Using **machine learning** for discovery of **mechanistic models from data** useable **in pharmacometrics**.

Data-Driven Discovery of Feedback Mechanisms in Acute Myeloid Leukaemia

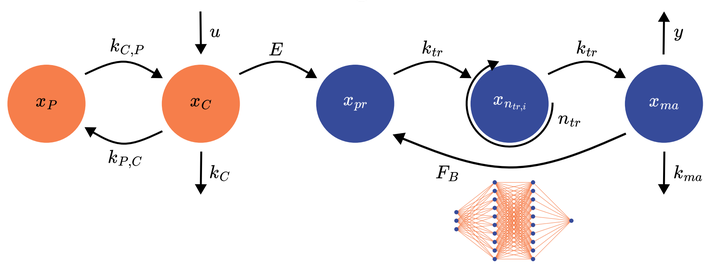
Carl Julius Martensen, Niklas Korsbo, Vijay Ivaturi, Sebastian Sager

PumasAI

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG

Motivation

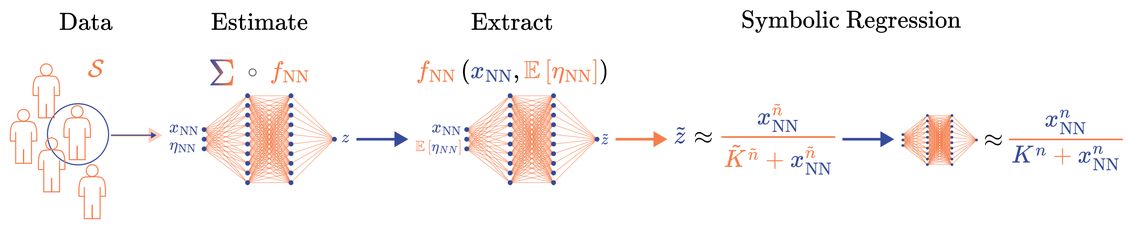
- We use Deep Nonlinear Mixed Effect Models to **effectively learn unknown submodels directly from data** in **DeepPumas** [1]
- We present a **method to distill symbolic expressions** from **longitudinal data** using the provided data-driven hypothesis of the machine learning model.



Graphical overview of the model using n_{tr} transition compartments as reported in [2]. We replace the feedback F_B using a neural network and recover an analytic function.

Methods and Materials

- We assume that all subjects share a **global** model structure, intersubject variability is fully captured by the **individual** parameters
- To recover the functional form, we perform symbolic regression on $\tilde{z}$, the mapping of the neural network using the individual input x_{NN} and the typical value random effects $\mathbb{E}[\eta_{\text{NN}}]$
- We investigate the validity of our approach on **a real-world dataset**, consisting of 23 patients who exhibit **high-dose Ara-C chemotherapy** [2] comparing it to the classical Friberg model [3] using **Pumas** [4]



Results

Baseline

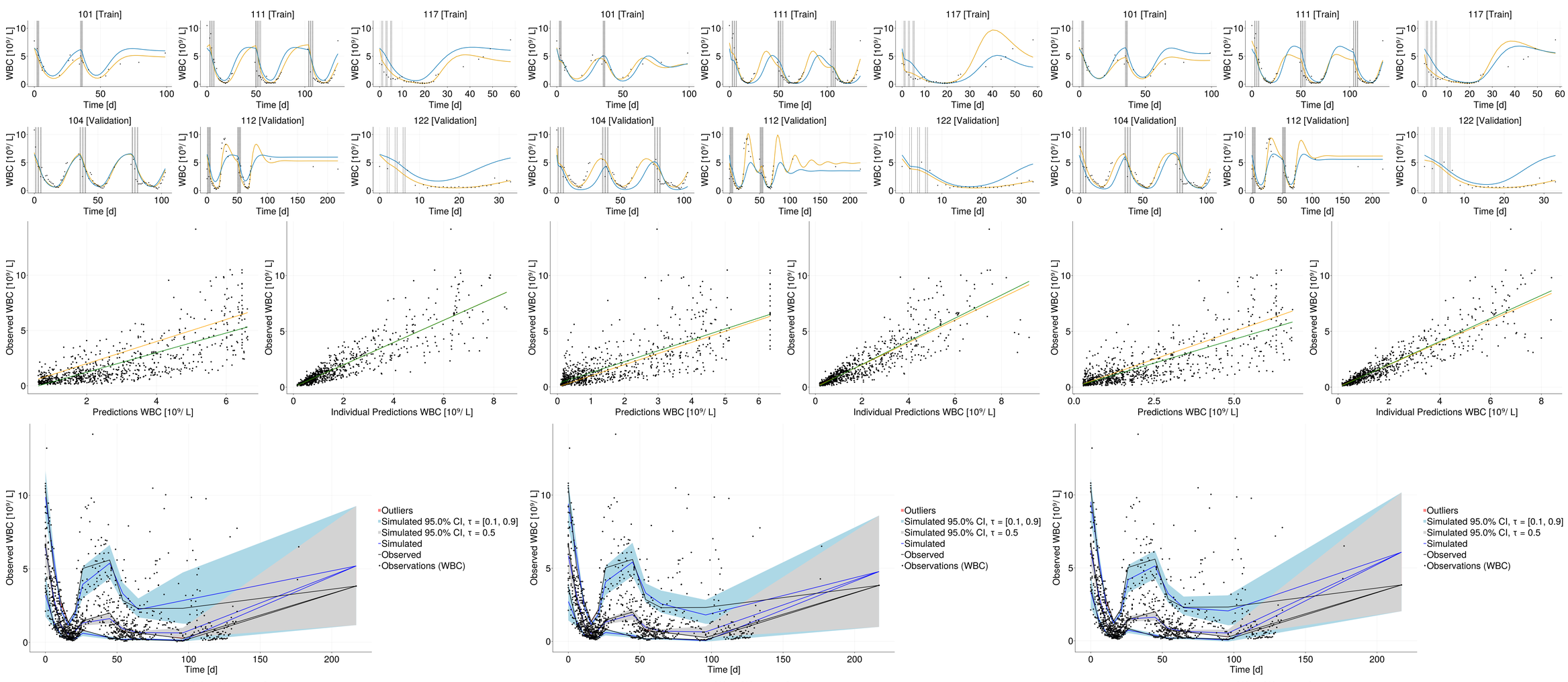
$F_B(x_{ma}) = \left(\frac{B}{x_{ma}}\right)^{\gamma}$ $n_{tr} = 1$

DeepNLME

$F_B(x_{ma}) = f_{\text{NN}}(x_{ma}, \eta_{\text{NN}})$ $n_{tr} = 3$

Discovered

$F_B(x_{ma}) = \frac{p_1}{p_2 + x_{ma}}$ $n_{tr} = 1$



	BIC	Loglikelihood
Training	1490.0	-705.0
Validation	571.0	-252.0

	BIC	Loglikelihood
Training	2032.7	-667.0
Validation	876.8	**-145.6**

	BIC	Loglikelihood
Training	**1404.4**	**-655.2**
Validation	**510.3**	-215.7

Experiments

- We used $n_{tr} = 1, 2, 3$ transition compartments for each model and selected the best
- The 23 subjects of the used dataset have been divided into a training and validation set (18, 5 subjects) stratified by treatment statistics
- Each DeepNLME model has been fitted ten times, varying the network parameters
- The best performing DeepNLME model on the training data has been used for symbolic regression for each n_{tr}
- The shown results have been selected over the different models via BIC of the validation set
- The shown VPC are derived using 300 simulations of the full population per model

Conclusion

- We showed that an **algorithmic recovery of mathematical expressions for longitudinal data** is possible **using DeepNLME**
- The found model shows a **similar structural properties** to the Friberg model but is **numerically favourable** and resembles a Hill equation
- The **performance of both the DeepNLME model and the discovered model is comparable to the baseline** in terms of loglikelihood and **underline the usefulness of both approaches**

References

[1] C. V. Rackauckas and V. Ivaturi, "Method and apparatus for automating models for individualized administration of medicaments," US20240087704A1, 2024.

[2] F. Jost, E. Schalk, K. Rinke, T. Fischer, and S. Sager, "Mathematical models for cytarabine-derived myelosuppression in acute myeloid leukaemia," *PLoS ONE*, vol. 14, no. 7, e0204540, 2019, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0204540.

[3] L. E. Friberg, A. Henningsson, H. Maas, L. Nguyen, and M. O. Karlsson, "Model of Chemotherapy-Induced Myelosuppression With Parameter Consistency Across Drugs," *JCO*, vol. 20, no. 24, pp. 4713–4721, 2002, ISSN: 0732-183X, 1527-7755. DOI: 10/d6jkkv.

[4] C. Rackauckas, Y. Ma, A. Noack, V. Dixit, P. K. Mogensen, S. Byrne, S. Maddhashiya, J. B. S. Calderón, J. Nyberg, J. V. S. Gobburu, and V. Ivaturi, *Accelerated Predictive Healthcare Analytics with Pumas, a High Performance Pharmaceutical Modeling and Simulation Platform*, 2020. DOI: 10.1101/2020.11.28.402297.

Preprint available