Creating an Automated Prediction Model for Chemotherapy Treatment Regimens through Molecular Feature Selection

Abstract

Introduction: Cancer is a leading cause of death around the world, and chemotherapy regimens are one of the most common cancer treatments. However, the current chemotherapy system is flawed because of limitations with *in vivo* testing that prevent physicians and patients from accurately identifying the most effective chemotherapy regimens. The objective of this study was to create a computational model to understand the molecular mechanisms responsible for regimen efficacy.

Methods: 14 molecular features were collected for 19 Chronic Myeloid Leukemia (CML) regimens. The regimens were grouped per each molecular feature through a k-means clustering algorithm. The cluster graphs were correlated to a network meta-analysis for CML.

Results: The computational model was able to isolate 2 molecular properties that appear to significantly influence the efficacy of chemotherapy regimens (polar surface area and acidic properties) and 2 properties that appear to have very little impact on determining efficacy (water solubility and polarizability).

Conclusion: This model allows us to understand the molecular mechanisms that are directly involved in determining regimen efficacy. The innovation of this study is that it can be applied to multiple types of cancer and can provide insights for drug design and regimen selection based on specific physicochemical traits.