Amsterdam UMC
Universitair Medische Centra

The Neural Mixed Effects Algorithm:

Leveraging Machine Learning for Pharmacokinetic Modelling

PAGE 2021 - Stuart Beal Methodology Session I

A. Janssen

symphony
NWO-NWA 2018

OPTICLOT

Classical Pharmacokinetic analysis

$$\mathcal{D} = \{\mathbf{x}_i, \mathbf{y}_i\}_{i \in \mathrm{N}}, \mathbf{y}_i \backsim \mathcal{N}(\mu_i, Var[\mathbf{Y}])$$

Solve a (compartment) model $\mu = A(t; \mathbf{p})$ using MLE:

$$\mathcal{L}(\mathbf{p}) = \sum_i \ln Var[\mathbf{Y}] + \frac{(\mu - \mathbf{y}_i)^2}{Var[\mathbf{Y}]}$$

How can we improve our accuracy?

Classical Pharmacokinetic analysis

$$\mathcal{D} = \{\mathbf{x}_i, \mathbf{y}_i\}_{i\in\mathrm{N}},\ \mathbf{y}_i \backsim \mathcal{N}(\mu_i, Var[\mathbf{Y}])$$

Solve a (compartment) model $\mu = A(t;\ \mathbf{p})$ using MLE:

$$\mathcal{L}(\mathbf{p}) = \sum_i \ln Var[\mathbf{Y}] + \frac{(\mu - \mathbf{y}_i)^2}{Var[\mathbf{Y}]}$$

How can we improve our accuracy?

1). Add population information to $\mathbf{p}$ as a function of $\mathbf{x}_i$ to generate $\mathrm{E}[\mathbf{p}_i]$ (Pop-PK).

Classical Pharmacokinetic analysis

$$\mathcal{D} = \{\mathbf{x}_i, \mathbf{y}_i\}_{i \in \mathrm{N}}, \ \mathbf{y}_i \backsim \mathcal{N}(\mu_i, Var[\mathbf{Y}])$$

Solve a (compartment) model $\mu = A(t; \ \mathbf{p})$ using MLE:

$$\mathcal{L}(\mathbf{p}) = \sum_i \ln Var[\mathbf{Y}] + \frac{(\mu - \mathbf{y}_i)^2}{Var[\mathbf{Y}]}$$

How can we improve our accuracy?

1). Add population information to $\mathbf{p}$ as a function of $\mathbf{x}_i$ to generate $\mathrm{E}[\mathbf{p}_i]$ (Pop-PK).

2). Generate individualized parameters $\widehat{\mathbf{p}}_i$ based on measurements $\mathbf{y}_i$ and IIV (NLME model).

Classical Pharmacokinetic analysis

- Typical PK parameters are estimated from data:

$$\mathrm{E}[\mathbf{p}_i] = f(\mathbf{x}_i;\ \theta)$$

- Choosing the correct representation of $\mathrm{E}[\mathbf{p}_i]$ is difficult.
- Based on basic functions, performing hypothesis tests for each covariate.
- Opportunity for Machine Learning (ML) based techniques.

Historical performance of ML methods

- In 1995[1]: Used Neural network for direct prediction of μ_i.
- "Neural networks made peak serum concentration predictions ... with statistically less bias and comparable precision" compared to a NLME model.
- Several limitations:
  - Only predicts concentrations at specific timepoints
  - Cannot handle complex dosing schemes
  - No measure of uncertainty or IIV
- Limitations essentially hold true for all ML methods

[1]: Brier et al. (1995). *Pharmaceutical Research*

Deep compartment models

- Chen et al. 2018[2]: automatic differentiation of ODE solvers.
- Instead of directly predicting μ a neural network ϕ predicts $\mathrm{E}[\mathbf{p}_i]$:

$$\mathrm{E}[\mathbf{p}_i] = \phi(\mathbf{x}_i; \mathcal{W})$$

$$\widehat{\mu}_i = A(t;\ \mathrm{E}[\mathbf{p}_i])$$

- Benefits:
  - Reliable solution at any timepoint
  - Supports any dosing scheme
  - Interpretable!

[2]: Chen et al. (2018). *32nd International Conference on Neural Information Processing Systems*

Neural mixed effects algorithm

Based on a model of IIV: $\widehat{\mathbf{p}}_i = \mathrm{E}[\mathbf{p}_i] \cdot \exp \eta_i\,,\ \eta_i \backsim \mathcal{N}_p(\Omega),$

and measurement errors: $\widehat{\mathbf{y}}_i = \mu_i + \epsilon_i,\ \epsilon_i \backsim \mathcal{N}(\sigma^2)$

the model parameters become $\Theta = [\mathcal{w},\ \Omega, \sigma^2].$

We minimize the first order approximation of $\mathcal{L}$:

$$\underset{\Theta}{\operatorname{argmin}}\, \mathcal{L}_{FO} = \sum_i \ln C_i + \frac{(\mu_i - \mathbf{y}_i)^2}{C_i}$$

$$Var[\mathbf{Y}] \approx C_i = G_i \Omega G_i{}' + H_i \sigma^2 H_i{}'$$

Simulation experiment

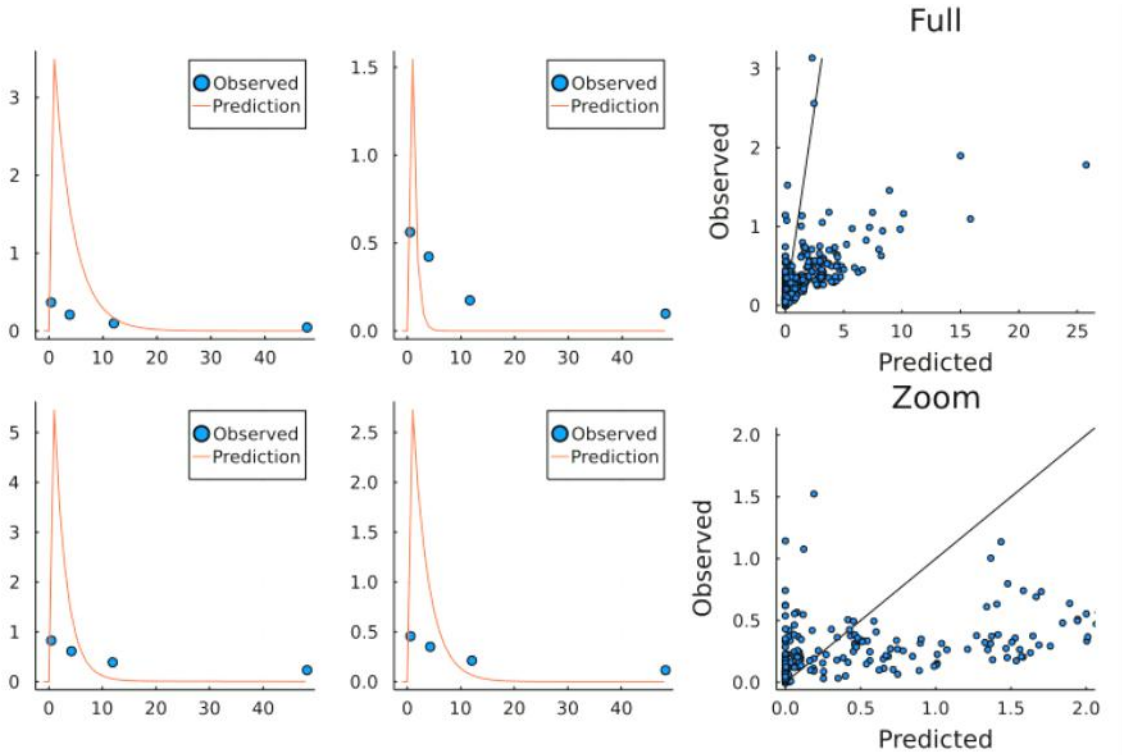
- Based on NLME model of 119 Haemophilia A patients treated with factor VIII (FVIII) perioperatively[3]. Concentrations were modeled using a two compartment model.
- Covariates in PK model were weight, age, blood group, and intensity of surgical procedure.
- Concentrations were based on $\widehat{\mathbf{p}}_i$ with $\Omega = \begin{bmatrix} 0.129 & 0.043 \\ 0.043 & 0.0705 \end{bmatrix}$ and additive error $\sigma^2 = 0.04^2$.
- Single bolus dose of 1500 IU at t=0
- Simulated concentrations at t=5min, 30min and every hour until t=48.

[3]: Hazendonk et al. (2016). *Haematologica*

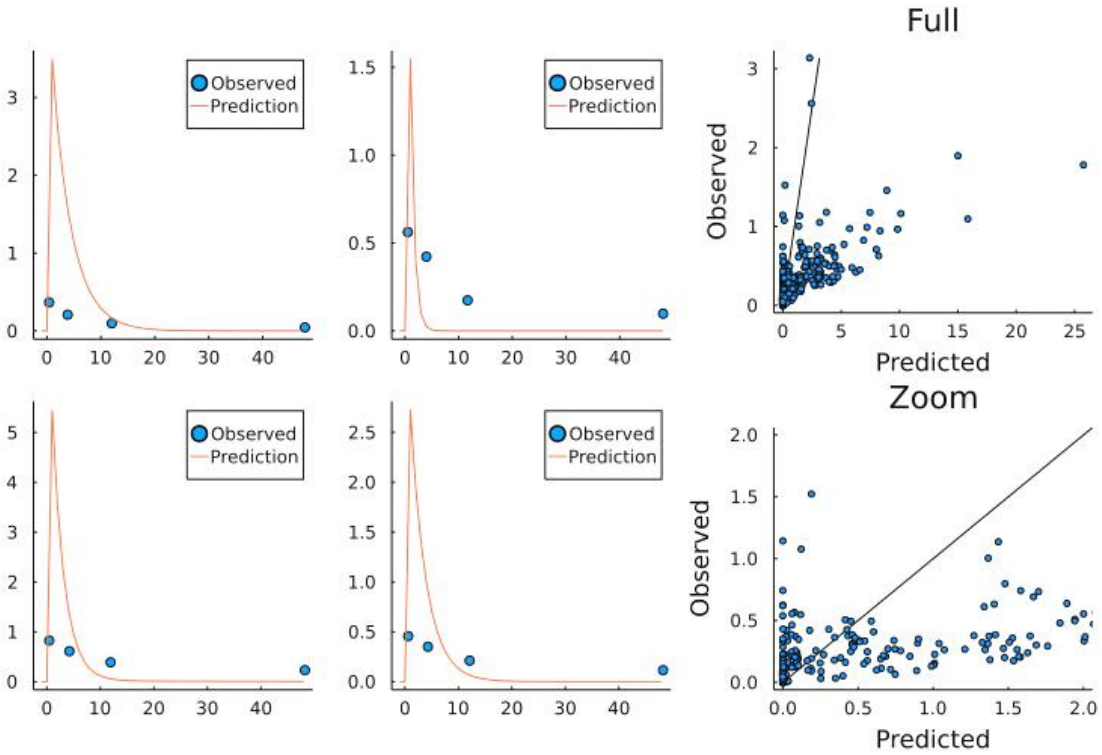
Training the NME model

- Train on 75 simulated individuals, validate on 425.
- Training measurements were limited to t = [0.5h, 4h, 24h, 48h]
- Trained using the same covariates as in the NLME model.
- Estimated IIV on clearance and central volume parameters
- Accuracy defined as %-age of predictions within 0.05 $IUmL^{-1}$ of ‘true’ simulated concentration.

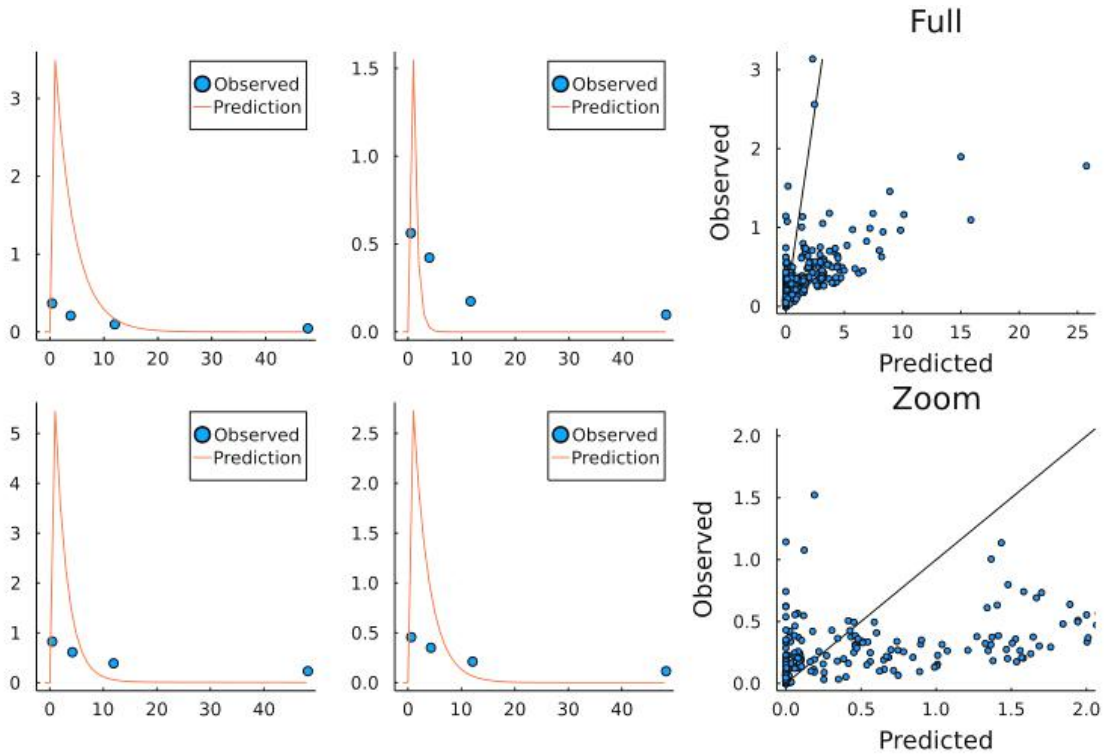
NME Simulation results



NME Simulation results



NME Simulation results



Final Accuracy:

Typical predictions: 68.5%

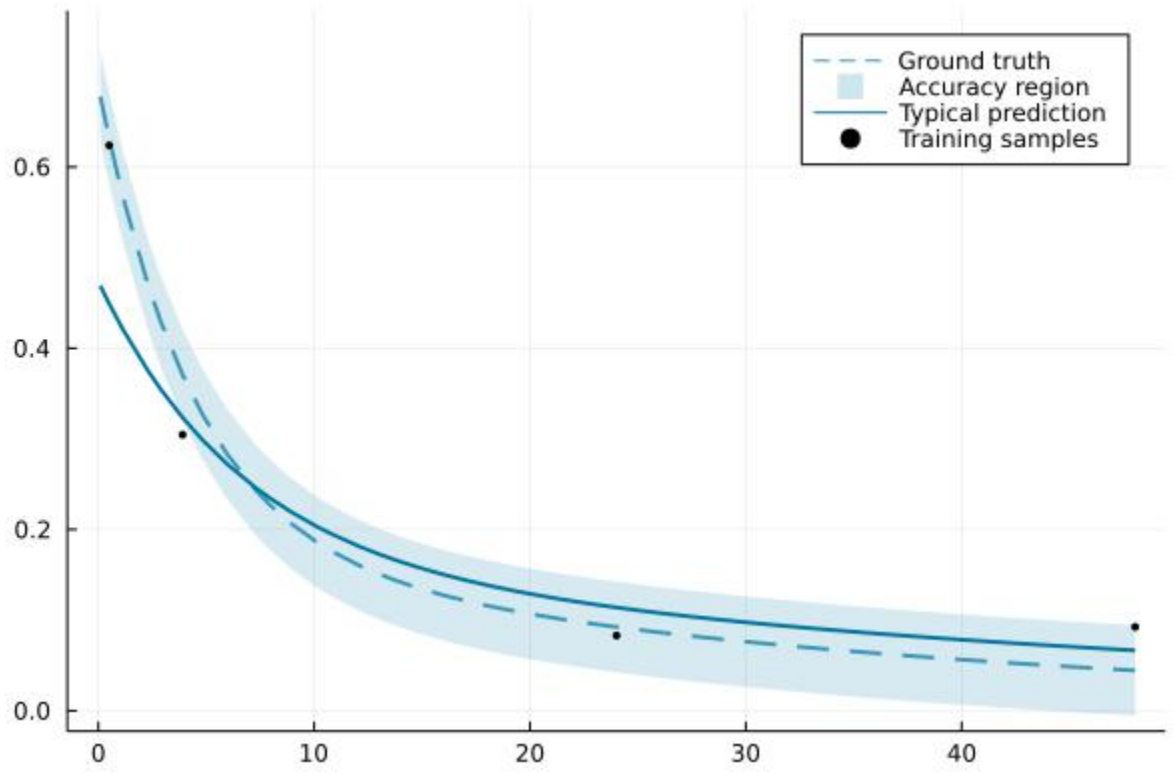
Individual predictions: 92%

$$\widehat{\Omega} = \begin{bmatrix} 0.126 & 0.0634 \\ 0.0634 & 0.0677 \end{bmatrix}, \widehat{\sigma}^2 = 0.0385^2$$

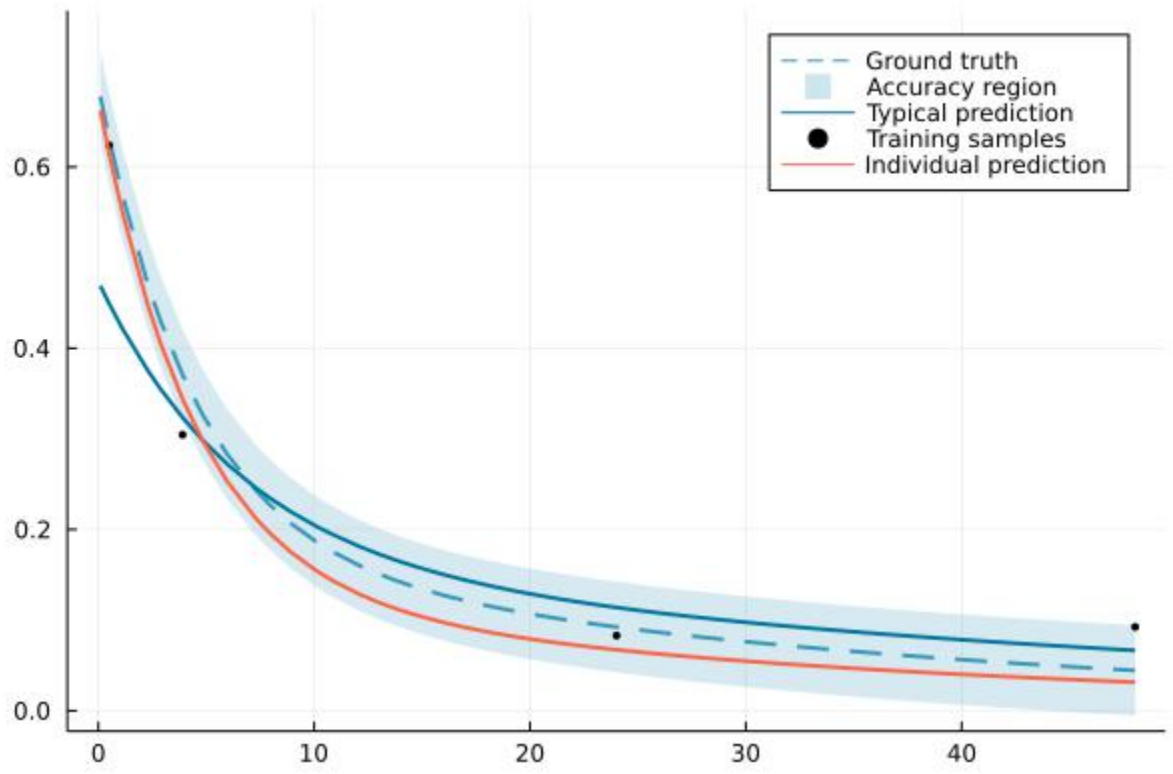
True parameters:

$$\Omega = \begin{bmatrix} 0.129 & 0.043 \\ 0.043 & 0.0705 \end{bmatrix}, \sigma^2 = 0.04^2$$

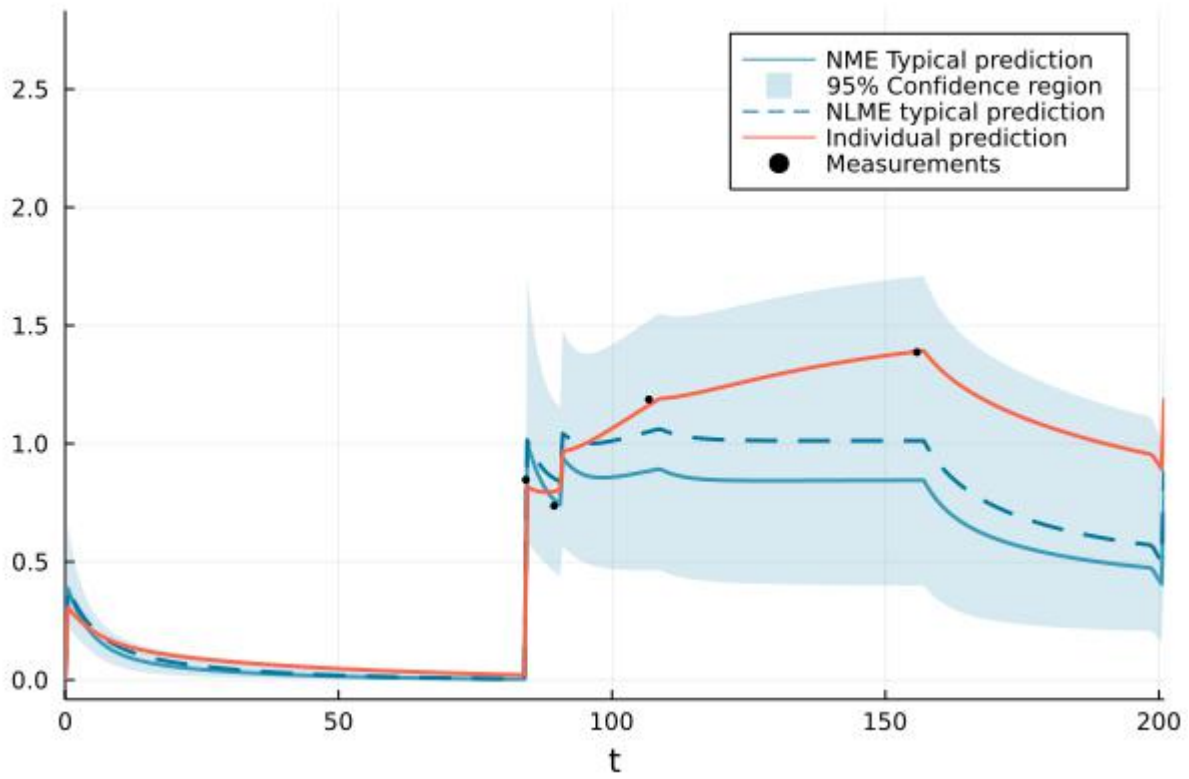
NME: Results for a typical individual



NME: Results for a typical individual



Easily transferred to more complex cases



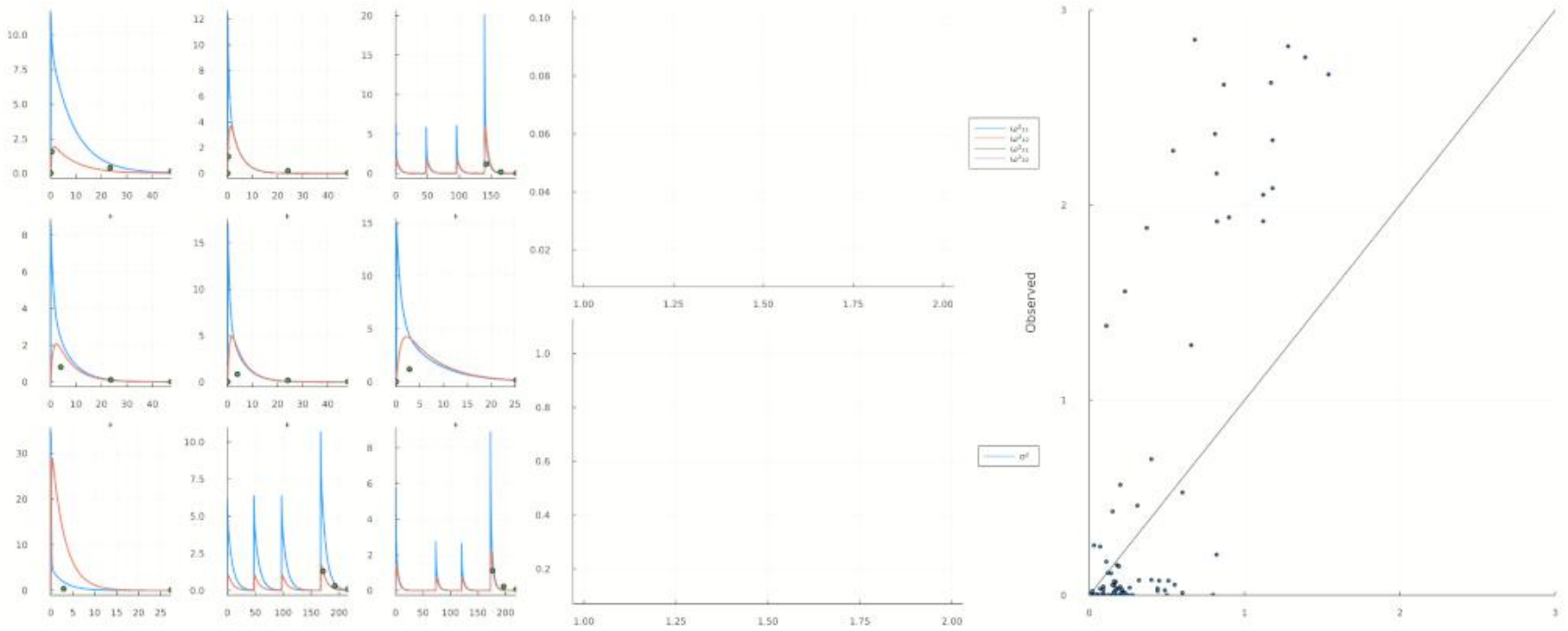
Complex dosing example:

- t=0: 500 IU
- t=84: 1250 IU
- t=84.08: 100 IU/hr (6.67hrs)
- t=90.75: 250 IU & 120 IU/hr (18hrs)
- t=108.75: 100 IU/hr (48.25hrs)
- etc...

Test on real world dataset

- PK profiles of 65 Haemophilia A patients
- Measurements at t=[4, 24, 48]
- Individuals received a single bolus dose of 50 $IUkg^{-1}$ FVIII concentrate
- Covariates used: weight, age, blood group, VWF antigen levels
- Split in train (N=46) and test (N=19) set

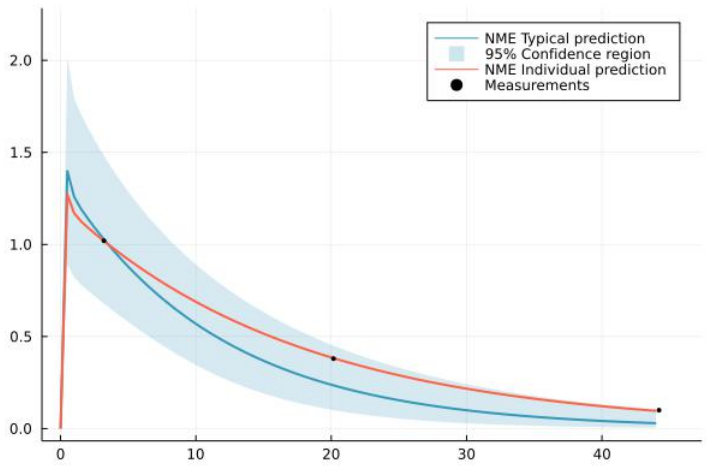
Monitoring training

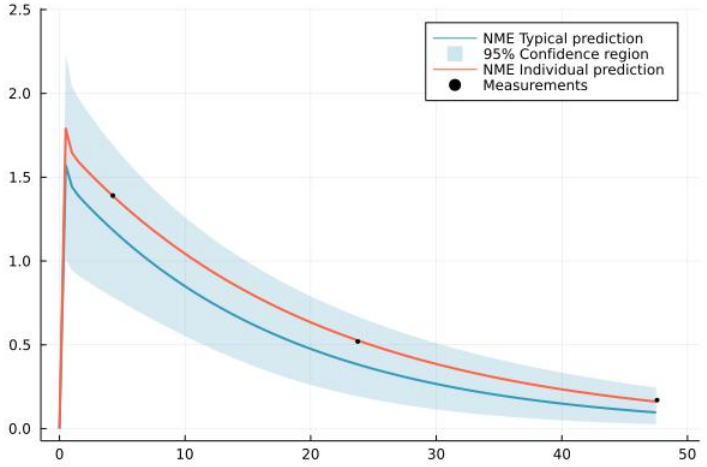


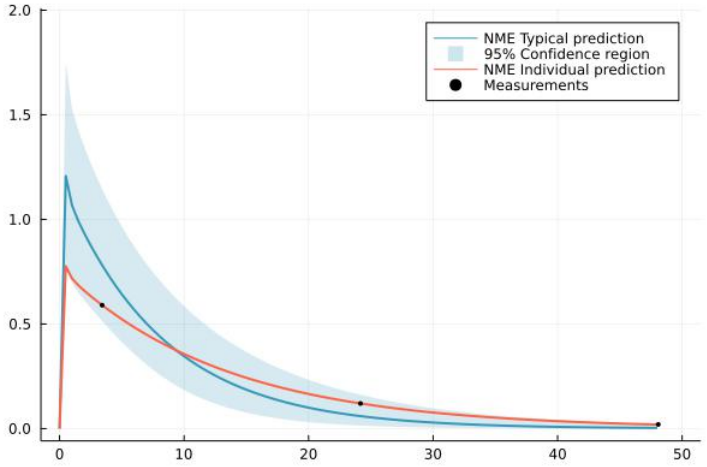
NME real world test result

- Final train accuracy: 47.6%, Test accuracy: 50.0%
- $\widehat{\Omega} = \begin{bmatrix} 0.0556 & 0.0414 \\ 0.0414 & 0.0781 \end{bmatrix}$, $\widehat{\sigma}^2 = 0.0597^2$

3 Examples from test set:

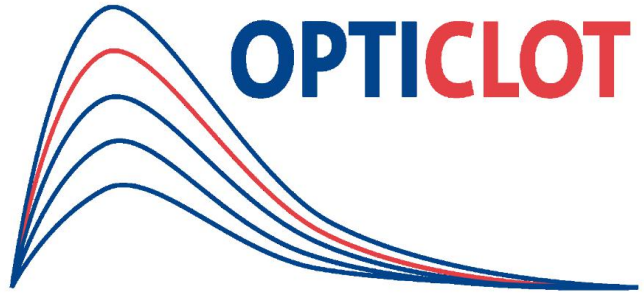






Conclusion

- NME algorithm can predict concentrations for any dosing scheme, and allows for reliable extrapolation of solution.
- Offers automated covariate implementation while also estimating IIV.
- Reduces data requirement and overfitting by using prior knowledge in the form of a compartment model.
- Model output is familiar to pharmacologists and interpretable!



Ask any questions in the Q&A chat or send me a mail at a.janssen@amsterdamumc.nl

Sponsors:

- Erasmus MC
- Sanquin Research & Sanquin Diagnostics
- Amsterdam UMC
- University Medical Center Groningen
- University Medical Center Utrecht
- Leiden University Medical Center
- Radboud UMC
- Netherlands Hemophilia Patient Society (NVHP)
- Dutch Hemophilia Treaters Society (NVHB)
  Dutch Society on Thrombosis and Hemostasis (NVTH)
- Bayer B.V. & CSL Behring B.V. & Sobi Benelux

Acknowledgments:

Authors: F. Leebeek, M.H. Cnossen, R. Mathtôt

Symphony consortium

OPTI-CLOT study group