

Analysis of Biomarkers as Indicators of Pancreatic Cancer

Joseph Mintz

December 7, 2021

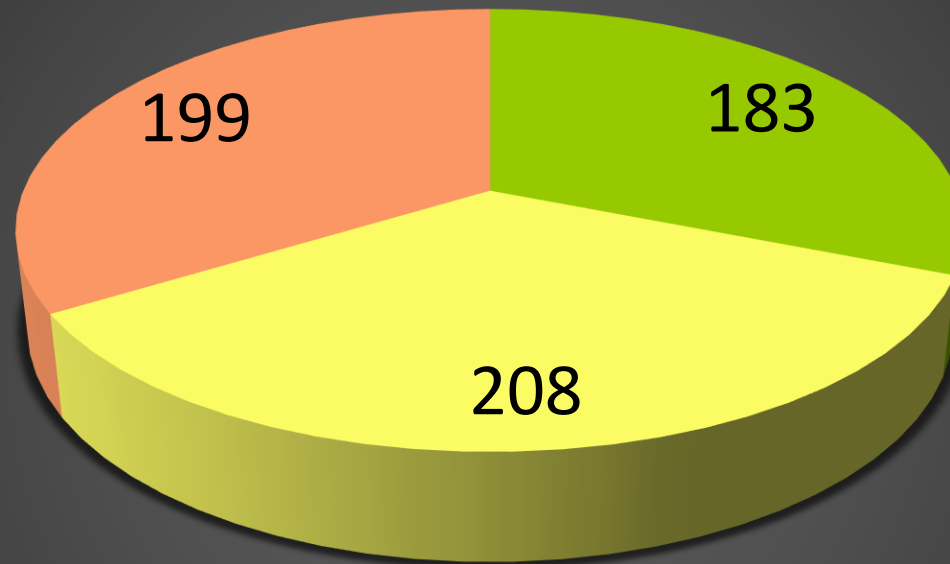
Overview and Problem Statement

- Pancreatic cancer (PDAC) is extremely deadly. Once symptoms arise, it is often already in advanced stages.
- Researchers have identified biomarkers that have previously shown promise in detecting PDAC earlier.
- How reliable are these biomarkers in predicting PDAC?
- On average, by how much do these biomarkers differ between patients with and without PDAC?

Data: Patient Groups

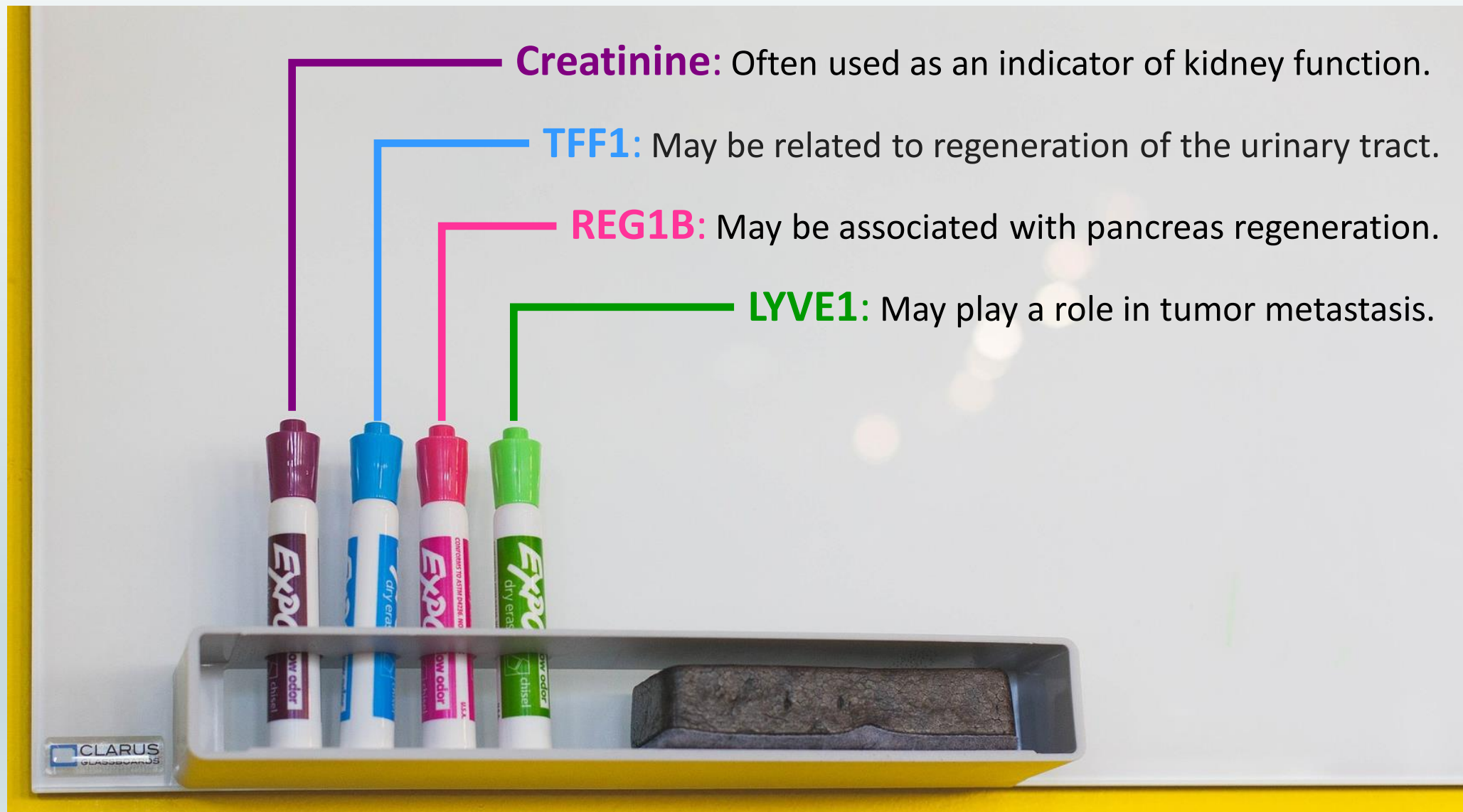
590 patients consisting of three groups based on diagnosis:

Number of Patients by Group

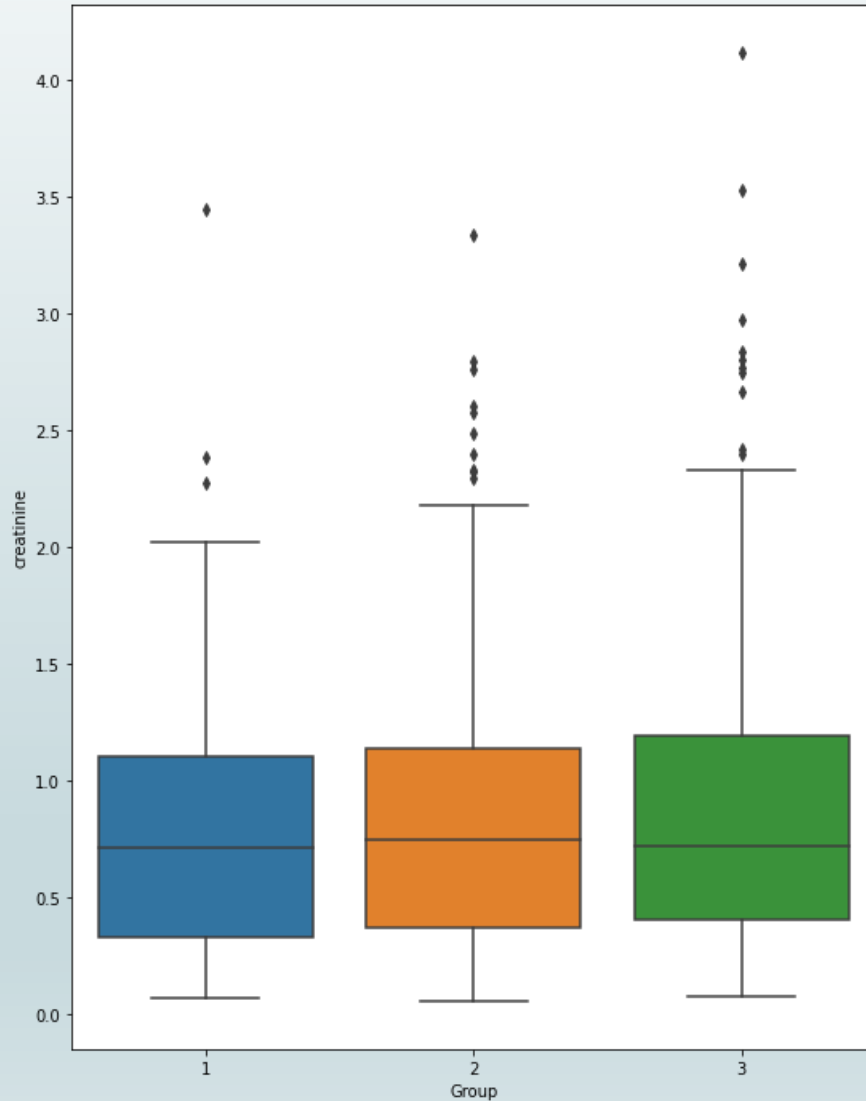


- Group 1: Healthy
- Group 2: With non-cancerous pancreatic condition
- Group 3: With PDAC

Data: Four Biomarkers to Test for Significance

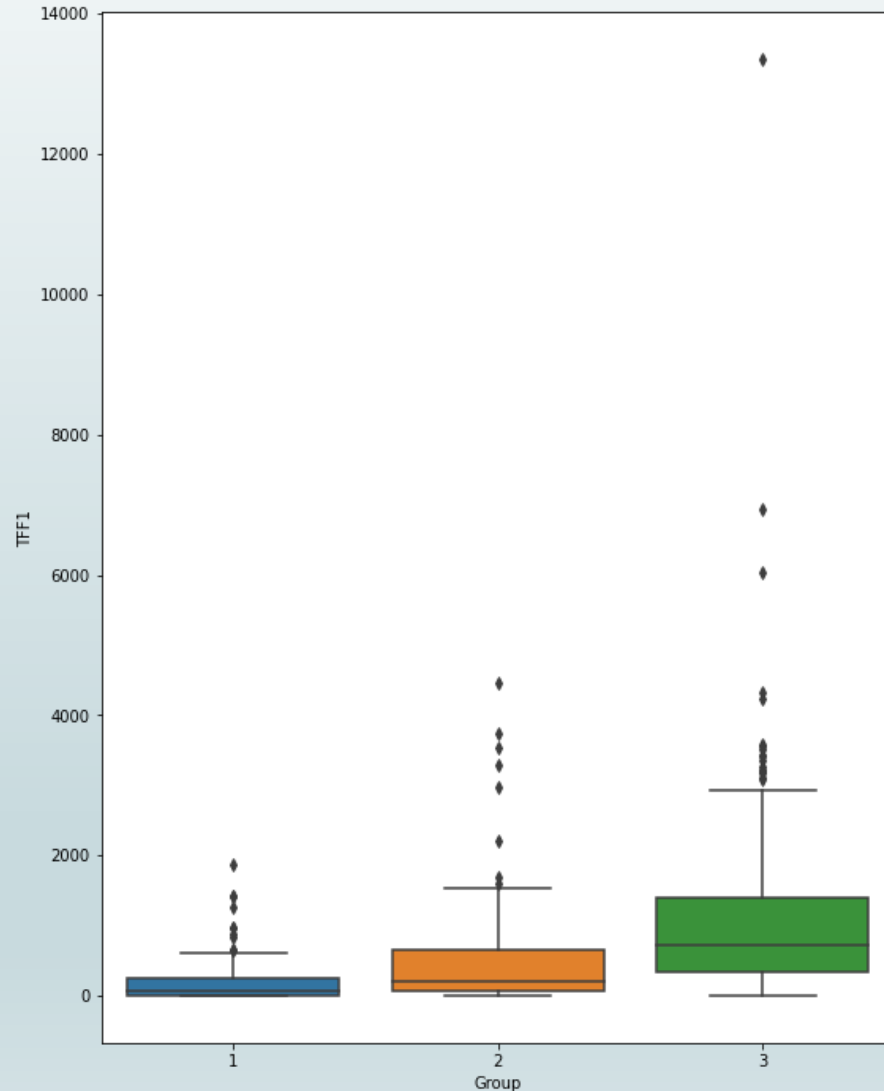


Creatinine Differences between Groups



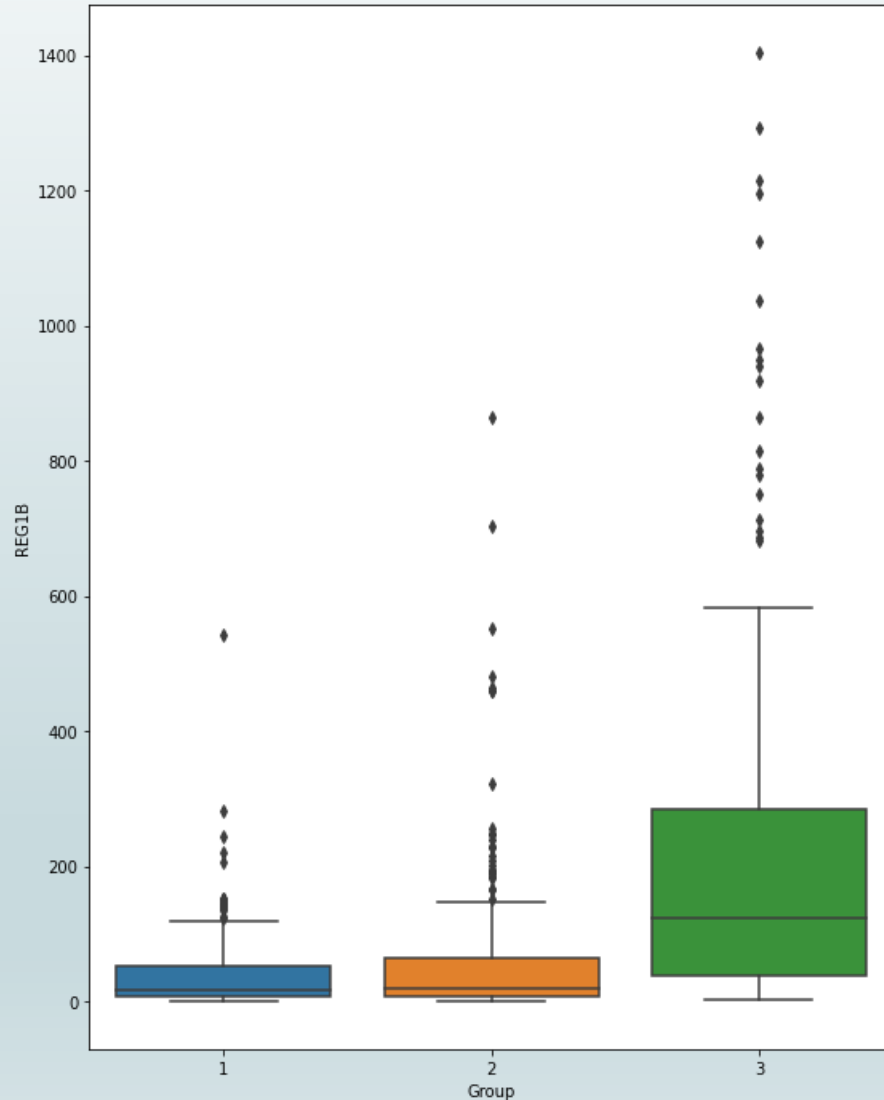
- No significant difference is found in the average creatinine levels between groups.

TFF1 Differences between Groups



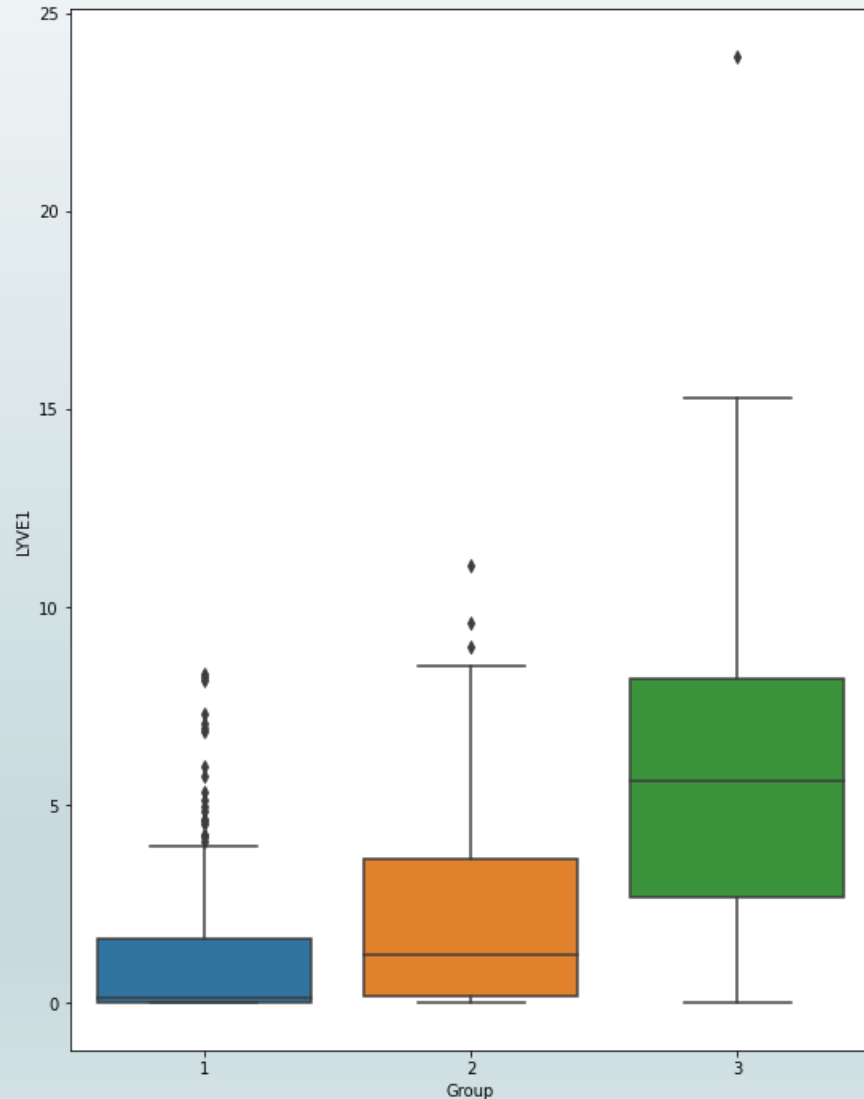
- With 95% confidence, patients with PDAC have TFF1 levels that are greater than those of healthy patients by an amount between 777 and 1182 on average.
- The PDAC patients also have some extreme TFF1 outliers on the high end.

REG1B Differences between Groups



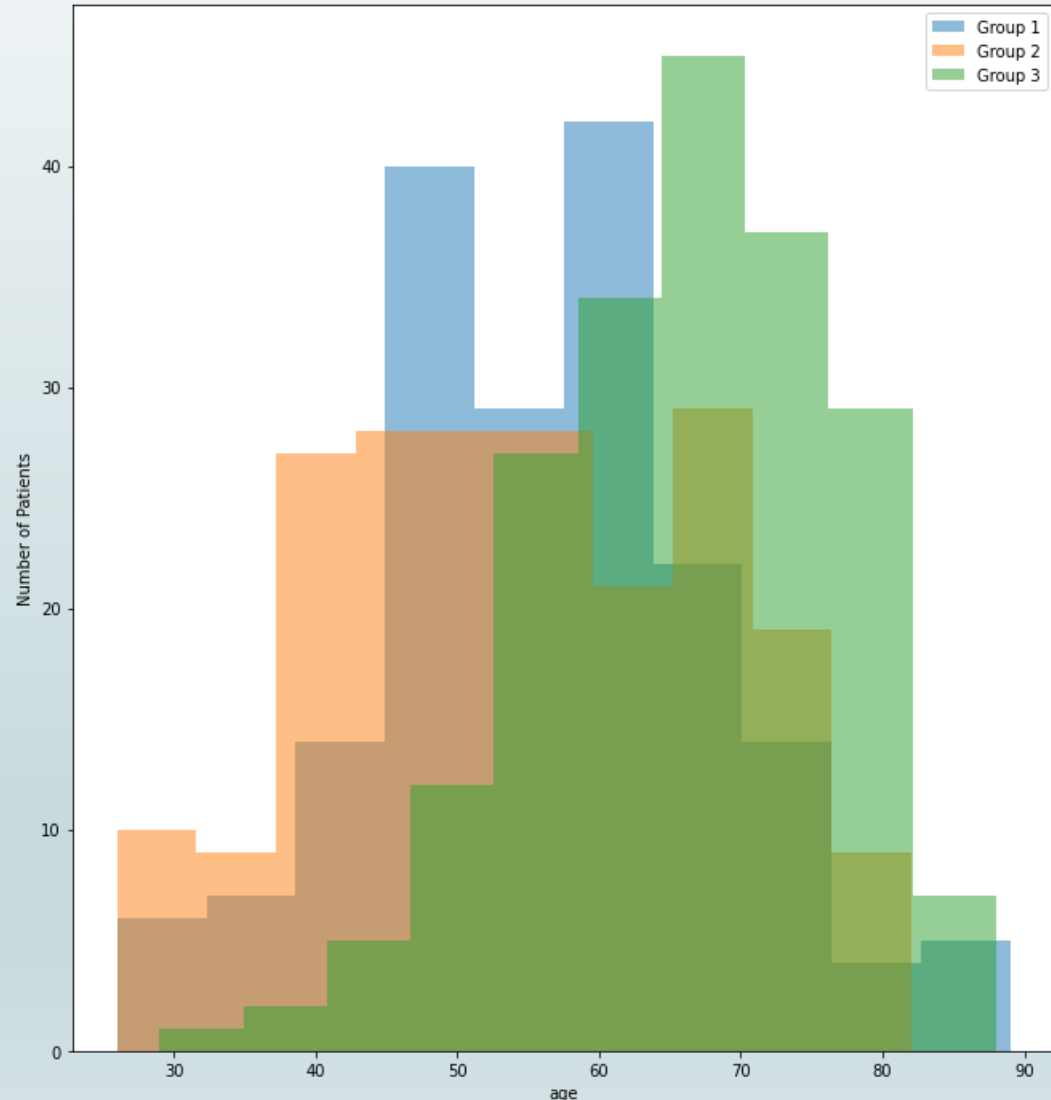
- With 95% confidence, patients with PDAC have REG1B levels that are greater than those of healthy patients by an amount between 145 and 225 on average.
- The PDAC patients also have many REG1B outliers on the high end.

LYVE1 Differences between Groups



- With 95% confidence, patients with PDAC have LYVE1 levels that are greater than those of healthy patients by an amount between 3.98 and 5.17 on average.

Checking for Bias in the Sampling by Age



- Patient age is another variable in the data.
- Group 3 (patients with PDAC) clearly has the oldest patients on average.
- Sampling bias? Or does old age truly correlate with higher rates of PDAC?

Checking for Bias in the Sampling by Sex

	Male	Female
Group 1	68	115
Group 2	107	101
Group 3	116	83

- Patient sex is another variable in the data.
- Group 1 has a significantly lower number of males, and Group 3 has a significantly lower number of females, indicating another source of potential bias.

Conclusions and Next Steps

- The analysis suggests that the biomarkers, TFF1, REG1B, and LYVE1, are reliable indicators of PDAC.
- Investigate the Group 3 biomarker outliers. Do these correlate with an advanced PDAC stage or something else? (stage is another interesting variable in the data)
- Recommend to protect against age and sex bias by resampling or subsetting existing samples to achieve similar distributions among groups.

