Title: Machine learning time-to-event analysis for prediction of 2-month culture conversion with phase 2a information in tuberculosis drug development

Authors: Huifang you (1), Lina Keutzer (1), Ulrika Simonsson (1)

Institutions:
(1) Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden

Introduction:
In clinical tuberculosis (TB) drug development, phase 2a trials assess early bactericidal activity (EBA) over two weeks using the time-to-positivity (TTP) biomarker time [1]. A standardized methodology has been developed for model-based EBA analysis of Phase 2a trials in TB [2]. The Phase 2a trial is followed by a Phase 2b study which is 8 weeks of treatment with the endpoint culture conversion. In this work, we investigated the informativeness of phase 2a TTP biomarker efficacy for predicting phase 2b TTE of culture conversion in TB using machine learning (ML). Additionally, the informativeness of different Phase 2a study lengths for the prediction of culture conversion was explored.

Methods:
Time-to-event culture conversion data from the phase 3 trial (REMoxTB, NCT00864383) was utilized [3]. The EBA analysis was applied using nonlinear mixed effects modeling in NONMEM (version 7.5.1, ICON Development Solutions, Hanover, MD, USA) and PsN (version 5.3.1, Department of Pharmacy, Uppsala University) to describe the decline in TTP over time. The model building was following Mockeliunas's process [2]. The predicted baseline TTP, predicted difference in TTP from 0 to 14 days ($TTP_{0\text{-}14}$ days) and 0 to 28 days ($TTP_{0\text{-}28}$ days) using the final EBA model were used as predictors of TTE of culture conversion in the ML analysis.

In the ML analysis, the TTE culture conversion data was divided into training and testing datasets for different scenarios. Extreme gradient boosting (Xgboost), Random forest (RF), K-nearest-neighbors algorithm (KNN) and C-support vector classification (SVC) were applied using Python (version 3.9.1) to predict the TTE of culture conversion at 2 months. Bayesian optimization and five-fold cross-validation were applied to the model-building process. Different ML models were evaluated using ROC-AUC, C-index metrics and VPC of Kaplan-Meier (KM) plots to guide the model selection.

Results:

In the EBA analysis, the median predicted baseline TTP, predicted $TTP_{0\text{-}14}$ and predicted $TTP_{0\text{-}28}$ were 126.8, 182, and 240.3 hours, respectively, with no significant differences among the three treatments in TTP. The ML analysis used various training/testing splits across scenarios: Scenario A utilized an 80%/20% split, achieving a ROC-AUC of 0.94 and C-index of 0.86 with Xgboost. Scenario B trained on data for 2 weeks and tested up to 2 months, selecting Xgboost as the best model (ROC-AUC 0.61, C-index 0.56). Scenario C extended training to 4 weeks, finding SVC as the optimal model (ROC-AUC 0.76, C-index 0.69). Scenario D, using data up to 2 weeks for a 4-week test, showed the best results with RF (ROC-AUC 0.73, C-index 0.64). All models demonstrated a good fit to the Kaplan-Meier plots, with model observations well within the 90% confidence interval. Notably, predictions in Scenario C were better than those in Scenario B from weeks 6 to 8.

Conclusions:

ML is a useful tool for the analysis of TTE data and can utilize pharmacometrics-based predicted biomarker information as part of the explanatory variables. Phase 2a TTP

information predicted the Phase 2b outcome and was even better when an extended Phase 2a trial of 4 weeks was used to predict the Phase 2b culture conversion.

Funding:
This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No 101007873. The JU receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA, Deutsches Zentrum für Infektionsforschung e. V. (DZIF), and Ludwig-Maximilians-Universität München (LMU). EFPIA/AP contribute to 50% of funding, whereas the contribution of DZIF and the LMU University Hospital Munich has been granted by the German Federal Ministry of Education and Research.

Reference:

[1] A. H. Diacon, J. S. Maritz, A. Venter, P. D. van Helden, R. Dawson, and P. R. Donald, "Time to liquid culture positivity can substitute for colony counting on agar plates in early bactericidal activity studies of antituberculosis agents," *Clinical Microbiology and Infection*, vol. 18, no. 7, pp. 711–717, Jul. 2012, doi: 10.1111/j.1469-0691.2011.03626.x.
[2] L. Mockeliunas *et al.*, "Standards for model-based early bactericidal activity analysis and sample size determination in tuberculosis drug development," *Front. Pharmacol.*, vol. 14, p. 1150243, Apr. 2023, doi: 10.3389/fphar.2023.1150243.
[3] S. H. Gillespie *et al.*, "Four-Month Moxifloxacin-Based Regimens for Drug-Sensitive Tuberculosis," *N Engl J Med*, vol. 371, no. 17, pp. 1577–1587, Oct. 2014, doi: 10.1056/NEJMoa1407426.