Investigation of Mechanisms of Response in Multiple Myeloma via Bayesian Causal Inference: An Early Analysis of the CoMMpass Study Data

Fred Gruber1, Boris Hayete2, Jonathan Keats3, Kyle McBride3, Karl Runge3, Mary DeRome3, Sagar Lonial3, Iya Khalil, Daniel Auclair4

1GNS Healthcare, 2Translational Genomics Research Institute, 3Instat, 4Multiple Myeloma Research Foundation, 5Department of Hematology & Medical Oncology, Emory University School of Medicine.

Introduction

Multiple myeloma (MM) is an incurable disease with a rapidly shifting treatment landscape that highlights the importance of a deeper understanding of drug response pathways to guide drug development and enable better drug targeting.

CoMMpass (NCT0145429) [1], a study by the Multiple Myeloma Research Foundation (MMRF), collects longitudinal data of newly diagnosed patients’ responses to treatment. The CoMMpass Interim Analysis 7 (IA7) dataset provides extensive clinical and molecular data on a population of almost 800 enrolled patients.

Objectives

- Build MM disease models to characterize the probabilistic network connections among variables from the multiple data modalities generated in the CoMMpass study.
- Run in silico simulations of the MM disease models to identify novel intervention targets for modulating MM clinical endpoints.
- Build a software interface for data analysis and simulation for the MMRF and CoMMpass trial collaborators.

Methodology

The GNS Healthcare REFS™ (Reverse Engineering, Forward Simulation) machine learning platform uses well documented mathematical techniques to infer causal relationships [2] in high dimensional datasets constrained by a minimal set of biological considerations but otherwise entirely de novo.

To capture variability in data and inference and to distinguish confident predictions from incidental ones, REFS™ returns an ensemble of models that are all consistent with the observed disease biology.

Simulations on this ensemble are then developed to find which variables are potential drivers of the outcomes.

Data

IA7 dataset includes clinical measurements (demographics, labs, treatment information, survival, etc), somatic single nucleotide variants (SNV), structural variants, somatic copy numbers (SCNV), and RNAseq gene expression.

The final dataset after preprocessing had 452 patients, and 28,200 variables.

Preprocessing:
- SNVs were first filtered with Strelka, MuTect, and Seurat and then aggregated into gene region burden scores.
- mRNA variables with zero expression in majority of samples were removed.
- SCNVs were segmented with the CBS algorithm.

Selected Results

Response Assessment: first treatment response that lasts at least 1 year before a progressive disease.

MM status: No response or sustained response

References


Conflict of Interest

There are no relevant conflicts of interest to disclose.