Development and Evaluation of an Automated Machine Learning Algorithm for In-Hospital Mortality Risk Adjustment Among Critical Care Patients

Ryan J. Delahanty, David Kaufman, Spencer S. Jones

David Soto, Fabien Merceron, Celine Hajjar, Dina El Zein
Biostatistics – MVA 2021

Overview

Contributions

- Risk adjustment algorithms are important for the improvement of healthcare quality.
- Many tools are available but are not widely adopted, due to high licensing costs and time-intensive data collection.
- Main contribution of this study: development of a highly accurate risk adjustment algorithm with minimal cost (data- and licensing- wise).

Materials

Hospital and ICU Selection

79 care hospitals across the United States participated

Selected Criteria:

- has at least one ICU that meets the National Healthcare criteria
- not children hospital
- a minimum of 100 ICU patients in the study period
- use the center Millennium EHR system

➔ 53 hospitals and 131 ICUs met these criteria

Patient Selection

- 18 years or older patients that spent a portion of their hospital stay in an ICU
- unknown discharge disposition patients were excluded from the study

➔ In total 237,173 patients included

Methods

Feature Selection

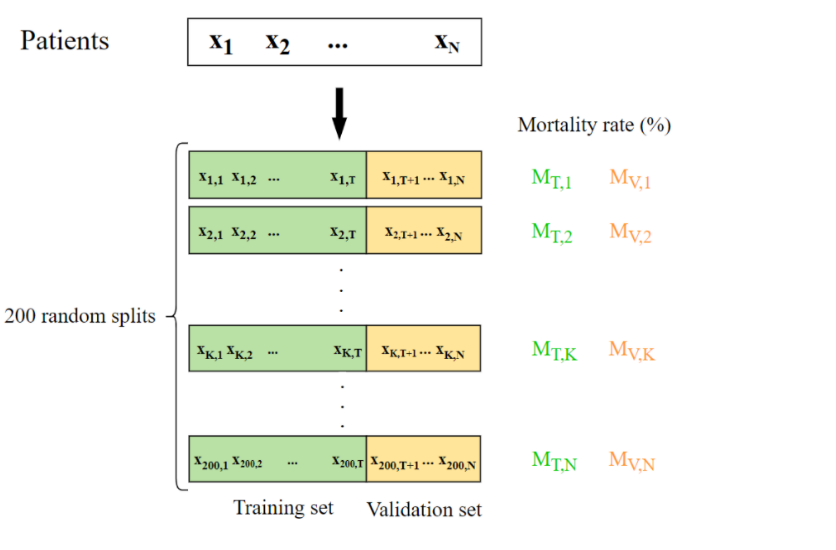
Clinical attributes available in the EHR and in administrative data amount to more than 15000 total features.

➔ Feature selection in 3 steps:

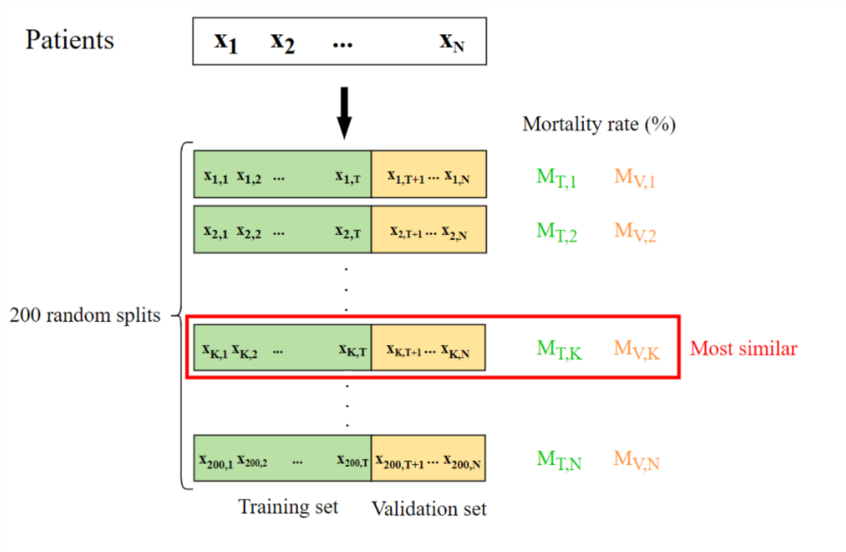
1. Environmental scan of existing algorithms
2. Consultation with a professional
3. Automated selection via Machine Learning models

➔ Steps 1 and 2 result in 215 selected features

Model Development and Evaluation



Model Development and Evaluation



Model Development and Evaluation

- After the 2 stages of feature selection, 215 features left
- XGBoost : Final **feature selection** and **risk adjustment model**

  Ensemble of simple decision trees

  Maximize Area Under Curve (AUC)

  Measure of features' relative influence : "gain"
- At each iteration ⅓ of the features is **removed**
- Reduced the number of features from 215 to 17

Results

RIPD Score Calibration

- Model accuracy and calibration were assessed using the adjusted Brier score and visual analysis of the calibration curves.
- The **adjusted Brier score** represents the percent reduction in deviation when using the RIPD model as opposed to assigning all patients a risk score equal to the average mortality for the entire population.

RIPD Score Calibration

- The adjusted Brier score for the model is 52.8% (good calibration !)

Visual Analysis

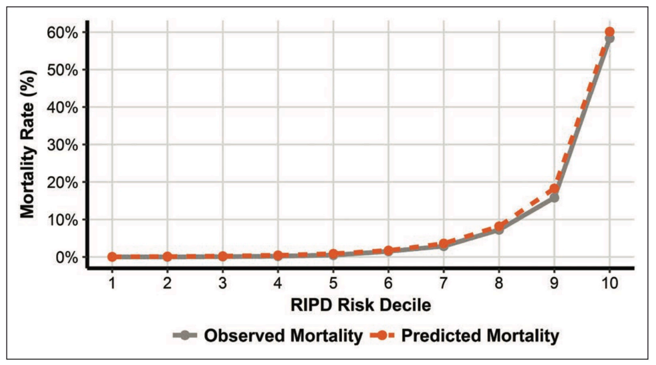


Figure 1: Calibration curves showing observed and predicted mortality rates across deciles of the Risk of Inpatient Death (RIPD) score distribution.

RIPD Score Discrimination

- Model discrimination was evaluated based on the area under the ROC curve (AUC)
- AUC is 0.94 (good discrimination !)

Discussion

Limitations

- **Data Leakage in the model development:** use the validation set to choose the best train-validation split
- Hospital split for train-validation sets is biased
- Ambiguous definition of the RIPD score

Thank you for your attention!