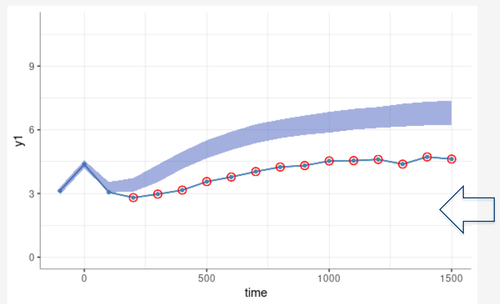
# A method to correct VPC bias due to non-random dropout via censored data addition, using the MonolixSuite

Agathe Yvinec-Tolmer[1], Monika Twarogowska[1], Pauline Traynard[1], Géraldine Cellière[1]

(1) Simulations Plus, Lixoft division, Antony, France.

## INTRODUCTION

**Non-random dropout,** where the **observation value and the probability of dropout are correlated,** is common in clinical studies focusing on **tumor growth dynamics**. This phenomenon results in an underestimation of the mean tumor size and can give a **bias in the Visual Predictive Check (VPC) plot**.




Visual predictive check (VPC): combines empirical percentiles and simulated predictions intervals

**Selected model can be correct, but the VPC is biased due to non-random dropout**
**VPC cannot be used for model diagnosis**

## GOAL: Recover diagnostic value of the VPC plot.

We present a **method to correct the VPC bias due to non-random dropout by adding censored observations after the time of dropout.** The method:

- Works without a dropout model
- Does not require precise predictions of the observation after a dropout
- Is integrated in Monolix workflow with simple data modification
- Preserves detection of a model misspecification

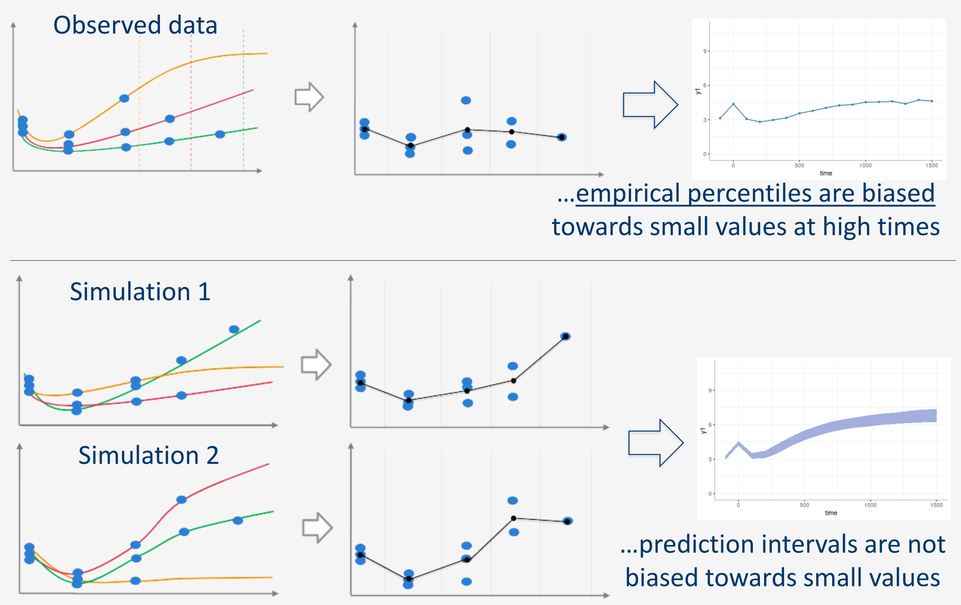
## ORIGIN OF THE BIAS

**OBSERVED DATA: empirical percentiles**
Percentiles, e.g. median in the figure, are computed from the original data across bins.

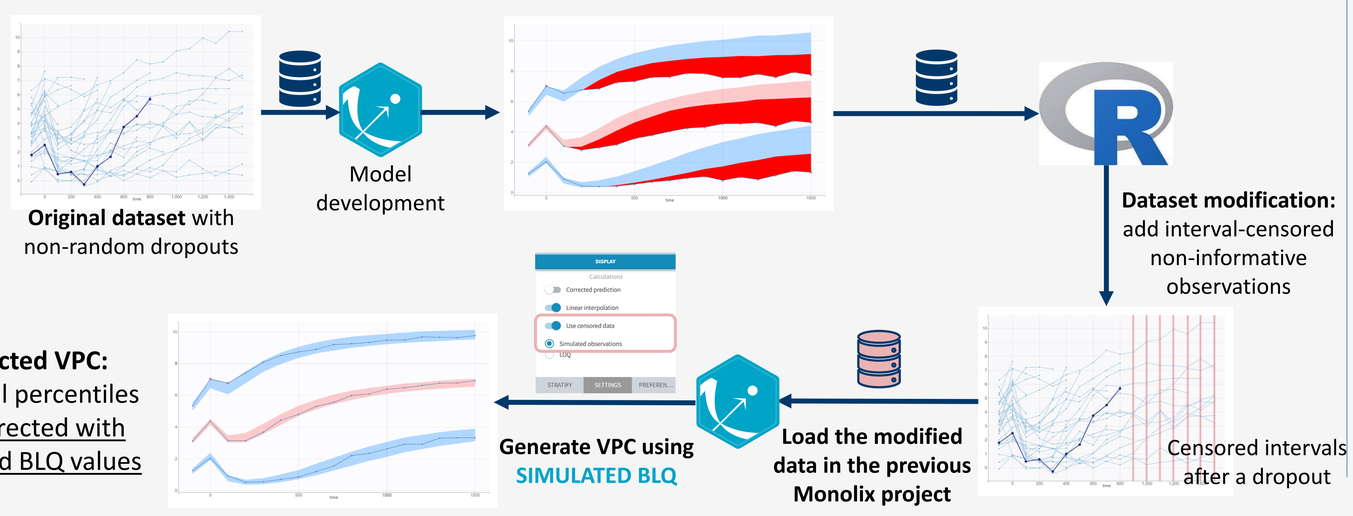
*If there is a correlation between the value of the observation and the probability of a dropout, e.g. the higher the value, the more likely that it is missing, then…*

**VPC SIMULATIONS: prediction intervals**
Prediction intervals are calculated over percentiles for each of the 500 datasets generated via Monte Carlo simulations in Monolix with the same design structure as the original dataset (measurements and doses).

*If a model used in the Monte Carlo simulations does not include the correlation between dropout and observation value, then…*




## NEW METHOD: WORKFLOW



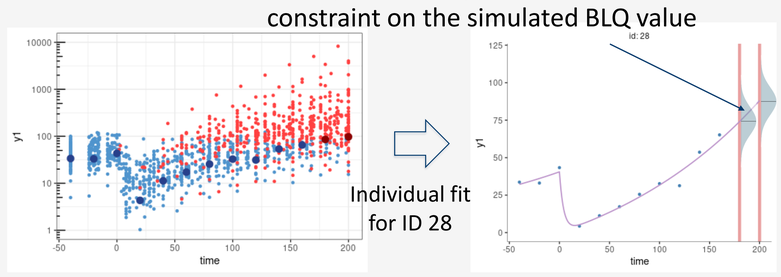

### Simulated BLQ

**Monolix simulates BLQ** values for each individual using:

- model prediction at time of the BLQ value and
- truncated residual error - truncation assures that the observation is in the censored interval

$$\mathrm{simBLQ_i} = \mathrm{f}(\psi_i, t_{ij}) + \left[b \cdot f(\psi_i, t_{ij}) \cdot \varepsilon\right]_{\mathrm{truncated}}$$

Large censoring interval removes constraint on the simulated BLQ value




## NEW METHOD: EFFECTIVENESS

A robust method corrects the bias in the VPC due to non-random dropouts and keeps its diagnostic value to detect a model misspecification. **We present a routine to study the effectiveness of a VPC correction method**.
The proposed VPC correction method has been tested using simple tumor growth models (exponential, logistic), Wang and Two-Population models from the MlxTran Tumor Growth Inhibition (TGI) library, and a mechanistic model of Prostate Specific Antibodies (PSA) kinetics[1].

### Method effectiveness assessment: procedure

**1. Simulate a dataset with non-random dropout according to a known model**

**2. Check the correction of the bias in VPC**

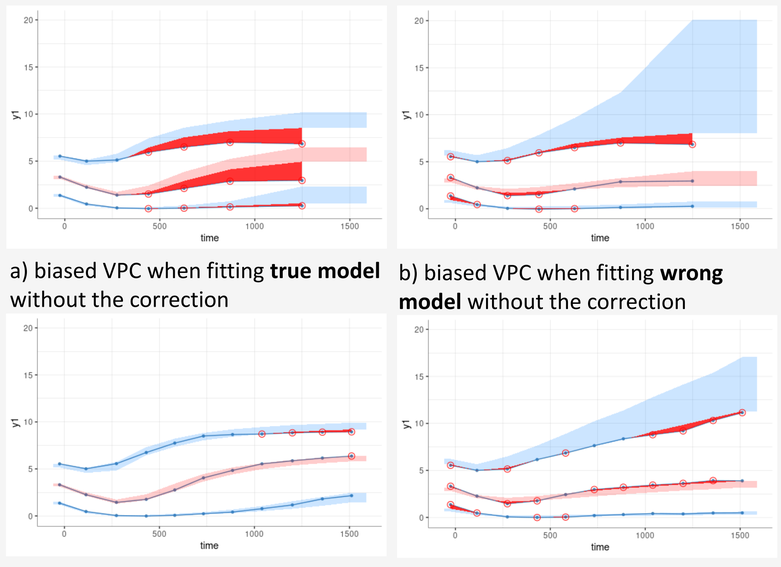
- Fit the model of interest to simulated data ⇒ VPC is biased (figure 4a)
- Apply proposed method to create a new dataset with censored observations
- Check that the bias is corrected in the VPC (figure 4c)

**3. Check for over-correction**

- Fit "wrong" models that don't capture the simulated data (figure 4b)
- Check that the bias-corrected VPC still shows a misspecification (figure 4d)

[1] Desmée, Solène et al. "Using the SAEM algorithm for mechanistic joint models characterizing the relationship between nonlinear PSA kinetics and survival in prostate cancer patients." Biometrics vol. 73,1 (2017): 305-312. doi:10.1111/biom.12537

**Example:** simulated data according to a model for PSA[1], and the Stein model as a "wrong" model for comparison.




a) biased VPC when fitting **true model** without the correction

b) biased VPC when fitting **wrong model** without the correction

c) VPC when fitting **true model** after the correction

d) VPC when fitting **wrong model** after the correction

## CONCLUSIONS

### ADVANTAGES OF THE METHOD

- **Feasible in complex case studies** because adding censored intervals does not require an accurate prediction of the observation, e.g. tumor size, beyond the dropout
- **Automated** because dataset modification, generation of Monolix projects and VPC plots can be scripted in R using lixoftConnectors functions.
- **Gives an immediate feedback** because the correction can be easily integrated withing a Monolix workflow of model development

**Dedicated R function:**

```
> correctVpc(projectName = "simulatedPSA.mlxtran",
  newDatasetName = "simulatedPSA_vpcdata.csv",
  timeIntervalBetweenObs = "uniform",
  uniformTimeIntervalValue = 84)
```

Contact information:
monika.twarogowska@simulations-plus.com