Heart Failure Survival Study

Group 6 Final Project Proposal

Huanyu Chen, Runze Cui, Jiahe Deng, Yi Huang, Xuesen Zhao

2023-10-09

Objective:

The goal of our course project is to perform a re-analysis of the survival data from the paper entitled "Survival Analysis of Heart Failure Patients: A Case Study". This paper focused on heart failure patients admitted to the Institute of Cardiology and Allied Hospital in Faisalabad, Pakistan between April and December 2015. The original study highlighted key risk factors, including age, renal impairment, blood pressure, and ejection fraction, which significantly contribute to mortality in heart failure patients. Through this re-analysis, we aim to further investigate and validate the identified risk factors associated with mortality in heart failure patients. Going beyond the foundation established by this paper and their study, our project aims to re-evaluate and potentially refine the predictive model from other new points of view.

Background:

According to the World Health Organization (WHO), cardiovascular heart disease (CHD) is the leading cause of 31% of global deaths, and Pakistan is witnessing a significant rise in CHD prevalence. A report from Al-Shifa Hospital revealed that 33% of the Pakistani population over 45 has hypertension, 25% suffers from diabetes mellitus, and CHD-related deaths have reached approximately 200,000 per year, translating to 410 per 100,000 of the population. These factors contribute to an increased prevalence of heart failure in Pakistan, with an estimated rate of 110 cases per million.

The study included 299 heart failure patients, comprising 105 women and 194 men, all aged over 40, with left ventricular systolic dysfunction falling into NYHA class III and IV. The follow-up period ranged from 4 to 285 days, with an average of 130 days. TAge, serum sodium, serum creatinine, gender, smoking, Blood Pressure (BP), Ejection Fraction (EF), anemia, platelets, Creatinine Phosphokinase (CPK) and diabetes were considered as potential variables explaining mortality caused by CHD. Age, serum sodium and CPK are continuous variables whereas EF, serum creatinine and platelets were taken as categorical variables.

Planned Analysis

1) Perform data pre-processing and exploratory data analysis to visualize the distributions of our variables of interest.
2) Compute Kaplan-Meier and Nelson-Aalen survival curves to estimate the survival probabilities over time for different groups.
3) Use the Log-rank test and Wilcoxon test to compare survival curves between different groups to assess if there are statistically significant differences in survival time.
4) Choose an appropriate survival model, comparing the sem-parametric model (Cox model) with parametric models (Exponential, Weibull, Gompertz, and Log Logistic Distribution).
5) Check the goodness of fit by applying Schoenfeld residuals test within the Cox model, Standardized residuals for the parametric models, and compute the confidence interval for the hazard ratio.

Team members' responsibilities:

Team Lead: Xuesen Zhao

Planned Analysis 1 & Background & Introduction (Runze Cui)
Planned Analysis 2 & Methodology & Data Collection (Yi Huang)
Planned Analysis 3 & Data Visualization (Xuesen Zhao)
Planned Analysis 4 & Results (Huanyu Chen)
Planned Analysis 5 & Conclusion & Discussion (Jiahe Deng)

Planned timelines:

Oct. 16: Proposal Final Draft
Oct. 30: EDA, Modeling, Graphs
Nov. 23: First Draft of Report
Dec. 10: Final Draft

Bibliography:

Ahmad, T., Munir, A., Bhatti, S. H., Aftab, M., & Raza, M. A. (2017). Survival analysis of heart failure patients: A case study. PLoS ONE, 12(7), e0181001. https://doi.org/10.1371/journal.pone.0181001