The Effect of Somatic Mutations in Glioblastoma Multiforme on Patient Survival

Pandey, Neelesh (School: Gulliver Preparatory School)

Glioblastoma multiforme is one of the deadliest forms of brain tumors. Many genes are associated with its formation, progression, and survival, but no consensus has been reached on which somatic mutations have the most significant impact on survival. This study hypothesized that patients with an increased age at diagnosis, EGRF mutations, and no mutations in the IDH1, ATRX, and RB1 genes would have lower survivability. A Cox Proportional Hazards model is used to predict survivability after diagnosis based on tumor mutational profiling. The model contains 499 genes and their respective somatic mutation status from 280 individuals using a derived dataset from the TCGA database of the National Cancer Institute posted on Kaggle (https://www.kaggle.com/datasets/palashio/glioblastomamutations). A univariate Cox Proportional Hazards model was first implemented for each gene and clinical category to find features that had a significant ($p \leq 0.05$) effect on survival. After splitting the dataset into training and testing groups using a 70-30 split, a multivariate Cox Proportional Hazards model was then composed of 11 significant genes (NLRP4, CALN1, DNAH17, ITGB4, ITGAD, PCDHA12, RGPD4, SPEG, DST, SRCAP, and RIPK3) and 2 clinical characteristics (Age at Diagnosis and Gender) from the univariate analysis. It was evaluated using the concordance index (0.67) and time-dependent ROC curves at 100 days (AUC=0.67), 300 days (AUC=0.74), and 500 days (AUC=0.76). The effect of each significant gene and clinical characteristic from the multivariate model was then evaluated using Kaplan-Meier curves.