Hierarchical Deep Compartment Modeling: A Workflow to Leverage Machine Learning for Hierarchical Pharmacometric Modeling

Ahmed Elmokadem[1], Kiersten Utsey[1], Eric Jordie[1], and Timothy Knab[1]
[1]*Metrum Research Group, Tariffville, CT, USA*

Abstract

Objectives: Population pharmacokinetic (PK) modeling is considered the standard approach to characterize the PK in a given population. This approach incorporates subject covariates to explain random variability in PK parameters and improve the predictive performance of the model. Covariate modeling can be a complex process since the underlying structural relationship between covariates and PK parameters is often unknown. Deep compartment modeling (DCM) was previously proposed to use machine learning techniques to automate the covariate modeling step [1]. DCM, however, does not evaluate any residual error or interindividual variability (IIV). This could lead to model misspecification and over-fitting. The work presented here demonstrates Hierarchical Deep Compartment Modeling (HDCM), which is an extension of DCM. These models utilize machine learning to characterize the relationship between covariates and PK parameters while evaluating varying levels of random effects using Bayesian inference.
Methods: Synthetic PK data were used to demonstrate the proposed HDCM framework. The data were generated from a two-compartment population PK model and were divided into training (10 subjects) and test (20 subjects) datasets. Standard covariate modeling was used to characterize the generated individual PK parameters. In the HDCM approach, an artificial neural network (ANN) was used to learn these covariate relationships. Random effects were introduced as measurement noise and IIV on the clearance (CL) parameter. The hierarchical Bayesian inference was carried out on the training dataset using the No-U-Turn Sampler (NUTS) algorithm. Further HDCM validation was carried out using the test PK dataset. Open-source Julia tools were utilized to build the model (SciML), train the ANN (Flux.jl), and run the Bayesian inference (Turing.jl).
Results: The proposed HDCM framework successfully inferred the PK model parameters as well as the random effects and ANN weights while quantifying the uncertainty around them. This was evident from the standard Bayesian model diagnostics, such as the convergence of multiple chains, the effective sample size (ESS), and the Gelman-Rubin statistic (Rhat). Posterior predictive checks (PPCs) demonstrated the ability of the model predictions to characterize the train PK dataset with an average normalized root mean square error (NRMSE) of 16.1%. PPCs also indicated that the model could characterize the test PK dataset, which improved confidence in the model predictions.
Conclusions: The HDCM framework introduced in this work was developed using open-source tools in Julia. Synthetic PK data were utilized to demonstrate the training of an ANN to learn the relationships between covariates and PK model parameters. The proposed approach was integrated with Bayesian inference to quantify the uncertainty around the model parameters. The convenience of the proposed HDCM framework makes it readily accessible to an audience interested in applying deep learning to hierarchical pharmacometric compartment models.

Methods

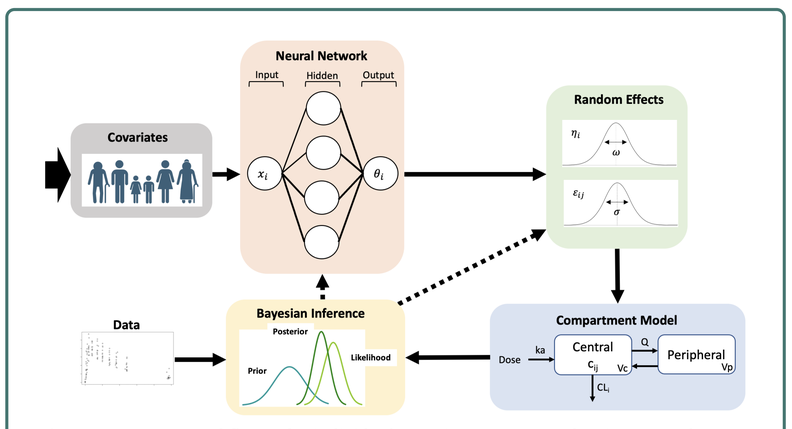

Figure 1. HDCM workflow. The individual covariates x_i are used as inputs to the ANN that would then output the typical individual parameters θ_i. IIV (η_i) and residual error (ϵ_{ij}) parameters are added to the hierarchical compartmental model structure. Model predictions are compared to the observed data within a Bayesian analysis framework to infer the posterior distributions of ANN parameters as well as the random effects.

Application

- Synthetic PK data for 30 subjects receiving a single oral dose of a drug. The PK was assumed to follow a two-compartment model distribution.
- The data was split into training (n = 10) and test (n = 20) datasets.
- The following 4 covariates were tested and used as inputs to the ANN: age, weight, EGFR, and albumin.
- The output from the ANN θ represented the typical individual values of the PK parameters: CL, V1, Q, V2, and ka.
- IIV was added to CL such that $CL_i = \theta_i.e^{\eta_i}$ where $\eta_i \sim N(0, \omega^2)$.
- The compartmental model was built using DifferentialEquations.jl.

ANN structure

- Input layer composed of 4 nodes (4 covariates).
- Hidden layer composed of 6 nodes.
- Output layer composed of 5 nodes (typical PK parameters).
- Activation functions were *swish* and *CELU* for the output layer.
- The output layer was initialized at the maximum a posteriori (MAP) Bayes estimate of the parameters from an initial naive pooled fit.
- Flux.jl and DiffEqFlux.jl were used to build the ANN and integrate with ODE solvers.

Statistical model

Likelihood:

$$c_{ij} \sim N(\hat{c}_{ij}, \sigma^2)$$

where c_{ij} = concentration of the drug for individual i at timepoint j and $\hat{c}_{ij}$ = the corresponding prediction.

Prior distributions:

$$\omega \sim half - Cauchy(0, 0.5)$$
$$\sigma \sim half - Cauchy(0, 0.5)$$
$$w \sim N(0, 0.75)$$

where ω and σ represent the standard deviations of IIV and residual error, respectively. w represents the ANN weights.

Bayesian inference

The No-U-Turn Sampler (NUTS) was used to draw 3 chains of the posterior samples (500 warmup and 500 sampling) with an acceptance ratio of 0.65. Turing.jl was used for Bayesian inference.

Results

The HDCM framework was successfully applied to train an ANN to characterize the functional relationship between the tested covariates and the typical PK parameters. The approach was evaluated using standard Bayesian inference diagnostics including trace and density plots showing convergence of all chains to the same distributions for the select parameters (Figure 2). The posterior estimates for the select parameters are shown in Table 1 with the Markoc chain Monte Carlo (MCMC) diagnostics: the Gelman-Rubin statistic ($\hat{R}$) showing values > 1.05 and effective sample size (ESS) showing reasonably large values for ESS_{bulk} and ESS_{tail}. Table 2 displays the interpretable posterior predictions for the typical PK parameter values together with the uncertainty around the estimates. PPCs showed good characterization of the train (Figure 3) and test (Figure 4) data. The latter indicated that there was no overfitting with HDCM.

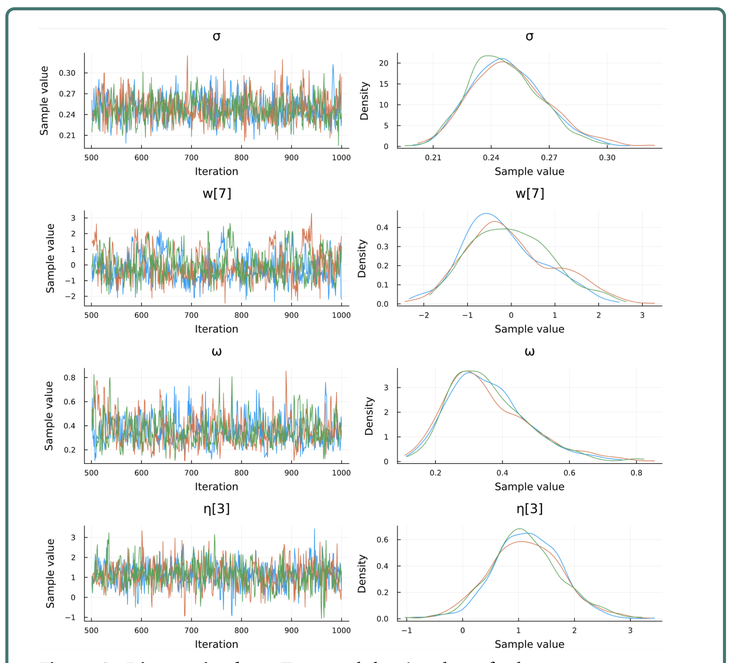

Figure 2. Diagnostic plots. Trace and density plots of select parameters: σ = residual error, w = NN weight, ω = standard deviation for IIV on CL, and η = individual random effect on CL. Number of samples = 500, number of chains = 3.

Parameter	Mean (95% CI)	$\hat{R}$	ESS_{bulk}	ESS_{tail}
σ	0.25 (0.22, 0.29)	1.002	887	1013
w[7]	-0.06 (-1.65, 1.96)	1.004	260	316
ω	0.356 (2.24, 4.65)	1.004	427	529
η[3]	1.14 (-0.07, 2.41)	1.007	819	891

Table 1. Parameter table. CI = credible interval, σ = residual error, w = NN weight, ω = standard deviation for IIV on CL, and η = individual random effect on CL. $\hat{R}$ = Gelman-Rubin statistic, ESS_{bulk} and ESS_{tail} are the effective sample size at bulk and tail of the distribution, respectively.

Parameter	Mean (95% CI)	SD (95% CI)
CL (L/h)	3.64 (2.79, 4.7)	1.25 (0.471, 2.29)
V1 (L)	69.3 (60.4, 76.5)	29.6 (19.7, 38.2)
Q (L/h)	3.3 (2.24, 4.65)	1.16 (0.303, 2.67)
V2 (L)	70.9 (50.9, 106.0)	18.6 (5.44, 41.9)
ka (/h)	2.27 (1.58, 3.61)	0.47 (0.144, 1.21)

Table 2. PK parameter table. SD = standard deviation, CI = credible interval, CL = central clearance, V1 = central volume, Q = intercompartmental clearance, V2 = peripheral volume, ka = absorption rate constant.

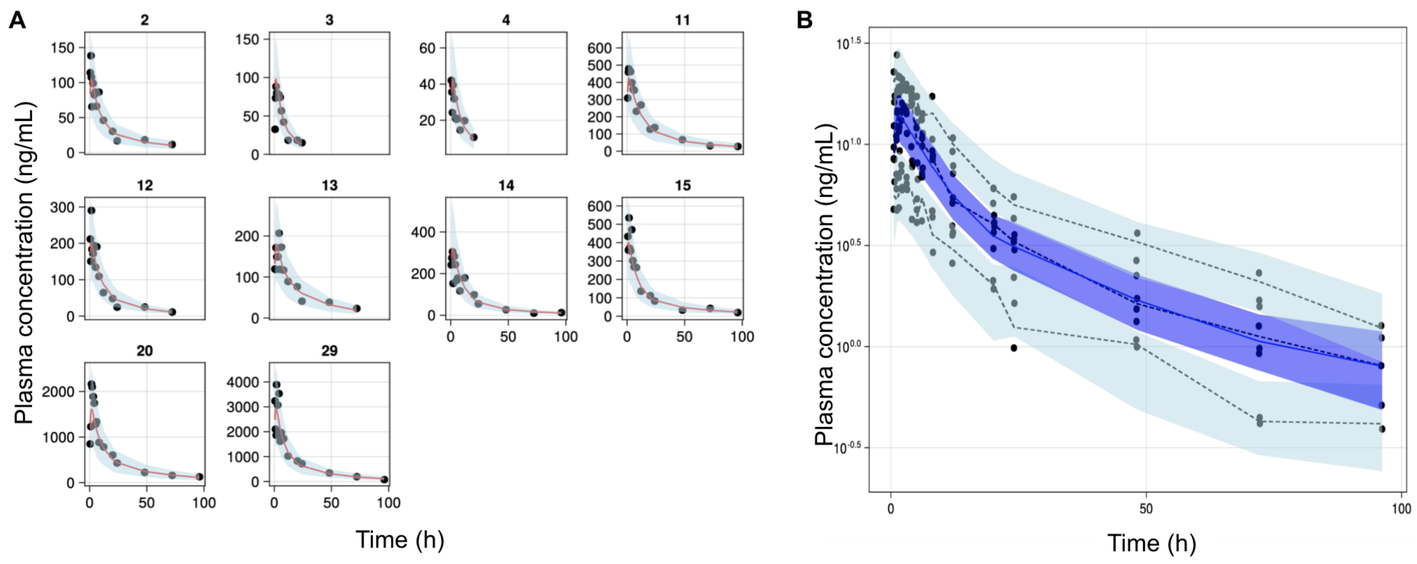

Figure 3. PPCs for train data. Individual (A) and summary (B) PPCs for train data. Black dots = observed data, red lines = median of predictions, blue bands = 95% CI for the predictions around distribution statistics, namely, median in (A) and (B), and 5% and 95% of predictions in (B). Dashed lines in (B) are medians through 5% and 95% of predictions.

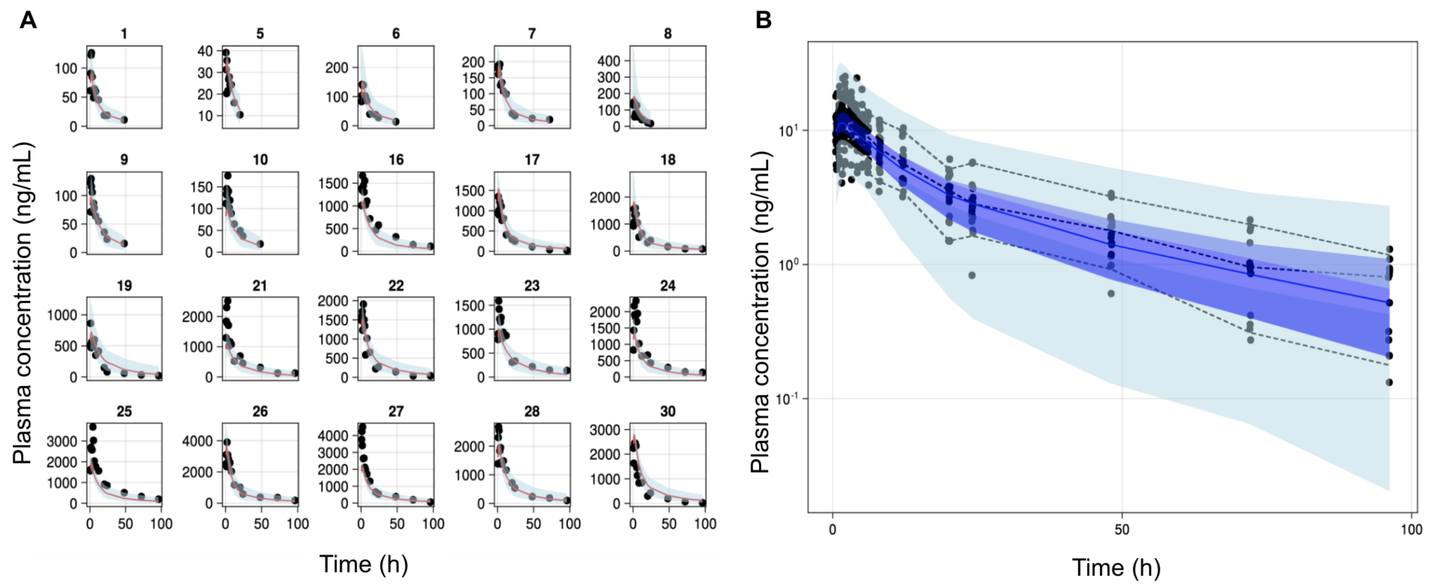

Figure 4. PPCs for test data. Individual (A) and summary (B) PPCs for test data. Black dots = observed data, red lines = median of predictions, blue bands = 95% CI for the predictions around distribution statistics, namely, median in (A) and (B), and 5% and 95% of predictions in (B). Dashed lines in (B) are medians through 5% and 95% of predictions.

Conclusion

- This work introduced the HDCM workflow that integrates population PK modeling with deep learning to learn the structural relationship between an individual's covariates and their PK parameter values.
- Bayesian inference was implemented to quantify the uncertainty around the trained neural network and to infer IIV and residual random effects in the hierarchical model structure.
- The proposed framework helps avoid overfitting and provides interpretability of model outputs.

References

[1] Janssen, A., Leebeek, F.W.G., Cnossen, M.H., Mathôt, R.A.A., Fijnvandraat, K., Coppens, M., Meijer, K., Schols, S.E.M., Eikenboom, H.C.J., Schutgens, R.E.G., Beckers, E.A.M., Ypma, P., Kruip, M.J.H.A., Polinder, S., Tamminga, R.Y.J., Brons, P., Fischer, K., Galen, K.P.M., Nieuwenhuizen, L., Driessens, M.H.E., Vliet, I., Lock, J., Hazendonk, H.C.A.M., Moort, I., Heijdra, J.M., Goedhart, M.H.J., Al Arashi, W., Preijers, T., Jager, N.C.B., Bukkems, L.H., Cloesmeijer, M.E., Collins, P.W., Liesner, R., Chowdary, P., Millar, C.M., Hart, D., Keeling, D. and for the OPTI-CLOT study group and SYMPHONY consortium. Deep compartment models: A deep learning approach for the reliable prediction of time-series data in pharmacokinetic modeling. *CPT Pharmacometrics Syst. Pharmacol.* (2022).