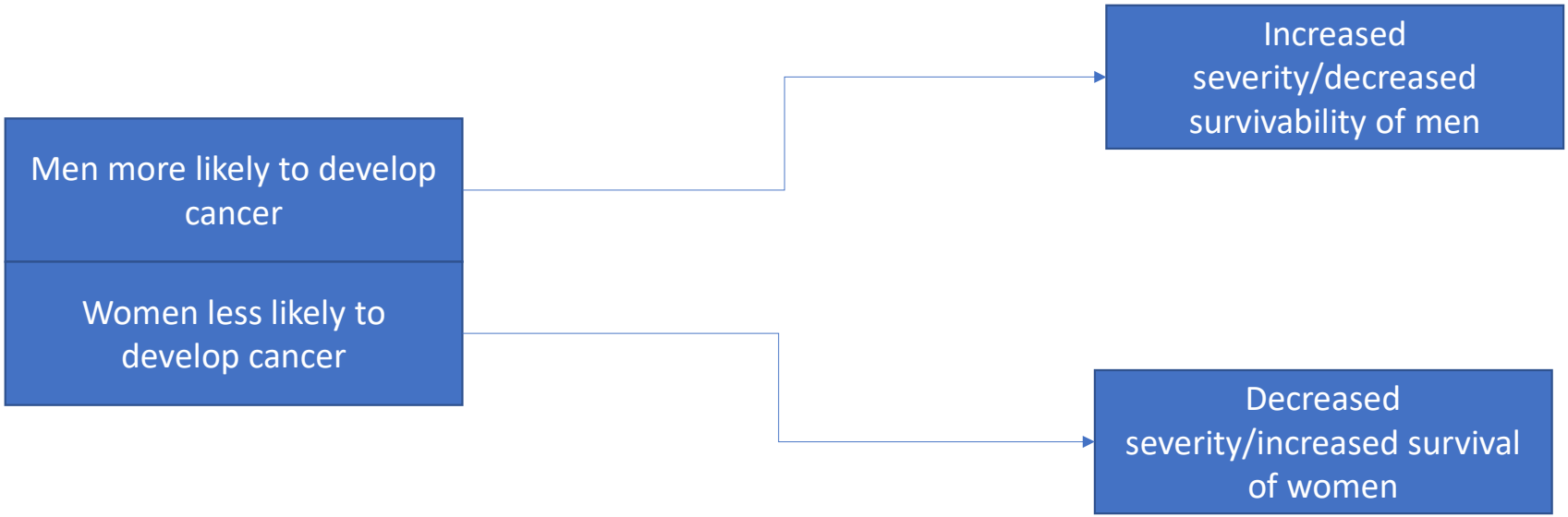
Acute Lymphocytic Leukemia: The Average Patient

By: Jordan Hendriksen

Hypothesis: The patient demographics profile will be different than the “average” American

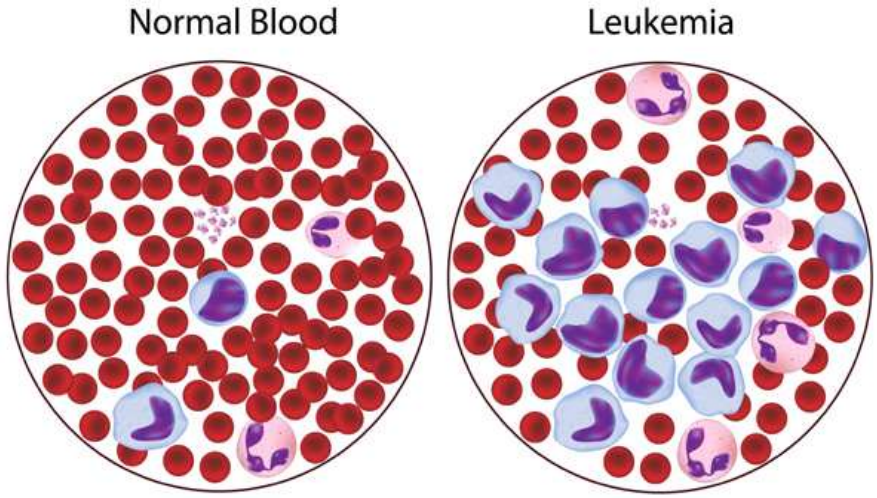
- Secondary: Demographics with differences will act as predictors of severity or survival



Acute Lymphocytic Leukemia - ALL

- Starts in bone marrow and develops from lymphocytes
- Approximately 6,000 new cases yearly in US
- Causes approximately 1,600 deaths in US yearly
- Lifetime risk is 1/1000
- If untreated often causes death within a few months

- Dataset:
  - Clinical data from Genomic Data Commons legacy archive (TCGA)
  - 1783 unique patients



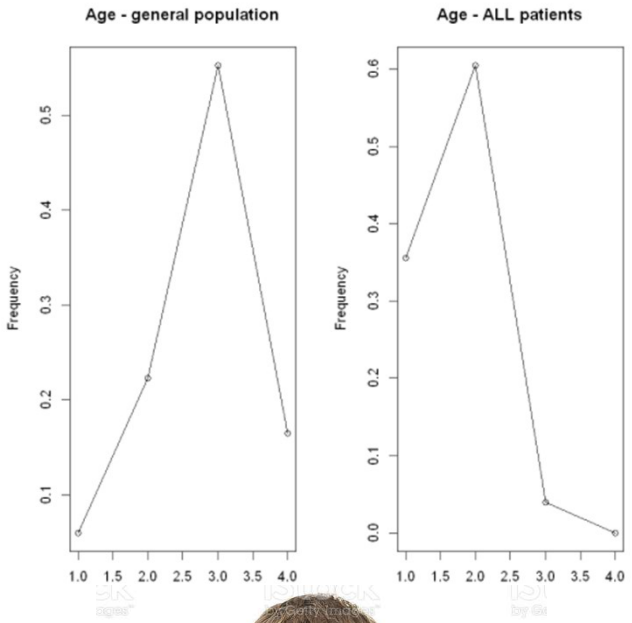
The Average American:

- Female – 50.8% vs 49.2%
- White – 78.6%
- Non-Latin – 76.5%
- Over 18 years old and under 65 years old – 55.2%

*data from 2019 census

The Average ALL Patient:

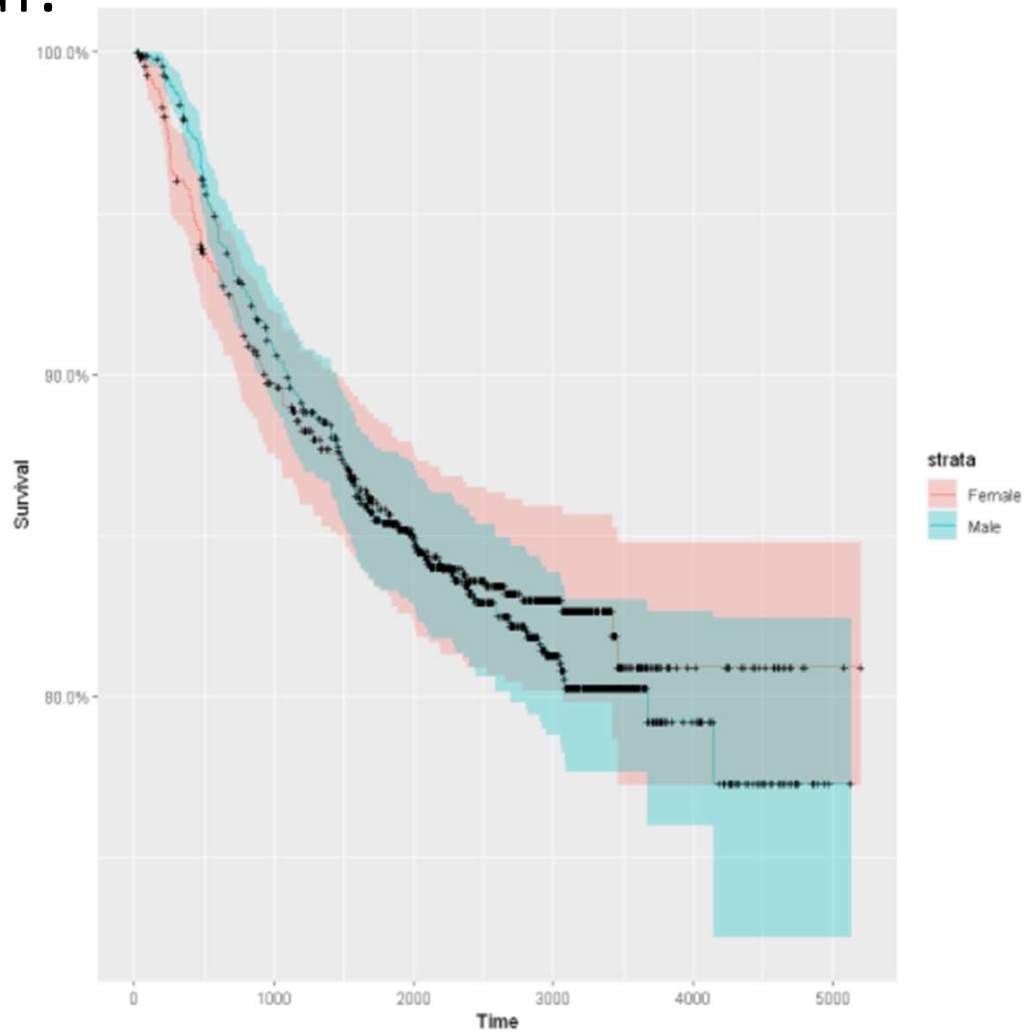
- Male – 59.7% vs 49.2% for general population
  - X^2 p-value < 2.2e-16
- White – 85.9% vs 78.6% for general population
  - X^2 p-value = 9.9e-16
- Non-Latin – 76.5% vs 81.5% for general population
  - X^2 p-value < 2.2e-16
- Young – Average age of 8.93 years – 5 years < 60.5% < 18 years
  - Vs 5 years < 22.3% < 18 years for general population
  - X^2 p-value < 2.2e-16



Demographics and Survival: Gender

Survival difference:

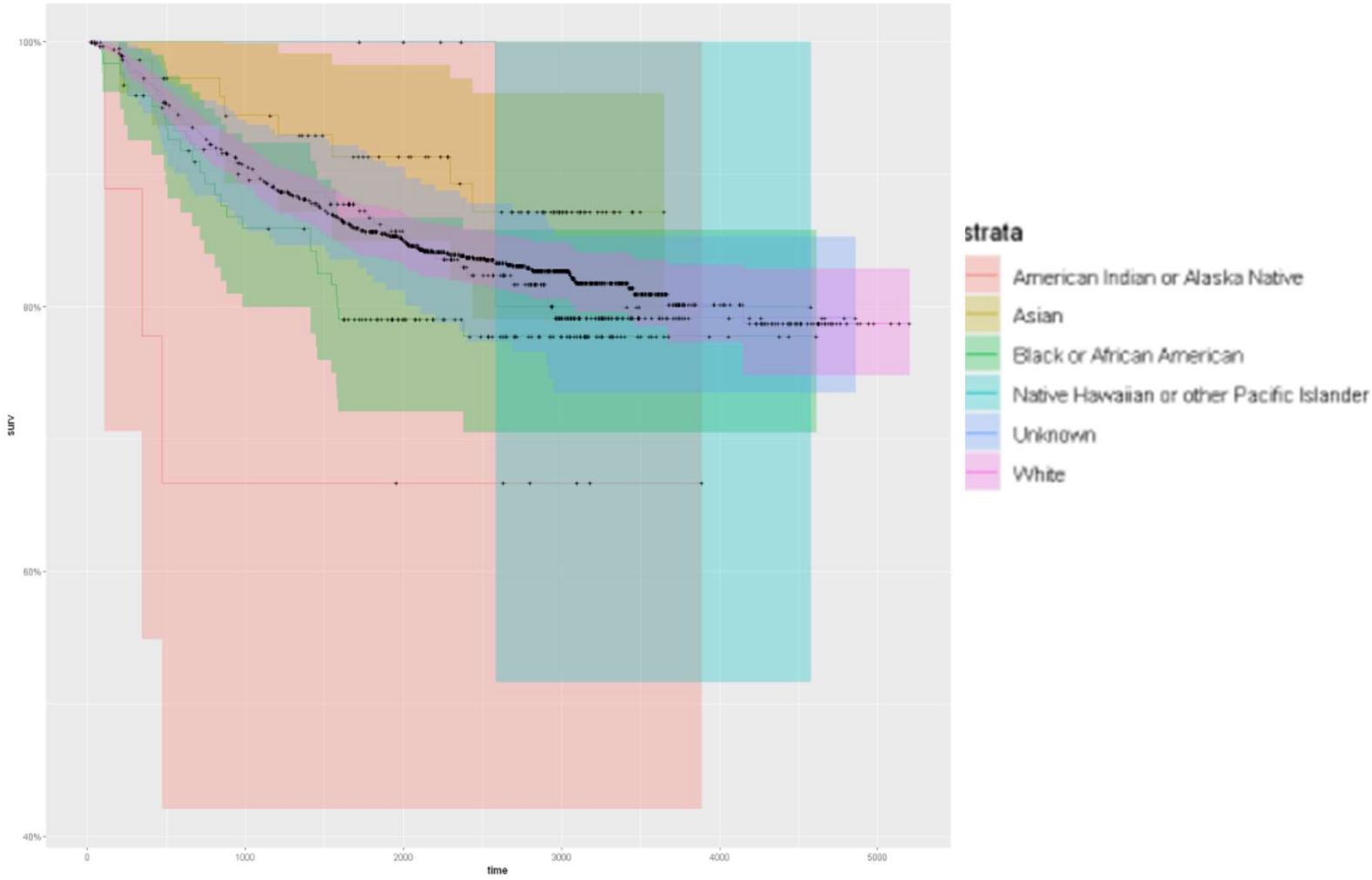
- 718 Female
- 1065 Male
- P-value = 0.6



Demographics and Survival: Race

Survival difference:

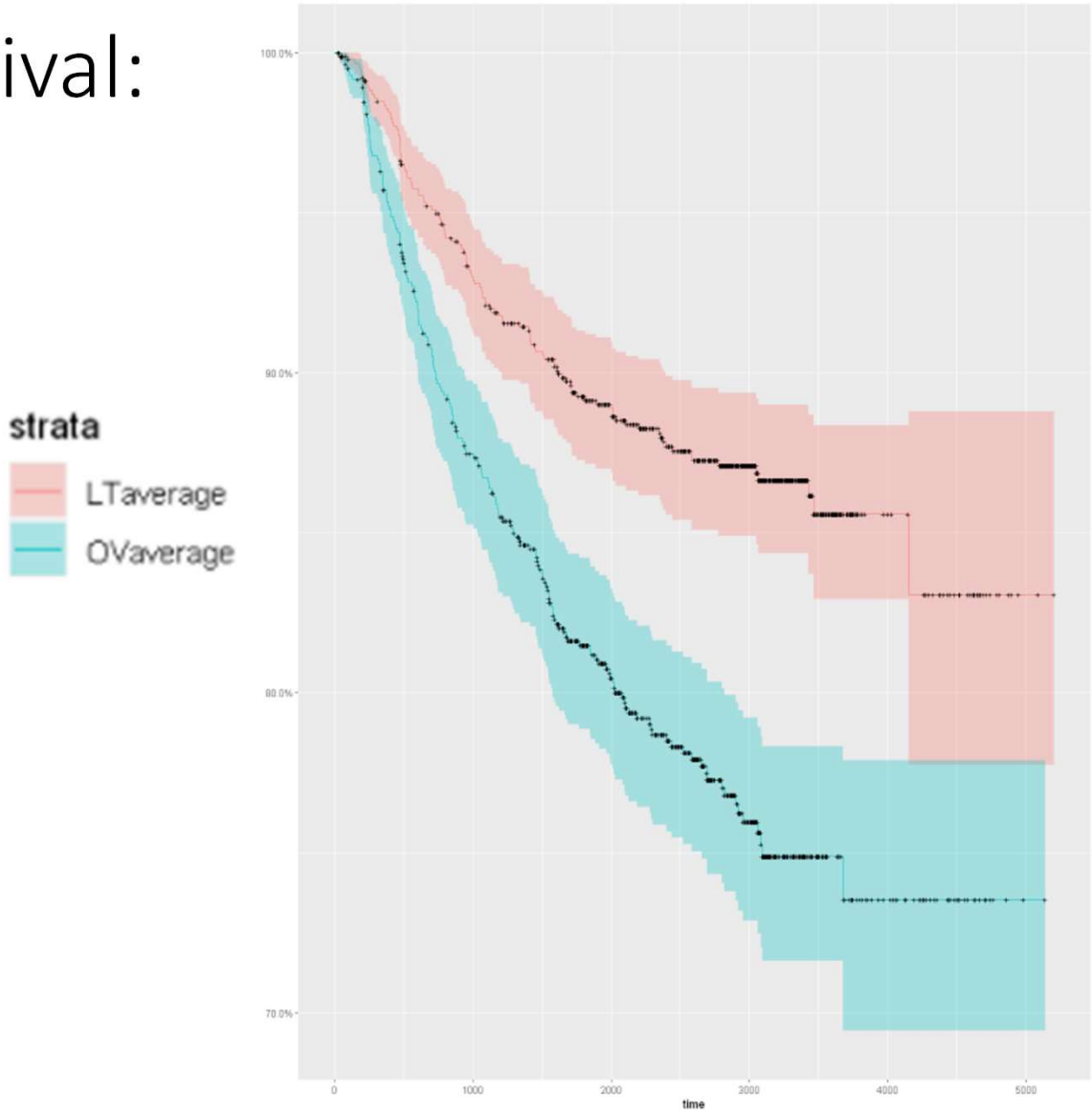
- P-value = 0.3



Demographics and Survival: Age at Diagnosis

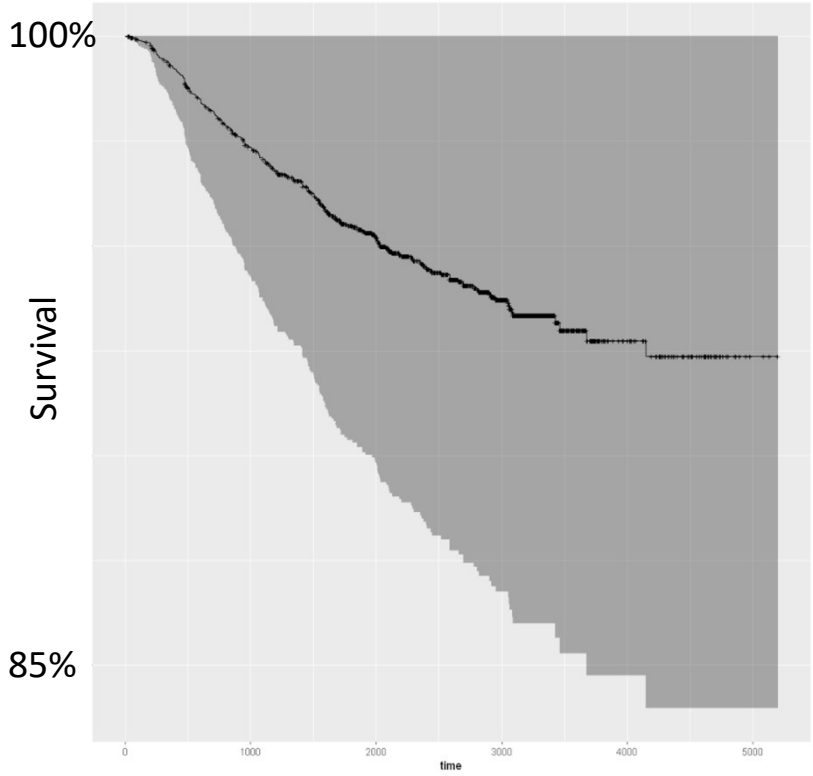
Survival difference:

- 934 younger than 3260 days
- 849 older than 3260 days
- P-value = 3e-8



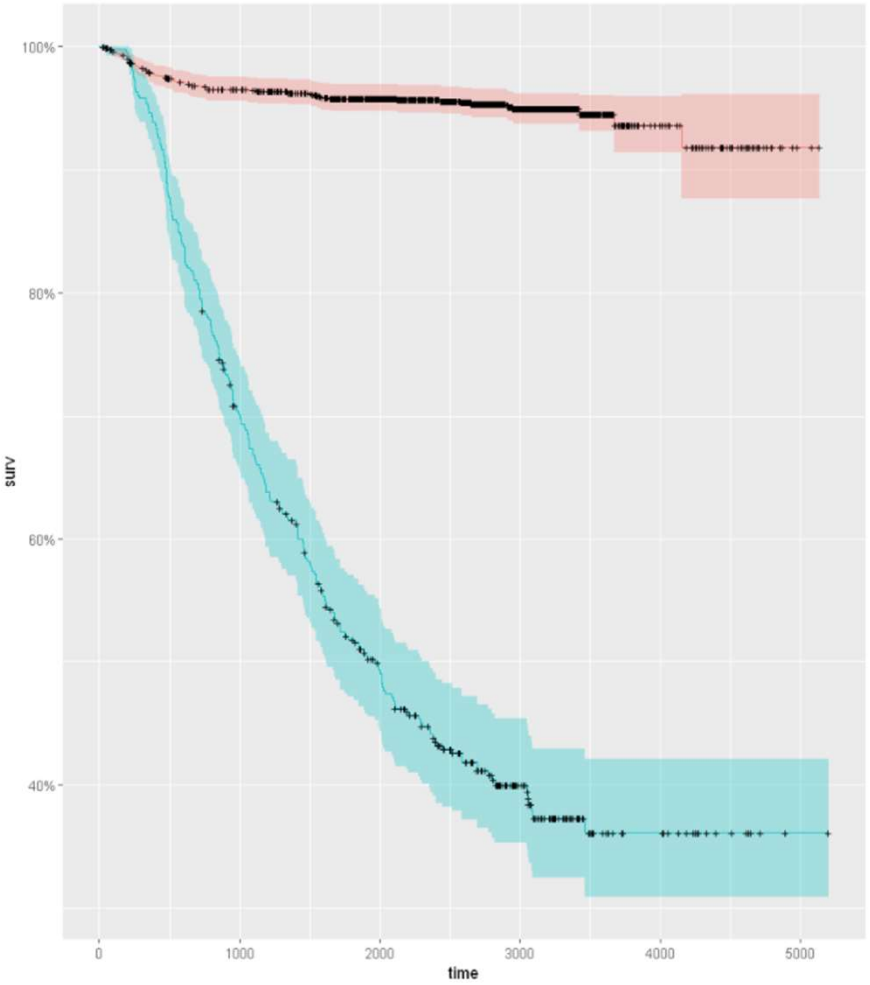
Cox PH Model

- Explanatory variables:
  - Gender, Race, Age at Dx, WBC at Dx, MRD Day 29, DNA Index, BMA blasts Day 8, BMA Blasts Day 29, Relapse
- Most Significant Variables:
  - Age at Dx: p-value = 4e-8
    - exp(coeff) = 1.0577
  - DNA Index: p-value = 5.7e-3
    - exp(coeff) = 0.182
  - BMA Blasts Day 29: p-value = 6e-9
    - exp(coeff) = 1.054
  - Relapse: p-value < 2.2e-16
    - exp(coeff) = 17.8



Relapse:

- Relapse acts as a major predictor overall survival time
- Built models to predict relapse in patients
  - Naïve Bayes classifier using kernel density function
    - Classified all patients as no relapse
  - Naïve Bayes using Gaussian distribution



Using all variables in Cox PH model

Prediction	Reference FALSE	TRUE
FALSE	295	75
TRUE	49	26

Using patient demographics

Prediction	Reference FALSE	TRUE
FALSE	344	101
TRUE	0	0

Conclusion – My hypothesis was both correct and incorrect

- Patient population is distinct from general population in all demographics
  - May not change which is most frequent but differences in frequencies are significant
- Demographics do not act as significant predictors of survival
  - Minimal correlation with severity indicators
  - Not useful in predicting relapse – most important predictor of survival
- Future Directions:
  - Compare to patient demographics distribution of other cancers
  - Correlate with gene expression data