Spatial Tumor-Immune Analysis: Insights From Pathology Slides and Breast Cancer Survival

Wang, Jonathan (School: Colleyville Heritage High School)

Cancer is the second leading cause of death in the US. Among the various forms of cancer, breast cancer and lung cancer are particularly significant due to their prevalence and impact. Breast cancer in particular contributes to around 30% of all new female cases each year, while also having some of the highest mortality rates. Scientists and doctors rely on pathology slides to aid in the discovery of a cure, diagnose patients, and provide treatment. These slides play a crucial role in examining samples and identifying any abnormalities. The primary goal of this project was to analyze pathology slides from 873 cancer patients in The Cancer Genome Atlas Breast Invasive Carcinoma (TCGA-BRCA). A spatial tumor and immune analysis was conducted with the hypothesis that a quantitative analysis of cell type specific clusters in the spatial context could lead to novel insights on patient survival. First, tumor and immune cells were identified using a HD-Yolo algorithm. Then tumor clusters and immune cell clusters were created with DBSCAN. Next, descriptive statistics such as Jaccard distance, Hausdorff distance, Earth Mover's distance, Jensen-Shannon divergence, Kernel Density Estimation, tumor density, and immune cell density were derived and correlated with the patients survival while adjusting for clinical attributes such as patient age and tumor stage using Cox Proportional Hazard models. Cox regression graphs, Kaplan-Meier graphs, and logrank tests were also used for univariate visual analysis. The results discovered spatial attributes and known clinical risk features associated with survival.

Awards Won: