AACR 2018 – Special Conference: AI, Big Data & Prediction in Cancer

Title: Precision medicine in lung cancer: towards predicting recurrence for early stage disease

Donald J. Johann Jr, Jason Muesse, IJ Shin, Konstantinos Arnaoutakis, Susanne K Jeffus, Sanjay Maraboyina, Erich Peterson, Matthew Steliga

Winthrop P Rockefeller Cancer Institute, University of Arkansas for Medical Sciences, Little Rock, AR 72205

An unmet medical need is to improve cure rates in lung cancer patients with early stage disease. Lung cancer is the leading cause of cancer related deaths for men and women in the United States and worldwide. In Arkansas the incidence of lung cancer has been higher than the national average for more than 20 years. In response to these issues, an advanced clinical trial has been initiated focused on fostering an improved understanding of lung cancer pathogenesis and recurrence. Cornerstones include novel molecular assays, model systems, and bioinformatics.

Specifically, this trial involves liquid biopsy methodologies, a co-clinical trial approach including tumor explant and organoid applications, along with advanced bioinformatics. In an effort to improve cell-free DNA understanding, in addition to lung cancer patients, other cohorts for liquid biopsy assays are being enrolled and include: i) normal volunteers having blood collection before and after ~30 – 45 min of cardio-level exercise, ii) smokers at risk for cancer (> 30 pack years) but having a negative low-dose CT (LDCT) screening scan, and iii) patients with inflammatory disease (Rheumatoid Arthritis, Hepatitis C) with a negative cancer history.

To date, for the lung cancer cohort, over 60 patients have been enrolled with early stage disease who have been operated on “for cure”. It is known that the majority of these patients will recur. For this group, solid tumor tissue is subjected to extensive molecular profiling, namely, low pass WGS for CNV, a high coverage panel based on TCGA lung cancer findings and utilizing universal molecular identifiers (UMIs) for mutations, RNA-seq, and methylation. Plasma and urine for liquid biopsy analysis are collected pre-op, post-op, and at every standard of care visit. A detailed molecular profiling database has been designed to allow for integration and annotation of the big data from these molecular assays along with comprehensive clinical, animal and model system data.

Out of necessity, advanced molecular assays are compelling a convergence of high performance computing (HPC) and machine intelligence approaches. An argument may be made that increasingly, prescience may be possible in more oncologic instances, especially regarding disease recurrence and/or *de novo* presentations. HPC and machine intelligence software subsystems are now commodities. What is not a commodity are the well annotated cancer datasets composed of detailed and temporal-based molecular assay results along with high-quality clinical data. These are sorely needed and our clinical trial is addressing this need. Approaches for the structuring of such data for optimal machine learning are now in progress and will be discussed.