RANDOM SURVIVAL FORESTS IN THE PRESENCE OF COMPETING RISKS

Freshteh osmani[1], Atefeh hajian[2], Ebrahim hajizadeh[3,*]

Department of Biostatistics, Faculty of Medical Sciences, Tarbiat Modares University,Tehran,Iran

Correspondence:
Tel: +98.82883080, E-mail: hajizadeh@modares.ac.ir

TYPE OF ARTICLE: CONFERENCE ABSTRACT

ABSTRACT

Introduction: Random forests in medical research are a popular method for creating models to predict event risk. This method requires no assumption, and one of the advantages of this method is election prediction variables in order of importance. This study compares the competing risks forests as a nonparametric method with cause-specific Cox as a semi-parametric method, along with specifying risk factors of colorectal cancer.

Methods: This study is a retrospective cohort study. Data were collected from 1000 patients with colorectal cancer in the Taleghani Hospital between the years 1996–2011. By using a cross-validation method, two criteria of the competing risks forests (splitting log-rang and gary) and cause-specific Cox regression models have been fitted to the data. The best model to predict was determined by using concordance index and integrated brier score (IBS) in the study.

Results: A total of 1000 patients 37.9% women with mean (±SD) age of 52.04 (±15.534) and 62.1% male with mean (±SD) age of 53.41 (±16.358) participated in this study. The IBS for splitting log-rang, gary, and cause-specific Cox regression is .175, .178, and .185, and the values of cross-validation index for the three models are .714, .711, and .649, respectively.

Conclusion: According to the results of this study, the competing risks forest model has better performance in predictability compared with the other two models.

KEYWORDS: Tree-Based models, Random survival forests, Competing risk forest, Colorectal cancer