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:

_____ _____

Jingru Zhou Date

The Effect of Radiation on Overall Survival among Metastatic Prostate Cancer Patients treated by ADT as the first-line treatment

By

Jingru Zhou
Master of Public Health

Biostatistics and Bioinformatics

Yuan Liu, PhD
Thesis Advisor

Xiangqin Cui, PhD
Reader

The Effect of Radiation on Overall Survival among Metastatic Prostate Cancer Patients treated by ADT as the first-line treatment

By

Jingru Zhou

Bachelor of Science
Hunan University
2018

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 Biostatistics and Bioinformatics
2020

Abstract

The Effect of Radiation on Overall Survival among Metastatic Prostate Cancer Patients treated by ADT as the first-line treatment

By Jingru Zhou

Background: STAMPEDE (Systemic Therapy in Advancing or Metastatic Prostate Cancer: Evaluation of Drug efficacy) is a large clinical trial that conducted to evaluate nova treatment methods for patients suffered from high-risk prostate cancer. Our goal is to identify a subgroup of these patients who may benefit from radiation therapy in the National Cancer Data Base (NCDB) to avoid unwarranted use of radiotherapy in patients.

Methods: Patients with newly diagnosed metastatic prostate cancer were identified according to STAMPEDE trial exclusion and inclusion criteria. We compared the overall survival (OS) of patients who received radiation therapy and those who did not using Kaplan-Meier method with log-rank test, Cox proportional hazard regression model, and propensity score matching. Subgroup analyses were conducted by fitting Cox proportional hazard models that include radiation treatment interacting with race, age, facility type, Charlson-Deyo score and metastasis. We built several models with each model containing one interaction term.

Results: A sample of 35177 patients were analyzed. From multivariate statistics for OS, radiation therapy showed a significant association with improved overall survival [HR 0.46, CI [0.43, 0.49], P <0.001]. The effect of radiation on overall survival is different in three patient subgroups (interaction p < 0.001, and they High-Risk group [HR 0.43, CI [0.39-0.47], p <0.001], Node-Positive group [HR 0.59, CI [0.52-0.67], p <0.001], and Metastasis group [HR 0.43, CI [0.39-0.48], p <0.001]). Radioation therapy also improved OS for two subgroups of facility type, Academic / Research Program [HR 0.53, CI [0.48-0.57], p <0.001], Non-Academic/Research Program [HR 0.43, CI [0.40-0.47], p <0.001]. These estimated stratified treatment effects were obtained after controlling for other patient demographics and disease characteristics variables.

Conclusions: This study supports the previous finding that beam radiation has a protective effect on prostate cancer patients. The protective effect is more prominent in High-Risk and Metastasis subpopulation than the Node-Positive group.

The Effect of Radiation on Overall Survival among Metastatic Prostate Cancer Patients treated by ADT as the first-line treatment

By

Jingru Zhou

Bachelor of Science
Hunan University
2018

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 Biostatistics and Bioinformatics
2020

Acknowledgment

My deeply gratitude goes first to my thesis advisor, Dr. Yuan Liu, whose encouragement, guidance and support from the initial to the final level enabled me to establish the comprehension of the subject. I am extremely grateful for your assistance and patience during my research period. I am also thankful to my thesis reader, Dr. Xiangqin Cui for her kindly help and guidance throughout my thesis period. Thanks to my families, my parents, my brother and my boyfriend for their unconditional support and love.

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

Jingru Zhou

Contents

1 Introduction

1.1 Prostate cancer and treatments

Prostate cancer (PCa) is one of the most common malignancies associated with bone metastases. Since 2007, the incidence of metastatic prostate cancer has risen significantly.[1] The American Cancer Society's 2019 estimate of prostate cancer in the United States shows that approximately one in nine people will be diagnosed with prostate cancer in their lifetime. Prostate cancer is the second leading cause of cancer deaths in American men, second only to lung cancer, and 1 of 41 men will die of prostate cancer.[2] Several treatment options have been developed for advanced prostate cancer, which can help slow its spread, prolong life, and control its symptoms. The primary approach for treating advanced prostate cancer is hormone therapy, which can help slow the growth and spread of prostate cancer by reducing the amount of testosterone in the body. The other way was used only to relieve symptoms associated with symptomatic metastatic disease in the past called chemotherapy. At last clinical trials test experimental treatments were introduced, new combinations of drugs or new methods of surgery or radiation therapy. [3]

1.2 Previous clinical trial

Christopher Parker, together with his colleagues, build on the body of work that incorporated within the Systemic Therapy for Advanced or Metastatic Prostate cancer: Evaluation of Drug Efficacy (STAMPEDE) initiative. The study screened the effects of prostate cancer radiation on the metastatic outcomes in men. According to previous findings, prostate cancer

radiotherapy does not improve survival in unselected newly diagnosed patients with metastatic prostate cancer, but in a predetermined subgroup analysis, the overall survival of men with low metastatic burden was indeed improved. The authors conclude that their findings establish a new standard of care that includes prostate radiation in men with oligometastatic cancer, in which case cancer cells in a primary (primary) tumor spread in the body and form a small number of new tumors (metastatic tumors) in one or two other parts of the body. These findings also raise the interest in looking for other optimal subgroups of patients who will benefit from the radiotherapy.

1.3 Study Objective

Prostate radiotherapy is a simple technology, relatively cheap and widely used, can be easily implemented. Research data may change clinical practice. We agree with the conclusion from the authors; however, further confirmation of these findings are required before offering radiation to the prostate in oligometastatic prostate cancer as a new standard for all.[6] In our research, we will use large database of NCDB and rich treatment information to not only verify the results but also find more subgroups that can benefit from radiation therapy. For some patients with specific characteristics, if radiotherapy is added to the treatment plan, their survival rate can be greatly improved, our analysis is meaningful.

The authors also noted that after receiving the chararted trial report, they could conclude the clear clinical significance of the burden of metastasis in prostate cancer patients. The shifting burden is determined by retrospective collection of searchable baseline scans, which

means they cannot cover all patients. In our retrospective survival analysis, the data came from NCDB data, which provided us with a wide range of patient information.

In this project, we analyzed prostate cancer patients with high-risk node-positive metastatic bone disease present at diagnosis from a large cancer registry database in the US, the National Cancer Database (NCDB).[1] The primary objective of this project is to identify a subgroup of these patients who may not benefit from radiation therapy, to avoid unwarranted use of radiotherapy in a substantial proportion of patients.

2 Methods

2.1 Introduction of National Cancer Database

The National Cancer Database (NCDB) is a nationally recognized clinical oncology database. The database is constructed based on data from hospital registration agencies, which come from more than 1,500 Cancer Council (CoC) accredited institutions. NCDB contains data from 21 million cancer patients diagnosed between 1985 and 2013 and is used to study cancer diagnosis, disease treatment, and overall survival factors, recognized as the world's largest clinical registry. NCDB collects approximately 70% of newly diagnosed invasive cancers in the United States each year, with a record of about 940,000 newly diagnosed invasive cancers.[14] These records are used to analyze and track the treatment methods and results of patients with malignant diseases. There are many types of cancer datasets in NCDB, such as non-small cell lung cancer dataset, breast cancer dataset, and prostate dataset. [14]

2.2 NCDB prostate cancer data set

The NCDB Participant User File prostate 2016 was used in this study. 1,535,577 cases with 102 variables are included in the database. The variables are classified into seven groups based on their features. Case key is a unique case identification number assigned to the case in the database. Facility Type provides a general classification of the structural characteristics of each reporting facility.[5][5] Patient Demographics provides patients' baseline characteristics like age, gender, race, insurance status, Medicaid expansion status state group cancer identification,

Charlson/Deyo score, et.al. status, et.al. Cancer Identification group provides serval cancer information, such as years of diagnosis, primary site, laterality, histology, grade, et.al. Stage of Disease indicates which stage of prostate cancer diagnosed by both clinical and pathologic methods. Treatment contains kinds of treatment information, treatment ways, status, start day, duration, et.al. Outcomes covers thirty/ ninety-day mortality, last contact or death, months from Dx, vital status.[5]

2.3 Inclusion and Exclusion

STAMPEDE's overall goal is to evaluate multiple therapeutic strategies in the management of high-risk locally advanced and metastatic hormone-naïve prostate cancer. The purpose is to determine whether it is possible to improve existing treatments by adding new therapies to standard therapies or changing the type of hormone therapy, thereby reducing the side effects of treatment and improving the quality of life (Table 1).

The primary purpose of this study is to mimic the trial design used in STAMPEDE and target a similar study population of the prostate in men with newly – diagnosed metastatic disease. Therefore, the following sample inclusion and exclusion criteria were utilized. Firstly, cases with age above or equal to 40 are included as facility type, and facility location are missing for those with cancer below 40. Secondly, cases with invasive tumor behavior, had no previous cancer diagnosis and with positive histology diagnosis confirmation are included. Thirdly, all cases were treated by hormone therapy within six months after diagnosis and had no surgery or chemotherapy or palliative care. Finally, for cases that were given radiation, it should be given to prostate and or Pelvis, had total radiation dose above 59 Gy, and started between -30

to 183 days relative to hormone therapy started. Last but not least, all cases with a missing outcome were excluded. The sample size of the final analytical data is 35,177. When STAMPEDE criteria were implemented, we obtained three eligible groups of cases, high-risk (High-Risk Newly Diagnosed Non-Metastatic Node-Negative Disease), node-positive (Newly Diagnosed Metastatic or Node-Positive Disease), and with metastasis.

2.4 Statistical Analysis

Statistical analysis was conducted in SAS version 9.4 by using SAS Marcos developed by the Biostatistics & Bioinformatics shared resource at Winship Cancer Institute (BBISR). The significant level was set at 0.05. [23]

2.4.1 Descriptive Statistics for Baseline Covariates

Descriptive statistics for all selected variables in the analytic dataset were generated by SAS macro %DESCRIPTIVE. For the categorical variables, the observation number and percentage of each level were generated. For continuous variable, mean, median, minimum, maximum, standard deviation, and missing quantities were generated. The purposes are to check potential code errors or additional exclusion and to decide how to categorize some variables.

2.4.2 Univariate Association with Radiation for each covariate

The categorical outcomes were generated by SAS macro %UNI_CAT. The observation number and row percent of each level for each covariate among two treatment groups (Radiation Yes vs. No) were reported. For categorical covariates, Chi-square test (parametric p-value) or Fisher's exact test (non-parametric p-value) were conducted on contingency tables. For continuous covariates, the sample size, mean, and median were calculated and ANOVA (parametric p-value) or Kruskal-Wallis test (non-parametric p-value) were conducted for testing. The purposes are to show the association between each covariate and Radiation individually and provides a comparison of baseline heterogeneity among study cohorts.

2.4.3 Univariate association with overall survival

The hazard ratio calculated by Cox proportional hazards regression model was chosen as an indicator. Cox proportional model is the most commonly used one when comparing overall survival between two groups. Univariate survival analysis for each variable in the dataset was generated by SAS macro %UNI_PHREG. The hazard ratio with 95% CI was presented along with the log-rank test p-value. For categorical variables, the reference group will be shown along with the number of observations in each category. The proportional hazard assumption can optionally be checked. Cox regression was utilized in this macro. The purposes are to show the association between each covariate, including study cohort variable and Radiation individually, and provides crude association.

2.4.4 Multivariate logistic regression with Radiation for each covariate

A binary outcome or ordinal outcome using a cumulative logit model can be used. Odds ratio with 95% confidence interval of primary exposure variable controlling for covariates among radiation was generated by SAS macro %LOGREG_SEL.

OS ~ Radiation + Age + Race + Median Income Quartiles 2008-2012 + Percent No High School Degree 2008-2012 + Urban/Rural 2013 + Primary Payor + Year of Diagnosis + Facility Type + Facility Location + Charlson-Deyo Score + Sequence Number + Grade + AJCC T + AJCC N + AJCC M + PSA + Gleason + Months of Hormone Start from Diagnosis

The purpose is to conduct backward selection on a logistic regression 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 when using their selection methods.

2.4.5 Multivariate association with overall survival

The main goal is to estimate the treatment effect (comparison between Radiation Yes vs. No) when holding other covariates consistent. Also called adjusted association that takes the confounding effect under control.

Fitting a multivariable model does not follow strict rules. The basic principle is to control all possible confounding effects. Hazard ratio with a 95% confidence interval of primary exposure variable between treatment group controlling for covariates was generated by SAS macro %PHREG_SEL. The model used here is the same as the one in logistic analysis. Backward selection on a Cox proportional Hazard model using the maximum possible sample

size at each stage of the selection process was conducted in the Marco, only variables with final p-value less than 0.05 would stay in the final model.

2.4.6 Multivariate association with overall survival stratified by covariates of interest

Hazard ratio, along with a 95% confidence interval for each level of interested covariates were generated by SAS macro %PHREG_SEL, which based on Cox proportional hazard regression model. Type 3 p-value was also reported to study if there exists a significantly different hazard ratio for each level of interested covariates. Here we consider race, age, facility type, Charlson-Deyo score, metastasis as covariates of interest. For example, the age model is:

OS ~ Radiation + Age + Race + Percent No High School Degree 2008-2012 + Urban/Rural 2013 + Primary Payor + Year of Diagnosis + Facility Type + Facility Location + Charlson-Deyo Score + Sequence Number + Grade + AJCC T + AJCC N + AJCC M + PSA + Gleason + Months of Hormone Start from Diagnosis + Radiation * Race

2.4.7 Propensity score analysis

The propensity score (PS) method is commonly applied in the public health area. PS analysis is a method utilized to minimize selection bias or confounding effects in observational studies. It is defined as the conditional probability of treatment assignment given the observed baseline covariates. [21] This study is an observational study in which treatment choices are often influenced by subject characteristics. Therefore, when assessing the impact of treatment on overall survival, systematic differences in baseline characteristics of radiation therapy or non-

radiation therapy should be considered. Logistic regression models are used to estimate propensity scores, where the treatment group was the dependent variable, and all covariates were independent variables. After propensity score matching with calipers of width equal to 0.2 implementation, we reconstructed a pseudo study population that meets ignorability assumption with an indication of causality, in which the comparison treatment groups have balanced baseline covariates.[18] We assessed the balance of baseline covariates between subjects in the two treatment groups in the propensity score analysis samples. Then the hazard ratio and Kaplan Meier curve between the two treatments are generated.

All of the processes are implemented through SAS Macro %CALC_PS, %STD_DIFF, and %KM_PLOT. %CALC_PS macro is used to estimate the propensity score (PS) for treatment assignment using the logistic regression model (binary treatment) or multinomial logistic regression model to assess and visualize the common support or the overlap of distribution of propensity score by treatment groups. %STD_DIFF is to check the covariate-balance or distribution of covariates across all levels of a categorical cohort using absolute standardized difference (ASD), perform covariate-balance-check before and after a propensity score adjustment, such as matching or weighting. Also, demonstrate the sample distribution change before and after PS adjustment samples. A hazard ratio with 95% confidence interval will be generated. %KM_PLOT is used to produce Kaplan Meier curves for matched samples and survival rate for each treatment group.

3 Results

3.1 Description of Study Population

Among 35,177 patients, 11,192 (31.8%) didn't receive radiation therapy and 23,985 (68.2%) received beam radiation. There are 13, 618 (38.7%) patients whose age at diagnosis are below or equal 65, and 21,559 (61.3%) whose are above 65. The average age at diagnosis year of the selected patients is 68.93 with a standard deviation of 10.30. The average months of hormone start from diagnosis is 1.10 with standard deviation of 1.12. There are 2,4986 (71.0%) patients are NH-White, 16,791 (19.3%) patients are NH-Black, 866 (2.5%) patients are Asian and 1,951 (5.5%) patients are Hispanic. Most patients are living in metro (28,351, 80.6%), 5,188 (14.7%) living in urban and 808 (2.3%) living in rural. Most of the selected patients are in High risk group (32368, 95.0%), and 1,657 (4.9%) patients are in intermediate risk group (Table 2).

3.2 Association between treatment and each covariate

The %UNICAT macro was used to calculate the distribution of each covariate in the treatment group (Radiation Yes vs. No). The observation number and row percent of each level for each covariate among two treatment groups were reported. If there is a significant association between the covariate and the treatment group, the covariate may influence decisions about treatment choices. Therefore, it may be a potential confounder, which should be controlled in the following cox regression model for survival analysis. We used Chi-square test for

categorical covariates and ANOVA for numerical covariates to evaluate each covariate's univariate association with two cohorts of Radiation Yes vs. No. We found that all of the covariates have a significant association with the treatment group. Patients with age below or equal 65 (40.38%) have a higher proportion to get radiation therapy rather than none radiation than those with age less than 65 (37.94%). NH-white (74.71%), living in metro (79.96%) with Medicare (53.65%) as primary payor patients are more willing to receive radiation therapy. Patients whose Charlson-Deyo Score equals 0 (85.67%), which means no comorbid conditions and grade is poorly or undifferentiated, which means they got severe tumor recorded tend to choose radiation therapy. Patients without metastasis (91.44%), without a spread of cancer to nearby lymph nodes (72.43%), with PSA levels above 20 ng/mL of blood (54.52%), with higher Gleason grade (8-10) (86.97%) are more willing to receive radiation therapy. Patients in the high-risk group (96.59%) tend to receive radiation therapy. MVA logistic regression model further demonstrated multiple factors associated with an increased odds of radiation usage (Table 2).

3.3 Univariate association with overall survival

Among 23,985 patients who didn't receive radiation therapy, 15,068 (63%) died at the end date, and 8,917 (37%) patients were censored. The five-year survival rate of patients without radiation therapy is 29.4%, and ten-year survival rate is 11.1%. Among 11,192 patients who received radiation therapy, 3,319 (30%) patients died at the end date, and 7,873 (70%) patients were censored. Five-year survival rate of patients with radiation therapy is 74.9%, and ten-year survival rate is 44.7%, which are better than patients without radiation therapy. (Table 5)

According to the log-rank test, the p-value is less than 0.001, which means the average overall survival of patients with radiation therapy is significantly better than patients without radiation therapy. Radiation therapy was associated with the improved overall survival (OS) when compared to none radiation therapy [HR 0.26, CI [0.25-0.27], p <0.001], which means at any particular time, 0.26 times as many patients in the radiation therapy group are experiencing an event compared to the none radiation therapy group. We further explore the univariate association with overall survival for each covariate. Except for urban/rural setting, year of diagnosis, all the other variables were significantly associated with OS. (Table 3)

3.4 Multivariate association with overall survival by treatment group

The results from univariate association analysis were confirmed with multivariate association analysis for OS; radiation therapy remained a significant factor associated with increased OS compared to none radiation therapy [HR 0.46, CI [0.44-0.48], p <0.001]. Apart from this, Charlson-Deyo Score above 2 holds [HR 1.73, CI [1.63-1.83], p <0.001] compared to score 0, Patients without metastasis holds [HR 0.53, CI [0.51-0.56], p <0.001] compared to those with metastasis. The backward selection was conducted in the model using the maximum possible sample size at each stage of the selection process instead of 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. Median Income Quartiles 2008-2012 was removed after backward selection. The final analysis model includes non-missing subjects of selected variables whose p-values are less than 0.05. The model is:

OS ~ Radiation + Age + Race + Percent No High School Degree 2008-2012 +

Urban/Rural 2013 + Primary Payor + Year of Diagnosis + Facility Type + Facility Location

+ Charlson-Deyo Score + Sequence Number + Grade + AJCC T + AJCC N + AJCC M + PSA + Gleason + Months of Hormone Start from Diagnosis

Based on the results above, we concluded that radiation therapy brought in long-term benefit in OS to patients compared with none radiation therapy at baseline of other covariates. (Table 3)

3.5 Multivariable interaction with treatment by subgroups

The hazard ratio may differ among each level of covariates of interest between Radiation Yes vs. No. Age, Charlson-Deyo score and metastasis have great impact on survival rate, we were also interested in the difference among race, facility type and patient group. We build several models, each model contains one interaction term and the main effect for other covariates, outcome comparisons stratified by each interaction term controlling for other covariates are shown in Table 4. All hazard ratios are significantly smaller than 1, which means, among each patient group, radiation therapy is significantly superior to none radiation therapy. The hazard ratios between radiation therapy and none radiation therapy are different among three patient groups with type 3 p-value less than 0.001. For patients in High-Risk [HR 0.42, CI [0.39-0.45], $p < 0.001$] and Metastasis group [HR 0.44, CI [0.40-0.49], $p < 0.001$], radiation therapy was a significant predictor on MVA that reduce the risk of death compared with patients in Node-Positive group [HR 0.58, CI [0.52-0.64], $p < 0.001$].

3.6 Propensity score matching analysis

Propensity score (PS) matching analysis attempts to repeat the characteristics of randomized trials to assess the effects of treatment. Based on results from UVA with radiation therapy, UVA, and MVA with OS, all the variables were used in the propensity score estimation model. After calculating PS for each subject, the histograms of PS distribution were generated and shown in Figure 3. The overlap area was acceptable, which may indicate a sufficient foundation to draw the causal inference. Balance check results were summarized in ; all the covariates' balance was achieved by ASD < 0.1, which is much improved than that before matching. shows the KM plot for the propensity score match cohorts, stratified by Radiation Yes vs. No. Among 8,129 patients who didn't receive radiation therapy, 4,493 (55%) died at the end date, and 3,636 (45%) patients were censored. Five-year survival rate of patients with radiation therapy is 37.4%, and ten-year survival rate is 15.3%. Among 8,129 patients who received radiation therapy, 2,464 (30%) patients died at the end date, and 5,665 (70%) patients were censored. Five-year survival rate of patients with radiation therapy is 73.9%, and ten-years survival rate is 44.0%, which is better than none radiation therapy (Table 5). The UVA showed that beam radiation was still significantly associated with improved OS [HR 0.34, CI [0.32, 0.36], P <0.001]. From MVA for OS, beam radiation remained a significant factor associated with increased OS compared to none radiation [HR 0.46, CI [0.43, 0.49], P <0.001] ().

showed the interaction analysis results conducted in the PS matched sample. We find that the hazard ratios between Radiation Yes vs. No is only different among three patient groups and facility type with type 3 p-value less than 0.001. For three levels of patients group, High-Risk group [HR 0.43, CI [0.39-0.47], p <0.001], Node-Positive group [HR 0.59, CI [0.52-0.67], p <0.001], Metastasis group [HR 0.43, CI [0.39-0.48], p <0.001]. For two levels of facility type,

Academic/Research Program [HR 0.53, CI [0.48-0.57], p <0.001], Non-Academic/Research Program [HR 0.43, CI [0.40-0.47], p <0.001]. All hazard ratios are significantly smaller than 1, which means, among each patient group, radiation therapy is significantly inferior to none radiation therapy. Among three patient groups, compared to the Node-Positive group, patients in the High-Risk group and Metastasis group were associated with improved overall survival, which means patients of the High-Risk group and Metastasis group could benefit from radiation therapy most.

4 Discussion

All patient's information was sourced from the 2016 prostate National Cancer Participant User File, include and exclude criteria was established based on study purpose and medical knowledge. The final analytical data contains 35,177 patients after applied the criteria. Therefore, the sample size is large enough to offer sufficient statistical power. The abundant treatment information of patients is another apparent strength, which detailed organized patient characteristics, cancer staging and tumor histological characteristics, type of first-course treatment administered, and outcomes information.[14] It provided us rich information to analyze the relationship between overall survival and treatments. Plenty of SAS Marcos were introduced during the data analysis process, which is a high-efficient time-saving tool that could easily generate readable tables and fancy graphs.

As previous STAMPED clinical mentioned, there is currently no optimal definition of low burden metastatic, because it only conveys a small number of metastatic definitions, so sometimes it can even be misleading. According to CHAARTED, if metastases are limited to lymph nodes and axial bones, there may be an unlimited number of metastases.[3] So in this study, we did not conduct a subgroup analysis of metastatic (high vs. low) but considered patients with metastatic in general. The author also mentioned that after receiving CHAARTED trial report, they could draw a clear clinical relevance of the metastatic burden of prostate cancer patients. Metastatic burden was determined by retrospectively gathering searchable baseline scans, which means they are unable to cover all patients. In our retrospective survival analysis, data sourced from NCDB data, which provided us with fully covered patient's information.

Based on previous STAMPEDE randomized controlled phase 3 trial, the author concluded that the survival for patients with node-positive prostate cancer (TanyN1M0), patients with low metastatic burden prostate cancer (TanyNanyM1) would be improved by radiation therapy. It is also wildly known that survival will be improved for patients with locally advanced prostate cancer (T3-4N0M0).[3] Our findings on radiation therapy have a protective effect on several prostate cancer patient groups are broadly consistent with results from previous clinical trials. We further conducted MVA to compare the survival rate between those three groups to identify which group will be associated with higher improved overall survival after receiving radiation therapy.

This study has several limitations. The existing selection bias is unallowable to be neglected, we did not put other non-cancer related health information into analysis, such as smoking status, alcohol consumption, diabetes, heart disease, high blood pressure, high blood sugar, etc. If these health conditions are protentional confounders, selection bias will exist, and our conclusion will need further proof. Although we conducted multivariate association analysis and propensity score analysis to minimize the known difference, the health status did not record in NCDB cannot be addressed. The speared site of metastatic was also inaccessible. The results could be verified by performing randomized clinical trial studies.

5 Conclusion

This study is a retrospective survival analysis that aimed to identify a subgroup of patients who may benefit from radiation therapy. Our results suggest that patients with newly diagnosed metastatic disease treated with radiation therapy have statistically significant superior overall survival than similar patients treat with none radiation therapy. Patients in the High-Risk group and Metastasis group were associated with higher survival when independently compared to patients in the Node-Positive group. This effect persisted after propensity score matching analysis to minimize the impact that selection bias brought.

In conclusion, our findings lend support that radiation therapy may confer a protective effect on prostate cancer patients; such a protective effect is more prominent in High-Risk or Metastasis subpopulation compared to the Node-Positive group.

6 Bibliography

[1] *Fischer-Valuck, Benjamin W., et al. "Palliative radiation therapy (RT) for prostate cancer patients with bone metastases at diagnosis: A hospital-based analysis of patterns of care, RT fractionation scheme, and overall survival." Cancer medicine 7.9 (2018): 4240-4250.*

[2] *Cancer.org. 2020. Prostate Cancer | Prostate Cancer Information And Overview. [online] Available at: <https://www.cancer.org/cancer/prostate-cancer.html> [Accessed 17 April 2020].*

[3] *Parker, C. C., et al. (2018). "Radiotherapy to the primary tumour for newly diagnosed, metastatic prostate cancer (STAMPEDE): a randomised controlled phase 3 trial." Lancet **392**(10162): 2353-2366.*

[4] *Cancer.Net. 2020. Prostate Cancer - Types Of Treatment. [online] Available at: <https://www.cancer.net/cancer-types/prostate-cancer/types-treatment> [Accessed 17 April 2020].*

[5] *Ncdbpuf.facs.org. 2020. Facility Type | National Cancer Data Base - Data Dictionary PUF 2016. [online] Available at: <http://ncdbpuf.facs.org/?q=content%2Fparticipant-use-file-facility-type> [Accessed 17 April 2020].*

[6] *Quilty, P. M., et al. "A comparison of the palliative effects of strontium-89 and external beam radiotherapy in metastatic prostate cancer." Radiotherapy and Oncology 31.1 (1994): 33-40.*

[7] *Liu, Y., Nickleach, D. C., Zhang, C., Switchenko, J. M., & Kowalski, J. (2018). Carrying out streamlined routine data analyses with reports for observational studies: introduction to a series of generic SAS® macros. F1000Research, 7.*

[8] *Parikh, R. R., et al. (2017). "Local therapy improves overall survival in patients with newly diagnosed metastatic prostate cancer." The Prostate **77**(6): 559-572.*

[9] *Culp, S. H., et al. (2014). "Might men diagnosed with metastatic prostate cancer benefit from definitive treatment of the primary tumor? A SEER-based study." European urology **65**(6): 1058-1066.*

[10] *Fossati, N., et al. (2015). "Identifying optimal candidates for local treatment of the primary tumor among patients diagnosed with metastatic prostate cancer: a SEER-based study." European urology **67**(1): 3-6.*

[11] *Gansler, T., et al. (2018). "Most Gleason 8 biopsies are downgraded at prostatectomy—does 4+ 4= 7?" The Journal of urology **199**(3): 706-712.*

[12] *Omarini, C., et al. (2017). "Impact of time to surgery after neoadjuvant chemotherapy in operable breast cancer patients." European Journal of Surgical Oncology (EJSO) **43**(4): 613-618.*

[13] *Rusthoven, C. G., et al. (2016). "Improved survival with prostate radiation in addition to androgen deprivation therapy for men with newly diagnosed metastatic prostate cancer." Journal of Clinical Oncology **34**(24): 2835-2842.*

[14] *Bilimoria, Karl Y., et al. "The National Cancer Data Base: a powerful initiative to improve cancer care in the United States." Annals of surgical oncology 15.3 (2008): 683-690.*

[15] *Rusthoven, C. G., et al. (2016). "Improved survival with prostate radiation in addition to androgen deprivation therapy for men with newly diagnosed metastatic prostate cancer." Journal of Clinical Oncology **34**(24): 2835-2842.*

[16] *Clinicaltrials.gov. 2020. Systemic Therapy In Advancing Or Metastatic Prostate Cancer: Evaluation Of Drug Efficacy - Full Text View - Clinicaltrials.Gov. [online] Available at: <https://clinicaltrials.gov/ct2/show/NCT00268476> [Accessed 17 April 2020].*

[17] *Sinha, S., et al. (2017). "Characteristics and national trends of patients receiving treatment of the primary tumor for metastatic prostate cancer." Prostate international **5**(3): 89-94.*

[18] *Yao, X., Wang, X., Speicher, P., Hwang, E., Cheng, P., Harpole, D., Berry, M., Schrag, D. and Pang, H., 2017. Reporting and Guidelines in Propensity Score Analysis: A Systematic Review of Cancer and Cancer Surgical Studies. JNCI: Journal of the National Cancer Institute, 109(8).*

[19] *Sun, X., et al. (2010). "Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses." Bmj **340**: c117.*

[20] *Zhou, X.-H., et al. (2017). "Statistical evaluation of adding multiple risk factors improves Framingham stroke risk score." BMC medical research methodology **17**(1): 58.*

[21] *Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate behavioral research, 46(3), 399-424.*

[22] *Bilimoria, K. Y., Stewart, A. K., Winchester, D. P., & Ko, C. Y. (2008). The National Cancer Data Base: a powerful initiative to improve cancer care in the United States. Annals of surgical oncology, 15(3), 683-690.*

[23] *Liu, Yuan, et al. "Carrying out streamlined routine data analyses with reports for observational studies: introduction to a series of generic SAS® macros." F1000Research 7 (2018).*

[24] *Sebastian, Nikhil T., et al. "Survival after radiotherapy vs. radical prostatectomy for unfavorable intermediate-risk prostate cancer." Urologic Oncology: Seminars and Original Investigations. Vol. 37. No. 11. Elsevier, 2019.*

[25] *Byrd, David R., et al. AJCC cancer staging manual. Ed. Stephen B. Edge. Vol. 7. New York: Springer, 2010.*

Table 1. Diagram of Study Population Selection

Selection and Exclusion Criteria	Sample Size	Excluded
NCDB Prostate PUF Cancer Cases	385039	-
Male with Age >=40	384926	113
Invasive tumor behavior	384884	42
Had no previous cancer diagnosis	352651	32233
Include Diagnostic Confirmation as positive histology	343461	9190
Include cases as high-risk, node positive, or metastasis	85390	258071
Include cases treated by hormone therapy within 6 months after diagnosis	73607	11783
Include cases had no surgery or chemotherapy or palliative care	51774	21833
If any, radiation volume as: Prostate and pelvis, Prostate, and Pelvis (NOS), and Total Radiation Dose > 59 Gy, started between -30 or 183 days relative to hormone therapy	41588	10186
Exclude cases had missing outcome	35177	6411

Table 2. Baseline characteristics for the entire study population and by the study cohorts and the factors predict the utilization of radiation therapy

Demographics and Clinical Characteristics	Total N (%) 35177 (100)	Univariate Association with Radiation Yes vs. No at any CoC Facility (All p < 0.001)			Logistic Regression Model for the Probability of Radiation Therapy Odds Ratio (95% CI)
		None Radiation N (%) 23985 (68.2)	Beam radiation N (%) 11192 (31.8)	ASD [1]	
Age at Diagnosis					
<=65	13618 (38.7)	9099 (37.94)	4519 (40.38)	0.14	1.32 (1.21-1.44)
>65	21559 (61.3)	14886 (62.06)	6673 (59.62)	**0.35**	Ref
Mean (SD)	69 (10.3)	69 (10.88)	68 (8.77)	**0.34**	
Race-Ethnic Groups					
NH-White	24986 (71.0)	16624 (69.31)	8362 (74.71)	0.12	1.15 (1.05-1.27)
NH-Black	6791 (19.3)	4905 (20.45)	1886 (16.85)	0.09	Ref
Asian	866 (2.5)	611 (2.55)	255 (2.28)	0.02	1.06 (0.84-1.33)
Hispanic	1951 (5.5)	1438 (6.00)	513 (4.58)	0.06	1.08 (0.91-1.28)
Other/Unknown	583 (1.7)	407 (1.70)	176 (1.57)	0.01	0.87 (0.67-1.12)
Median Income Quartiles 2008-2012					
<$38,000	7088 (20.2)	5093 (21.34)	1995 (17.92)	0.09	0.85 (0.74-0.98)
$38,000-$47,999	8157 (23.3)	5574 (23.35)	2583 (23.20)	0.01	0.87 (0.78-0.98)
$48,000-$62,999	9322 (26.6)	6283 (26.32)	3039 (27.29)	0.02	0.88 (0.79-0.97)
$63,000 +	10439 (29.8)	6921 (28.99)	3518 (31.59)	0.06	Ref
Percent No High School Degree 2008-2012					
>=21%	6548 (18.7)	4789 (20.05)	1759 (15.78)	0.11	0.91 (0.80-1.03)
13-20%	9028 (25.8)	6098 (25.53)	2930 (26.29)	0.02	1.07 (0.98-1.18)
7.0-12.9%	11255 (32.1)	7512 (31.44)	3743 (33.59)	0.05	Ref
<7%	8203 (23.4)	5491 (22.98)	2712 (24.34)	0.03	0.90 (0.81-0.99)
Urban/Rural 2013					
Metro	28351 (80.6)	19402 (80.89)	8949 (79.96)	0.02	Ref
Urban	5188 (14.7)	3444 (14.36)	1744 (15.58)	0.03	0.93 (0.84-1.02)
Rural	808 (2.3)	564 (2.35)	244 (2.18)	0.01	0.84 (0.67-1.05)

Unknown	830 (2.4)	575 (2.40)	255 (2.28)	0.01	0.77 (0.61-0.98)
Primary Payor					
Other Government/Not Insured/Unknown	5274 (15.0)	4044 (16.86)	1230 (10.99)	0.2	0.84 (0.75-0.95)
Private	10533 (29.9)	6576 (27.42)	3957 (35.36)	0.17	1.25 (1.14-1.37)
Medicare	19370 (55.1)	13365 (55.72)	6005 (53.65)	0.04	Ref
Year of Diagnosis					
2004-2006	6003 (17.1)	3893 (16.23)	2110 (18.85)	0.07	NS[2]
2007-2009	7156 (20.3)	4712 (19.65)	2444 (21.84)	0.05	
2010-2012	9154 (26.0)	6210 (25.89)	2944 (26.30)	0.01	
2013-2015	12864 (36.6)	9170 (38.23)	3694 (33.01)	0.11	
Facility Type					
Non-Academic/Research Program	21033 (59.8)	13728 (57.24)	7305 (65.27)	0.17	Ref
Academic/Research Program	14144 (40.2)	10257 (42.76)	3887 (34.73)	0.17	0.59 (0.55-0.63)
Facility Location					
East	14802 (42.1)	9679 (40.35)	5123 (45.77)	0.11	1.25 (1.16-1.34)
Central/Mountain	16456 (46.8)	11593 (48.33)	4863 (43.45)	0.10	Ref
West	3919 (11.1)	2713 (11.31)	1206 (10.78)	0.02	0.84 (0.75-0.94)
Charlson-Deyo Score					
0	28448 (80.9)	18860 (78.63)	9588 (85.67)	0.19	Ref
1	4773 (13.6)	3513 (14.65)	1260 (11.26)	0.10	0.75 (0.68-0.83)
2+	1956 (5.6)	1612 (6.72)	344 (3.07)	0.17	0.54 (0.46-0.64)
Sequence Number					
0	33115 (94.1)	22658 (94.47)	10457 (93.43)	0.04	NS[2]
1	2062 (5.9)	1327 (5.53)	735 (6.57)	0.04	
Grade					
Well/Moderately Differentiated	2020 (5.7)	1477 (6.16)	543 (4.85)	0.06	1.05 (0.89-1.24)
Poorly/Undifferentiated	27900 (79.3)	17632 (73.51)	10268 (91.74)	**0.50**	Ref
Unknown	5257 (14.9)	4876 (20.33)	381 (3.40)	**0.54**	0.77 (0.65-0.91)
CLIN_T					
T1	16960 (48.2)	12107 (50.48)	4853 (43.36)	0.14	Ref
T2	9072 (25.8)	3658 (15.25)	5414 (48.37)	**0.76**	1.95 (1.81-2.10)
T3	3654 (10.4)	2956 (12.32)	698 (6.24)	**0.21**	0.81 (0.72-0.91)
Unknown	5491 (15.6)	5264 (21.95)	227 (2.03)	**0.64**	0.54 (0.45-0.65)
TNM_N					

0	18399 (52.3)	10293 (42.91)	8106 (72.43)	**0.63**	Ref
1	11828 (33.6)	8843 (36.87)	2985 (26.67)	**0.22**	0.46 (0.43-0.50)
Unknown	4950 (14.1)	4849 (20.22)	101 (0.90)	**0.66**	0.40 (0.32-0.50)
TNM_M					
0	14910 (42.4)	4676 (19.50)	10234 (91.44)	**2.10**	0.62 (0.40-0.83)
1	20267 (57.6)	19309 (80.50)	958 (8.56)	**2.10**	Ref
PSA					
<10	5778 (16.4)	2761 (11.51)	3017 (26.96)	**0.40**	1.99 (1.81-2.17)
10-20	4709 (13.4)	2802 (11.68)	1907 (17.04)	0.15	1.51 (1.37-1.67)
>20	22420 (63.7)	16318 (68.03)	6102 (54.52)	**0.28**	Ref
Unknown	2270 (6.5)	2104 (8.77)	166 (1.48)	**0.34**	0.37 (0.30-0.45)
Gleason					
2-7	4341 (12.3)	3118 (13.00)	1223 (10.93)	0.06	0.98 (0.87-1.10)
8-10	24645 (70.1)	14911 (62.17)	9734 (86.97)	**0.59**	Ref
Unknown	6191 (17.6)	5956 (24.83)	235 (2.10)	**0.71**	0.35 (0.29-0.42)
Risk Group					
Low	54 (0.2)	39 (0.17)	15 (0.13)	0.01	NS[2]
Intermediate	1657 (4.9)	1292 (5.64)	365 (3.27)	0.12	
High	32368 (95.0)	21589 (94.19)	10779 (96.59)	0.12	
Months of Hormone Start from Diagnosis (quartiles)					
<=0.3	16960 (48.2)	12107 (50.48)	4853 (43.36)	0.14	
>0.3-0.76	9072 (25.8)	3658 (15.25)	5414 (48.37)	0.06	
>0.76-1.48	3654 (10.4)	2956 (12.32)	698 (6.24)	0.05	
>1.48	5491 (15.6)	5264 (21.95)	227 (2.03)	0.18	
Mean (SD)	1.1 (1.12)	0.93 (1.04)	1.48 (1.21)	0.17	1.14 (1.11-1.17)

[1]ASD: absolute standardized difference. A value of > 0.2 is considered a substantial difference in distribution.
[2]NS: not selected by the variable backward elimination at a significance level of 0.05.
* P-value < 0.05.

Table 3. Univariate and Multivariable Cox regression analysis of overall survival

Covariate	Months from Treatment					
	Univariate analysis			Multivariate analysis		
	Hazard Ratio (95% CI)	HR P-value	Type3 P-value	Hazard Ratio (95% CI)	HR P-value	Type3 P-value
Radiation Therapy at any CoC Facility						
Beam radiation	0.26 (0.25-0.27)	**<.001**	**<.001**	0.46 (0.44-0.48)	**<.001**	**<.001**
None	Ref	Ref		Ref	Ref	
Age at Diagnosis						
<=65	0.71 (0.69-0.74)	**<.001**	**<.001**	0.80 (0.77-0.84)	**<.001**	**<.001**
>65	Ref	Ref		Ref	Ref	
Race-Ethnic Groups						
NH-White	0.97 (0.93-1.00)	0.065	**<.001**	1.05 (1.01-1.10)	**0.018**	**<.001**
Asian	0.79 (0.71-0.88)	**<.001**		0.83 (0.74-0.92)	**<.001**	
Hispanic	0.82 (0.76-0.89)	**<.001**		0.80 (0.74-0.87)	**<.001**	
Other/Unknown	0.72 (0.63-0.81)	**<.001**		0.84 (0.73-0.95)	**0.008**	
NH-Black	Ref	Ref		Ref	Ref	
Median Income Quartiles 2008-2012						
<$38,000	1.18 (1.13-1.23)	**<.001**	**<.001**	--	--	
$38,000-$47,999	1.11 (1.07-1.16)	**<.001**		--	--	--
$48,000-$62,999	1.09 (1.05-1.13)	**<.001**		--	--	
$63,000 +	Ref	Ref		--	--	
Percent No High School Degree 2008-2012						
>=21%	1.06 (1.02-1.10)	**0.007**	**<.001**	1.03 (0.99-1.08)	0.179	**<.001**
13-20%	1.01 (0.97-1.05)	0.675		1.03 (0.99-1.07)	0.205	
<7%	0.92 (0.88-0.95)	**<.001**		0.94 (0.90-0.98)	**0.002**	
7.0-12.9%	Ref	Ref		Ref	Ref	
Urban/Rural 2013						
Urban	0.99 (0.95-1.03)	0.58	0.769	0.95 (0.91-0.99)	**0.017**	**0.011**
Rural	1.01 (0.91-1.11)	0.909		0.88 (0.80-0.97)	**0.013**	
Unknown	0.95 (0.86-1.05)	0.353		0.96 (0.86-1.06)	0.402	
Metro	Ref	Ref		Ref	Ref	
Primary Payor						

Other Government/Not Insured/Unknown	0.90 (0.86-0.94)	**<.001**	**<.001**	0.99 (0.94-1.04)	0.645	**<.001**
Private	0.66 (0.64-0.69)	**<.001**		0.83 (0.79-0.87)	**<.001**	
Medicare	Ref	Ref		Ref	Ref	
Year of Diagnosis						
2004-2006	0.95 (0.91-1.00)	**0.032**	0.062	1.09 (1.04-1.14)	**<.001**	**<.001**
2007-2009	0.98 (0.94-1.02)	0.353		1.09 (1.04-1.14)	**<.001**	
2010-2012	0.95 (0.91-0.99)	**0.021**		1.02 (0.97-1.06)	0.425	
2013-2015	Ref	Ref		Ref	Ref	
Facility Type						
Academic/Research Program	0.85 (0.82-0.88)	**<.001**	**<.001**	0.83 (0.81-0.86)	**<.001**	**<.001**
Non-Academic/Research Program	Ref	Ref		Ref	Ref	
Facility Location						
East	0.94 (0.91-0.97)	**<.001**	**<.001**	0.99 (0.96-1.03)	0.751	**0.021**
West	0.88 (0.84-0.92)	**<.001**		0.93 (0.88-0.98)	**0.006**	
Central/Mountain	Ref	Ref		Ref	Ref	
Charlson-Deyo Score						
1	1.44 (1.39-1.50)	**<.001**	**<.001**	1.26 (1.20-1.31)	**<.001**	**<.001**
2+	2.24 (2.12-2.37)	**<.001**		1.73 (1.63-1.83)	**<.001**	
0	Ref	Ref		Ref	Ref	
Sequence Number						
1	1.15 (1.09-1.21)	**<.001**	**<.001**	1.16 (1.10-1.23)	**<.001**	**<.001**
0	Ref	Ref		Ref	Ref	
Grade						
Well/Moderately Differentiated	0.88 (0.82-0.94)	**<.001**	**<.001**	0.91 (0.84-0.99)	**0.023**	**<.001**
Unknown	2.17 (2.09-2.26)	**<.001**		1.18 (1.11-1.25)	**<.001**	
Poorly/Undifferentiated	Ref	Ref		Ref	Ref	
CLIN_T						
T2	0.66 (0.63-0.69)	**<.001**	**<.001**	0.97 (0.94-1.02)	0.231	**<.001**
T3	1.69 (1.62-1.77)	**<.001**		1.37 (1.31-1.44)	**<.001**	
Unknown	2.00 (1.93-2.08)	**<.001**		1.16 (1.11-1.22)	**<.001**	
T1	Ref	Ref		Ref	Ref	
TNM_N						
1	1.23 (1.19-1.27)	**<.001**	**<.001**	1.02 (0.99-1.06)	0.249	**<.001**
Unknown	2.47 (2.38-2.57)	**<.001**		1.11 (1.06-1.16)	**<.001**	

0	Ref	Ref		Ref	Ref	
TNM_M						
0	0.29 (0.28-0.30)	**<.001**	**<.001**	0.53 (0.51-0.56)	**<.001**	**<.001**
1	Ref	Ref		Ref	Ref	
PSA						
<10	0.60 (0.58-0.63)	**<.001**	**<.001**	0.88 (0.83-0.92)	**<.001**	**<.001**
10-20	0.76 (0.72-0.79)	**<.001**		0.93 (0.89-0.98)	**0.006**	
Unknown	1.34 (1.27-1.41)	**<.001**		1.00 (0.94-1.08)	0.914	
>20	Ref	Ref		Ref	Ref	

* Backward selection with an alpha level of removal of 0.05 was used. The following variables were removed from the model: Median Income Quartiles 2008-2012.

Table 4. Multivariable interaction model for overall survival for the effect of radiation in subgroups

Covariates	Months from Treatment		
	Multivariate Association with Radiation Yes vs. No		
	Hazard Ratio (95% CI)	HR P-value	Type 3 P-value
Race-Ethnic Groups			0.075
NH-White	0.47 (0.44-0.50)	**<.001**	-
Asian	0.37 (0.27-0.50)	**<.001**	-
Hispanic	0.38 (0.30-0.48)	**<.001**	-
Other/Unknown	0.55 (0.40-0.75)	**<.001**	-
NH-Black	0.43 (0.39-0.48)	**<.001**	-
Age at Diagnosis			0.307
<=65	0.47 (0.44-0.51)	**<.001**	-
>65	0.45 (0.43-0.48)	**<.001**	-
Facility Type			0.067
Academic/Research Program	0.49 (0.45-0.52)	**<.001**	-
Non-Academic/Research Program	0.45 (0.42-0.47)	**<.001**	-
Charlson-Deyo Score			0.920
1	0.45 (0.40-0.51)	**<.001**	-
2+	0.45 (0.38-0.54)	**<.001**	-
0	0.46 (0.44-0.49)	**<.001**	-
TNM_M			0.374
0	0.47 (0.44-0.49)	**<.001**	-
1	0.44 (0.40-0.49)	**<.001**	-
Patient Group			**<.001**
High-Risk	0.42 (0.39-0.45)	**<.001**	-
Node-Positive	0.58 (0.52-0.64)	**<.001**	-
Metastasis	0.44 (0.40-0.49)	**<.001**	-

Table 5. The estimated 5-yr and 10-yr survival rate by study cohorts in original and matched sample

Radiation Therapy at any CoC Facility		N	Event	Censored	Median Survival (95% CI)	60 Mo Survival	120 Mo Survival
Original sample	Beam radiation	11192	3319 (30%)	7873 (70%)	107.8 (105, 110.7)	74.9% (73.9%, 75.8%)	44.7% (43.1%, 46.3%)
	None	23985	15068 (63%)	8917 (37%)	33.4 (32.8, 34.1)	29.4% (28.7%, 30.1%)	11.1% (10.4%, 11.8%)
Matched sample	Beam radiation	8129	2464 (30%)	5665 (70%)	106.5 (103.3, 109.7)	73.9% (72.7%, 75.0%)	44.0% (42.1%, 45.9%)
	None	8129	4493 (55%)	3636 (45%)	42.9 (41.6, 44.4)	37.4% (36.1%, 38.7%)	15.3% (13.9%, 16.8%)

Table 6. Univariate and Multivariable cox regression analysis of overall survival in PS matched sample

Covariate	N	Months from Treatment					
		Univariate analysis			Multivariate analysis		
		Hazard Ratio (95% CI)	HR P-value	Type3 P-value	Hazard Ratio (95% CI)	HR P-value	Type3 P-value
Radiation Therapy at any CoC Facility							
Beam radiation	8048	0.34 (0.32-0.36)	**<.001**	**<.001**	0.46 (0.43-0.49)	**<.001**	**<.001**
Beam radiation	8048	Ref	Ref		Ref	Ref	
Age at Diagnosis							
<=65	6706	0.78 (0.74-0.82)	**<.001**	**<.001**	0.78 (0.74-0.84)	**<.001**	**<.001**
>65	9390	Ref	Ref		Ref	Ref	
Race-Ethnic Groups							
NH-White	11724	1.10 (1.03-1.17)	**0.003**	**<.001**	1.04 (0.97-1.11)	0.300	**<.001**
Asian	398	0.80 (0.67-0.96)	**0.015**		0.78 (0.65-0.93)	**0.006**	
Hispanic	766	0.81 (0.71-0.93)	**0.003**		0.78 (0.67-0.89)	**<.001**	
Other/Unknown	277	0.72 (0.58-0.89)	**0.002**		0.81 (0.66-1.00)	0.054	
NH-Black	2931	Ref	Ref		Ref	Ref	
Median Income Quartiles 2008-2012							
<$38,000	3058	1.05 (0.98-1.13)	0.135	0.064	--	--	--
$38,000-$47,999	3802	1.09 (1.02-1.16)	**0.009**		--	--	
$48,000-$62,999	4325	1.06 (0.99-1.13)	0.079		--	--	
$63,000 +	4911	Ref	Ref		--	--	
Percent No High School Degree 2008-2012							
>=21%	2731	0.97 (0.91-1.04)	0.438	0.201	--	--	--
13-20%	4171	1.00 (0.94-1.06)	0.983		--	--	
<7%	3890	0.94 (0.88-1.00)	0.055		--	--	
7.0-12.9%	5304	Ref	Ref		--	--	
Urban/Rural 2013							
Urban	2533	1.01 (0.95-1.08)	0.736	0.648	--	--	--
Rural	382	1.04 (0.90-1.21)	0.605		--	--	

Unknown	325	0.90 (0.76-1.08)	0.27		--	--	
Metro	12856	Ref	Ref		--	--	
Primary Payor							
Other Government/Not Insured/Unknown	1957	0.88 (0.81-0.95)	**<.001**	**<.001**	1.12 (1.03-1.23)	**0.008**	**<.001**
Private	5492	0.73 (0.70-0.77)	**<.001**		0.85 (0.80-0.91)	**<.001**	
Medicare	8647	Ref	Ref		Ref	Ref	
Year of Diagnosis							
2004-2006	2731	1.01 (0.93-1.09)	0.865	0.721	--	--	--
2007-2009	3264	1.04 (0.96-1.12)	0.33		--	--	
2010-2012	4287	1.02 (0.95-1.10)	0.561		--	--	
2013-2015	5814	Ref	Ref		--	--	
Facility Type							
Academic/Research Program	6227	0.80 (0.76-0.84)	**<.001**	**<.001**	0.81 (0.77-0.85)	**<.001**	**<.001**
Non-Academic/Research Program	9869	Ref	Ref		Ref	Ref	
Facility Location							
East	6976	0.98 (0.93-1.03)	0.366	**<.001**	0.99 (0.96-1.03)	0.751	**0.021**
West	1785	0.85 (0.79-0.93)	**<.001**		0.93 (0.88-0.98)	**0.006**	
Central/Mountain	7335	Ref	Ref		-	-	
Charlson-Deyo Score							
1	2002	1.32 (1.23-1.41)	**<.001**	**<.001**	1.26 (1.18-1.35)	**<.001**	**<.001**
2+	603	1.79 (1.61-2.00)	**<.001**		1.74 (1.55-1.94)	**<.001**	
0	13491	Ref	Ref		Ref	Ref	
Sequence Number							
1	1061	1.43 (1.32-1.55)	**<.001**	**<.001**	1.33 (1.23-1.44)	**<.001**	**<.001**
0	15035	Ref	Ref		Ref	Ref	
Grade							
Well/Moderately Differentiated	915	0.76 (0.68-0.86)	**<.001**	**<.001**	--	--	**--**
Unknown	588	1.17 (1.03-1.33)	**0.017**		--	--	
Poorly/Undifferentiated	14593	Ref	Ref		--	--	
CLIN_T							
T2	5403	0.97 (0.92-1.02)	0.291	**<.001**	1.03 (0.97-1.08)	0.365	**--**
T3	1305	1.67 (1.54-1.81)	**<.001**		1.65 (1.52-1.79)	**<.001**	
Unknown	392	0.99 (0.86-1.15)	0.941		1.11 (0.96-1.28)	0.16	

T1	8996	Ref	Ref		Ref	Ref	
TNM_N							
1	5375	1.02 (0.97-1.07)	0.503	**<.001**	--	--	**--**
Unknown	183	1.80 (1.51-2.15)	**<.001**		--	--	
0	10538	Ref	Ref		--	--	
TNM_M							
0	9705	0.33 (0.32Ref0.35)	**<.001**	**<.001**	--	--	**--**
1	6391	Ref	Ref		--	--	
PSA							
<10	3192	0.85 (0.80-0.91)	**<.001**	**<.001**	0.84 (0.79-0.89)	**<.001**	**<.001**
10-20	2550	0.93 (0.87-0.99)	**0.025**		0.88 (0.82-0.94)	**<.001**	
Unknown	288	1.14 (0.98-1.34)	0.099		1.04 (0.89-1.22)	0.618	
>20	10066	Ref	Ref		Ref	Ref	
Gleason							
2-7	2105	0.70 (0.65-0.75)	**<.001**	**<.001**	0.65 (0.60-0.70)	**<.001**	**<.001**
Unknown	389	1.22 (1.07-1.39)	**0.004**		1.10 (0.96-1.25)	0.184	
8-10	13602	Ref	Ref		Ref	Ref	
Risk Group							
Low	6422	0.34 (0.32-0.36)	**<.001**	**<.001**	--	--	**--**
Intermediate	3283	0.33 (0.30-0.35)	**<.001**		--	--	
High	6391	Ref	Ref		--	--	
Patient Group							
High-Risk	6422	--	--	**--**	0.47 (0.44-0.51)	**<.001**	**<.001**
Node-Positive	3283	--	--		0.54 (0.50-0.58)	**<.001**	
Metastasis	6391	--	--		-	-	
Months of Hormone Start from Diagnosis	16096	0.82 (0.80-0.84)	**<.001**	**<.001**	0.88 (0.86-0.90)	**<.001**	**<.001**

* Backward selection with an alpha level of removal of 0.05 was used. The following variables were removed from the model: Facility Location, Grade, Median Income Quartiles 2008-2012, Percent No High School Degree 2008-2012, Urban/Rural 2013, and Year of Diagnosis.

Table 7. Multivariable interaction model for overall survival for the effect of radiation in subgroups (matched sample)

Covariates	Months from Treatment Multivariate Association with Radiation Yes vs. No		
	Hazard Ratio (95% CI)	**HR P-value**	**Type3 P-value**
Race-Ethnic Groups			0.075
NH-White	0.47 (0.44-0.50)	**<.001**	-
Asian	0.39 (0.37-0.50)	**<.001**	-
Hispanic	0.37 (0.28-0.50)	**<.001**	-
Other/Unknown	0.56 (0.37-0.84)	**<.001**	-
NH-Black	0.45 (0.39-0.51)	**<.001**	-
Age at Diagnosis			0.307
<=65	0.47 (0.43-0.51)	**<.001**	-
>65	0.46 (0.43-0.49)	**<.001**	-
Facility Type			0.067
Academic/Research Program	0.53 (0.48-0.57)	**<.001**	-
Non-Academic/Research Program	0.43 (0.40-0.47)	**<.001**	-
Charlson-Deyo Score			0.920
1	0.46 (0.40-0.53)	**<.001**	-
2+	0.47 (0.37-0.59)	**<.001**	-
0	0.46 (0.43-0.49)	**<.001**	-
Patient Group			**<.001**
High-Risk	0.43 (0.39-0.47)	**<.001**	-
Node-Positive	0.59 (0.52-0.67)	**<.001**	-
Metastasis	0.43 (0.39-0.48)	**<.001**	-

Figure 1: Kaplan-Meier curves by Radiation Yes vs. No

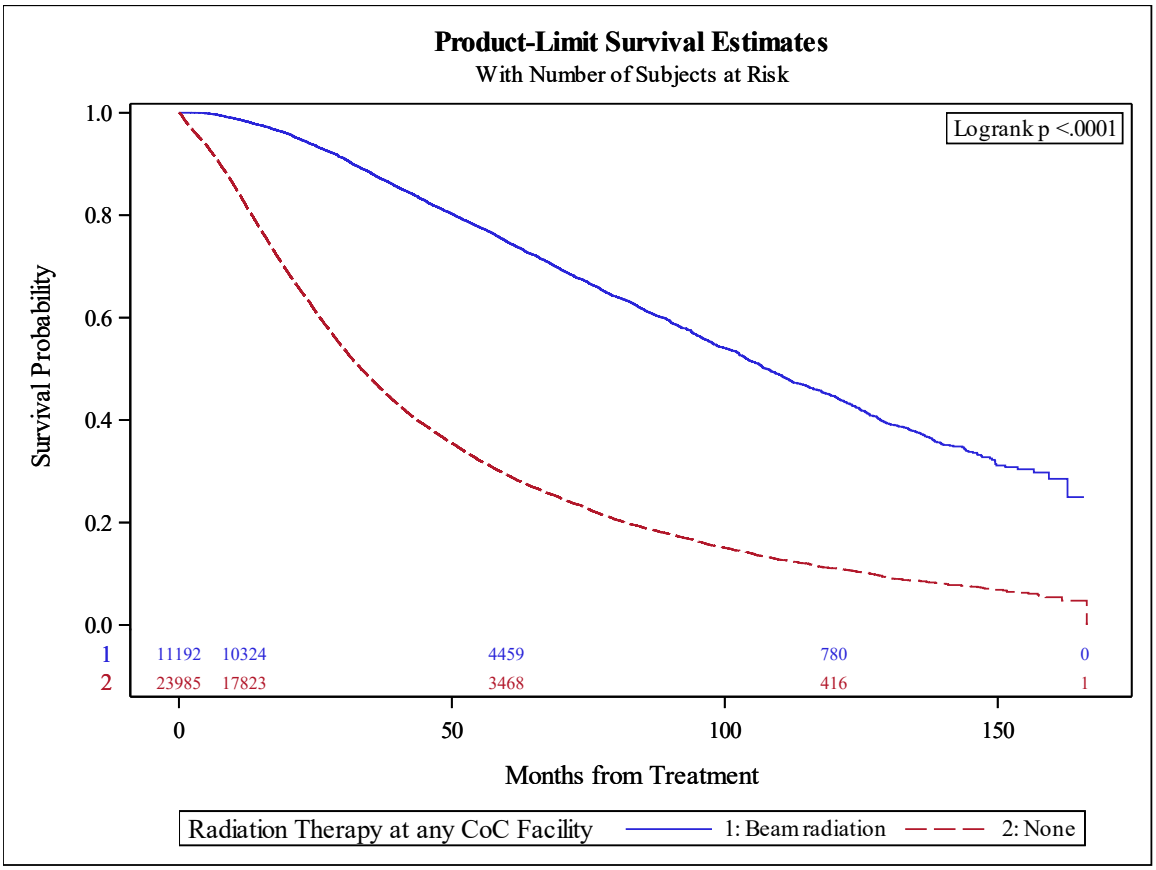

Figure 2: Kaplan-Meier curves by Radiation Yes vs. No in PS matched sample

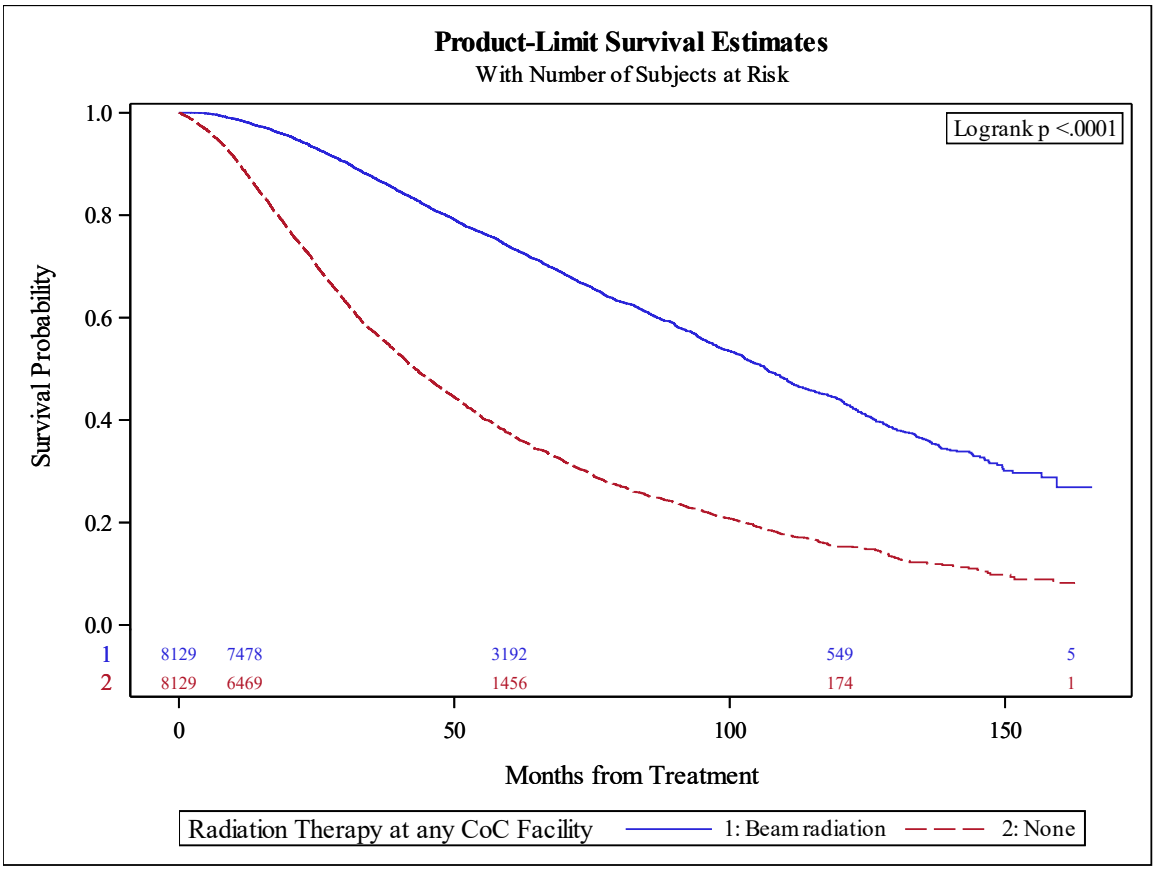

Figure 3. PS matching overlap

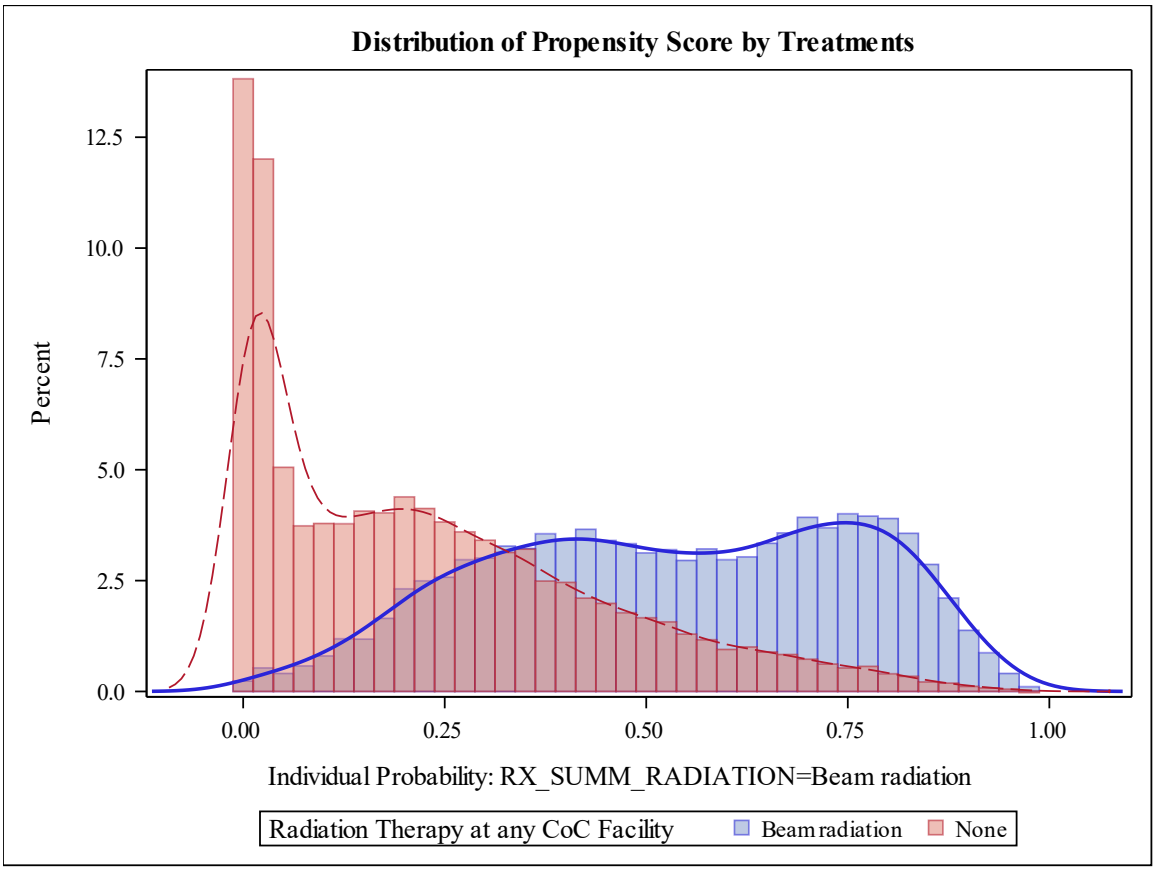

Figure 4. Covariate balance check before and after PS matching