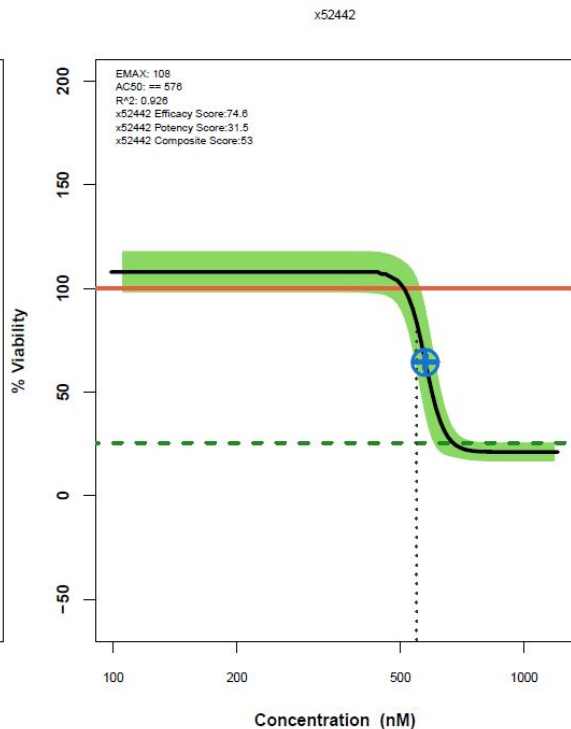
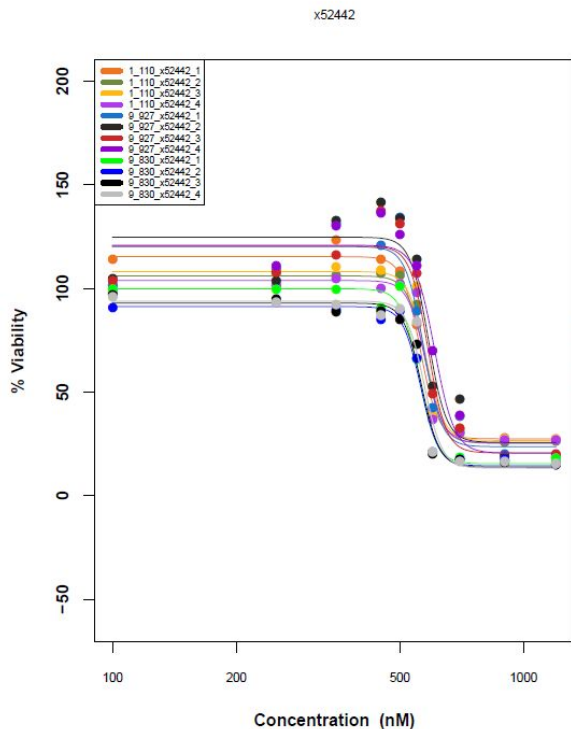


Progress Report

Jeremy Ash and Alex Larsen

Data

- Dose response measurements for approx. 90 cell lines collected from breast cancer patients
 - Percent viability measured at 10 different doses
 - Extensive QC pipeline



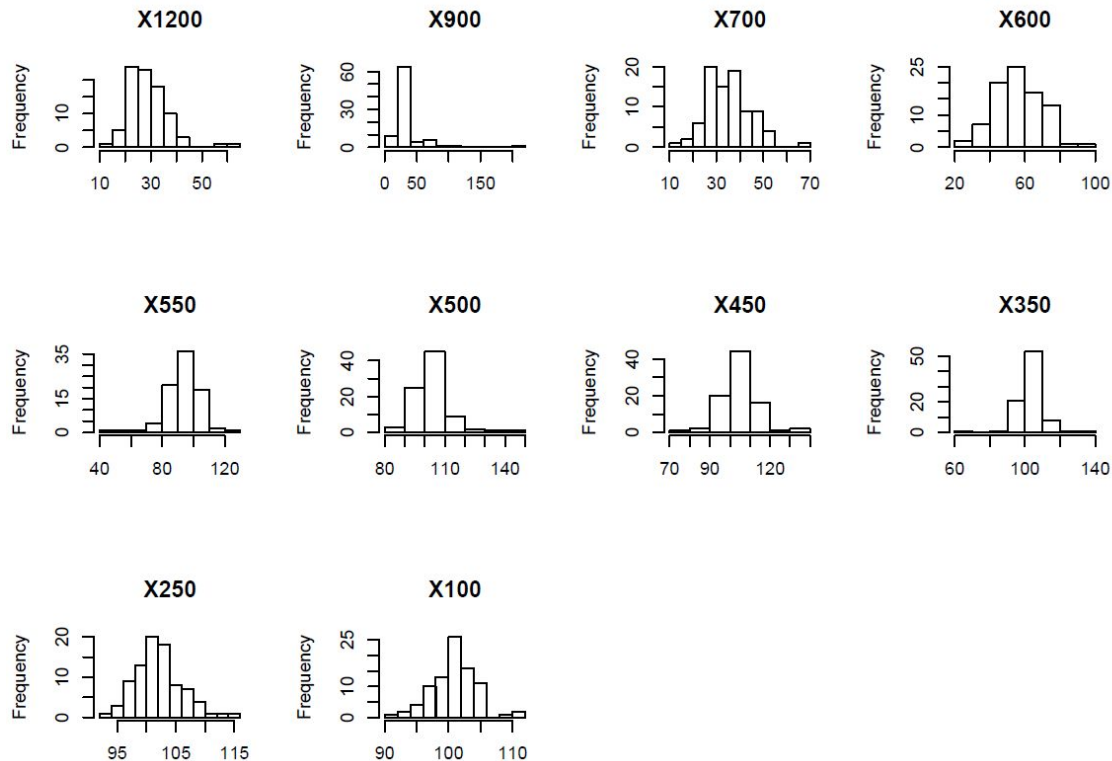
Data

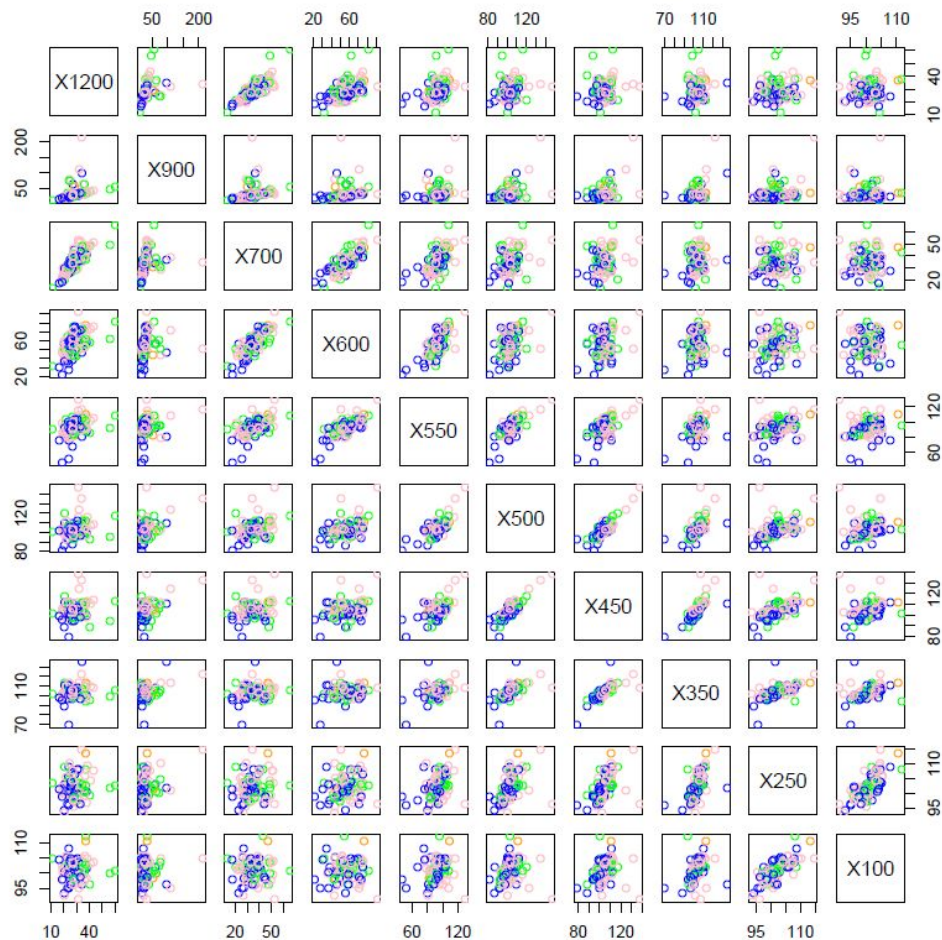
- Additional predictors variables
 - Demographic variables
 - Treatment regimens
 - Initial stage of cancer
 - etc.
- Response variable
 - Categorical variable for patient response to treatment (no response, partial response, etc)

Research Questions

- Is in-vitro exposure of lymphoblastoid cell lines to Paclitaxel a good predictor of cancer patient responses to chemotherapy?
 - Currently ordinal regression models
- Are there significant differences in the dose response profiles for patients in different demographic groups
 - MANOVA
 - 10 dose response variables are response

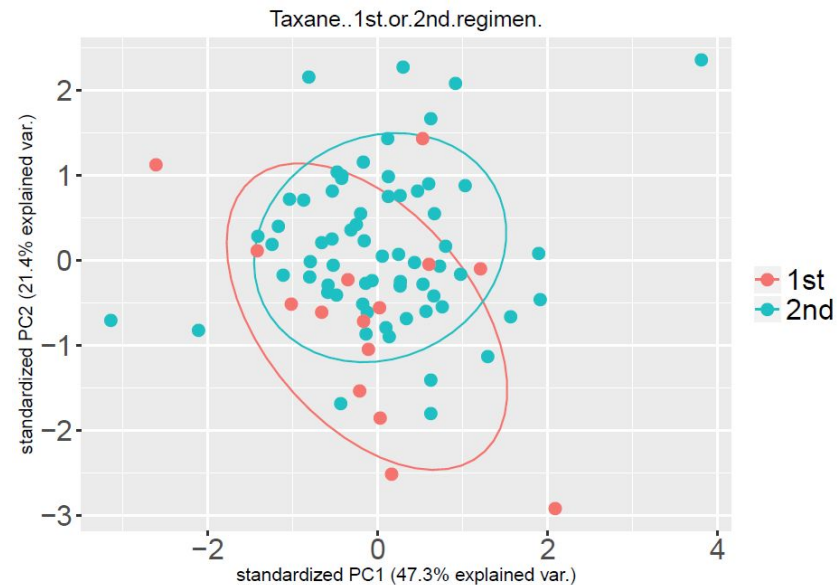
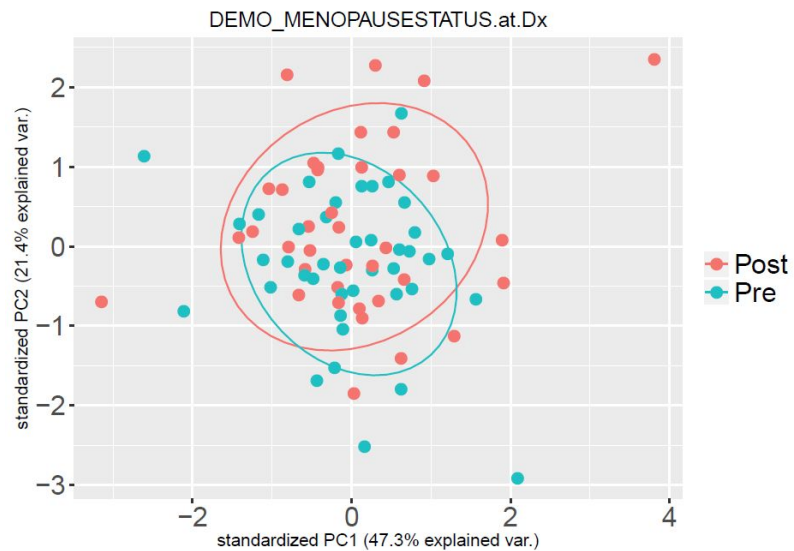
Initial Data Exploration



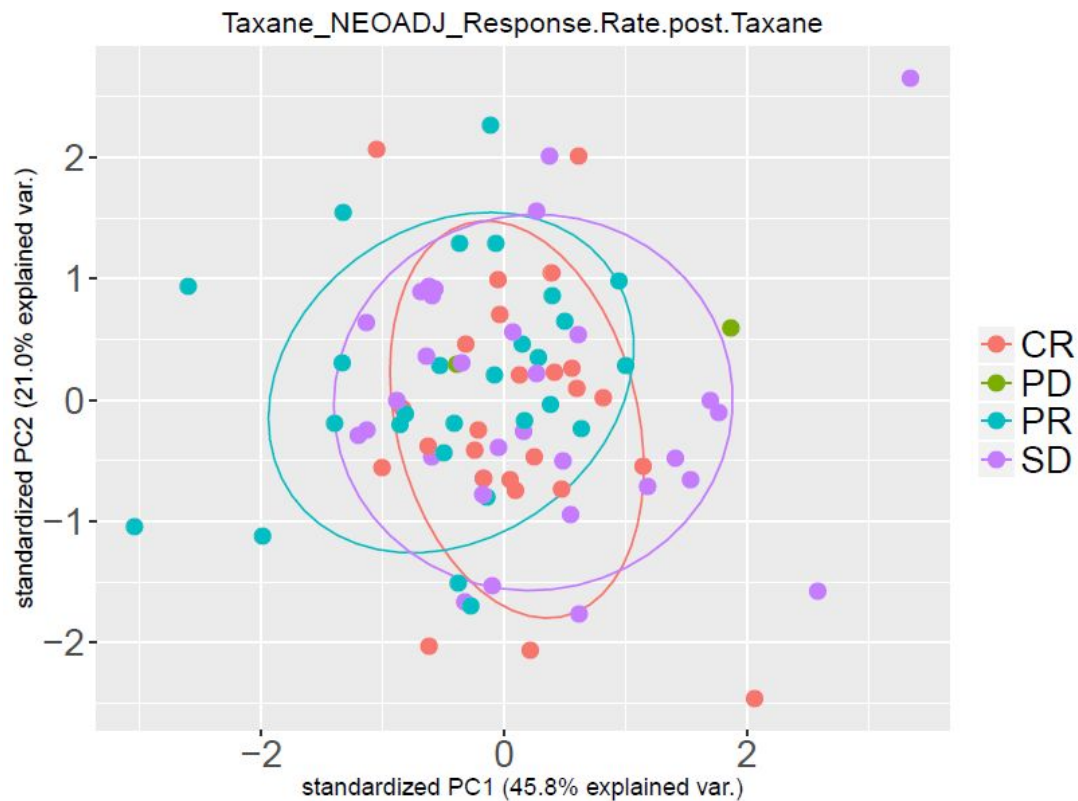


- Correlation between response will be an issue for MANOVA
 - How do we account for correlation structure?
- Highest pearson's correlation between one of the dose response variables and an interval version of our response is .25

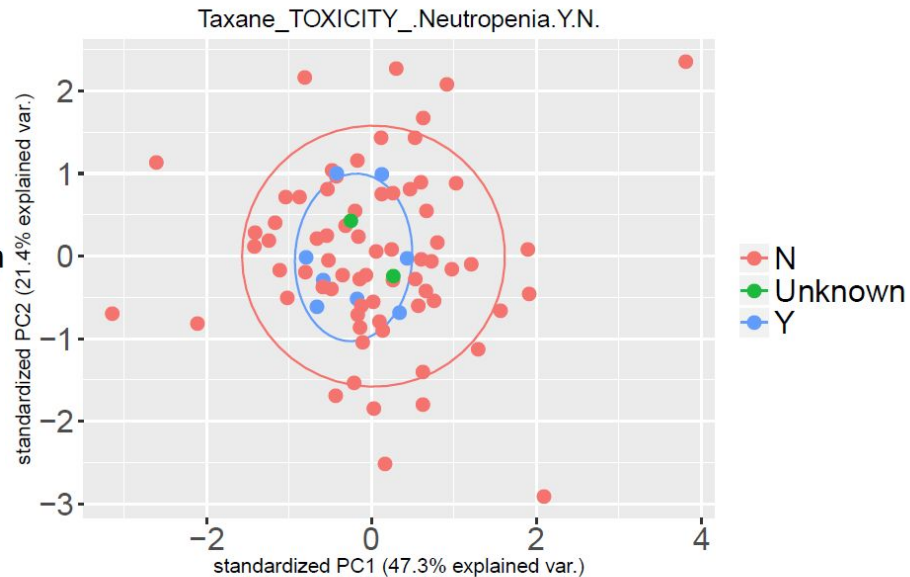
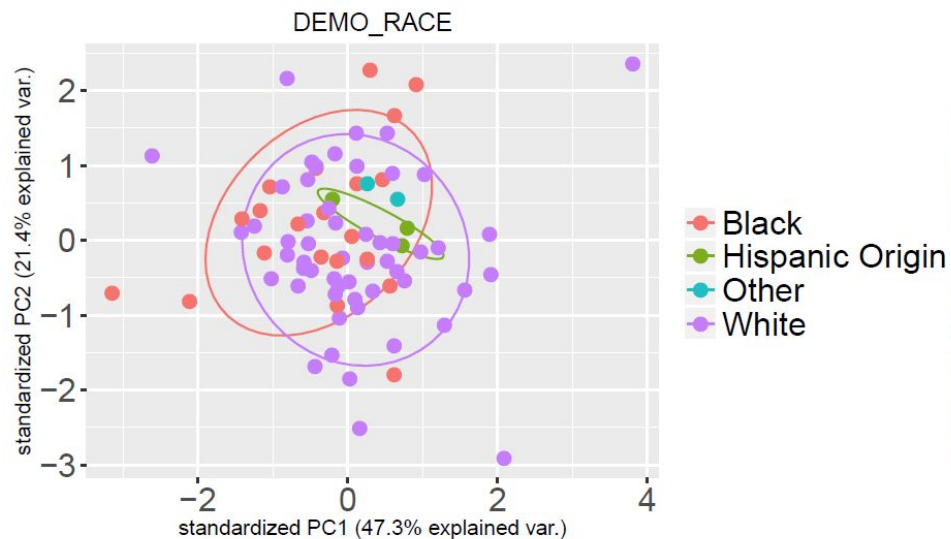
Initial Data Exploration



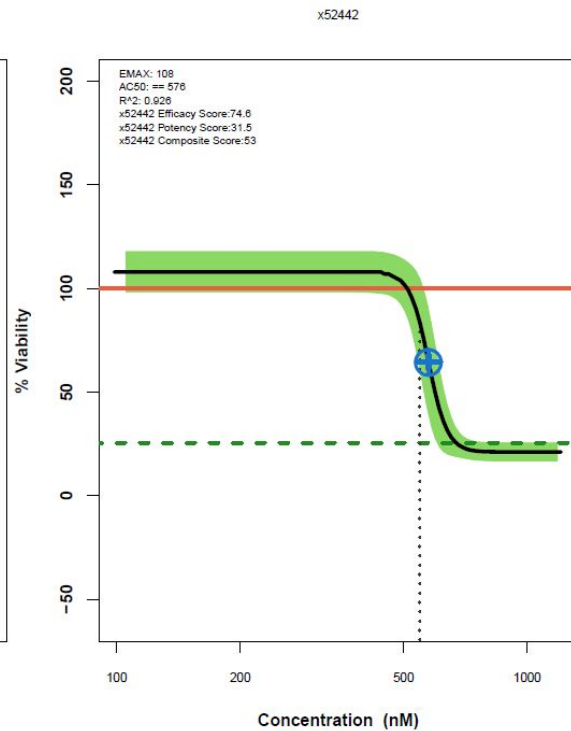
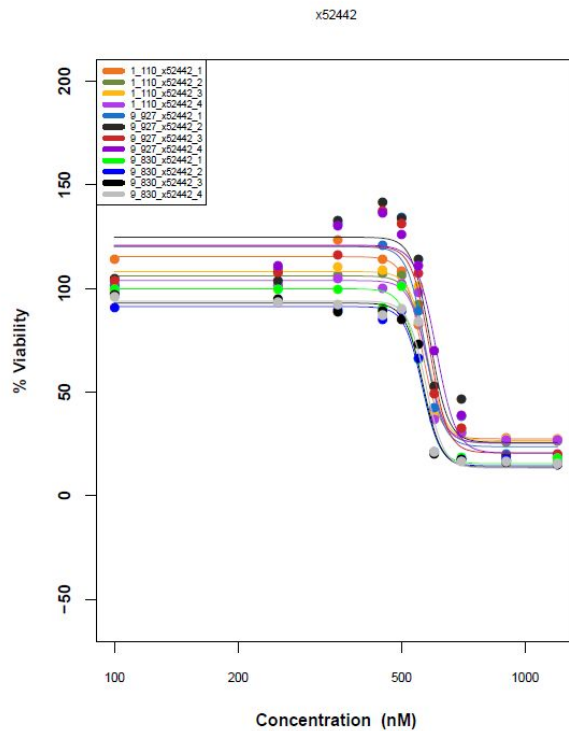
Initial Data Exploration



Initial Data Exploration



Curve Fit Data



QC Pipeline

- Data normalized to positive and negative controls (100% viability at negative control, 0% viability at positive control)
- Coefficient of variation smoothing
- Regression smoothing
 - Would do ideally, but could not get to work
- Monotonicity filter
- Removed plates with nearly all dead cells across all concentrations
 - Obviously sick
- Took the mean of remaining replicates for each dose response

Preliminary Results

Question: Is in-vitro exposure of lymphoblastoid cell lines to Paclitaxel a good predictor of cancer patient responses to chemotherapy?

Ordinal Logistic Regression:

Response: Recovery status (none, partial, full)

Baseline covariates: age, menopause, estrogen receptor, treatment regimen, number of cycles, total weeks, pre-experiment cancer stage, race, smoking status, Her2 (human epidermal growth factor receptor 2)

Additional covariates of interest: IC-50, dose response curve values at each of 10 doses (100, 250, ..., 1200)

Preliminary Results

Evaluate models via $AIC = 2p - 2\ln(L)$

Model	AIC
Baseline	311.695
Add IC-50 to the model	269.822
Add dose-response curve values to the model	281.681

Future Directions

- Variable selection (lasso, etc.)
- Ordinal regression with other response variables
- MANOVA
- Other machine learning techniques (Weka)