Individualized treatment effect was predicted best by modeling baseline risk in interaction with treatment assignment

# Abstract

Our objective was to compare different risk-based methods for optimal prediction of individualized treatment effects from RCTs.

We simulated RCT data using diverse assumptions for the average treatment effect, a baseline prognostic index of risk (PI), the shape of its interaction with treatment (none, linear, quadratic or non-monotonic) and the magnitude of treatment-related harms (none or constant independent of the PI). The combination of these settings resulted in the definition of 648 simulation scenarios. In each sample we predicted absolute benefit using: models with a constant relative treatment effect; stratification in quarters of the PI; models including a linear interaction of treatment with the PI; models including an interaction of treatment with a restricted cubic spline (RCS) transformation of the PI with 3, 4, or 5 knots; an adaptive approach using Akaike’s Information Criterion. We evaluated predictive performance using root mean squared error and measures of discrimination and calibration for benefit.

The linear-interaction model and the RCS-interaction (3 knots) model outperformed the constant treatment effect model in many simulation scenarios. The RCS-model was optimal when quadratic or non-monotonic deviations from a constant treatment effect were stronger, and when sample size was larger. Larger sample size also supported the use of the adaptive approach. Increasing the number of knots in the RCS models did not provide any improvement in predictive performance, while often tended to be overfitted, especially in scenarios with smaller sample sizes. All the simulation results can be explored at <https://arekkas.shinyapps.io/simulation_viewer/>. We also illustrated the application of the considered methods in the GUSTO-I trial which compared tissue plasminogen activator treatment to streptokinase in patients with acute myocardial infarction. The constant treatment effect method, the linear interaction method and the RCS with 3 knots had very comparable performance.

In conclusion, an interaction between baseline risk and treatment assignment should be considered to improve treatment effect predictions. Non-linear interactions should be considered only in larger sample sizes.