Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Yaqi Jia Date

SBRT vs. surgery in elderly early stage non-small cell lung cancer patients

By

Yaqi Jia
Master of Public Health

Global Environmental Health

Yuan Liu, PhD
Committee Chair

Barry Ryan, PhD
Committee Member

Paige Tolbert, PhD
Committee Member

SBRT vs. surgery in elderly early stage non-small cell lung cancer patients

By

Yaqi Jia

Bachelor of Medicine
Peking University
2013

Thesis Committee Chair: Yuan Liu, Ph.D.

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Environmental Health
2016

Abstract

SBRT vs surgery in elderly early stage non-small cell lung cancer patients
By Yaqi Jia

Background: Lung cancer remains the leading cause of cancer death worldwide, also in the United State, of which more than 85 percent are classified as non-small cell lung cancer (NSCLC). For operable early stage NSCLC, the surgical treatment is still considered as the gold standard treatment. However, for 70+ years old elderly patients with increasing risk of comorbidities, SBRT has been proved to be an excellent alternative to surgery for inoperable patients and show less treatment morbidity and treatment related mortality than surgery.
Purpose: To determine if SBRT offers equivalent overall survival compared to surgery in elderly early stage non-small cell lung cancer patients.
Methodology: Comparing hazard ratio between SBRT and surgery using cox proportional hazard regression model along with propensity score matching method. If proportional hazard assumption is violated, Cox non-proportional hazard regression model will be used.
Results: The hazard ratio between SBRT and surgery is 2.04 (1.92-2.16) after controlling covariates. It indicates that the hazard for a 70+ early stage NSCLC patient who got SBRT is around 2 times high as the hazard for one who got surgery. However, when comparing SBRT with Pneumonectomy, the hazard ratio is non-proportional. At 6th month after treatment, the hazard ratio between SBRT and pneumonectomy is 0.76 (0.50, 0.81), which indicates that SBRT offers better overall survival than pneumonectomy at 6th month after treatment. After one year (12 months), the hazard ratio between SBRT and Pneumonectomy is greater than 1, which indicates that SBRT offers worse overall survival compared with pneumonectomy (HR=2.00 (1.09, 2.71) at 36th month and HR=4.35 (2.74, 11.52) at 60th month).
Conclusion: SBRT offers worse overall survival comparing with surgery in elderly early stage NSCLC patients, especially compared with lobectomy and resection. But SBRT offers better overall survival comparing with pneumonectomy at less than 12 months after treatment and worse overall survival at more than 12 months after treatment.

SBRT vs. surgery in elderly early stage non-small cell lung cancer patients

By

Yaqi Jia

Bachelor of Medicine
Peking University
2013

Thesis Committee Chair: Yuan Liu, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Environmental Health
2016

Acknowledgment

I am heartily thankful to my thesis advisor, Dr. Yuan Liu, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. I am also thankful to my friend, Chao Zhang, one senior biostatistician at Winship Cancer Institute, for his kindly help and guidance on statistical methods. Thank to my faulty advisor Dr. Barry Ryan and Dr. Dana Barr and the chair Dr. Page Tolbert for their great advice on my thesis. Thanks to my families, my parents, brothers and fiancée for their unconditional support and positive encouragement.

Lastly, I offer my regards and blessings to all of those who supported me in any respect during the completion of the thesis.

Yaqi Jia

Table of Contents

CHAPTER 1

INTRODUCTION

1. Introduction to non-small cell lung cancer

Lung cancer remains the leading cause of cancer death worldwide, also in the United State, of which more than 85 percent are classified as non-small cell lung cancer (NSCLC). Number of New Cases and Deaths per 100,000: It has been estimated that there will be 221,200 new cases of lung and bronchus cancer, with an estimated 158,040 people death due to this disease in 2015[1]. With great improvement of screening techniques for lung cancer, for example, the increased use of low-dose computed tomography (CT) scans, more lung cancer cases are found on their early stage. Currently, approximately 15%-20% are classified as early stage NSCLC[2].

2. Surgical treatment for NSCLC patients

For operable early stage NSCLC, the surgical treatment like lobectomy or pneumonectomy with mediastinal lymph node dissection is still considered as the gold standard treatment. The working group of the International Association for the Study of Lung Cancer (IASLC) defined that the goal of surgical treatment for NSCLC is to obtain a complete resection. For patients with clinical stage at I to III and good health condition to tolerate a major operation, surgical

resection has been a great method to get long-term survival rate [3-7]. Until now, Different types of operative procedures are currently available to the thoracic surgeon, and some of these interventions can be performed by video-assisted thoracic surgery (VATS) with the same oncological results as those by open thoracotomy. However, some patients with comorbidity have a higher surgical risk. For these cases, we usually suggest to get a sub-lobar resection, even though its outcome is not as good as gold standard treatment according to a randomized study[8] .For adenocarcinoma in situ and minimally invasive adenocarcinoma, the role of sublobar resection along with wide wedge resection and segmental resection, is considered as a better surgical treatment[9]. In addition, many retrospective non-randomised studies indicate that sublobar resection may be an acceptable surgical treatment for early lung cancers, also when performed by VATS. In summary, most of clinical stage I-II of NSCLC select surgical resection and it has been a very important treatment [10, 11]. However, for 70 plus years old elderly patients with increasing risk of comorbidities, they may tend to select radiation rather than surgery.

3. Introduction to SBRT

Stereotactic body radiotherapy (SBRT), also called stereotactic ablative radiotherapy (SABR) now is a most used radiation therapy method. The utilization of Lung SBRT is rapidly growing and its consensus documents has been greatly developed recently [12, 13]. Some stage I NSCLC patients can’t get surgical resection due to many bad health conditions, including

multiple comorbidities, older age, and poor performance [10, 14]. SBRT has been proved to be an excellent alternative to surgery for inoperable patients and show less treatment morbidity and treatment related mortality compared to surgery [15-18]. There are also evidences to say that SBRT has become an effective treatment option for patients with stage I NSCLC[19]. Ezer et al performed a retrospective research to compare the outcomes after SBRT versus limited resection in 65 plus years old patients with early stage NSCLC. The results showed that 16 percent (362/2243) patents received SBRT and adjusted overall survival has no difference between the two treatment ($p>0.05$). In this study, SBRT-treated patients had higher comorbidity scores and older than those treated by limited resection, so the latter one has a better non-adjusted overall survival. At the same time, however, limited resection has a more frequent respiration-related adverse event[20]. In fact, the difference of health condition between SBRT and surgery exists in multiple studies. The patients in the SBRT group often have older ages, worse lung function, and more comorbidities comparing with those in surgery group [21]. Until now, there are three randomized trials comparing SBRT with golden standard treatment, but all closed very early because of very slow accrual[19]. Therefore, Whether SBRT is an ideal alternative treatment for elderly, early stage NSCLC needs more evidences to support.

4. Surgery types and their operation criteria

Surgery is commonly used for NSCLC patients, if they meet all surgical conditions. But different types of surgery should be selected based on patients' tumor size and its stage. Basically, there are three types of surgery for non-small cell lung cancer.

- **Wedge and segmental resection:** This is a small operation also called sublobar resection which is to remove only cancerous tissue from the lung. Segmental resection may remove more tissues than wedge resection, but still not a whole lobe. This operation is performed when the cancer is in the early stages and is only present in one very small area of the lung. This operation could spare more pulmonary function for stage I NSCLC patients compared with lobectomy and pneumonectomy [22]. However, this approach might be used if a person has very bad health condition, for example, s/he doesn't have enough lung function to withstand removing the whole lobe.
- **Lobectomy:** This operation removes an entire lobe containing the tumors from the lung. The lungs are made up of 5 lobes (3 on the right and 2 on the left). This is a very common type of surgery for non-small cell lung cancer. This operation is preferred even if a patient has a small tumor. Lobectomy has been proved to be a standard treatment for stage I NSCLC even though sublobar resection spares more pulmonary function [23]. A Meta-Analysis conducted by Yang Zhang is to compare outcomes of lobectomy with wedge and segmental resection for stage I NSCLC patients. The combined hazard ratio between sublobar resection and lobectomy is 1.53 with confidence interval 1.42 to 1.67, which indicates that lobectomy offers better overall survival than wedge and segmental resection for stage I NSCLC patients [24].

- **Pneumonectomy:** This operation removes an entire lung. This might be needed if the tumor is spread to more than one lobe and even close to the center of the chest. Most patients are not expected to get pneumonectomy due to its adverse outcomes, especially among elderly patients [25]. However, recent evidence shows that pneumonectomy among elderly patients doesn't offer inferior outcomes compared with among younger patients [26, 27]. Therefore, old age should not be a factor to refuse pneumonectomy for early stage NSCLC patients.

CHAPTER 2

STUDY PURPOSE

1. Study Purpose

1.1 To determine if SBRT offers equivalent overall survival compared to surgery in elderly early stage non-small cell lung cancer patients.
1.2 To determine if SBRT offers equivalent overall survival compared with different types of surgery (lobectomy, pneumonectomy, wedge and Segmental Resection) in elderly early stage non-small cell lung cancer patients.

CHAPTER 3

METHODOLOGY

1. Introduction to National Cancer Database

National Cancer Data Base (NCDB) is a nationally-recognized clinical oncology database sourced from hospital registry data that are collected in more than 1,500 Commission on Cancer (CoC)-accredited facilities. NCDB data collect approximately 70 percent of newly diagnosed cancer cases in United States each year and 30 million historical records, which are used to analyze and track patients with malignant neoplastic diseases, their treatments, and outcomes. The NCDB now contains data from more than 21 million cancer patients diagnosed between 1985 and 2013, and has been recognized as the largest clinical registry in the world. This database has been used to study factors related to cancer diagnosis, disease treatment, and overall survival [28]. There are many types of cancer data sets in the NCDB, like non-small cell lung cancer data set, breast data set, and prostate data set, et. al.[29].

2. Non-small cell lung cancer data set

Non-small cell lung cancer (NSCLC) data set (NCDB Lung PUF 2013) is one of the largest data sets in the NCDB. There are 1,163,309 cancer cases (observations) in this data set and 123 variables. All the variables are categorized into seven groups according to their features for NSCLC: Case Key, Facility, Patients Demographics, Cancer Identification, Stage of Disease, Treatment, and Outcomes.

Case key variable represents the patient ID, thus this variable is used to identify subject ID for data analysis. Facility group is to identify where the patient registered once s/he got cancer. Patients' demographics contain patients' basic information, like age, race, gender, insurance

status, et.al. For Cancer Identification group, there are five cancer information included, such as Years of Diagnosis, Primary Site, Laterality, Histology and Grade. The group of Stage of Disease indicates which stage of lung cancer diagnosed by both clinical and pathologic methods. For treatment group, there are different kinds of treatment methods included, such as surgery, radiation, hormone, chemotherapy, et.al. For this study, we select patients who get either surgery or SBRT. Outcome group contains patients' disease outcome at last contact, like vital status, days from diagnosis to last contact or death, et.al. (Table 1).

3. Include/Exclude Criteria

The NCDB Lung PUF 2013 data set includes 1,163,309 patients. But we only want patients who are at early stage of NSCLC with age above 70. Therefore, the following sample selection criteria were used to choose study population of interest. In NCDB lung cancer data set, the patients are diagnosed between 2004 and 2013. We only select those with diagnosis year from 2004 to 2012, since the vital status for those with diagnosis year in 2013 is missing. Then we exclude the patients whose treatment type is neither surgery nor SBRT. However, we include those who also get other treatment type, like chemotherapy, apart from surgery or SBRT. Then we exclude those with age under 70. Finally, we exclude the patients whose cancer is not at early stage. We identify 'early stage' through both surgery methods and pathologic methods. For either method, we strictly rely on medical definition for early stage of lung cancer. Generally, we

identify stage I and stage II as early stage. We also exclude few patients who get chemotherapy, before surgery or SBRT. The final analytical data contains 38,899 patients (table 2).

4. Statistical Analysis

4.1 Description of each selected variable

Statistical analysis was conducted using SAS version 9.4 and SAS macros or software developed at the biostatistics & bioinformatics Winship Cancer Institute. The significant level was set at 0.05. Descriptive statistics for each selected variable were reported using DESCRIPTIVE macro. For categorical variable, the observation number and percentage of each level were reported. For continuous variable, mean, standard deviation, median, minimum, maximum and missing numbers were reported.

4.2 Associate with treatment group for each covariate

The distribution of each covariate among treatment group (Surgery vs. SBRT) was calculated using %UNICAT macro. For categorical covariates, a contingency table along with the chi-square test (parametric p-value) or Fisher's exact test (non-parametric p-value) can be produced. For numerical covariates the sample size,

mean, and median along with ANOVA test (parametric p-value) or Kruskal-Wallis test (non-parametric p-value) can be produced. The observation number and row percent of each level for each covariate among two treatment groups were also reported.

4.3 Univariate association with overall survival

When comparing overall survival between two groups, we use hazard ratio as indicator calculated by Cox proportional hazards regression model, since this model is the most used model in a survival analysis [30, 31]. Univariate association with overall survival for each selected variables was studied. Hazard ratio with 95% CI was calculated for each selected variable using %UNI_PHREG macro. Cox proportional-hazard model was used in this macro. Only one variable was added into the model for each calculation. Hazard ratio p-value and log rank p-value are both reported.

4.4 Multivariate association with overall survival

Multivariate association with overall survival for treatment group was calculated using %PHREG_SEL macro. Hazard ratio with 95% confidence interval of primary exposure variable treatment group controlling for covariates was generated by using cox proportional hazard regression model (Proc PHREG procedure in SAS). Backward

selection was conducted in the model using the maximum possible sample size at each stage of the selection process instead of restricting to the sample size from the first step as SAS does. The final sample size in the analysis model is the number of total non-missing subjects of selected variables whose p-values are less than SLSTAY values.

4.5 Multivariate association with overall survival stratified by covariates of interest

Hazard ratio for treatment group stratified covariates of interest was generated by %PHREG_SEL macro. For each level of covariates, hazard ratio along with 95% confidence interval was calculated using cox proportional hazard regression model (Proc PHREG procedure with SLICE statement in SAS). Type 3 p-value was reported to see if there is significantly different hazard ratio for each level of selected covariates. In this study, we are interested in studying these covariates, such as patients' age, charlson-deyo score and tumor stage.

4.6 Propensity score analysis

This study is an observational study, in which treatment selection is often influenced by subject characteristics. Therefore, systematic differences in baseline characteristics between SBRT and surgery subjects should be taken into account when estimating the

effect of treatment on overall survival[32]. Propensity score has been a popular method to reduce the impact of treatment selection bias on overall survival of an observational study[33]. This method allows us to replicate some characteristics of a randomized controlled trial, which is always considered the gold standard method for estimating the effects of two treatment group on overall survival[34]. The propensity score is estimated using a logistic regression model with treatment group as the dependent variable and all covariates as the independent variables. Matched data sets of two treatment group subjects are formed by matching on the propensity score. We assess the balance in baseline covariates between two treatment group subjects in the propensity-score-matched sample when the sample is formed. Then we compare the hazard ratio and Kaplan-Meier curves between the two treatments. All the analyses are finished by using several SAS macros including %CALC_PS, %GREEDMTCH, %STD_DIFF, %UNI_PHREG, and %KM_PLOT. %CALC_PS macro is used to calculate the propensity score through logistic model. The values of linear predictor will be produced for each subject in new data set called PS. %GREEDMTCH is used to select 1:1 matched samples for treatment group (SBRT vs Surgery) based on calculated values of linear predictor. One new data set called MATCHED will be produced, which contains selected subjects. Each selected SBRT patient has one surgery patient matched. %STD_DIFF is used to check balance of matched samples. Selected matched samples are satisfying if all parametric p-values for each covariate are greater than 0.1. %UNI_PHREG is used to calculate the

hazard ratio of SBRT versus surgery in matched samples. We add only treatment group in PROC PHREG model, we don't need to control other covariates for matched samples. Hazard ratio with 95% confidence interval will be generated. %KM_PLOT macro is used to produce Kaplan-Meier curves for matched samples and one year and five years survival rate for each treatment group.

4.7 Comparison between SBRT and four types of surgery

Our second aim is to determine if SBRT offers equivalent overall survival compared with different types of surgery (lobectomy, pneumonectomy, wedge and Segmental Resection). We divide surgery into four types: lobectomy, pneumonectomy, wedge and segmental resection, and other surgery types. Hazard ratio between SBRT and each type of surgery was calculated using cox proportional hazard model. Kaplan-Meier curves along with log rank p-value were also reported.

4.8 Check proportional hazard regression model assumption

Most of analysis in this study use cox proportional hazard regression model, so its assumption needs to be checked, especially for proportional assumption. First, for each analysis, Kaplan-Meier curve was produced and could be an indicator of proportional assumption check. The proportional assumption may be violated if there is a cross between two curves[35]. Next, plot of Schoenfeld residuals method is used to further check proportional hazard assumption. In this method, Schoenfeld residuals are saved in output dataset by using 'output' statement with RESSCH option followed. Plot of Schoenfeld residuals for treatment group as function of time is generated by using PROC GPLOT procedure [36]. There exists a proportional hazard assumption violation

if the plot shows the residuals are associated with time. Finally, interaction term between treatment and overall survival along with TEST statement was added in the model to test if there is violation for proportional assumption. The proportional assumption is violated if p-value generated by test statement is less than 0.05[37].

4.9 Cox non-proportional hazard regression model

Cox non-proportional hazard regression model was used when proportional assumption is violated. When using SAS to do cox non-proportional hazard regression model, the interaction term of treatment and overall survival should be added into MODEL statement. Produced parameter estimation (β) for both treatment and interaction could be used to calculate the hazard ratio between treatment groups over time. Figure of change of hazard ratio over time could also be reported.

CHAPTER 4

STUDY RESULTS

1. Description of data set

Among 38,899 patients, 3,466 (8.9%) got SBRT treatment and 35,433 (91.1%) got surgery. There are 18,363 (47.2%) male patients and 20.536 (52.8%) females. The average age at diagnosis year of the selected patients is 76.4 with standard deviation 4.69. There are 3,466 (8.9%) patients between 70 and 74 years old and 1,3185 (33.9%) patients between 75 and 79 and 9905 (25.4%) patients are equal or greater than 80 years old. Most patients are white (35,537, 91.4%). Most of the selected patients have Medicare insurance (34,109, 84.7%) and 3,908 (10.0%) patients have private insurance. (Table 3).

2. Associate with treatment group for each covariate

The distribution of each covariate among treatment group (Surgery vs. SBRT) was calculated using %UNICAT macro. The observation number and row percent of each level for each covariate among two treatment groups were reported. The covariate may affect the decision on treatment selection, if there is a significant association between covariates and treatment group. Therefore, it may be a potential confounder, which should be controlled in the following cox regression model for survival analysis. We assessed the univariate association of each covariate with two cohorts of Surgery and SBRT by using the chi-square test for categorical covariates and ANOVA for numerical covariates. We found that most of covariates have a significant associate with treatment group. We have more concern on the distribution of patients' age. We found that patients with age above 80 have higher chance to get SBRT rather than surgery than those with age less than 80. Age is a critical factor for overall survival time after treatment. Therefore, age is a confounder of relationship between treatment group and overall survival. Age is also an indication of surgery. That's why aged patients have high probability to receive SBRT rather than surgery. (Table 4)

3. Univariate association with overall survival

Among 3,466 70+-year-old early stage NSCLC patients who got SBRT treatment, 1,908(55%) patients died at the end date and 1,558 (45%) patients are censored. One year survival rate of SBRT is 81.2% and 5 years survival rate is 24.9%. For 35,433 70+-year-old early stage NSCLC patients, 14,451 (41%) died at the end date and 20,982 (59%) patients are censored. One year survival rate of surgery treatment is 87.5% and 5 years survival rate is 53.5%, which are both better than SBRT perspective. According to log-rank test, the p value is less than 0.001, which means the average overall survival of surgery is significantly better than SBRT. The hazard ratio without controlling covariates between SBRT and surgery is 2.01 (1.91-2.11) with p-value less than 0.001 (table 6). We further divide surgery into four types: Lobectomy Wedge, Segmental Resection, Pneumonectomy and Other Surgery Types. We already know that lobectomy wedge is the gold standard for an early stage NSCLC. Whether other types of surgery is better than SBRT should be studied. We further compare the outcome of different surgery type with SBRT. We found that Pneumonectomy is not significantly better than SBRT. We further explore univariate association with overall survival for each covariate. As we mentioned above, age has a significant association with overall survival. The older, the higher hazard ratio one patient will get.

4. Multiunivariate association with overall survival by treatment group

Multiunivariate association with overall survival by treatment group was analyzed in

this study, which explore the relationship between overall survival and treatment group after controlling other covariates of interest. Hazard ratio with 95% confidence interval of treatment group controlling for covariates was generated by using %PHREG_SEL macro, in which cox proportional hazard regression model (Proc PHREG procedure in SAS) is used to generate results. Backward selection was conducted in the model using the maximum possible sample size at each stage of the selection process instead of restricting to the sample size from the first step as SAS does. The final sample size in the analysis model is the number of total non-missing subjects of selected variables whose p-values are less than SLSTAY values. The hazard ratio between SBRT and surgery is 2.04 with confidence interval 1.92 to 2.16 after controlling following covariates: Case Key/ Patient ID, Facility Key, Facility Type, Facility Location, Age at Diagnosis, Race, Sex, Insurance Status, Median Income Quartiles, Urban/Rural 2003, Charlso-Deyo Score, Years of Diagnosis, Primary Site, Laterality, Histology, Grade, NCDB Analytic Stage Group, AJCC Clinical N, AJCC Clinical M, Regional Lymph Nodes Positive, AJCC Clinical Stage Group, AJCC Pathologic M, Definitive Surgical Procedure, Chemotherapy. We can conclude that surgery offers better overall survival compared with SBRT at base line of other covariates.

5. Propensity score matching

Propensity score 1:1 matching is used to balance covariates for treatment group. The

macros used in the analysis

contain %CALC_PS, %GREEDMTCH, %STD_DIFF, %UNI_PHREG,

and %KM_PLOT. %CALC_PS macro is used to calculate the propensity score by logistic model. The values of linear predictor will be produced for each subject in new data set called PS. %GREEDMTCH is used to select 1:1 matched samples for treatment group (SBRT vs Surgery) based on calculated values of linear predictor. One new data set called MATCHED will be produced, which contains selected subjects. We have 3202 patients selected for each treatment group. Therefore, 264 SBRT patients are excluded, since we have overall 3466 patients who get SBRT in our data set. Each selected SBRT patient has one surgery patient matched. %STD_DIFF is used to check balance of matched samples (Table 8). All parametric p-values for each covariate are greater than 0.1, which means matched samples is satisfying. %UNI_PHREG is used to calculate the hazard ratio of SBRT vs surgery in matched samples. We add only treatment group in PROC PHREG model, we don't need to control other covariates for matched samples. Hazard ratio between SBRT and surgery is 2.00 with confidence interval between 1.86 and 2.16, which is a little smaller than 2.04 calculated in multivariate association with overall survival but still indicates that SBRT is significantly worse than surgery. %KM_PLOT macro is used to produce Kaplan-Meier curves for matched samples (Figure 2) and one year and five years survival rate for each treatment group (Table 10). Both one-year survival rate of SBRT and surgery is above 80%. For SBRT patients, one-year survival rate is 81.4 (CI: 80.0%-82.7%), while

surgery patients are 88.4% (CI: 87.2%-89.5%). For five-year survival rate, SBRT patients are much worse that surgery ones. For SBRT patients, five-year survival rate is 24.9% (CI: 22.5%-27.3%), while surgery patients are 52.0% (CI: 49.4%-54.5%).

6. Multivariate association with overall survival by treatment interaction with covariates

The hazard ratio may be different between SBRT and surgery among each level of covariates of interest. Age, charlson-deyo score and tumor stage have a great effect on survival rate, so their interaction items with treatment are added into model. The outcome comparisons stratified by age at diagnosis (Categorical) controlling for other covariates are showed in table 8. The hazard ratio between SBRT and surgery decreases as age increases with type 3 p-value less than 0.001, but all three hazard ratio is significantly greater than 1, which means, among each age level, SBRT is significantly inferior to surgery for early stage NSCLC patients. Similar results are produced after stratified by charlson-deyo score and tumor stage (table 9, table 10). All results indicate that outcomes from SBRT is significantly worse than surgery.

7. Comparing overall survival between SBRT and three types of surgery

From table 7, we already know that older patients have a higher hazard ratio than

younger patients. Patients with higher charlson-deyo score have a higher hazard ratio than those with lower charlson-deyo score. Patients at tumor stage II have a higher hazard ratio than those at stage I. These patients with higher-hazard-ratio covariates have a worse health condition. Although outcomes of SBRT is significantly worse than surgery in each level of covariates of interest, we can still find that hazard ratio between SBRT and surgery is decreasing with worse health condition. Health condition is also a factor for patients to receive which type of surgery. Therefore, patients with different surgery type may have different survival rate.

The overall survival between SBRT and each type of surgery is generated for two time points: 12 moth survival and 60 month survival. When comparing SBRT with lobectomy and comparing SBRT with wedge and segmental resection, the results don't change. At 12-month-survival point, the survival rate for SBRT is 81.2%, which is less than the survival rate for wedge and segmental resection (87.5%) and for lobectomy (88.3%). At 60-month-survival point, the survival rate for SBRT is 24.9%, which is also less than the survival rate for wedge and segmental resection (46.2%) and lobectomy (56.8%) (Table 15, table 16). Therefore, the outcomes of SBRT are significantly worse than wedge and segmental resection and also worse than lobectomy (log rank p value<0.001) (Figure 4 and Figure 5).

When comparing outcomes of SBRT and pneumonectomy, the results are complicated.

At 12-month-survival point, the survival rate for SBRT is 81.2%, which is much greater than pneumonectomy (67.3%). But at 60-month-survival point, the survival rate for SBRT is 24.9%, which is much less than pneumonectomy (38.7%) (Table 14). From KM plot curves, there is a crossing between two lines (Figure 3). This kind of result indicates the proportional hazard assumption may be violated, which means we can't get average hazard ratio between SBRT and pneumonectomy. The Cox non-proportional hazard model should be used if proportional hazard assumption is violated. Therefore, proportional hazard assumption need to be tested next.

8. Test the proportional hazard assumption in Cox model

Kaplan-Meier curve indicates that there may exist proportional hazard assumption violation for Cox model used in comparing outcomes of SBRT with pneumonectomy, because there is a crossing between these two lines (Figure 3). Then, plot of Schoenfeld residuals method is used to further check proportional hazard assumption. Firstly, Schoenfeld residuals for each observation are generated in output dataset by using 'output' statement with RESSCH option followed. Next, plot of Schoenfeld residuals by treatment group as function of time is generated by using PROC GPLOT procedure (Y axis is Schoenfeld residual by treatment group, X axis is time).There exists a proportional hazard assumption violation, because the plot shows the residuals are associated with time, thus, there is a relationship between residuals and time but not

randomized (Figure 7). Finally, interaction term between treatment and overall survival was added in the model followed by TEST statement to test if there is violation for proportional assumption. As a result, the proportional assumption is violated since p-value generated by test statement is less than 0.05 (Table 18).

9. Comparing outcomes between SBRT and pneumonectomy using Cox non-proportional hazard regression model

Cox non-proportional hazard regression model need to be used since proportional assumption is violated when comparing outcomes between SBRT and pneumonectomy. When using SAS to do cox non-proportional hazard regression model, the interaction term between treatment and overall survival was added into MODEL statement. Parameter estimations (β) for both treatment and interaction generated by the model were used to calculate the hazard ratio between treatment groups against time [38]. The parameter estimation for treatment group (β_1) is -0.47164, which indicates that hazard ratio between SBRT and pneumonectomy is less than 1 at base line. The parameter estimation (β_2) for interaction item (treatment group * overall survival) is 0.03238 greater than 0, which indicates hazard ratio between SBRT and pneumonectomy increases over time (Table 19). The hazard ratio between SBRT and pneumonectomy against time could be calculated by below formula [39]:

$$\text{Hazard Ratio} = e^{(\beta 1+\beta 2*os)} = e^{(-0.47164+0.03238*os)}$$

Based on above formula, four time points of interest were calculated including 6 months, 12 months, 36 months and 60 months. At 6th month after treatment, the hazard ratio between SBRT and pneumonectomy is 0.75778 with 95% confidence interval between 0.50276 and 0.8119, which indicates that SBRT offer better overall survival

than pneumonectomy at 6th month after treatment. After one year (12 months), the hazard ratio between SBRT and Pneumonectomy is greater than 1, which indicates that SBRT offers worse overall survival compared with pneumonectomy (HR=2.00 (1.09, 2.71) at 36th month and HR=4.35 (2.74, 11.52) at 60th month) (Table 20). Spline plot is used to show the change of hazard ratio over time between SBRT vs Pneumonectomy and its 95% confidence interval (Figure 8) [40].

CHAPTER 5

DISCUSSION

1. Outcome of SBRT versus surgery

This study is a retrospective survival analysis. All patients collected from National Cancer Database are non-small cell lung cancer patients. Their demographics information, cancer information, treatment information, and outcomes were collected. The purpose of this study is to determine if SBRT give equivalent overall survival compared with surgery in elderly, early stage non-small cell lung cancer patients. Therefore, we only choose 70-year-old and plus patients who were treated either SBRT or surgery and were at early tumor stage. Include and exclude criteria was set based on our purpose and medical knowledge. After exclusion, 38,899 patients were left, among which 35,433(91.1%) patients got surgery and 3,466(8.9%) got SBRT. Even though most patients got surgery, the percent of SBRT increases by 10.72% from year 2004 to year 2012 and its sample size is enough to do data analysis. SAS version 9.4 is used to

do analysis. The rude hazard ratio between SBRT and surgery is 2.01 (1.91-2.11) without controlling covariates. It indicates that the hazard for a 70+ early stage NSCLC patient who got SBRT is around 2 times high as the hazard for one who got surgery. The hazard ratio changes a little from 2.01 to 2.04 (1.92-2.16) after controlling covariates, including facility type, facility location, age, sex, race, median income, urban or rural, Charlson-Deyo score, year of diagnosis, primary site, laterality, histology, grade, AJCC analytic stage group, chemotherapy or not, and tumor size. We can conclude that, at the 5 percent level of significance, we reject the null hypothesis that SBRT gives equivalent overall survival for elderly, early stage NSCLC patients compared with surgery.

Propensity score 1:1 matching is used to compare hazard ratio between SBRT and surgery in matched samples. Propensity score is generated by using logistic regression model with treatment as dependent variable and covariates as independent variables. Based on propensity score, matched samples are selected. 3202 patients for each treatment group are selected. The sample size is large enough to do analysis. The balance of the matched data was checked and the p-value for all covariates are greater than 0.1, which indicates the balance is satisfying. The average hazard ratio between SBRT and surgery is 2.00 (CI: 1.86, 2.16), which has a same conclusion as multivariate analysis above. Therefore, we conclude that, for elderly, early stage non-small cell lung cancer patients, surgery is a better choice than SBRT. However, for special surgery types, although SBRT offers inferior overall survival compared with lobectomy and

wedge and segmental resection, SBRT could be a potential alternative to pneumonectomy.

2. Strengths and limitations

The data of this study were collected by National Cancer Database, which is a nationally-recognized clinical oncology database sourced from hospital registry data that are collected in more than 1,500 Commission on Cancer (CoC)-accredited facilities. The lung cancer data set is one of the largest data sets in NCDB with 1,163,309 patients and 123 related variables. The final analytical data contains 38,899 patients after exclude those who don't meet the include criteria. Therefore, the sample size is large enough to do data analysis. Another strength is rich treatment information of patients. The information on cancer identity and stage and its treatment methods were detailed collected, which is very helpful for research on relationship between treatment methods and overall survival. In addition, numerous macros were used in this study and it saves a lot of time and generates standard tables. In these macros, both basic and advanced statistical methods and models are used. T-test and chi-square were used to do basic analysis. For categorical variable, the observation number and percentage of each level were reported. For continuous variable, mean, standard deviation, median, minimum, maximum and missing numbers were reported. When comparing outcomes of treatment group, Cox proportional hazard regression model was used to compare average hazard ratio between SBRT and surgery. Propensity score

matching was used to minimize confounders' effect on outcomes by balancing the each level of covariates for each treatment group. Cox non-proportional hazard regression model was used if proportional hazard assumption is violated. For example, when comparing the outcomes between SBRT and Pneumonectomy, change of hazard ratio against time was reported since proportional hazard assumption is violated and average hazard ratio is not fit.

One big limitation in this study is selection bias may exist. There is no any else health information of patients other than demographics information and cancer-related information. We don't know anything about patients' other disease, like diabetes, high blood pressure, smoking status and other infectious diseases. If these characteristics for different treatment group are significantly different, selection bias exists and the results will be affected by those confounders. However, we have large sample size and there are not evidences to show that the patients' characteristics are significantly different between two treatment groups, so the result is still meaningful. We need more studies, especially randomized studies, to prove this conclusion,

Another potential problem should be the overall 5-year survival rate for patients who received SBRT in this study seems to be lower. One study conducted by Dr. Onishi in 2011 shows that overall 5-year survival rate for total cases who received SBRT was 69.5% with confidence interval 58.55 to 80.1% [41]. This result is different from the

one of this study in which overall 5-year survival rate for all cases who received SBRT was 24.9% with confidence interval 22.7% to 27.3%. The difference of the two results for SBRT results in a different conclusion in regard of whether SBRT is an efficient alternative to standard treatment-lobectomy for early stage lung small cell lung cancer patients. The reasons for the difference of results should be the difference of patients' health condition in two studies. The first difference is the age. The median age out of 87 patients in Dr. Onishi's study is 74, which is much smaller than it in this study in which the median age for SBRT patients is 79. This difference may explain why the results from these two studies are so different.

CHAPTER 6

CONCLUSION AND RECOMMENDATIONS

1. Study conclusion

Based on our results, surgery is basically better than SBRT for elderly early stage NSCLC patients, but not any types of surgery is better than SBRT. Generally, surgery is categorized to three big types: lobectomy, pneumonectomy, and wedge and segmental resection. Which type should be selected is based on patients' health condition and tumor identification. After comparison between SBRT and each type of surgery, both lobectomy and wedge & segmental resection are still much better than SBRT. But for pneumonectomy patients, the result is much complicated. Cox

proportional hazard regression model can't be used since proportional assumption is violated based on plot of Schoenfeld residual and p-value generated by TEST statement in SAS. Therefore, cox non-proportional hazards regression model is used to generate functional relationship between hazard ratio and overall survival time. Less than 12 months after treatment, SBRT offers better overall survival than pneumonectomy, but after 12^{th} month, the outcome of SBRT is worse than pneumonectomy.

2. Recommendations

Both lobectomy and wedge and segmental resection offer better overall survival compared with SBRT. So SBRT may not an excellent alternative to lobectomy and wedge and segmental resection. However, for patients who have to receive pneumonectomy, they need to take SBRT into consideration. We recommend patients whose health condition is too weak to get pneumonectomy or those who have high risk for pneumonectomy should consider to get SBRT instead of pneumonectomy.

In addition, when we explore multivariate association with overall survival by treatment group interaction with covariates, we found that hazard ratio between SBRT and surgery is decreasing with worse health condition. Therefore, SBRT is a potential alternative to surgery for inoperable patients due to older age or with multiple comorbidities. We need more studies to prove it.

CHAPTER 7

REFERENCES

1. Surveillance, E., and End Results (SEER) Program, *SEER Stat Fact Sheets: Lung and Bronchus Cancer.*
2. Uzel, E.K. and U. Abacioglu, *Treatment of early stage non-small cell lung cancer: surgery or stereotactic ablative radiotherapy?* Balkan Med J, 2015. **32**(1): p. 8-16.
3. Albain, K.S., et al., *Concurrent cisplatin/etoposide plus chest radiotherapy followed by surgery for stages IIIA (N2) and IIIB non-small-cell lung cancer: mature results of Southwest Oncology Group phase II study 8805.* J Clin Oncol, 1995. **13**(8): p. 1880-92.
4. Albain, K.S., et al., *Radiotherapy plus chemotherapy with or without surgical resection for stage III non-small-cell lung cancer: a phase III randomised controlled trial.* Lancet, 2009. **374**(9687): p. 379-86.
5. Boffa, D.J., et al., *Now or later: evaluating the importance of chemotherapy timing in resectable stage III (N2) lung cancer in the National Cancer Database.* Ann Thorac Surg, 2015. **99**(1): p. 200-8.
6. Howington, J.A., et al., *Treatment of stage I and II non-small cell lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines.* Chest, 2013. **143**(5 Suppl): p. e278S-313S.
7. Ramnath, N., et al., *Treatment of stage III non-small cell lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines.* Chest, 2013. **143**(5 Suppl): p. e314S-40S.
8. Whitson, B.A., et al., *Survival after lobectomy versus segmentectomy for stage I non-small cell lung cancer: a population-based analysis.* Ann Thorac Surg, 2011. **92**(6): p. 1943-50.
9. Van Schil, P.E., et al., *Surgical treatment of early-stage non-small-cell lung cancer.* EJC Suppl, 2013. **11**(2): p. 110-22.
10. Raz, D.J., et al., *Natural history of stage I non-small cell lung cancer: implications for early detection.* Chest, 2007. **132**(1): p. 193-9.
11. Rosen, J.E., et al., *The Natural History of Operable Non-Small Cell Lung Cancer in the National Cancer Database.* Ann Thorac Surg, 2016.
12. Brada, M., A. Pope, and M. Baumann, *SABR in NSCLC--the beginning of the end or the end of the beginning?* Radiother Oncol, 2015. **114**(2): p. 135-7.
13. Louie, A.V., et al., *Management of early-stage non-small cell lung cancer using stereotactic ablative radiotherapy: controversies, insights, and changing horizons.* Radiother Oncol, 2015. **114**(2): p. 138-47.
14. Liu, H.W., et al., *Outcomes in stage I non-small cell lung cancer following the introduction of stereotactic body radiotherapy in Alberta - A population-based study.* Radiother Oncol, 2015. **117**(1): p. 71-6.
15. Timmerman, R., et al., *Stereotactic body radiation therapy for inoperable early stage lung cancer.* JAMA, 2010. **303**(11): p. 1070-6.
16. Chang, J.Y., et al., *Stereotactic ablative radiotherapy versus lobectomy for operable stage I*

non-small-cell lung cancer: a pooled analysis of two randomised trials. Lancet Oncol, 2015. **16**(6): p. 630-7.

17. Nguyen, N.P., et al., *Can stereotactic fractionated radiation therapy become the standard of care for early stage non-small cell lung carcinoma.* Cancer Treat Rev, 2008. **34**(8): p. 719-27.
18. Solda, F., et al., *Stereotactic radiotherapy (SABR) for the treatment of primary non-small cell lung cancer; systematic review and comparison with a surgical cohort.* Radiother Oncol, 2013. **109**(1): p. 1-7.
19. Roesch, J., N. Andratschke, and M. Guckenberger, *SBRT in operable early stage lung cancer patients.* Transl Lung Cancer Res, 2014. **3**(4): p. 212-24.
20. Ezer, N., et al., *Outcomes after Stereotactic Body Radiotherapy versus Limited Resection in Older Patients with Early-Stage Lung Cancer.* J Thorac Oncol, 2015. **10**(8): p. 1201-6.
21. Crabtree, T.D., et al., *Stereotactic body radiation therapy versus surgical resection for stage I non-small cell lung cancer.* J Thorac Cardiovasc Surg, 2010. **140**(2): p. 377-86.
22. Keenan, R.J., et al., *Segmental resection spares pulmonary function in patients with stage I lung cancer.* Ann Thorac Surg, 2004. **78**(1): p. 228-33; discussion 228-33.
23. Ginsberg, R.J. and L.V. Rubinstein, *Randomized trial of lobectomy versus limited resection for T1 N0 non-small cell lung cancer. Lung Cancer Study Group.* Ann Thorac Surg, 1995. **60**(3): p. 615-22; discussion 622-3.
24. Zhang, Y., et al., *Meta-analysis of lobectomy, segmentectomy, and wedge resection for stage I non-small cell lung cancer.* J Surg Oncol, 2015. **111**(3): p. 334-40.
25. Kim, T.H., et al., *Pneumonectomy for Clinical Stage I Non-Small Cell Lung Cancer in Elderly Patients over 70 Years of Age.* Korean J Thorac Cardiovasc Surg, 2015. **48**(4): p. 252-7.
26. Mizushima, Y., et al., *Survival and prognosis after pneumonectomy for lung cancer in the elderly.* Ann Thorac Surg, 1997. **64**(1): p. 193-8.
27. Speicher, P.J., et al., *Survival in the elderly after pneumonectomy for early-stage non-small cell lung cancer: a comparison with nonoperative management.* J Am Coll Surg, 2014. **218**(3): p. 439-49.
28. Lerro, C.C., et al., *Comparison of cases captured in the national cancer data base with those in population-based central cancer registries.* Ann Surg Oncol, 2013. **20**(6): p. 1759-65.
29. Bilimoria, K.Y., et al., *The National Cancer Data Base: a powerful initiative to improve cancer care in the United States.* Ann Surg Oncol, 2008. **15**(3): p. 683-90.
30. Sargent, D.J., *A general framework for random effects survival analysis in the Cox proportional hazards setting.* Biometrics, 1998. **54**(4): p. 1486-97.
31. Zhang, M.J., *Cox proportional hazards regression models for survival data in cancer research.* Cancer Treat Res, 2002. **113**: p. 59-70.
32. Austin, P.C., *An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies.* Multivariate Behav Res, 2011. **46**(3): p. 399-424.
33. Austin, P.C., *A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003.* Stat Med, 2008. **27**(12): p. 2037-49.
34. Rosenbaum, P.R. and D.B. Rubin, *Reducing Bias in Observational Studies Using Subclassification on the Propensity Score.* Journal of the American Statistical Association, 1984. **79 (387): 516-524**.

35. Wilson, M.G., *Using SAS® to Assess and Model Time-to-Event Data with Non-Proportional Hazards.* MWSUG, 2010. **125**.
36. Kohl, M., et al., *PSHREG: A SAS macro for porportional and nonproportional subdistribution hazards regression.* Computer methods and programs in biomedicine, 2014. **II8 (2015): 218-233**.
37. Ata, N. and M.T. SÄozer, *COX REGRESSION MODELS WITH NONPROPORTIONAL HAZARDS APPLIED TO LUNG CANCER SURVIVAL DATA.* Journal of Mathematics and Statistics, 2007. **36(2):157-167**.
38. Devarajan, K. and N. Ebrahimi, *On penalized likelihood estimation for a non-proportional hazards regression model.* Stat Probab Lett, 2013. **83**(7): p. 1703-1710.
39. Borucka, J., *Extensions of Cox model for non-proportional hazards purpose.* PhUSE, 2013. **07**.
40. Giorgi, R., et al., *A relative survival regression model using B-spline functions to model non-proportional hazards.* Stat Med, 2003. **22**(17): p. 2767-84.
41. Onishi, H., et al., *Stereotactic body radiotherapy (SBRT) for operable stage I non-small-cell lung cancer: can SBRT be comparable to surgery?* Int J Radiat Oncol Biol Phys, 2011. **81**(5): p. 1352-8.

CHAPTER 8

TABLES AND FIGURES

Table 1: Variables Selected from NACLC data set

Variable Group Name	**Variable Label**	**Variable Name**
Case Key	Case Key/ Patient ID	PUF_CASE_ID
Facility	Facility Key	PUF_FACILITY_ID
	Facility Type	FACILITY_TYPE_CD
	Facility Location	FACILITY_LOCATION_CD
Patients Demographics	Age at Diagnosis	AGE
	Race	RACE
	Sex	SEX
	Insurance Status	INSURANCE_STATUS
	Median Income Quartiles	MED_INC_QUAR_00
	Urban/Rural 2003	UR_CD_03
	Percent No High School Degree Quartiles 2000	NO_HSD_QUAR_00
	Great Circle Distance	CROWFLY
	Charlso-Deyo Score	CDCC_TOTAL
Cancer Identification	Years of Diagnosis	YEAR_OF_DIAGNOSIS

	Primary Site	PRIMARY_SITE
	Laterality	LATERALITY
	Histology	HISTOLOGY
	Grade	GRADE
Stage of Disease	NCDB Analytic Stage Group	ANALYTIC_STAGE_GROUP
	AJCC Clinical N	TNM_CLIN_N
	AJCC Clinical M	TNM_CLIN_M
	Regional Lymph Nodes Positive	REGIONAL_NODES_POSITIVE
	AJCC Clinical Stage Group	TNM_CLIN_STAGE_GROUP
	AJCC Pathologic M	TNM_PATH_M
Treatment	Definitive Surgical Procedure, Days from Dx	DX_DEFSURG_STARTED_DAYS
	Radiation, Days from Dx	DX_RAD_STARTED_DAYS
	Chemotherapy	RX_SUMM_CHEMO
Outcomes	Last Contact or Death, Months from Dx	DX_LASTCONTACT_DEATH_MONTHS
	Vital Status	VITAL_STATUS

Table 2: Diagram for Study Population Selection

Selection and Exclusion Criteria	Sample Size	Excluded
NCDB Lung PUF Cancer Cases	1163309	-
Include diagnosis year from 2004 to 2012	1039990	123319
Include treatment type either SBRT or Surgery	277272	762718
Keep sequence numbers with 00 and 01	199169	78103
Exclude cases with class of case equal 0	193633	5536
Include cases with invasive behavior (BEHAVIOR=3)	193282	351
Select cases at early stage: (REGIONAL_NODES_POSITIVE equal 0 98 99; TNM_CLIN_N equal 0 ;TNM_CLIN_M equal 0 ; TNM_CLIN_STAGE_GROUP not equal 3 3A 3B 4; TNM_PATH_M not equal 1 1A 1B ;ANALYTIC_STAGE_GROUP not equal 3 4; CS_METS_AT_DX not between 15 and 50; CS_METS_DX_BONE not equal 1; CS_METS_DX_BRAIN not equal 1; CS_METS_DX_LIVER not equal 1; CS_METS_DX_LUNG not equal 1).	88456	104826
Include patients with radiation only on chest or lung (RAD_TREAT_VOL equal 0 10 11 40 99).	88373	83
Exclude cases with additional palliative care	87953	420
Exclude cases with chemotherapy ahead of surgery or SBRT	87256	697

Selection and Exclusion Criteria	Sample Size	Excluded
Include patients with histology in (8012 8013 8014 8020 8021 8022 8030 8031 8032 8033 8046 8050 8052 8070 8071 8072 8073 8074 8076 8078 8082 8083 8084 8123 8140 8141 8200 8240 8244 8246 8249 8250 8251 8252 8253 8254 8255 8260 8310 8320 8323 8430 8480 8481 8490 8550 8560 8570 8574 8575)	87103	153
Exclude cases with missing overall survival	85497	1606
Exclude cases with delayed treatment. (If the number of days from diagnosis date to treatment date (Surgery or SBRT) is greater than 180 days, we will exclude those patients).	84594	903
Excluede cases with age under 70	38899	45695

Table 3: Descriptive statistics for each selected variable

Variable Name	Variable Label		Level	N (%) = 38899
FACILITY_TYPE_CD	Facility Type	1	Community Cancer Program/Other	2947 (7.6)
		2	Comprehensive Community Cancer Program	20042 (51.5)

Variable Name	Variable Label		Level	N (%) = 38899
		3	Academic/Research Program	13314 (34.2)
		4	Integrated Network Cancer Program	2596 (6.7)
FACILITY_LOCATION_CD	Facility Location	1	Northeast	8265 (21.2)
		3	South	14672 (37.7)
		4	Midwest	10554 (27.1)
		8	West	5408 (13.9)
AGE	Age at Diagnosis		Mean	76.40
			Median	76.00
			Minimum	70.00
			Maximum	90.00
			Std Dev	4.69
			Missing	0.00
SEX	Sex	1	Male	18363 (47.2)
		2	Female	20536 (52.8)

Variable Name	Variable Label		Level	N (%) = 38899
RACE	Race	1	White	35537 (91.4)
		2	Black	2163 (5.6)
		3	Others	1199 (3.1)
INSURANCE_STATUS	Primary Payor	0	No Insured	97 (0.2)
		1	Private Insurance	3908 (10.0)
		2	Medicaid Insurance	372 (1.0)
		3	Medicare or Other Government Insurance	34109 (87.7)
		9	Insurance Status Unknown	413 (1.1)
MED_INC_QUAR_00	Median Income Quartiles 2000		Not Available	1492
		1	< $30,000	4506 (12.0)
		2	$30,000 - $35,999	7066 (18.9)
		3	$36,000 - $45,999	10630 (28.4)
		4	$46,000 +	15205 (40.6)
NO_HSD_QUAR_00	Percent No High School Degree Quartiles 2000		Not Available	1499
		1	>=29%	5463 (14.6)
		2	20-28.9%	8737 (23.4)

Variable Name	Variable Label		Level	N (%) = 38899
		3	14-19.9%	9556 (25.6)
		4	< 14%	13644 (36.5)
UR_CD_03	Urban/Rural 2003	1	Metro	30310 (80.8)
		4	Urban	6271 (16.7)
		8	Rural	914 (2.4)
			Missing	1404
CDCC_TOTAL	Charlson-Deyo Score	0	0	18737 (48.2)
		1	1	13971 (35.9)
		2	2+	6191 (15.9)
YEAR_OF_DIAGNOSIS	Year of Diagnosis	2004	2004	2012 (5.2)
		2005	2005	2342 (6.0)
		2006	2006	2475 (6.4)
		2007	2007	3005 (7.7)
		2008	2008	4762 (12.2)
		2009	2009	5620 (14.4)
		2010	2010	6043 (15.5)
		2011	2011	6240 (16.0)

Variable Name	Variable Label		Level	N (%) = 38899
		2012	2012	6400 (16.5)
PRIMARY_SITE	Primary Site	C340	C340 - Main Bronchus	100 (0.3)
		C341	C341 - Upper lobe, Lung	22440 (57.7)
		C342	C342 - Middle lobe, Lung	2051 (5.3)
		C343	C343 - Lower lobe, Lung	13399 (34.4)
		C348	C348 - Overlapping lesion of lung	333 (0.9)
		C349	C349 - Lung, NOS	576 (1.5)
LATERALITY	Laterality	0	Origin of primary is bilateral,midline or unknown	133 (0.3)
		1	Origin of primary is right	22538 (57.9)
		2	Origin of primary is left	16228 (41.7)
HISTOLOGY	Histology	8012	Others	12448 (32.0)
		8070	Squamous cell carcinoma of lung	12534 (32.2)
		8140	Adenocarcinoma of lung	13917 (35.8)

Variable Name	Variable Label		Level	N (%) = 38899
GRADE	Grade	1	Well differentiated, differentiated, NOS	6492 (16.7)
		2	Moderately differentiated, moderately well differentiated, intermediate differentiation	16166 (41.6)
		3	Poorly differentiated	11424 (29.4)
		4	Undifferentiated, anaplastic	581 (1.5)
		9	Cell type not determined, not stated or not applicable, unknown primaries, high grade dysplasia	4236 (10.9)
DX_DEFSURG_STARTED_DAYS	Definitive Surgical Procedure, Days from Dx		Mean	32.75
			Median	28.00
			Minimum	0.00
			Maximum	180.00
			Std Dev	32.85
			Missing	3466.00

Variable Name	Variable Label		Level	N (%) = 38899
DX_RAD_STARTED_DAYS	Radiation, Days from Dx		Mean	65.93
			Median	59.00
			Minimum	0.00
			Maximum	180.00
			Std Dev	34.66
			Missing	35433.00
RX_SUMM_CHEMO	Chemotherapy at any CoC Facility	0	No Chemothrapy	35621 (91.6)
		1	Chemotherapy Administered	2227 (5.7)
		88	Unknown	1051 (2.7)
TREATMENT	Treatment Group (SBRT/Surgery)	1	SBRT	3466 (8.9)
		2	Surgery	35433 (91.1)
agec	Age at Diagnosis (Categorical)	1	70-74	15809 (40.6)
		2	75-79	13185 (33.9)
		3	80+	9905 (25.5)

Variable Name	Variable Label		Level	N (%) = 38899
crowfly10	Great Circle Distance in 10miles Unit		Mean	2.98
			Median	1.01
			Minimum	0.00
			Maximum	263.34
			Std Dev	10.21
			Missing	644.00
tumor_size_cm	Tumor Size in cm Unit		Mean	2.89
			Median	2.50
			Minimum	0.05
			Maximum	98.90
			Std Dev	2.68
			Missing	360.00
ana_stage	AJCC Analytic Stage Group	1	Stage 0, Stage 1 or Occult(lung only)	35411 (91.0)
		2	Stage II	3208 (8.2)
		3	AJCC Stage Group Unknown	280 (0.7)

Variable Name	Variable Label		Level	N (%) = 38899
trt_type	Treatment Type	1	Pneumonectomy	475 (1.2)
		2	Wedge and Segmental Resection	7858 (20.2)
		3	Lobectomy	26212 (67.4)
		4	Other Surgery Types	888 (2.3)
		5	SBRT	3466 (8.9)

Table 4: Univariate association of each covariate with two treatment group

			Treatment Group (SBRT/Surgery)		
Covariate	**Statistics**	**Level**	**SBRT N=3466**	**Surgery N=35433**	**Parametric P-value***
Facility Type	N (Row %)	Community Cancer Program/Other	67 (2.27)	2880 (97.73)	**<.001**
	N (Row %)	Comprehensive Community Cancer Program	1503 (7.5)	18539 (92.5)	
	N (Row %)	Academic/Research Program	1580 (11.87)	11734 (88.13)	
	N (Row %)	Integrated Network Cancer Program	316 (12.17)	2280 (87.83)	
Facility Location	N (Row %)	Northeast	622 (7.53)	7643 (92.47)	**<.001**

Covariate	Statistics	Level	Treatment Group (SBRT/Surgery) SBRT N=3466	Surgery N=35433	Parametric P-value*
	N (Row %)	South	1223 (8.34)	13449 (91.66)	
	N (Row %)	Midwest	1133 (10.74)	9421 (89.26)	
	N (Row %)	West	488 (9.02)	4920 (90.98)	
Age at Diagnosis	N (Row %)	70-74	914 (5.78)	14895 (94.22)	**<.001**
(Categorical)	N (Row %)	75-79	1013 (7.68)	12172 (92.32)	
	N (Row %)	80+	1539 (15.54)	8366 (84.46)	
Sex	N (Row %)	Male	1476 (8.04)	16887 (91.96)	**<.001**
	N (Row %)	Female	1990 (9.69)	18546 (90.31)	
Race	N (Row %)	White	3137 (8.83)	32400 (91.17)	**0.021**
	N (Row %)	Black	228 (10.54)	1935 (89.46)	
	N (Row %)	Others	101 (8.42)	1098 (91.58)	
Primary Payor	N (Row %)	No Insured	9 (9.28)	88 (90.72)	**0.002**
	N (Row %)	Private Insurance	281 (7.19)	3627 (92.81)	

Covariate	Statistics	Level	Treatment Group (SBRT/Surgery) SBRT N=3466	Surgery N=35433	Parametric P-value*
	N (Row %)	Medicaid Insurance	40 (10.75)	332 (89.25)	
	N (Row %)	Medicare or Other Government Insurance	3097 (9.08)	31012 (90.92)	
	N (Row %)	Insurance Status Unknown	39 (9.44)	374 (90.56)	
Median Income	N (Row %)	< $30,000	419 (9.3)	4087 (90.7)	**0.045**
Quartiles 2000	N (Row %)	$30,000 - $35,999	678 (9.6)	6388 (90.4)	
	N (Row %)	$36,000 - $45,999	973 (9.15)	9657 (90.85)	
	N (Row %)	$46,000 +	1296 (8.52)	13909 (91.48)	
Percent No High School	N (Row %)	>=29%	424 (7.76)	5039 (92.24)	**0.002**
Degree Quartiles 2000	N (Row %)	20-28.9%	798 (9.13)	7939 (90.87)	
	N (Row %)	14-19.9%	922 (9.65)	8634 (90.35)	
	N (Row %)	< 14%	1222 (8.96)	12422 (91.04)	
Urban/Rural 2003	N (Row %)	Metro	2722 (8.98)	27588 (91.02)	0.896
	N (Row %)	Urban	572 (9.12)	5699 (90.88)	
	N (Row %)	Rural	85 (9.3)	829 (90.7)	

Covariate	Statistics	Level	Treatment Group (SBRT/Surgery) SBRT N=3466	Surgery N=35433	Parametric P-value*
Charlson-Deyo Score	N (Row %)	0	2084 (11.12)	16653 (88.88)	**<.001**
	N (Row %)	1	871 (6.23)	13100 (93.77)	
	N (Row %)	2+	511 (8.25)	5680 (91.75)	
Year of Diagnosis	N (Row %)	2004	30 (1.49)	1982 (98.51)	**<.001**
	N (Row %)	2005	53 (2.26)	2289 (97.74)	
	N (Row %)	2006	110 (4.44)	2365 (95.56)	
	N (Row %)	2007	237 (7.89)	2768 (92.11)	
	N (Row %)	2008	298 (6.26)	4464 (93.74)	
	N (Row %)	2009	487 (8.67)	5133 (91.33)	
	N (Row %)	2010	589 (9.75)	5454 (90.25)	
	N (Row %)	2011	762 (12.21)	5478 (87.79)	
	N (Row %)	2012	900 (14.06)	5500 (85.94)	
Primary Site	N (Row %)	C340 - Main Bronchus	14 (14)	86 (86)	**<.001**
	N (Row %)	C341 - Upper lobe, Lung	2070 (9.22)	20370 (90.78)	

Covariate	Statistics	Level	Treatment Group (SBRT/Surgery) SBRT N=3466	Surgery N=35433	Parametric P-value*
	N (Row %)	C342 - Middle lobe, Lung	143 (6.97)	1908 (93.03)	
	N (Row %)	C343 - Lower lobe, Lung	1139 (8.5)	12260 (91.5)	
	N (Row %)	C348 - Overlapping lesion of lung	7 (2.1)	326 (97.9)	
	N (Row %)	C349 - Lung, NOS	93 (16.15)	483 (83.85)	
Laterality	N (Row %)	Origin of primary is bilateral,midline or unknown	15 (11.28)	118 (88.72)	**<.001**
	N (Row %)	Origin of primary is right	1905 (8.45)	20633 (91.55)	
	N (Row %)	Origin of primary is left	1546 (9.53)	14682 (90.47)	
Histology	N (Row %)	Others	956 (7.68)	11492 (92.32)	**<.001**
	N (Row %)	Squamous cell carcinoma of lung	1202 (9.59)	11332 (90.41)	
	N (Row %)	Adenocarcinoma of lung	1308 (9.4)	12609 (90.6)	
Grade	N (Row %)	Well differentiated, differentiated, NOS	301 (4.64)	6191 (95.36)	**<.001**
	N (Row %)	Moderately differentiated, moderately well differentiated, intermediate differentiation	581 (3.59)	15585 (96.41)	

Covariate	Statistics	Level	Treatment Group (SBRT/Surgery) SBRT N=3466	Surgery N=35433	Parametric P-value*
	N (Row %)	Poorly differentiated	696 (6.09)	10728 (93.91)	
	N (Row %)	Undifferentiated, anaplastic	14 (2.41)	567 (97.59)	
	N (Row %)	Cell type not determined, not stated or not applicable, unknown primaries, high grade dysplasia	1874 (44.24)	2362 (55.76)	
AJCC Analytic Stage Group	N (Row %)	Stage 0, Stage 1 or Occult(lung only)	3274 (9.25)	32137 (90.75)	**<.001**
	N (Row %)	Stage II	170 (5.3)	3038 (94.7)	
	N (Row %)	AJCC Stage Group Unknown	22 (7.86)	258 (92.14)	
Chemotherapy at any CoC Facility	N (Row %)	No Chemothrapy	3324 (9.33)	32297 (90.67)	**<.001**
	N (Row %)	Chemotherapy Administered	65 (2.92)	2162 (97.08)	
	N (Row %)	Unknown	77 (7.33)	974 (92.67)	
Great Circle Distance in 10miles Unit	N		3433	34822	0.660
	Mean		3.05	2.97	
	Median		1.16	0.99	

Covariate	Statistics	Level	Treatment Group (SBRT/Surgery) SBRT N=3466	Surgery N=35433	Parametric P-value*
Tumor Size in cm Unit	N		3353	35186	**<.001**
	Mean		2.62	2.91	
	Median		2.4	2.5	

* The parametric p-value is calculated by ANOVA for numerical covariates

and chi-square test for categorical covariates.

Table 5: Overall Survival of SBRT vs. Surgery

Treatment Group (SBRT Surgery)	No. of Subject	Event	Censored	Median Survival (95% CI)	12 Month Survival	60 Month Survival
SBRT	3466	1908 (55%)	1558 (45%)	33.4 (31.8, 35.1)	81.2% (79.9%, 82.5%)	24.9% (22.7%, 27.3%)
Surgery	35433	14451 (41%)	20982 (59%)	65.3 (64.3, 66.8)	87.5% (87.2%, 87.9%)	53.5% (52.8%, 54.1%)

Table 6: Univariate association with overall survival (Months)

			Overall Survival (Months)		

Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Treatment Group	SBRT	3466	2.01 (1.91-2.11)	**<.001**	**<.001**
(SBRT/Surgery)	Surgery	35433	-	-	
Treatment Type	SBRT	3466	1.13 (0.99-1.28)	0.063	**<.001**
	Other Surgery Types	888	1.04 (0.90-1.20)	0.620	
	Lobectomy	26212	0.51 (0.45-0.58)	**<.001**	
	Wedge and Segmental Resection	7858	0.66 (0.59-0.75)	**<.001**	
	Pneumonectomy	475	-	-	
Facility Type	Community Cancer Program/Other	2947	1.17 (1.08-1.27)	**<.001**	**<.001**
	Comprehensive Community Cancer Program	20042	1.05 (0.98-1.12)	0.167	
	Academic/Research Program	13314	0.94 (0.88-1.00)	0.057	
	Integrated Network Cancer Program	2596	-	-	

			Overall Survival (Months)		

Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Facility Location	Northeast	8265	0.96 (0.90-1.01)	0.104	**<.001**
	South	14672	1.12 (1.07-1.17)	**<.001**	
	Midwest	10554	1.18 (1.12-1.24)	**<.001**	
	West	5408	-	-	
Age at Diagnosis	70-74	15809	0.61 (0.59-0.63)	**<.001**	**<.001**
(Categorical)	75-79	13185	0.77 (0.74-0.80)	**<.001**	
	80+	9905	-	-	
Sex	Male	18363	1.44 (1.39-1.48)	**<.001**	**<.001**
	Female	20536	-	-	
Race	Others	1199	0.79 (0.71-0.87)	**<.001**	**<.001**
	Black	2163	0.96 (0.90-1.03)	0.248	
	White	35537	-	-	

			Overall Survival (Months)		

Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Primary Payor	No Insured	97	1.14 (0.80-1.63)	0.475	0.153
	Private Insurance	3908	1.06 (0.90-1.24)	0.478	
	Medicaid Insurance	372	1.03 (0.83-1.30)	0.768	
	Medicare or Other Government Insurance	34109	1.12 (0.96-1.30)	0.150	
	Insurance Status Unknown	413	-	-	
Median Income Quartiles 2000	< $30,000	4506	1.29 (1.23-1.36)	**<.001**	**<.001**
	$30,000 - $35,999	7066	1.22 (1.17-1.28)	**<.001**	
	$36,000 - $45,999	10630	1.19 (1.14-1.23)	**<.001**	
	$46,000 +	15205	-	-	
Percent No High School Degree Quartiles 2000	>=29%	5463	1.20 (1.15-1.26)	**<.001**	**<.001**
	20-28.9%	8737	1.19 (1.14-1.24)	**<.001**	
	14-19.9%	9556	1.16 (1.12-1.21)	**<.001**	
	< 14%	13644	-	-	

			Overall Survival (Months)		

Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Urban/Rural 2003	Metro	30310	0.85 (0.77-0.94)	**0.001**	**<.001**
	Urban	6271	1.01 (0.91-1.12)	0.889	
	Rural	914	-	-	
Charlson-Deyo Score	0	18737	0.66 (0.64-0.69)	**<.001**	**<.001**
	1	13971	0.77 (0.74-0.80)	**<.001**	
	2+	6191	-	-	
Year of Diagnosis	2004	2012	1.08 (1.00-1.17)	0.063	**0.002**
	2005	2342	1.12 (1.03-1.21)	**0.007**	
	2006	2475	1.12 (1.04-1.21)	**0.004**	
	2007	3005	1.17 (1.08-1.26)	**<.001**	
	2008	4762	1.07 (0.99-1.15)	0.069	
	2009	5620	1.10 (1.03-1.18)	**0.008**	
	2010	6043	1.05 (0.97-1.12)	0.220	
	2011	6240	1.04 (0.96-1.11)	0.362	
	2012	6400	-	-	

			Overall Survival (Months)		

Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Primary Site	C340 - Main Bronchus	100	1.28 (0.95-1.71)	0.104	**<.001**
	C341 - Upper lobe, Lung	22440	0.79 (0.70-0.88)	**<.001**	
	C342 - Middle lobe, Lung	2051	0.73 (0.63-0.83)	**<.001**	
	C343 - Lower lobe, Lung	13399	0.82 (0.73-0.93)	**0.001**	
	C348 - Overlapping lesion of lung	333	1.00 (0.83-1.21)	0.993	
	C349 - Lung, NOS	576	-	-	
Laterality	Origin of primary is bilateral,midline or unknown	133	1.55 (1.23-1.96)	**<.001**	**<.001**
	Origin of primary is right	22538	1.00 (0.97-1.03)	0.969	
	Origin of primary is left	16228	-	-	
Histology	Others	12448	0.94 (0.90-0.97)	**0.001**	**<.001**
	Squamous cell carcinoma of lung	12534	1.32 (1.27-1.37)	**<.001**	
	Adenocarcinoma of lung	13917	-	-	

			Overall Survival (Months)		

Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Grade	Well differentiated, differentiated, NOS	6492	0.49 (0.46-0.53)	**<.001**	**<.001**
	Moderately differentiated, moderately well differentiated, intermediate differentiation	16166	0.76 (0.72-0.80)	**<.001**	
	Poorly differentiated	11424	0.98 (0.93-1.03)	0.467	
	Undifferentiated, anaplastic	581	1.04 (0.93-1.17)	0.492	
	Cell type not determined, not stated or not applicable, unknown primaries, high grade dysplasia	4236	-	-	
AJCC Analytic Stage Group	AJCC Stage Group Unknown	280	1.27 (1.07-1.50)	**0.005**	**<.001**
	Stage II	3208	1.68 (1.59-1.77)	**<.001**	
	Stage 0, Stage 1 or Occult(lung only)	35411	-	-	
Chemotherapy at any CoC	Unknown	1051	0.75 (0.68-0.83)	**<.001**	**<.001**

			Overall Survival (Months)		

Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Facility	Chemotherapy Administered	2227	0.86 (0.80-0.92)	**<.001**	
	No Chemothrapy	35621	-	-	
Great Circle Distance in 10miles Unit		38255	1.00 (1.00-1.00)	0.084	-
Tumor Size in cm Unit		38539	1.03 (1.03-1.03)	**<.001**	-

Table 7: Multivariate association with overall survival by treatment group

		Overall Survival (Months)		

Covariate	**Level**	**Hazard Ratio**	**HR P-value**	**Type3 P-value**
Treatment Group (SBRT/Surgery)	SBRT	2.04 (1.92-2.16)	**<.001**	**<.001**
	Surgery	-	-	

		Overall Survival (Months)		

Covariate	**Level**	**Hazard Ratio**	**HR P-value**	**Type3 P-value**
Facility Type	Community Cancer Program/Other	1.19 (1.09-1.30)	**<.001**	**<.001**
	Comprehensive Community Cancer Program	1.06 (0.99-1.14)	0.077	
	Academic/Research Program	0.95 (0.89-1.02)	0.184	
	Integrated Network Cancer Program	-	-	
Facility Location	Northeast	1.02 (0.96-1.08)	0.569	**<.001**
	South	1.06 (1.00-1.11)	**0.040**	
	Midwest	1.12 (1.06-1.18)	**<.001**	
	West	-	-	
Age at Diagnosis (Categorical)	70-74	0.63 (0.60-0.66)	**<.001**	**<.001**

		Overall Survival (Months)		

Covariate	**Level**	**Hazard Ratio**	**HR P-value**	**Type3 P-value**
	75-79	0.80 (0.77-0.83)	**<.001**	
	80+	-	-	
Sex	Male	1.33 (1.29-1.38)	**<.001**	**<.001**
	Female	-	-	
Race	Others	0.86 (0.78-0.95)	**0.004**	**0.010**
	Black	0.96 (0.89-1.03)	0.259	
	White	-	-	
Median Income Quartiles 2000	< $30,000	1.21 (1.14-1.28)	**<.001**	**<.001**
	$30,000 - $35,999	1.12 (1.06-1.17)	**<.001**	
	$36,000 - $45,999	1.13 (1.08-1.17)	**<.001**	
	$46,000 +	-	-	
Urban/Rural 2003	Metro	1.03 (0.93-1.14)	0.600	**0.027**
	Urban	1.09 (0.98-1.21)	0.109	

		Overall Survival (Months)		

Covariate	**Level**	**Hazard Ratio**	**HR P-value**	**Type3 P-value**
	Rural	-	-	
Charlson-Deyo Score	0	0.68 (0.65-0.71)	**<.001**	**<.001**
	1	0.80 (0.76-0.83)	**<.001**	
	2+	-	-	
Year of Diagnosis	2004	1.22 (1.12-1.33)	**<.001**	**<.001**
	2005	1.24 (1.14-1.35)	**<.001**	
	2006	1.26 (1.16-1.36)	**<.001**	
	2007	1.27 (1.17-1.37)	**<.001**	
	2008	1.14 (1.06-1.23)	**<.001**	
	2009	1.19 (1.10-1.28)	**<.001**	
	2010	1.07 (0.99-1.15)	0.074	
	2011	1.01 (0.94-1.09)	0.765	
	2012	-	-	
Primary Site	C340 - Main Bronchus	0.59 (0.33-1.03)	0.065	**<.001**

		Overall Survival (Months)		

Covariate	**Level**	**Hazard Ratio**	**HR P-value**	**Type3 P-value**
	C341 - Upper lobe, Lung	0.84 (0.74-0.95)	**0.005**	
	C342 - Middle lobe, Lung	0.79 (0.69-0.91)	**0.001**	
	C343 - Lower lobe, Lung	0.88 (0.78-1.00)	**0.048**	
	C348 - Overlapping lesion of lung	1.05 (0.86-1.28)	0.652	
	C349 - Lung, NOS	-	-	
Laterality	Origin of primary is bilateral,midline or unknown	1.75 (1.10-2.79)	**0.017**	**0.004**
	Origin of primary is right	1.04 (1.01-1.08)	**0.015**	
	Origin of primary is left	-	-	
Histology	Others	0.95 (0.91-0.99)	**0.025**	**<.001**
	Squamous cell carcinoma of lung	1.10 (1.06-1.14)	**<.001**	
	Adenocarcinoma of lung	-	-	

		Overall Survival (Months)		

Covariate	**Level**	**Hazard Ratio**	**HR P-value**	**Type3 P-value**
Grade	Well differentiated, differentiated, NOS	0.67 (0.62-0.72)	**<.001**	**<.001**
	Moderately differentiated, moderately well differentiated, intermediate differentiation	0.95 (0.89-1.01)	0.095	
	Poorly differentiated	1.14 (1.07-1.21)	**<.001**	
	Undifferentiated, anaplastic	1.33 (1.18-1.51)	**<.001**	
	Cell type not determined, not stated or not applicable, unknown primaries, high grade dysplasia	-	-	
AJCC Analytic Stage Group	AJCC Stage Group Unknown	1.14 (0.94-1.38)	0.174	**<.001**
	Stage II	1.64 (1.55-1.74)	**<.001**	

		Overall Survival (Months)		
Covariate	**Level**	**Hazard Ratio**	**HR P-value**	**Type3 P-value**
	Stage 0, Stage 1 or Occult(lung only)	-	-	
Chemotherapy at any CoC Facility	Unknown	0.74 (0.66-0.82)	**<.001**	**<.001**
	Chemotherapy Administered	0.77 (0.72-0.83)	**<.001**	
	No Chemothrapy	-	-	
Tumor Size in cm Unit		1.03 (1.02-1.03)	**<.001**	**<.001**

* Number of observations in the original data set = 38899. Number of observations used = 36352.

** Backward selection with an alpha level of removal of .15 was used. The following variables were removed from the model: Primary Payor, Percent No High School Degree Quartiles 2000, and Great Circle Distance in 10miles Unit.

Table 8: Balance check after 1:1 matching of SBRT and Surgery

Covariate	Level	Statistics	Treatment Group (SBRT/Surgery) SBRT N=3202	Surgery N=3202	Parametric P-value*	Standardized Difference
Facility Type	Community Cancer Program/Other	N (Col%)	62 (1.94)	55 (1.72)	0.574	0.016
	Comprehensive Community Cancer Program	N (Col%)	1400 (43.72)	1358 (42.41)		0.026
	Academic/Research Program	N (Col%)	1463 (45.69)	1493 (46.63)		0.019
	Integrated Network Cancer Program	N (Col%)	277 (8.65)	296 (9.24)		0.021
Facility Location	Northeast	N (Col%)	587 (18.33)	566 (17.68)	0.902	0.017
	South	N (Col%)	1101 (34.38)	1099 (34.32)		0.001
	Midwest	N (Col%)	1066 (33.29)	1079 (33.7)		0.009
	West	N (Col%)	448 (13.99)	458 (14.3)		0.009
Sex	Male	N (Col%)	1367 (42.69)	1362 (42.54)	0.899	0.003

Covariate	Level	Statistics	Treatment Group (SBRT/Surgery) SBRT N=3202	Surgery N=3202	Parametric P-value*	Standardized Difference
	Female	N (Col%)	1835 (57.31)	1840 (57.46)		0.003
Race	White	N (Col%)	2893 (90.35)	2904 (90.69)	0.876	0.012
	Black	N (Col%)	214 (6.68)	204 (6.37)		0.013
	Others	N (Col%)	95 (2.97)	94 (2.94)		0.002
Primary Payor	No Insured	N (Col%)	9 (0.28)	8 (0.25)	0.680	0.006
	Private Insurance	N (Col%)	259 (8.09)	237 (7.4)		0.026
	Medicaid Insurance	N (Col%)	37 (1.16)	45 (1.41)		0.022
	Medicare or Other Government Insurance	N (Col%)	2862 (89.38)	2871 (89.66)		0.009
	Insurance Status Unknown	N (Col%)	35 (1.09)	41 (1.28)		0.017
Median Income Quartiles 2000	< $30,000	N (Col%)	397 (12.4)	405 (12.65)	0.946	0.008
	$30,000 - $35,999	N (Col%)	647 (20.21)	661 (20.64)		0.011
	$36,000 - $45,999	N (Col%)	931 (29.08)	916 (28.61)		0.010

Covariate	Level	Statistics	Treatment Group (SBRT/Surgery) SBRT N=3202	Surgery N=3202	Parametric P-value*	Standardized Difference
	$46,000 +	N (Col%)	1227 (38.32)	1220 (38.1)		0.004
Percent No High School	>=29%	N (Col%)	403 (12.59)	387 (12.09)	0.757	0.015
Degree Quartiles 2000	20-28.9%	N (Col%)	759 (23.7)	772 (24.11)		0.010
	14-19.9%	N (Col%)	877 (27.39)	851 (26.58)		0.018
	< 14%	N (Col%)	1163 (36.32)	1192 (37.23)		0.019
Urban/Rural 2003	Metro	N (Col%)	2596 (81.07)	2597 (81.11)	0.911	0.001
	Urban	N (Col%)	526 (16.43)	520 (16.24)		0.005
	Rural	N (Col%)	80 (2.5)	85 (2.65)		0.010
Year of Diagnosis	2004	N (Col%)	27 (0.84)	38 (1.19)	0.678	0.034
	2005	N (Col%)	47 (1.47)	46 (1.44)		0.003
	2006	N (Col%)	87 (2.72)	93 (2.9)		0.011
	2007	N (Col%)	216 (6.75)	203 (6.34)		0.016
	2008	N (Col%)	275 (8.59)	305 (9.53)		0.033

Covariate	Level	Statistics	Treatment Group (SBRT/Surgery) SBRT N=3202	Surgery N=3202	Parametric P-value*	Standardized Difference
	2009	N (Col%)	457 (14.27)	462 (14.43)		0.004
	2010	N (Col%)	555 (17.33)	576 (17.99)		0.017
	2011	N (Col%)	694 (21.67)	677 (21.14)		0.013
	2012	N (Col%)	844 (26.36)	802 (25.05)		0.030
Primary Site	C340 - Main Bronchus	N (Col%)	11 (0.34)	10 (0.31)	0.701	0.005
	C341 - Upper lobe, Lung	N (Col%)	1910 (59.65)	1892 (59.09)		0.011
	C342 - Middle lobe, Lung	N (Col%)	133 (4.15)	117 (3.65)		0.026
	C343 - Lower lobe, Lung	N (Col%)	1062 (33.17)	1103 (34.45)		0.027
	C348 - Overlapping lesion of lung	N (Col%)	6 (0.19)	3 (0.09)		0.025
	C349 - Lung, NOS	N (Col%)	80 (2.5)	77 (2.4)		0.006
Laterality	Origin of primary is bilateral,midline or unknown	N (Col%)	11 (0.34)	10 (0.31)	0.213	0.005
	Origin of primary is right	N (Col%)	1753 (54.75)	1684 (52.59)		0.043
	Origin of primary is left	N (Col%)	1438 (44.91)	1508 (47.1)		0.044

Covariate	Level	Statistics	Treatment Group (SBRT/Surgery) SBRT N=3202	Surgery N=3202	Parametric P-value*	Standardized Difference
Histology	Others	N (Col%)	877 (27.39)	915 (28.58)	0.571	0.026
	Squamous cell carcinoma of lung	N (Col%)	1120 (34.98)	1101 (34.38)		0.012
	Adenocarcinoma of lung	N (Col%)	1205 (37.63)	1186 (37.04)		0.012
AJCC Analytic Stage Group	Stage 0, Stage 1 or Occult(lung only)	N (Col%)	3038 (94.88)	3021 (94.35)	0.640	0.024
	Stage II	N (Col%)	150 (4.68)	166 (5.18)		0.023
	AJCC Stage Group Unknown	N (Col%)	14 (0.44)	15 (0.47)		0.005
Age at Diagnosis		Mean (Std)	78.76 (5.46)	78.82 (5.01)	0.673	0.011
Great Circle Distance in 10miles Unit		Mean (Std)	2.99 (8.43)	3 (8.86)	0.964	0.001

Covariate	Level	Statistics	Treatment Group (SBRT/Surgery) SBRT N=3202	Surgery N=3202	Parametric P-value*	Standardized Difference
Tumor Size in cm Unit		Mean (Std)	2.63 (2.38)	2.68 (2.13)	0.343	0.024

* The parametric p value is calculated by ANOVA for numerical covariates and Chi-Square test for categorical covariates.

Table 9: Association with overall survival in matched samples

Covariate	Level	N	Overall Survival (Months) Hazard Ratio (95% CI)	HR P-value	Log-rank P-value
Treatment Group (SBRT/Surgery)	SBRT	3202	2.00 (1.86-2.16)	**<.001**	**<.001**
	Surgery	3202	-	-	

Table 10: Overall survival for treatment group in matched sample

Treatment Group (SBRT/Surgery)	**No. of Subject**	**Event**	**Censored**	**Median Survival (95% CI)**	**12 Months Survival**	**60 Months Survival**
SBRT	3202	1759 (55%)	1443 (45%)	33.3 (31.8, 35)	81.4% (80.0%, 82.7%)	24.9% (22.5%, 27.3%)
Surgery	3202	1106 (35%)	2096 (65%)	63.6 (59.3, 67.2)	88.4% (87.2%, 89.5%)	52.0% (49.4%, 54.5%)

Table 11: Multivariate Association with Overall Survival by Treatment interaction with Age

		Overall Survival (Months)		
Covariate	**Level**	**Hazard Ratio**	**HR P-value**	**Type3 P-value**
Comparisons Stratified by Age at Diagnosis (Categorical):		-	-	**<.001**
70-74	SBRT vs. Surgery	2.35 (2.12-2.60)	**<.001**	-
75-79	SBRT vs. Surgery	2.23 (2.03-2.46)	**<.001**	-

		Overall Survival (Months)		
			HR	Type3
Covariate	Level	Hazard Ratio	P-value	P-value
80+	SBRT vs. Surgery	1.78 (1.64-1.93)	<.001	-

* Number of observations in the original data set = 38899. Number of observations used = 36352.

** Backward selection with an alpha level of removal of .20 was used. The following variables were removed from the model: Primary Payor, Percent No High School Degree Quartiles 2000, and Great Circle Distance in 10miles Unit.

*** The estimated stratified treatment effect was controlled by: Charlson-Deyo Score, Facility Location, Facility Type, Grade, Histology, Laterality, Median Income Quartiles 2000, Primary Site, Race, Chemotherapy at any CoC Facility, Sex, Treatment Group (SBRT/Surgery), Urban/Rural 2003, Year of Diagnosis, AJCC Analytic Stage Group, and Tumor Size in cm Unit

Table 12: Multivariate Association with Overall Survival by Treatment interaction with Charlson-Deyo Score

		Overall Survival (Months)		
Covariate	**Level**	**Hazard Ratio**	**HR P-value**	**Type3 P-value**
Comparisons Stratified by Charlson-Deyo Score		-	-	**0.006**
0	SBRT vs. Surgery	2.18 (2.03-2.35)	**<.001**	-
1	SBRT vs. Surgery	1.80 (1.63-2.00)	**<.001**	-
2+	SBRT vs. Surgery	1.99 (1.76-2.25)	**<.001**	-

* Number of observations in the original data set = 38899. Number of observations used = 36352.

** Backward selection with an alpha level of removal of .20 was used. The following variables were removed from the model: Primary Payor, Percent No High School Degree Quartiles 2000, and Great Circle Distance in 10miles Unit.

*** The estimated stratified treatment effect was controlled by: Facility Location, Facility Type, Grade, Histology, Laterality, Median Income Quartiles 2000, Primary Site, Race, Chemotherapy at any CoC Facility, Sex, Treatment Group (SBRT/Surgery), Urban/Rural 2003, Year of Diagnosis, Age at Diagnosis (Categorical), AJCC Analytic Stage Group, and Tumor Size in cm Unit

Table 13: Multivariate Association with Overall Survival by Treatment interaction with Tumor Stage

		Overall Survival (Months)		
Covariate	**Level**	**Hazard Ratio**	**HR P-value**	**Type3 P-value**
Comparisons Stratified by AJCC		-	-	0.251
Analytic Stage Group:				
Stage 0, Stage 1 or Occult	SBRT vs. Surgery	2.06 (1.94-2.19)	**<.001**	-
Stage II	SBRT vs. Surgery	1.71 (1.37-2.13)	**<.001**	-

* Number of observations in the original data set = 38899. Number of observations used = 36352.

** Backward selection with an alpha level of removal of .20 was used. The following variables were removed from the model: Primary Payor, Percent No High School Degree Quartiles 2000, and Great Circle Distance in 10miles Unit.

*** The estimated stratified treatment effect was controlled by: Charlson-Deyo Score, Facility Location, Facility Type, Grade, Histology, Laterality, Median Income Quartiles 2000, Primary Site, Race, Chemotherapy at any CoC Facility, Sex, Treatment Group (SBRT/Surgery), Urban/Rural 2003, Year of Diagnosis, Age at Diagnosis (Categorical), and Tumor Size in cm Unit

Table 14: Overall survival for treatment type SBRT vs pneumonectomy

Treatment Type	No. of Subject	Event	Censored	Median Survival (95% CI)	12 Month Survival	60 Month Survival
SBRT	3466	1908 (55%)	1558 (45%)	33.4 (31.8, 35.1)	81.2% (79.9%, 82.5%)	24.9% (22.7%, 27.3%)
Pneumonectomy	475	283 (60%)	192 (40%)	39.3 (29.4, 46.1)	67.3% (62.9%, 71.4%)	38.7% (33.6%, 43.7%)

Table 15: Overall survival for treatment type SBRT vs wedge and segmental resection

Treatment Type	No. of Subject	Event	Censored	Median Survival (95% CI)	12 Month Survival	60 Month Survival
SBRT	3466	1908 (55%)	1558 (45%)	33.4 (31.8, 35.1)	81.2% (79.9%, 82.5%)	24.9% (22.7%, 27.3%)
Wedge and Segmental Resection	7858	3593 (46%)	4265 (54%)	55.2 (53.4, 56.8)	87.5% (86.7%, 88.2%)	46.2% (44.8%, 47.7%)

Table 16: Overall survival for treatment type SBRT vs lobectomy

Treatment Type	No. of Subject	Event	Censored	Median Survival (95% CI)	12 Month Survival	60 Month Survival
SBRT	3466	1908 (55%)	1558 (45%)	33.4 (31.8, 35.1)	81.2% (79.9%, 82.5%)	24.9% (22.7%, 27.3%)
Lobectomy	26212	10024 (38%)	16188 (62%)	71.5 (70.1, 73.4)	88.3% (87.9%, 88.6%)	56.8% (56.0%, 57.5%)

Table 17: Overall Survival by Treatment SBRT vs. Pneumonectomy in matched samples

Treatment Type	No. of Subject	Event	Censored	Median Survival (95% CI)	12 Months Survival	60 Months Survival
Pneumonectomy	244	124 (51%)	120 (49%)	48.6 (37.2, 63.3)	71.1% (65.0%, 76.4%)	45.3% (37.8%, 52.5%)
SBRT	244	177 (73%)	67 (27%)	25.9 (21.8, 34.1)	77.6% (71.8%, 82.4%)	18.7% (13.0%, 25.2%)

Table 18: Linear hypotheses testing results

Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
Treatment*os	68.7396	1	<.0001

Table 19: Analysis of Maximum Likelihood Estimates

Parameter	Parameter Estimate	Pr > ChiSq	Hazard Ratio	Label
TRT	-0.47164	**<.0001**	0.624	SBRT/Pneu.
TRTOS	0.03238	**<.0001**	1.033	Trt*os
FACILITY_TYPE_CD	-0.07661	0.3742	0.926	Facility Type Academic/Research Program
FACILITY_TYPE_CD	0.1912	0.2321	1.211	Facility Type Community Cancer Program/Other
FACILITY_TYPE_CD	-0.02951	0.7235	0.971	Facility Type Comprehensive Community Cancer Program
FACILITY_LOCATION_CD	0.02848	0.7021	1.029	Facility Location Midwest
FACILITY_LOCATION_CD	0.17209	0.0454	1.188	Facility Location Northeast
FACILITY_LOCATION_CD	0.06956	0.3521	1.072	Facility Location South

SEX	-0.20146	<.0001	0.818	Sex Female
RACE	-0.25672	0.0111	0.774	Race Black
RACE	0.004	0.978	1.004	Race Others
INSURANCE_STATUS	0.22652	0.2982	1.254	Primary Payor Insurance Status Unknown
INSURANCE_STATUS	-0.26046	0.2972	0.771	Primary Payor Medicaid Insurance
INSURANCE_STATUS	0.13874	0.0848	1.149	Primary Payor Medicare or Other Government Insurance
INSURANCE_STATUS	-0.11591	0.7824	0.891	Primary Payor No Insured
MED_INC_QUAR_00	-0.11159	0.2006	0.894	Median Income Quartiles 2000 $30,000 - $35,999
MED_INC_QUAR_00	-0.00957	0.9144	0.99	Median Income Quartiles 2000 $36,000 - $45,999
MED_INC_QUAR_00	-0.20289	0.039	0.816	Median Income Quartiles 2000 $46,000 +
NO_HSD_QUAR_00	-0.1291	0.1506	0.879	Percent No High School Degree Quartiles 2000 14-19.9%

NO_HSD_QUAR_00	-0.09556	0.2508	0.909	Percent No High School Degree Quartiles 2000 20-28.9%
NO_HSD_QUAR_00	-0.05651	0.5573	0.945	Percent No High School Degree Quartiles 2000 < 14%
UR_CD_03	0.11324	0.0784	1.12	Urban/Rural 2003 Metro
UR_CD_03	0.14257	0.3386	1.153	Urban/Rural 2003 Rural
CDCC_TOTAL	-0.28195	<.0001	0.754	Charlson-Deyo Score 0
CDCC_TOTAL	-0.30211	<.0001	0.739	Charlson-Deyo Score 1
YEAR_OF_DIAGNOSIS	0.20697	0.1762	1.23	Year of Diagnosis 2004
YEAR_OF_DIAGNOSIS	0.12455	0.3431	1.133	Year of Diagnosis 2005
YEAR_OF_DIAGNOSIS	0.09482	0.4274	1.099	Year of Diagnosis 2006
YEAR_OF_DIAGNOSIS	0.09381	0.3264	1.098	Year of Diagnosis 2007
YEAR_OF_DIAGNOSIS	0.03101	0.738	1.031	Year of Diagnosis 2008
YEAR_OF_DIAGNOSIS	0.11313	0.1716	1.12	Year of Diagnosis 2009
YEAR_OF_DIAGNOSIS	0.0948	0.2426	1.099	Year of Diagnosis 2010
YEAR_OF_DIAGNOSIS	0.05994	0.4509	1.062	Year of Diagnosis 2011
PRIMARY_SITE	0.14263	0.5716	1.153	Primary Site C340 - Main Bronchus
PRIMARY_SITE	-0.0225	0.8665	0.978	Primary Site C341 -

				Upper lobe, Lung
PRIMARY_SITE	0.17966	0.2722	1.197	Primary Site C342 - Middle lobe, Lung
PRIMARY_SITE	0.04505	0.7415	1.046	Primary Site C343 - Lower lobe, Lung
PRIMARY_SITE	-0.02765	0.9142	0.973	Primary Site C348 - Overlapping lesion of lung
LATERALITY	-0.04106	0.3721	0.96	Laterality Origin of primary is left
HISTOLOGY	-0.16946	0.0021	0.844	Histology Adenocarcinoma of lung
HISTOLOGY	-0.07989	0.1636	0.923	Histology Others
GRADE	0.34223	0.0002	1.408	Grade Cell type not determined, not stated or not applicable, unknown primaries, high grade dysplasia
GRADE	0.3369	0.0008	1.401	Grade Moderately differentiated, moderately well

				differentiated, intermediate differentiation
GRADE	0.43237	<.0001	1.541	Grade Poorly differentiated
GRADE	0.60206	0.0106	1.826	Grade Undifferentiated, anaplastic
ANA_STAGE	0.00752	0.9772	1.008	AJCC Analytic Stage Group AJCC Stage Group Unknown
ANA_STAGE	-0.34103	0.0001	0.711	AJCC Analytic Stage Group Stage 0, Stage 1 or Occult(lung only)
RX_SUMM_CHEMO	0.03317	0.8685	1.034	Chemotherapy at any CoC Facility Chemotherapy Administered
RX_SUMM_CHEMO	0.36186	0.0192	1.436	Chemotherapy at any CoC Facility No Chemothrapy

Table 20: Hazard ratio between SBRT and pneumonectomy against time

Overall Survival (Month)	HR (95% CI)
6	0.75778 (0.50276, 0.8119)
12	0.92028 (0.58647, 1.0333)
36	2.00179 (1.08588, 2.7105)
60	4.35428 (2.73582, 11.5152)

Figure 1:Survival Curves by Treatment SBRT vs. Surgery

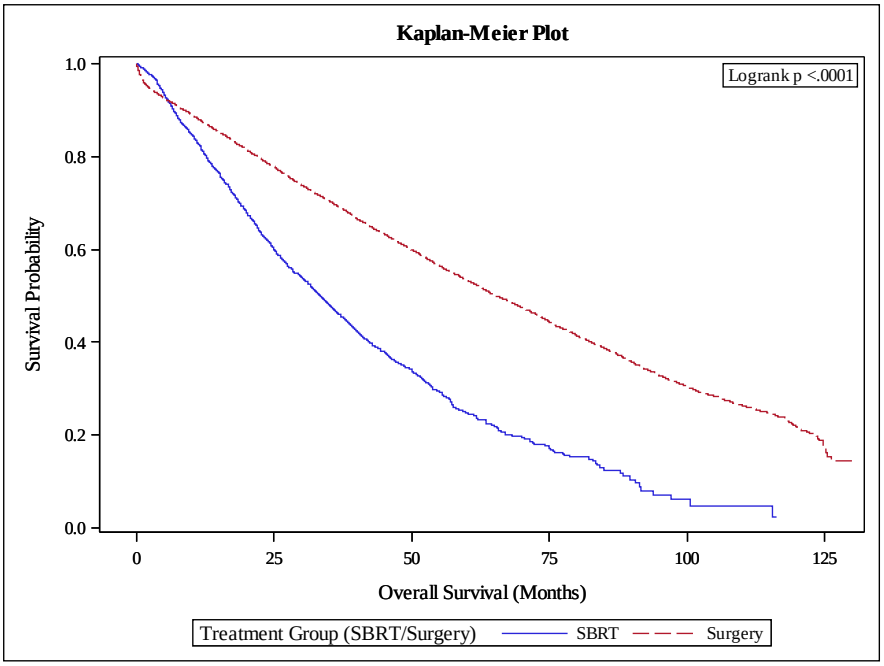

Figure 2:Kaplan-Meier curves in mached samples

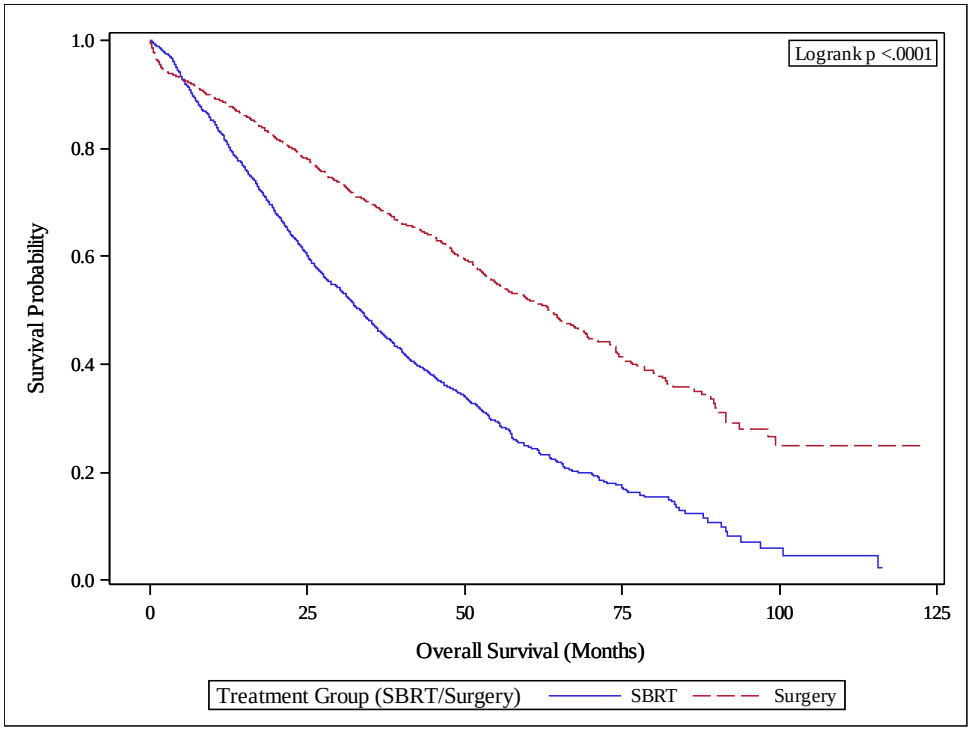

Figure 3:Survival curves by SBRT versus Pneumonectomy

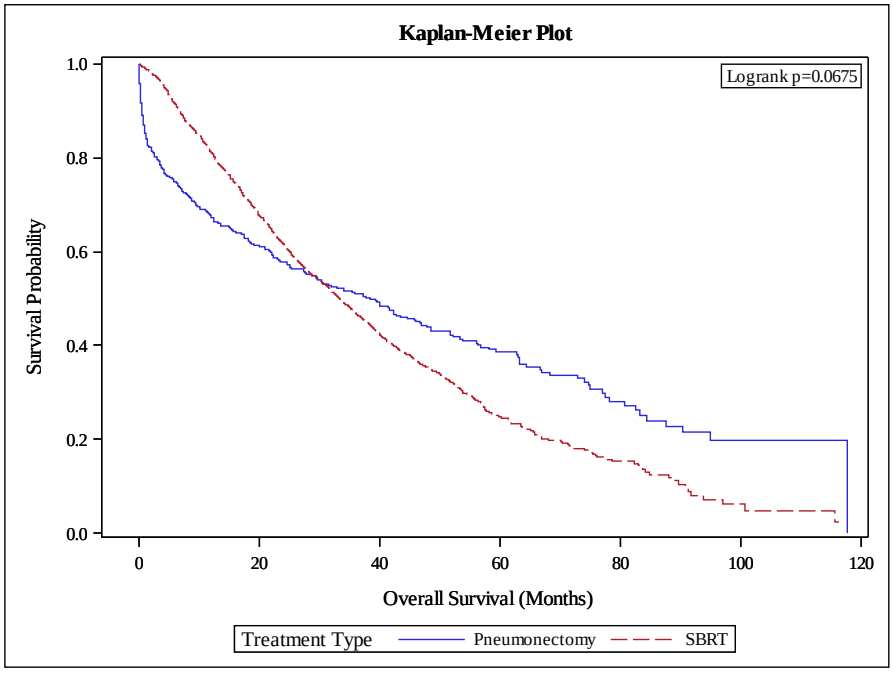

Figure 4: Survival curves by SBRT versus wedge and segmental resection

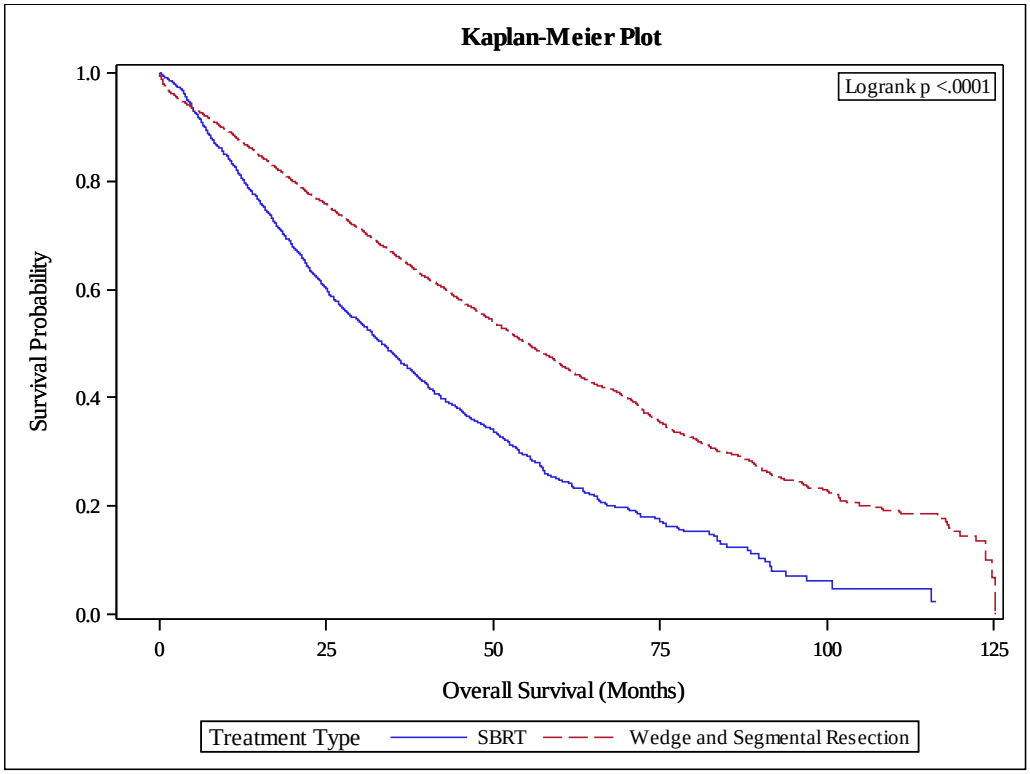

Figure 5: Survival curves by SBRT versus lobectomy

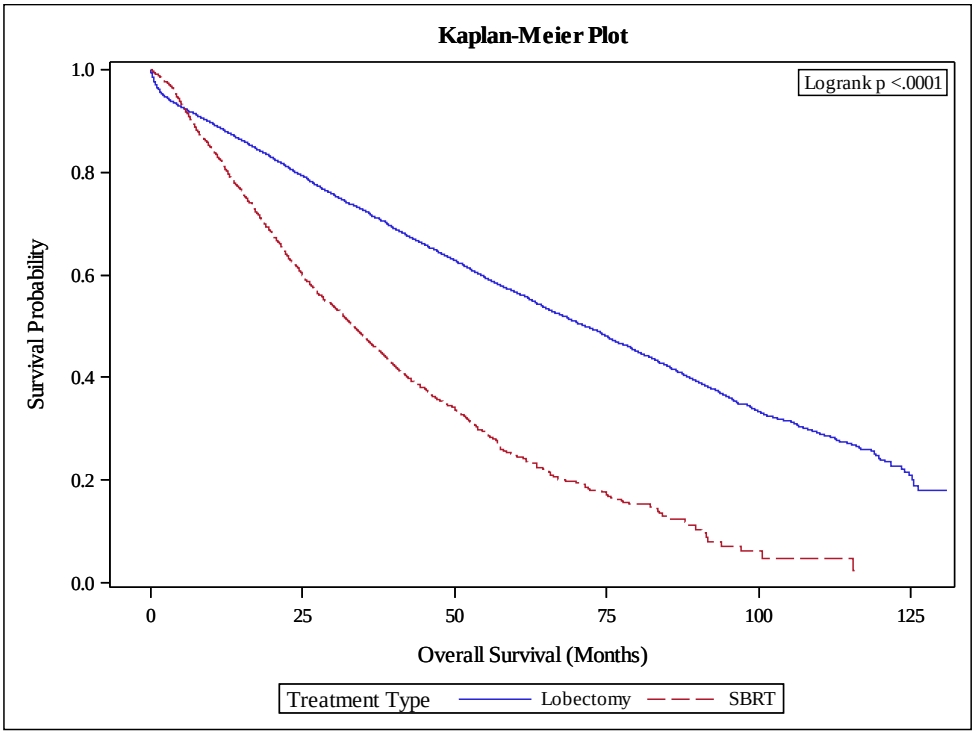

Figure 6:Survival Curves by Treatment SBRT vs. Pneumonectomy

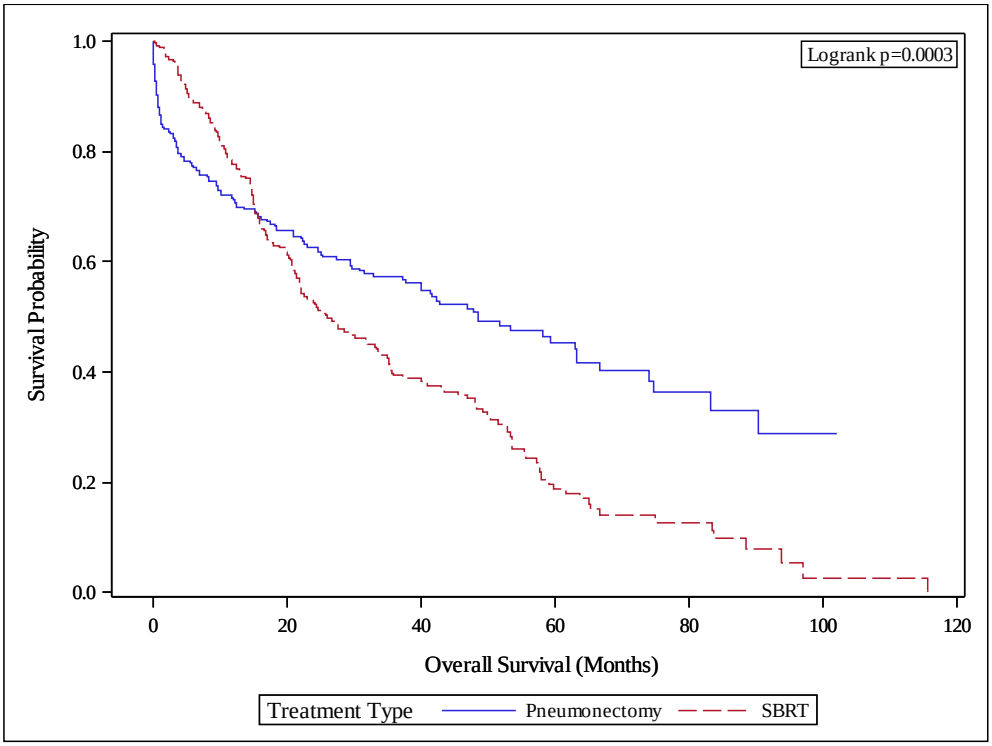

Figure 7: Plot of Schoenfeld residual for treatment type against overall survival

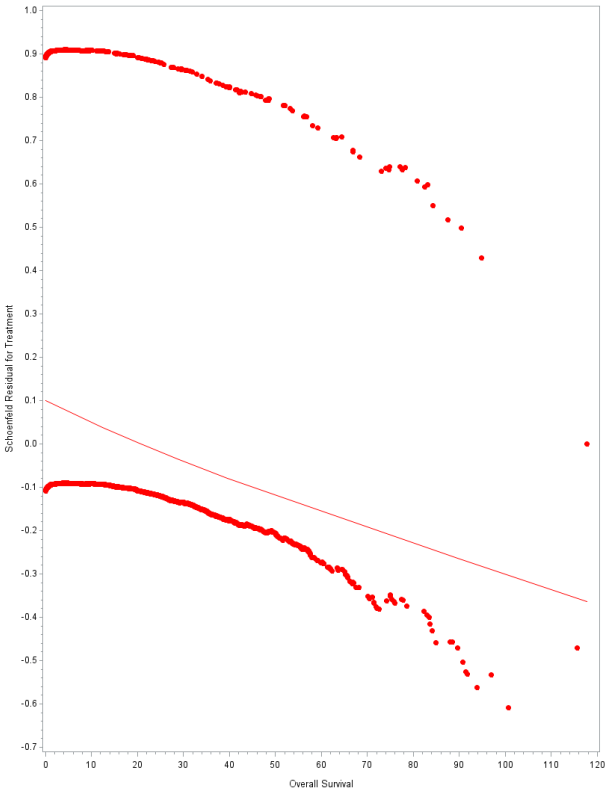

Figure 8:Hazard ratio changes over time between SBRT vs Pneumonectomy