Understanding sources of variability: The variability attribution plot

Sebastian Ueckert, Svetlana Freiberga, Gunnar Yngman, Mats O. Karlsson
Pharmacometrics group, Department of Pharmaceutical Biosciences, Uppsala University

Objective

Understanding and explaining sources of variability in the data is a significant component of any analysis. Clinical trial data with its many levels of hierarchy is particularly complex. Fortunately, nonlinear mixed effect models (NLMEM) have proven to be well suited to handle this complexity, and have become the quasi-standard for pharmacometric trial data analysis. However, it can be challenging in practice to understand the complex interactions of parameter variability, residual variability, covariates and the non-linear function.

In light of these complexities, the objective of this work was to introduce a novel model visualization that allows attributing the different constituents of response variability over time.

Methods

This work considers NLMEM for continuous data with normal distribution in the two-stage representation, i.e. observation j from subject i follows:

$$y_{ij} \sim \mathcal{N}\left(f(t_{ij}, \phi_i), h(t_{ij}, \phi_i)\right) \qquad \phi_i = g(\theta, \eta_i, a_i)$$

where $f(\cdot)$ and $h(\cdot)$ are non-linear functions of the observation time t_{ij} and individual model parameters ϕ_i. The individual model parameters are described by the non-linear function $g(\cdot)$ of the fixed effects θ, the random effects η_i ($\eta_i \sim \mathcal{N}(0, \omega)$) and covariates a_i.

Derivation

From the law of total variance, the model variability can be split into residual unexplained variability (RUV) and between-subject variability (BSV), i.e.

$$\mathrm{Var}(y_{ij}) = \overbrace{\mathrm{E}\left(\mathrm{Var}\left(y_{ij}|\phi_i\right)\right)}^{\text{RUV}} + \overbrace{\mathrm{Var}(E(y_{ij}|\phi_i))}^{\text{BSV}}.$$

The BSV can then be further decomposed by successively conditioning the remaining variability on groups of variables that contribute variability γ_i (random effects η_i and/or covariates a_i) [1]:

$$\begin{aligned}\mathrm{Var}(E(y_{ij}|\phi_i)) = &\mathrm{Var}(\mathrm{E}[y_{ij} \mid \gamma_i^1]) \\ &+ \mathrm{E}\{\mathrm{Var}[\mathrm{E}(y_{ij} \mid \gamma_i^1, \gamma_i^2) \mid \gamma_i^1]\} \\ &+ \mathrm{E}\{\mathrm{Var}[\mathrm{E}(y_{ij} \mid \gamma_i^1, \gamma_i^2, \gamma_i^3) \mid \gamma_i^1, \gamma_i^2\} \\ &+ \dots\end{aligned}$$

Each term in the previous equation, as well as, the RUV constitute a variability component. They can be approximated by linearizing the NLMEM [2] and then deriving the conditional variability for the resulting linear mixed effect model.

Implementation

The visualization method was implemented using NONMEM, PsN, and R. First, PsN is used to generate an augmented NONMEM control stream extracting the necessary derivative information for the linearization. Finally, R is used to calculate the conditional variability expressions and plot each variability component as stacked ribbon versus the independent variable.

Illustration model

The properties of this visualization are illustrated using a PKPD model with one-compartment first-order absorption PK and an immediate effect Emax PD model [3], i.e.

$$y_{ij}^{\mathsf{PK}} = C(t_{ij}) = \frac{D}{V_i}\frac{k_{ai}}{k_{ai} - CL_i/V_i}\left(\exp\left(-CL_i/V_i t_{ij}\right) - \exp\left(-k_{ai} t_{ij}\right)\right) \cdot \left(1 + \varepsilon_{ij}^1\right) + \varepsilon_{ij}^2$$

$$y_{ij}^{\mathsf{PD}} = E_{0i} + \frac{E_{maxi} \cdot C(t_{ij})}{C(t_{ij}) + EC_{50i}} + \varepsilon_{ij}^3$$

$$CL_i = \theta_1 \cdot \left(\frac{a_{wt}}{70}\right)^{0.75} \cdot \exp(\eta_{1i}) \qquad V_i = \theta_2 \cdot \frac{a_{wt}}{70} \cdot \exp(\eta_{2i}) \qquad k_{ai} = \theta_3 \cdot \exp(\eta_{3i})$$

$$E_{0i} = \theta_4 \cdot \exp(\eta_{4i}) \qquad E_{maxi} = \theta_5 \cdot \exp(\eta_{5i}) \qquad EC_{50i} = \begin{cases} \theta_6 \cdot \exp(\eta_{6i}) & a_{gen} = 0 \\ \theta_7 \cdot \exp(\eta_{6i}) & a_{gen} = 1 \end{cases}$$

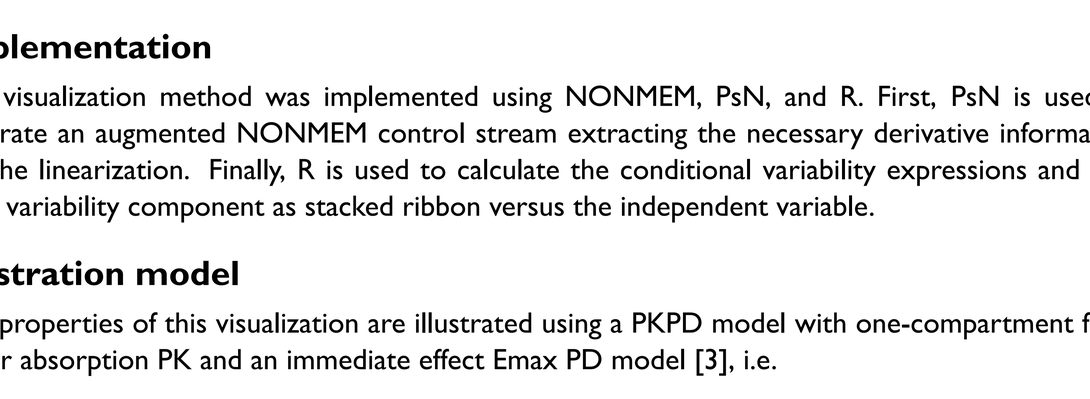

Figure 1: Model predicted PK and PD response (mean plus 95% prediction interval).

Results

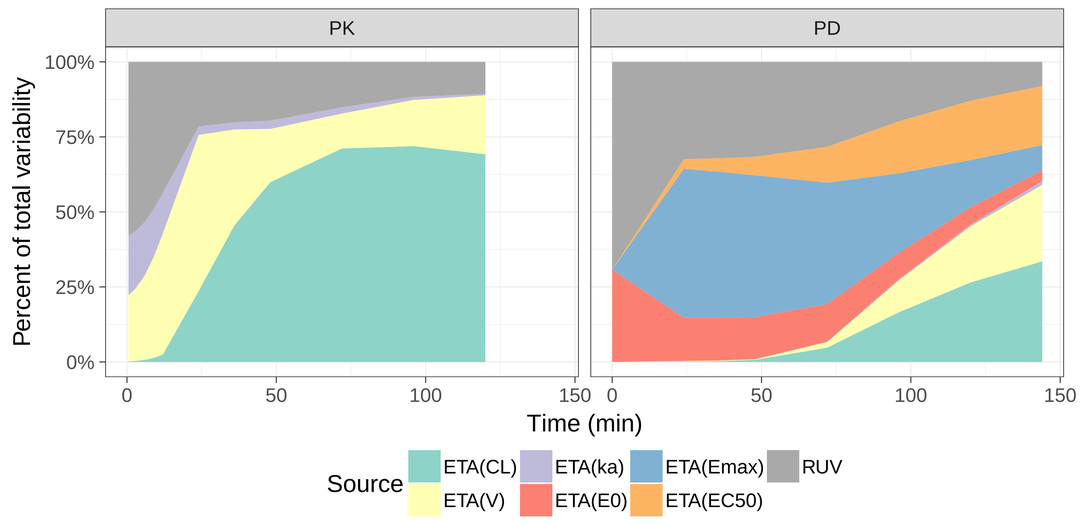

Figure 2: Variability attribution plot for PK and PD outcomes of the example model. The plot visualizes the variability contribution from the RUV as well as all BSV random effects. The creation of this plot requires merely a model file and is fully automatized. A version of this graph is part of the report for the PsN QA tool [4].

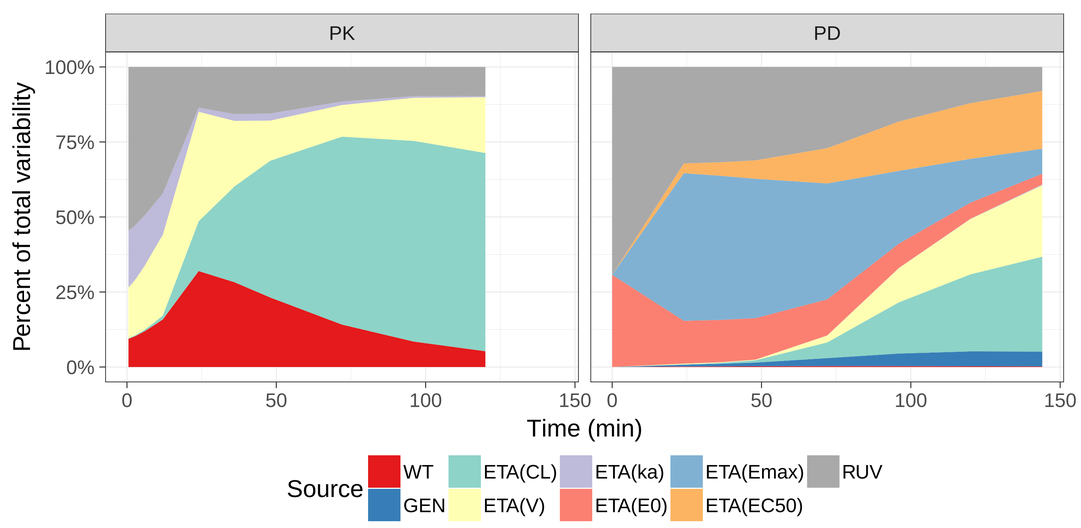

Figure 3: Variability attribution plot including the effect of covariates. Covariates can be taken into account by providing the mathematical expressions for ϕ_i to the algorithm. Covariates in FREM-type models are automatically taken into account.

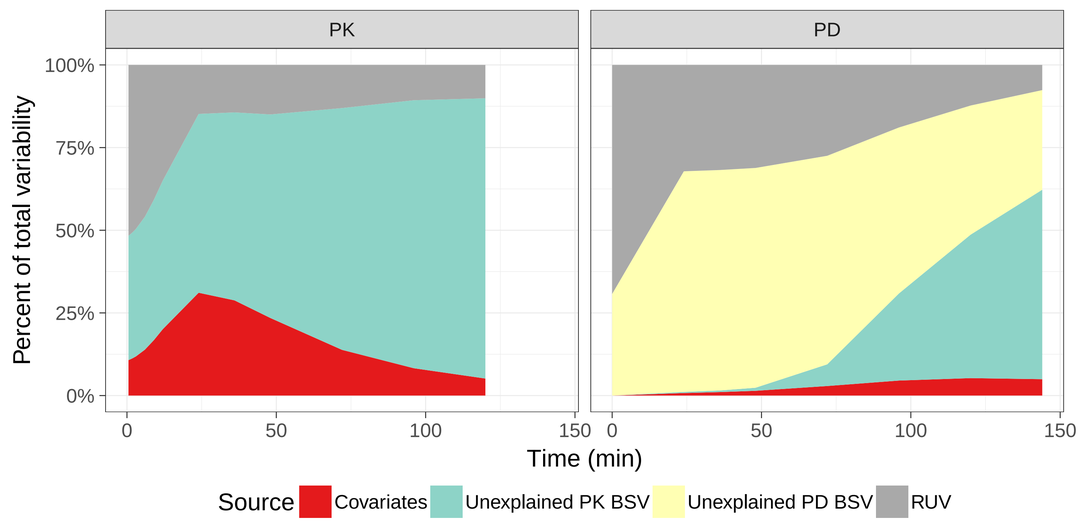

Figure 4: Variability attribution plot with grouped sources. The algorithm permits to freely choose the conditioning order as well as to group sources of variability, allowing to highlight particular aspects of a model. This version of the previous plot shows a band for covariates, PK and PD random effects.

Conclusion

This new visualization provides insights into the importance of the many sources of variability in an NLMEM. The gained understanding can guide model building decisions and assist in communicating model properties.

Acknowledgement: We thank Nick Holford for providing the model code for the example.

[1] Clive G Bowsher and Peter S Swain. Identifying sources of variation and the flow of information in biochemical networks. *Proceedings of the National Academy of Sciences*, 109(20):E1320–E1328, 2012.

[2] Akash Khandelwal, Kajsa Harling, E Niclas Jonsson, Andrew C Hooker, and Mats O Karlsson. A fast method for testing covariates in population pk/pd models. *The AAPS journal*, 13(3):464, 2011.

[3] Mats O Karlsson and Nick Holford. A tutorial on visual predictive checks. *PAGE 17 (2008) Abstr 1434*, 2008.

[4] Mats O Karlsson, Svetlana Freiberga, Gunnar Yngman, Rikard Nordgren, and Sebastian Ueckert. Extensive and automatic assumption assessment of pharmacometric models. *PAGE 27 (2018) Abstr 8754*, 2018.