****

**The Effect of Surgical Treatment in Groups of Patients with Different Stages of Lung and Bronchus Cancer**

I. Arvelo, J. DeLano, K. Gregory, and N. Seetharaman

**Introduction**

*Scientific Question:* This study will examine if the effect of available surgical treatments differs for individuals depending on how far their cancer has spread from its point of origin. Cancer is the second leading cause of death in the United States (CDC, 2021), and within this broader category Lung and Bronchus Cancer make up about a fifth of deaths each year. In fact, Lung and Bronchus cancer is projected to be responsible for 131,880 out of the 608,570 people expected to die from cancer in 2021 (U.S. Department). With over 235,000 new Lung and Bronchus cancer diagnoses made each year, it is worthwhile to identify which medical interventions are effective in each stage of cancer to help healthcare providers efficiently allocate resources available for cancer treatment.

*Hypotheses:* A priori, the hypothesis is that resection and tumor destruction surgeries are more effective treatments (at reducing the risk of death) in patients with in-situ or localized cancer than in patients with distant or regional cancer. Statistically speaking, our hypothesis is that stage of cancer is an effect modifier for the association between treatment and mortality. We predict that the ratio of the instantaneous risk of death between resection surgery, tumor destruction surgeries and no surgery is lowest for individuals with in-situ cancer and the highest for individuals with regional cancer.

*Potential Implications:* The findings of this study could help healthcare providers and patients make more cost effective-treatment decisions. If the available types of surgery are not associated with large changes in the risk of death for the stage of cancer a patient has, then it may not be worth the potential risks and physical stress on the body. Time and resources may be better allocated on alternative treatment options such as radiation.

**Methods**

*Population Description*: The study population includes male and females with Lung and Bronchus cancer in the general U.S. population between the ages of 20 and 64.

*Eligibility Criteria:* The study only includes young to middle aged adults between the ages of 20 to 64 that were diagnosed with Lung and Bronchus cancer between 2008 and 2018. The cases observed in this study are required to have information on if and which surgery they received within the time of observation, time to event data, and their median household income.

*Study Design:* This study looks back in time at archived SEER data to examine whether effect of resection surgery, tumor destruction surgery and no surgery was different between groups of patients with different stages of cancer. To avoid competing risks, the binary outcome was an individual’s mortality at the end of the decade. The exposure was the type of surgery each patient received: resection, tumor destruction, or no surgery. However, the focus in this study is the effect modification by Stage. Stage describes how far a patient’s cancer has spread from its point of origin and is divided into In-Situ, Localized, Regional, Distant, and Unknown.

*Data Preparation:* The data for this study was collected from the “Incidence - SEER Research Data - 18 Registries, Nov 2020 Sub” database. Death was coded an indicator variable for whether the subject was alive or dead the last time we had available data on them. Different surgery types were put into three broader categories: no surgery, tumor destruction, and resection. Cases with unknown median household income, county rural urban continuum code, or without time to event data were excluded. Cases listed as unknown if surgery performed or surgery site not otherwise specified were also excluded. Income and county codes were converted from categorical to numeric variables with levels.

*Data Exploration:* Data gathered from SEER is collected by cancer registries so all the variables in our study reflect the data on patients’ medical records. We discovered that only one case out of 168,942 in our dataset had in-situ cancer and this individual did not have surgery (Appendix A, Table 1). Considering that our dataset had zero cases in the groups with in-situ cancer stage and resection or tumor destruction surgery, we decided to remove the one case (and thus, the entire in-situ group) from our dataset before running our analyses. A priori, we did not expect to lack data for these groups, so it did not impact our eligibility criteria; however, our analyses about stage of cancer as an effect modifier for the association between surgery type and mortality would not have any power with zero individuals in those groups.

As for the distribution of other covariates among the subjects in our dataset, 47.4% were female, 77.2% were white, 54.7% lived in counties with metropolitan areas greater than or equal to one million in population, 23.9% lived where the median household income is greater than $75,000 per year, and the median age was 58 years old (Appendix A, Table 2 A, B, C). Of particular note, the distribution of the age variable was very left skewed, and the participants ranged from 20 to 64 (Appendix A, Figure 1).

Considering the relationship between the covariates and mortality, a higher proportion of those who were alive at the end of the study were female (Appendix A, Figure 2), lived in homes with higher median household incomes, and lived in counties with metropolitan areas greater than or equal to one million in population. Proportional to the starting distribution of individuals by race, more Blacks and American Indian/Alaska Natives were deceased by the end of the study while more Asian or Pacific Islanders and those with unknown race were alive. The distribution of age, on the other hand, seemed to be approximately the same among those alive and dead at the end of the study.

*Inferential Procedure and Assumptions:* This study uses the Cox-Proportional Hazards Model to understand the effect of surgery between groups of patients that have cancer in different stages. The first assumption we made is that censored observations were similar to those still in study, which was corroborated by the similar distribution of demographic variables across censored and complete cases (Appendix A, Figure 3). Each case appears only once, and SEER uses rigorous sampling techniques, so it is safe to assume that observations were independent from each other. This study estimates the best *linear* (in log hazard) association across modeled predictor value and may not be entirely accurate. The last major assumption is that the effect of a predictor on the hazard is about the same at every point in time.

*Model Descriptions:* The unadjusted and adjusted analyses use a Cox-Proportional Hazards Model where the response variable is a survival object using the number of months survived as the unit of time and death as the event indicator. The unadjusted model contained 11 coefficients for the interaction between stage of cancer (4 categories: distant, localized, regional, and unknown) and type of surgery (3 categories: no surgery, resection, and tumor destruction). The adjusted model contained 9 additional coefficients: race (5 categories: White, Black, Asian or Pacific Islander, American Indian/Alaska Native, and Unknown), age, year of diagnosis, median household income, and county rural urban continuum code. No variables were transformed.

*Covariates:* The model controls for certain patient case attributes to increase accuracy and precision of the estimated treatment effect in each group. We controlled for sex and age to gain precision because potential biological differences between these groups may be associated with differences in mortality. We also included race as a potential confounding variable to account for racial differences in medical care that may be associated with which surgeries patients are offered, how well/carefully they are treated, and the risk of death. Other variables accounted for were median household income and country rural urban code. Median household income is a potential confounder because it may limit which surgeries patients can afford and their overall health, which would in turn be associated with mortality. Country urban code may also be associated with both treatment and mortality due to pollution levels, accessibility of healthcare services, comorbidities such as stress, SES and many other factors that could differ between rural and metropolitan areas.

**Results***:*

*Unadjusted Model:* Based on Cox proportional hazards regression analysis of mortality versus surgery type with stage as an interaction term, we estimate that among patients with distant cancer the instantaneous risk of death is 0.351 times as high for patients that receive resection surgery and 0.881 times as high for patients that receive tumor destruction surgery than patients that do not receive surgery at all holding the stage of cancer constant. With 95% confidence, we can say that the true hazard ratio is between 0.338 and 0.365 for Resection surgery and 0.774 and 1.002 for Tumor destruction surgery. The hazard ratio for resection surgery is significantly different from 1 (p < 0.001) and the hazard ratio for tumor destruction does not appear to be significantly different from 1 (p = 0.054).

Holding the surgery site constant, we estimate that patients with localized cancer have an instantaneous risk of death that is 0.352 times as high as patients with distant cancer. With 95% confidence, this estimate is consistent with a true hazard ratio between 0.342 and 0.363, which is significantly different from 1 (p < 0.001). Holding the surgery site constant, we estimate that patients with regional cancer have an instantaneous risk of death that is 0.505 times as high as patients with distant cancer. With 95% confidence, this estimate is consistent with a true hazard ratio between 0.496 and 0.513, which is significantly different from 1 (p < 0.001). Holding the surgery site constant, we estimate that patients with unstaged cancer have an instantaneous risk of death that is 0.617 times as high as patients with distant cancer. With 95% confidence, this estimate is consistent with a true hazard ratio between 0.594 and 0.64, which is significantly different from 1 (p < 0.001).

Compared to those with distant cancer and no surgery (baseline), we estimate that patients with regional cancer undergoing resection surgery have a 0.976 times as high instantaneous risk of death and patients with regional cancer undergoing tumor destruction surgery have a 1.172 times as high instantaneous risk of death. With 95% confidence, these true hazard ratios are within 0.93 and 1.025 for resection surgery and 0.934 and 1.47 for tumor destruction surgery. Neither of these ratios are significantly different from 1 (p=0.172 for resection surgery and p=0.371 for tumor destruction surgery). The effect of each surgery does not seem to be significantly different between patients with distant and regional cancer.

Compared to those with distant cancer and no surgery (baseline), we estimate that patients with localized cancer undergoing resection surgery have 0.7 times as high instantaneous risk of death and patients with localized cancer undergoing tumor destruction surgery have a 0.73 times as high instantaneous risk of death. With 95% confidence, these true hazard ratios are within 0.661 and 0.74 for resection surgery and 0.539 and 0.992 for tumor destruction surgery. These ratios are both significantly different from 1 (p < 0.001 for resection surgery and p=0.044 for tumor destruction surgery), which shows that for patients with localized cancer, resection and tumor destruction surgery are associated with significantly larger reductions in the instantaneous risk of death as compared to no surgery than they are for patients with distant cancer.

Compared to those with distant cancer and no surgery (baseline), we estimate that patients with unstaged cancer undergoing resection surgery have 0.92 times as high instantaneous risk of death and patients with unstaged cancer undergoing tumor destruction surgery have a 0.642 times as high instantaneous risk of death. With 95% confidence, these true hazard ratios are within 0.766 and 1.105 for resection surgery and 0.239 and 1.728 for tumor destruction surgery. Neither of these ratios are significantly different from 1 (p =0.371 for resection surgery and p=0.381 for tumor destruction surgery). There does not appear to be a significant difference in the effect of surgery between distant and unstaged cancer.

We have strong statistical evidence for effect modification by stage, as shown by a Wald test (p ≈ 0.0).

*Adjusted Model:* Based on Cox proportional hazards regression analysis of mortality versus surgery type with stage as an interaction term, we estimate that among patients with distant cancer the instantaneous risk of death is 0.357 times as high for patients that that receive Resection surgery and 0.867 times as high for patients that receive Tumor destruction surgery than for patients that do not have surgery, holding stage of cancer, sex, age, race, year of diagnosis, median household income, and county rural urban constant. Both hazard ratios are significantly different from one (p < 0.001 and p = 0.031, respectively). With 95% confidence, we can say that the true hazard ratio is between 0.344 and 0.372 for resection surgery and 0.761 and 0.987 for tumor destruction surgery.

Holding the surgery type constant, we estimate that patients with localized cancer have an instantaneous risk of death that is 0.335 times as high as patients with distant cancer. This estimate is consistent, with 95% confidence, with a true hazard ratio between 0.325 and 0.345, and is significantly different from 1 (p < 0.001). Holding surgery type constant, we estimate that patients with regional cancer have an instantaneous risk of death that is 0.0484 times as high as patients with distant cancer. This is significantly different from 1 (p < 0.001) and with 95% confidence this is consistent with a true hazard ratio between 0.475 and 0.492. Holding surgery type constant, we estimate that patients with unstaged cancer have an instantaneous risk of death that is 0.616 times as high as those with distant cancer. This is consistent with a true hazard ratio between 0.593 and 0.64 with 95% confidence, and it is significantly different from a hazard ratio of 1 (p < 0.001).

Compared to a baseline of patients with no surgery and distant cancer, while still holding other covariates constant, we estimate that patients with regional cancer have a 0.985 times as high instantaneous risk of death when they underwent resection surgery, and 1.201 times as high instantaneous risk of death when they underwent tumor destruction surgery. With 95% confidence, this is consistent with a true hazard ratio between 0.939 and 1.034 for resection surgery and 0.956 and 1.507 for tumor destruction surgery. Neither of these ratios is significantly different from a hazard ratio of 1 (p = 0.549 for resection surgery and p = 0.115 for tumor destruction surgery).

Compared to the same baseline of patients with distant cancer that did not undergo surgery and holding other covariates constant, we estimate that patients with localized cancer undergoing resection surgery have 0.727 times as high instantaneous risk of death and those undergoing tumor destruction surgery have a 0.764 times as high instantaneous risk of death. Both of these ratios are significantly different from 1 (p < 0.001 for both resection and tumor destruction surgery), and with 95% confidence the true hazard ratio falls between 0.687 and 0.769 for resection surgery and 0.563 and 1.036 for tumor destruction surgery.

Compared to patients with distant cancer and no surgery, patients with unstaged cancer undergoing resection surgery have a 0.899 times as high instantaneous risk of death, and patients undergoing tumor destruction surgery have a 0.605 times as high instantaneous risk of death. With 95% confidence, this is consistent with a true hazard ratio between 0.749 and 1.08 for resection surgery and 0.225 and 1.627 for tumor destruction surgery. Neither of these ratios is significantly different from a hazard ratio of 1 (p = 0.256 for resection surgery and 0.319 for tumor destruction surgery).

We have strong statistical evidence for effect modification by stage, as shown by a Wald test (p ≈ 0.0).

**Discussion**

*Major Findings:* Our findings are generalizable to middle aged adults in the general U.S. population with Lung and Bronchus cancer that has spread past the in-situ stage. Among this population, resection surgery is, in general, the most effective way to reduce a patient’s instantaneous risk of death, and resection surgery in an earlier stage leaves the patient with the lowest instantaneous risk of death.

Based on the Kaplan-Meier curves, the rate of death is a lot more gradual with Resection Surgery than in both other treatment groups (Appendix A, Figure 4). Over the course of the study, approximately 10 years of follow-up, the curve for no surgery is the steepest (higher hazard, higher risk of death) for all stages of cancer. We estimate that 50% of adults with resection surgery survived at least 106 months (8.8 years), whereas 50% of adults with tumor destruction surgery survived at least 11 months and adults that did not receive surgery survived at least 9 months after they were enrolled in the study.

This result makes sense intuitively, as generally we know that the earlier cancer can be caught the more likely a patient is to survive, especially when treatment can be applied early on. Cancers that have not spread very far can be much more easily removed (i.e., resection surgery), and leave the best chances for a patient to recover and survive. This reinforces the knowledge that early intervention is key to give patients the best chance at survival. To allow patients to be screened for cancer early and stage early interventions, issues in health equity in the U.S. need to be addressed, as well as the way the healthcare system is set up with regards to insurance and access to affordable care.

Other interesting findings included the distribution of risk across different demographics in the population of interest. Holding all else equal, males had higher instantaneous risk of death than females. American Indian /Alaska natives had the highest risk by any racial group, although they made up only about 0.667% of the dataset. Lastly, groups with higher median household income had lower risk of death, which makes sense since they likely have increased access to high quality medical resources and care.

*Limitations:* When considering the findings of this study, it is important to consider its several potential limitations. To begin, the study only controls for six other covariates and there may be many other factors associated with stage of cancer and/or mortality. To make a robust model that does not overfit the data, we had to choose to control for the variables we thought to be the most important, given the data that was available. We do not have any data or ways to control for comorbidities, overall health, or preexisting conditions that may have major influences on mortality and the risk of surgery. Another limitation is that this study compares two general types of surgery and does not analyze the specific surgical procedures performed for Lung and Bronchus cancer.

*Secondary Comparisons:* Potential follow up studies could repeat our analyses with Lung and Bronchus specific categories for different types of surgery to compare procedures such as pneumonectomies, lobectomies, segmentectomies, and sleeve resections (American Cancer Society, 2019) between patients with different stages of cancer. This data would be more actionable for doctors and patients. Other follow up studies could use the spread of cancer as the response variable, perhaps by comparing the proportion of positive nodes and nodes examined at the beginning and end of the study in order to quantify the “effectiveness” of surgery. Limiting the spread is not only important for reducing the risk of death, but it also plays an important role in decreasing the pain and physical deterioration of cancer patients. It would also be interesting to look at other types of cancer to see if type of surgery and stage changes the efficacy of the surgery.

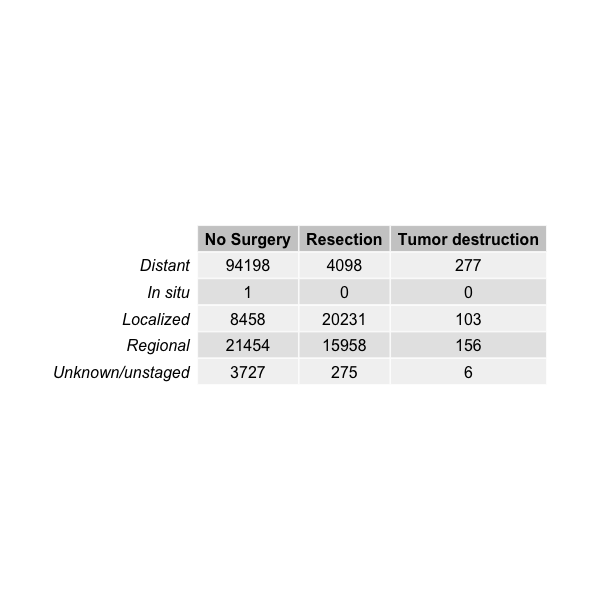
**Summary**

Through this analysis, we aim to determine if there is a difference in surgical treatment effectiveness depending on how far the cancer has spread from its point of origin. We hope that the findings can be used by patients and healthcare professionals to make more cost-effective treatment decisions.

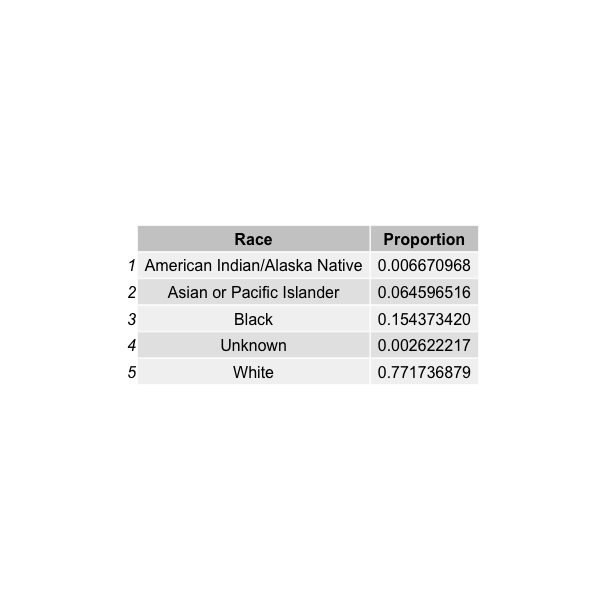
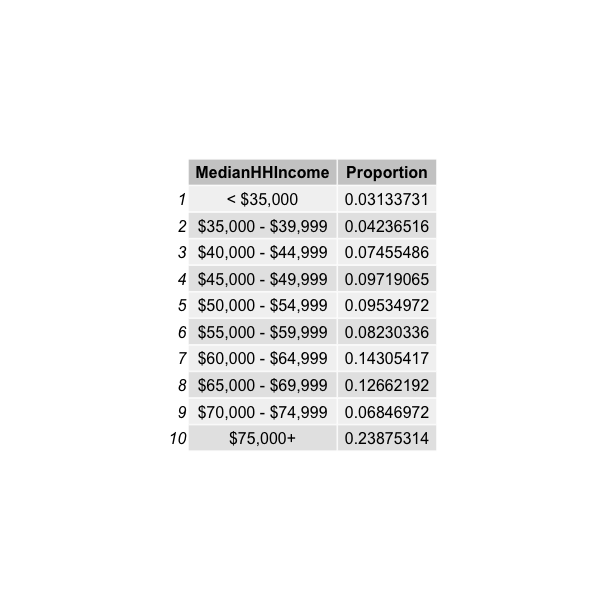
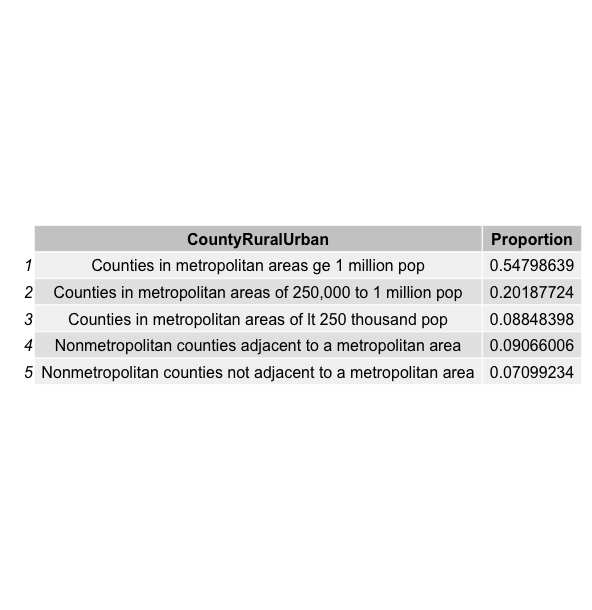
We constructed a Cox-Proportional Hazards Model to understand the effect of surgical treatment between groups of patients with cancer in varying stages. The study population consisted of male and females with Lung and Bronchus cancer in the U.S. population between the ages of 20 and 64. The model controls for sex, age, race, median household income, and country urban code in order to increase accuracy and precision of the estimated surgical treatment effect in each patient group.

Based on our model, we assert that survival chances are higher for patients with cancers that have not spread very far. Furthermore, survival chances increase with earlier treatment. Resection surgeries (procedures that remove the cancerous tissue/organ) are in general the most effective treatment to reduce a patient’s instantaneous risk of death; additionally, the earlier the resection surgery is performed, the lower the instantaneous risk of death is for the patient. These conclusions are generalizable to middle aged adults in the general U.S. population with Lung and Bronchus cancer that has spread past the in-situ stage. However, this analysis was limited by the specific covariates we chose to control for. We could not control for comorbidities or other pre-existing health conditions that would influence risk of surgery. We also only had data to compare two general types of cancer treatment surgeries (resection and tumor destruction) so we could not look into the effects of procedures specific to Lung and Bronchus cancer.

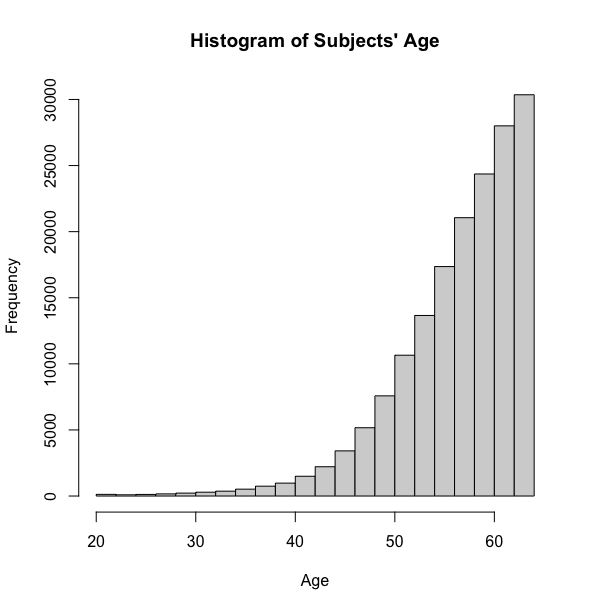
**Appendix A**

****

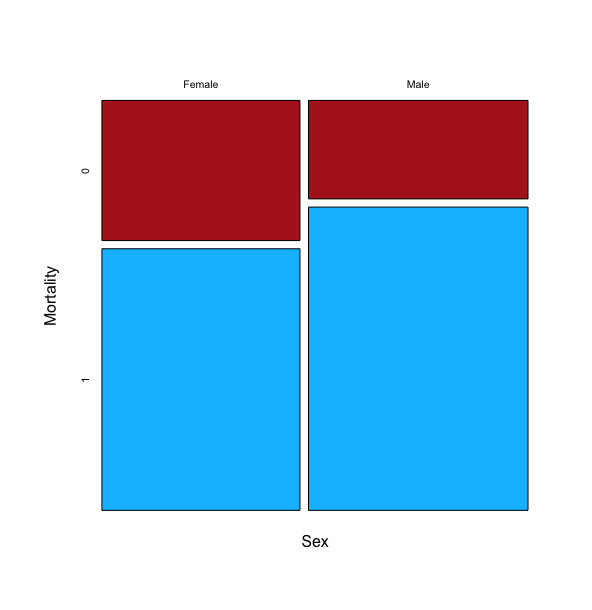
***Table 1:*** Counts of cases by surgery type (coded as no surgery, resection or tumor destruction) and stage of cancer for the full data.

**A.** **B.**   
**C.** 

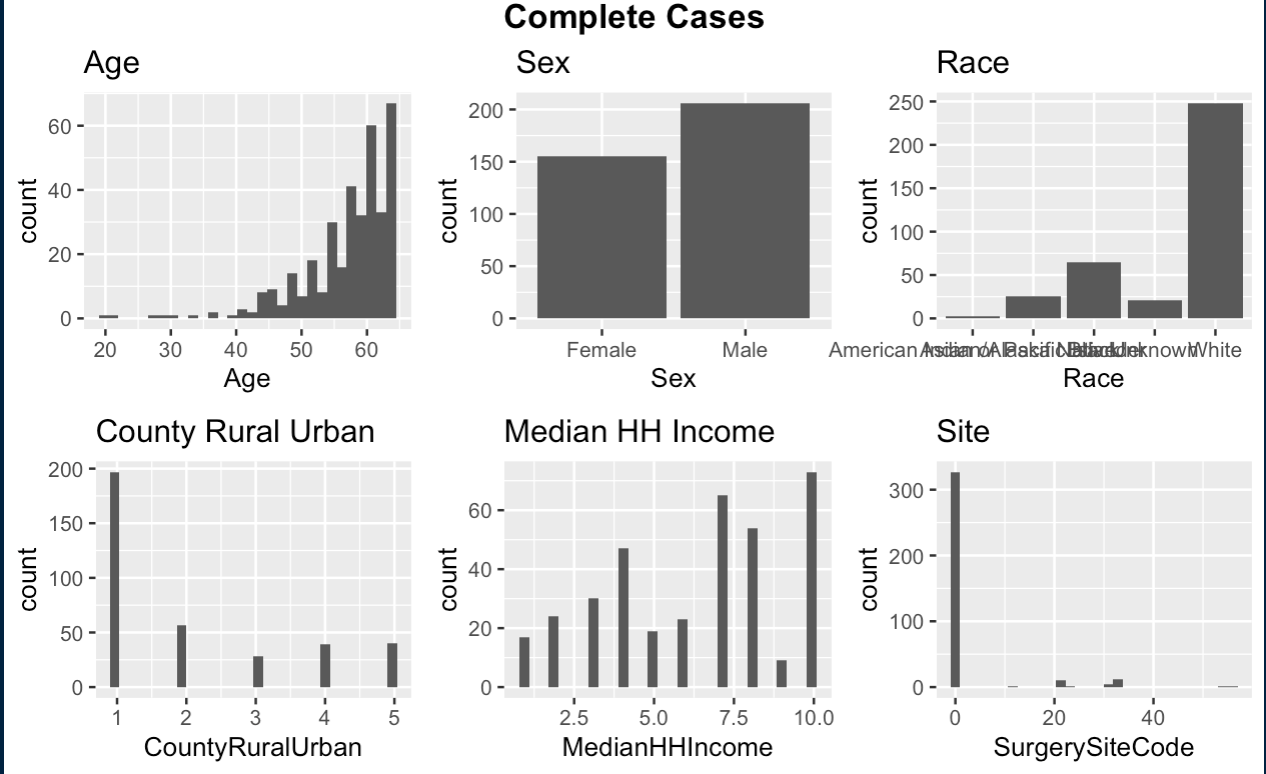
***Table 2:*** Tables summarizing the distribution of demographic characteristics of subjects.

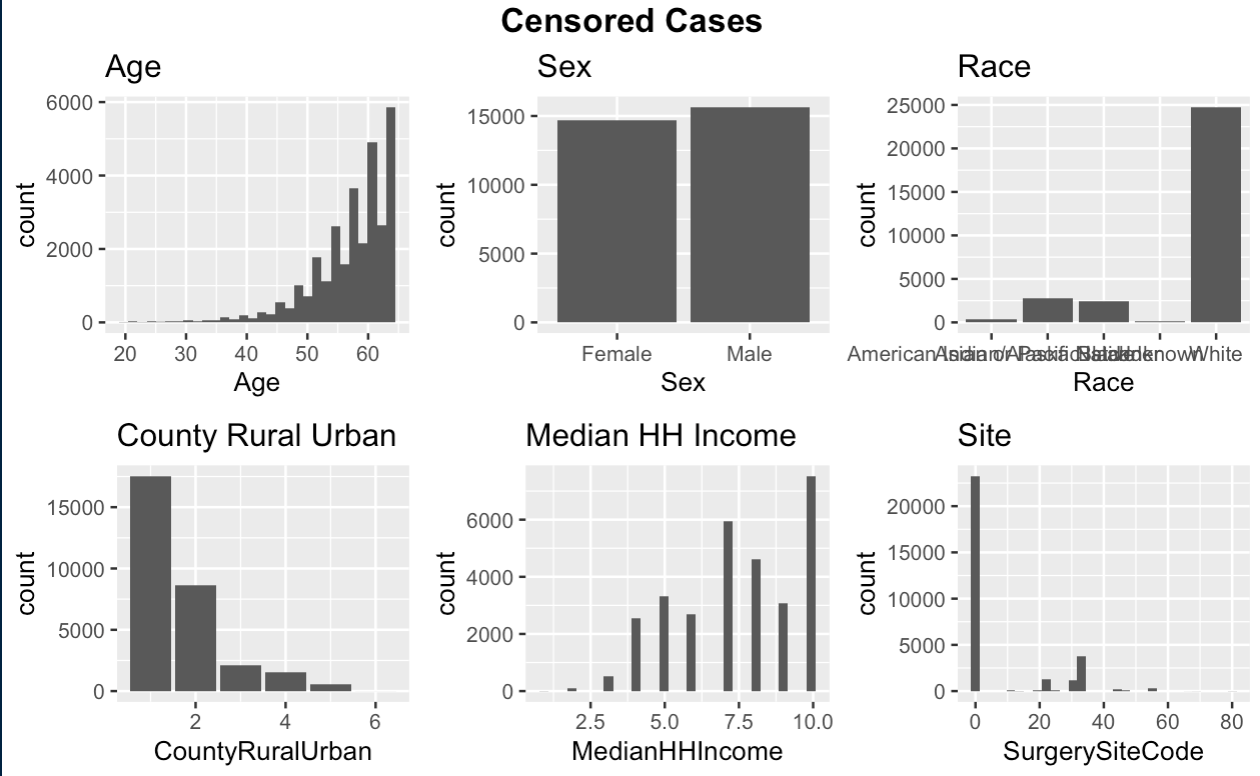


***Figure 1:*** Histogram showing the distribution of the subjects’ age.

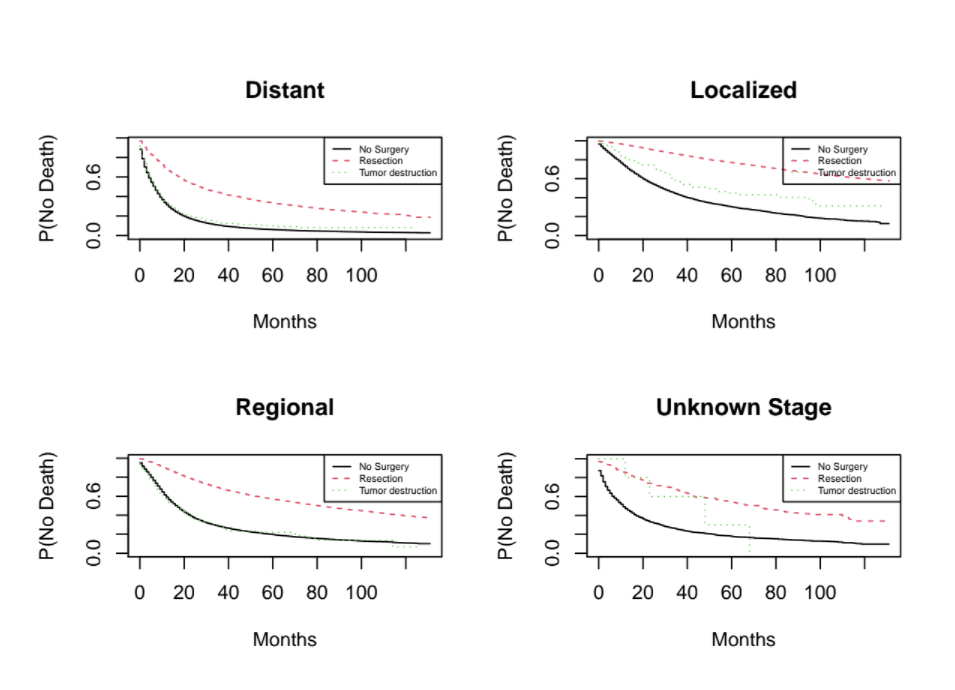


***Figure 2:*** Mosaic plot of distribution of sex and mortality.

****

****

***Figure 3:*** Distribution of demographic variables for cases for which we have complete data(A) and cases that were censored (B).



***Figure 4***. Kaplan-Meier survival curves separated by stage of cancer.

**References**

American Cancer Society. (2019, October 1). Non-small Cell Lung Cancer Surgery: Lung

Cancer Surgery. American Cancer Society. https://www.cancer.org/cancer/lung

-cancer/treating-non-small-cell/surgery.html.

Centers for Disease Control and Prevention. (2021, March 1). *FastStats - Leading Causes of*

*Death*. CDC. https://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm.

U.S. Department of Health and Human Services, National Institutes of Health, National Cancer

Institute. (n.d.). *Cancer Stat Facts: Common Cancer Sites*. SEER .

https://seer.cancer.gov/statfacts/html/common.html.