Matt Castellano

Survivability Analysis of Patients with Lymphatic Invasion, Venous Invasion, and Both: Comparing Males and Females to Confirm Sexual Dimorphism

Introduction:

Colorectal cancer is the second leading cause of cancer deaths in the United States just behind lung cancer. However, due to the use of precision medicine and other therapeutic approaches, the cancer survival rates have improved in high-income countries, but remain stagnant in low-income countries with inadequate cancer prevention, education, and access to screening tests and treatment. Since the Human Genome Project in 2001, cancer deaths have gone down by 27%. Likewise, the cost to decode RNA and obtain the necessary data to make precise and effective medicine has gone down drastically, from roughly $300 million dollars to less than $1000 dollars. That being said, given the technological advancements that have sparked innovation and discovery, we are able to analyze the many different aspects of cancer diagnosis easier and more efficiently than ever before. The two data points that we will focus on are venous invasion and lymphatic invasion: two well-established independent prognostic indicators for colorectal cancer (CRC), specifically during stage II for the decision to administer adjuvant therapy. We will look at the two factors and how they contribute to the survivability rates of patients, comparing how they influence those rates amongst males and females.
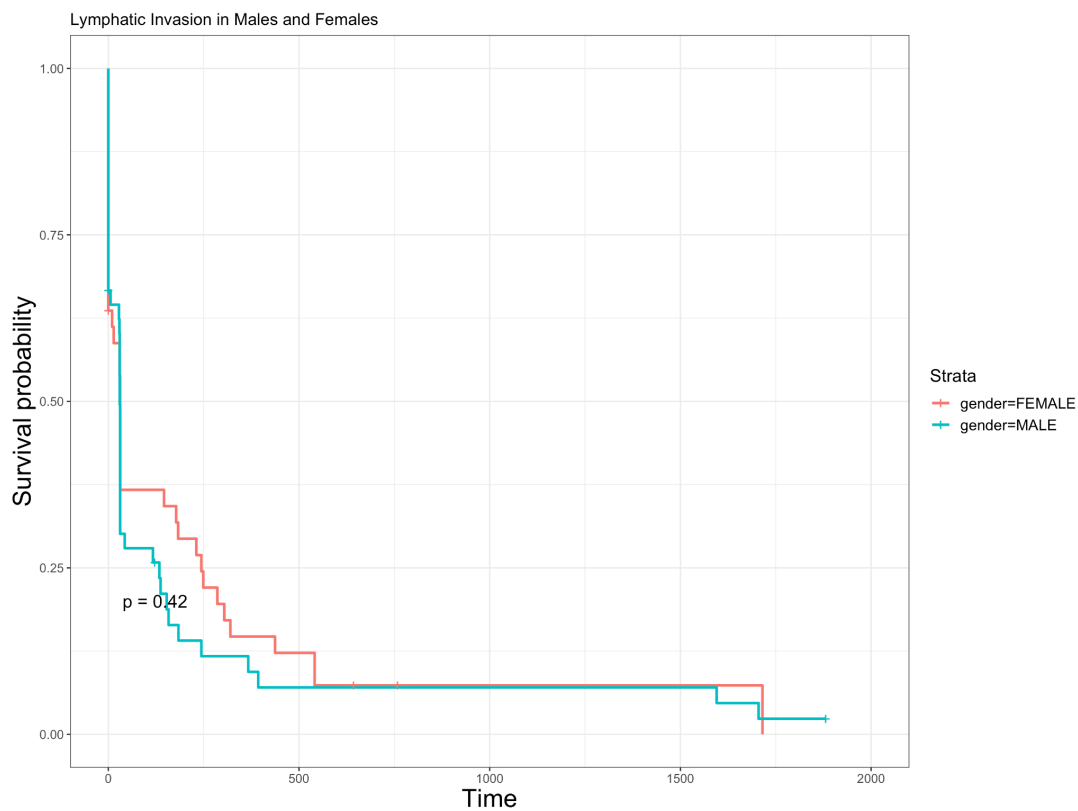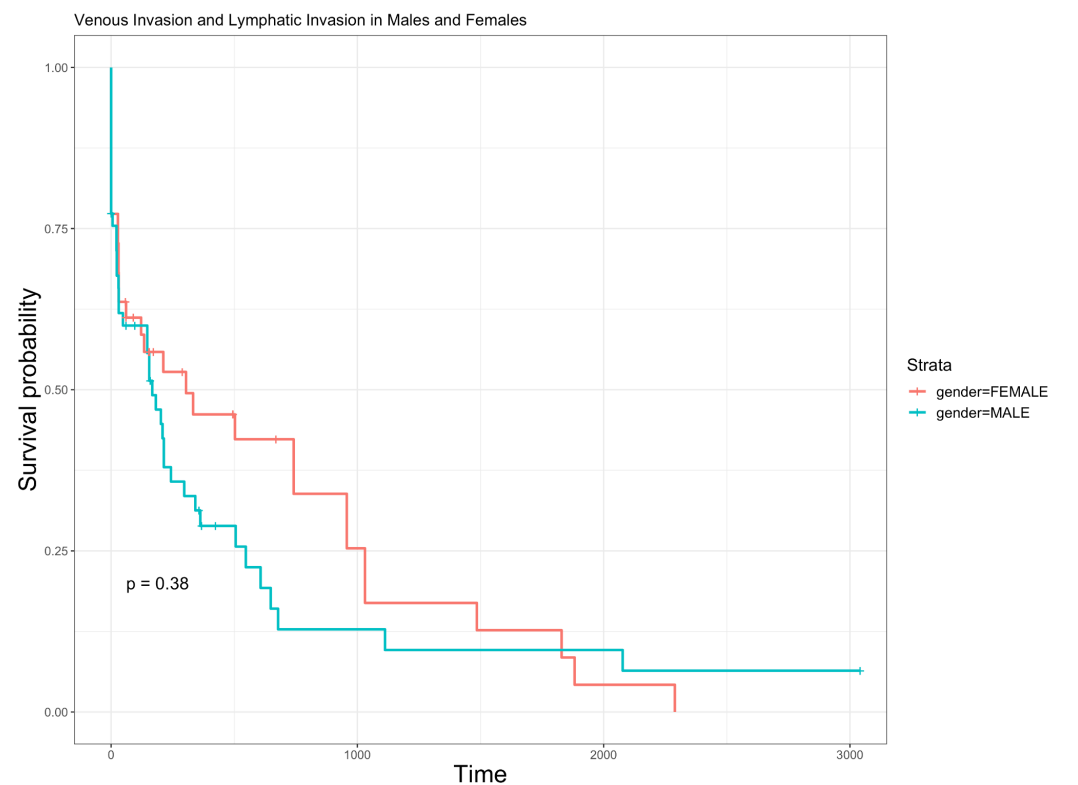
Methods:

The following analysis was performed by accessing colon cancer clinical data from The Cancer Genome Atlas (TCGA) using the R package TCGAbiolinks with accession code "COAD". After installing the packages and downloading the data, the data was processed and loaded into RStudio to parse and analyze the clinical data. The original data was reduced into three data sets: one with only patients that had only venous invasion (VI), another of patients
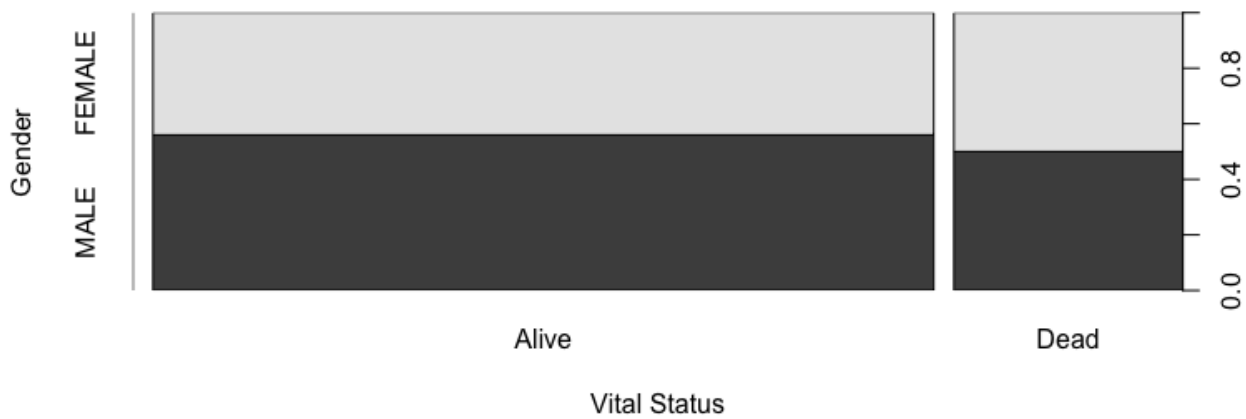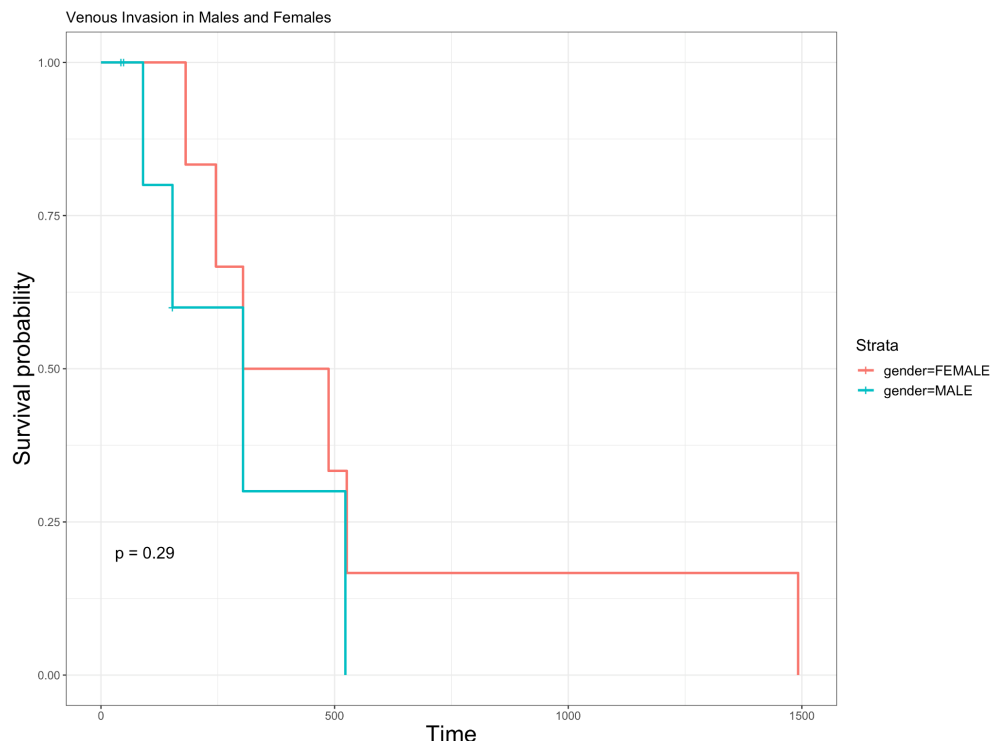
with only lymphatic invasion (LI), and patients with both VI and LI. By applying indexing, the extraneous data points, like a blank space or NA, were disregarded. The data was then slightly altered in order to make it more readable and easier to include with statistical analysis. Once the data was prepared for survival analysis, the appropriate packages to accomplish so were installed and loaded into RStudio. A survival object was created in order to narrow the data even more to look directly at the mortality information. Using the ggplot function, three different graphs were created in order to visualize and compare the survivability rates of colorectal cancer men and women patients that had either venous invasion, lymphatic invasion, or both. The three graphs were further analyzed to help rationalize whether there is a sexual dimorphism present.

Results:

We found that amongst the 524 CRC patients, (53% males), in the TCGA-COAD data set there were 189 patients with venous invasion, 13 of which did not also have lymphatic invasion. Conversely, there were 110 patients with lymphatic invasion and 92 of them did not have venous invasion. 97 patients had both venous invasion and lymphatic invasion, with 53 of them being males. Given these totals, 36% of the total patients had venous invasion and only 6% of those patients did not also have lymphatic invasion. Of the 13 patients with only venous invasion, 7 were males (54%) and 6 were females (46%). Of the 13 patients, 3 died and they were all male. Of the 92 patients, 48 of them were males. In addition, the p value for the survivability rates of patients with only LI was .42. The p value for the survivability rates of patients with only VI was .29. The p value for the survivability rates of patients with both LI and VI was .38. An ideal confidence interval would be about 95%, making the p value .05.

Figures:

Venous Invasion and Lymphatic Invasion in Males and Females

p = 0.38

Strata
— gender=FEMALE
— gender=MALE

Lymphatic Invasion in Males and Females

p = 0.42

Strata
— gender=FEMALE
— gender=MALE

Venous Invasion in Males and Females

Survival probability

p = 0.29

Strata
— gender=FEMALE
— gender=MALE

Time

Gender

FEMALE

MALE

Alive

Dead

Vital Status

Discussion:

In conclusion, the analysis of the influence of venous invasion and lymphatic invasion had some shortcomings. To begin, the degree of lymphatic invasion is not included in the data, affecting the survivability statistics. Although the data suggested that women had better survivability than men across all three tested metrics, there is a mathematical problem with the results. The p value was way too high in the analysis, making it difficult to conclude whether there is sexual dimorphism or if the typically differing environmental factors that men and women are exposed to is to blame for the discrenpincy. Typically, the p value is around .05 for a high level of confidence and little to no randomness. All of the p values were above .28, meaning that the data is not as clear as it should be to confidently draw supportable conclusions. Crucial factors in cancer development, such as dangerous jobs, smoking, and obesity are drastically more applicable to men than women. That being said, there is also a very important genetic component that cannot be ignored. Overall, the analysis was inconclusive and would need more specific data and a much higher confidence interval in order to support the hypothesis that there is sexual dimorphism for venous invasion, lymphatic invasion, and its connection to CRC survivability.

References:

https://www.icr.ac.uk/news-features/latest-features/how-the-human-genome-project-shook-the-world-of-cancer-research

https://www.cdc.gov/cancer/dcpc/research/update-on-cancer-deaths/index.htm#:~:text=In%20the%20past%2020%20years,144.1%20deaths%20per%20100%2C000%20population.

https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4283716/#:~:text=Venous%20invasion%20(VI)%20is%20a,decision%20to%20administer%20adjuvant%20therapy.