

## Supplementary Figure S2

Supplement to:

Title: Machine learning and pharmacometrics for prediction of pharmacokinetic data: Differences, similarities and challenges illustrated with rifampicin

Authors: Lina Keutzer<sup>1‡</sup>, Huifang You<sup>1‡</sup>, Ali Farnoud<sup>2</sup>, Joakim Nyberg<sup>3</sup>, Sebastian G Wicha<sup>4</sup>, Gareth Maher-Edwards<sup>5</sup>, Georgios Vlasakakis<sup>5</sup>, Gita Khalili Moghaddam<sup>5</sup>, Elin M Svensson<sup>3,8</sup>, Michael Menden<sup>3,6,7</sup>, Ulrika SH Simonsson<sup>1\*</sup>

### Affiliations:

<sup>1</sup>Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden, Ulrika.Simonsson@farmbio.uu.se;

<sup>2</sup>Institute of Computational Biology, Helmholtz Zentrum München, Munich, Germany;

<sup>3</sup>Department of Pharmacy, Uppsala University, Uppsala, Sweden;

<sup>4</sup>Department of Clinical Pharmacy, Institute of Pharmacy, University of Hamburg, Hamburg, Germany;

<sup>5</sup>Research, Clinical Pharmacology Modelling & Simulation, GlaxoSmithKline, London, UK;

<sup>6</sup>Department of Biology, Ludwig-Maximilians University Munich, Martinsried, Germany;

<sup>7</sup>German Center for Diabetes Research (DZD e.V.), Neuherberg, Germany;

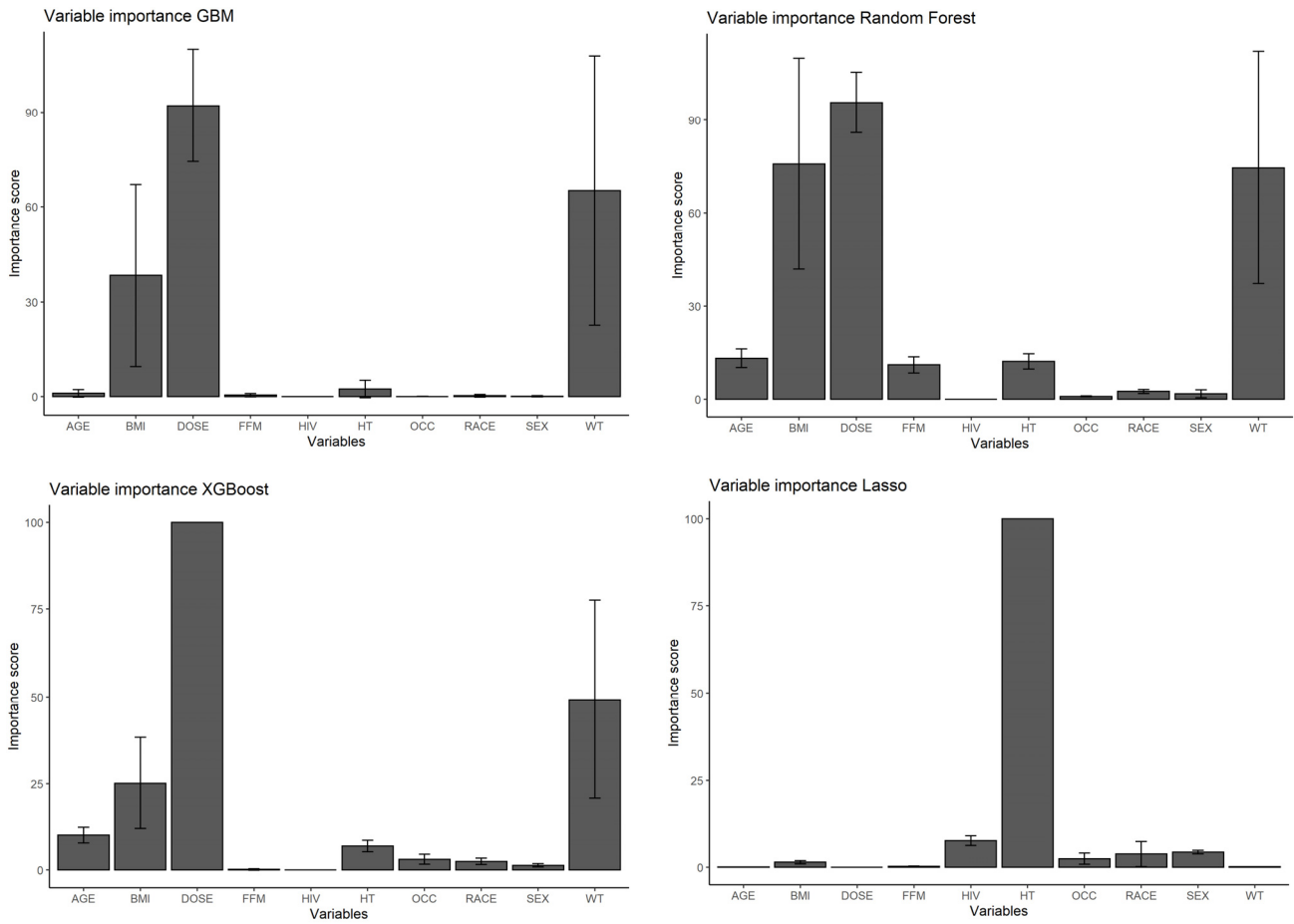
<sup>8</sup> Department of Pharmacy, Radboud institute of health sciences, Radboud university medical center, Nijmegen, The Netherlands

<sup>‡</sup> These authors contributed equally

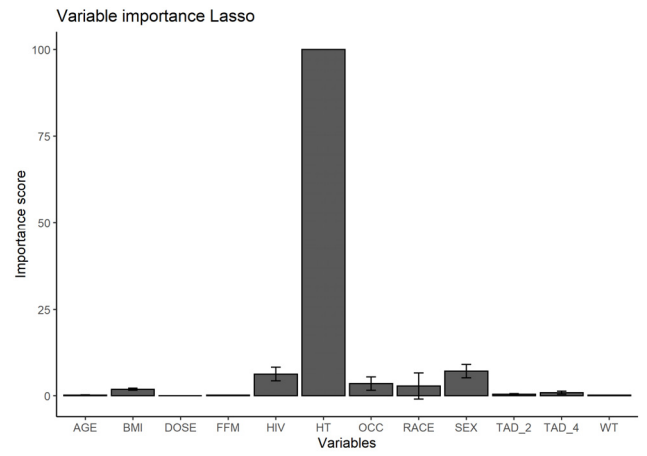
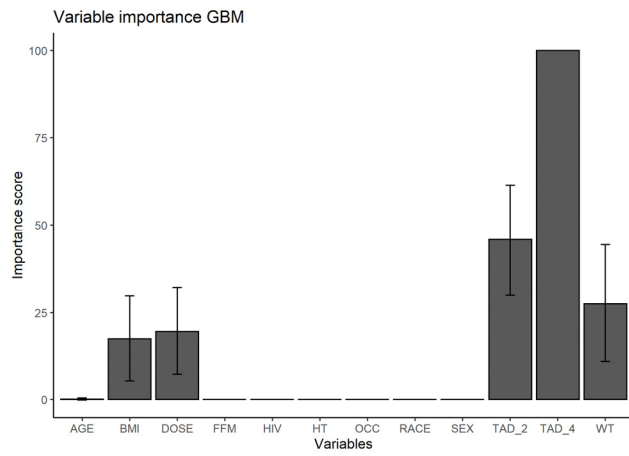
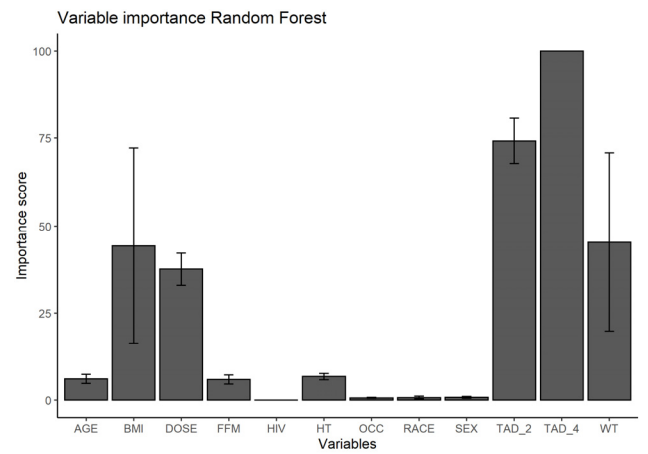
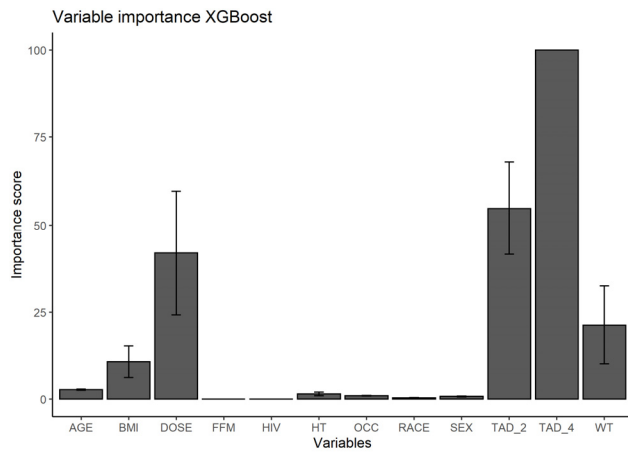
\*Correspondence: Ulrika.Simonsson@farmbio.uu.se, Department of Pharmaceutical Biosciences, Uppsala University, Box 591, 75124 Uppsala, Sweden;

Supplementary Figure S2

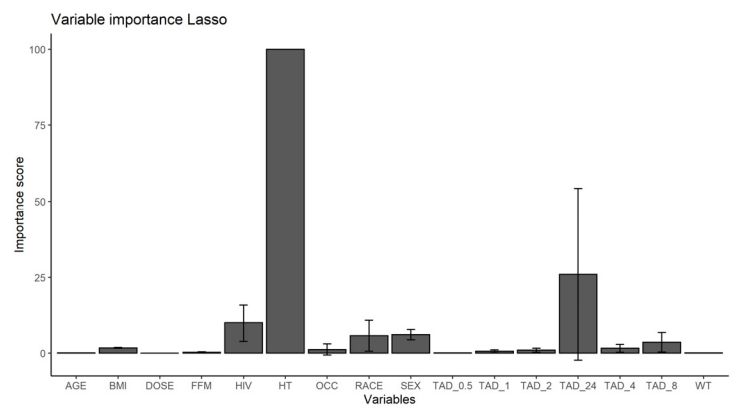
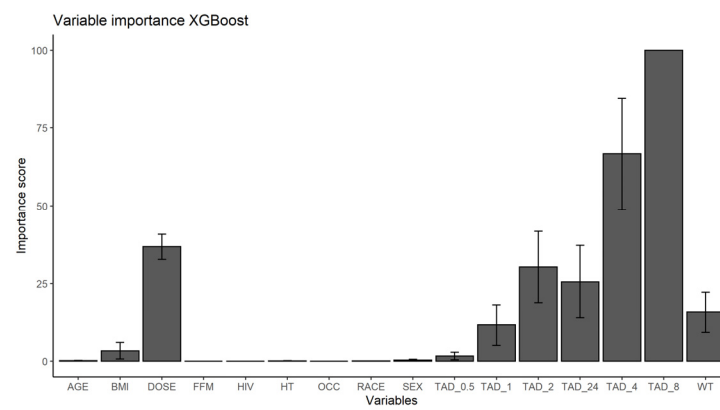
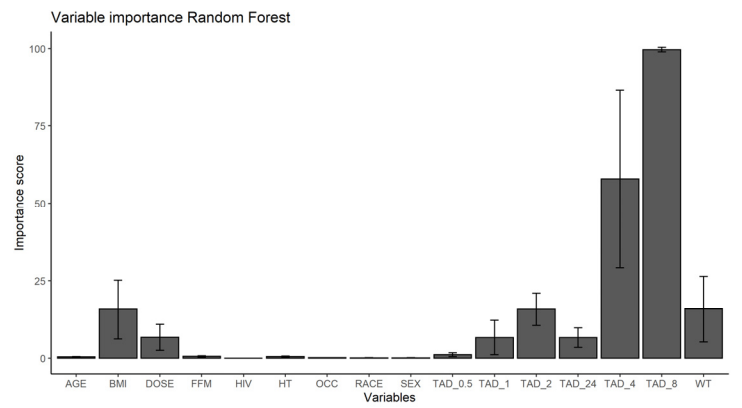
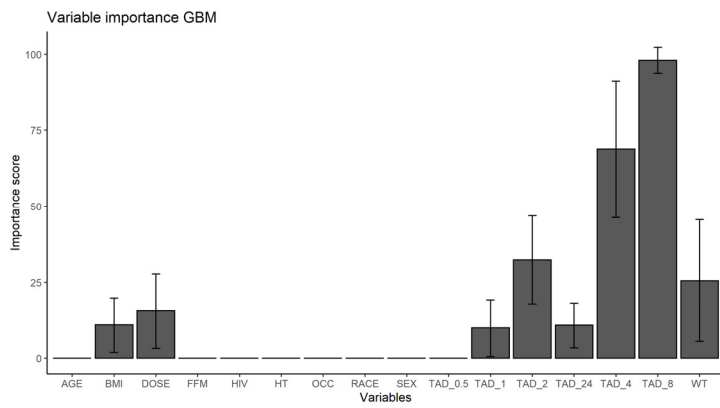
A)



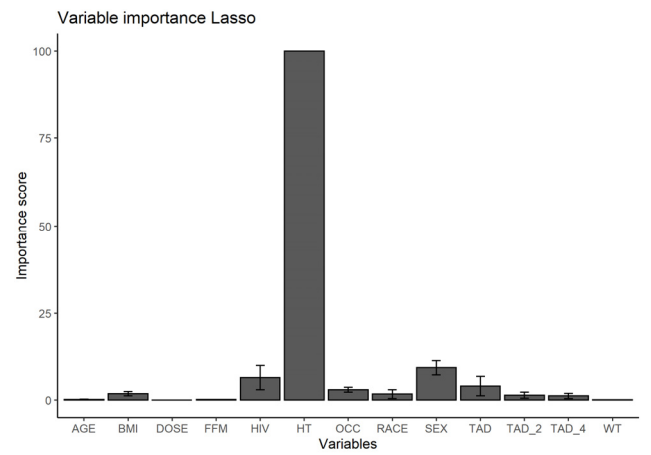
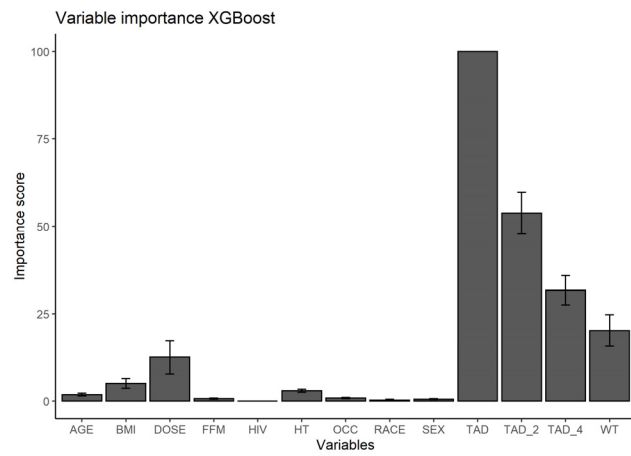
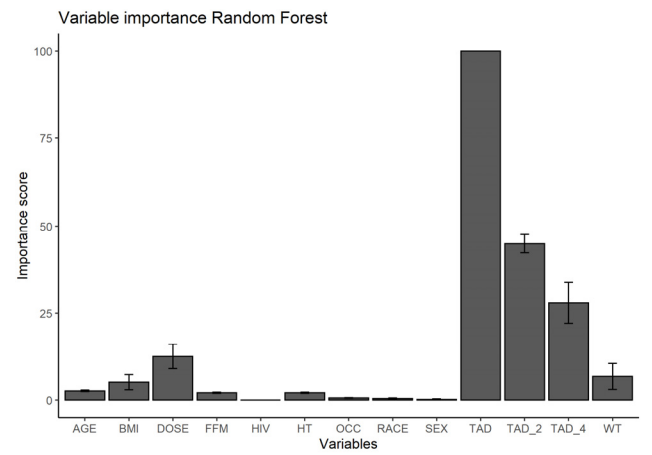
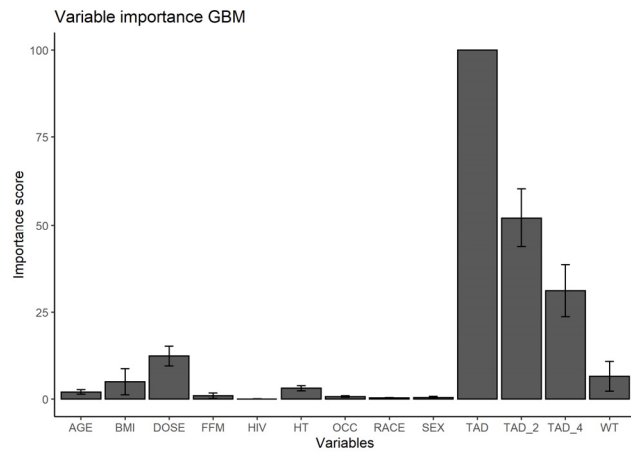
**B)**



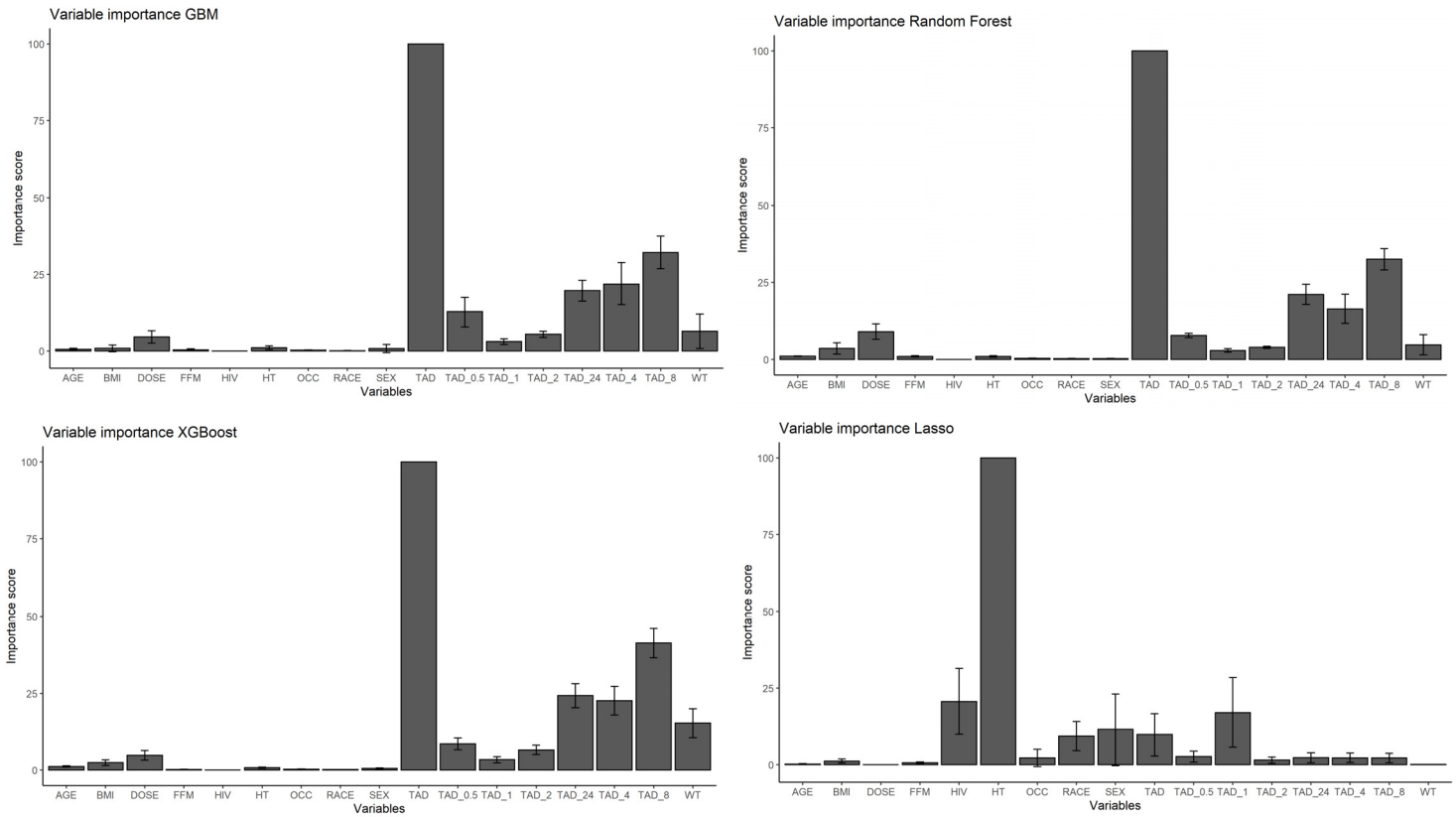
C)



D)



E)



**Figure S2.** Importance scores for evaluated features shown for the machine learning algorithms GBM, Random forest, XGBoost and LASSO for A)  $AUC_{0-24h}$  predictions using features only as input (scenario 4), B)  $AUC_{0-24h}$  predictions using 2 plasma concentrations as input (scenario 5), C)  $AUC_{0-24h}$  predictions using 6 plasma concentrations as input (scenario 6), D) prediction of the plasma concentration-time series using 2 plasma concentrations as input (scenario 2), E) prediction of the plasma concentration-time series using 6 plasma concentrations as input (scenario 3). AGE, age (years); BMI, body mass index ( $kg/m^2$ ); DOSE, daily rifampicin dose (mg); FFM, fat-free mass (kg); HIV, HIV-coinfection; HT, body height (cm); OCC, treatment week, RACE, race; SEX, gender; TAD, time after dose (h); WT, bodyweight (kg); TAD, time after dose (h); TAD\_0.5, rifampicin plasma concentration at 0.5 hours post-dose; TAD\_1, rifampicin plasma concentration at 1 hours post-dose; TAD\_2, rifampicin plasma concentration at 2 hours post-dose; TAD\_4, rifampicin plasma concentration at 4 hours post-dose; TAD\_8, rifampicin plasma concentration at 8 hours post-dose; TAD\_24, rifampicin plasma concentration at 24 hours post-dose.