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

Mihaly LEIWOLF

with the guidance of

Nicolas GREGOIRE,

Vincent ARANZANA-CLIMENT &

Jean-Baptiste WOILLARD

A priori precision dosing

Body weight → 72 kg

Renal function → 67 mL.min^{-1}

Burn patient

Intensive care patient

Ensembling of model predictions based on patient characteristics

Administer predicted 1st dose

Measure drug concentration

In target interval

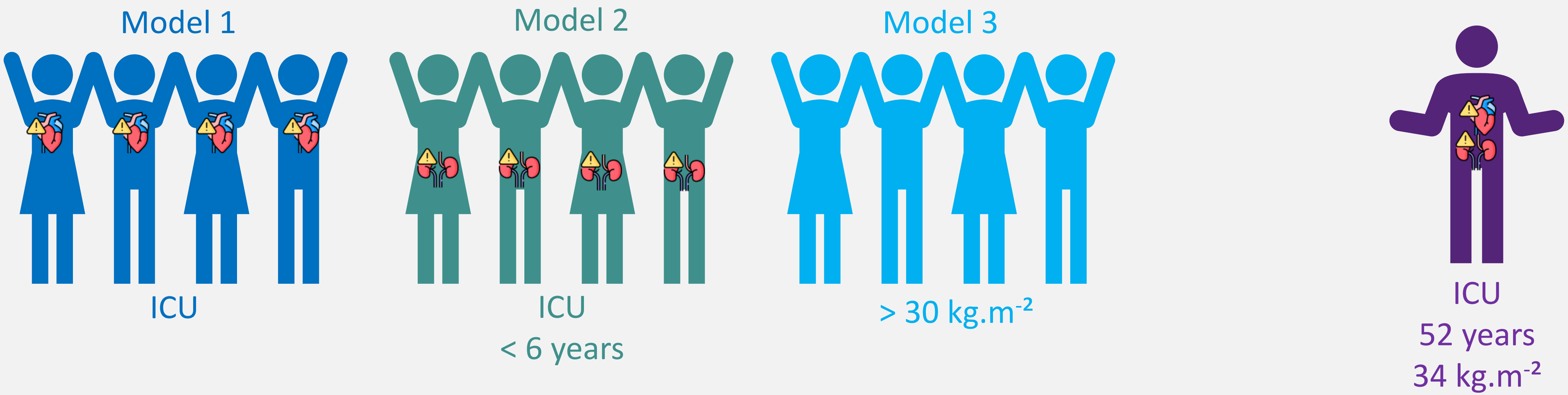
Precision dosing done

Not in target interval

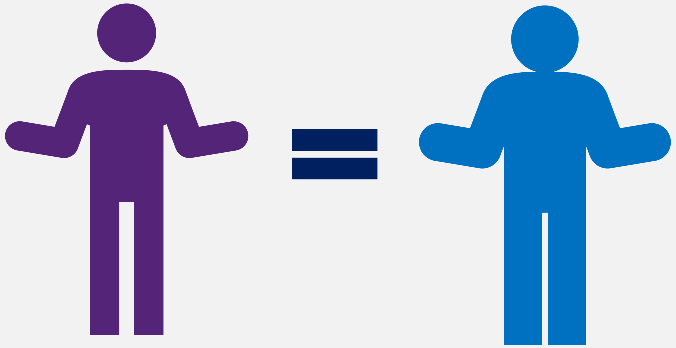
Continue with *a posteriori* precision dosing

- ✓ Faster target attainment
- ✓ Fewer concentration measurements

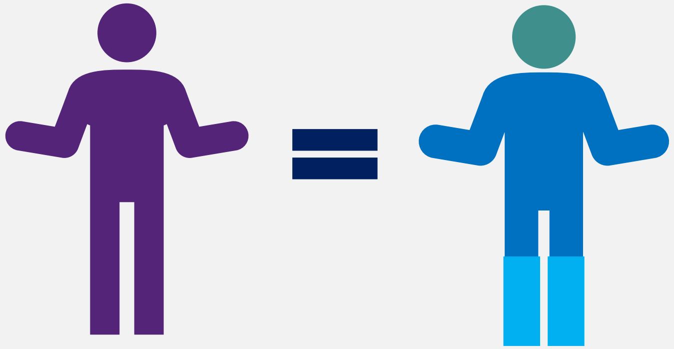
PopPK model ensembling



Model selection



Model ensembling



Objective & workflow

Predict a personnalized **amoxicillin dose** based on plasmatic trough concentrations to reach the concentration range of **40-80 mg.L^{-1}**[*].

Collection of **bibliographic PopPK models**

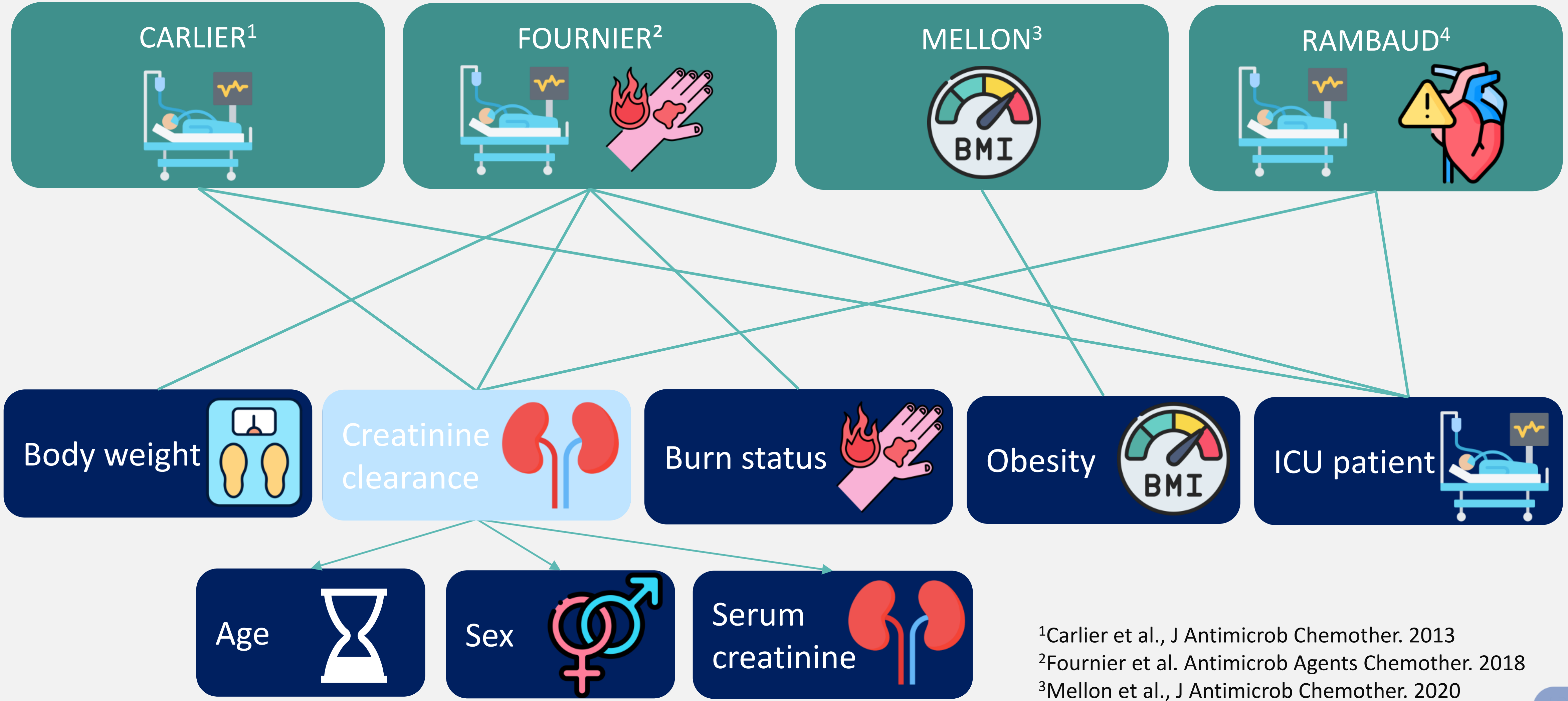
Simulation of subjects and concentrations

Development of *a priori* approaches
- Model ensembling
- Machine learning

Validation using clinical data from 2 ICUs

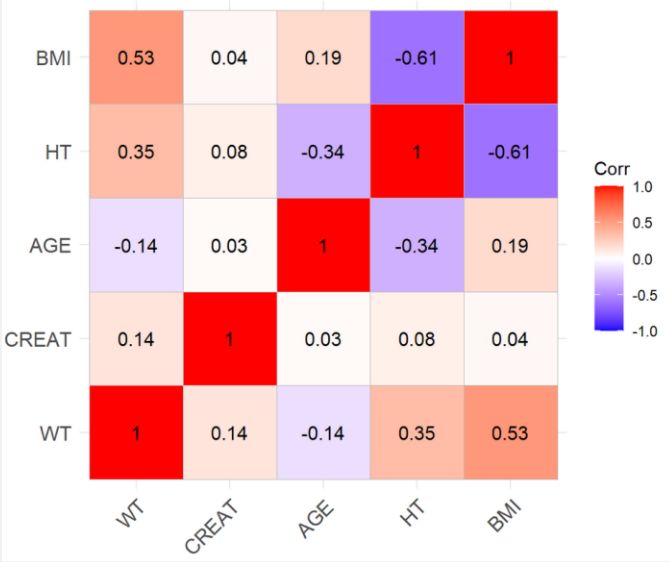
[*]Guilhaumou et al., Crit Care. 2019

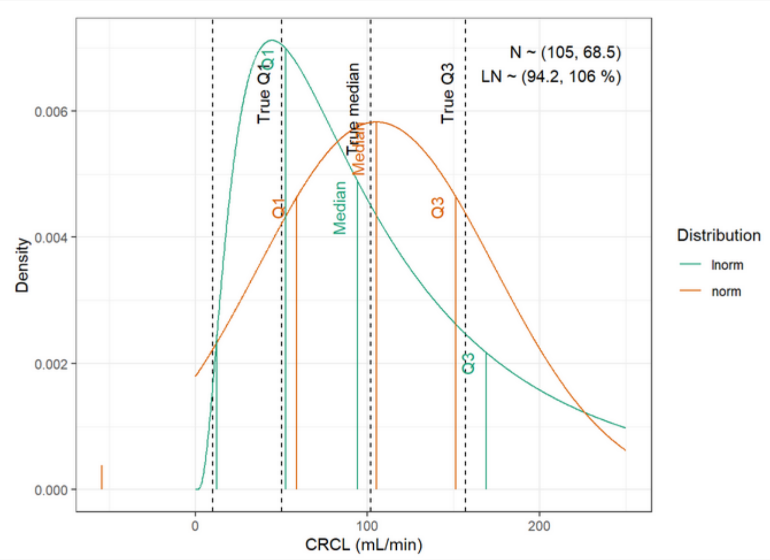
Models and covariates



[1]Carlier et al., J Antimicrob Chemother. 2013
[2]Fournier et al. Antimicrob Agents Chemother. 2018
[3]Mellon et al., J Antimicrob Chemother. 2020
[4]Rambaud et al. J Antimicrob Chemother. 2020

Simulation of virtual subjects & concentrations





Correlations from MIMIC-IV* (database of ~16000 ICU patients)

Best fitting distribution to the median & IQR

Covariate sampling from the multivariate distribution

true value

Carlier IPRED

Carlier cohort subject

Carlier PRED

Fournier PRED

Mellon PRED

Rambaud PRED

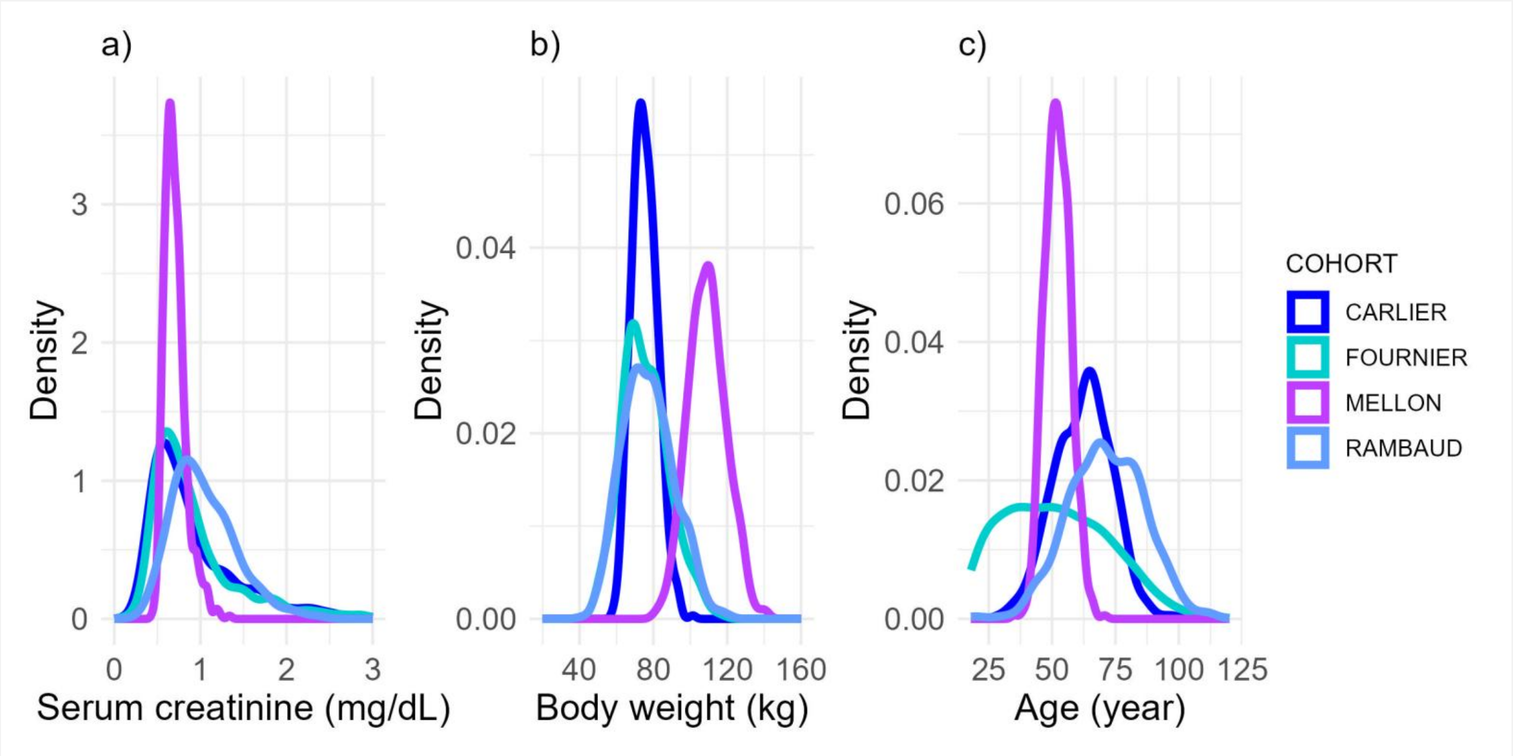
predicted

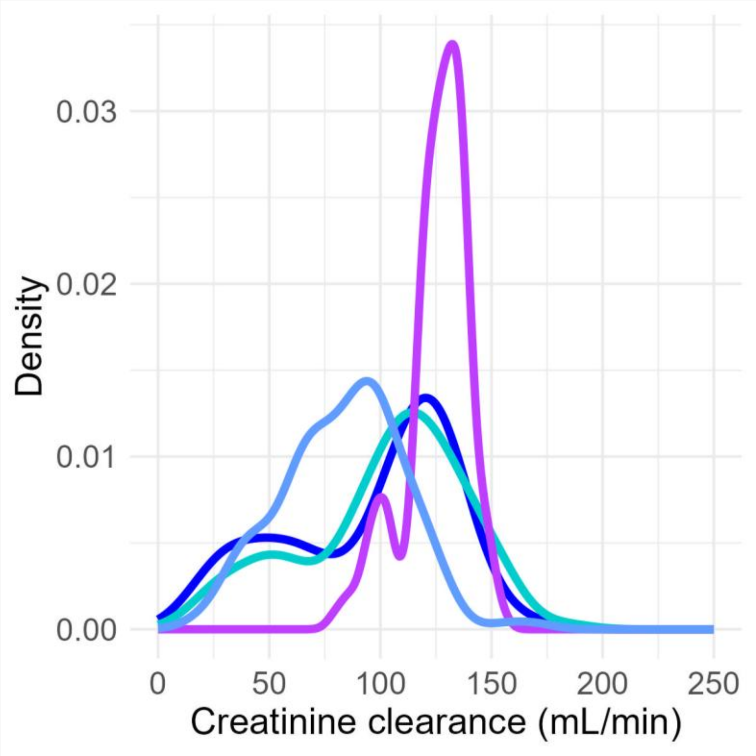
*Johnson et al., Sci Data. 2023

Models and covariates – simulated data

Distribution of continuous covariates

2500 virtual subjects

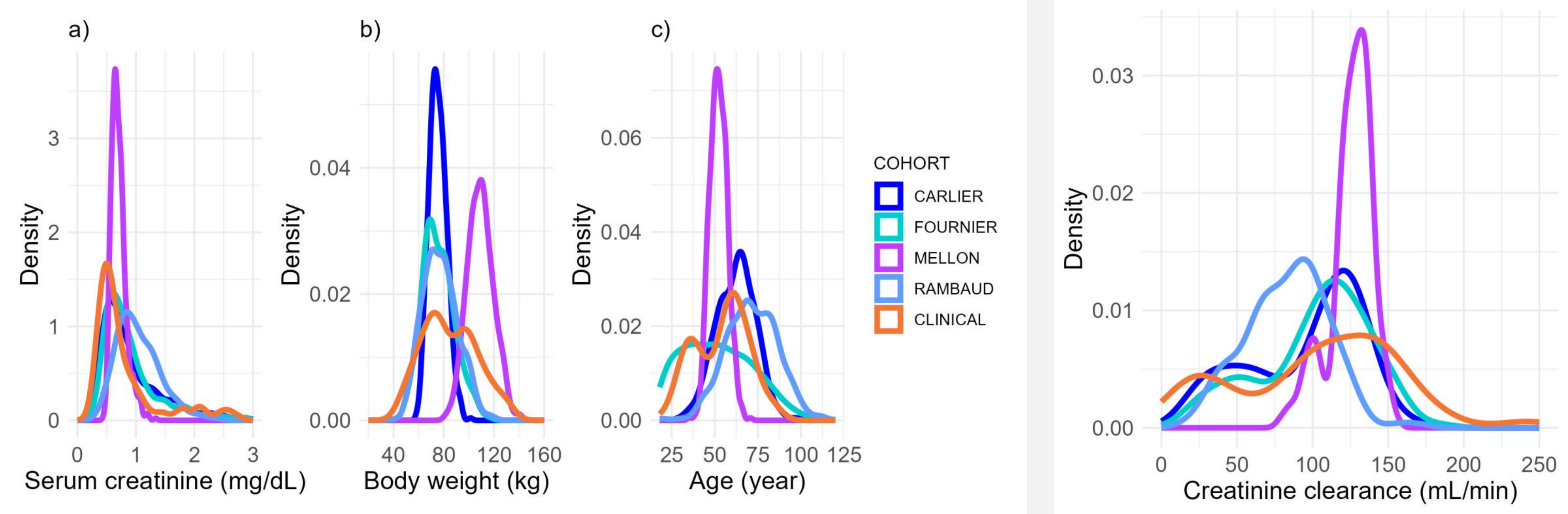




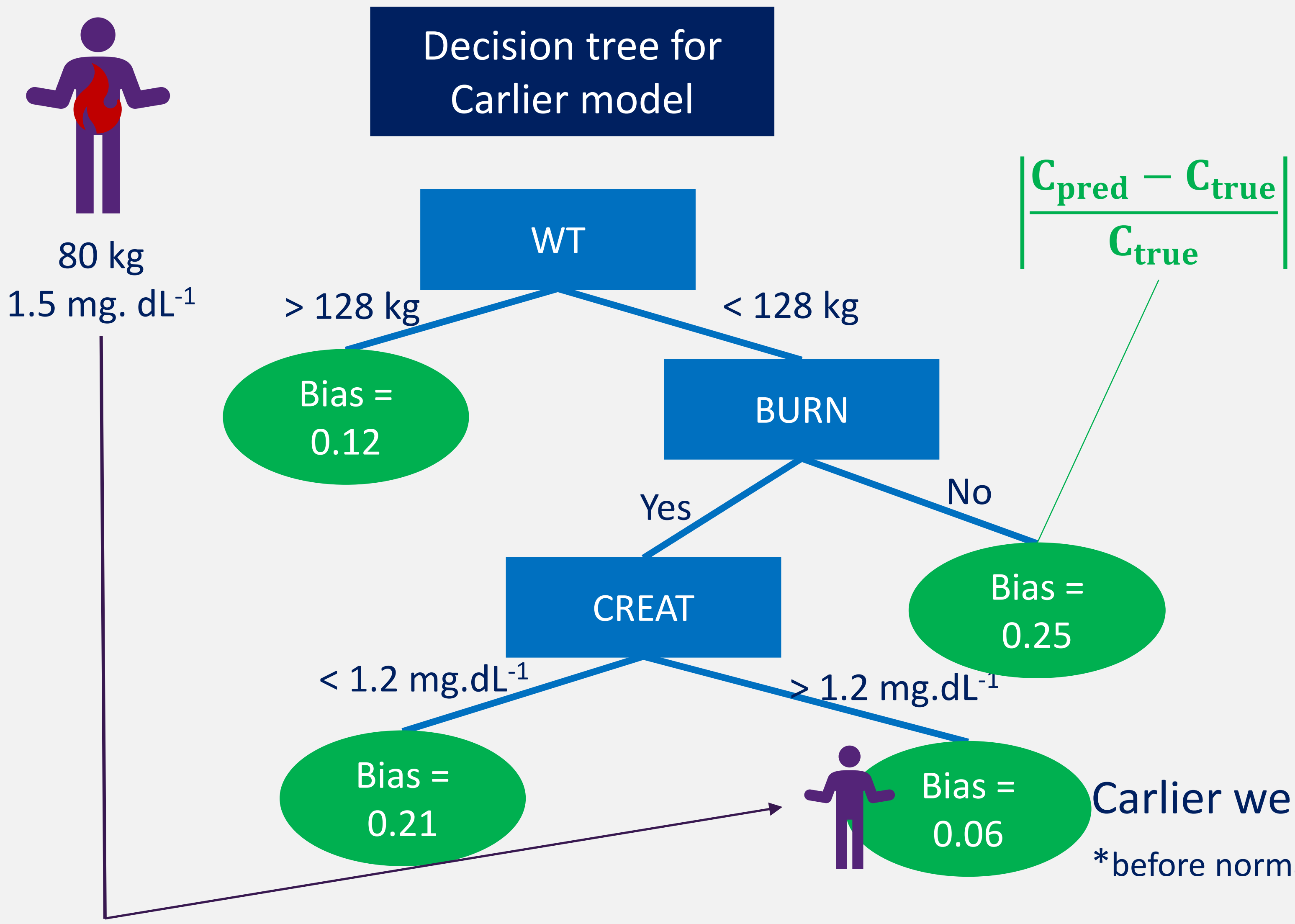
Models and covariates – clinical data

Distribution of continuous covariates

121 patients
from Jean-Verdier Hospital (Bondy) & Poitiers



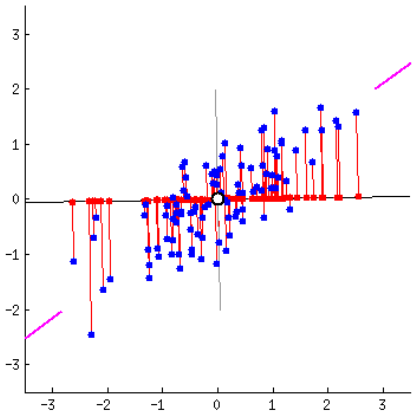
Regression tree (RT)-informed ensembling

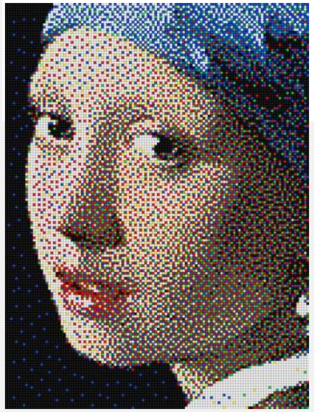


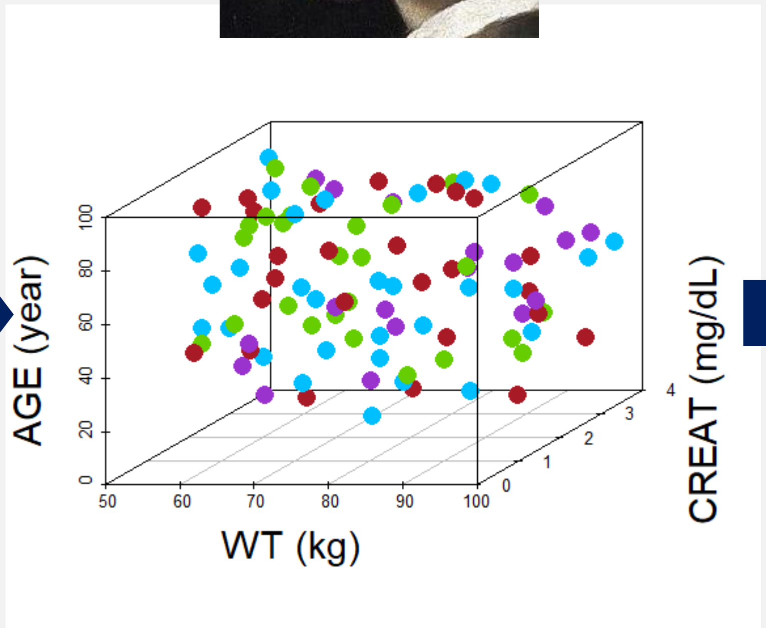
Carlier weight = $\frac{1}{0.06}$*

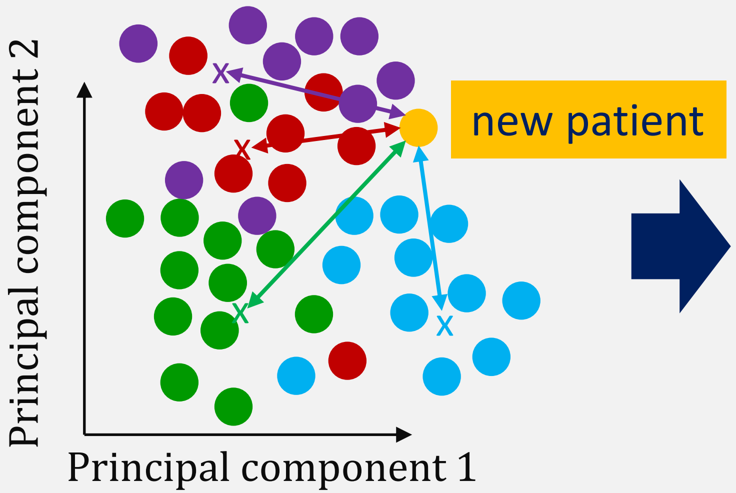
*before normalization

Factor Analysis of Mixed Data (FAMD)



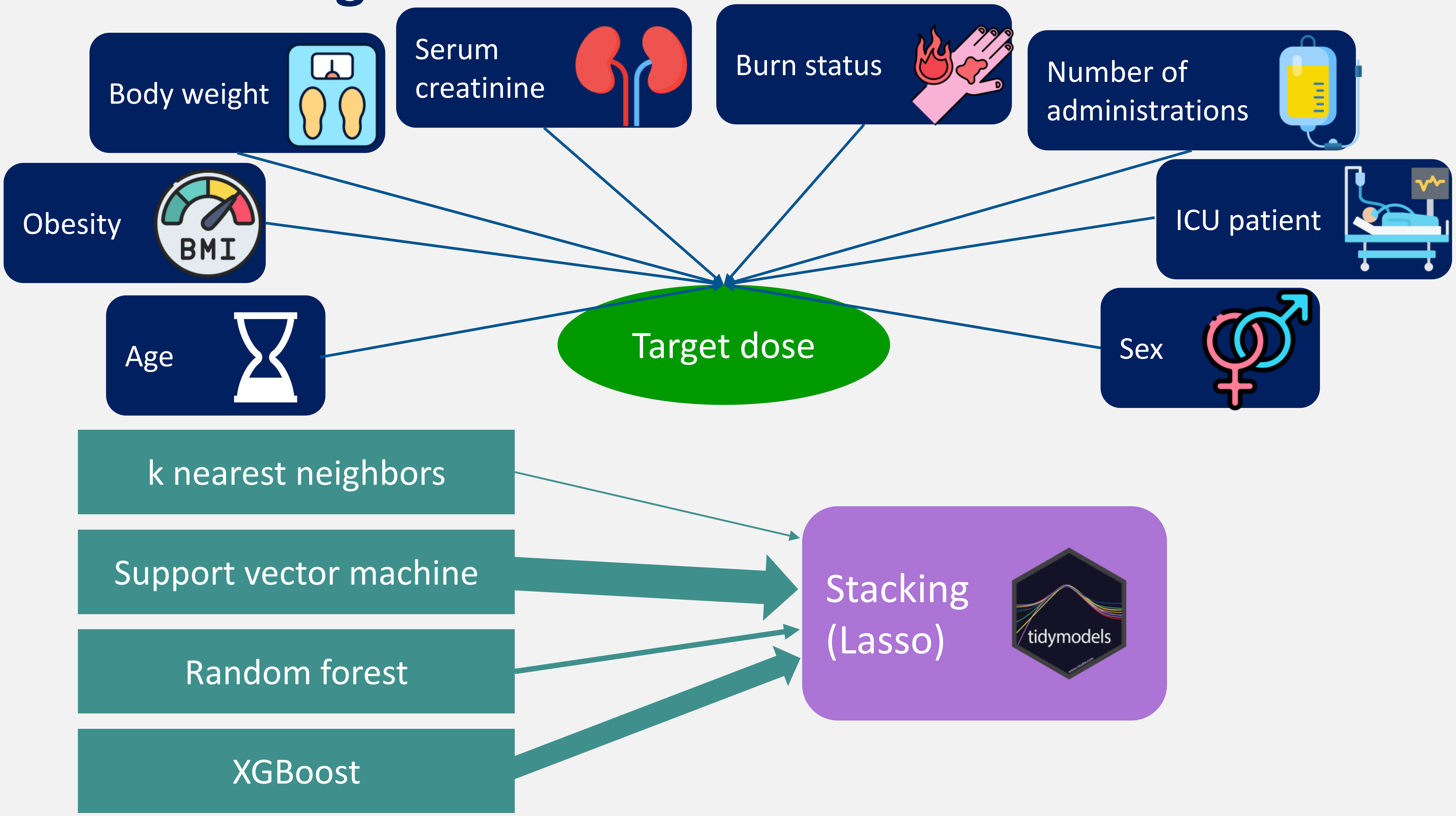




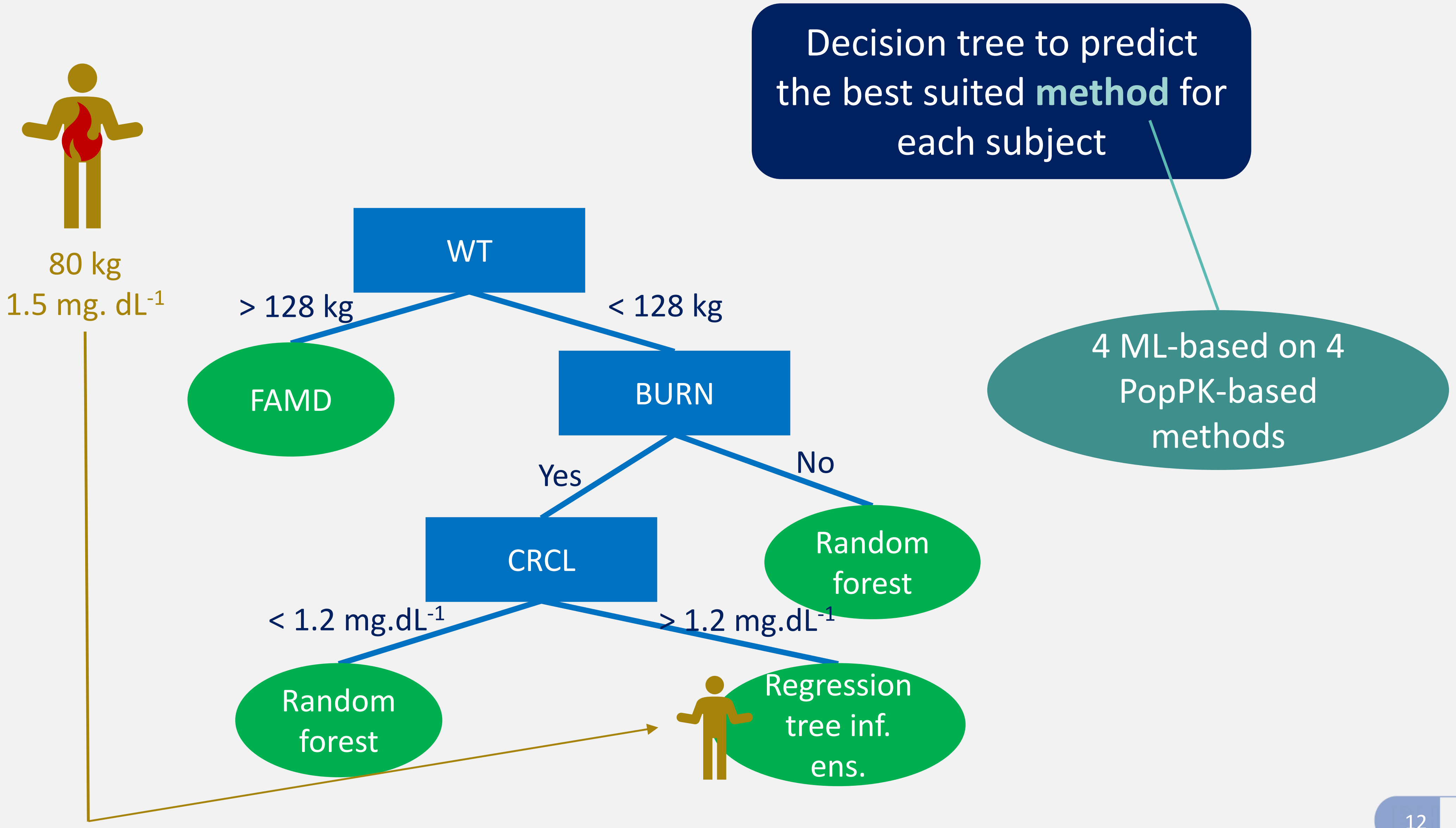


$$\text{Model weight} = \frac{1}{\text{Mahalanobis distance}}$$

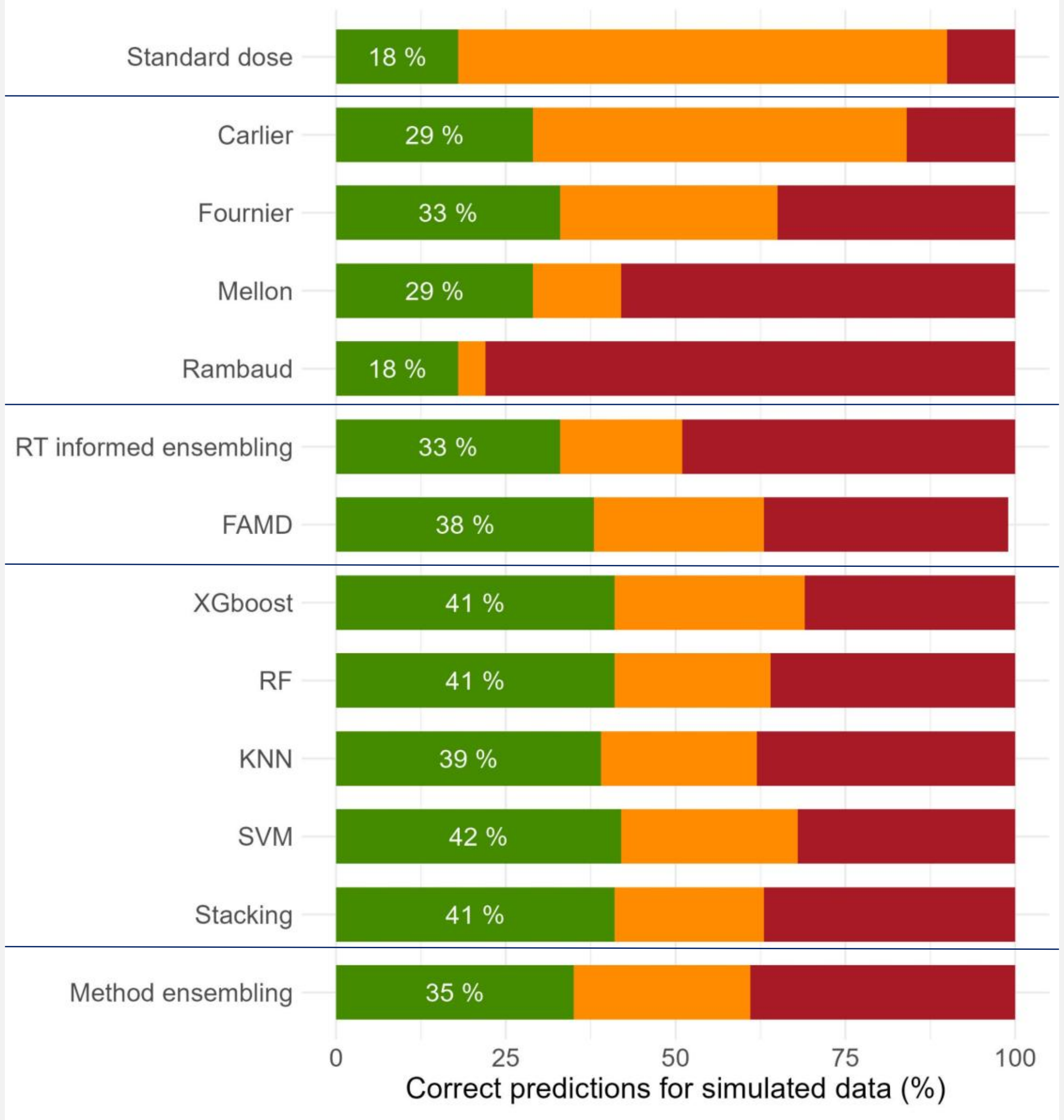
Machine learning



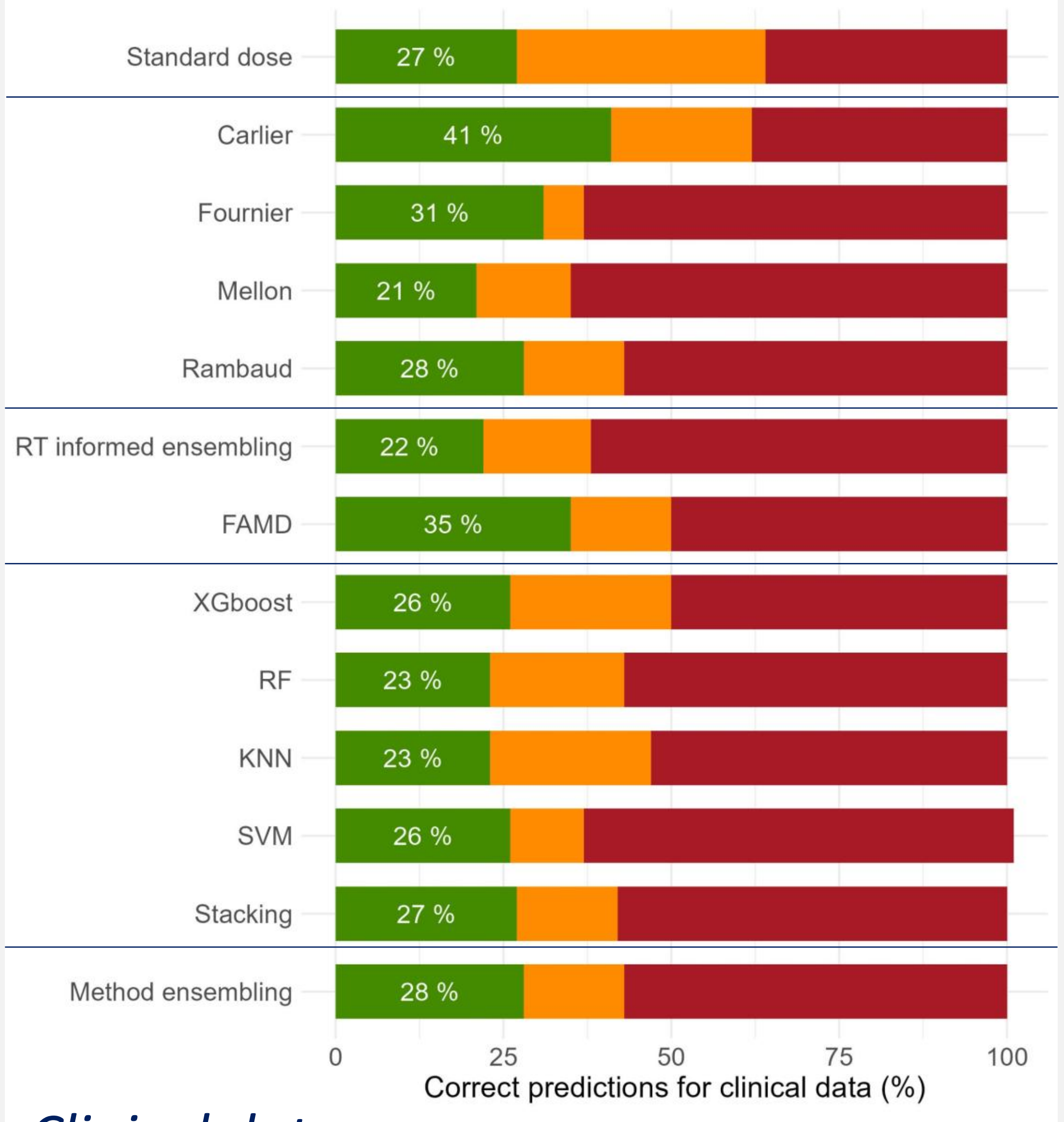
Method ensembling



Results

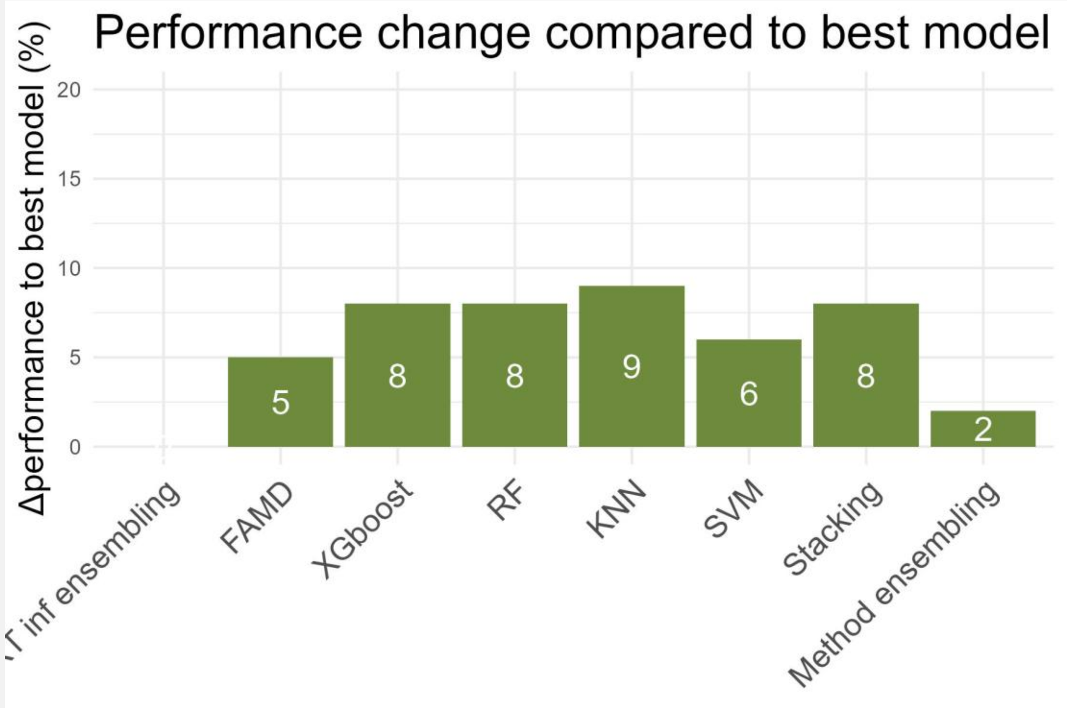


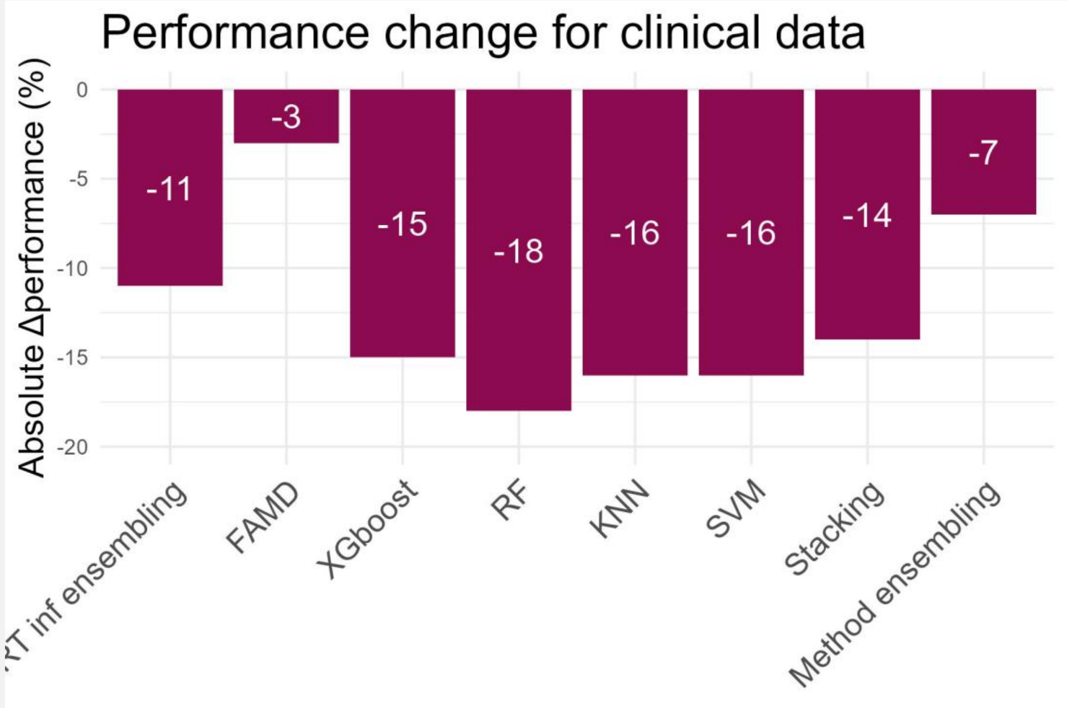
Simulated data



Clinical data

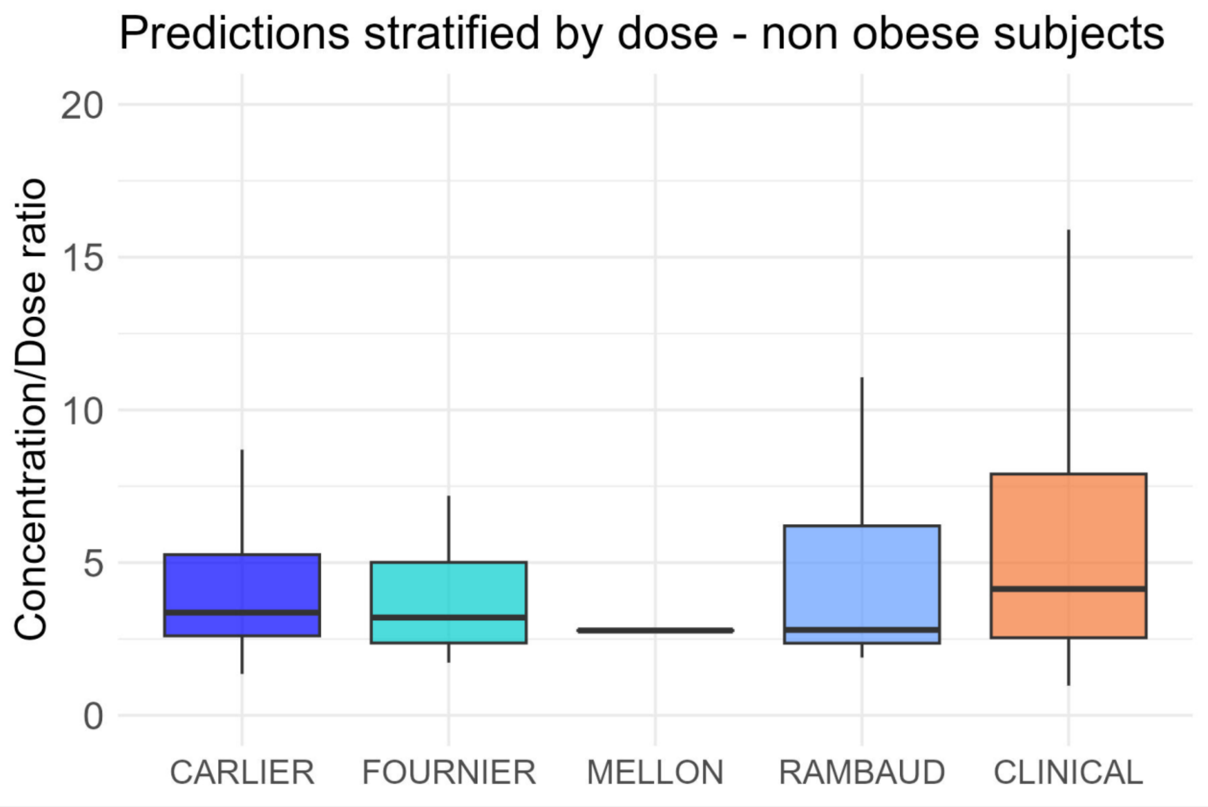
Results

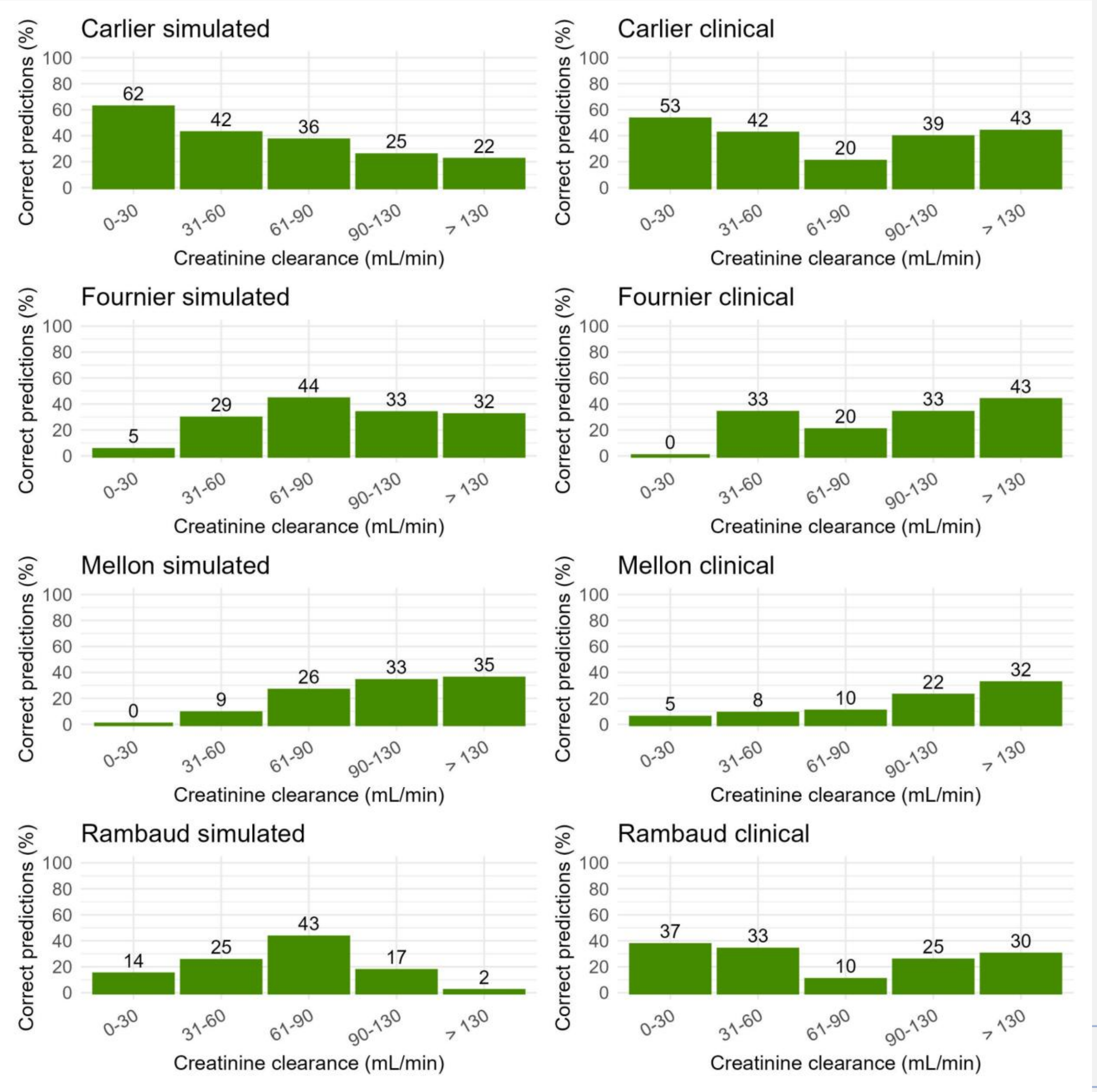

Performance change compared to best model
Δperformance to best model (%)
20
15
10
5
0
5
8
8
9
6
8
2
T inf ensembling
FAMD
XGboost
RF
KNN
SVM
Stacking
Method ensembling


Performance change for clinical data
Absolute Δperformance (%)
0
-5
-10
-15
-20
-11
-3
-15
-18
-16
-16
-14
-7
T inf ensembling
FAMD
XGboost
RF
KNN
SVM
Stacking
Method ensembling

Discussion

Kidney failure patients
Simulated: 3.5 %
Clinical: 16 %





Conclusion

No need for model selection

Methods sensitive to overfitting → retraining with local clinical data

FAMD: reliable & extrapolable, but sensitive to models with lower performance in their own cohort

Better applicable to a molecule with a larger number & more diverse models

Remerciements

Nicolas Gregoire

Vincent Aranzana-Climent

Jean-Baptiste Woillard

Sophie Magréault

Vincent Jullien

Benedicte Franck

Thank you for your attention!

A priori precision dosing method repertoire

Empirical

- Standard dose
- Nomogram

Single model approach

- Carlier
- Fournier
- Mellon
- Rambaud
- Meta model

PopPK model ensembling

- Uninformed
- Weighed
- Classification tree informed
- Regression tree informed
- FAMD

ML + PopPK ensembling ensembling

- Method ensembling (decision tree)

Machine learning

- Support Vector Machine
- k nearest neighbors
- Random forest
- XGBoost

ML ensembling

- Stacking
- ML ensembling (decision tree)