

Development of a population PK model using MONOLIX - Case Study I

September 9, 2019

Contents

1	Introduction	2
2	Data file and data exploration	2
3	Development of the structural model	4
3.1	Choosing and initializing a model	4
3.2	Running Monolix tasks.	7
3.3	Diagnostic plots	10
3.4	Convergence assessment	13
4	Development of the statistical model	14
4.1	Exploration of covariate relationships	14
4.2	Covariate model	16
4.3	Results and interpretation	17
4.4	Automatic Covariate Model Building	18
5	Final plots	20
6	Summary of results	20

1 Introduction

The aim of this tutorial is to develop a population PK model and explore covariate relationships using MonolixSuite applications of a hypothetical test drug used in the treatment of atrial fibrillation. The study was a double blind, parallel, single dose intravenous bolus trial, in which 100 patients were randomized to either of two doses namely: 100 mg or 250 mg. The data set contains concentration measurements from blood samples collected at 0.25,0.5, 0.75,1,1.5,2,2.5,3,4,6,8,10,12,16,20,24 hours post-dose, the dose regimen and covariates: age, weight, creatinine clearance and gender.

The workflow consists of the preparation of a data set, choosing a PK model, estimating parameters and analyzing goodness of fits, building a statistical model and generating results. In this case study it takes 2 minutes in the Monolix interface to perform the whole workflow - from loading the data set to displaying plots and results - and get an output ready for the basic analysis.

The following MonolixSuite applications are used:

- [Datxplorer](#) to visualize and explore the data to give the first insights into the structural model and covariates relationships.
- [Monolix](#) to develop models, estimate parameters and generate statistical tests and diagnostic plots.
- [Sycomore](#) to manage the modeling workflow and compare models.

2 Data file and data exploration

Monolix accepts only one header line and data in the long format: each variable is in a separate column and each line corresponds to one individual and one time point. The header, the first line of the file, contains in each column the names of the variables. The choice of the columns names is free.

A sample of a csv file, compatible with all applications of the MonolixSuite, is shown in Figure 1, and Table 1 contains a complete description of all variables of the data set under study. All column types available in the MonolixSuite are described in the [data set documentation](#).

1	CID	TIME	CONC	AMT	DOSE	MDV	AGE	WT	SCR	ISM	CLCR
2	1	0	0	100	100	1	34.823	38.212	1.1129	0	42.635
3	1	0.25	13.026	0	100	0	34.823	38.212	1.1129	0	42.635
4	1	0.5	14.984	0	100	0	34.823	38.212	1.1129	0	42.635
5	1	0.75	14.16	0	100	0	34.823	38.212	1.1129	0	42.635
6	1	1	19.316	0	100	0	34.823	38.212	1.1129	0	42.635
7	1	1.5	13.146	0	100	0	34.823	38.212	1.1129	0	42.635
8	1	2	12.921	0	100	0	34.823	38.212	1.1129	0	42.635
9	1	2.5	8.485	0	100	0	34.823	38.212	1.1129	0	42.635
10	1	3	16.437	0	100	0	34.823	38.212	1.1129	0	42.635
11	1	4	10.724	0	100	0	34.823	38.212	1.1129	0	42.635
12	1	6	8.7352	0	100	0	34.823	38.212	1.1129	0	42.635
13	1	8	7.697	0	100	0	34.823	38.212	1.1129	0	42.635
14	1	12	4.479	0	100	0	34.823	38.212	1.1129	0	42.635
15	1	16	2.4183	0	100	0	34.823	38.212	1.1129	0	42.635
16	1	20	4.7586	0	100	0	34.823	38.212	1.1129	0	42.635
17	1	24	2.3655	0	100	0	34.823	38.212	1.1129	0	42.635
18	2	0	0	100	100	1	32.765	74.838	0.8846	1	126
19	2	0.25	7.4203	0	100	0	32.765	74.838	0.8846	1	126
20	2	0.5	4.9869	0	100	0	32.765	74.838	0.8846	1	126
21	2	0.75	13.63	0	100	0	32.765	74.838	0.8846	1	126

Figure 1: A sample of a csv file with concentration measurements, dose regimen and covariates compatible with MonolixSuite applications.

Column name	Description	Requirement
CID	identifier of the individual tagged as ID	mandatory
TIME	time of a dose administration or observation record tagged as TIME	mandatory
CONC	observation for continuous data (drug concentration) tagged as OBSERVATION	mandatory
AMT	dose amount at dosing time (zero for observation records) tagged as AMOUNT	
DOSE	dose received by a patient (for data visualization) tagged as CATEGORICAL COVARIATE	
MDV	identifier to ignore an observation information tagged as IGNORED OBSERVATION	
AGE, WT, CLCR, SCR	AGE - age in yrs, WT - weights in kg, CLCR - creatinine clearance in ml/min, SCR - serum creatinine in mg/dL tagged as CONTINUOUS COVARIATES	
ISM	ISM - is male (=1), female (=0) tagged as CATEGORICAL COVARIATE	

Table 1: Description of the data set variables. Full description is available in the [data set documentation](#).

Data file exploration starts with opening Datxplre, choosing a "New project" and loading the csv file. Because the columns names are free, the user must indicate to the application their meanings. In this example, most of the columns are automatically recognized, but the remaining ones have to be tagged. Then the data set is accepted by clicking the button ACCEPT, see Figure 2 for a Datxplre interface screenshot.

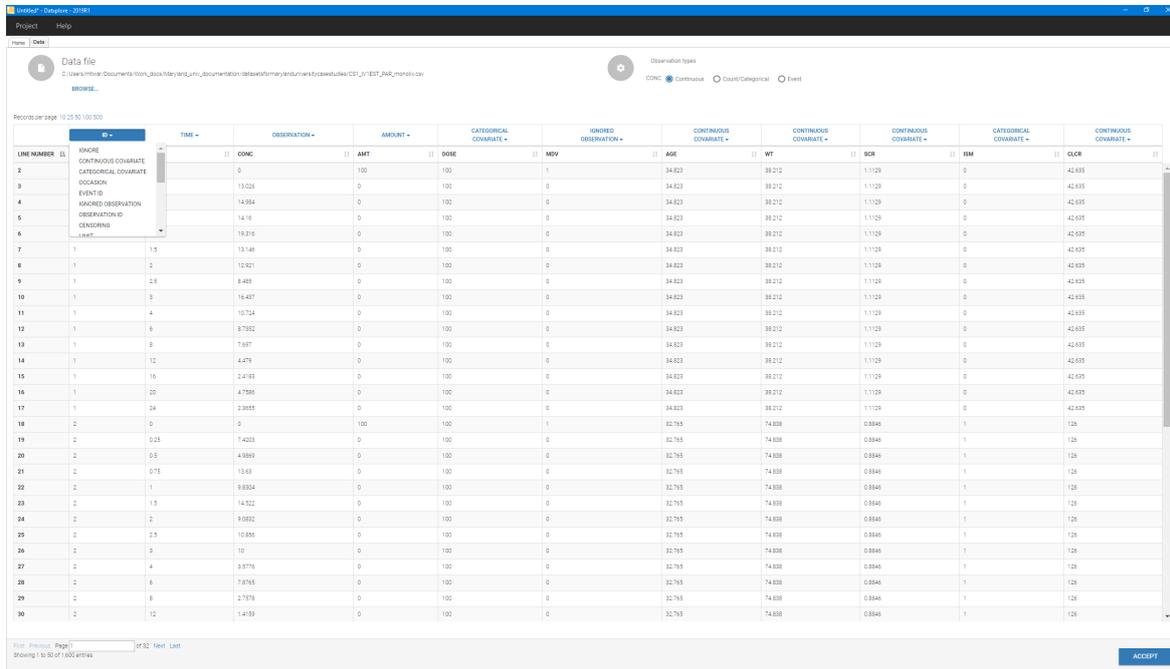


Figure 2: Preparing the data set in Datxplre: tagging columns (variable types) and accepting the data by clicking the button ACCEPT.

Datxplre allows for a fast and easy preview of the data to help in the identification of a first structural model, influential covariates and correlations between covariates. Figure 3 (on the left) presents semi-log plot of the time evolution of the concentrations. Constant elimination slope suggests a one-compartment model, while stratification by ceratinine clearance explains part of the variability

in concentration levels. The covariate viewer of Datxplore (on the right) allows to see possible co-linearities between covariates, such as for example between weight and creatinine clearance.

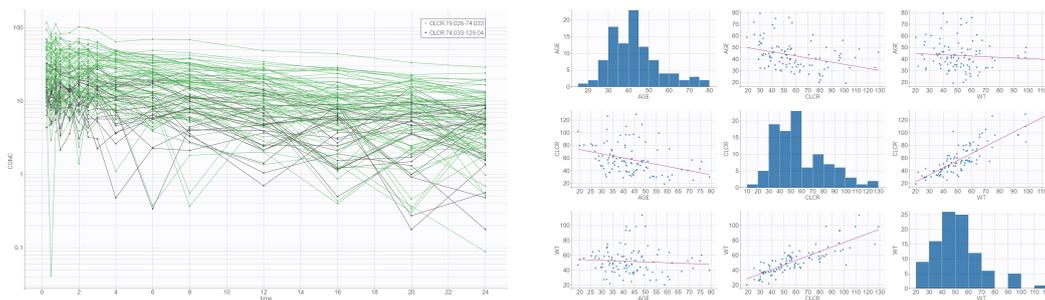


Figure 3: Datxplore viewers: time evolution of concentration stratified by ceratinine clearance (on the left) and covariates relationships (on the right).

The data set visualization can be saved as a new datxplore project, and the plots can be saved as images. The data set information is then exported directly to Monolix from the top panel Export > Export to Monolix. You will see the Monolix application opening with all required data information already setup.

3 Development of the structural model

This section describes how to build a model for this data set using the Monolix model library, how to run different tasks to estimate parameters, analyze the model and interpret the results.

3.1 Choosing and initializing a model

Choosing a model from the Monolix PK library

In Monolix, the structural model is defined using the [Mlxtran language](#) as a separate text file. In this case study the model is loaded from the Monolix library, see Figure 4. Many typical models are available by clicking LOAD FROM LIBRARY in the "Structural model" tab: PK, PD, PKPD, PK Double Absorption, TMDD, TTE and Count.

PK library is selected and the top panel allows to filter the list of models, see Figure 5. After choosing bolus as administration method, 1 compartment distribution - according to the data visualization, the drug seems to follow a one-compartment model behaviour - and a linear elimination, two parametrizations are available: (V,Cl) and a (V,k). File bolus_1cpt_VCl.txt is selected.

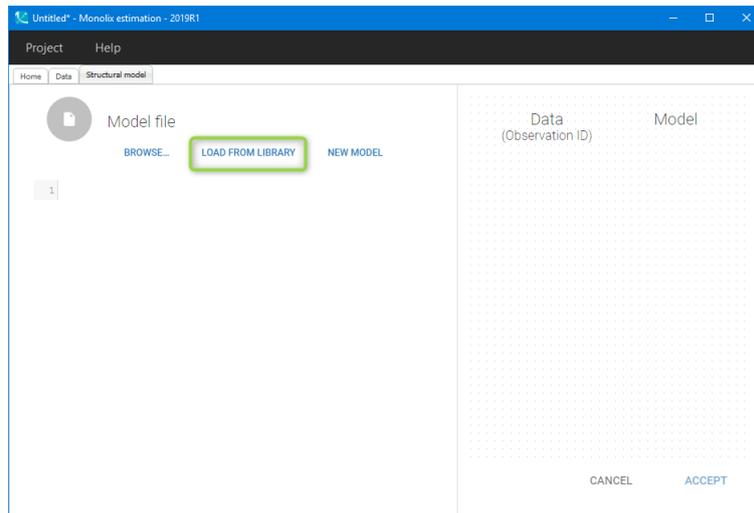


Figure 4: Options in Monolix to define a structural model. BROWSE - load a txt file with model written in the Mlxtran language; LOAD FROM LIBRARY - load a model from the Monolix library; NEW MODEL - write a new, customized model using Mlxtran language (MlxEditor opens automatically).

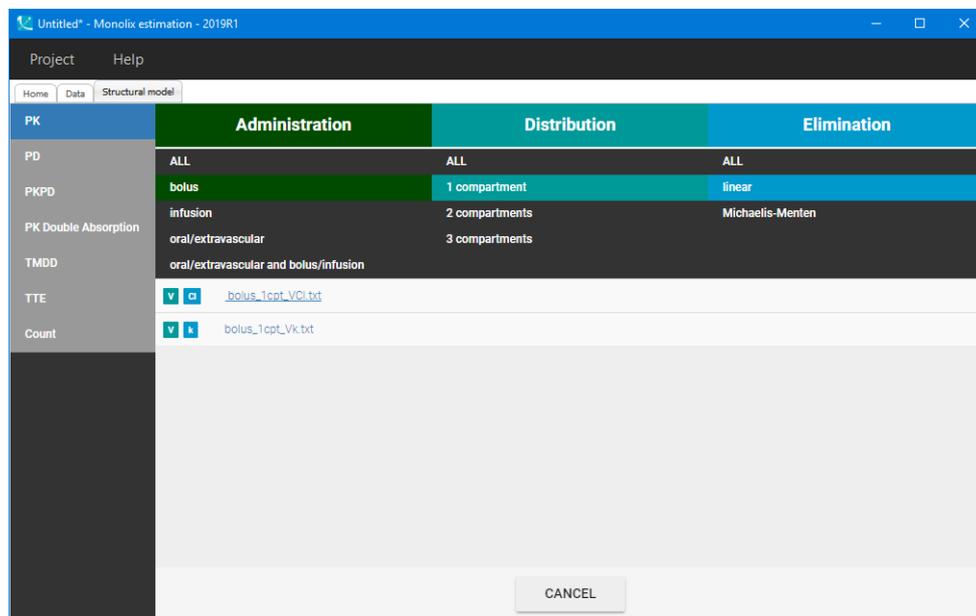


Figure 5: Using Monolix PK library to define a structural model.

Remark: Development of a structural model in Monolix is flexible. Each model from the library can be customized by a user. MlxEditor facilitates the editing process and allows for a quick check for potential errors. Additionally, any txt file with model description, written in the mlxtran language, is accepted. Clicking on the NEW MODEL button opens automatically MlxEditor to write a new model, or a model already written in a txt file can be loaded by clicking on the BROWSE button.

Initialization

At the next step the "Check initial estimates" tab is displayed, where initial parameter values have to be chosen. Although Monolix is very robust with respect to the initial parameter choice, it is a good practice to set reasonable values to speed up the convergence. By default, all values are set to one. They can be modified manually or by selecting "AUTO-INIT" button - a built-in algorithm finds the best fit for the pooled data. Figure 6 shows concentration data points for several individuals (in blue) and the model prediction (in red) for the population parameters, which are indicated at the bottom, obtained using the auto-init option.

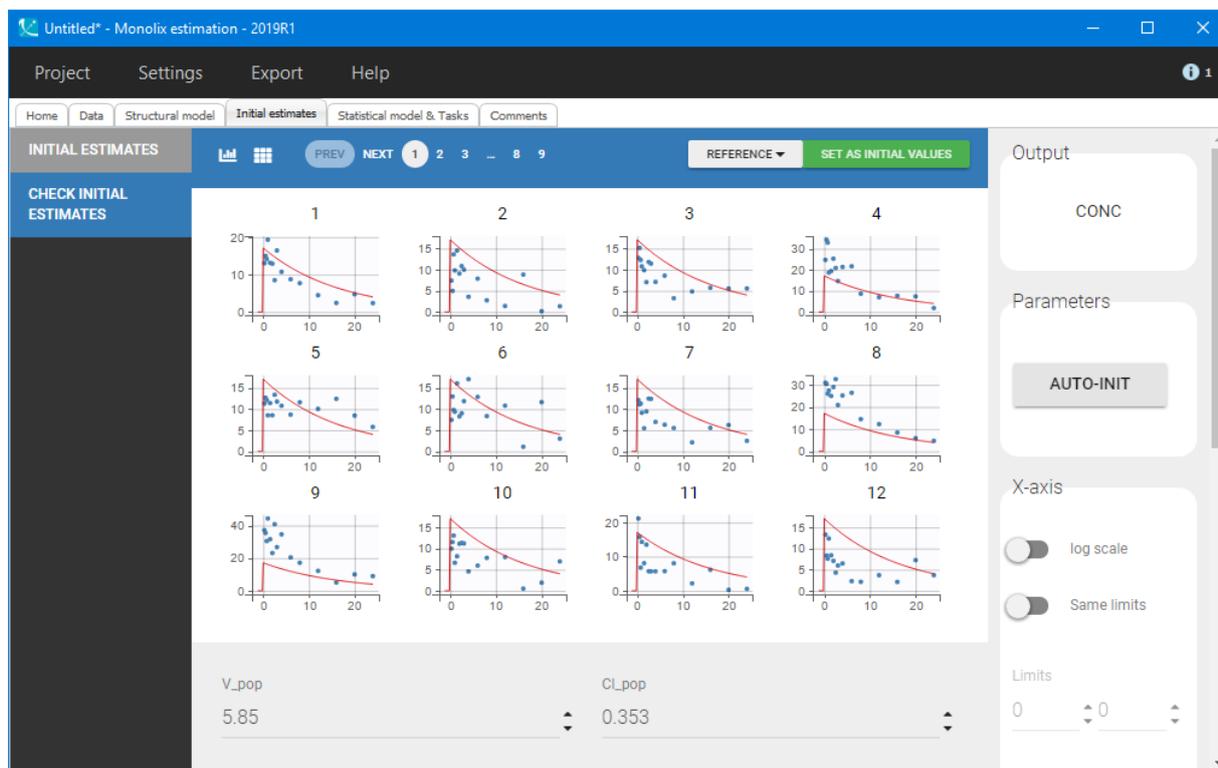


Figure 6: Graphical guide to choose initial parameters: concentration data points for several individuals (in blue) and model prediction (in red).

Initial values are accepted by clicking the green box "SET AS INITIAL VALUES" and this brings the user to the next tab, where the statistical part of the model has to be defined. For the base model the following options are set, see Figure 7:

- **Residual error model** - default combined1 model with an additive and a proportional term; it will be checked later if both terms are necessary.
- **Parameter distributions** - default log-normal distributions for V and Cl, and no correlations between random effects
- **Covariates** - ignored during the development of the structural model.

Before running any Monolix task, the project has to be saved, for example as "project_01.mlxtran" (menu: Project > Save), so that result files will be saved in a folder next to the project file.

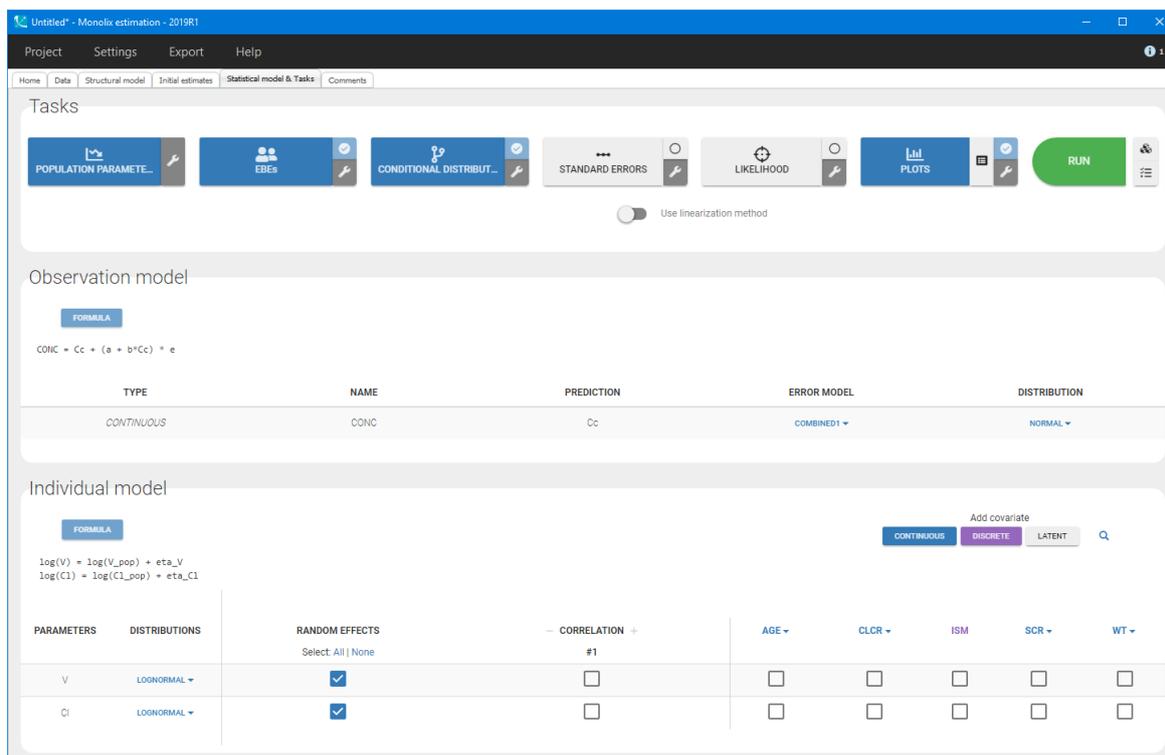


Figure 7: Defining the statistical model and running Monolix tasks.

3.2 Running Monolix tasks.

Running a task in Monolix is done by clicking on the corresponding button in the "Tasks" section of the "Statistical model and tasks" tab, see Figure 7. It is also possible to chose (tick) several tasks and then RUN. For each task, settings are available via the wrench icon. They are for advanced use only; running typical PKPD models does not require to adjust them.

The following estimation tasks can be run:

- **Estimation of population parameters**

The goal of this task is to calculate the population parameters that characterize the population: typical volume and clearance value in the population (called fixed effects), inter-individual variability (standard deviation of the random effects) and residual error parameters. When clicking on the task, a pop-up window shows the progression of the parameter estimation over the iterations of the stochastic approximation expectation (SAEM) algorithm. In the first phase (before the red line), the algorithm converges to a neighborhood of the solution, while in the second phase (after the red line) it converges to the maximum likelihood. By default, automatic termination criteria are used.

Here, the algorithm shows a good convergence: after an initial transition phase, the parameter estimates stabilize, see Figure 8. Once the task has finished, the estimated population parameters are displayed in a table in the "Result" tab, copied values are presented in Figure 9. For the volume, the result can be phrased as follows: "The typical volume is 7.14 L with 44.6% variability between individuals". Monolix does not report units, user has to deduce the units of the volume based on the units of the concentration and the amount in the data set.



Figure 8: Estimation of population parameters using Stochastic Approximation Expectation (SAEM) algorithm.

	VALUE	S.E. (S.E.%)
Fixed Effects		
V_{pop}	7.14	0.325 (4.55%)
Cl_{pop}	0.408	0.0193 (4.73%)
Standard Deviation of the Random Effects		
ω_V	0.446	0.033 (7.41%)
ω_{Cl}	0.415	0.0377 (9.09%)
Error Model Parameters		
a	2.37	0.114 (4.81%)
b	0.0814	0.00678 (8.33%)

Figure 9: Results: estimated population parameters (called fixed effects), inter-individual variability (standard deviation of the random effects) and residual error parameters.

- **Estimation of the EBEs.**

This task calculates the most probable parameter values for each individual (also called EBEs for empirical bayes estimates). It outputs the results, the estimates of individual parameters and a summary (min, median, max), in a table displayed in the "Results" tab, Figure 10. The EBEs are useful to generate the individual predictions used in plots such as the "Individual fits".

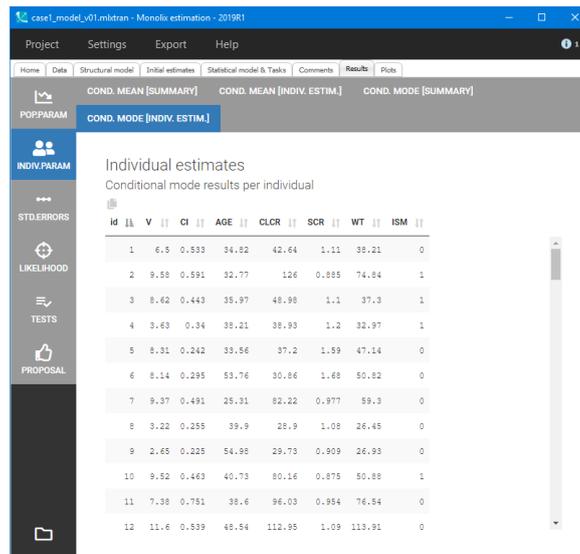


Figure 10: Results: individual parameters estimates.

- **Estimation of the conditional distribution.**

This task permits to sample individual parameter values from the conditional distribution of each individual using a MCMC procedure. The conditional distribution represents the uncertainty of the individual parameters. These values improve the diagnostic power of the plots. In the "Results" tab, the conditional mean (i.e average from the sampled values) is displayed for each individual.

- **Estimation of the standard errors via the Fisher Information Matrix (FIM).**

There are two options available: via linearization (toggle "Use linearization method" on) or via stochastic approximation (toggle off). The linearization method is fast, because it uses an approximate linearized model to calculate the standard errors. The stochastic approximation method uses the exact model and takes longer to run. The one-compartment model at hand

runs fast, so the stochastic approximation option is chosen.

Analysis of the results:

- The estimated standard errors and relative standard errors are displayed in the "POP.PARAM" sub-tab, see Figure 11 on the left. They represent the uncertainty of the estimated population parameters. In this run, all parameters are estimated with a good confidence (RSE < 10%).
- In the "STD.ERRORS" tab, the correlation matrix of the estimates is shown, see Figure 11 on the right. Values close to 1 or -1 on off-diagonal terms indicate that the corresponding population parameter values are highly correlated. Strong correlations usually suggest a poorly identifiable model. In this example, the error model parameters "a" and "b" are negatively correlated, which means that an almost equally good likelihood could be obtained by increasing a and decreasing b, or doing the opposite. As none of the value for "a" and "b" is small, a "combined2" error model, which has a different formula, could be tried.
- The MAX/MIN eigenvalue is called the condition number. Low value (i.e < 100) suggests that the model is not over-parameterized.

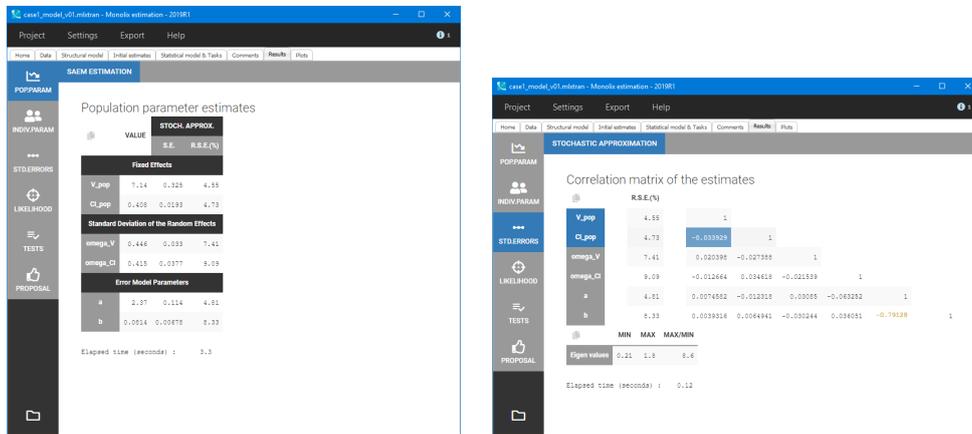


Figure 11: Results: population parameters with standard errors (on the left) and correlation matrix of the estimates (on the right).

- **Estimation of the log-likelihood**

The log-likelihood is the objective function optimized during the population parameter estimation. The construction of the SAEM algorithm makes it necessary to use a separate task to calculate the final value of the objective function. The estimation of the log-likelihood is available in two options: via linearization (faster but with approximate linearized model, toggle "Use linearization method" on) or via the importance sampling (exact model but slower, toggle off). The latter gives a -2log-likelihood, AIC and BIC criteria, which are displayed in the "LIKELIHOOD" sub-tab, see Figure 12. The AIC, BIC and BICc criteria take into account the number of estimated parameters, and can be used to compare models with each other.

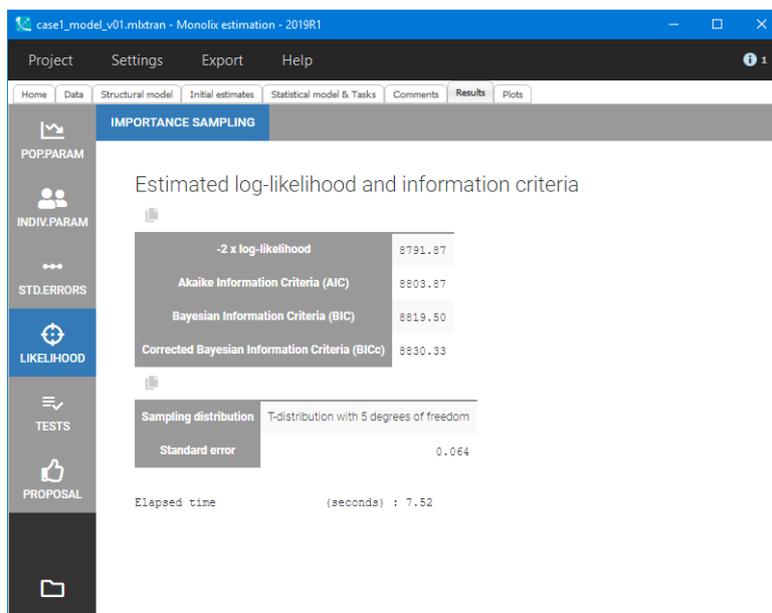


Figure 12: Results: estimated log-likelihood and information criteria.

- **Generation of the diagnostic plots**

This task generates automatically selected built-in graphics. By default, a small set of rapid-to-generate plots are selected. Clicking on the "list" icon of the right of the "PLOTS" button shows the full list (including VPC, etc), see Figure 13

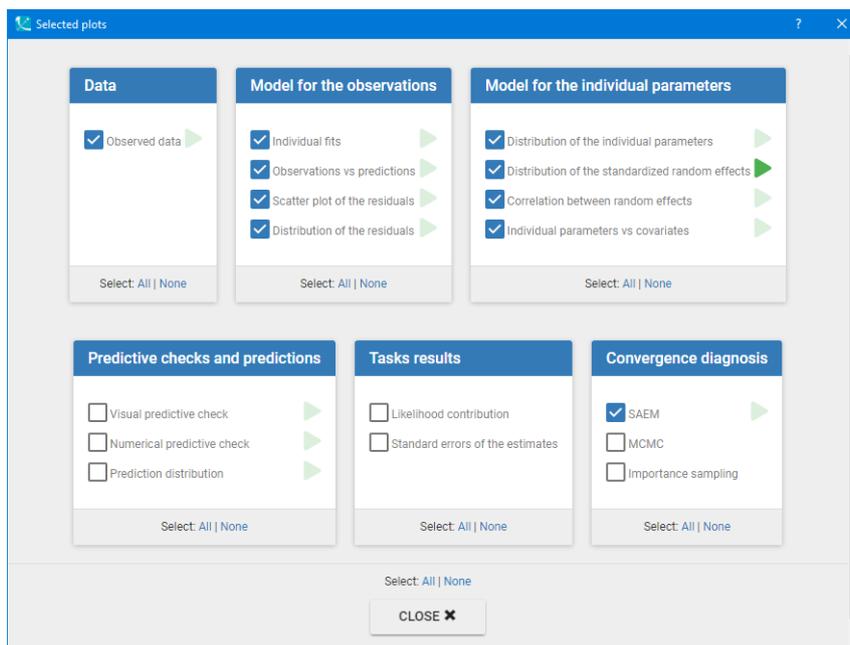


Figure 13: List of built-in diagnostic plots.

3.3 Diagnostic plots

The primary analysis of the goodness of the model is done through the diagnostic plots. The panel on the left of the "Plots" tab is divided into sections that correspond to model aspects that can be assessed by particular plots. The panel on the right allows to set many options for all plots: what to display (tab "Settings"), the layout such as colors, font size, etc (tab "Preferences") and stratification via splitting/coloring/filtering based on covariates (tab "Stratify").

At first the assessment of the structural model using the "Model for the observations" section is performed.

- **Individual fits** compare model predictions using the individual parameters with the observed data. Results, presented in Figure 14 for several individuals, show good agreement between predictions and observations. Pages for all individuals can be browsed with the buttons at the top.

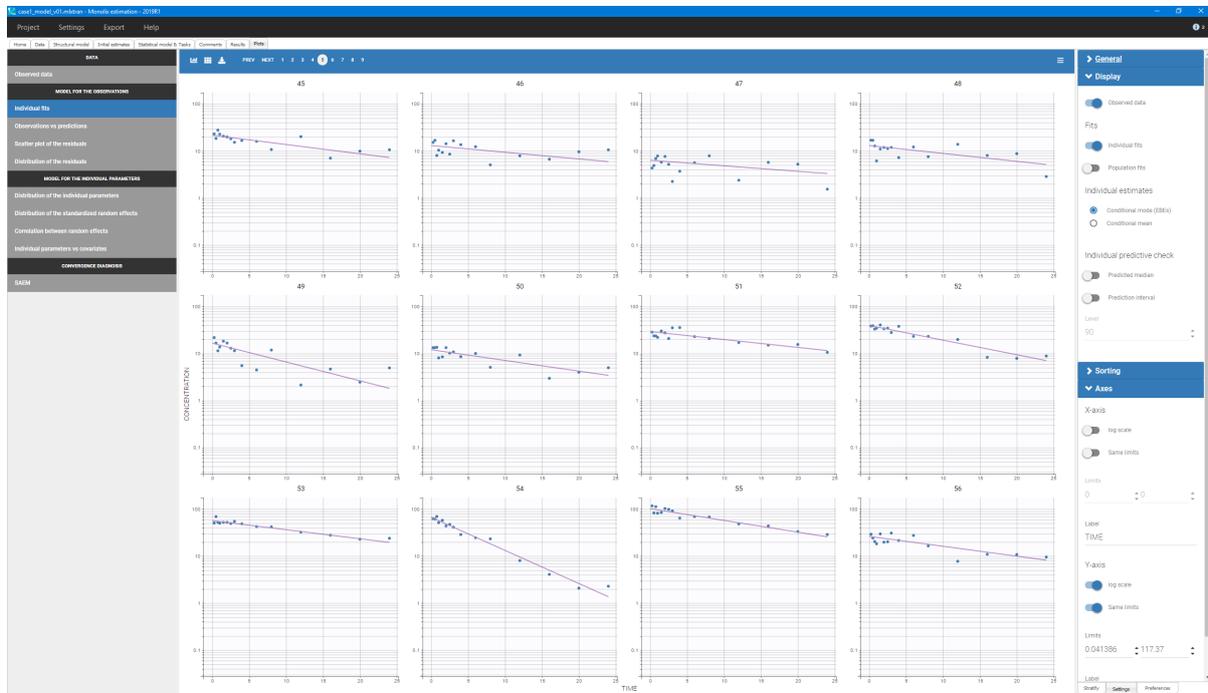


Figure 14: Diagnostic plot: individual fits for the one-compartment model without covariates. Log-scale and same y-axis limits for all individuals have been chosen in the tab "Settings".

- **Observations vs. predictions** plot is used to analyze possible mis-specifications of the structural model. In Figure 15 individual estimates sampled from the conditional distribution are used to generate the individual predictions. The 90% prediction interval helps to recognize outliers. Here points are well balanced on each side of the $y = x$ line and no mis-specification is detected.

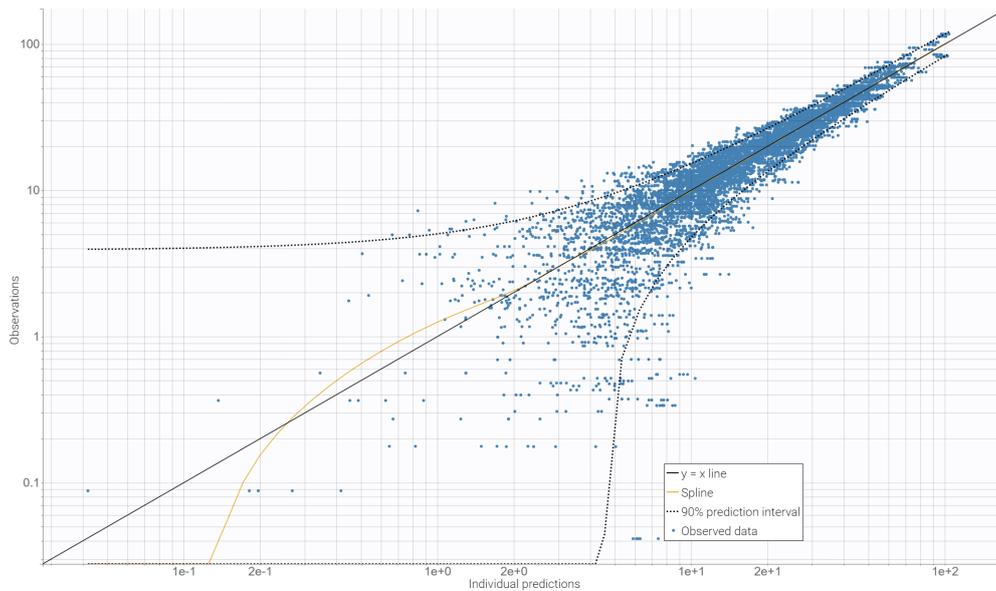


Figure 15: Diagnostic plot: observations vs. predictions for the one-compartment model without covariates. Log-log scale has been chosen, as well as the overlay of the 90% prediction interval.

- **Scatter plot of the residuals** displays the PWRES (population weighted residuals) and the IWRES (individual weighted residuals) w.r.t. the prediction. Symmetry of the distributions, see Figure 16, indicates that residuals behave as independent standardized normal random variables. No tendency with respect to time (not shown in the figure) or concentration is seen (see spline in orange).

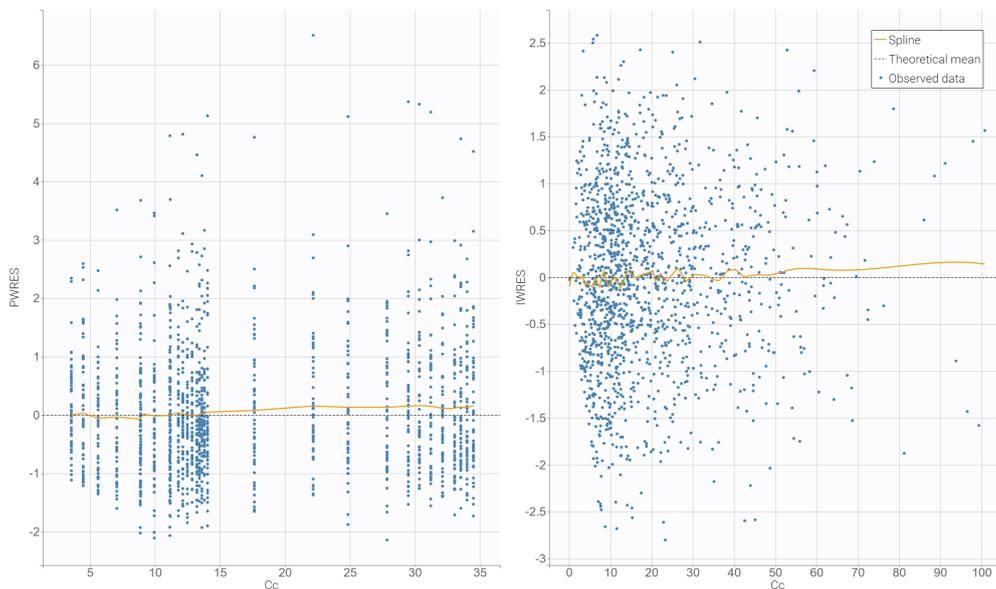


Figure 16: Diagnostic plot: scatter plots of PWRES and IWRES. The spline has been overlayed.

- **Distribution of the residuals** helps to detect mis-specification in the model. Figure 17 shows that the empirical and theoretical probability density function (PDF) of the PWRES and IWRES are in agreement.

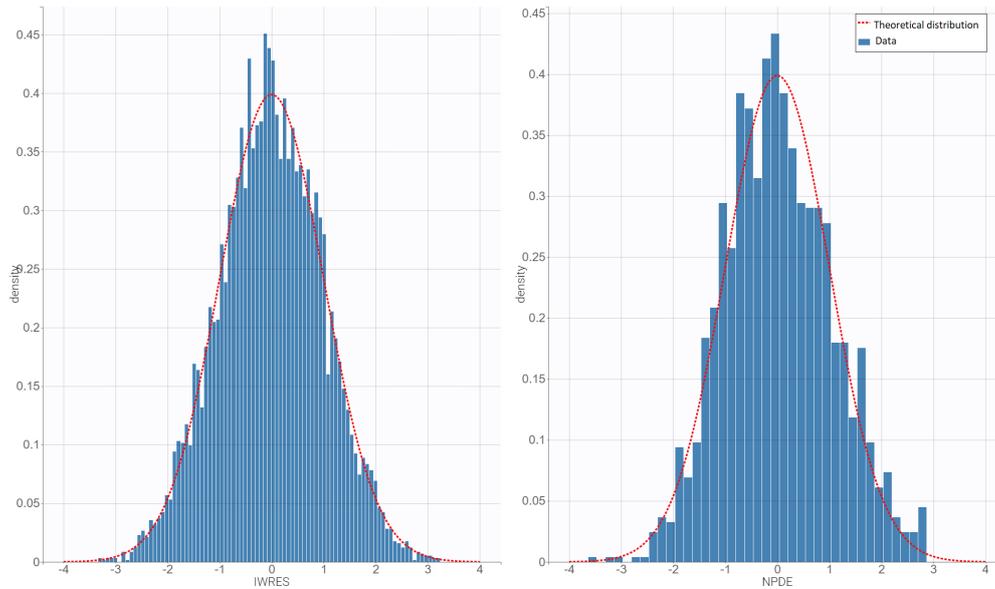


Figure 17: Diagnostic plot: distribution of the empirical (histogram) and theoretical (dashed line) PDF of the PWRES and IWRES.

In conclusion, diagnostic plots do not show any mis-specification of the structural model. However, Monolix offers more tools to analyze the goodness of fits and they are going to be described in the further sections.

3.4 Convergence assessment

The SAEM algorithm is a stochastic method, which means that the exact values of the estimated parameters depend on the sequence of random numbers set via the seed (the seed value can be seen in the global menu Settings > Project settings). The estimated parameters may also depend on the initial values, if different local minima are found when starting from different initial values. The **CONVERGENCE ASSESSMENT ALGORITHM** allows to execute a workflow of estimation tasks several times, with different initial values of fixed effects, and/or different seeds. The goal is to assess the robustness of the convergence. In this case study, there are no doubts about the good convergence of the SAEM algorithm (see Figure 8), and the convergence assessment is performed for a didactic purpose.

The convergence test can be run from the "Tasks" section - by clicking the button under the "run" arrow, see Figure 18 (on top) - or from the "Home" tab (on bottom left). After defining the number of runs, type of the assessment and initial parameters in a pop-up window (on bottom right), the assessment tool can be run. The progress of the SAEM algorithm convergence, parameters estimation, FIM and Importance sampling is monitored during the run.

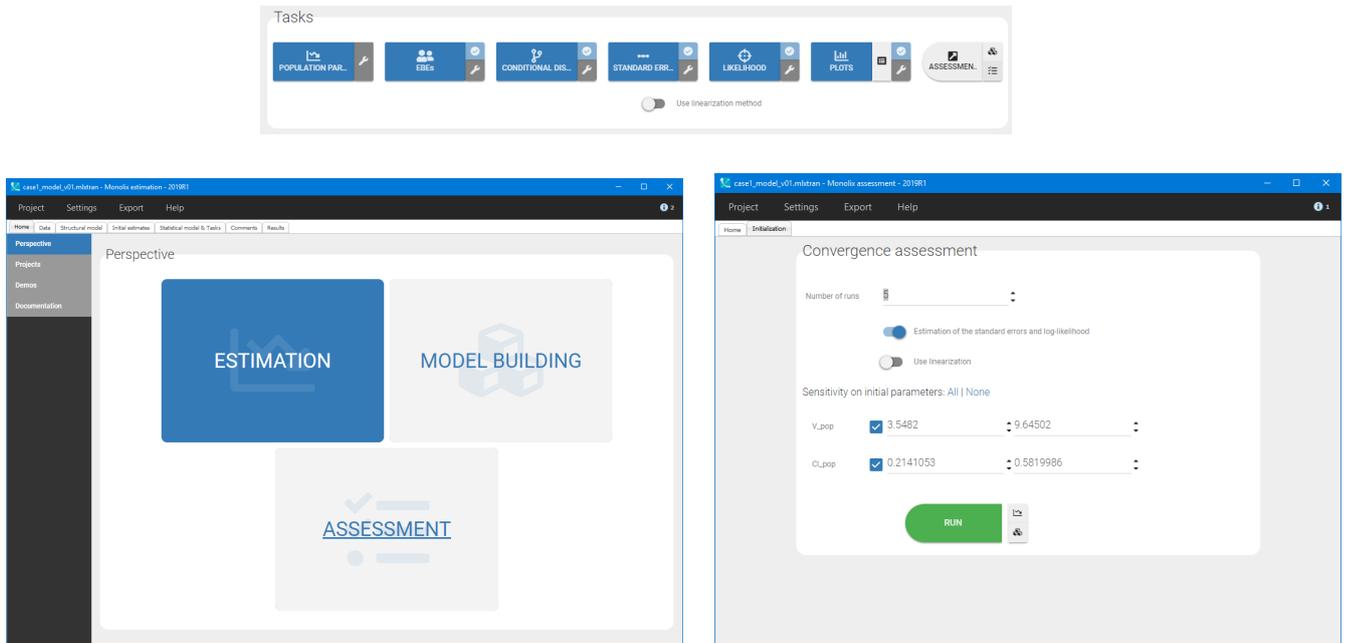


Figure 18: Convergence assessment tool: settings.

Here, the convergence assessment uses 5 runs with initial parameters sampled uniformly between the displayed bounds. The results, presented in Figure 19, show a very good convergence (as expected).

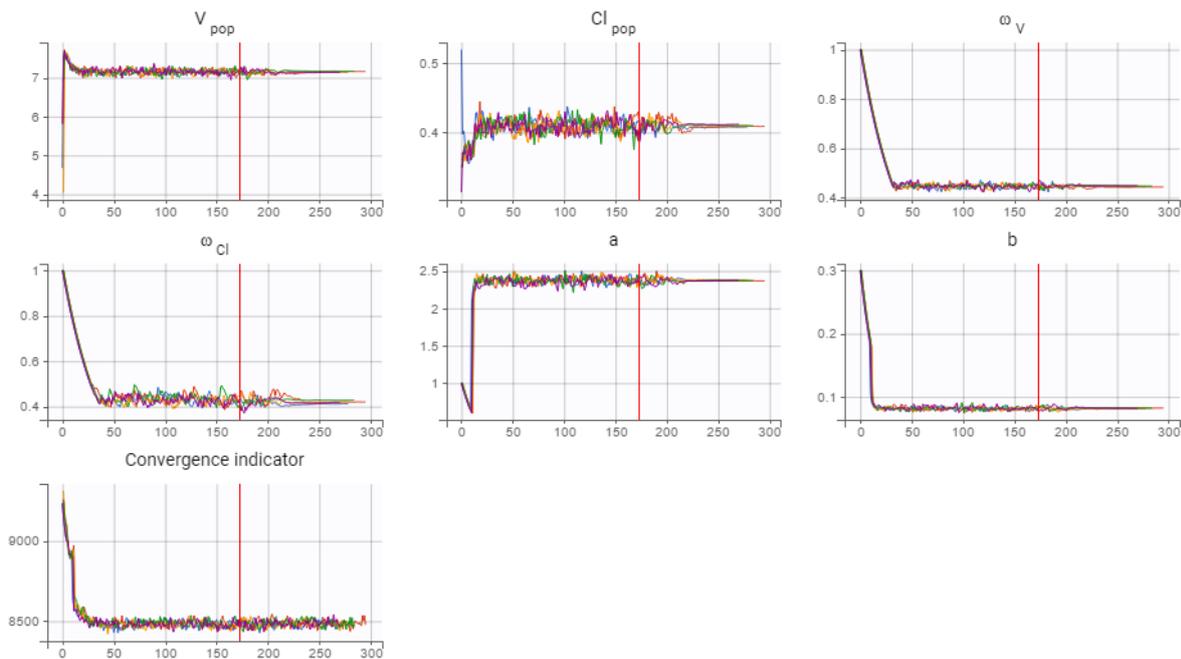


Figure 19: Convergence assessment tool: convergence of the SAEM algorithm for model parameters.

4 Development of the statistical model

4.1 Exploration of covariate relationships

Goodness of individual fits, studied in the previous section, is satisfactory. The next step is to assess if part of the inter-individual variability captured via the random effects can be explained by the

covariates.

In Monolix there are several tools to guide the user in the choice of covariates: the diagnostic plots, the tests and the proposals displayed in the "Results" section. All these tools give hints about possible parameter-covariate relationships from a statistical perspective. Adding covariates should also take into account physiological plausibility.

- **Random effects vs. covariates plots** show a strong correlations between the covariates age, creatinine clearance and weight with clearance and volume, see Figure 20.

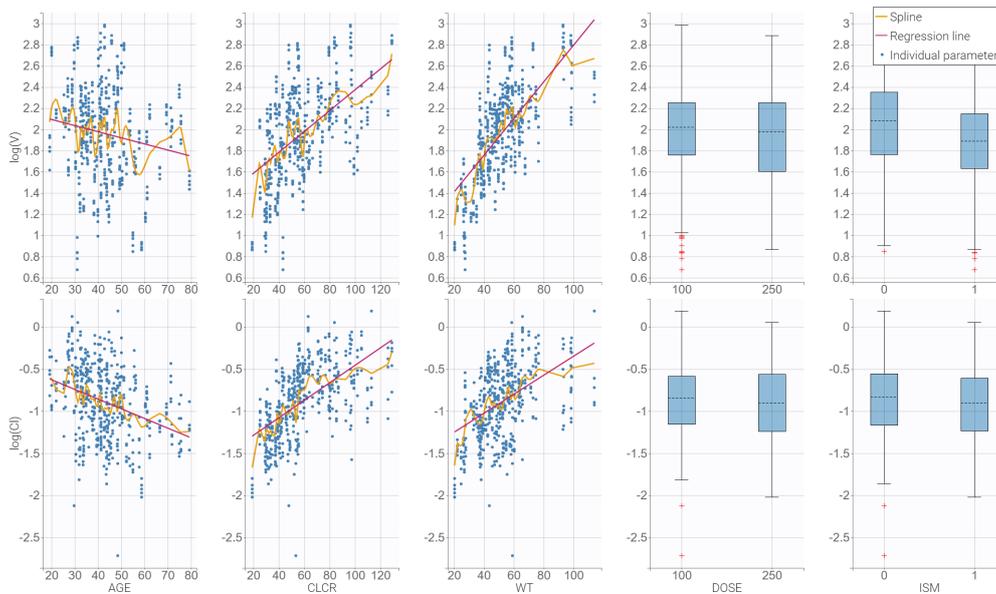


Figure 20: Individual, transformed volume and clearance parameters vs. covariates. Random effects can be chosen in the Settings. By default the random effects are calculated using the individual parameters sampled from the conditional distribution, which are not subject to shrinkage, in opposition to EBEs.

- **Pearson's test** checks if a covariate should be added to the model. A small p-value indicates that the correlation between the random effects and the covariate values is significant and so it is probably worth considering to add it in the model.

In case of the volume, p-values for creatinine clearance and weight are very small (colored in red), whereas in case of the clearance all covariates apart from gender and dose are strongly correlated, see Figure 21 on the left. Yet, the data exploration in Datxplre showed that weight and creatinine clearance are correlated too, so it may not be necessary to add both. For this reason, it is helpful to use the Proposal, see Figure 21 on the right, because it can better handle colinear covariates.

- **Proposal** is based on a linear regression between the individual parameters sampled from the conditional distribution and the covariates. Linear regressions with/without each covariate are ranked according to the number of covariates included and goodness of the regression. The best model is displayed in the first sub-tab. It suggests to add CLCR on clearance and WT on the volume.

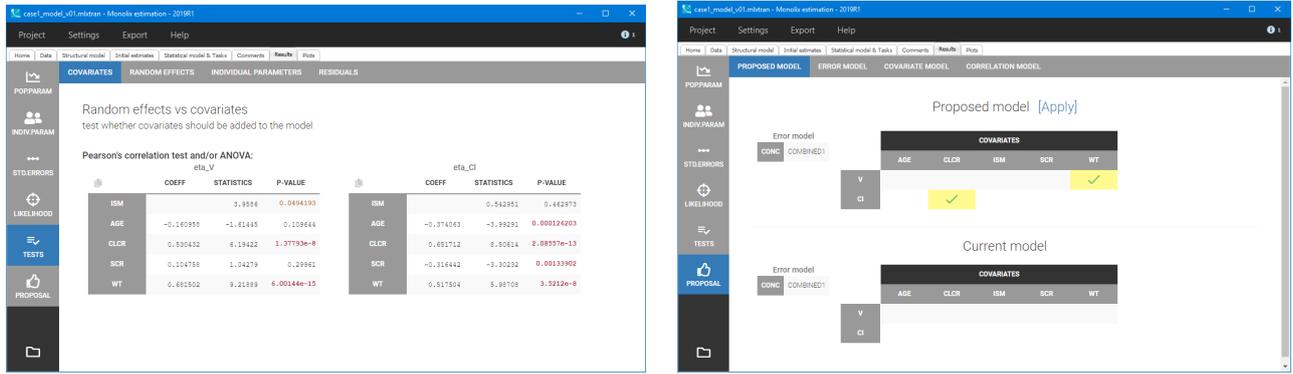


Figure 21: Results: statistical tests to analyze the covariate relationships (on the left) and Monolix "Proposal" to improve the model.

4.2 Covariate model

According to the diagnostic plots, tests and the Monolix proposal the statistical model is modified by adding CLCR covariate on clearance and WT covariate on the volume, as these two relationships make sense also physiologically. Before adding any covariates, it is a good practice to save the current estimated parameters as initial values of the next run. It can be done in the "Initial estimates" tab by clicking "Use last estimates" in the right top angle. It will accelerate the convergence of the next Monolix run.

Covariates are added in the "Statistical model & tasks" tab by clicking corresponding boxes in the "Individual model" section, see Figure 22. First centered, log-transformed covariates are created by clicking on the small blue arrow next to the covariate name. The covariates are then added to the model by ticking the corresponding box: creatinine clearance on the clearance and weight on the volume. The individual parameters model becomes:

$$\log(V) = \log(V_{pop}) + \beta_V \log tWT + \eta_V,$$

$$\log(Cl) = \log(Cl_{pop}) + \beta_{Cl} \log tCLCR + \eta_{Cl}.$$

Taking the exponential and replacing the $\log tWT$ and $\log tCLCR$ covariates by their definition (which can be seen by clicking on the small blue arrow > edit), the above set is equivalent to:

$$V = V_{pop} \left(\frac{WT}{49.16} \right)^{\beta_V} e^{\eta_V},$$

$$Cl = Cl_{pop} \left(\frac{CLCR}{53.57} \right)^{\beta_{Cl}} e^{\eta_{Cl}}.$$

Note that adding the non-transformed covariates would have resulted in an exponential relation between the covariate and the parameter. This is also an option, although less common.

Then the project is saved under the new name and the Monolix tasks can be run again. It is advised to always save a new project under a new name before running, otherwise the previous results will be overwritten.

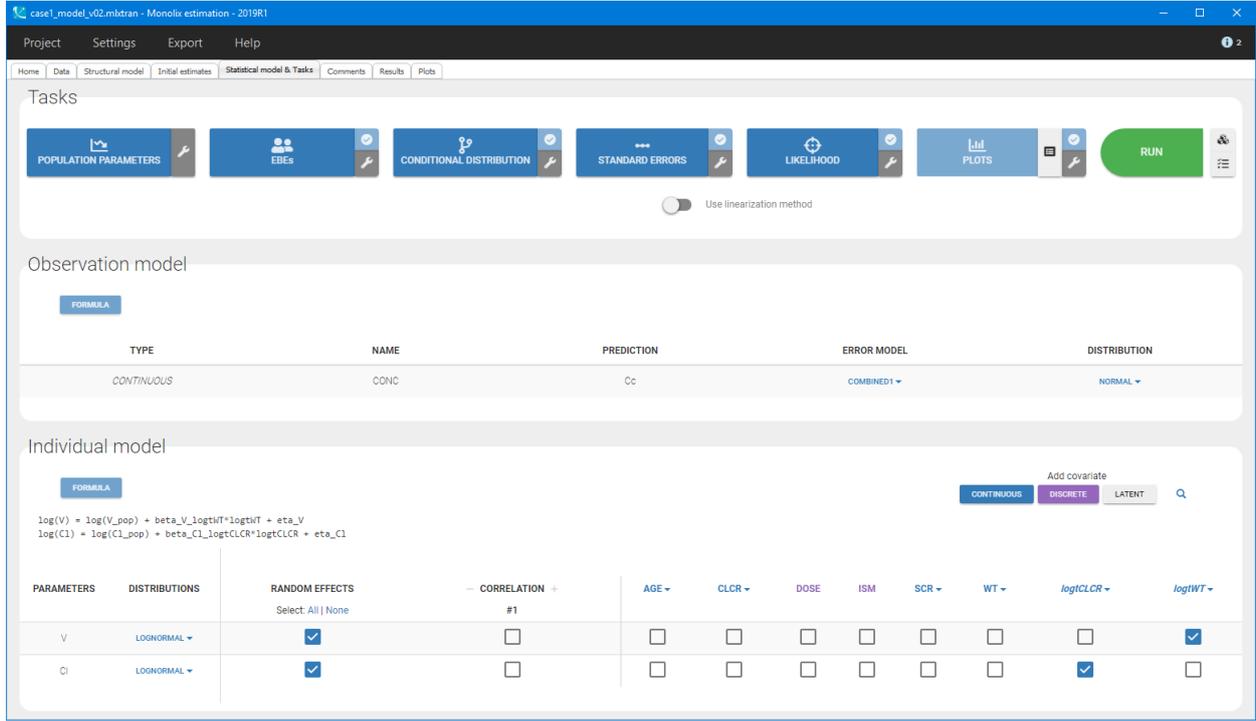


Figure 22: Covariate model development: adding log-transformed covariates to the statistical model.

4.3 Results and interpretation

As previously, parameter estimates, standard deviations and log-likelihood are displayed in the "Results" tab. Compared to the previous run, two additional beta parameters are estimated. The unexplained inter-individual variability (represented by the omega parameters) has decreased from 44.6% to 29.8% for the volume and from 41.5% to 24% for the clearance, see Table 2

	VALUE	S.E. (S.E.%)
Fixed Effects		
V_{pop}	7.12	0.221 (3.11%)
$\beta_V \log tWT$	0.993	0.0944 (9.51%)
Cl_{pop}	0.428	0.0134 (3.13%)
$\beta_{Cl} \log tCLCR$	0.872	0.0777 (8.91%)
Standard Deviation of the Random Effects		
ω_V	0.298	0.0237 (7.96%)
ω_{Cl}	0.24	0.0258 (10.7%)
Error Model Parameters		
a	2.39	0.111 (4.63%)
b	0.0801	0.00656 (8.19%)

Table 2: Results: estimated population parameters (called fixed effects), inter-individual variability (standard deviation of the random effects) and residual error parameters.

Updated p-values of the Pearson's correlation tests are shown in Figure 23 on the left. Small p-values in case of individual parameters confirm that the chosen covariates should be kept in the model, because there is a significant correlation with the individual parameters sampled from the conditional distribution. For the random effects, see Figure 23 on the right, the large p-values indicate that there are no other significant relationships between random effects and covariates that should be added to the model. Additionally, the Wald test, which relies on standard errors, checks if the beta parameter is equal to zero. A small p-value, as in the case of $\beta_{V \cdot \log(tWT)}$ and $\beta_{Cl \cdot \log(tCLCR)}$, indicates that the null hypothesis can be rejected and the covariate should be kept in the model.

Individual parameters vs covariates
test whether covariates should be removed from the model

Pearson's correlation test and/or ANOVA:

	COEFF	STATISTICS	P-VALUE
V			
logtWT	0.756711	11.4586	<2.2e-16
CI			
logtCLCR	0.880016	18.3426	<2.2e-16

Wald test (stochastic approximation):

	STATISTICS	P-VALUE		STATISTICS	P-VALUE
V			CI		
beta_V_logtWT	10.5145	<2.2e-16	beta_CI_logtCLCR	11.2206	<2.2e-16

Random effects vs covariates

test whether covariates should be added to the model

Pearson's correlation test and/or ANOVA:

	COEFF	STATISTICS	P-VALUE
eta_V			
DOSE		1.40683	0.238451
ISM		4.07085	0.0463633
AGE	-0.115136	-1.14742	0.254001
CLCR	-0.0953041	-0.947776	0.345575
SCR	0.127082	1.26833	0.207685
WT	-0.0465069	-0.460894	0.648896
logtCLCR	-0.066428	-0.65906	0.511403
logtWT	-0.000916747	-0.00907534	0.992777

	COEFF	STATISTICS	P-VALUE
eta_CI			
DOSE		0.870988	0.352976
ISM		1.86999	0.174604
AGE	-0.0529415	-0.52483	0.600887
CLCR	-0.0289174	-0.286388	0.775186
SCR	-0.106516	-1.06048	0.291531
WT	-0.0527556	-0.522882	0.602167
logtCLCR	-0.00475255	-0.0470484	0.96257
logtWT	-0.0357532	-0.354165	0.723976

Figure 23: Results: Pearson's test and Wald test for the model with covariates.

New scatter plots of random effects vs. covariates, presented in Figure 24, do not show any other correlations.

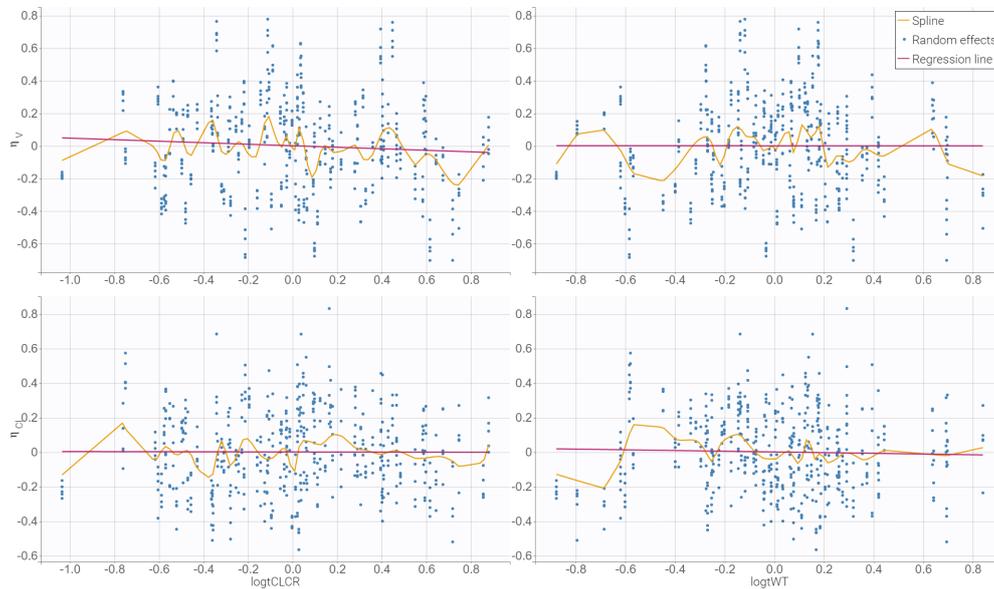


Figure 24: Scatter plots of random effects vs. covariates for the model with covariates.

4.4 Automatic Covariate Model Building

While in the above section covariate analysis was based on the tests and diagnostic plots, it is also possible to let Monolix automatically build the covariate model. The "AUTOMATIC MODEL BUILDING" contains routines exploring automatically, in a stepwise manner, parameter-covariate relationships that can improve the model. It is found in the "Home" tab, similarly to the Convergence Assessment tool, see Figure 25 on the left.

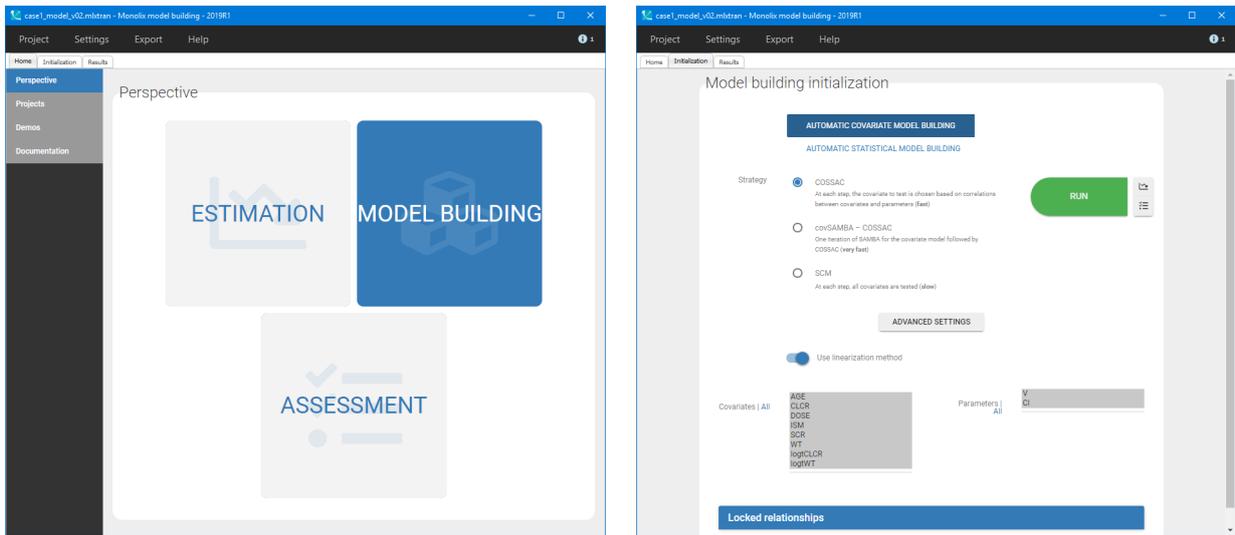


Figure 25: Automatic covariate model building

For example, the **COSSAC** routine (Conditional Sampling for Stepwise Approach based on Correlation tests) makes use of the information contained in the "Tests" of the base model run to choose which covariate to try first (instead of trying all covariates "blindly"). The results of each iteration are displayed in the tab Results, see Figure 26. By default, the runs are displayed by order of iteration, except the best model which is displayed in the first position and is highlighted in green. Differences with the previous iteration are highlighted in yellow.

In this case the best model obtained by the automatic covariate search is the same as the one found via the analysis of diagnostic plots, tests and proposals, that is creatinine clearance explains variability in the value of the clearance parameter, while weight explains variability in the volume.

	AGE	CLCR	DOSE	ISM	SCR	WT
Iteration: 3 BICc: 8712.33 -2LL: 8664.65						Export and load
V						✓
Cl		✓				
Iteration: 1 BICc: 8831.58 -2LL: 8793.11						Export and load
V						
Cl						
Iteration: 2 BICc: 8770.81 -2LL: 8727.74						Export and load
V						✓
Cl						
Iteration: 4 BICc: 8712.65 -2LL: 8660.37				✓		Export and load
V						✓
Cl		✓				
Iteration: 5 BICc: 8714.44 -2LL: 8662.17						Export and load
V	✓					✓
Cl		✓				
Iteration: 6 BICc: 8716.02 -2LL: 8663.74						Export and load
V						✓
Cl		✓	✓			

Figure 26: Results of the Automatic Covariate Model Building

5 Final plots

The VPC (Visual Predictive Check) plot, presented in Figure 27, gives a graphical assessment whether simulations from a model of interest are able to reproduce both the central trend and the variability in the observed data, when plotted versus an independent variable (typically time). The observed empirical percentiles remain within the prediction intervals, which means that the model properly captures the data.

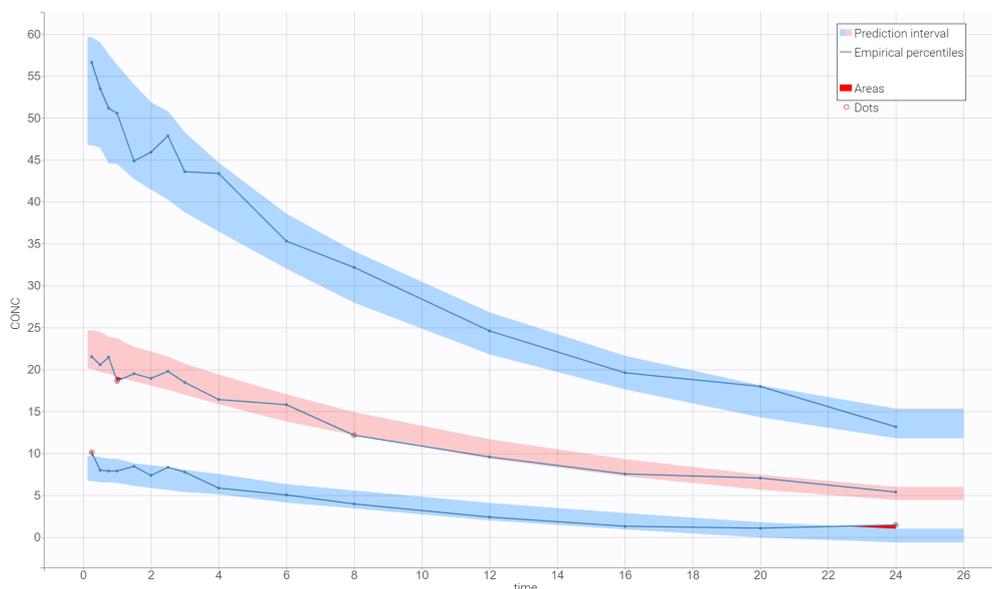


Figure 27: Visual predictive check (VPS): empirical percentiles and prediction intervals for the final model with covariates.

6 Summary of results

This case study focused on a development of a population PK model of a hypothetical drug. The structural part is a one compartment model with linear elimination and is parametrized by volume and clearance. Model parameters were estimated using the SAEM algorithm implemented in Monolix. Subsequently, through the covariate exploration using diagnostic plots, statistical tests and automated covariate model building tools, significant covariate relationship explaining the variability were found. In particular, the statistical model was enriched with the correlation between creatinine clearance and clearance and between weight and volume. Final results and diagnostic plots confirm that the model correctly captures the data.

The Sycomore application, part of the MonolixSuite, is designed to organize modeling workflow. It clearly displays summary of results, see Figure 28, and allows to compare different models side by side, see Figure 29.

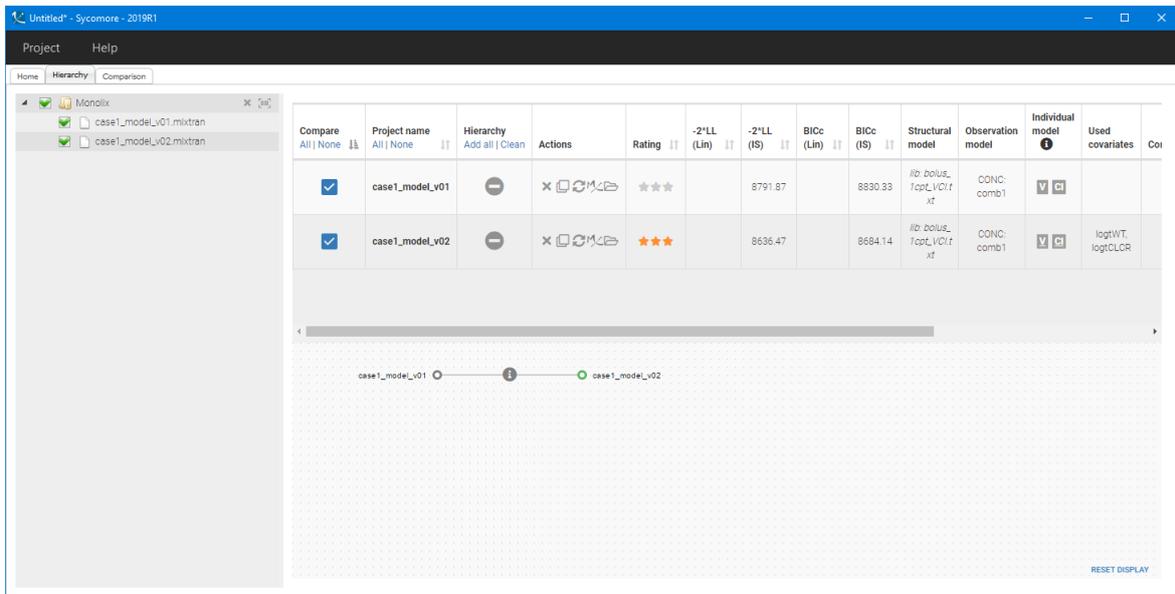


Figure 28: Sycomore application to manage modeling workflow: summary of the results.

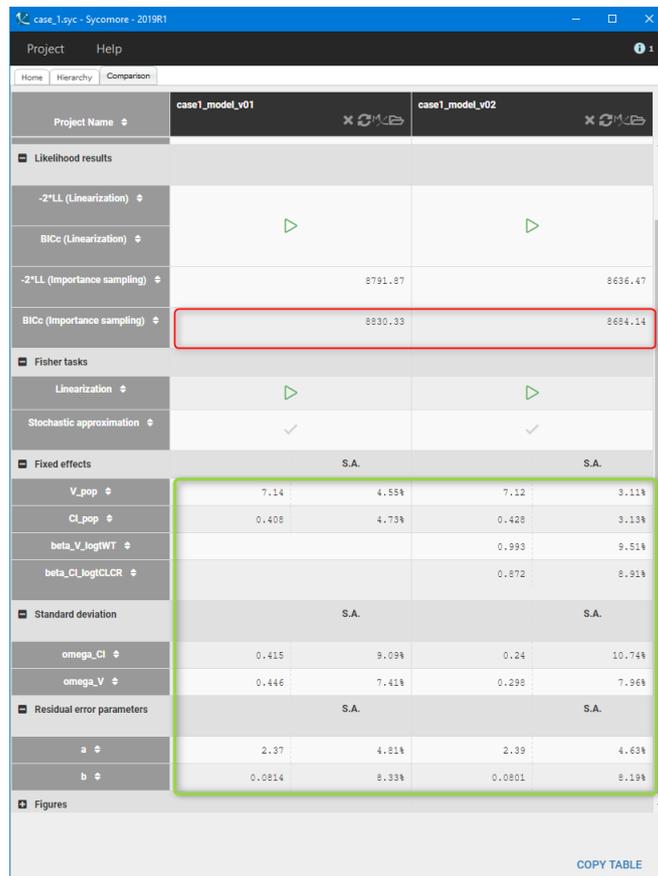


Figure 29: Sycomore application to manage modeling workflow: side by side model comparison.