

TOWARD AUTONOMOUS PHARMACOMETRICS: A MULTI AGENT FRAMEWORK FOR STRUCTURE DISCOVERY, COVARIATE SELECTION AND GRAY-BOX MODELING

Berrah R 1 , Woillard JB 1,2,3

1 Pharmacology & Transplantation, INSERM U1248, Limoges University, Limoges, FRANCE
2 Department of Pharmacology, Toxicology and Pharmacovigilance, CHU Limoges, Limoges, FRANCE
3 Fédération Hospitalo-Universitaire Survival Optimization in Organ Transplantation (FHU SUPORT), Limoges, FRANCE

Background

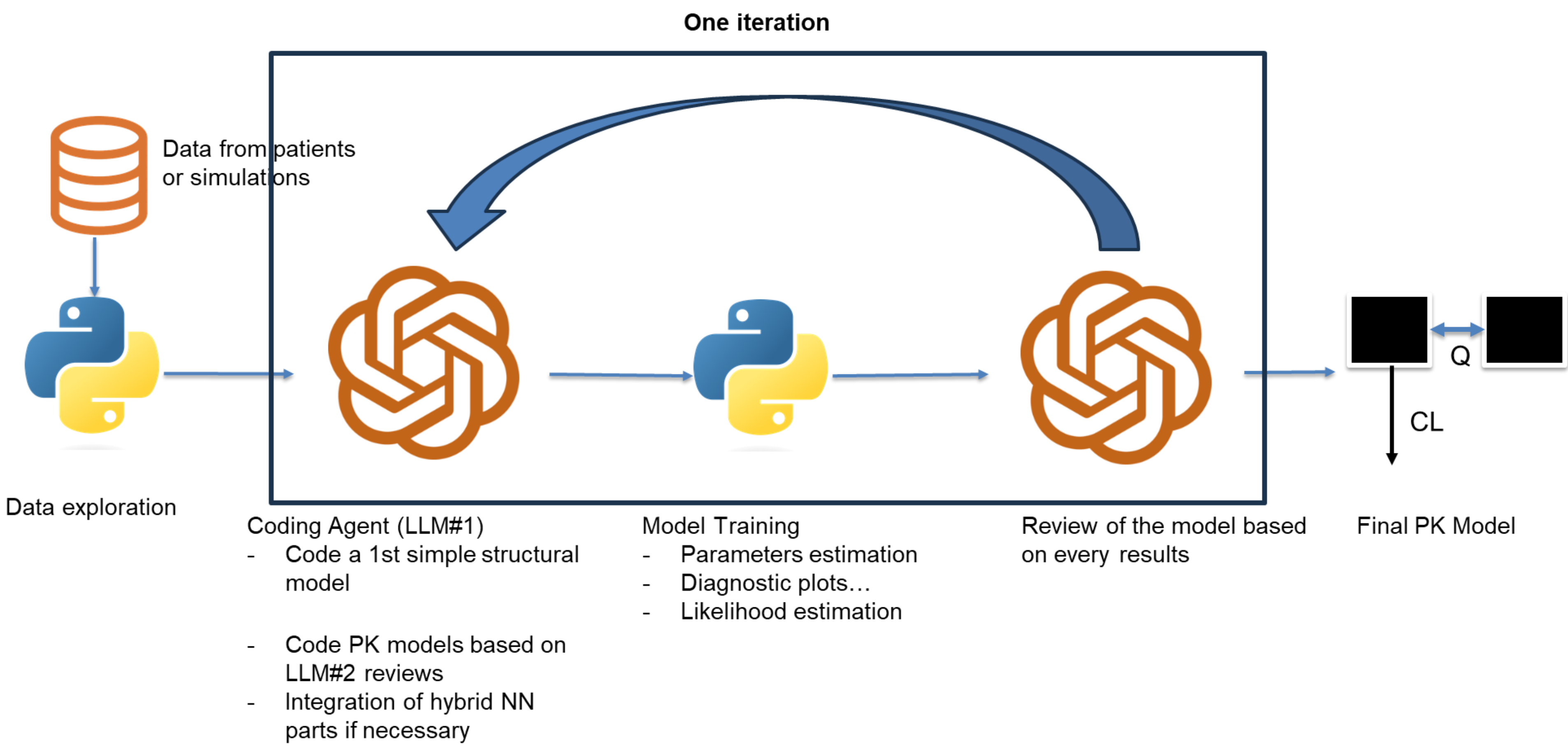
PopPK modeling is central to model-informed precision dosing, but model development remains iterative and expert-dependent. Classical compartmental models are interpretable but rigid, whereas pure ML is flexible but less mechanistic. We propose an autonomous multi-agent framework built around a custom SAEM engine.

Objective :

Turn the PopPK loop into an autonomous pipeline

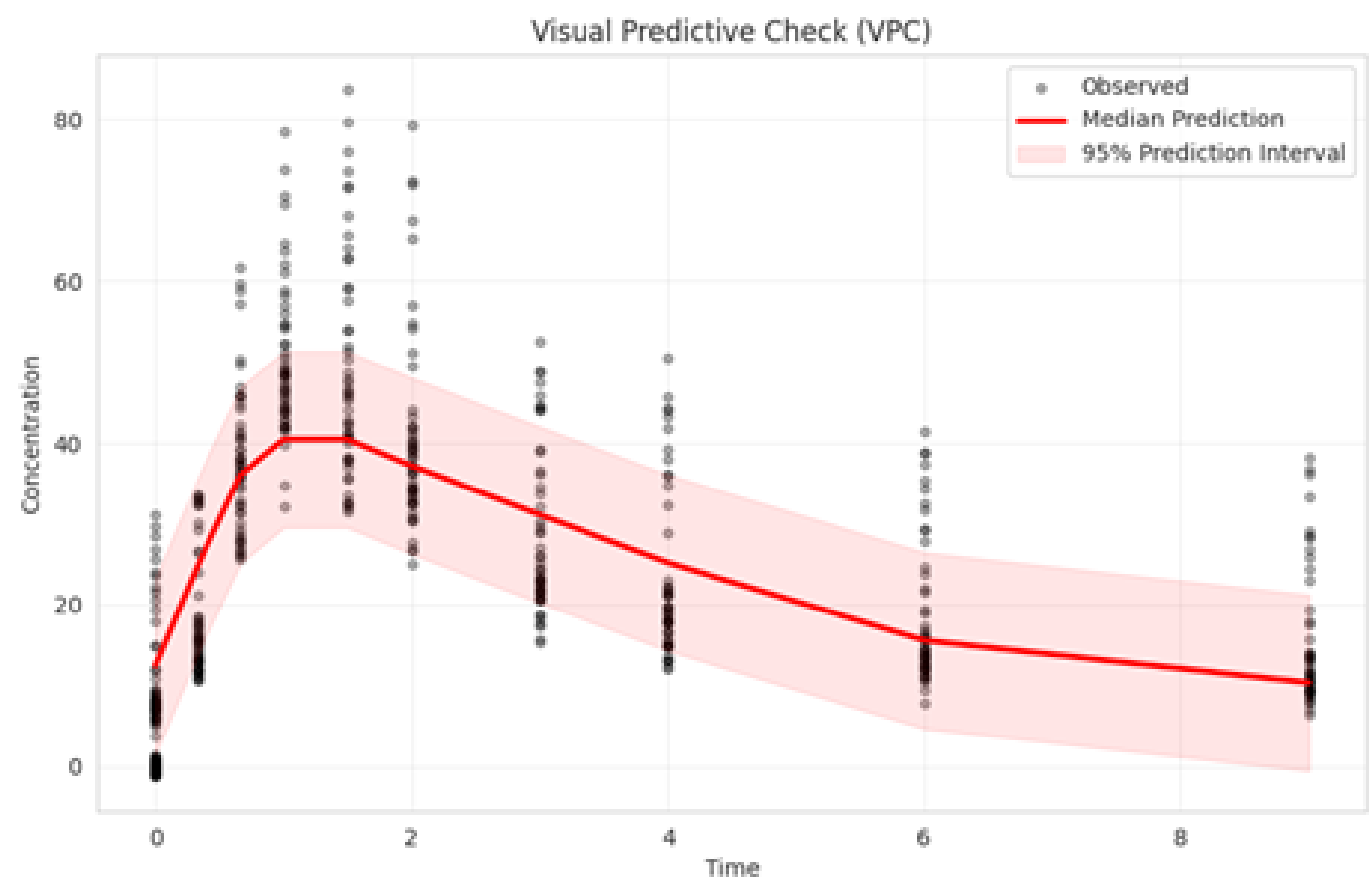
- Build a multi-agent LLM framework executing the full PopPK cycle: EDA structure code SAEM diagnostics.
- Develop a JAX-accelerated SAEM engine
- Validate the framework on simulated and real-world tacrolimus data
- Enable future bounded gray-box extensions while preserving interpretability.

Architectural Pipeline

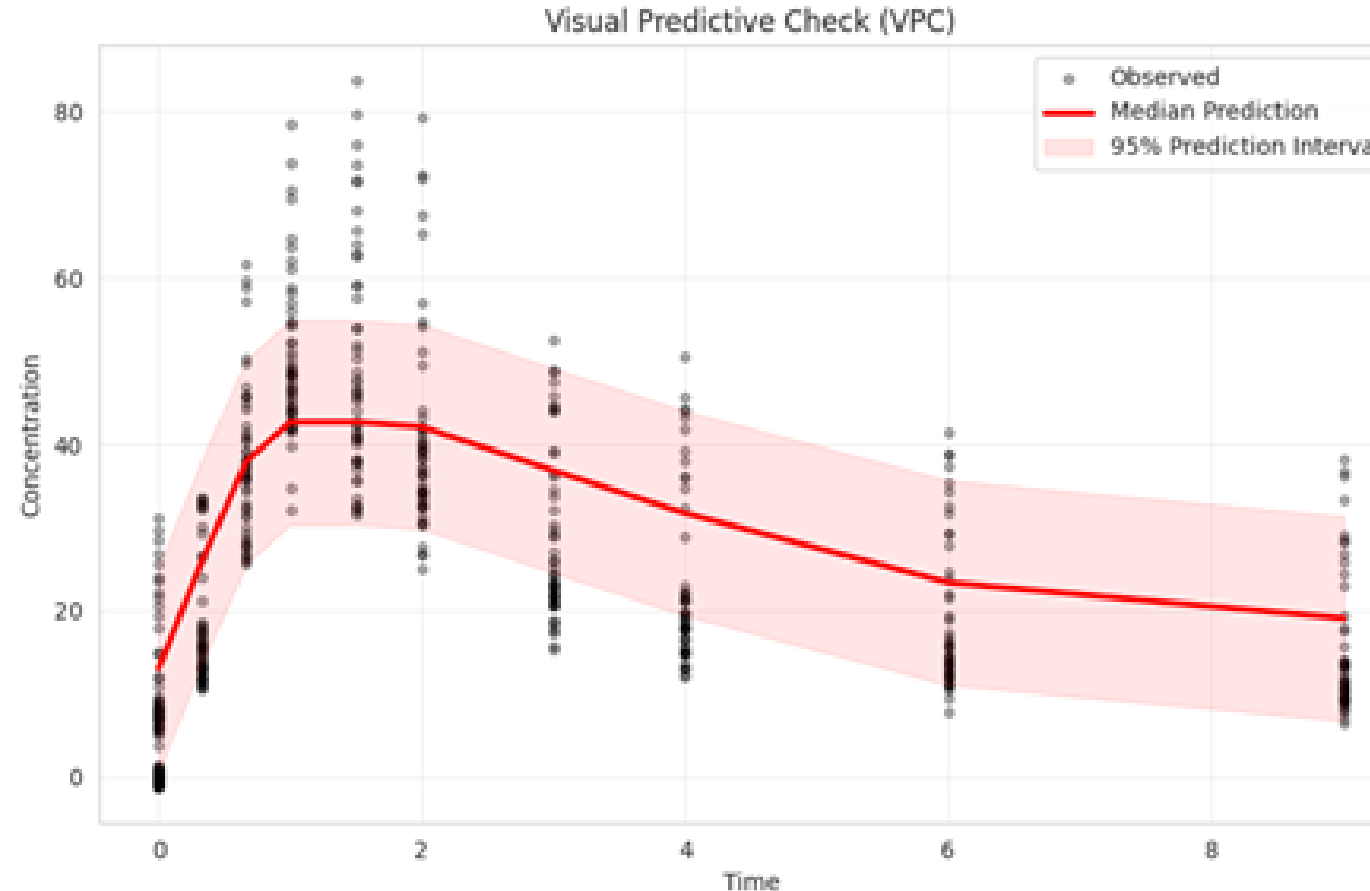


Pipeline example

LLM Proposal based on EDA



LLM Review : “Add a second compartment to have a slower elimination at the end”



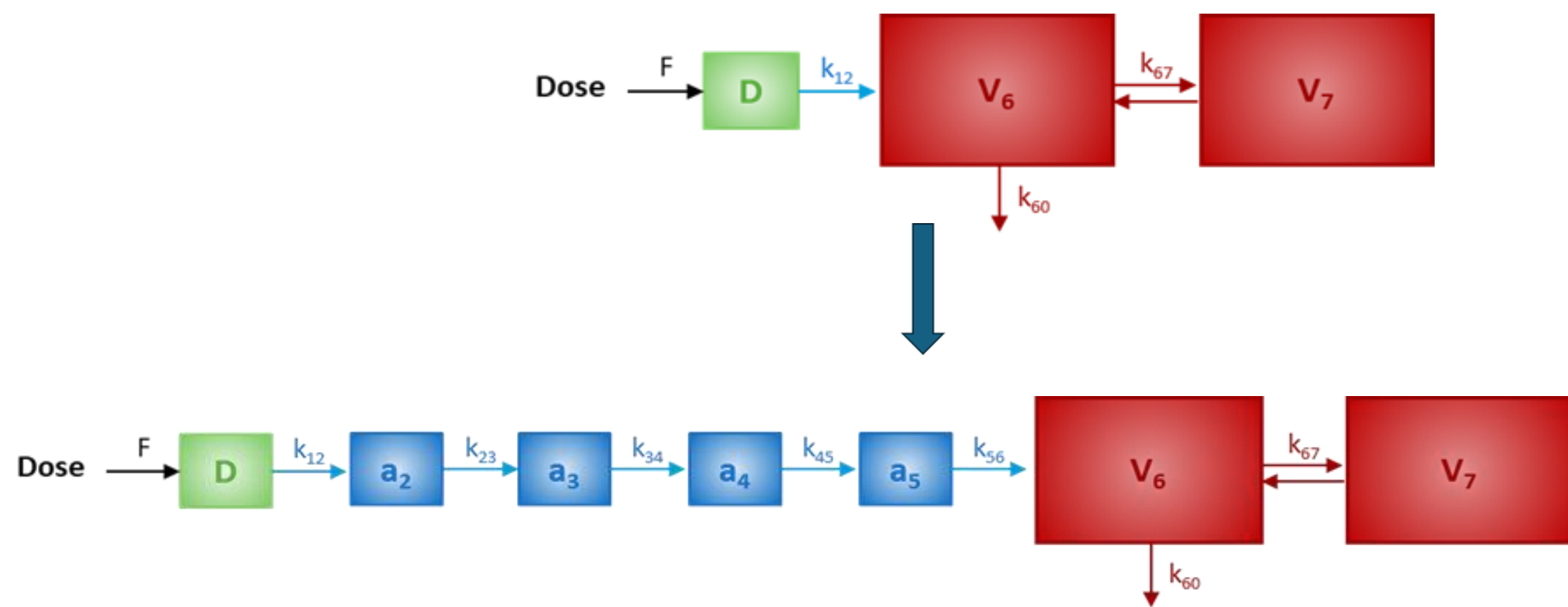
SAEM Validation : Theophylline Dataset

Parameter	PyTorch-SAEM	Monolix
CL	0.046	0.040
V	0.48	0.45
KA	1.69	1.53

These results support the numerical validity of the in-house estimation backend as a basis for autonomous model exploration.

Pipeline Validation : Simulated Dataset & Real Dataset

- On simulated tacrolimus data, the multi-agent loop generally recovered the ground-truth structure and rejected a random dummy SEX covariate.
- When convergence was achieved, the correct PopPK structure was typically recovered within 5–8 refinement iterations.
- The workflow was also transferable to 73 real-world tacrolimus profiles with 10–12 observations per patient.



Gray-box Modeling

- The PyTorch-based engine enables bounded Neural-ODE extensions in which a mechanistic PK backbone is preserved, while a neural residual captures unexplained dynamics under pharmacological constraints.
- More broadly, the custom SAEM engine provides an extensible backbone allowing neural components to propose missing covariate effects or structural refinements within a formal population-estimation framework.

Limitations

- Transfer to real-world tacrolimus data was feasible, but robustness remains limited by LLM-generated code instability.
- Frequent syntax or runtime failures still require stronger execution guardrails.
- JAX + L40 reduced runtime to ~7 min/iteration, but the system remains slower than optimized software.

Conclusion

- This proof of concept shows that LLM agents can act as effective junior pharmacometricians for iterative PopPK structure exploration, candidate coding, and model revision.
- The framework combines automated hypothesis generation, custom SAEM estimation, and model refinement.
- Full autonomy is not yet achieved: strict numerical safeguards, validation layers, and pharmacometric supervision remain essential.

Perspective

- Full JAX port: accelerate SAEM and autonomous structure search on GPU/HPC infrastructures.
- Robust orchestration: add an Aviary-inspired safety layer for code, ODE, and execution validation.
- Hybrid discovery: use the SAEM backbone to test neural proposals for missing covariate effects or structural refinements.

Related Literature

Tacrolimus Model (1): Woillard JB et al. Br J Clin Pharmacol. 2011;71(3):391–402.
Sam Holt et al. Automatically Learning Hybrid Digital Twins of Dynamical Systems

Funding This work is part of the DIGPHAT project, which was supported by a grant from the French government, managed by the National Research Agency (ANR), under the France2030 program, with reference ANR-22-PESN-0017.