Model Ensembling and Machine Learning Approaches to Predict the First Dose of Amoxicillin in Intensive Care

Mihály Leiwolf[1,2], Nicolas Gregoire[1,3], Sophie Magréault[4,5], Bénédicte Franck[6,7], Ombeline Krekounian[1,8], Jean-Baptiste Woillard[2,9], Vincent Aranzana-Climent[1]

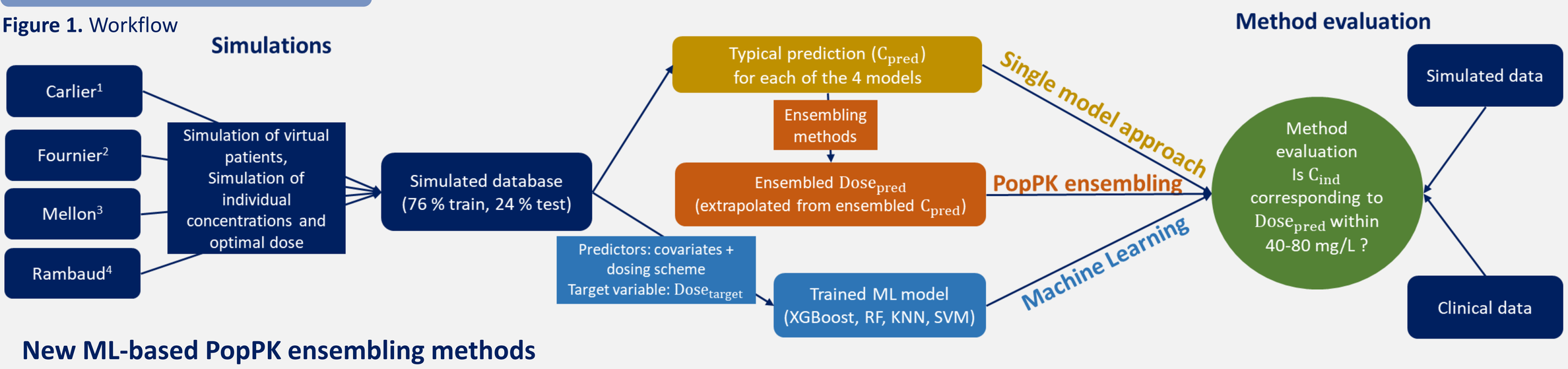
[1]Inserm U1070 PHAR2, University of Poitiers, Poitiers, France; [2]Inserm U1248 P&T, University of Limoges, Limoges, France; [3]Department of Toxicology and Pharmacokinetics, University Hospital of Poitiers, Poitiers, France; [4]Department of Pharmacology, AP-HP, Groupe Hospitalier Paris Seine Saint-Denis, Bondy, France; [5]Inserm UMR1137, IAME, University Paris Cité & Sorbonne University, Paris, France; [6]Univ Rennes, CHU Rennes, EHESP, Irset – UMR S 1085, 35000 Rennes, France; [7]Inserm, CIC 1414, Rennes, France; [8]Department of Clinical Pharmacology, Medical University of Vienna, Vienna, Austria; [9]Department of Pharmacology, Toxicology, and Pharmacovigilance, University Hospital of Limoges, Limoges, France

Objective

To **develop new machine learning (ML)-based model ensembling** and **ML methods** for **amoxicillin** to predict an **optimal first dose in ICU based solely on covariates** (*i.e.* before any individual concentrations are available thus *a priori*) and to compare them to existing methods.

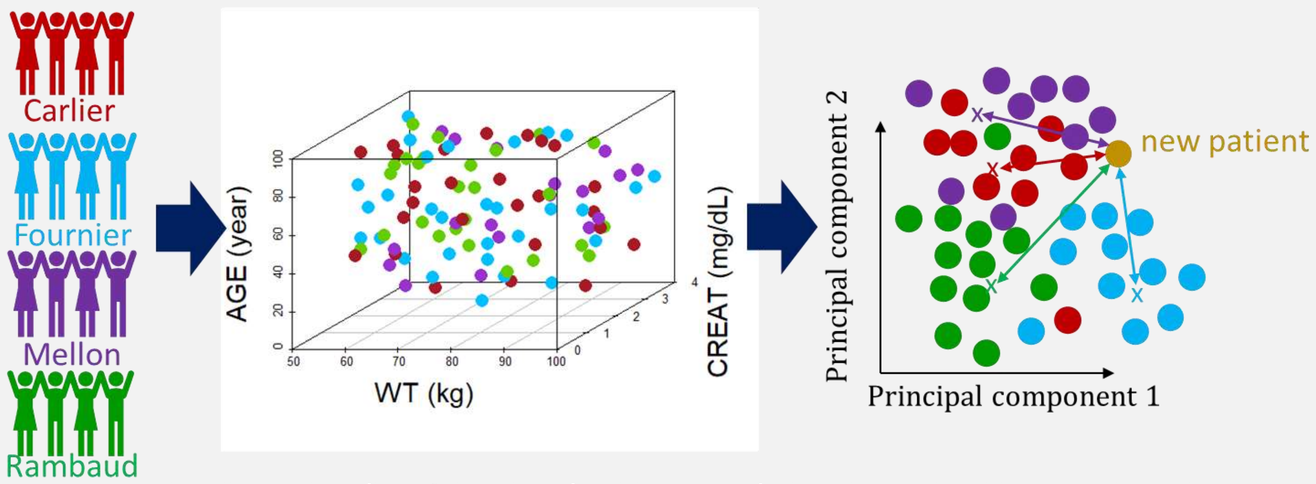
Materials / Methods

Figure 1. Workflow



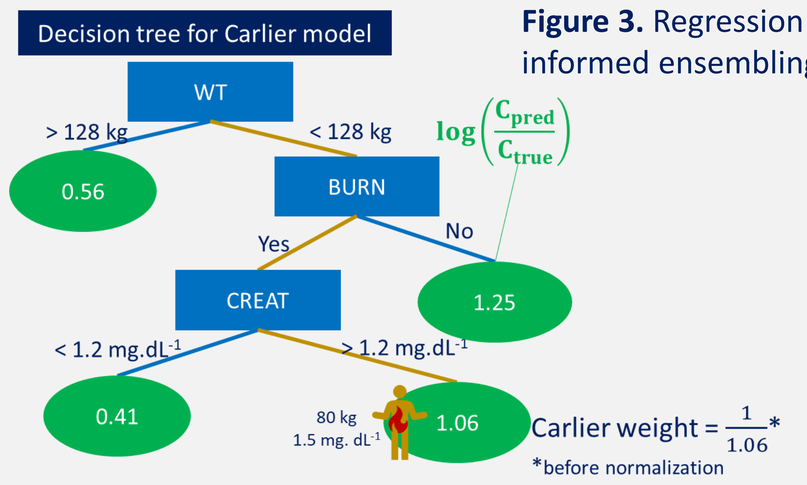
New ML-based PopPK ensembling methods

Figure 2. Factor Analysis of Mixed Data (FAMD)



The dimensionality of the virtual subject population determined by covariates is reduced. A new patient is projected into this latent space, and Mahalanobis distances are computed between the patient and the centroids of each model's cohort to identify the closest matching populations.

Figure 3. Regression tree (RT) informed ensembling



A separate tree is developed for each model with the covariates as predictors and the $\log\left(\frac{C_{pred}}{C_{true}}\right)$ as target.

Results

Figure 4. Target attainment across methods on simulated test data and continuous (ci) clinical data expressed as the % of patients with correctly predicted doses (green).

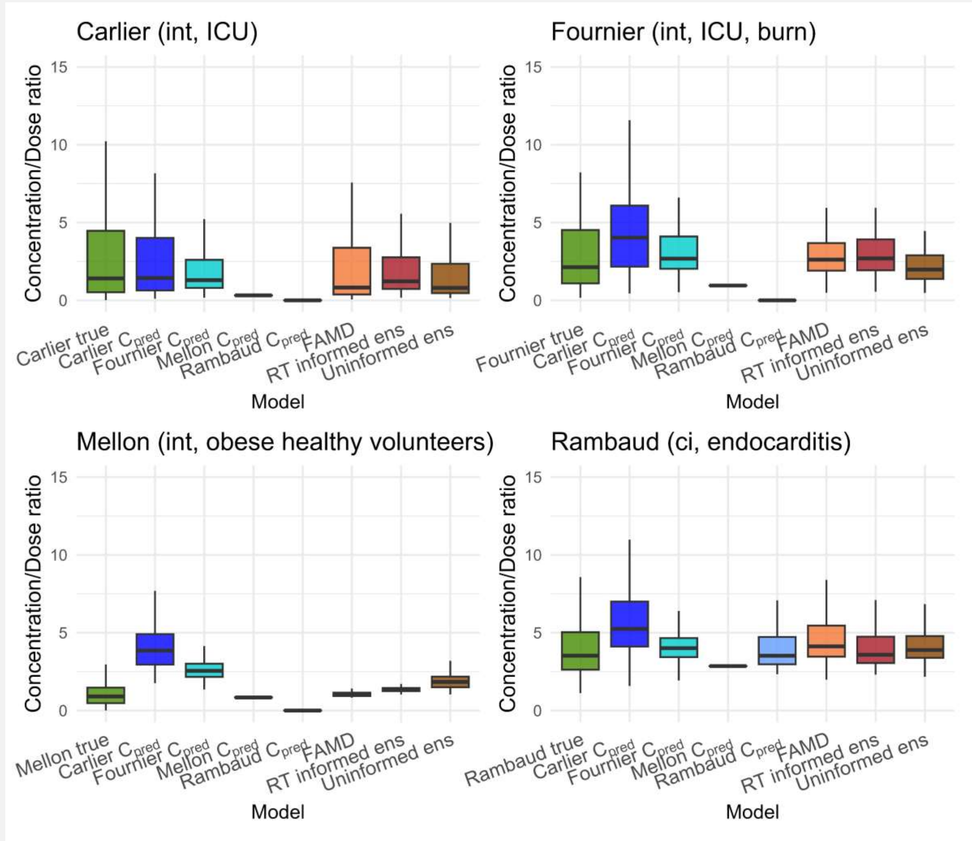


Figure 5. $Cmin_{pred}$ and ground truth trough concentrations (mg/L) normalized by the dose (g) for simulated subjects.

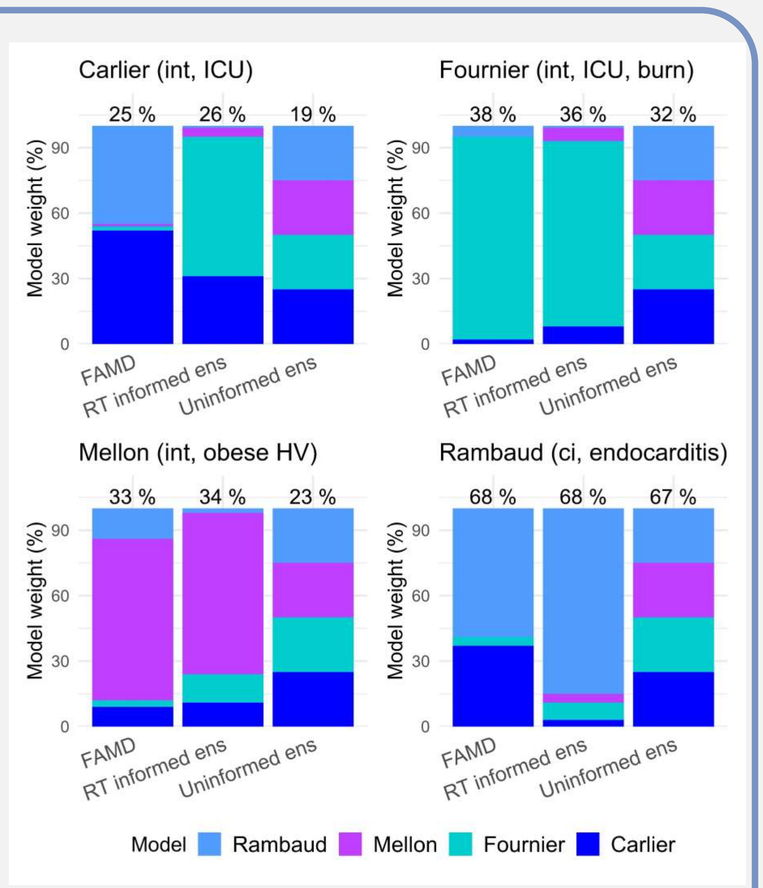


Figure 6. Average model weight attributed to simulated subjects in *Figure 5* by ensembling methods stratified by model cohort. The percentages indicate target attainment.

Conclusion

- ML-based PopPK ensembling outperforms single models by compensating their individual biases: high unexplained inter-individual variability or structural misspecifications
- PopPK ensembling outperforms uninformed ensembling (same weight to all models) on both simulated & clinical data
- Lower extrapolability of pure ML methods due to their data-driven nature

openMIPD

R package for the easy application of 8 different precision dosing methods

Article preprint

openMIPD

References: 1 – Carlier et al. 2013. J Antimicrob Chemother
2 – Fournier et al. 2018. Antimicrob Agents Chemother
3 – Mellon et al. 2020. J Antimicrob Chemother
4 – Rambaud et al. 2020. J Antimicrob Chemother