**Deep Learning Prediction of Chemotherapy Response using Multi-Omics Features**

Shicheng Guo1, Steven J. Schrodi1, 2

1Center for Precision Medicine Research, Marshfield Clinic Research Institute

2Computation and Informatics in Biology and Medicine, University of Wisconsin-Madison

**Abstract**

Combining deep learning algorithms with multi-omics data provide an opportunity to improve precision medicine, especially for personalized therapy treatment. In this study, we investigated the performance of deep-learning algorithms applied to multi-omics genomic data from the TCGA project. Drug response status including complete response (n=815), partial response (n=102), stable disease (n=182) and clinical progressive (n=892) across 32 cancers were evaluated. For computational efficiency and classifier performance, we proposed an approach using Sure Independence Screening (SIS) and Gini impurity index for feature selection from ultrahigh dimensional genomic feature set followed by conditional generative adversarial networks (CGAN) to produce predictive models for chemotherapy response, internally validated by 10-fold cross-validation. mRNA, miRNA, and methylation datasets were used both independently and jointly. For the miRNA-based prediction, we identified 23 highly informative features, including miR-141, miR-200c, miR-205, miR-9 and miRR-338, which, in combination, produced an AUC=0.64 for discriminating complete and partial response from stable disease and clinical progressive endpoints. Using mRNA-seq data, we identified 264 highly informative features including *MYCBP*, *KLF15*, *IGIP*, and *GRIA3*. The mRNA-based predictive model yielded an average AUC=0.71. The DNA methylation model attained a higher level of performance with AUC=0.81 (95% CI: 0.78-0.84). Combining mRNA-seq, miRNA, and methylation data improved the predictive performance with accuracy of 84.2% and an AUC=0.86 (95% CI: 0.82-0.90). Functional enrichment analysis for mRNA and DNA methylation selected features showed enrichment for the antioxidant response pathway and basal transcription factors associated with the platinum drug resistance (PDR) pathway. In summary, this work demonstrates that applying deep-learning algorithms to multi-omics data can generate informative predictive models for chemotherapy response across several cancers.

32 cancer types can be found here:

<https://github.com/Shicheng-Guo/HowtoBook/blob/master/TCGA/drug_response/Pid.drugResponse.txt>