International Journal of Population Data Science

Journal Website: www.ijpds.org

Ensemble-based Classification Models for Predicting Post-Operative Mortality Risk in Coronary Artery Disease

Brobbey, A[1], Faris, P[2], Nettel-Aguirre, A[1], Williamson, T[1], Wiebe, S[1], Peng, M[1], Quan, H[1], and Sajobi, TT[1]

[1]University of Calgary

[2]Alberta Health Services

Introduction

There has been an increased demand for more accurate prediction tools to aid clinical decision-making regarding disease diagnosis prognosis for coronary artery disease(CAD) patients. Patients undergoing CABG surgery are older and a larger number have had previous heart surgery. Consequently, mortality after CABG is expected to increase despite procedural advances.

Objectives and Approach

This study aims to compare the predictive performance of random forest(RF) and logistic regression(LR) classifiers for predicting 30-day and 1-year post-operative mortality risk in CAD patients who underwent CABG. Data was obtained by linking the Alberta Provincial Project for Outcome Assessment in Coronary Heart Disease(APPROACH) registry, a prospective longitudinal data of patients undergoing cardiac catheterization in Alberta, Canada, to vital statistics database. All patients who underwent first-time isolated CABG between January 1, 2007 and December 31, 2012 were included in the analysis. Area under the receiver operating curve(AUC) was used to compare the predictive performance of LR and RF regression.

Results

Of the 4,908 eligible subjects who underwent isolated CABG during the study period, mortality estimates of 30-day and 1-year post CABG surgery were 1.59% and 3.85%, respectively. Descriptive analysis revealed that age, sex, hypertension, dialysis, cerebrovascular disease, chronic obstructive pulmonary disease, and chronic heart failure were associated with 30-day and 1-year mortality. The accuracy of the LR and RF regression classifiers in predicting 30-day mortality were 74.1, and 99.7%, respectively. While the accuracy of the former and latter classifiers in predicting 1-year post CABG mortality were 74% and 97.4%, respectively.

Conclusion/Implications

This study shows that RF classifier results in better predictive accuracy than LR in predicting post-operating mortality risk in CAD patients. Machine learning models are potentially usefully for developing clinical prediction models that can be used to aid the monitoring of post-discharge outcomes in the management of cardiovascular diseases.

http://dx.doi.org/10.23889/ijpds.v3i4.877