Machine Learning Tool for Accurate, Cost-Effective, and Rapid Prediction of Small Cell Lung Cancer Stage: Significantly Increase Patient Survival in One Minute

Baral, Jessika

Small cell lung cancer (SCLC), the most fatal subtype of lung cancer, has a five-year survival rate of 2%. Currently, patients must get expensive computerized tomography and positron emission tomography scans, then wait weeks for a complete diagnosis. There exist no rapid, cost-effective, and comprehensive tools to predict SCLC stage using just the standard biopsy. Nuclear Factor I/B (NFIB) is an oncogene in SCLC and is important for lung maturation in human embryos. Alamar blue growth assays and western blots were used to establish the importance of Nfib to SCLC metastasis and growth. When Nfib expression was knocked down using shRNA, there were fewer metastatic liver tumors in comparison to the control. Immunohistochemistry assays were carried out to stain Nfib in human lung tissue biopsies. A novel bioinformatics image-processing tool, OncoDetector, was created to analyze digital images of these biopsies. A series of machine learning logistic regression classifiers were implemented for binary classification. First, the model was trained to determine whether the biopsy was cancerous or not; then it was trained to predict whether the biopsy represented limited stage SCLC or extensive stage SCLC. The overall tool proved to be highly accurate -- 95.11% accuracy. Doctors can directly use OncoDetector to accurately predict stage of SCLC in less than one minute, saving patients both time and tens of thousands of dollars. This model will allow doctors to better guide their patients' treatments to treat SCLC in the hundreds of thousands afflicted worldwide and it is currently being tested by oncologists.

Awards Won:
Fourth Award of $500