted in Figure 2 and a scheme of the run series is provided in Figure 3a.

Initialization

- Run 1 (base model): LL = 653.7

Iteration 1 – 1 run

For the first run, we start with the backward step. As the model does not include covariates, the backward step is skipped, and we proceed to the forward step. The p values of the correlation tests between the random effects and covariates are shown in Figure 2. Four relationships' p values are below the 0.3 threshold and logWT on the volume V has the smallest p value. This relationship is thus added in the model for the next run.

- Run 2 with logWT on V: LL = 627.5.

This model leads to an LL of 627.5, which corresponds to a decrease of 26.2 points, larger than the 3.841 LRT threshold. The model is accepted, and a new iteration starts with this run as a basis.

Iteration 2 – 1 run

As the previous step was forward, we start analyzing run 2 with the backward step. The p value for beta_V_logWT, the only relationship included in the model is 2.45e-8, below the 0.01 threshold, so we do not try to remove it.

We proceed to the forward step of run 2. Two relationships are below the 0.3 threshold and logWT on the Cl has the smallest p value. This relationship is thus added in the model for the next run.

- Run 3 with logWT on V, logWT on Cl: LL = 622.4.

Run 3 corresponds to a decrease in LL of 5.1 points, larger than the 3.841 LRT threshold. The logWT on Cl relationship is accepted and a new iteration starts.

Iteration 3 – 1 run

We start with the backward check of run 3. One p value for the correlations of relationships already included in the model is above 0.01 threshold, that for logWT on Cl. Removing logWT on Cl from run 3 actually corresponds to run 2, which has already run and was shown to be worse than run 3.

No untested candidates remain for the backward step, so we proceed to the forward step. One relationship is below the 0.3 threshold: logAGE on Cl. This relationship is thus added in the model for the next run.

- Run 4 with logWT on V, logWT on Cl, logAGE on Cl: LL = 618.1.

This model leads to an LL of 618.1, which corresponds to an improvement of 4.3 points, larger than the 3.841 LRT threshold. The logAGE on Cl relationship is accepted.

Iteration 4 – 3 runs

As the previous step was forward, we start with the backward step on run 4. Two p values for the correlations of relationships already included in the model are above the 0.01 threshold, that for logAGE on Cl and for logWT on Cl. We start with the larger one, logAGE on Cl. Removing it would correspond to run 3, which was shown to be worse than run 4. We thus try to remove the second candidate, logWT on Cl, as the next run.

- Run 5 with logWT on V, logAGE on Cl: LL = 623.9.

Compared with run 4, the LL has increased by 5.8 points, which corresponds to a significant worsening,

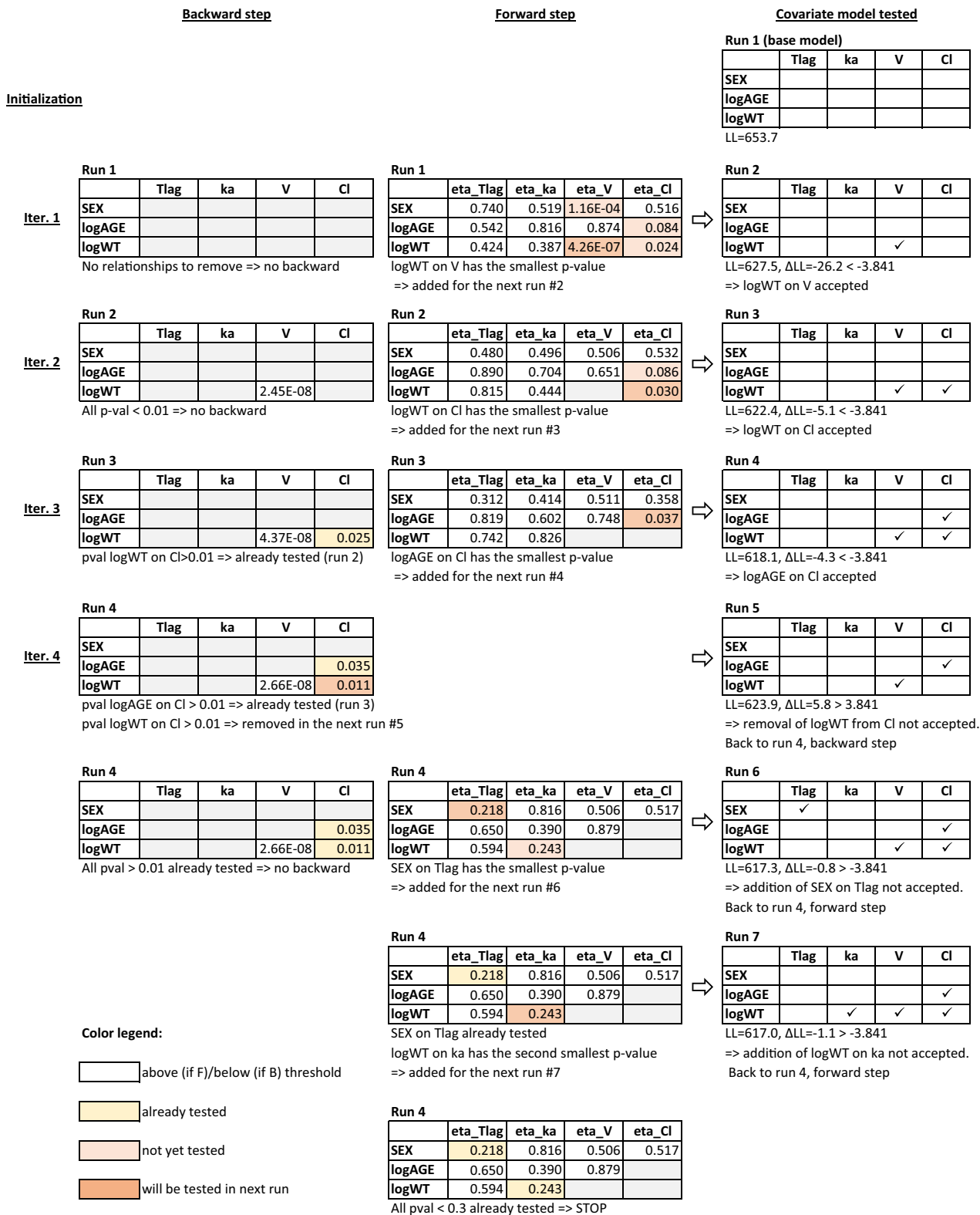
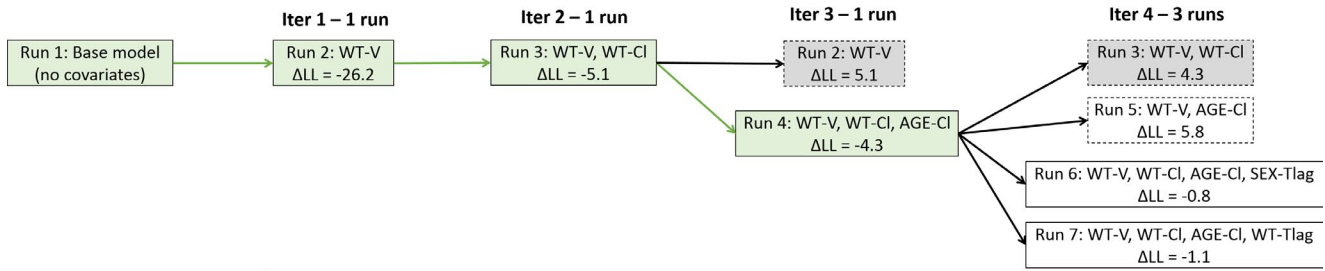


FIGURE 2 Step-by-step Conditional Sampling use for Stepwise Approach based on Correlation tests (COSSAC) procedure on the warfarin example with three covariates (log weight [logWT], logAGE, and SEX) and a four-parameter model (lag time of absorption [Tag], absorption rate constant [ka], volume [V], and clearance [Cl]). At each iteration, covariates are checked for removal (backward step), and covariates are checked for addition (forward step). At each new run, the log-likelihood (LL) change is assessed. If a run is accepted, a new iteration starts. The correlation test *p* value thresholds are 0.3 in forward and 0.01 in backward. The likelihood ratio test *p* value threshold is 0.05 in both directions, corresponding to a difference of 3.841 LL points. Correlation test *p* values below the forward threshold or above the backward threshold are colored in yellow if it has already been tested, in light orange if it has not yet been tested, and in dark orange if it will be tested in the next run

(a) COSSAC procedure



(b) SCM procedure

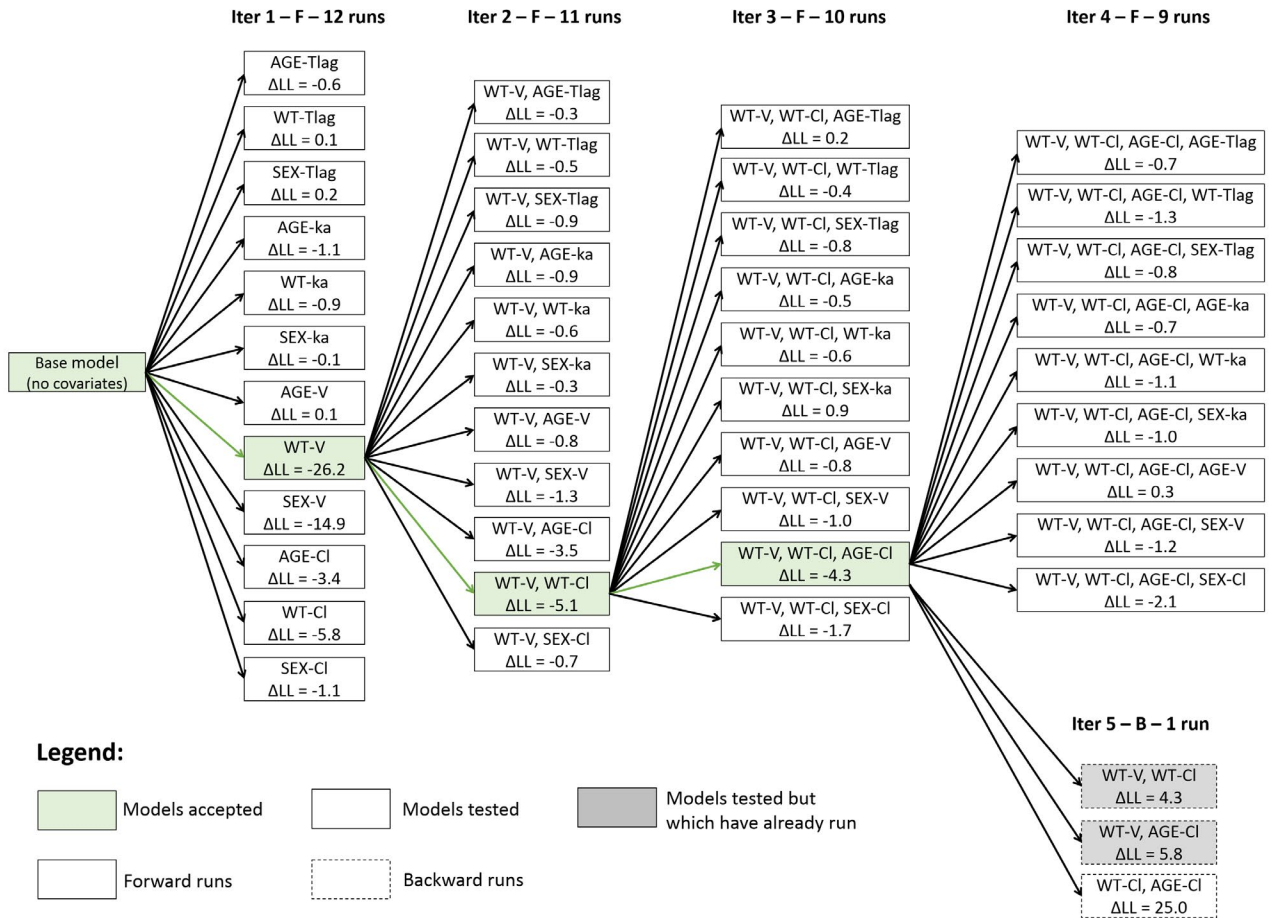


FIGURE 3 Tree-view of the runs performed on the warfarin example with the COSSAC procedure (a) and the Stepwise Covariate Modeling (SCM) procedure (b). Selected models are colored in green, tested models in white, and models tested but which have already run in grey

above the 3.841 threshold. The removal of logWT from CI is not accepted and we go back to run 4 (which has already run). In run 4, all possible backward candidates have been tested.

We proceed to the forward step. Two relationships are below the 0.3 threshold: SEX on Tlag and logWT on ka. We start with the smallest *p* value and add SEX on Tlag for the next run.

- Run 6 logWT on V, logWT on CI, logAGE on CI, SEX on Tlag: LL = 617.3.

This model leads to an LL of 617.3, which corresponds to an improvement of 0.8 point, smaller than the 3.841 LRT threshold. The SEX on Tlag relationship is not accepted and we go back to run 4. In run 4, the second relationship to test is logWT on ka, it is added for the next run.

- Run 7 logWT on V, logWT on CI, logAGE on CI, logWT on ka: LL = 617.0.

The LL improvement is only 1.1 points, not enough to accept logWT on ka. There are no remaining relationships

with correlation p value below 0.3. All backward and forward candidates have been tested, so the COSSAC procedure stops. The final model is run 4.

SCM procedure

The SCM procedure is described in the Appendix S2 and depicted in Figure 3b. In brief, during the first iteration, all 12 possible models are run and logWT on V is accepted, then the remaining 11 candidates are run and logWT on Cl is accepted, etc.

Comparison

Both procedures lead to the same final covariate model and have taken the same path of accepted models. The COSSAC procedure requires 6 runs (base run excluded) whereas 43 are needed with SCM.

Performance evaluation of COSSAC versus SCM

To assess the performance of the COSSAC procedure compared to the SCM procedure, we have applied both routines to a collection of 17 representative datasets. They comprise PK, PK/PD, and disease models. Most of them are continuous data but we have also included one time-to-event data set and one count data example. Some data sets are dense, whereas others are sparse. The number of parameters vary from two to nine, whereas the number of covariates vary from two to eight. Many examples contain covariates that are strongly correlated with each other. An overview of the data set characteristics is given in Table 1 and Monolix projects and settings are provided in the Appendix S2.

All data sets are real (not simulated) for which the “true” covariate model is unknown. The COSSAC and SCM procedures are compared in terms of number of runs, final covariate model obtained, final LL, and BICc.¹⁵ The run duration varies only very slightly with the number of covariate effects to estimate (Appendix S1). The total computational time of the covariate search procedure is thus proportional to the number of runs.

All results are presented in Table 1. The first observation is that the COSSAC procedure completes in much fewer runs than SCM. The reduction in the number of runs varies from twice fewer for models with few possible relationships to more than 20 times fewer for realistic data sets with many covariates, such as the data set provided by the ASCPT Grand Prix modeling challenge. An approximate formula for the difference in the number of runs can be derived (Appendix S1).

The difference grows linearly with the number of parameters and the number of covariates, and decreases with the square of the number of relationships.

Out of the 17 data sets, 11 of them result in the exact same final covariate model after COSSAC and after SCM. They correspond to data sets with few relationships in the final model. Four further datasets lead to slightly different models, but for which the difference in LL is small, less than 3.84 points.

For one run (remifentanyl seqPD), the SCM method finds a model with one additional relationship (logAGE on E0) compared to COSSAC, which leads to an 8.4 points better LL. This better model is not tested by COSSAC because logAGE on E0 improves the LL only once covariates have been added on the gamma parameter, which has no variability and is tested after all others only. Running COSSAC again at the end would resolve the discrepancy but comes at a substantial cost in terms of runs. A similar situation happens for the Verapamil PK example.

On the opposite, for the model-informed drug development (MIDD) dataset, COSSAC finds a model that is 40 points better than SCM. The path of accepted runs taken by both methods is the same for the four first covariate additions. For the fifth, SCM adds nDiseases on the first order degradation rate (k_{deg}) (largest LL decrease) and no further addition leads to a sufficient LL improvement. COSSAC adds logWT on Clr (smallest correlation p value). The LL improvement of this addition is smaller than that of nDiseases on k_{deg} , but this turns to be an advantage afterward, as the end-stage renal disease (ESRD) on Cl and logALB on renal clearance (Clr) can be added as additional significant covariates.

DISCUSSION

This paper presents a novel covariate model building procedure, which offers many advantages. First, it is systematic and does not depend on a subjective preselection of covariates. Second, its implementation is relatively easy and, contrary to the lasso or FREM, does not require to modify the covariate encoding. Finally, it requires only a limited number of runs, much fewer than the widely used SCM method.

One of the key characteristics of COSSAC is the use of correlation tests between the individual parameters sampled from the conditional distribution and the covariates. Importantly, these correlation tests are fast to calculate and necessitate only the base (or current) run. In addition, these tests are not subject to shrinkage bias and are reliable, because they use samples from the conditional distributions rather than the modes (EBEs).¹⁰ In COSSAC, the correlation test p -values are used to select which parameter-covariate relationship will be added and evaluated first. This is very efficient because the p -values are good predictors

TABLE 1 Comparison of the COSSAC and SCM results on 17 representative data sets

Data set	Characteristics	Parameters (<i>italic</i> : <i>no variability</i>)	Covariates	COSSAC		SCM		ΔLL	ΔBIC_c	Ratio # runs
				No. runs	Final model	No. runs	Final model			
Remifentanyl PK	Linear PK - SD - dense 65 indiv - 1992 obs	6 - CI, V1, Q2, V2, Q3, V3	6 - SEX, logAGE, logBSA, logHT, logLBM, logWT	13	SEX - V3 logAGE - CI, Q2, V2, V3 logBSA - CI logLBM - V1, V2	295	logAGE - CI, Q2, Q3, V2, V3 logBSA - CI logHT - V2 logLBM - V1	-3.8	0.4	22.7
Theophylline PK	Linear PK - SD - dense 12 indiv - 120 obs	3 - ka, V, CI	2 - SEX, logWT	6	None	7	None	Identical		1.2
Verapamil PK	Linear PK - SD - dense 22 indiv - 330 obs	6 - Tlag, ka, CI, V1, Q, V2	7 - SEX, RACE, logAGE, logHT, logWT, logDIABP, logSYSBP	34	SEX - CI, V1, ka logAGE - ka logWT - Q, V2	241	SEX - CI, ka logAGE - ka logWT - Q, V2	-2.6	0.5	7.1
GBR12909 PK	Linear PK - MD - dense 12 indiv - 232 obs	5 - ka, V, k, k12, k21	2 - SEX, logWT	5	logWT - ka	20	logWT - ka	Identical		4
Quinidine PK	Linear PK - SD - dense 21 indiv - 315 obs	6 - Tlag, ka, CI, V1, Q, V2	7 - SEX, RACE, logAGE, logHT, logWT, logDIABP, logSYSBP	20	SEX - CI, V1	124	SEX - CI, V1	Identical		6.2
Quinidine sparse PK	Linear PK - MD - sparse 136 indiv - 361 obs	3 - ka, V, CI	7 - RACE, HEART, ETHANOL, SMOKE, logAGE, logHT, logWT	11	None	22	None	Identical		2
Tobramycin sp. PK	Linear PK - MD - sparse 97 indiv - 322 obs	2 - V, CI	4 - SEX, logAGE, logCLCR, logWT	7	logCLCR - CI logWT - V	22	logCLCR - CI logWT - V	Identical		3.1
Cisplatin PK	Linear PK - MD - dense 23 indiv - 524 obs	6 - CI, V1, Q2, V2, Q3, V3	5 - SEX, logAGE, logBSA, logHT, logWT	16	logBSA - V1	60	logBSA - V1	Identical		3.8
Theophylline ER PK	Linear PK - SD - dense 18 indiv - 362 obs	7 - ka1, ka2, F1, Tlag1, diffTlag2, V, CI	3 - logAGE, logHT, logWT	17	logWT - Tlag1, V	61	logAGE - ka2 logWT - Tlag1	0.5	0.5	3.6
IgG1 mAb PK	TMDD PK - SD - dense 28 indiv - 263 obs	7 - V, <i>k_{int}</i> , kon, R0, CI, Q, V2	2 - RA, logWT	13	RA - CI, V, kon logWT - V	63	RA - CI, V, kon logWT - V	Identical		4.8
Remifentanyl seqPD	PD - SD - dense 61 indiv - 3989 obs	5 - ke0, E0, I _{max} , IC ₅₀ , <i>gam</i> (indiv. PK param fixed)	6 - SEX, logAGE, logBSA, logHT, logLBM, logWT	29	SEX - gam logAGE - IC ₅₀ , gam, ke0 logHT - gam	194	SEX - gam logAGE - E0, IC ₅₀ , gam, ke0 logHT - gam	8.4	4.3	6.7

(Continues)

TABLE 1 (Continued)

Data set	Characteristics	Parameters (<i>italic</i> : <i>no variability</i>)	COSSAC		SCM		Ratio # runs		
			No. runs	Final model	No. runs	Final model		ΔLL	$\Delta BICc$
Dofetilide PK/PD	Joint PK/PD - SD - dense 22 indiv - 328x2 obs	8 - Tlag, ka, Cl, V1, Q, V2, <i>intercept</i> , slope logHT, logWT, logDIABP, logSYSBP	60	RACE - <i>intercept</i> logWT - V1	220	RACE - <i>intercept</i> logSYSBP - <i>intercept</i> logWT - V1	0.5	0.2	3.7
MIDD (ASCPT Gran Prix)	Joint model parent/metab/urine/PD 176 indiv - 2664 + 2723 + 147 + 2600 obs	9 - ka, V, Cl, Clr, Clm, Vm, R0, kdeg, IC ₅₀	20	ESRD - Cl logAGE - Cl logALB - Cl, Clr, V, ka logWT - Clr	421	logAGE - Cl logALB - Cl, V, ka nDiseases - k _{deg}	-40	-29	21.1
Warfarin PK/PD	Joint PK/PD - SD - dense 32 indiv - 247 + 232 obs	8 - Tlag, ka, V, Cl, R0, kout, I _{max} , IC ₅₀	11	logWT - V	48	logWT - V	Identical		4.4
Cholesterol	Disease progression 200 indiv - 1044 obs	2 - Chol0, slope	5	logAGE - Chol0, slope SEX - slope	12	logAGE - Chol0, slope SEX - slope	Identical		2.4
Alzheimer	Disease - count data 896 indiv - 3707 obs	2 - p0, slope	8	APOE - alpha, p0 logAGE, logBMI, logHT, logWT	82	APOE - alpha, p0 logAGE - alpha, p0 logBMI - alpha logWT - p0	Identical		10.3
Lung cancer survival	Time-to-event 228 indiv - 165 events	2 - Te, k	13	AGE - k ecogPH - Te SEX - Te	36	AGE - k ecogPH - Te SEX - Te	identical		2.8

Note: The ratio number of runs is defined as the number of runs for SCM divided by the number of runs for COSSAC. Differences in the covariate models are highlighted in bold. $\Delta BICc$, $BICc(\text{cossac}) - BICc(\text{scm})$; ΔLL , $LL(\text{cossac}) - LL(\text{scm})$; ALB, albumin; APOE, apolipoprotein E genotype (0/1/2); $BICc$, corrected BIC¹⁵; BMI, body mass index; BSA, body surface area; COSSAC, Conditional Sampling use for Stepwise Approach based on Correlation tests; Cl, clearance; CLCR, creatinine clearance; DIABP, diastolic blood pressure; ecogPH, Eastern Cooperative Oncology Group (ECOG) performance status by physician; ER, extended release; ESRD, end-stage renal disease (yes/no); HEART, congestive heart failure (mild/moderate/severe); ETHANOL, alcohol abuse (none/former/current); HT, height; IC₅₀, half-maximal inhibitory concentration; IgG1, immunoglobulin G1; I_{max}, maximal inhibition; indiv., individual; ka, absorption rate constant; kamoPAT, Kamofsky performance status by patient; kamoPH, Kamofsky performance status by physician; LBM, lean body mass; LL, -2 times the log-likelihood; mAb, monoclonal antibody; MD, multiple doses; MIDD, model-informed drug development; obs., observed; PD, pharmacodynamic; PK, pharmacokinetic; RA, rheumatoid arthritis patient (yes/no); SCM, Stepwise Covariate Modeling; SD, single dose; SMOKE, smoking status (yes/no); SYSBP, systolic blood pressure; TMDD, target-mediated drug disposition; V, volume; WT, weight.

of the LL improvement obtained when the corresponding relationship is added to the model. Note that ranking the p -values is equivalent to ranking the correlation coefficients (Appendix S1). Tested relationships are thus accepted most of the time. When the p values are sufficiently different from each other, their order reflects the order of LL improvement obtained when adding the covariates in a univariate manner. In this case, COSSAC and SCM are taking the same path of accepted models, except that COSSAC tries only the best relationship, whereas SCM tests all possible relationships. Thus, in the MIDD example, COSSAC adds the first four covariates in four runs. SCM adds the same 4 covariates but requires 282 runs. When the two smallest p -values are close, their order can by chance be the same or opposite than that of the LL improvement. This is, for instance, the case when two covariates are strongly correlated with each other. In this case, the paths taken by SCM and COSSAC can be different but with very similar LL improvements. At a given step, SCM will always select the addition leading to the largest LL improvement whereas COSSAC may not. This may (or may not) lead to a better final model, as in the MIDD example.

We have tested the COSSAC procedure on all nonconfidential data sets in our hands. Among the 17 data sets, the majority leads to the same final covariate model proposed by COSSAC and SCM (11 cases) or to very similar models with almost the same LL and BIC¹⁵ (4 cases). Yet, COSSAC requires, on average, seven times fewer runs to complete. The gain in number of runs is especially large when there are many possible relationships (i.e., number of parameters times number of covariates) but only a few are retained in the final model. On the opposite, parameters without random effects cannot be assessed via COSSAC. In the Monolix implementation, SCM is used for them. Therefore, models having parameters with and without variability show only an intermediate speedup compared to SCM. To avoid the costly SCM on parameters without variability, one option is to consider a small fixed variability (for instance 5%) on these parameters.

As SCM, COSSAC is a stepwise procedure, which repeatedly applies a likelihood ratio test to assess the benefits of the added covariates. In general, stepwise procedures are known to be prone to selection bias (i.e., overestimation of the effects of the selected covariates). In the population PK/PD field, some authors have reported a high selection bias with SCM, especially for small data sets,¹⁸ whereas others have concluded it is only minor in a typical realistic covariate search setting.¹⁹ One can expect the selection bias to be similar with SCM and COSSAC. Its precise quantification deserves further investigation but is out of the scope of this article. We have also not attempted to evaluate type I and type II errors using simulated data sets. This will be addressed in a separate work.

The large speedup in computation time offered by COSSAC compared to the standard SCM method makes it an

appealing method, especially for complex models with long run time or data sets with many covariates. We believe that its efficient implementation in Monolix will contribute to its spreading in the community.

ACKNOWLEDGEMENTS

The authors are thankful to Pauline Traynard and Marc Lavielle for their proofreading of the manuscript and their suggestions.

CONFLICT OF INTEREST

All authors are employees of Lixoft. Lixoft licenses the MonolixSuite software.

AUTHOR CONTRIBUTIONS

G.A. and C.M. wrote the manuscript. G.A., J.-F.S.A. and J.C. designed the research. G.A., J.-F.S.A. and J.C. performed the research. G.A., J.-F.S.A., and J.C. analyzed the data.

REFERENCES

1. Hutmacher MM, Kowalski KG. Covariate selection in pharmacometric analyses: A review of methods. *Br J Clin Pharmacol*. 2015;79:132-147.
2. Mould DR, Upton RN. Basic concepts in population modeling, simulation, and model-based drug development-part 2: introduction to pharmacokinetic modeling methods. *CPT Pharmacometrics Syst Pharmacol*. 2013;2:e38.
3. Jonsson EN, Karlsson MO. Automated covariate model building within NONMEM. *Pharm Res*. 1998;15:1463-1468.
4. Kowalski KG, Hutmacher MM. Efficient screening of covariates in population models using Wald's approximation to the likelihood ratio test. *J Pharmacokinet Pharmacodyn*. 2001;28:253-275.
5. Gastonguay M. A full model estimation approach for covariate effects: Inference based on clinical importance and estimation precision. *AAPS J*. 2004;6:AbstractW4354.
6. Karlsson MO. A full model approach based on the covariance matrix of parameters and covariates. *PAGE Abstr Annu Meet Popul Approach Gr Eur*. 2012;21:Abstr 2455.
7. Ribbing J, Nyberg J, Caster O, Jonsson EN. The lasso - a novel method for predictive covariate model building in nonlinear mixed effects models. *J Pharmacokinet Pharmacodyn*. 2007;34:485-517.
8. Mandema JW, Verotta D, Sheiner LB. Building population pharmacokinetic pharmacodynamic models. I. Models for covariate effects. *J Pharmacokinet Biopharm*. 1992;20:511-528.
9. Nguyen THT, et al. Mouksassi M-S, Holford N, Model evaluation of continuous data pharmacometric models: Metrics and graphics. *CPT Pharmacometrics Syst Pharmacol*. 2017;6:87-109.
10. Lavielle M, Ribba B. Enhanced method for diagnosing pharmacometric models: random sampling from conditional distributions. *Pharm Res*. 2016;33(12):2979-2988.
11. covSAMBA-COSSAC. <http://monolix.lixoft.com/model-building/automatic-covariate-model-building/>.
12. SAMBA procedure. <http://monolix.lixoft.com/model-building/samba/>.
13. Efroymson MA. Multiple regression analysis. In *Mathematical Methods for Digital Computers*. Hoboken, NJ: John Wiley; 1960.

14. Derksen S, Keselman HJ. Backward, forward and stepwise automated subset selection algorithms: frequency of obtaining authentic and noise variables. *Br J Math Stat Psychol*. 1992;45:265-282.
15. Delattre M, Lavielle M, Poursat MA. A note on BIC in mixed-effects models. *Electron J Stat*. 2014;8:456-475.
16. O'Reilly RA, Aggeler PM. Studies on coumarin anticoagulant drugs. Initiation of warfarin therapy without a loading dose. *Circulation*. 1968;38:169-177.
17. Lindbom L, Pihlgren P, Jonsson N. PsN-Toolkit - A collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. *Comput Meth Programs Biomed*. 2005;79:241-257.
18. Ribbing J, Niclas Jonsson E. Power, selection bias and predictive performance of the population pharmacokinetic covariate model. *J Pharmacokinet Pharmacodyn*. 2004;31:109-134.
19. Wählby U, Jonsson EN, Karlsson MO. Comparison of stepwise covariate model building strategies in population

pharmacokinetic-pharmacodynamic analysis. *AAPS PharmSci*. 2002;4:E27.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: 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 Pharmacometrics Syst. Pharmacol*. 2021;10:318–329. <https://doi.org/10.1002/psp4.12612>