Can we Predict Rectal Cancer Outcomes using Clinical Data? A Comparative Analysis of Different Techniques.

by Karina Krishnan

Project Summary

Purpose

The aim of this study is to identify pre-surgery and Magnetic Resonance Imaging (MRI) variables that are significantly associated with rectal cancer outcomes, specifically pathologic TNM (Tumor, Node, Metastasis) stages that describe the extent of the primary tumor and invasion into surrounding tissues, and cancer recurrence. Using high dimensional data, predictors critical for prognosis, surgical planning, and treatment evaluation in rectal cancer care were identified.

Hypothesis

The extent of lymph node involvement by cancer cells, determined by imaging variables, is significantly associated with rectal cancer outcomes.

Data

Data includes the pathologic TNM-stage classification and recurrence, as well as several pre-surgery and imagining variables of 55 patients.

Procedure

A comparative analysis was conducted using different regression techniques to analyze both MRI variables, such as lymphovascular invasion, and pre-surgery variables, such as race and initial clinical staging, to determine their associations with rectal cancer outcomes. I used Tobit and Logit

regressions for ordered outcomes, Lasso regression for optimal variable selection, Ridge regression for dealing with multicollinearity, as well as a combination of the Ridge and Lasso, called ElasticNet regression.

Summary of Findings and Conclusion

Results indicate that certain variables, specifically, the initial metastatic stage (pre-surgery variable) and the number of positive lymph nodes (imaging variable) supporting the hypothesis, are consistently linked to poor outcomes across different techniques. These findings suggest how data-driven insights can have implications for personalized treatment planning in rectal cancer. As imaging technology advances, the accuracy and predictive power of such models will grow, allowing for optimal patient-centered care.