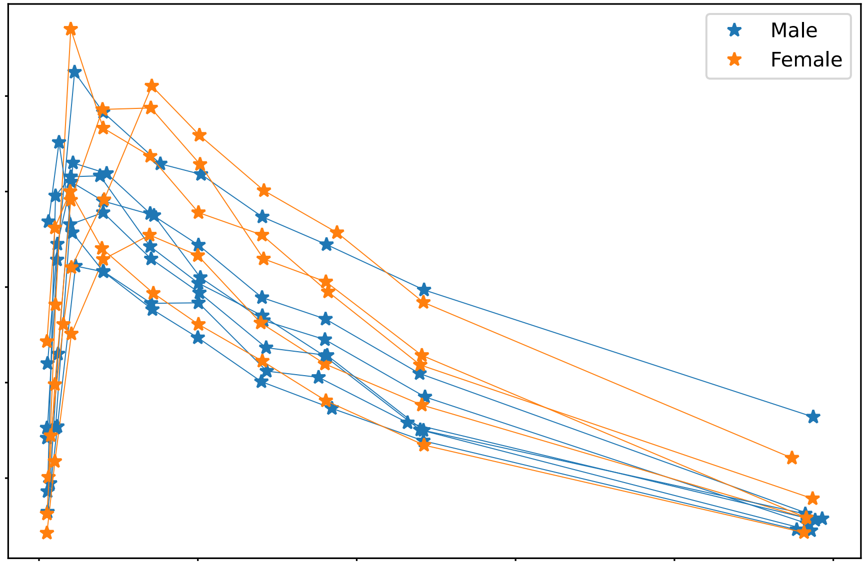
Redefining Parameter Estimation and Covariate Selection Via Variational Autoencoders: One run is all you need

Jan Rohleff, D. Bräm, F. Bachmann, U. Nahum, B. Steffens, M. Pfister, G. Koch and **J. Schropp**

University of Konstanz, 05.06.2025
Page 2025, Thessaloniki, Greece

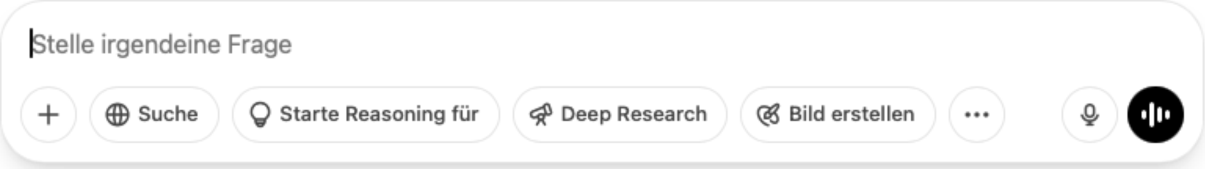


Machine Learning and Artificial Intelligence (AI)

ChatGPT

How powerful is Generative AI?

Generative AI is very powerful — and it's getting stronger fast, growing at an unprecedented pace that's transforming entire industries in real time.



1

Generative AI

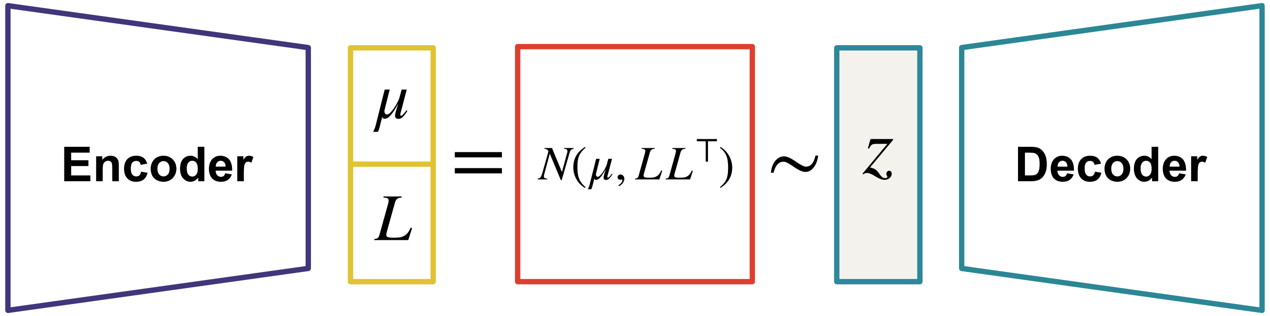
Generative AI has grown rapidly in the last few years.

Machine Learning and Artificial Intelligence (AI)

1

Generative AI

Generative AI has grown rapidly in the last few years.

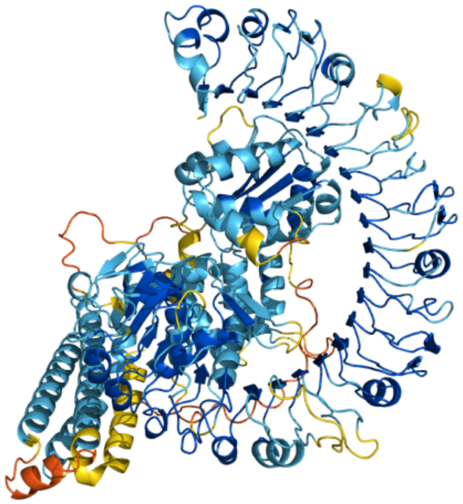


2

Variational Autoencoder

Bayesian framework within generative AI.

Machine Learning and Artificial Intelligence (AI)



Article

Highly accurate protein structure prediction with AlphaFold

https://doi.org/10.1038/s41586-021-03819-2

Received: 11 May 2021

Accepted: 12 July 2021

Published online: 15 July 2021

Open access

Check for updates

John Jumper[1,4]✉, Richard Evans[1,4], Alexander Pritzel[1,4], Tim Green[1,4], Michael Figurnov[1,4], Olaf Ronneberger[1,4], Kathryn Tunyasuvunakool[1,4], Russ Bates[1,4], Augustin Žídek[1,4], Anna Potapenko[1,4], Alex Bridgland[1,4], Clemens Meyer[1,4], Simon A. A. Kohl[1,4], Andrew J. Ballard[1,4], Andrew Cowie[1,4], Bernardino Romera-Paredes[1,4], Stanislav Nikolov[1,4], Rishub Jain[1,4], Jonas Adler[1], Trevor Back[1], Stig Petersen[1], David Reiman[1], Ellen Clancy[1], Michal Zielinski[1], Martin Steinegger[2,3], Michalina Pacholska[1], Tamas Berghammer[1], Sebastian Bodenstein[1], David Silver[1], Oriol Vinyals[1], Andrew W. Senior[1], Koray Kavukcuoglu[1], Pushmeet Kohli[1] & Demis Hassabis[1,4]✉

1

Generative AI

Generative AI has grown rapidly in the last few years.

2

Variational Autoencoder

Bayesian framework within generative AI.

3

AlphaFold

VAE that predicts 3D structure of proteins from their amino acid sequence

Bayesian Framework within AI: Variational Autoencoder

AI

Non-linear Mixed Effects (NLME) Modeling

f

Structural Model and Observations

Consider N subjects, for $i \in \{1,\ldots,N\}$:

$$\frac{d}{dt} y_i(t) = f(t, y_i(t), \phi_i),$$
$$y_i(0) = y_{i,0}\,.$$

Observations:

$$x_{ij} = g(y_i(t_{ij}), \phi_i) + \epsilon_{ij}, \quad \epsilon \sim \mathcal{N}(0, a^2)\,.$$

β

Population Approach

$$h(\phi_i) = z_i, \quad z_i = z_{\text{pop}} + \beta c_i + \eta_i \quad \text{with} \quad \eta_i \sim \mathcal{N}(0, \Omega)\,.$$

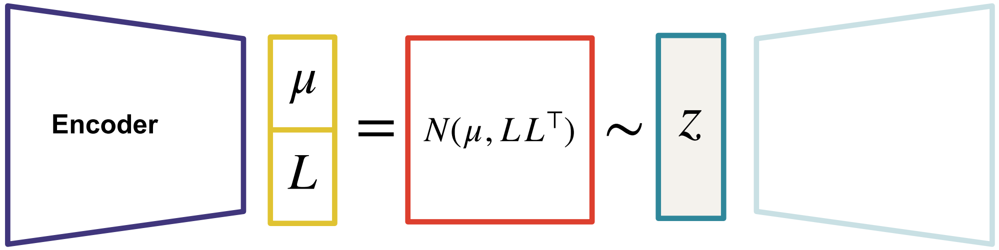
$\mathcal{L}\mathcal{L}$

Population fit

Maximize the Log-Likelihood $\log p(x)$ with respect to $(z_{pop}, \beta, \Omega, a)$.

Variational Autoencoders in NLME Modeling

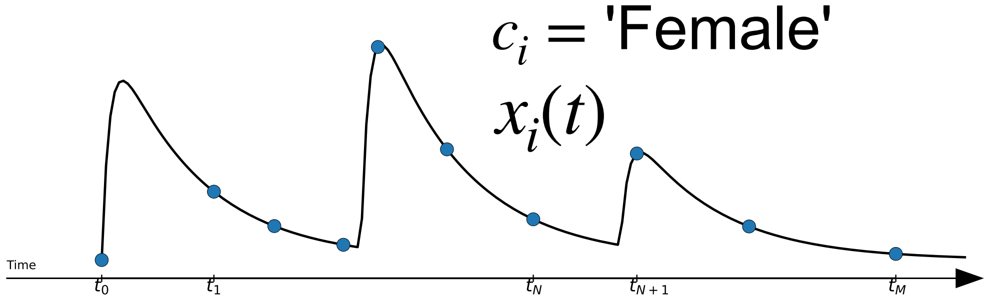
1. Encode the data of Individual i



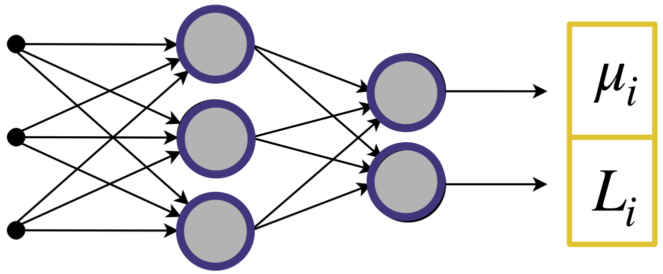
The encoder is parametrized by a Long Short-Term Memory (LSTM) neural network.

$$(\mu_i, L_i) = \text{LSTM}(x_i, c_i), \quad \text{for } i = 1,\ldots,N$$

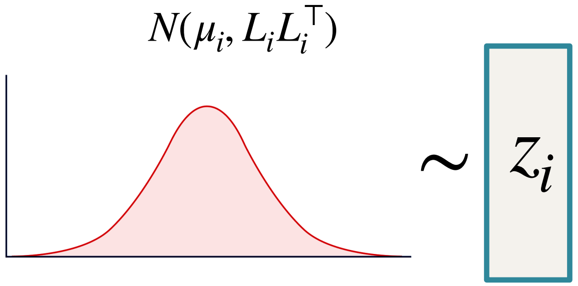
Data



LSTM Encoder



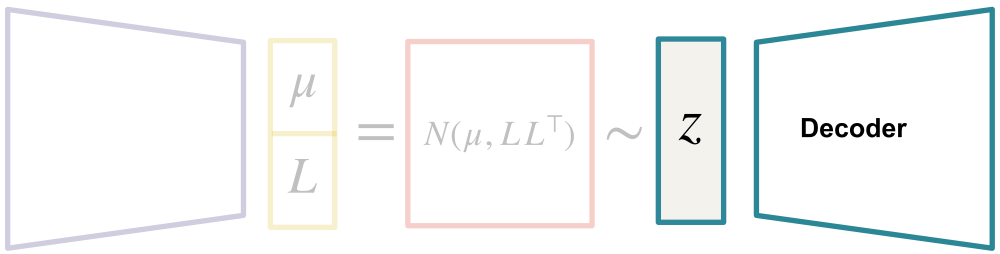
Distribution of individual Parameters

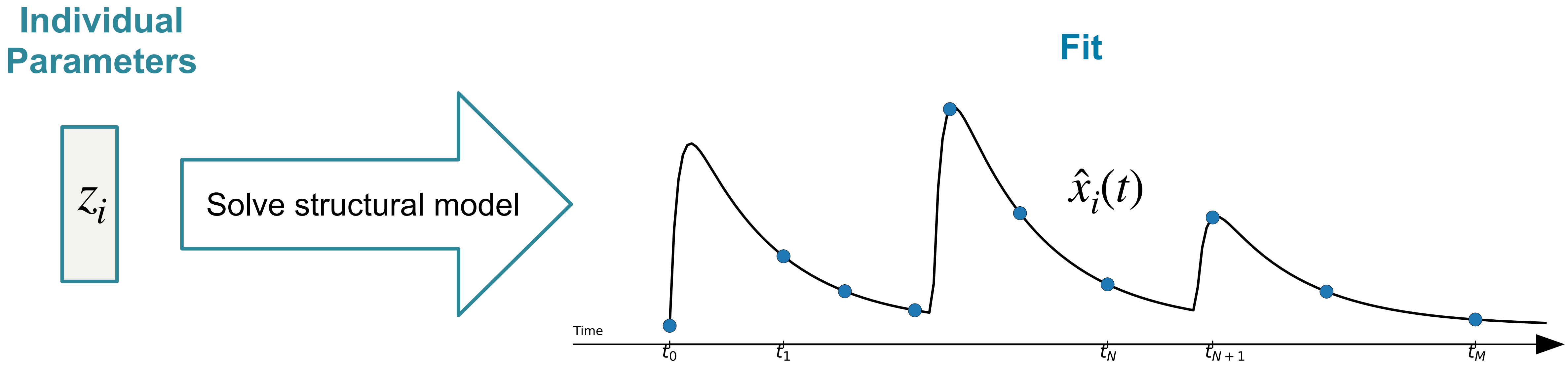


Variational Autoencoders in NLME Modeling

2. Decode individual Parameter z_i by solving the Model

The decoder takes the individual parameter z_i and solves the model equations.

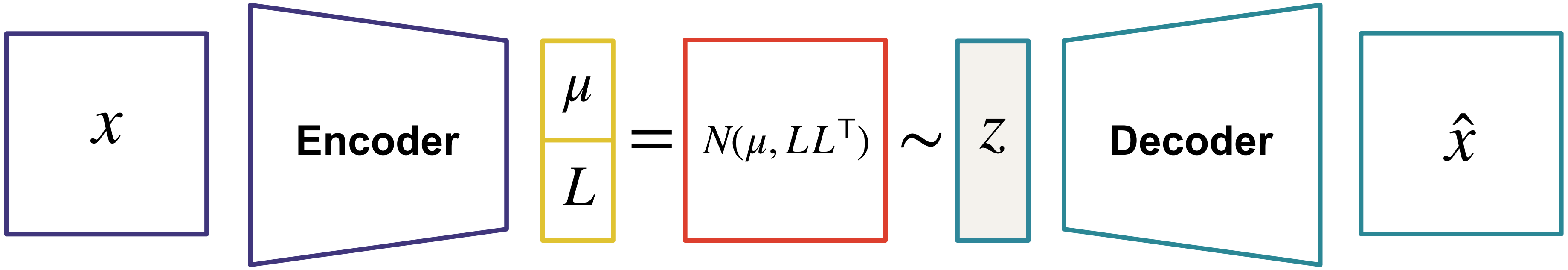




Variational Autoencoders in NLME Modeling

3. Maximize the Evidence Lower Bound (ELBO) Loss function

$$\mathscr{L}_{\psi}^{ELBO}(x) = \sum_{i=1}^{N} \mathbb{E}_{z_i \sim q_{\psi}(\cdot | x_i)}[\log p(x_i | z_i)] - \sum_{i=1}^{N} D_{KL}(q_{\psi}(z_i | x_i) \,||\, p(z_i))$$



Theophylline - Example

Data

Consider a population of $N = 12$ patients.

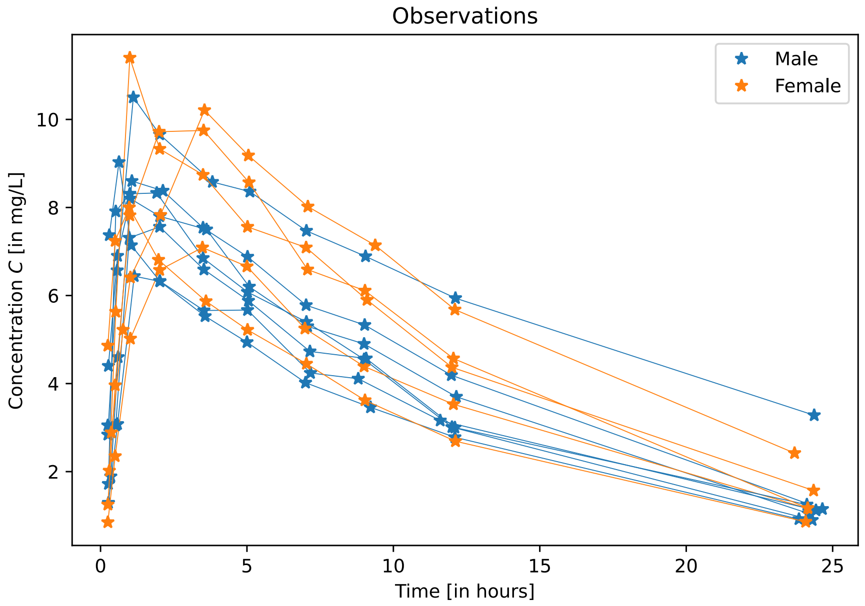
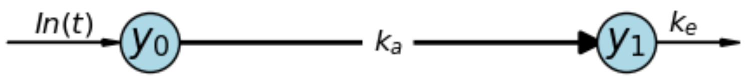


Figure: Theophyilline Data Set by Boeckmann, A. J., Sheiner et al. (1994). NONMEM Users Guide: Part V., University of California, San Francisco.

Model

Linear **1-compartment** PK model with **absorption**.



The model is given by:

$$C(t) = \frac{Dk_a}{V(k_a - k_e)} \left(e^{-k_e t} - e^{-k_a t} \right)$$

Log-Normal distribution:

$$V_i = V_{pop} \cdot e^{\eta_{V,i}}, \quad k_{a,i} = k_{a,pop} \cdot e^{\eta_{k_a,i}}, \quad k_{e,i} = k_{e,pop} \cdot e^{\eta_{k_e,i}}.$$

VAE produces Parameter Estimates consistent with SAEM

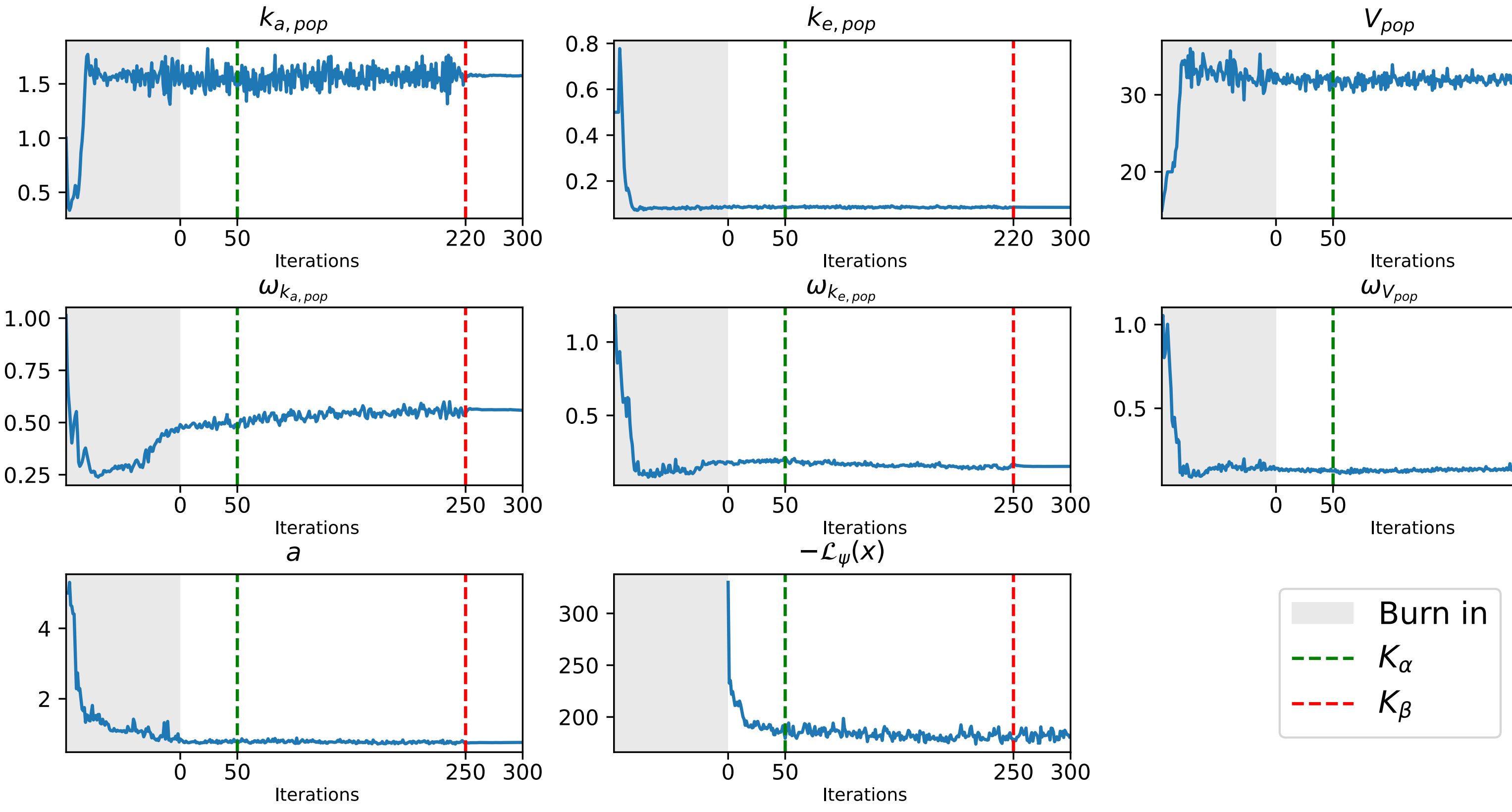
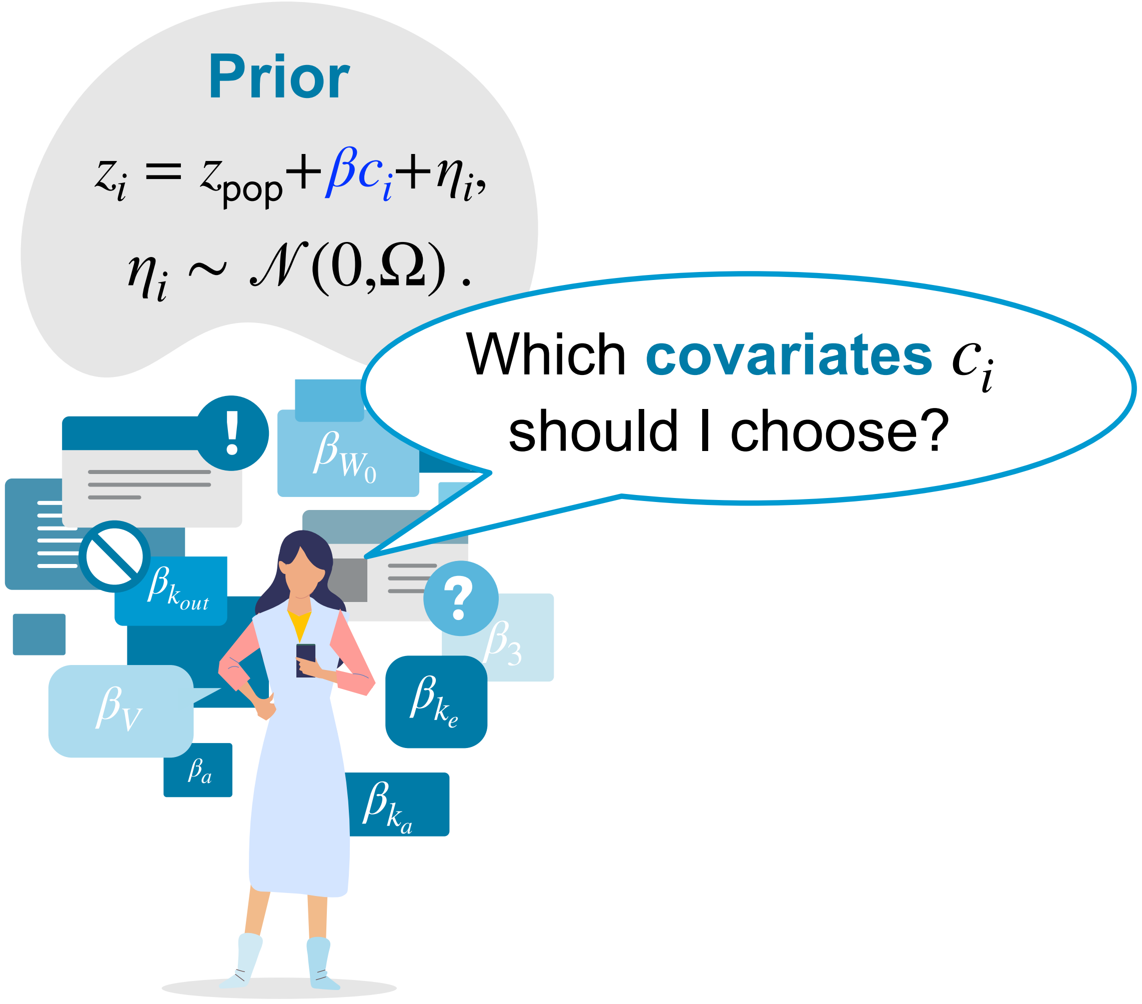


Figure: Convergence of the **VAE**.

	VAE	SAEM
Fixed Effects		
$k_{a,pop}$	1.60	1.61
$k_{e,pop}$	0.085	0.085
V_{pop}	32.00	31.98
Variance		
ω_{k_a}	0.57	0.63
ω_{k_e}	0.15	0.15
ω_V	0.14	0.15
Error Model		
a	0.72	0.73
Stat. Criteria		
$-2\mathscr{LL}$	338.9	338.6

Table: Theophylline Model results

Covariate Selection



Covariate Selection

Standard

- SAEM + Selection Tools, i.e. **SAMBA, COSSAC, SCM**, etc.
- SAEM solves NLME problems for a fixed choice of covariates.
- BICc is computed afterwards:

$$-2\mathscr{L}\mathscr{L}(x) + \log(N)k$$

- k number of covariates,
- N subjects.

- Iteratively **multiple runs.**

Prior

$$z_i = z_{\text{pop}} + \beta c_i + \eta_i,$$

$$\eta_i \sim \mathcal{N}(0,\Omega)\,.$$

Covariate Selection - Dream

Standard

- SAEM + Selection Tools, i.e. **SAMBA, COSSAC, SCM**, etc.
- SAEM solves NLME problems for a fixed choice of covariates.
- BICc is computed afterwards:

$$-2\mathscr{L}\mathscr{L}(x) + \log(N)k$$

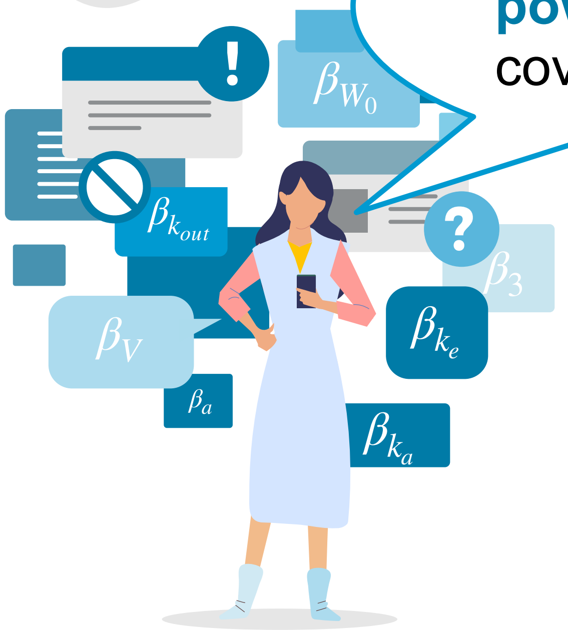
- k number of covariates,
- N subjects.

- Iteratively **multiple runs.**

Prior

$$z_i = z_{\text{pop}} + \beta c_i + \eta_i,$$

$$\eta_i \sim \mathcal{N}(0, \Omega)$$



Is it really necessary to **solve** the NLME problem **multiple times**?

Why not use the **AI power** of the VAE for covariate selection?

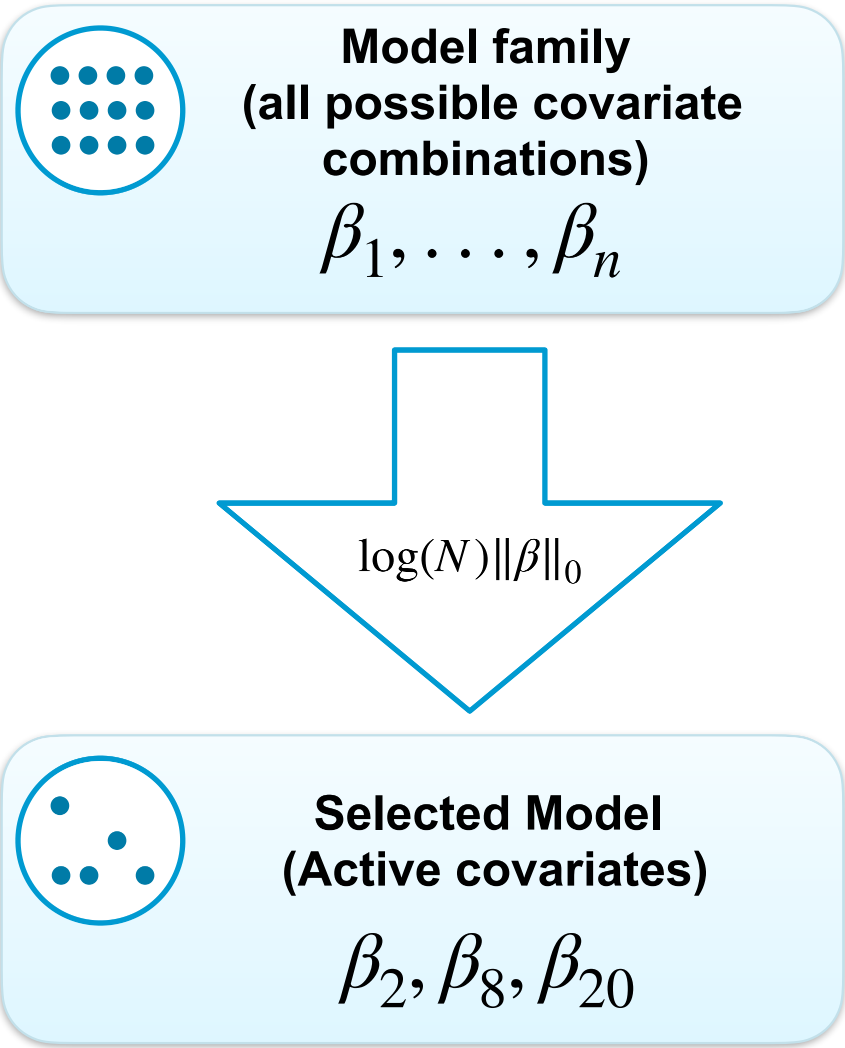
Can we estimate **parameters** and select **covariates simultaneously** in **one run**?

Using **VAEs**, we can offer determination of **population parameters** and select **covariates** **all at ones.**

One Run is all you need.

Simultaneously determination of Population Parameters and Covariates

Analyze a family of models.



1. Initialization: Model family, β vector of covariate effects.
Number of active covariates: $\|\beta\|_0 = \#\{i,\ \beta_i \neq 0\}$.

2. Optimize the BICc-ELBO:

$$z_{pop}^{(k+1)}, \beta^{(k+1)} = \arg\max_{z_{pop}^{(k)}, \beta^{(k)}} \{2\mathscr{L}_{\psi}^{ELBO}(x) + \log(N)\|\beta^{(k)}\|_0\}$$

One run is all you need.

Weight of Neonates - Example

Data

Consider a population of $N = 2425$ patients.

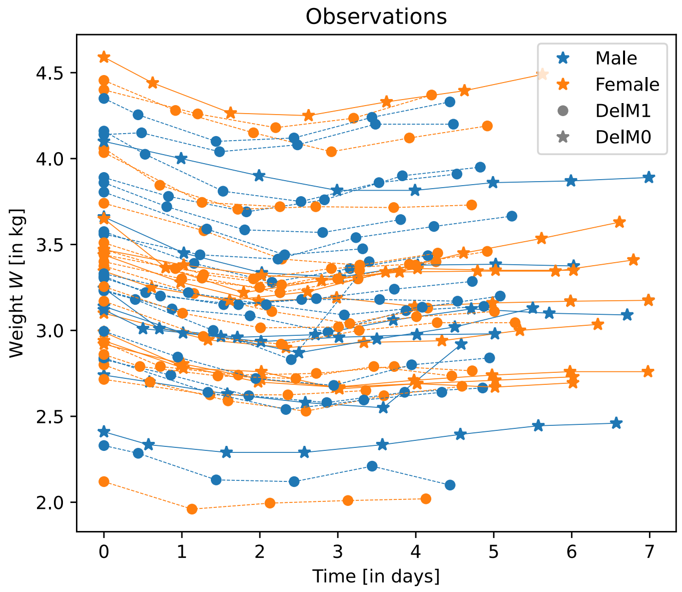


Figure: Neonates Dataset by Wilbaux M., Kasser S., Gromann J. et al. (2019). Personalized Weight Change Prediction in the First Week of Life. Clin Nutr. 38(2):689-696.

Model

The model is given by

$$\frac{d}{dt}W(t) = k_{prod}(t) - k_{el}(t)W(t), \text{ for } (t, T], \quad W(0) = W_0$$

where

$$k_{prod}(t) = \frac{k_{in}}{1 + \exp(-2(t - T_{lag}))} \quad \text{and} \quad k_{el}(t) = k_{out}\left(1 - \frac{t}{T_{50} + t}\right).$$

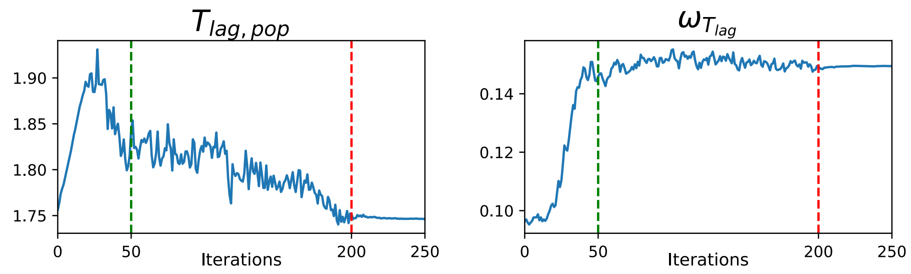
Five Parameters $(k_{in}, k_{out}, T_{50}, T_{lag}, W_0)$ **log-normally** distributed.

Five covariates $(Sex, DelM, GA, Mage, Para_2)$
=> 2^{25} possible covariate model combinations.

All at Once: Populations Parameters and Covariate Selection

In every iteration the **population parameters and covariate model** are updated.

Population Parameters



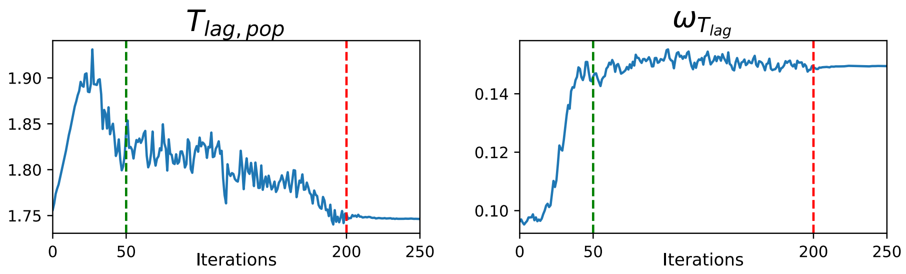
All at Once: Populations Parameters and Covariate Selection

In every iteration the **population parameters** **and** **covariate model** are updated.

Population Parameters

+

Covariate Model



Inactive Covariates

Active Covariates

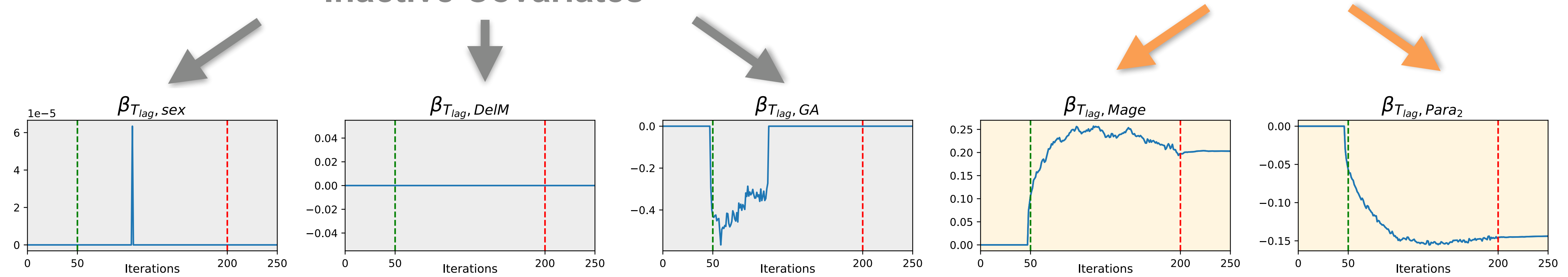


Figure: Convergence of the Covariates of T_{lag}.

VAE Convergence - Covariates

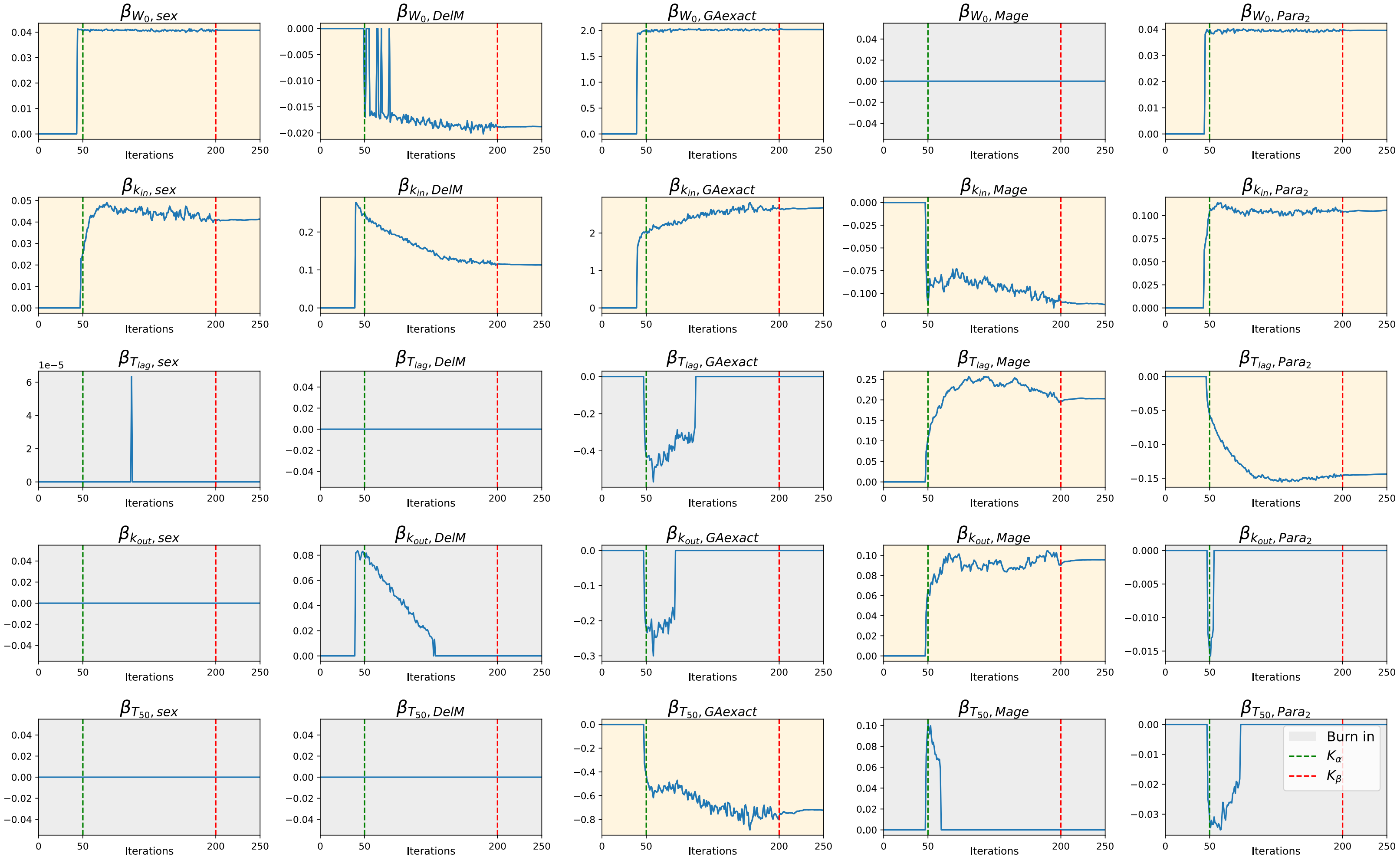


Figure: Convergence of the Covariates.

Neonates - Selected Covariates

	VAE	COSSAC	SAMBA	SCM
$-2\mathscr{LL}$	146154	146132	146177	**146123**
BICc	146351	146329	146374	**146305**

Table: Statistic Criteria for the VAE, COSSAC, SAMBA and SCM Covariates Selection.

	β_{sex}	β_{DelM}	$\beta_{GAexact}$	β_{Mage}	β_{Para2}
W_0	$\color{magenta}\checkmark$	$\color{magenta}\checkmark$	$\color{magenta}\checkmark$		$\color{magenta}\checkmark$
k_{in}	$\color{green}\checkmark\color{red}\checkmark$	$\color{magenta}\checkmark$	$\color{magenta}\checkmark$	$\color{green}\checkmark\color{red}\checkmark\color{orange}\checkmark$	$\color{magenta}\checkmark$
T_{lag}			$\color{red}\checkmark\color{orange}\checkmark\color{blue}\checkmark$	$\color{green}\checkmark\color{red}\checkmark\color{orange}\checkmark$	$\color{magenta}\checkmark$
k_{out}		$\color{orange}\checkmark$		$\color{green}\checkmark\color{red}\checkmark$	$\color{blue}\checkmark$
T_{50}		$\color{orange}\checkmark$	$\color{green}\checkmark$	$\color{blue}\checkmark$	

- $\color{green}\checkmark$ VAE
- $\color{red}\checkmark$ COSSAC
- $\color{orange}\checkmark$ SAMBA
- $\color{blue}\checkmark$ SCM
- $\color{magenta}\checkmark$ All

Table: Selected Covariates for the VAE, COSSAC, SAMBA and SCM.

Neonates - Selected Covariates

	VAE	COSSAC	SAMBA	SCM
$-2\mathscr{LL}$	146154	146132	146177	**146123**
BICc	146351	146329	146374	**146305**
Runs	**1**	33	2	244

Table: Statistic Criteria for the VAE, COSSAC, SAMBA and SCM Covariates Selection.

	β_{sex}	β_{DelM}	$\beta_{GAexact}$	β_{Mage}	β_{Para2}
W_0	✓ All	✓ All	✓ All		✓ All
k_{in}	✓ VAE ✓ COSSAC	✓ All	✓ All	✓ VAE ✓ COSSAC ✓ SAMBA	✓ All
T_{lag}			✓ COSSAC ✓ SAMBA ✓ SCM	✓ VAE ✓ COSSAC ✓ SAMBA	✓ All
k_{out}		✓ SAMBA		✓ VAE ✓ COSSAC	✓ SCM
T_{50}		✓ SAMBA	✓ VAE	✓ SCM	

✓ VAE
✓ COSSAC
✓ SAMBA
✓ SCM
✓ All

Table: Selected Covariates for the VAE, COSSAC, SAMBA and SCM.

Conclusion

Generative AI → VAEs ← NLME Modeling

Present

$z_{pop} + \beta$ — Simultaneously Parameter Estimation + automated Covariate Selection

VAE-based NLME framework in Python

z_i — Prediction of parameters of new individuals

Future

Analysis of complex covariates

Totally or partially unknown ODE Models

Automated Modeling

Acknowledgments

Universität Konstanz

UKBB
kompetent & menschlich

Supervisor

Prof. Dr. J. Schropp
Department of Mathematics and Statistics · University of Konstanz

Research Team: UKBB and University of Konstanz

F. Bachmann
Department of Mathematics and Statistics · University of Konstanz

D. Bräm, B. Steffens, M. Pfister, G. Koch
Pediatric Pharmacology and Pharmacometrics, University Children's Hospital Basel (UKBB), University of Basel, Switzerland

U. Nahum
Institute of Biomedical Engineering and Medical Informatics, University of Applied Sciences and Arts, Northwestern Switzerland