MSR17

CAN MACHINE LEARNING SUPPORT SURVIVAL MODEL SELECTION TO INFORM ECONOMIC EVALUATIONS? EXPLORING K-FOLD CROSS VALIDATION BASED MODEL SELECTION IN SEVEN DATASETS

Bermejo I,[1] **Grimm S**[2]

[1]*Hasselt University, Liege, Belgium,* [2]*Maastricht University Medical Centre (MUMC+), Maastricht, Limburg, Netherlands*

Objectives: The selection of survival models for informing economic evaluations of innovative therapies with limited long-term data traditionally relies on metrics of statistical goodness of fit in the full trial data. However, models selected based on full trial data might underperform in the target population due to overfitting. K-fold cross validation (CV), commonly used in machine learning, splits the data allowing be estimation of fit in unseen data. We explore whether k-fold CV improves model selection. ***Methods:*** We used seven publicly available long-term survival datasets covering a range of diseases. We simulated 100 artificial data locks by sampling 250 patients without replacement, and right-censoring once median sur-vival was reached. We fitted standard parametric and flexible survival models to each simulated dataset and selected models with lowest AIC/BIC as estimated using 10-fold CV and traditional methods. We then estimated the restricted mean survival time (RMST) error of best-fitting models relative to the RMST calculated from the full dataset's Kaplan-Meier. ***Results:*** K-fold CV led to lower mean RMST errors compared to traditional model selection methods in six (all seven) datasets when selecting models based on AIC (and BIC). On average, the RMST error was 27% higher (when based on AIC) and 40% (BIC) higher using traditional model selection compared to CV-based model selection. CV never selected complex models (3+ parameters) whilst the traditional method resulted in complex models being selected in 51% (AIC) and 12% (BIC) of simulations. ***Conclusions:*** In the first study exploring k-fold CV for survival model selection, we show that it can regularly outperform traditional methods. Notably, k-fold CV favors less complex models compared to traditional methods, which may hint at their better generalizability. We conclude that k-fold CV may be an important addition to the modeler's toolbox when performing survival analysis. Further research should explore whether these findings hold in additional settings.