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:

_____	_____
Xiting Zhu	Date

Propensity Score Matched Analysis on Survival Profile of Head and Neck Cancer in the U.S. Hispanic Population

By

Xiting Zhu
Master of Public Health

Biostatistics and Bioinformatics

_____

Yuan Liu, PhD
(Thesis Advisor)

_____

C. Christina Mehta, PhD, MSPH
(Reader)

Propensity Score Matched Analysis on Survival Profile of Head and Neck Cancer in the U.S. Hispanic Population

By

Xiting Zhu

B.S.
Beijing Forestry University
2015

Thesis Committee Chair: Yuan Liu, PhD
Reader: C. Christina Mehta, PhD, MSPH

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
2017

Abstract

Propensity Score Matched Analysis on Survival Profile of Head and Neck Cancer in the U.S. Hispanic Population

By Xiting Zhu

Introduction: Head and neck cancer (HNC) ranks as the 10th most common cancer worldwide and the 7th most common cancer inducing mortality. It causes 300,000 deaths each year worldwide among 600,000 newly diagnosed cases. A better understanding of potential risk factors for the incidence of HNC should be completed as well as thorough survival analyses for each racial and ethnic population. Despite the continuous increase in the U.S. Hispanic population in the recent decade, however, limited data was published on the survival of HNC in the Hispanic population. The purpose of this paper is to compare overall survival (OS) and disease-specific survival between Hispanic and the general HNC patients of the United States and to identify prognostic factors that might affect the survival outcomes of HNC patients.

Methods: The Surveillance, Epidemiology, and End Results (SEER) database was queried for HNC patients diagnosed between 2000 and 2013. Kaplan-Meier and log-rank test were performed to compare OS and disease-specific survival by racial-ethnical groups. A multivariable Cox proportional hazard model was fit to select prognostic factors. In addition, a multivariable model was fit to test for the impact of interaction between race-ethnic groups and age on survival outcomes. A method of propensity score matching was also implemented.

Results: A total of 113885 patients were identified in the study; 8132 (7.1%) were Hispanic, 12669 (11.1%) were non-Hispanic Black, 85995 (75.5%) were non-Hispanic White, and 7089 (6.2%) were non-Hispanic other. 31720 (27.9%) patients were diagnosed with laryngeal cancer, 49850 (43.8%) were oral cavity cancer, and 32315 (28.4%) were pharyngeal cancer. In patients with HNC, Hispanic race was associated with poor median OS and 5-year overall survival rate compared with non-Hispanic White (72 vs. 80 months; 53% vs. 56%; $P < 0.0001$). In the propensity score matched sample, a poorer OS for Hispanic patients versus non-Hispanic White patients was obtained (HR, 1.09; 95% CI, 1.01-1.16; $P = 0.021$). Inferior median OS and 5-year overall survival rate remained associated with Hispanic race (compared with non-Hispanic White: 64 months vs. 78 months; 51% vs. 55.6%; $P < 0.0001$).

Conclusions: Hispanic race was associated with poor OS and disease-specific survival in HNC patients. A series of prognostic factors resulting in inferior survival were found in Hispanic HNC patients.

Propensity Score Matched Analysis on Survival Profile of Head and Neck Cancer in the U.S. Hispanic Population

By

Xiting Zhu

B.S.
Beijing Forestry University
2015

Thesis Committee Chair: Yuan Liu, PhD
Reader: C. Christina Mehta, PhD, MSPH

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
2017

Acknowledgements

I would like to thank the Rollins School of Public Health at Emory University for the professional training and guidance over the past two years. I would especially like to thank Yuan Liu for all of the help she gave me on my thesis as my thesis advisor. I would like to thank Renjian Jiang and Melissa Oh for answering all of my questions throughout this project. I would especially like to thank Christina Mehta for taking the time to read my thesis and giving me guidance. Finally, I would like to thank my family and friends for their support and encouragement on the completion of this project.

Table of Contents

1. INTRODUCTION

Head and neck cancer (HNC) ranks as the 10th most common cancer worldwide and the 7th most common cancer inducing mortality (World Health Organization, 2008). It causes 300,000 deaths each year worldwide among 600,000 newly diagnosed cases, with the most common cancer site being the larynx, the oral cavity, and the pharynx (Boyle & Levin, 2008). Smoking, alcohol use, genetic factors, and viral infection are known as the most prominent risk factors identified so far to induce the incidence of HNC, with the most important prognostic factors being cancer site and tumor TNM stage (Mehanna et al., 2010a). The standard treatment which has been widely accepted for HNC is combined surgery and radiotherapy along with rehabilitation (Mehanna et al., 2010a). However, the racial disparities in HNC and other socioeconomically disadvantaged factors of the American population adversely affect the treatment and its outcome (Goodwin et al., 2007). The high prevalence and mortality rates of HNC urge the need of thorough survival analyses for each racial and ethnic population as well as better understanding of potential risk factors to the incidence of HNC.

1.1 Problem Statement

In the United States, 50.5 million people (or 16 percent) among the total population self-reported their ethnicity as Hispanic or Latino according to the 2010 census with an increase of 15.2 million people since 2000 (Ennis et al., 2011). This accounts for over half of the 27.3 million increase in the total population of the United States (Ennis et al., 2011). However, there is little information published regarding the survival of HNC in the Hispanic population given the fact

that the Hispanic population is less likely to be involved in clinical trial studies (Murthy et al., 2004).

1.2 Purpose Statement

The purpose of this paper is to compare overall survival (OS) and disease-specific survival between four race/ethnicity groups: Hispanic, non-Hispanic White, non-Hispanic Black and non-Hispanic other patients of the United States and to identify prognostic factors that are associated with the survival outcomes of HNC patients. To reduce potential bias from its retrospective design, a propensity score analysis was performed by balancing the baseline demographic and socioeconomic factors among different racial groups.

1.3 Significance Statement

This paper will provide understanding of survival in HNC in the Hispanic population and how the survival varies by site. This will assist in diagnosis and treatment for these particular cancers in the Hispanic population. Moreover, this project could investigate different risk factors in this population to help with developing new preventive interventions, evidence of early detection or innovative therapies for HNC.

2. BACKGROUND/LITERATURE REVIEW

Study Topic: Head and Neck Cancer in Hispanic Population

2.1 Head and Neck Cancer Overview

Head and neck cancers include cancers of the upper aerodigestive tract (including oral cavity, nasopharynx, oropharynx, hypopharynx and larynx), the paranasal sinuses, and the salivary glands (Mehanna et al., 2010b). In the United States, the estimated incidence and mortality in 2015 were 59,340 new cases and 12,290 deaths, of which more than 90% are squamous cell carcinomas (Wyss et al., 2016). The epidemic of HNC demonstrates a trend that men and elders have higher incidence rates compared with women and younger population. Large scale epidemiological studies and national cancer registries reported that the incidence rate for male versus female varies from 2:1 to 15:1 depending on the site of cancer (Mehanna et al., 2010b). In Europe, 98% and 50% newly diagnosed patients are middle-aged and elderly people over 40 and 60 years of age, respectively (Mehanna et al., 2010b). The Indian subcontinent, Australia, France, Brazil, and Southern Africa suffer from a higher incidence of HNC while the incidence of oral cavity, laryngeal and other tobacco use related cancers is declining in North America and western Europe likely due to decreased exposure to tobacco (Boyle & Levin, 2008). Some developing countries which have the highest tobacco prevalence due to lagged tobacco free policy or regulation, have relative high incidence of HNC (Mehanna et al., 2010b). Although each particular cancer site has different courses and histopathological types of HNC, squamous cell carcinoma is the most common one (Mehanna et al., 2010b). The cancers and their treatments

heavily affect the corresponding anatomical sites on functions such as speech, swallowing, taste, and smell so that the patients' life quality cannot be assured (Mehanna et al., 2010b).

Known risk factors include tobacco and alcohol use, HPV infection of the mouth and throat, and genetic factors such as family history of HNC (Mehanna et al., 2010b). Many studies suggested a synergistic relationship between smoking and alcohol use that results in a 30-fold increased risk for individuals who both smoke and drink heavily (American Cancer Society, 2017). Non-smokers who have three or more alcoholic drinks per day have double the risk of developing the disease compared with non-drinkers, reported by pooled analyses of 15 case-control studies (odds ratio: 2.04, 95% confidence interval: 1.29 to 3.21) (Mehanna et al., 2010b). Individuals who have family history of HNC and who are infected by HPV have 1.7-fold (95% CI, 1.2 to 2.3) and 63-fold (95% CI, 14 to 480) increased risk of developing HNC, respectively. (Mehanna et al., 2010b).

Patients with HNC usually present a variety of symptoms, depending on the function of the cancer original site (Mehanna et al., 2010a). Laryngeal cancers usually present with hoarseness, while pharyngeal and oral cavity cancers normally present lesion or lump in the throat or mouth, ear pain, a neck mass or coughing up blood (American Cancer Society, 2017). Prevention of HNC is closely linked to the success in control of tobacco use. Previous studies demonstrated that quitting from smoking for 1-4 years reduces the risk of developing HNC (0.70, 95% CI 0.61 to 0.81) compared with current smoking, and further reduces the risk for 20 years or more (0.23,

95% CI 0.18 to 0.31) (Oxford Cancer Intelligence Unit, 2010). Benefit of quitting alcohol use, however, was only seen for 20 years or more of quitting time (Marron et al., 2010).

2.2 Prognosis and Treatment

The most important prognostic factors are site and tumor TNM (tumor, node, metastasis) stage (Mehanna et al., 2010a). The average survival rate for HNC at 5-years is 34.4%, with a survival rate of 44.9% for women and 31.0% for men, respectively (Gatta et al., 2015). The survival rates for HNC on each of the three major cancer sites (oral cavity, larynx and pharynx), however, vary study by study. Generally, the overall 5-year survival rate was in the range of 20-43% for oral cancer, 8-25% for pharyngeal cancers and 25-62% for laryngeal cancer (Rao et al., 1998). Another prominent prognostic factor is comorbid illness, which was proved to be associated with poor prognosis (Mehanna et al., 2010a).

Combined surgery and radiotherapy along with rehabilitation have now become the standard treatment for HNC, often encouraged by national guidelines and protocols (National Institute for Health and Clinical Excellence, 2004). Radiotherapy and cancer-directed surgery are two common and widely accepted treatments for HNC. The choice of treatment depends on individual factors related to the site of the tumor and cancer stage, but also patient preference (Mehanna et al., 2010a). For early stage tumors, radiotherapy is more superior than cancer-directed surgery for better organ preservation (Scottish Intercollegiate Guidelines Network). For advanced tumors, on the other hand, combined use of surgery and postoperative radiotherapy, or combined

chemotherapy and radiotherapy, ensure the highest chance of achieving the best outcome (Bhalavat et al., 2003).

2.3 Influence of Race/Ethnicity

According to the 2010 Census, 308.7 million people resided in the United States, of which 50.5 million people (or 16 percent) were of Hispanic or Latino origin (Ennis et al., 2011). The Hispanic population increased by 15.2 million from 35.3 million (13 percent) in 2000, which accounted for over half of the increase in the total population of the United States (Ennis et al., 2011). Between 2000 and 2010, the Hispanic population grew by 43 percent, which was four times the growth of the total population (Ennis et al., 2011).

Given the huge increase in the U.S. Hispanic population, however, there is very little data published on the survival of cancer in the Hispanic population, particularly in HNC. Compared with a 1.8% cancer research enrollment proportion among White patients, a relatively low enrollment rate was seen in Hispanic patients (1.3%; odds ratio vs. whites, 0.72; 95% CI, 0.68 to 0.77) (Murthy et al., 2004). Moreover, Shavers et al studied patterns of care, socioeconomic status and comorbidity in 471 patients with HNC and found racial disparities for the receipt of cancer treatment among patients with particular types of HNC (Shavers et al., 2003). Shavers et al also found that the differences in cancer stage and the lower rates of cancer screening for Hispanic patients compared with White patients result in different rates of early detection, which in turn contributes to racial differences in survival (Shavers et al., 2003).

Statistical Method: Propensity Score

2.4 Overview

Propensity score methods are widely used to analyze causal association between treatment and outcome in observational studies to reduce confounding that may affect the estimated effects of treatment. The propensity score is the probability of treatment assignment conditional on observed baseline covariates: $e_i = \Pr(Z_i = 1|X_i)$ where Z_i is an indicator variable denoting whether or not the ith subject was treated and e_i denotes the propensity score of the ith subject given his/her set of baseline covariates X_i (Austin, 2011). The propensity score allows one to design and analyze an observational (nonrandomized) study since it mimics some particular characteristics of a randomized control trial (Austin, 2011). After matching based on the propensity scores, the distributions of measured baseline covariates are similar between treatment and control subjects (Austin, 2011). Consequently, the distribution of the observed baseline covariates will be the same between the treatment and control groups in a set of subjects who have the same propensity score (Austin, 2011). The implementation is based on a logistic regression model, in which treatment status is regressed on selected observed baseline covariates. The predicted probability of treatment derived from the logistic regression model is the propensity score.

2.5 Methodological decisions

There are four different propensity score methods used for removing the effects of confounding: propensity score matching, stratification (or subclassification) on the propensity

score, inverse probability of treatment weighting (IPTW) using the propensity score, and covariate adjustment using the propensity score (Rosenbaum & Rubin, 1983). Among these methods, propensity score matching and inverse probability of treatment weighting using propensity score allow one to estimate the marginal effect of treatment on survival when estimating the effect of treatment on time-to-event outcomes using observational or non-randomized data (Austin, 2012). However, IPTW outperforms propensity score matching in estimates with lower MSE, proving that the estimates resulting from IPTW have improved precision than those resulting from propensity score matching (Austin, 2012).

Matching on the propensity score generates matched pairs of treated and untreated subjects whose values of propensity score are similar (Rosenbaum & Rubin, 1985). Similarity in the values of propensity score results in homogeneous distribution of the observed baseline covariates. Once the matched sample has been formed, the treatment effects can be directly estimated by comparing the outcomes from treated and untreated subjects in the matched sample (Austin, 2011). Several decisions must be made prior to forming matched pairs of treated and untreated subjects.

First, a decision of matching without replacement or matching with replacement needs to be considered. When using matching without replacement, once an untreated subject has been selected to be matched to a treated subject, that untreated subject is no longer available as a potential match for subsequent treated subjects (Austin, 2011). In contrast, matching with

replacement allows a given untreated subject to be included in more than one sets of matching (Austin, 2011).

A second choice is between greedy and optimal matching (Rosenbaum, 2002). The philosophy of greedy matching is that a treated subject is first selected at random. The untreated subject whose propensity score is closest to that of this randomly selected treated subject is chosen for matching to this treated subject (Austin, 2011). This process is then repeated until untreated subjects have been matched to all treated subjects or until the list of treated subjects for whom a list of untreated subjects can be found has been exhausted (Austin, 2011). An alternative to greedy matching is optimal matching, in which matches are formed in order to minimize the total within-pair difference of propensity score (Austin, 2011). However, Gu & Rosenbaum found that optimal matching did not perform better than greedy matching in producing balanced matched sample (Gu & Rosenbaum, 1993).

Third, a choice between nearest neighbor matching (NNM) and nearest neighbor matching within a specified caliper distance is then raised (Rosenbaum & Rubin, 1985). NNM selects for matching to a given treated subject that untreated subject whose propensity score is closest to that of the treated subject (Austin, 2011). One important setting here is that there are no restrictions upon the maximum acceptable distance between the propensity scores of two matched subjects (Austin, 2011). Nearest neighbor matching within a specified caliper distance resembles NNM with the further restriction that the absolute difference in the propensity scores of matched subjects must be below the pre-specified arbitrary caliper distance (Austin, 2011). If no untreated

subjects whose propensity scores lay within the pre-specified caliper distance of the propensity score of the treated subject, that treated subject would not be matched with any untreated subject (Austin, 2011). Then unmatched treated subject would be then excluded from the matched sample so that bad matches are avoided and the matching quality is ensured.

Inverse probability of treatment weighting (IPTW) using propensity score uses weights calculated by propensity score to create a synthetic sample in which the distribution of measured baseline covariates is independent of treatment groups (Austin, 2011). Rosenbaum first proposed that weights can be defined as $w_i = \frac{z_i}{e_i} + \frac{(1-z_i)}{1-e_i}$ (Rosenbaum, 1987). Then the weights based on the inverse probability of treatment can be used on regression models to estimate causal effects of treatments (Joffe et al., 2004).

In terms of the balance diagnostics, comparing the similarity of baseline covariates for treated and untreated subjects in the propensity score sample should begin with a comparison of the means or medians of continuous covariates and the distribution of categorical covariates between treated and untreated subjects (Austin, 2011). The standardized difference is widely used in comparing the mean of continuous and binary variables between treatment groups (multilevel categorical variables can be represented as several binary indicator variables) (Flury & Riedwyl, 1986). For a continuous covariate, the standard difference is given by:

$$d = \frac{(\bar{x}_{treatment} - \bar{x}_{control})}{\sqrt{\frac{{s^2}_{treatment} + {s^2}_{control}}{2}}}$$

where $\bar{x}_{treatment}$ and $\bar{x}_{control}$ denote the sample mean of the covariate in treated and untreated subjects, whereas ${s^2}_{treatment}$ and ${s^2}_{control}$ denote the sample variance of the covariate in

treated and untreated subjects (Austin, 2011). For dichotomous variables, the standardized difference is defined as:

$$d = \frac{(\hat{p}_{treatment} - \hat{p}_{control})}{\sqrt{\frac{\hat{p}_{treatment}(1-\hat{p}_{treatment}) + \hat{p}_{control}(1-\hat{p}_{control})}{2}}}$$

where $\hat{p}_{treatment}$ and $\hat{p}_{control}$ denote the prevalence of mean of the dichotomous variable in treated and untreated subjects (Austin, 2011). The standardized difference compares the difference in means in units of pooled sample variance. Moreover, it ignores the influence of sample size and allows for comparison between variables measured in different units (Austin, 2011). Even though there is no universal agreement upon what threshold of standardized difference can be used to indicate significant imbalance, a standardized difference that is less than 0.1 is considered as a criteria indicating a negligible difference in the mean or prevalence of a covariate between treatment and control groups (Austin, 2011).

2.6 Propensity Score Estimation for Multiple Treatments

The method of propensity score has been mostly used for 2 treatment settings. However, there are many cases in clinical research that one might be interested in comparison of more than 2 treatments. For nominal treatments, Imbens proposed the use of generalized propensity score, defined as the conditional probability of receiving a particular level of the treatment given a set of measured baseline covariates (Imbens, 2000). The generalized propensity score can be estimated with a multinomial logistic or probit regression (Spreeuwenberg et al., 2010). It predicts the probability that an outcome (e.g., treatment assignments, Z_i) equals each of its possible values ($t = 1, \dots, M$) as a function of a linear combination of the covariates and their products and cross

products:

$$P(Z_i = t|X_i) = \frac{e^{\beta_t' X_i}}{1 + \sum_{t'=1}^{M-1} e^{\beta_{t'}' X_i}}, t = 1, \ldots, M-1,$$

$$P(Z_i = M|X_i) = \frac{1}{1 + \sum_{t'=1}^{M-1} e^{\beta_{t'}' X_i}}$$

where β_t are unknown coefficients and estimated from the data (McCaffrey, 2013). For each subject, the probability of receiving each treatment given the observed set of baseline covariates is estimated (Spreeuwenberg et al., 2010). Imbens proved that, the generalized propensity score, like the simple propensity score, is a balancing score which can be used for sample matching (Imbens, 2000). Therefore, the generalized propensity score can be used to balance the measured baseline differences and allows for valid estimates in multiple treatment comparisons (Spreeuwenberg et al., 2010).

Once generalized propensity score is estimated for each patient on M levels of treatment assignments, the matching is generalized from a one-dimenstional line to a $M-1$ dimentional space formed by propensity score vectors. The algorithm tries to search for clusters of subjects-one receiving each of treatments $t = 1, \ldots, M$-with the smallest within-cluster distance (Rassen et al., 2013). The goal is to find an optimal within-cluster distance to limit required computations for all possible combinations of subjects (Rassen et al., 2013). Pairwise standard differences are used for balance check for baseline covariates across treatment groups after matching (Austin, 2009).

3. MATERIALS & METHODS

3.1 Study Design

The present study utilized the Surveillance, Epidemiology, and End Results (SEER) database, a National Cancer Institute program that provides information on the incidence and survival rates of cancer in the United States and that includes population-based cancer registries covering approximately 28% of the population of the United States.

Eligible patients were those who were diagnosed with HNC between 2000 and 2013 and whose cancer diagnosis is the only one or the first during the patient's life time. This helps reduce the confounding effect from the treatment for previous cancer diagnosis. Furthermore, patients with unknown survival information were excluded from the study. From the SEER database, a total of 113,994 patients were diagnosed with HNC and 109 patients were excluded from the study because of history of multiple synchronous cancers. As a result, 113,885 patients (Hispanic, n = 8132; non-Hispanic White, n = 85995; non-Hispanic Black, n = 12669; non-Hispanic other, n = 7089) were included and analyzed for this study.

Basic demographics, socioeconomic status, characteristics of disease, survival information and underlying prognostic factors were examined, including age, race, gender, marital status, insurance status and type, education, unemployment rate, area of residence (rural or urban county), movement history (moved from outside US), smoking history, year of diagnosis (2000-2006 or 2007-2013), cancer-directed surgery status (accepted or not), site of cancer, tumor grade, tumor size and combined SEER summary stage. Continuous variables education, ever smoker,

moved from outside U.S., unemployed and urban were defined as the proportion of adults in the patient's area of residence that did not graduate high school, that used to smoke or currently smoking, that moved from outside U.S., that were unemployed, and that lived in urban area. All continuous variables were collapsed based on their quartiles and only collapsed variables were analyzed in the study. Sites of HNC were classified as larynx, oral cavity (lip, tongue, floor of mouth, gum and other mouth) and pharynx (nasopharynx, tonsil, oropharynx, hypopharynx). Due to the small sample size in certain levels in the "race and origin code" variable, levels of "non-Hispanic American Indian/Alaska Native", "non-Hispanic Asian or Pacific Islander" and "non-Hispanic unknown race" were collapsed in "non-Hispanic other".

The primary outcome of the study was overall survival, which was defined as months from date of registration in SEER to death or last follow-up. Secondary outcome was disease-specific survival, which was defined as months from date of registration in SEER to death of any types of head and neck cancers or last follow-up. Death due to other reason was censored at date of death.

3.2 Statistical Analysis

Statistical analysis was conducted using SAS version 9.4 and SAS macros or software developed at the Biostatistics and Bioinformatics at Winship Cancer Institute (Nickleach et al., 2013). Descriptive statistics for each variable were reported. The univariate association with racial subgroups (Hispanic, non-Hispanic White, non-Hispanic Black, non-Hispanic other) was assessed using the chi-square test for categorical covariates and analysis of variance for continuous covariates. The univariate association between each covariate including race groups

and study outcomes (overall survival and disease-specific survival) were assessed using Cox proportional hazards models and log-rank tests. A multivariable Cox proportional hazard model was fit by a backward variable selection method applying an alpha =.05 removal criteria in order to select covariates. Insurance status was removed from multivariate analysis due to missing information for all subjects who were diagnosed before 2007. Kaplan-Meier plots were produced to compare the survival curves by subgroups along with log-rank *P*-value. In addition, a multivariable model was fit to test for the impact of interaction between race and age on survival outcomes to explore how the impact of race on survival outcomes varies by different age groups. The survival information of each race/ethnicity subgroup stratified by age were estimated along with *P*-value. All *P*-values were two-sided, and $P < .05$ was considered statistically significant. One thing to note is that the *P*-values are very significant in this study due to large sample size. Thus, interpretation with focusing on clinical relevance is highly recommended.

Propensity score matching between Hispanic, non-Hispanic White, non-Hispanic Black and non-Hispanic other groups based on the estimated propensity score for each subject was performed to reduce heterogeneity in distributions of measured baseline covariates. A multinomial logistic regression model predicting Hispanic, non-Hispanic Black and non-Hispanic other groups, which together determine the propensity score of non-Hispanic White, was used to calculate propensity scores including the covariates age, gender, marital status, year of diagnosis, cancer-directed surgery status, site of cancer, tumor grade, tumor size, combined SEER summary stage, education, unemployment rate, area of residence, movement history, and smoking history.

Cases from the four race groups were matched to each other based on generalized propensity scores using a caliper-based greedy matching without replacement algorithm (Austin, 2011). The sums of within-cluster distance were calculated and ploted versus their corresponding matching ID. The caliper for the matching algorithm was identified as the caliper for the matching ID as to when the sum of within-cluster distance began increasing dramatically. Then the matching process was repeated for each of three types (laryngeal, oral cavity, and pharyngeal) of HNC. The effectiveness of the matching was evaluated by calculating the standardized differences of the covariates on the matched sample.

4. RESULTS

A total of 113885 patients were registered between 2000 and 2013 and met study entry criteria. 8132 (7.1%) patients were Hispanic (all races), 12669 (11.1%) were non-Hispanic Black, 85995 (75.5%) were non-Hispanic White, and 7089 (6.2%) were non-Hispanic other. A total of 3092 (2.7%) patients were 20-39 years of age, 47943 (42.1%) were 40-59 years of age, 52617 (46.2%) were 60-79 years of age, and 10233 (9.0%) were 80 and above years of age. The majority of patients had oral cavity cancer (n=49850;43.8%), while 31720 (27.9%) patients had laryngeal cancer and 32315 (28.4%) patients had pharyngeal cancer. The median tumor size was 2.7 cm (Table 1).

A comparison of patient, disease, and prognostic characteristics by race is presented in Table 2. Younger Hispanic population (20-39 years of age and 40-59 years of age) was more likely to be involved in this study compared with non-Hispanic White. Hispanic population was more likely to be not covered by insurance compared with non-Hispanic White. Furthermore, low education level and high unemployment rate were found in Hispanic population compared with non-Hispanic White. Finally, proportion of smokers in Hispanic population was significantly lower than that in non-Hispanic White population.

Factors associated with improved OS on univariate analysis were younger age, non-Hispanic other races, male gender, married status, being covered by insurance, later years of diagnosis, cancer-directed surgery received, lower tumor grade, smaller tumor size, localized and regional combined SEER summary stage, higher education level, lower unemployment rate, and lower

proportion of smokers in community (Table 3.1). In addition, the result showed that Hispanic population had a poorer OS compared with non-Hispanic White (HR, 1.07; 95% CI, 1.03-1.10; $P < 0.001$). However, for patients with laryngeal cancer, the OS for Hispanic population was better compared with non-Hispanic White (HR, 0.93; 95% CI, 0.87-0.99; $P = 0.015$; Table 3.2). In terms of the disease-specific survival, factors associated with improved performance were similar to that of OS except for the exclusion of non-Spanish-Hispanic-Latino population and larynx and oral cavity cancer sites (Table 3.3).

Factors that remained associated with improved OS on multivariate analysis were younger age, non-Hispanic other races, male gender, married status, later years of diagnosis, cancer-directed surgery received, pharynx cancer site, tumor grade II and III, smaller tumor size, localized combined SEER summary stage, higher education level, unemployment rate ≥ 1%, ≤ 5%, moved from outside U.S. > 5%, ≤ 10%, ever smoker ≥ 29%, ≤ 45% (Table 4.1). Again, a poorer OS was obtained for Hispanic population versus non-Hispanic White (HR, 1.06; 95% CI, 1.03-1.10; $P < 0.001$) while a better OS for Hispanic laryngeal cancer patients was found compared with non-Hispanic White laryngeal cancer patients (HR, 0.91; 95% CI, 0.85-0.97; $P = 0.004$; Table 4.2). For disease specific survival, similar prognostic factors of significance were found (Table 4.3). There was a significant interaction between age and race ($P < 0.001$). When the range of patient's age was in 40-59 or 60-79, OS for Hispanic population was significantly inferior to the for non-Hispanic White or non-Hispanic other (Table 5.1). The same interaction effect was also found in the multivariate analysis for disease-specific survival (Table 5.2).

The median OS for Hispanic population was 72 months, compared with 80 months for non-Hispanic White and 102 months for non-Hispanic other races ($P < 0.0001$; Figure 1.1). The overall 5-year survival rate was 53% for Hispanic population, 56% for non-Hispanic White, 38.2% for non-Hispanic Black and 59.2% for non-Hispanic other races. Similar trends were obtained from the survival analysis for oral cavity and pharyngeal cancer. A comparable survival trend was obtained for disease-specific survival ($P < 0.0001$; Figure 1.2). For oral cavity cancer, the 5-year survival rate was 52% for Hispanic patients, 56.8% for non-Hispanic White, 34.2% for non-Hispanic Black and 58.1 for non-Hispanic other races ($P < 0.0001$; Figure 1.3). And for pharyngeal cancer, 5-year survival rates of 51.8% for Hispanic, 56% for non-Hispanic White, 32% for non-Hispanic Black and 59.2% for non-Hispanic other were seen ($P < 0.0001$; Figure 1.4). However, for patients with laryngeal cancer, the median OS was 80 months for Hispanic population versus 73 months for non-Hispanic White ($P < 0.0001$; Figure 1.5). Furthermore, the 5-year survival rate for laryngeal cancer was also higher for Hispanic patients than that for non-Hispanic White, with 55.6% for Hispanic versus 54.6% for non-Hispanic White.

As demonstrated in Table 2, there was statistical difference in all covariates included in the multivariable models across the four groups. To adjust the unequal distributions of measured baseline covariates and to mimic the setting in a randomized study, a propensity score matching was performed to generate subgroups of 3230 patients from each group (Table 6). A caliper was applied for the matching algorithm to optimize the within-cluster distances and to limit the required computations (Figure 2.1). The result showed no difference in the distributions of some

important factors such as age, gender, marital status, cancer-directed surgery status, oral cavity and pharynx of cancer site, tumor grade, tumor size, combined SEER summary stage, education, and etc., in the four subgroups. Propensity score matching was also implemented for 3 major sites of HNC to balance the baseline covariates, with 969 laryngeal cancer patients, 1220 oral cavity cancer patients, and 1016 pharyngeal cancer patients from each subgroup. Calipers for each of the three matching algorithms were determined based on the plots of sums of within-cluster distance versus matching ID (Figure 2.2-2.4).

The impact of race on OS was investigated within the propensity score matched sample (Table 7.1). A poorer OS for Hispanic patients versus non-Hispanic White patients was obtained, which was consistent with the result from multivariate analysis before matching (HR, 1.09; 95% CI, 1.01-1.16; $P = 0.021$). The result of disease-specific survival analysis for Hispanic patients compared with non-Hispanic White patients was similar to that of OS (HR, 1.13; 95% CI, 1.04-1.23; $P = 0.005$; Table 7.2). The median OS was 64 months for Hispanic patients versus 78 months for non-Hispanic White patients, and the 5-year survival rate was 51% for Hispanic patients and was 55.6% for non-Hispanic White ($P < 0.0001$; Figure 3.1). A similar trend was seen in terms of the disease-specific survival, with 5-year survival rate of 60.5% for Hispanic patients and 65.3% for non-Hispanic White patients ($P < 0.0001$; Figure 3.2). In terms of oral cavity ($P = 0.0001$; Figure 3.3) and pharyngeal cancer ($P < 0.0001$; Figure 3.4), poor median OS (oral cavity cancer, 63 months; pharyngeal cancer, 64 months) and 5-year survival rate (oral cavity cancer, 51%; pharyngeal cancer, 64%) were found for Hispanic population. In contrast, for

laryngeal cancer patients, OS for Hispanic population was found to be better than that for non-Hispanic White (HR, 0.84; 95% CI, 0.74-0.96; $P = 0.01$; Table 7.3). Similarly, the median OS was higher for Hispanic laryngeal cancer patients than that for non-Hispanic White laryngeal cancer patients, with 91 months for Hispanic and 71 months for non-Hispanic White ($P < 0.0001$; Figure 3.5).

5. DISCUSSION

The purpose of this analysis was to compare overall survival and disease-specific survival between Hispanic (all races) and non-Hispanic HNC patients in the United States. Another purpose of this study was to identify prognostic factors that might affect the survival of HNC patients. Although the multidisciplinary care of combined surgery and radiotherapy along with rehabilitation has gained significant improvement in patient's prognosis and survival time, its accessibility and reliability for minor races/ethnicities such as Hispanic patients remain unknown due to the lack of relevant clinical evidences. Given this uncertainty, it is necessary to compare the survival between Hispanic HNC patients and the major race, non-Hispanic White.

It was hypothesized that the survival for Hispanic HNC patients would vary significantly from the general population in the U.S. given the socioeconomic and other demographic disparities in Hispanic population. In the present study, survival data from the SEER database was utilized and it was demonstrated that the median survival and the 5-year survival rate for Hispanic HNC patients were significantly lower than those for non-Hispanic White and for non-Hispanic other races. Furthermore, the trends for survival data for each cancer group were also collected and analyzed. In respect to the oral cavity and the pharyngeal cancers, Hispanic patients demonstrated poorer OS as well as disease-specific survival than those for non-Hispanic White patients. These results coincided with our hypothesis and were consistent with literature which states that racial disparities result in inferior OS in patients with HNC (Shin & Truong, 2015). However, one surprising finding was that, for laryngeal cancer, Hispanic patients showed better

OS than that for non-Hispanic White patients, even though this superiority disappeared when comparing disease-specific survival between two racial groups. This illustrates that the Hispanic population outperforms the non-Hispanic White population in some specific prognostic factors such as tobacco use, especially smoking.

A multivariate analysis was implemented to identify the impact of important prognostic factors on both OS and disease-specific survival in Hispanic and non-Hispanic patients. It was found that Hispanic race is associated with inferior OS and disease-specific survival in patients with HNC. Hispanic patients were associated with relatively lower prevalence of being married, lower insurance cover rate, advanced tumor size, lower education, and higher unemployment rate, all of which were predictors of inferior OS and disease-specific survival in both univariate analyses and multivariable models. These socioeconomic and demographic factors account for the poor survival in Hispanic HNC patients. Besides, elder age, non-Hispanic Black race, advanced tumor grade, and higher prevalence of smoking were independent prognostic factors for worse survival rate, which have been published by previous study (Mehanna et al., 2010b).

Apart from the prognostic factors that have been discovered, the interaction between race and age has been proved to have significant impact on two survival outcomes (OS and disease-specific survival). It can be seen that among patients with 40-59 and 60-79 years of age, Hispanic race had significant poorer survival versus non-Hispanic White and non-Hispanic other races. However, this significant effect of interaction no longer existed in the youngest group (20-39 years of age) and in the oldest group (80+ years of age). For the youngest group, having a better

prognosis of HNC was conjectured based on the finding. On the other hand, the fact that elderly Hispanic patients were less likely to be enrolled in the study compared with non-Hispanic White and non-Hispanic other races accounted for the absence of significant interaction in the oldest group.

The results of propensity score matching based on complete sample and samples for each cancer site further illustrate that Hispanic HNC patients have poorer OS and disease-specific survival when comparing with non-Hispanic White or non-Hispanic other races. Hispanic patients presented lower hazard ratio in OS compared with non-Hispanic White for laryngeal cancer and higher hazard ratio in disease-specific survival for pharyngeal cancer. However, unlike the hazard ratio *P*-values in multivariate analyses before propensity score matching, there were many *P*-values calculated based on matched sample being not statistically significant as sample size of each racial sub-group decreased. From Kaplan-Meier curves generated from matched sample, it can be seen that the survival of Hispanic patients fell behind that of non-Hispanic White patients during the earlier phase of prognosis while it caught up during the later phase. This "catching-up" property of survival in Hispanic patients can be found in both OS and disease-specific survival Kaplan-Meier curves. A pathological reason for the presence of the "catching-up" property needs to be explored in future studies.

5.1 Implications

The results of this study illustrate that OS and disease-specific survival in Hispanic HNC patients were inferior to those in the general population in the U.S. The prognostic factors that

resulted in higher hazard ratio in Hispanic patients' survival reflect a variety of disparities in socioeconomic status and accessibility to public health well being among Hispanic population. These results also support the revision of standard treatment for HNC in Hispanic population and the improvement of public health welfare such as insurance coverage for Hispanic population in the U.S. These findings also identify the impact of particular behavior such as smoking on the prognosis of HNC, which can help with policy making in public health.

5.2 Strengths and Limitations

This study included data from a population-based cancer registry, which can be generalized to the entire U.S. population. Furthermore, the implementation of propensity score matching which allows for balanced baseline covariates between subgroups censored some drawbacks of a retrospective study and made the direct comparison of the effect of race on survival outcomes feasible. Despite the strengths, there are also some limitations and defects to the study. For example, there are still some measured baseline covariates that did not achieve a satisfactory balance between subgroups in the matched sample. This would make imprecision estimates for the impact of race on survival outcomes. However, the results of this analysis can still be considered valid because most covariates with levels were balanced perfectly. In addition, inadequate tumor grade information made it impossible to use a traditional TNM stage in the analysis. A combined SEER summary stage was used instead. Furthermore, insurance coverage was taken out from the multivariable models due to limited insurance coverage information

before 2007. These two are both important prognostic factors which could affect survival outcomes significantly.

5.3 Recommendations

These results can provide information on future studies regarding the survival of Hispanic HNC patients in the U.S. and the impact of prognostic factors. The implementation of prospective randomized studies is recommended to further validate the results of this study and to investigate more precise estimates for the impact of race on OS and disease-specific survival of HNC patients. In terms of the utilization of propensity score methods, a dataset with more comprehensive information at all longitudinal points could be used to further eliminate confounding from analysis. Additionally, more interaction effects between race and other baseline covariates such as gender, marital status, and year of diagnosis could be included in multivariate analysis in future studies to explore the effect of race on survival outcomes at different levels.

REFERENCES

American Cancer Society. (2017). Cancer Facts & Figures 2017. from www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2017/cancer-facts-and-figures-2017.pdf

Austin, P. C. (2009). Balance Diagnostics for Comparing the Distribution of Baseline Covariates Between Treatment Groups in Propensity-score Matched Samples. *Statistics in Medicine, 28(25),* 3083-3107.

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavior Research, 46,* 399-424.

Austin, P. C. (2012). The performance of different propensity score methods for estimating marginal hazard ratios. *Statistics in Medicine, 32,* 2837-2849.

Bhalavat, R. I., Fakih, A. R., Mistry, R. C. & Mahantshetty, U. (2003). Radical radiation vs surgery plus post-operative radiation in advanced (resectable) supraglottic larynx and pyriform sinus cancers: a prospective randomized study. European Journal of Surgical Oncology, 29, 750-756.

Bolye, P. & Levin B. (2008). World Cancer Report. International Agency for Research on Cancer.

Ennis, S. R., Rios-Vargas, M. & Albert, N. G. The Hispanic population: 2010. From www.census.gov/prod/cen2010/briefs/c2010br-04.pdf

Flury, B. K., & Riedwyl, H. (1986). Standard distance in univariate and multivariate analysis. *The American Statistician, 40,* 249–251.

Gatta, G., Botta, L., Sanchez, M. J., Anderson, L. A., Pierannunzio, D., Licitra, L. & the EUROCARE Working Group. (2015). Prognoses and improvement for head and neck cancers diagnosed in Europe in early 2000s: The EUROCARE-5 population-based study. *European Journal of Cancer, 51*, 2130-2143.

Goodwin, W. J., Thomas, G. R., Parker, D. F., Joseph, D., Levis, S., Franzmann, E., … Hu, J. J. Unequal Burden of Head and Neck Cancer in the United States. *Wiley InterScience, 30*, 358-371.

Gu, X. S. & Rosenbaum, P. R. (1993). Comparison of multivariate matching methods: Structures, distances, and algorithms. *Journal of Computational and Graphical Statistics, 2,* 405-420.

Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. Biometrika, 87, 706–710.

Intercollegiate Guidelines Network. (2006). Head and neck cancers management. from www.sign.ac.uk/guidelines/fulltext/90/index.html

Joffe, M. M., Ten Have, T. R., Feldman, H. I., & Kimmel, S. E. (2004). Model selection, confounder control, and marginal structural models: Review and new applications. *The American Statistician, 58,* 272–279.

Marron, M., Boffetta, P., Zhang, Z. F., Zaridze, D., Wunsch-Filho, V., Winn, D. M. & Hashibe, M. (2010). Cessation of Alcohol Drink, Tobacco Smoking and the Reversal of Head and Neck Cancer Risk. *International Journal of Epidemiology, 39*, 182-196.

McCaffrey, D. F., Griffin, B. A., Almirall, B. A., Slaughter, M. E., Ramchand, R. & Burgette, L. F. (2013). A Tutorial on Propensity Score Estimation for Multiple Treatments Using Generalized Boosted Models. *Statistics in Medicine, 32(19),* 3388-3414.

Mehanna, H., West, C. M. L., Nutting, C. & Paleri, V. (2010a). Head and Neck Cancer-Part 2: Treatment and Prognostic Factors. *BMJ, 341*, c4690.

Mehanna, H., Paleri, V., West, C. M. L., & Nutting, C. (2010b). Head and Neck Cancer-Part 1: Epidemiology, Presentation and Prevention. *BMJ, 340*, c4684.

Murthy, V. H., Krumholz, H. M. & Gross, C. P. (2004). Participation in Cancer Clinical Trials, Race-, Sex-, and Age-Based Disparities. *JAMA, 291*(22), 2720-2726.

National Institute for Health and Clinical Excellence. (2004). Improving Outcomes in Head and Neck Cancers- the manual. from www.guidance.nice.org.uk/csghn/guidance/pdf/English

Nickleach, D., Liu, Y., Shrewsberry, A., Ogan, K., Kim, S. & Wang, Z. (2013). SAS® Macros to Conduct Common Biostatistical Analyses and Generate Reports. *SESUG 2013: The Proceeding of the SouthEast SAS User Group*. from http://analytics.ncsu.edu/sesug/2013/PO-05.pdf.

Oxford Cancer Intelligence Unit. (2010). Profile of Head and Neck Cancers in England: Incidence, Mortality and Survival. from www.ncin.org.uk/cancer_type_and_topic_specific_work/cancer_type_specific_work/head_and_neck_cancers/head_and_neck_cancer_hub/resources

Rao, D. N., Shroff, P. D., Chattopadhyay, G. & Dinshaw, K. A. (1998). Survival analysis of 5595 head and neck cancers- results of conventional treatment in a high-risk population. *British Journal of Cancer, 77(9)*, 1514-1518.

Rassen, J. A., Shelat, A. A., Franklin, J. M., Glynn, R. J., Solomon, D. H. & Schneeweiss, S. (2013). Matching by Propensity Score in Cohorts Studies with Three Treatment Groups. *Epidemiology, 24(3),* 401-409.

Rosenbaum, P. R. & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika, 70(1),* 41-55.

Rosenbaum, P. R. & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician, 39,* 33-38.

Rosenbaum, P. R. (1987). Model-Based Direct Adjustment. *The Journal of the American Statistician, 82,* 387–394.

Rosenbaum, P. R. (2002). *Observational Studies* (2nd ed.). New York, NY: Springer-Verlag.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol, 66,* 688 –701.

Shavers, V. L., Harlan, L. C., Winn, D. & Davis, W. W. (2003). Racial/ethnic patterns of care for cancers of the oral cavity, pharynx, larynx, sinuses, and salivary glands. *Cancer Metastasis Rev, 22(1),* 25-38.

Shin, J. Y. & Truong, M. T. (2015). Racial Disparities in Laryngeal Cancer Treatment and Outcome: A Population-Based Analysis of 24,069 Patients. *The Laryngoscope, 125,* 1667-1674.

Spreeuwenberg, M. D., Bartak, A., Croon, M. A., Hagenaars, J. A., Busschbach, J. J., Andrea, H., … Stijnen, T. (2010). The multiple propensity score as control for bias in the comparison of more than two treatment arms: an introduction from a case study in mental health. *Medical Care, 48(2),* 166–174.

World Health Organization. (2008). The Global Burden of Disease: 2004 Update. from www.who.int/evidence/bod

Appendix: TABLES AND FIGURES

Table 1 Patient and Treatment Characteristics

Demographic or Prognostic Characteristics	Level	N = 113885 N(%)
Year of diagnosis	2000-2006	53691 (47.1)
	2007-2013	60194 (52.9)
Cancer site	Larynx	31720 (27.9)
	Oral cavity	49850 (43.8)
	Pharynx	32315 (28.4)
Race	Hispanic (All Races)	8132 (7.1)
	Non-Hispanic Black	12669 (11.1)
	Non-Hispanic Others	7089 (6.2)
	Non-Hispanic White	85995 (75.5)
Origin recode NHIA (Hispanic, Non-Hisp)	Non-Spanish-Hispanic-Latino	105753 (92.9)
	Spanish-Hispanic-Latino	8132 (7.1)
Sex	Female	27723 (24.3)
	Male	86162 (75.7)
Age	20-39 years	3092 (2.7)
	40-59 years	47943 (42.1)
	60-79 years	52617 (46.2)
	80+ years	10233 (9.0)
Tumor grade	Moderately differentiated; Grade II	45915 (40.3)
	Poorly differentiated; Grade III	28173 (24.7)
	Undifferentiated; anaplastic; Grade IV	1441 (1.3)
	Unknown	23626 (20.7)
	Well differentiated; Grade I	14730 (12.9)
Marital status at diagnosis	Married	58812 (51.6)
	Other	55073 (48.4)

Demographic or Prognostic Characteristics	Level	N = 113885 N(%)
Insurance status	Any Medicaid	9002 (15.0)
	Insurance status unknown	2688 (4.5)
	Insured	35700 (59.3)
	Insured/No specifics	9620 (16.0)
	Uninsured	3184 (5.3)
	Missing	53691
Combined SEER summary stage	Distant	18885 (16.6)
	Localized	41940 (36.8)
	Regional	48222 (42.3)
	Unknown	4838 (4.2)
Cancer-directed surgery accepted	No	57882 (50.8)
	Unknown	1194 (1.0)
	Yes	54809 (48.1)
CS tumor size in cm	>=0, <=2	13440 (11.8)
	>2, <=3	13852 (12.2)
	>3, <=4	15097 (13.3)
	>4, <=99	10672 (9.4)
	Unknown	60824 (53.4)
% Ever Smoker (males age 18+)	>=29, <=45	28500 (25.0)
	>45, <=51	29320 (25.7)
	>51, <=56	28012 (24.6)
	>56, <=73	28040 (24.6)
	Unknown	13 (0.0)
% < High school education	>=3, <=16	28695 (25.2)
	>16, <=19	28500 (25.0)
	>19, <=27	28810 (25.3)
	>27, <=51	27867 (24.5)
	Unknown	13 (0.0)
% Moved from outside US	>=0, <=2	28698 (25.2)
	>2, <=3	28399 (24.9)
	>3, <=5	29194 (25.6)
	>5, <=10	27581 (24.2)
	Unknown	13 (0.0)
% Unemployed	>=1, <=5	26621 (23.4)
	>5, <=6	33112 (29.1)
	>6, <=9	32733 (28.7)
	>9, <=18	21406 (18.8)
	Unknown	13 (0.0)

Demographic or Prognostic Characteristics	Level	N = 113885 N(%)
% Urban	>=0, <=76	28578 (25.1)
	>76, <=95	29757 (26.1)
	>95, <=99	27285 (24.0)
	>99, <=100	28252 (24.8)
	Unknown	13 (0.0)

Table 2 Univariate Association with Race

Demographic or Prognostic Characteristic	Level	Hispanic (All Races) N=8132 N(%)	Non-Hispanic Black N=12669 N(%)	Non-Hispanic Others N=7089 N(%)	Non-Hispanic White N=85995 N(%)	Parametric P-value*
Age	20-39 years	412 (5.07)	296 (2.34)	466 (6.57)	1918 (2.23)	**<.001**
	40-59 years	3411 (41.95)	6444 (50.86)	2881 (40.64)	35207 (40.94)	
	60-79 years	3656 (44.96)	5368 (42.37)	3017 (42.56)	40576 (47.18)	
	80+ years	653 (8.03)	561 (4.43)	725 (10.23)	8294 (9.64)	
Sex	Female	1799 (22.12)	2901 (22.9)	2008 (28.33)	21015 (24.44)	**<.001**
	Male	6333 (77.88)	9768 (77.1)	5081 (71.67)	64980 (75.56)	
Marital status at diagnosis	Married	4188 (51.5)	4040 (31.89)	4211 (59.4)	46373 (53.93)	**<.001**
	Other	3944 (48.5)	8629 (68.11)	2878 (40.6)	39622 (46.07)	
Insurance status	Any Medicaid	1147 (24.95)	1896 (29.96)	751 (18.75)	5208 (11.51)	**<.001**
	unknown	190 (4.13)	177 (2.8)	482 (12.03)	1839 (4.06)	
	Insured	2206 (47.98)	2621 (41.42)	2103 (52.5)	28770 (63.56)	
	Insured/No specifics	694 (15.09)	1059 (16.74)	513 (12.81)	7354 (16.25)	
	Uninsured	361 (7.85)	575 (9.09)	157 (3.92)	2091 (4.62)	
Year of diagnosis	2000-2006	3534 (43.46)	6341 (50.05)	3083 (43.49)	40733 (47.37)	**<.001**
	2007-2013	4598 (56.54)	6328 (49.95)	4006 (56.51)	45262 (52.63)	
Cancer-directed surgery accepted	No	4123 (50.7)	8102 (63.95)	3697 (52.15)	41960 (48.79)	**<.001**
	Unknown	71 (0.87)	146 (1.15)	74 (1.04)	903 (1.05)	
	Yes	3938 (48.43)	4421 (34.9)	3318 (46.8)	43132 (50.16)	
Cancer site	Larynx	2461 (30.26)	4850 (38.28)	1272 (17.94)	23137 (26.91)	**<.001**
	Oral cavity	3335 (41.01)	3822 (30.17)	3196 (45.08)	39497 (45.93)	
	Pharynx	2336 (28.73)	3997 (31.55)	2621 (36.97)	23361 (27.17)	
Tumor grade	Grade II	3257 (40.05)	5768 (45.53)	2503 (35.31)	34387 (39.99)	**<.001**
	Grade III	2120 (26.07)	3097 (24.45)	1791 (25.26)	21165 (24.61)	
	Grade IV	92 (1.13)	151 (1.19)	364 (5.13)	834 (0.97)	
	Unknown	1521 (18.7)	2610 (20.6)	1514 (21.36)	17981 (20.91)	
	Grade I	1142 (14.04)	1043 (8.23)	917 (12.94)	11628 (13.52)	

Demographic or Prognostic Characteristic	Level	Hispanic (All Races) N=8132 N(%)	Non-Hispanic Black N=12669 N(%)	Non-Hispanic Others N=7089 N(%)	Non-Hispanic White N=85995 N(%)	Parametric P-value*
CS tumor size in cm (quartile)	>=0, <=2	834 (10.26)	722 (5.7)	875 (12.34)	11009 (12.8)	**<.001**
	>2, <=3	995 (12.24)	1173 (9.26)	877 (12.37)	10807 (12.57)	
	>3, <=4	1140 (14.02)	1682 (13.28)	930 (13.12)	11345 (13.19)	
	>4, <=99	953 (11.72)	1722 (13.59)	702 (9.9)	7295 (8.48)	
	Unknown	4210 (51.77)	7370 (58.17)	3705 (52.26)	45539 (52.96)	
Combined SEER summary stage	Distant	1568 (19.28)	3293 (25.99)	1342 (18.93)	12682 (14.75)	**<.001**
	Localized	2847 (35.01)	3405 (26.88)	2510 (35.41)	33178 (38.58)	
	Regional	3373 (41.48)	5554 (43.84)	2749 (38.78)	36546 (42.5)	
	Unknown	344 (4.23)	417 (3.29)	488 (6.88)	3589 (4.17)	
% < High school education (quartile)	>=3, <=16	987 (12.14)	1450 (11.45)	2285 (32.23)	23973 (27.88)	**<.001**
	>16, <=19	1780 (21.89)	2788 (22.01)	2034 (28.69)	21898 (25.46)	
	>19, <=27	2036 (25.04)	4957 (39.13)	1085 (15.31)	20732 (24.11)	
	>27, <=51	3329 (40.94)	3474 (27.42)	1681 (23.71)	19383 (22.54)	
	Unknown	0 (0)	0 (0)	4 (0.06)	9 (0.01)	
% Unemployed (quartile)	>=1, <=5	850 (10.45)	1176 (9.28)	1260 (17.77)	23335 (27.14)	**<.001**
	>5, <=6	2120 (26.07)	2894 (22.84)	2186 (30.84)	25912 (30.13)	
	>6, <=9	3230 (39.72)	4076 (32.17)	2748 (38.76)	22679 (26.37)	
	>9, <=18	1932 (23.76)	4523 (35.7)	891 (12.57)	14060 (16.35)	
	Unknown	0 (0)	0 (0)	4 (0.06)	9 (0.01)	
% Urban (quartile)	>=0, <=76	513 (6.31)	2526 (19.94)	632 (8.92)	24907 (28.96)	**<.001**
	>76, <=95	2335 (28.71)	2656 (20.96)	1280 (18.06)	23486 (27.31)	
	>95, <=99	1926 (23.68)	2490 (19.65)	2473 (34.89)	20396 (23.72)	
	>99, <=100	3358 (41.29)	4997 (39.44)	2700 (38.09)	17197 (20)	
	Unknown	0 (0)	0 (0)	4 (0.06)	9 (0.01)	
% Moved from outside US (quartile)	>=0, <=2	547 (6.73)	3303 (26.07)	376 (5.3)	24472 (28.46)	**<.001**
	>2, <=3	1498 (18.42)	3579 (28.25)	1049 (14.8)	22273 (25.9)	
	>3, <=5	2285 (28.1)	2015 (15.9)	2509 (35.39)	22385 (26.03)	
	>5, <=10	3802 (46.75)	3772 (29.77)	3151 (44.45)	16856 (19.6)	
	Unknown	0 (0)	0 (0)	4 (0.06)	9 (0.01)	

Demographic or Prognostic Characteristic	Level	Hispanic (All Races) N=8132 N(%)	Non-Hispanic Black N=12669 N(%)	Non-Hispanic Others N=7089 N(%)	Non-Hispanic White N=85995 N(%)	Parametric P-value*
% Ever Smoker	>=29, <=45	3876 (47.66)	3050 (24.07)	2888 (40.74)	18686 (21.73)	**<.001**
(males age 18+)	>45, <=51	2607 (32.06)	2632 (20.78)	2062 (29.09)	22019 (25.6)	
(quartile)	>51, <=56	1266 (15.57)	3262 (25.75)	1736 (24.49)	21748 (25.29)	
	>56, <=73	383 (4.71)	3725 (29.4)	399 (5.63)	23533 (27.37)	
	Unknown	0 (0)	0 (0)	4 (0.06)	9 (0.01)	

Table 3.1 Univariate Association with Overall Survival

			Survival months		
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Age	40-59 years	47943	1.72 (1.60-1.85)	**<.001**	**<.001**
	60-79 years	52617	2.65 (2.47-2.84)	**<.001**	
	80+ years	10233	5.14 (4.77-5.53)	**<.001**	
	20-39 years	3092	-	**-**	
Race	Hispanic (All Races)	8132	1.07 (1.03-1.10)	**<.001**	**<.001**
	Non-Hispanic Black	12669	1.63 (1.59-1.67)	**<.001**	
	Non-Hispanic Others	7089	0.87 (0.84-0.91)	**<.001**	
	Non-Hispanic White	85995	-	**-**	
Origin recode NHIA (Hispanic, Non-Hisp)	Non-Spanish-Hispanic-Latino	105753	0.99 (0.96-1.02)	0.592	0.591
	Spanish-Hispanic-Latino	8132	-	**-**	
Sex	Female	27723	1.10 (1.08-1.12)	**<.001**	**<.001**
	Male	86162	-	**-**	
Marital status at diagnosis	Married	58812	0.60 (0.59-0.61)	**<.001**	**<.001**
	Other	55073	-	**-**	
Insurance status	Any Medicaid	9002	1.28 (1.20-1.37)	**<.001**	**<.001**
	Insurance status unknown	2688	0.78 (0.72-0.86)	**<.001**	
	Insured	35700	0.62 (0.58-0.66)	**<.001**	
	Insured/No specifics	9620	0.90 (0.85-0.96)	**0.002**	
	Uninsured	3184	-	**-**	
Year of diagnosis	2000-2006	53691	1.14 (1.11-1.16)	**<.001**	**<.001**
	2007-2013	60194	-	**-**	
Cancer-directed surgery accepted	No	57882	1.74 (1.71-1.77)	**<.001**	**<.001**
	Unknown	1194	1.75 (1.62-1.89)	**<.001**	
	Yes	54809	-	**-**	

Cancer site	Larynx	31720	1.00 (0.98-1.02)	0.912	**<.001**
	Oral cavity	49850	0.95 (0.93-0.97)	**<.001**	
	Pharynx	32315	-	**-**	
Tumor grade	Moderately differentiated; Grade II	45915	1.37 (1.34-1.41)	**<.001**	**<.001**
	Poorly differentiated; Grade III	28173	1.35 (1.31-1.39)	**<.001**	
	Undifferentiated; anaplastic; Grade IV	1441	1.02 (0.94-1.11)	0.583	
	Unknown	23626	1.27 (1.23-1.31)	**<.001**	
	Well differentiated; Grade I	14730	-	**-**	
CS tumor size in cm (quartile)	>2, <=3	13852	1.64 (1.57-1.72)	**<.001**	**<.001**
	>3, <=4	15097	2.16 (2.07-2.25)	**<.001**	
	>4, <=99	10672	3.46 (3.31-3.61)	**<.001**	
	Unknown	60824	2.11 (2.03-2.19)	**<.001**	
	>=0, <=2	13440	-	**-**	
Combined SEER summary stage	Distant	18885	1.59 (1.53-1.66)	**<.001**	**<.001**
	Localized	41940	0.55 (0.53-0.57)	**<.001**	
	Regional	48222	0.87 (0.84-0.91)	**<.001**	
	Unknown	4838	-	**-**	
% < High school education (quartile)	>16, <=19	28500	1.12 (1.09-1.14)	**<.001**	**<.001**
	>19, <=27	28810	1.24 (1.21-1.27)	**<.001**	
	>27, <=51	27867	1.26 (1.23-1.29)	**<.001**	
	Unknown	13	0.88 (0.37-2.11)	0.776	
	>=3, <=16	28695	-	**-**	
% Unemployed (quartile)	>5, <=6	33112	1.07 (1.05-1.10)	**<.001**	**<.001**
	>6, <=9	32733	1.18 (1.15-1.21)	**<.001**	
	>9, <=18	21406	1.32 (1.28-1.35)	**<.001**	
	Unknown	13	0.86 (0.36-2.07)	0.741	
	>=1, <=5	26621	-	**-**	
% Urban (quartile)	>76, <=95	29757	0.97 (0.94-0.99)	**0.006**	**<.001**
	>95, <=99	27285	0.93 (0.91-0.95)	**<.001**	
	>99, <=100	28252	1.03 (1.01-1.06)	**0.006**	
	Unknown	13	0.75 (0.31-1.81)	0.523	

	>=0, <=76	28578	-	**-**	
% Moved from	>2, <=3	28399	0.99 (0.96-1.01)	0.219	**<.001**
outside US	>3, <=5	29194	0.90 (0.88-0.92)	**<.001**	
(quartile)	>5, <=10	27581	0.96 (0.94-0.99)	**0.002**	
	Unknown	13	0.74 (0.31-1.77)	0.492	
	>=0, <=2	28698	-	**-**	
% Ever Smoker	>45, <=51	29320	1.00 (0.97-1.02)	0.770	**<.001**
(males age 18+)	>51, <=56	28012	1.07 (1.04-1.09)	**<.001**	
(quartile)	>56, <=73	28040	1.11 (1.08-1.13)	**<.001**	
	Unknown	13	0.80 (0.33-1.91)	0.611	
	>=29, <=45	28500	-	**-**	

Table 3.2 Univariate Association with Overall Survival for Laryngeal Cancer

			Survival months		
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Age	40-59 years	11370	2.93 (2.33-3.68)	**<.001**	**<.001**
	60-79 years	17132	3.93 (3.13-4.93)	**<.001**	
	80+ years	2812	7.37 (5.85-9.28)	**<.001**	
	20-39 years	406	-	**-**	
Race	Hispanic (All Races)	2461	0.93 (0.87-0.99)	**0.015**	**<.001**
	Non-Hispanic Black	4850	1.26 (1.21-1.31)	**<.001**	
	Non-Hispanic Others	1272	0.80 (0.73-0.87)	**<.001**	
	Non-Hispanic White	23137	-	**-**	
Origin recode NHIA (Hispanic, Non-Hisp)	Non-Spanish-Hispanic-Latino	29259	1.11 (1.05-1.18)	**<.001**	**<.001**
	Spanish-Hispanic-Latino	2461	-	**-**	
Sex	Female	6017	1.06 (1.02-1.10)	**0.007**	**0.006**
	Male	25703	-	**-**	
Marital status at diagnosis	Married	16499	0.64 (0.62-0.67)	**<.001**	**<.001**
	Other	15221	-	**-**	
Insurance status	Any Medicaid	2863	1.28 (1.14-1.45)	**<.001**	**<.001**
	Insurance status unknown	411	0.85 (0.69-1.04)	0.116	
	Insured	8718	0.74 (0.66-0.83)	**<.001**	
	Insured/No specifics	2798	0.99 (0.88-1.13)	0.924	
	Uninsured	899	-	**-**	
Year of diagnosis	2000-2006	16031	1.02 (0.98-1.05)	0.308	0.304
	2007-2013	15689	-	**-**	
Cancer-directed surgery accepted	No	19020	1.33 (1.29-1.38)	**<.001**	**<.001**
	Unknown	413	1.26 (1.11-1.44)	**<.001**	
	Yes	12287	-	**-**	

			Survival months		
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Tumor grade	Moderately differentiated; Grade II	15051	1.44 (1.37-1.52)	**<.001**	**<.001**
	Poorly differentiated; Grade III	5685	2.06 (1.95-2.19)	**<.001**	
	Undifferentiated; anaplastic; Grade IV	146	1.94 (1.56-2.41)	**<.001**	
	Unknown	6431	1.37 (1.29-1.45)	**<.001**	
	Well differentiated; Grade I	4407	-	**-**	
CS tumor size in cm (quartile)	>2, <=3	2330	1.84 (1.67-2.03)	**<.001**	**<.001**
	>3, <=4	2443	2.24 (2.04-2.46)	**<.001**	
	>4, <=99	1572	2.89 (2.62-3.19)	**<.001**	
	Unknown	23049	1.46 (1.35-1.57)	**<.001**	
	>=0, <=2	2326	-	**-**	
Combined SEER summary stage	Distant	5740	1.81 (1.66-1.97)	**<.001**	**<.001**
	Localized	18147	0.65 (0.60-0.71)	**<.001**	
	Regional	6692	1.38 (1.27-1.50)	**<.001**	
	Unknown	1141	-	**-**	
% < High school education (quartile)	>16, <=19	7660	1.08 (1.03-1.13)	**0.002**	**<.001**
	>19, <=27	8515	1.12 (1.07-1.17)	**<.001**	
	>27, <=51	8437	1.15 (1.10-1.20)	**<.001**	
	Unknown	3	2.22 (0.31-15.73)	0.426	
	>=3, <=16	7105	-	**-**	
% Unemployed (quartile)	>5, <=6	9040	1.06 (1.01-1.11)	**0.014**	**<.001**
	>6, <=9	9099	1.10 (1.05-1.15)	**<.001**	
	>9, <=18	6532	1.22 (1.17-1.28)	**<.001**	
	Unknown	3	2.22 (0.31-15.72)	0.426	
	>=1, <=5	7046	-	**-**	
% Urban (quartile)	>76, <=95	7892	1.00 (0.96-1.04)	0.897	0.650
	>95, <=99	6948	1.01 (0.96-1.05)	0.734	
	>99, <=100	7832	1.03 (0.98-1.07)	0.246	
	Unknown	3	2.05 (0.29-14.53)	0.473	
	>=0, <=76	9045	-	**-**	

Covariate	Level	N	Survival months Hazard Ratio (95% CI)	HR P-value	Log-rank P-value
% Moved from outside US (quartile)	>2, <=3	8112	1.03 (0.98-1.07)	0.233	0.442
	>3, <=5	7204	0.99 (0.94-1.03)	0.534	
	>5, <=10	7073	1.01 (0.97-1.05)	0.673	
	Unknown	3	2.04 (0.29-14.50)	0.474	
	>=0, <=2	9328	-	-	
% Ever Smoker (males age 18+) (quartile)	>45, <=51	7549	0.96 (0.92-1.01)	0.129	0.077
	>51, <=56	8140	1.00 (0.95-1.04)	0.877	
	>56, <=73	9268	1.03 (0.98-1.07)	0.267	
	Unknown	3	2.03 (0.29-14.40)	0.479	
	>=29, <=45	6760	-	-	

Table 3.3 Univariate Association with Disease-Specific Survival

Covariate	Level	N	Survival months Hazard Ratio (95% CI)	HR P-value	Log-rank P-value
Age	40-59 years	47943	1.44 (1.34-1.56)	**<.001**	**<.001**
	60-79 years	52617	1.84 (1.70-1.99)	**<.001**	
	80+ years	10233	2.95 (2.72-3.21)	**<.001**	
	20-39 years	3092	-	**-**	
Race	Hispanic (All Races)	8132	1.18 (1.13-1.22)	**<.001**	**<.001**
	Non-Hispanic Black	12669	1.83 (1.78-1.88)	**<.001**	
	Non-Hispanic Others	7089	0.99 (0.95-1.04)	0.732	
	Non-Hispanic White	85995	-	**-**	
Origin recode NHIA (Hispanic, Non-Hisp)	Non-Spanish-Hispanic-Latino	105753	0.92 (0.89-0.96)	**<.001**	**<.001**
	Spanish-Hispanic-Latino	8132	-	**-**	
Sex	Female	27723	1.10 (1.08-1.13)	**<.001**	**<.001**
	Male	86162	-	**-**	
Marital status at diagnosis	Married	58812	0.59 (0.58-0.60)	**<.001**	**<.001**
	Other	55073	-	**-**	
Insurance status	Any Medicaid	9002	1.20 (1.12-1.29)	**<.001**	**<.001**
	Insurance status unknown	2688	0.69 (0.62-0.76)	**<.001**	
	Insured	35700	0.55 (0.52-0.59)	**<.001**	
	Insured/No specifics	9620	0.81 (0.75-0.87)	**<.001**	
	Uninsured	3184	-	**-**	
Year of diagnosis	2000-2006	53691	1.13 (1.11-1.16)	**<.001**	**<.001**
	2007-2013	60194	-	**-**	
Cancer-directed surgery accepted	No	57882	2.02 (1.97-2.06)	**<.001**	**<.001**
	Unknown	1194	2.04 (1.86-2.24)	**<.001**	
	Yes	54809	-	**-**	

Covariate	Level	N	Survival months Hazard Ratio (95% CI)	HR P-value	Log-rank P-value
Cancer site	Larynx	31720	0.85 (0.83-0.87)	**<.001**	**<.001**
	Oral cavity	49850	0.85 (0.83-0.88)	**<.001**	
	Pharynx	32315	-	**-**	
Tumor grade	Moderately differentiated; Grade II	45915	1.63 (1.57-1.69)	**<.001**	**<.001**
	Poorly differentiated; Grade III	28173	1.71 (1.64-1.78)	**<.001**	
	Undifferentiated; anaplastic; Grade IV	1441	1.33 (1.21-1.47)	**<.001**	
	Unknown	23626	1.48 (1.42-1.54)	**<.001**	
	Well differentiated; Grade I	14730	-	**-**	
CS tumor size in cm (quartile)	>2, <=3	13852	2.04 (1.93-2.17)	**<.001**	**<.001**
	>3, <=4	15097	2.95 (2.79-3.12)	**<.001**	
	>4, <=99	10672	5.11 (4.83-5.41)	**<.001**	
	Unknown	60824	2.78 (2.65-2.93)	**<.001**	
	>=0, <=2	13440	-	**-**	
Combined SEER summary stage	Distant	18885	1.80 (1.71-1.90)	**<.001**	**<.001**
	Localized	41940	0.39 (0.37-0.41)	**<.001**	
	Regional	48222	0.91 (0.87-0.96)	**<.001**	
	Unknown	4838	-	**-**	
% < High school education (quartile)	>16, <=19	28500	1.14 (1.11-1.18)	**<.001**	**<.001**
	>19, <=27	28810	1.28 (1.24-1.32)	**<.001**	
	>27, <=51	27867	1.30 (1.26-1.33)	**<.001**	
	Unknown	13	1.15 (0.43-3.05)	0.786	
	>=3, <=16	28695	-	**-**	
% Unemployed (quartile)	>5, <=6	33112	1.10 (1.07-1.13)	**<.001**	**<.001**
	>6, <=9	32733	1.24 (1.20-1.28)	**<.001**	
	>9, <=18	21406	1.37 (1.33-1.42)	**<.001**	
	Unknown	13	1.13 (0.43-3.02)	0.801	
	>=1, <=5	26621	-	**-**	

			Survival months		
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
% Urban	>76, <=95	29757	1.00 (0.97-1.03)	0.971	**<.001**
(quartile)	>95, <=99	27285	0.98 (0.95-1.01)	0.228	
	>99, <=100	28252	1.10 (1.07-1.13)	**<.001**	
	Unknown	13	0.99 (0.37-2.65)	0.991	
	>=0, <=76	28578	-	**-**	
% Moved from	>2, <=3	28399	1.01 (0.98-1.04)	0.633	**<.001**
outside US	>3, <=5	29194	0.92 (0.90-0.95)	**<.001**	
(quartile)	>5, <=10	27581	1.02 (0.99-1.05)	0.117	
	Unknown	13	0.96 (0.36-2.56)	0.938	
	>=0, <=2	28698	-	**-**	
% Ever Smoker	>45, <=51	29320	0.99 (0.96-1.02)	0.433	**<.001**
(males age	>51, <=56	28012	1.05 (1.02-1.08)	**0.001**	
18+) (quartile)	>56, <=73	28040	1.04 (1.01-1.07)	**0.006**	
	Unknown	13	0.99 (0.37-2.65)	0.989	
	>=29, <=45	28500	-	**-**	

Table 4.1 Multivariate Association with Overall Survival

Covariate	Level	Survival months ---------- Hazard Ratio (95% CI)	HR P-value
Age	40-59 years	1.57 (1.47-1.69)	**<.001**
	60-79 years	2.65 (2.47-2.85)	**<.001**
	80+ years	5.73 (5.32-6.17)	**<.001**
	20-39 years	-	-
Race	Hispanic (All Races)	1.06 (1.03-1.10)	**<.001**
	Non-Hispanic Black	1.34 (1.30-1.37)	**<.001**
	Non-Hispanic Others	0.91 (0.88-0.95)	**<.001**
	Non-Hispanic White	-	-
Marital status at diagnosis	Married	0.68 (0.66-0.69)	**<.001**
	Other	-	-
Year of diagnosis	2000-2006	1.19 (1.17-1.22)	**<.001**
	2007-2013	-	-
Cancer-directed surgery accepted	No	1.53 (1.50-1.55)	**<.001**
	Unknown	1.51 (1.39-1.63)	**<.001**
	Yes	-	-
Cancer site	Larynx	1.03 (1.00-1.05)	**0.045**
	Oral cavity	1.12 (1.09-1.14)	**<.001**
	Pharynx	-	-
Tumor grade	Moderately differentiated; Grade II	1.19 (1.16-1.23)	**<.001**
	Poorly differentiated; Grade III	1.08 (1.05-1.11)	**<.001**
	Undifferentiated; anaplastic; Grade IV	0.86 (0.79-0.94)	**<.001**
	Unknown	1.03 (1.00-1.07)	**0.042**
	Well differentiated; Grade I	-	-

Covariate	Level	Survival months Hazard Ratio (95% CI)	HR P-value
CS tumor size in cm (quartile)	>2, <=3	1.25 (1.20-1.31)	**<.001**
	>3, <=4	1.40 (1.34-1.46)	**<.001**
	>4, <=99	1.95 (1.87-2.04)	**<.001**
	Unknown	1.54 (1.48-1.60)	**<.001**
	>=0, <=2	-	-
Combined SEER summary stage	Distant	1.86 (1.78-1.95)	**<.001**
	Localized	0.67 (0.65-0.70)	**<.001**
	Regional	1.13 (1.08-1.18)	**<.001**
	Unknown	-	-
% < High school education (quartile)	>16, <=19	1.05 (1.02-1.08)	**<.001**
	>19, <=27	1.07 (1.04-1.10)	**<.001**
	>27, <=51	1.12 (1.08-1.16)	**<.001**
	Unknown	1.22 (0.51-2.94)	0.653
	>=3, <=16	-	-
% Unemployed (quartile)	>5, <=6	1.00 (0.98-1.03)	0.888
	>6, <=9	1.05 (1.02-1.08)	**<.001**
	>9, <=18	1.08 (1.05-1.12)	**<.001**
	Unknown	-	-
	>=1, <=5	-	-
% Urban (quartile)	>76, <=95	0.97 (0.94-1.00)	0.055
	>95, <=99	0.97 (0.94-1.01)	0.112
	>99, <=100	0.99 (0.95-1.02)	0.437
	Unknown	-	-
	>=0, <=76	-	-
% Moved from outside US (quartile)	>2, <=3	1.01 (0.98-1.04)	0.436
	>3, <=5	0.97 (0.94-1.01)	0.130
	>5, <=10	0.95 (0.91-0.99)	**0.014**
	Unknown	-	-
	>=0, <=2	-	-

Covariate	Level	Survival months Hazard Ratio (95% CI)	HR P-value
% Ever Smoker (males age 18+) (quartile)	>45, <=51	1.01 (0.98-1.04)	0.465
	>51, <=56	1.06 (1.03-1.10)	**<.001**
	>56, <=73	1.06 (1.02-1.10)	**0.001**
	Unknown	-	-
	>=29, <=45	-	-

* Number of observations in the original data set = 113885. Number of observations used = 113885.

** Backward selection with an alpha level of removal of 0.05 was used. The following variables were removed from the model: Origin recode NHIA (Hispanic, Non-Hisp).

Table 4.2 Multivariate Association with Overall Survival for Laryngeal Cancer

Covariate	Level	Survival months ---------- Hazard Ratio (95% CI)	HR P-value
Age	40-59 years	2.72 (2.17-3.42)	**<.001**
	60-79 years	4.27 (3.40-5.36)	**<.001**
	80+ years	9.20 (7.30-11.59)	**<.001**
	20-39 years	-	-
Race	Hispanic (All Races)	0.91 (0.85-0.97)	**0.004**
	Non-Hispanic Black	1.08 (1.03-1.12)	**0.001**
	Non-Hispanic Others	0.79 (0.72-0.86)	**<.001**
	Non-Hispanic White	-	-
Sex	Female	0.92 (0.89-0.96)	**<.001**
	Male	-	-
Marital status at diagnosis	Married	0.71 (0.69-0.73)	**<.001**
	Other	-	-
Year of diagnosis	2000-2006	1.12 (1.08-1.16)	**<.001**
	2007-2013	-	-
Cancer-directed surgery accepted	No	1.46 (1.41-1.51)	**<.001**
	Unknown	1.31 (1.15-1.50)	**<.001**
	Yes	-	-
Tumor grade	Moderately differentiated; Grade II	1.21 (1.15-1.28)	**<.001**
	Poorly differentiated; Grade III	1.44 (1.36-1.53)	**<.001**
	Undifferentiated; anaplastic; Grade IV	1.38 (1.11-1.72)	**0.004**
	Unknown	1.21 (1.14-1.28)	**<.001**
	Well differentiated; Grade I	-	-

Covariate	Level	Survival months Hazard Ratio (95% CI)	HR P-value
CS tumor size in cm (quartile)	>2, <=3	1.21 (1.09-1.33)	**<.001**
	>3, <=4	1.28 (1.16-1.40)	**<.001**
	>4, <=99	1.54 (1.39-1.71)	**<.001**
	Unknown	1.24 (1.15-1.34)	**<.001**
	>=0, <=2	-	-
Combined SEER summary stage	Distant	2.13 (1.95-2.33)	**<.001**
	Localized	0.74 (0.68-0.80)	**<.001**
	Regional	1.57 (1.44-1.72)	**<.001**
	Unknown	-	-
% < High school education (quartile)	>16, <=19	1.00 (0.94-1.05)	0.863
	>19, <=27	1.03 (0.98-1.09)	0.217
	>27, <=51	1.07 (1.01-1.13)	**0.030**
	Unknown	1.63 (0.23-11.61)	0.624
	>=3, <=16	-	-
% Unemployed (quartile)	>5, <=6	1.03 (0.98-1.08)	0.315
	>6, <=9	1.05 (1.00-1.11)	0.063
	>9, <=18	1.11 (1.04-1.17)	**<.001**
	Unknown	-	-
	>=1, <=5	-	-
% Urban (quartile)	>76, <=95	1.01 (0.96-1.06)	0.785
	>95, <=99	1.01 (0.95-1.07)	0.727
	>99, <=100	0.94 (0.89-1.00)	**0.035**
	Unknown	-	-
	>=0, <=76	-	-
% Ever Smoker (males age 18+) (quartile)	>45, <=51	0.95 (0.90-1.00)	0.058
	>51, <=56	0.99 (0.94-1.05)	0.781
	>56, <=73	1.01 (0.95-1.06)	0.815
	Unknown	-	-
	>=29, <=45	-	-

Covariate	Level	Survival months Hazard Ratio (95% CI)	HR P-value

* Number of observations in the original data set = 31720. Number of observations used = 31720.

** Backward selection with an alpha level of removal of 0.05 was used. The following variables were removed from the model: Origin recode NHIA (Hispanic, Non-Hisp).

Table 4.3 Multivariate Association with Disease-Specific Survival

Covariate	Level	Survival months Hazard Ratio (95% CI)	HR P-value
Age	40-59 years	1.30 (1.20-1.41)	**<.001**
	60-79 years	1.87 (1.73-2.02)	**<.001**
	80+ years	3.48 (3.20-3.78)	**<.001**
	20-39 years	-	-
Race	Hispanic (All Races)	1.12 (1.08-1.17)	**<.001**
	Non-Hispanic Black	1.39 (1.34-1.43)	**<.001**
	Non-Hispanic Others	1.00 (0.95-1.04)	0.866
	Non-Hispanic White	-	-
Sex	Female	1.06 (1.03-1.08)	**<.001**
	Male	-	-
Marital status at diagnosis	Married	0.68 (0.67-0.70)	**<.001**
	Other	-	-
Year of diagnosis	2000-2006	1.20 (1.17-1.23)	**<.001**
	2007-2013	-	-
Cancer-directed surgery accepted	No	1.65 (1.62-1.69)	**<.001**
	Unknown	1.68 (1.53-1.84)	**<.001**
	Yes	-	-
Cancer site	Larynx	1.02 (0.99-1.05)	0.150
	Oral cavity	1.14 (1.11-1.17)	**<.001**
	Pharynx	-	-

Covariate	Level	Survival months Hazard Ratio (95% CI)	HR P-value
Tumor grade	Moderately differentiated; Grade II	1.26 (1.21-1.31)	**<.001**
	Poorly differentiated; Grade III	1.14 (1.09-1.19)	**<.001**
	Undifferentiated; anaplastic; Grade IV	0.86 (0.77-0.95)	**0.003**
	Unknown	1.05 (1.01-1.10)	**0.018**
	Well differentiated; Grade I	-	-
CS tumor size in cm (quartile)	>2, <=3	1.38 (1.30-1.46)	**<.001**
	>3, <=4	1.60 (1.51-1.70)	**<.001**
	>4, <=99	2.33 (2.20-2.47)	**<.001**
	Unknown	1.86 (1.76-1.95)	**<.001**
	>=0, <=2	-	-
Combined SEER summary stage	Distant	2.09 (1.98-2.20)	**<.001**
	Localized	0.51 (0.49-0.54)	**<.001**
	Regional	1.18 (1.12-1.24)	**<.001**
	Unknown	-	-
% < High school education (quartile)	>16, <=19	1.07 (1.03-1.10)	**<.001**
	>19, <=27	1.08 (1.04-1.12)	**<.001**
	>27, <=51	1.14 (1.09-1.18)	**<.001**
	Unknown	1.57 (0.59-4.19)	0.368
	>=3, <=16	-	-
% Unemployed (quartile)	>5, <=6	1.01 (0.98-1.05)	0.390
	>6, <=9	1.07 (1.04-1.11)	**<.001**
	>9, <=18	1.08 (1.04-1.13)	**<.001**
	Unknown	-	-
	>=1, <=5	-	-

Covariate	Level	Survival months Hazard Ratio (95% CI)	HR P-value
% Moved from outside US (quartile)	>2, <=3	1.00 (0.97-1.03)	0.966
	>3, <=5	0.95 (0.92-0.99)	**0.007**
	>5, <=10	0.95 (0.91-0.99)	**0.018**
	Unknown	-	-
	>=0, <=2	-	-
% Ever Smoker (males age 18+) (quartile)	>45, <=51	1.02 (0.98-1.05)	0.288
	>51, <=56	1.07 (1.02-1.11)	**0.002**
	>56, <=73	1.03 (0.99-1.08)	0.167
	Unknown	-	-
	>=29, <=45	-	-

* Number of observations in the original data set = 113885. Number of observations used = 113885.

** Backward selection with an alpha level of removal of 0.05 was used. The following variables were removed from the model: Origin recode NHIA (Hispanic, Non-Hisp).

Table 5.1 Interaction Effect Between Race and Age on Overall Survival

			Survival months		
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Type3 P-value**
Comparisons Stratified by Age :	**Race :**		-	-	**<.001**
20-39 years	Non-Hispanic Black vs. Hispanic (All Races)	296 vs. 412	1.17 (0.91-1.51)	0.215	-
	Non-Hispanic Others vs. Hispanic (All Races)	466 vs. 412	0.84 (0.65-1.09)	0.202	
	Non-Hispanic White vs. Hispanic (All Races)	1918 vs. 412	0.73 (0.59-0.90)	**0.003**	
40-59 years	Non-Hispanic Black vs. Hispanic (All Races)	6444 vs. 3411	1.43 (1.34-1.53)	**<.001**	-
	Non-Hispanic Others vs. Hispanic (All Races)	2881 vs. 3411	0.80 (0.73-0.87)	**<.001**	
	Non-Hispanic White vs. Hispanic (All Races)	35207 vs. 3411	0.89 (0.84-0.94)	**<.001**	
60-79 years	Non-Hispanic Black vs. Hispanic (All Races)	5368 vs. 3656	1.13 (1.07-1.20)	**<.001**	-
	Non-Hispanic Others vs. Hispanic (All Races)	3017 vs. 3656	0.86 (0.80-0.93)	**<.001**	
	Non-Hispanic White vs. Hispanic (All Races)	40576 vs. 3656	0.95 (0.91-1.00)	**0.042**	
80+ years	Non-Hispanic Black vs. Hispanic (All Races)	561 vs. 653	0.95 (0.83-1.09)	0.466	-
	Non-Hispanic Others vs. Hispanic (All Races)	725 vs. 653	0.90 (0.79-1.02)	0.103	
	Non-Hispanic White vs. Hispanic (All Races)	8294 vs. 653	1.01 (0.92-1.11)	0.821	

Covariate	Level	N	Survival months Hazard Ratio (95% CI)	HR P-value	Type3 P-value

* Number of observations in the original data set = 113885. Number of observations used = 113885.

** Backward selection with an alpha level of removal of .05 was used. The following variables were removed from the model: Origin recode NHIA (Hispanic, Non-Hisp).

*** The estimated stratified treatement effect was controlled by: % Ever Smoker (males age 18+ - sae 2000-) (quartile), % Moved from outside US 2000 (quartile), % Unemployed 2000 (quartile), % Urban 2000 (quartile), Age_80*RaceandoriginrecodeNHWNHBNHAIA, CS tumor size in cm(2004+) (quartile), Cancer-directed surgery, Combined SEER summary stage, Grade, Marital status at diagnosis, Sex, Site recode ICD-O-3/WHO 2008, Year of diagnosis

Table 5.2 Interaction Effect Between Race and Age on Disease-Specific Survival

			Survival months		
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Type3 P-value**
Comparisons Stratified by Age :	**Race :**		-	-	**<.001**
20-39 years	Non-Hispanic Black vs. Hispanic (All Races)	296 vs. 412	1.06 (0.81-1.39)	0.686	-
	Non-Hispanic Others vs. Hispanic (All Races)	466 vs. 412	0.85 (0.64-1.11)	0.232	
	Non-Hispanic White vs. Hispanic (All Races)	1918 vs. 412	0.76 (0.61-0.95)	**0.014**	
40-59 years	Non-Hispanic Black vs. Hispanic (All Races)	6444 vs. 3411	1.40 (1.30-1.50)	**<.001**	-
	Non-Hispanic Others vs. Hispanic (All Races)	2881 vs. 3411	0.87 (0.78-0.96)	**0.004**	
	Non-Hispanic White vs. Hispanic (All Races)	35207 vs. 3411	0.88 (0.83-0.94)	**<.001**	
60-79 years	Non-Hispanic Black vs. Hispanic (All Races)	5368 vs. 3656	1.11 (1.04-1.19)	**0.003**	-
	Non-Hispanic Others vs. Hispanic (All Races)	3017 vs. 3656	0.87 (0.80-0.95)	**0.001**	
	Non-Hispanic White vs. Hispanic (All Races)	40576 vs. 3656	0.89 (0.84-0.94)	**<.001**	
80+ years	Non-Hispanic Black vs. Hispanic (All Races)	561 vs. 653	1.03 (0.88-1.22)	0.695	-
	Non-Hispanic Others vs. Hispanic (All Races)	725 vs. 653	1.02 (0.87-1.21)	0.788	
	Non-Hispanic White vs. Hispanic (All Races)	8294 vs. 653	0.97 (0.86-1.11)	0.692	

Covariate	Level	N	Survival months Hazard Ratio (95% CI)	HR P-value	Type3 P-value

* Number of observations in the original data set = 113885. Number of observations used = 113885.

** Backward selection with an alpha level of removal of .05 was used. The following variables were removed from the model: Origin recode NHIA (Hispanic, Non-Hisp).

*** The estimated stratified treatement effect was controlled by: % < High school education 2000 (quartile), % Ever Smoker (males age 18+ - sae 2000-) (quartile), % Moved from outside US 2000 (quartile), % Unemployed 2000 (quartile), Age_80*RaceandoriginrecodeNHWNHBNHAIA, CS tumor size in cm(2004+) (quartile), Cancer-directed surgery, Combined SEER summary stage, Grade, Marital status at diagnosis, Sex, Site recode ICD-O-3/WHO 2008, Year of diagnosis

Table 6 Demographics of the Patient After Propensity Score Matching

Covariate	Level	Hispanic (All Races) N=3230 N(%)	Non-Hispanic Black N=3230 N(%)	Non-Hispanic Others N=3230 N(%)	Non-Hispanic White N=3230 N(%)	Standardized Difference
Age	20-39 years	95 (2.94)	89 (2.76)	91 (2.82)	76 (2.35)	0.037
	40-59 years	1428 (44.21)	1442 (44.64)	1444 (44.71)	1427 (44.18)	0.011
	60-79 years	1478 (45.76)	1479 (45.79)	1450 (44.89)	1494 (46.25)	0.027
	80+ years	229 (7.09)	220 (6.81)	245 (7.59)	233 (7.21)	0.030
Sex	Female	756 (23.41)	712 (22.04)	798 (24.71)	712 (22.04)	0.063
	Male	2474 (76.59)	2518 (77.96)	2432 (75.29)	2518 (77.96)	0.063
Marital status at diagnosis	Married	1653 (51.18)	1637 (50.68)	1629 (50.43)	1675 (51.86)	0.028
	Other	1577 (48.82)	1593 (49.32)	1601 (49.57)	1555 (48.14)	0.028
Year of diagnosis	2000-2006	1515 (46.9)	1454 (45.02)	1513 (46.84)	1715 (53.1)	**0.162**
	2007-2013	1715 (53.1)	1776 (54.98)	1717 (53.16)	1515 (46.9)	**0.162**
Cancer-directed surgery accepted	No	1800 (55.73)	1750 (54.18)	1777 (55.02)	1719 (53.22)	0.050
	Unknown	30 (0.93)	34 (1.05)	28 (0.87)	16 (0.5)	0.064
	Yes	1400 (43.34)	1446 (44.77)	1425 (44.12)	1495 (46.28)	0.059
Cancer site	Larynx	816 (25.26)	955 (29.57)	801 (24.8)	953 (29.5)	**0.107**
	Oral cavity	1317 (40.77)	1238 (38.33)	1370 (42.41)	1254 (38.82)	0.083
	Pharynx	1097 (33.96)	1037 (32.11)	1059 (32.79)	1023 (31.67)	0.049
Tumor grade	Grade II	1261 (39.04)	1251 (38.73)	1285 (39.78)	1327 (41.08)	0.048
	Grade III	914 (28.3)	921 (28.51)	838 (25.94)	880 (27.24)	0.058
	Grade IV	28 (0.87)	22 (0.68)	23 (0.71)	24 (0.74)	0.021
	Unknown	691 (21.39)	649 (20.09)	729 (22.57)	634 (19.63)	0.072
	Grade I	336 (10.4)	387 (11.98)	355 (10.99)	365 (11.3)	0.050
CS tumor size in cm (quartile)	>=0, <=2	278 (8.61)	285 (8.82)	289 (8.95)	226 (7)	0.072
	>2, <=3	352 (10.9)	391 (12.11)	367 (11.36)	326 (10.09)	0.064
	>3, <=4	422 (13.07)	462 (14.3)	420 (13)	388 (12.01)	0.068
	>4, <=99	357 (11.05)	381 (11.8)	374 (11.58)	330 (10.22)	0.050
	Unknown	1821 (56.38)	1711 (52.97)	1780 (55.11)	1960 (60.68)	**0.156**

Covariate	Level	Hispanic (All Races) N=3230 N(%)	Non-Hispanic Black N=3230 N(%)	Non-Hispanic Others N=3230 N(%)	Non-Hispanic White N=3230 N(%)	Standardized Difference
Combined SEER summary stage	Distant	613 (18.98)	591 (18.3)	623 (19.29)	594 (18.39)	0.025
	Localized	1061 (32.85)	1099 (34.02)	1075 (33.28)	1119 (34.64)	0.038
	Regional	1399 (43.31)	1392 (43.1)	1344 (41.61)	1378 (42.66)	0.034
	Unknown	157 (4.86)	148 (4.58)	188 (5.82)	139 (4.3)	0.069
% < High school education (quartile)	>=3, <=16	547 (16.93)	603 (18.67)	526 (16.28)	601 (18.61)	0.063
	>16, <=19	959 (29.69)	977 (30.25)	993 (30.74)	1025 (31.73)	0.044
	>19, <=27	690 (21.36)	611 (18.92)	705 (21.83)	587 (18.17)	0.091
	>27, <=51	1034 (32.01)	1039 (32.17)	1006 (31.15)	1017 (31.49)	0.022
% Unemployed (quartile)	>=1, <=5	438 (13.56)	426 (13.19)	500 (15.48)	427 (13.22)	0.065
	>5, <=6	1004 (31.08)	942 (29.16)	1038 (32.14)	1037 (32.11)	0.064
	>6, <=9	1328 (41.11)	1388 (42.97)	1174 (36.35)	1293 (40.03)	**0.136**
	>9, <=18	460 (14.24)	474 (14.67)	518 (16.04)	473 (14.64)	0.050
% Urban (quartile)	>=0, <=76	191 (5.91)	221 (6.84)	182 (5.63)	153 (4.74)	0.090
	>76, <=95	711 (22.01)	703 (21.76)	709 (21.95)	748 (23.16)	0.033
	>95, <=99	946 (29.29)	960 (29.72)	840 (26.01)	973 (30.12)	0.092
	>99, <=100	1382 (42.79)	1346 (41.67)	1499 (46.41)	1356 (41.98)	0.096
% Moved from outside US (quartile)	>=0, <=2	184 (5.7)	134 (4.15)	155 (4.8)	116 (3.59)	**0.100**
	>2, <=3	539 (16.69)	553 (17.12)	551 (17.06)	525 (16.25)	0.023
	>3, <=5	938 (29.04)	1035 (32.04)	827 (25.6)	956 (29.6)	**0.143**
	>5, <=10	1569 (48.58)	1508 (46.69)	1697 (52.54)	1633 (50.56)	**0.117**

Covariate	Level	Hispanic (All Races) N=3230 N(%)	Non-Hispanic Black N=3230 N(%)	Non-Hispanic Others N=3230 N(%)	Non-Hispanic White N=3230 N(%)	Standardized Difference
% Ever Smoker (males age 18+) (quartile)	>=29, <=45	1453 (44.98)	1396 (43.22)	1502 (46.5)	1433 (44.37)	0.066
	>45, <=51	1060 (32.82)	1123 (34.77)	1060 (32.82)	1145 (35.45)	0.056
	>51, <=56	548 (16.97)	508 (15.73)	486 (15.05)	467 (14.46)	0.069
	>56, <=73	169 (5.23)	203 (6.28)	182 (5.63)	185 (5.73)	0.045

Table 7.1 The Effect of Race on Overall Survival in Matched Sample

			Survival Months		
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Type3 P-value**
Race	Hispanic (All Races)	3230	1.09 (1.01-1.16)	**0.021**	**<.001**
	Non-Hispanic Black	3230	1.26 (1.18-1.35)	**<.001**	
	Non-Hispanic Others	3230	0.90 (0.84-0.97)	**0.005**	
	Non-Hispanic White	3230	-	**-**	

Analysis was taken the clustering effect within Patient_ID into account, and N represented number of Patient_ID-times.

Table 7.2 The Effect of Race on Disease-Specific Survival in Matched Sample

			Survival Months		
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Type3 P-value**
Race	Hispanic (All Races)	3230	1.13 (1.04-1.23)	**0.005**	**<.001**
	Non-Hispanic Black	3230	1.31 (1.21-1.42)	**<.001**	
	Non-Hispanic Others	3230	0.98 (0.90-1.07)	0.648	
	Non-Hispanic White	3230	-	**-**	

Analysis was taken the clustering effect within Patient_ID into account, and N represented number of Patient_ID-times.

Table 7.3 The Effect of Race on Overall Survival in Laryngeal Cancer Matched Sample

			Survival Months		

Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Type3 P-value**
Race	Hispanic (All Races)	969	0.84 (0.74-0.96)	**0.010**	**<.001**
	Non-Hispanic Black	969	1.02 (0.90-1.15)	0.812	
	Non-Hispanic Others	969	0.75 (0.66-0.86)	**<.001**	
	Non-Hispanic White	969	-	**-**	

Analysis was taken the clustering effect within Patient_ID into account, and N represented number of Patient_ID-times.

Figure 1.1 Kaplan-Meier Overall Survival Estimate of HNC Patients by Race

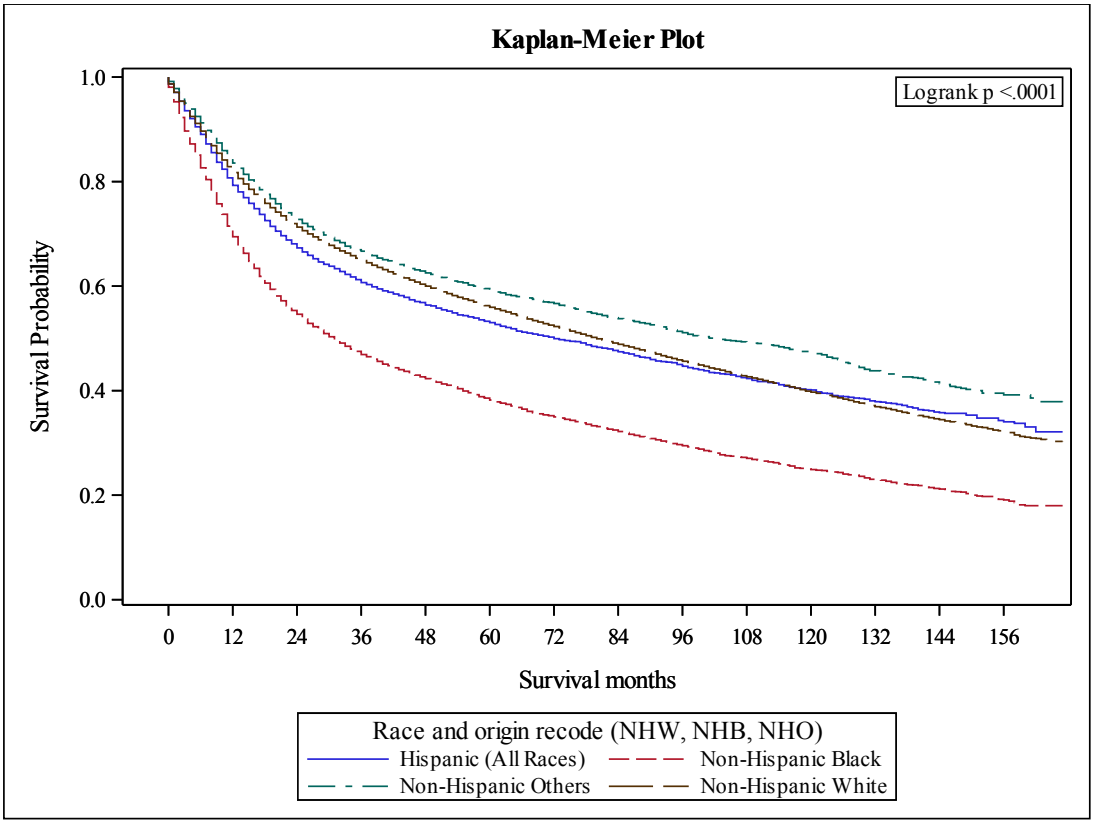

Figure 1.2 Kaplan-Meier Disease-Specific Survival Estimate of HNC Patients by Race

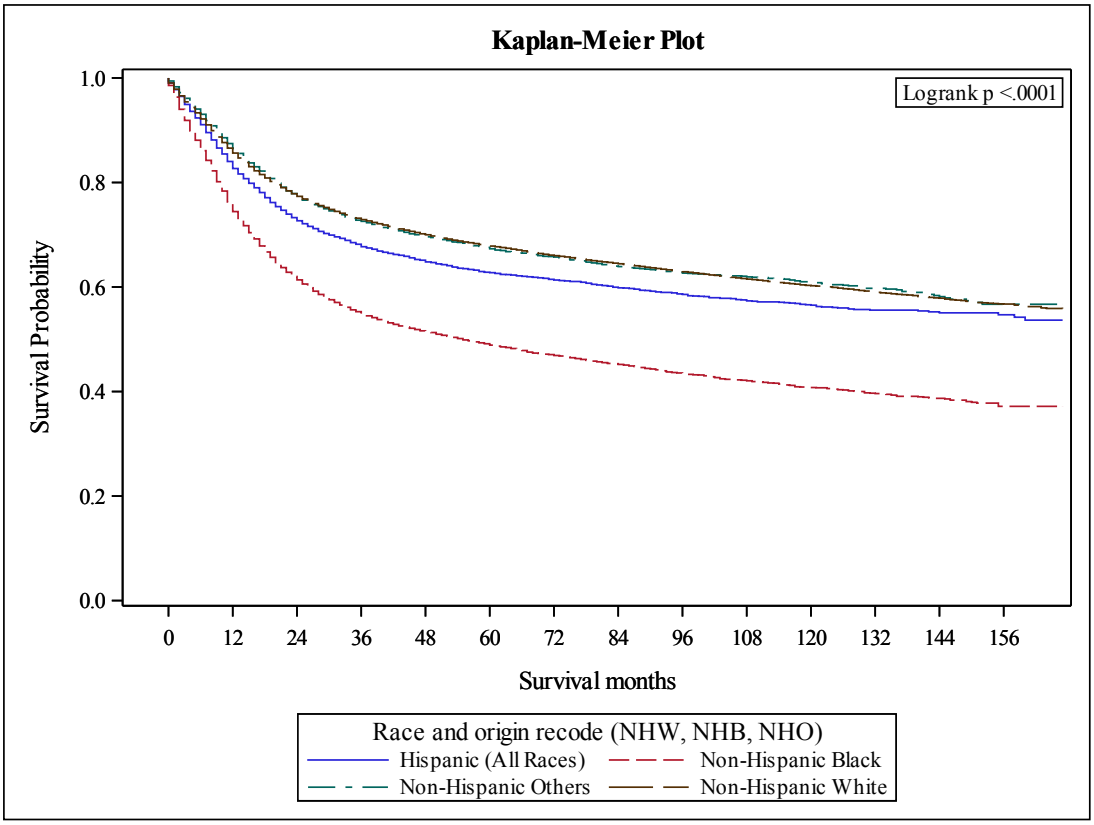

Figure 1.3 Kaplan-Meier Overall Survival Estimate of Oral Cavity Cancer Patients by Race

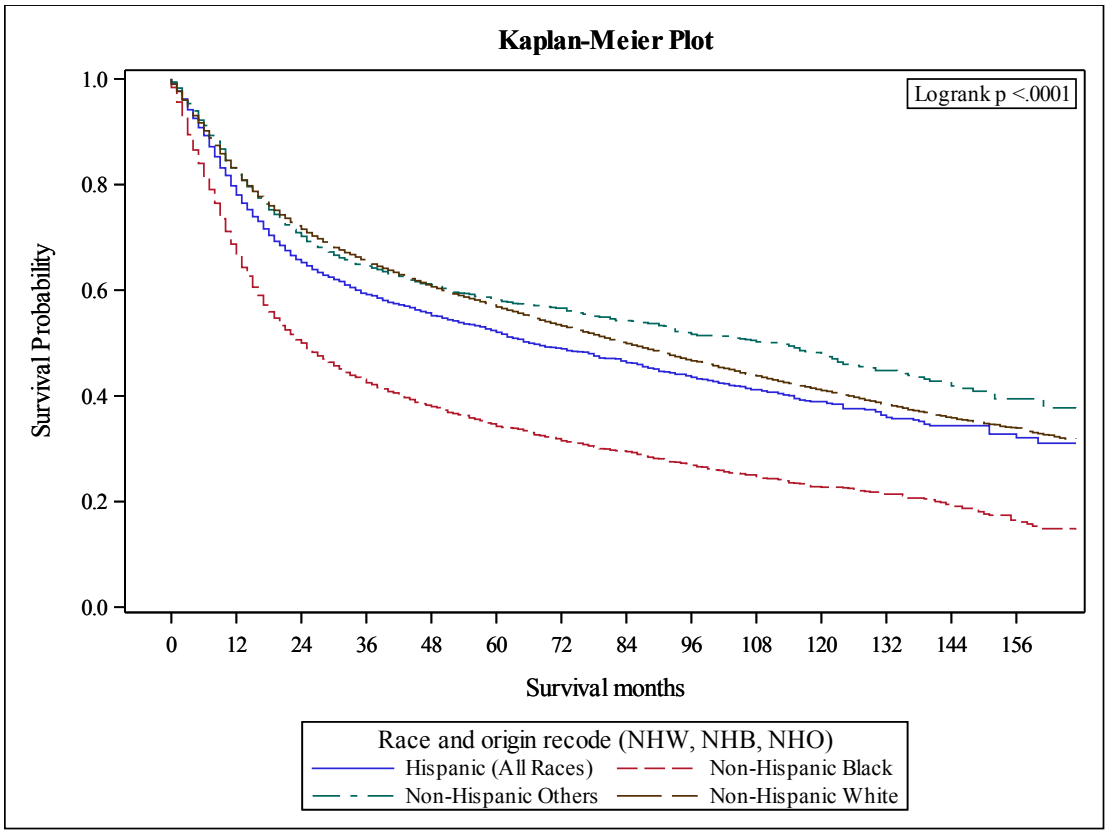

Figure 1.4 Kaplan-Meier Overall Survival Estimate of Pharyngeal Cancer Patients by Race

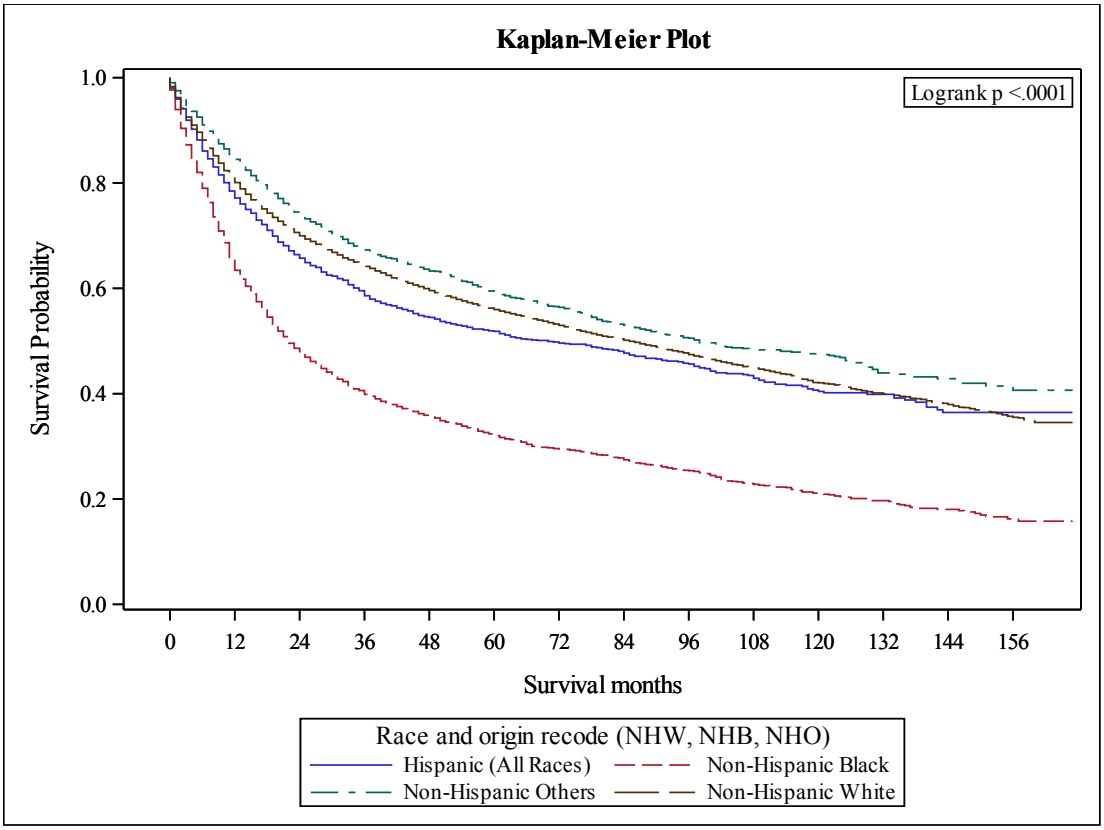

Figure 1.5 Kaplan-Meier Overall Survival Estimate of Laryngeal Cancer Patients by Race

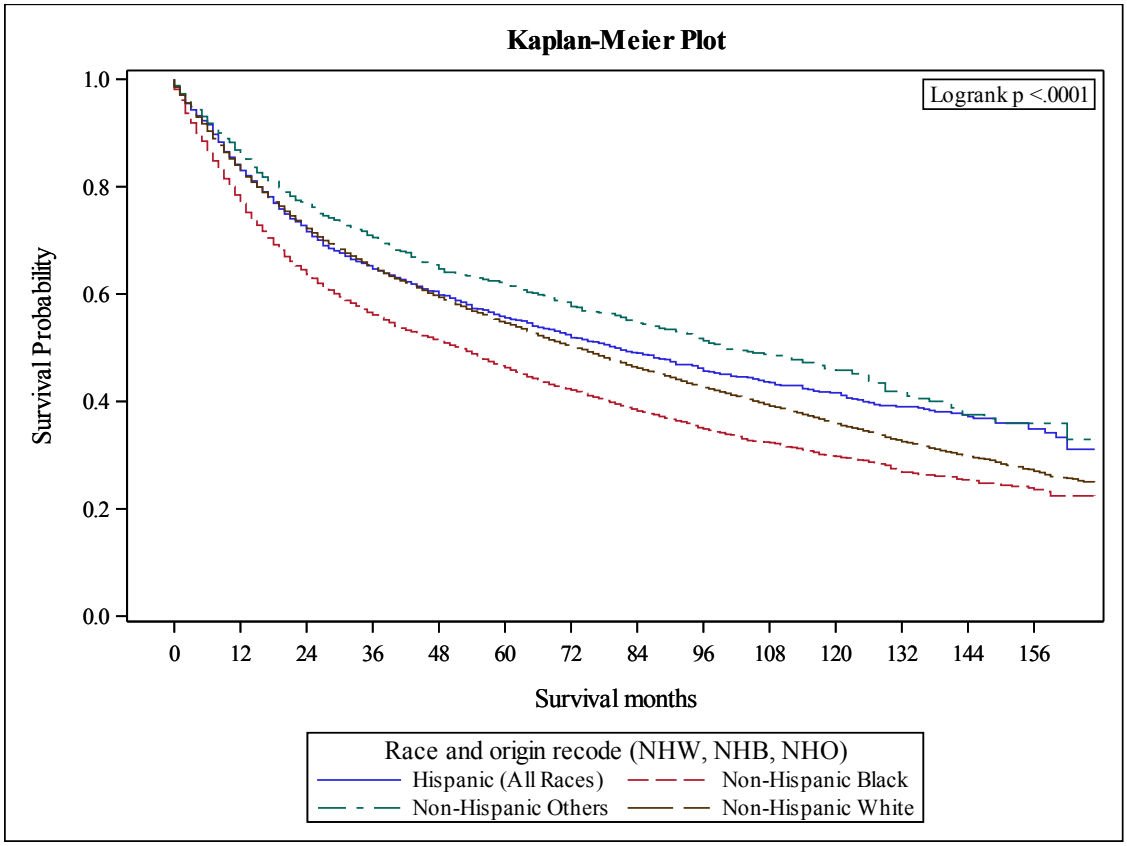

Figure 2.1 The Sum of Within-cluster Distance vs. Matching ID

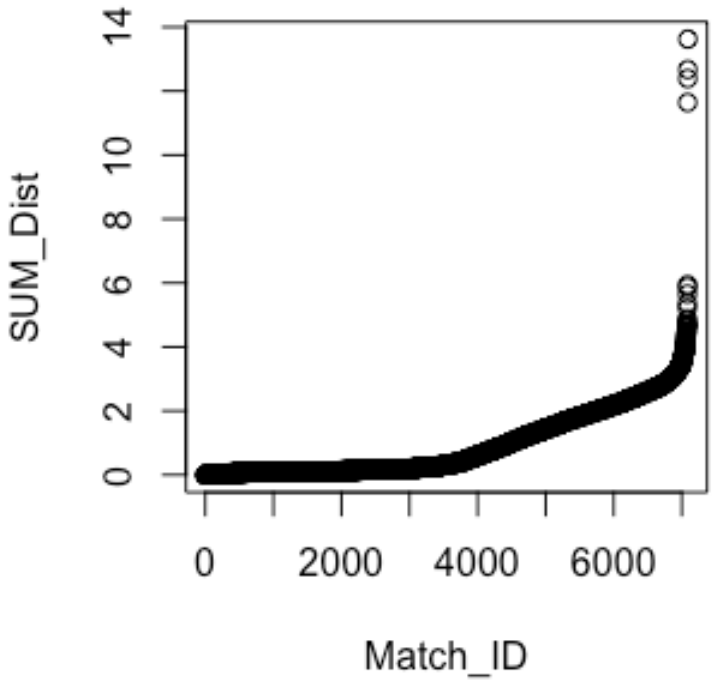

Figure 2.2 The Sum of Within-cluster Distance vs. Matching ID for Laryngeal Cancer Sample

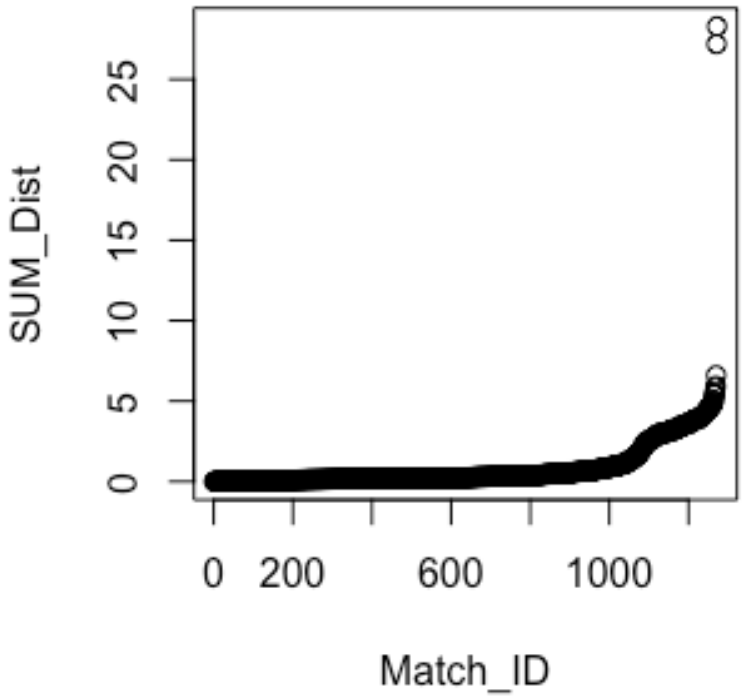

Figure 2.3 The Sum of Within-cluster Distance vs. Matching ID for Oral Cavity Cancer Sample

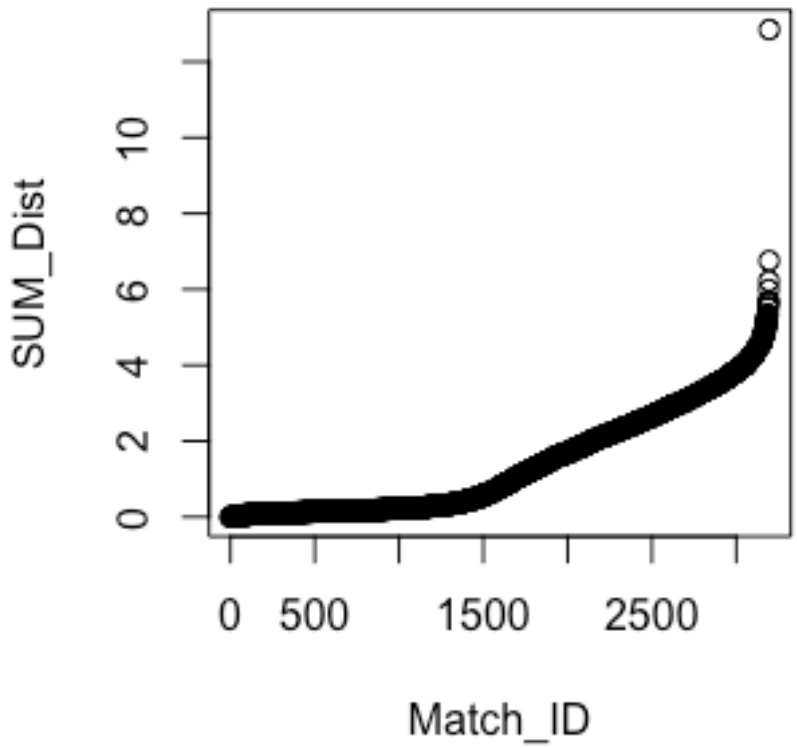

Figure 2.4 The Sum of Within-cluster Distance vs. Matching ID for Pharyngeal Cancer Sample

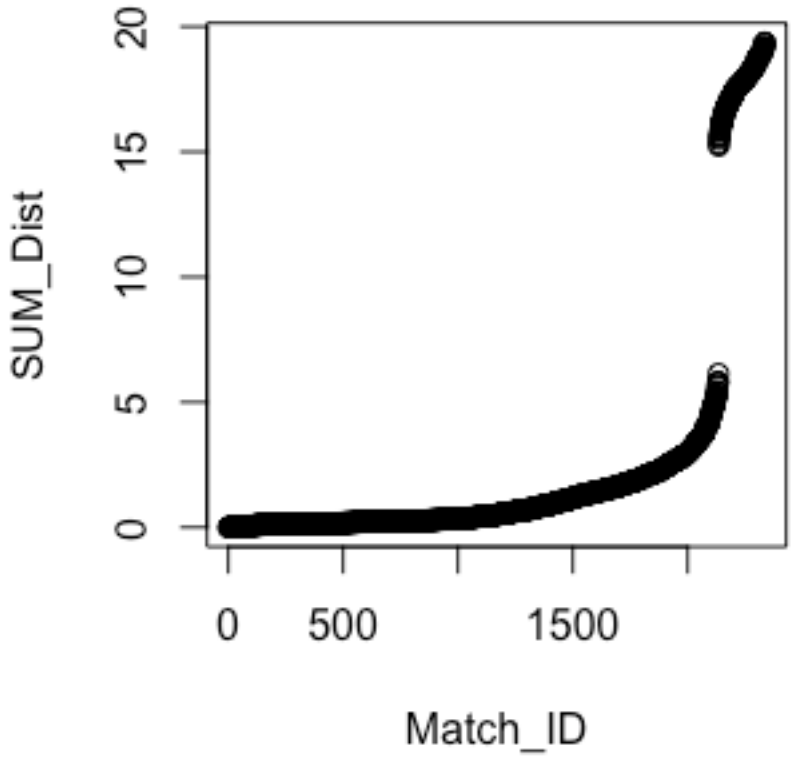

Figure 3.1 Kaplan-Meier Overall Survival Estimate of HNC Patients by Race in Matched Sample

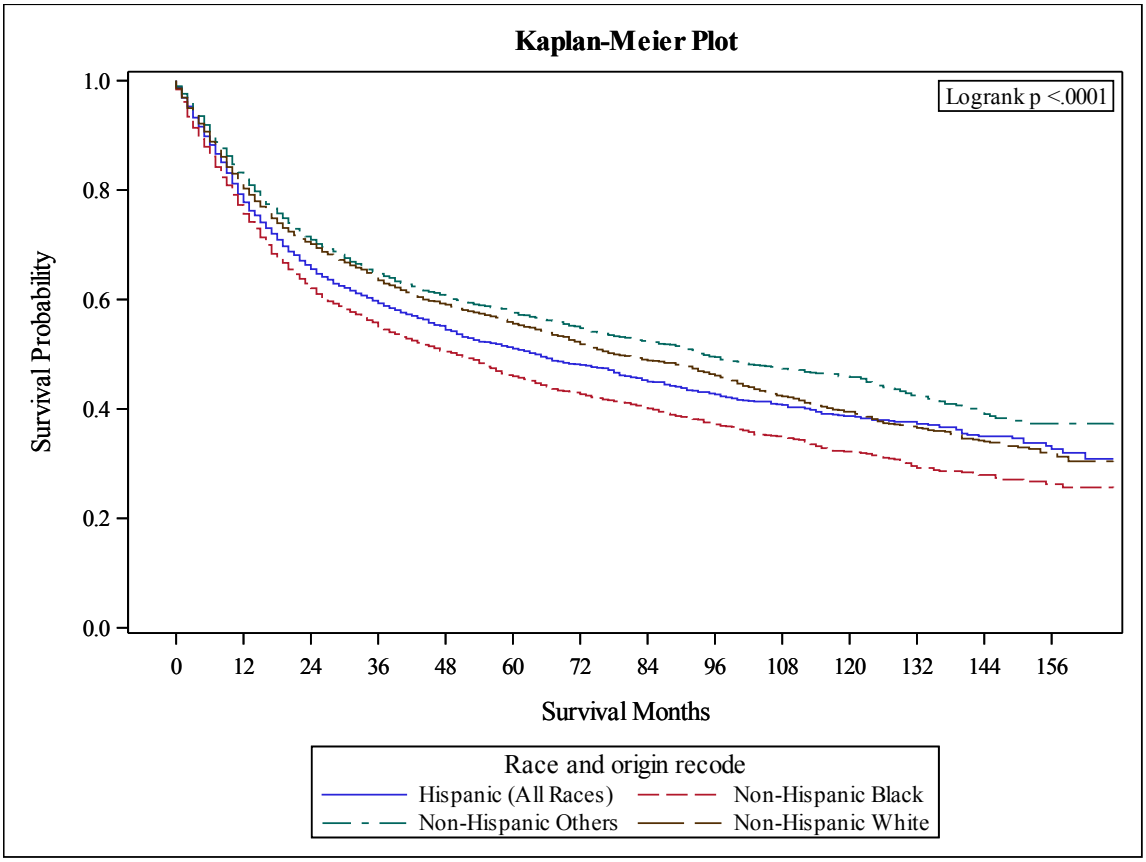

Figure 3.2 Kaplan-Meier Disease-Specific Survival Estimate of HNC Patients by Race in Matched Sample

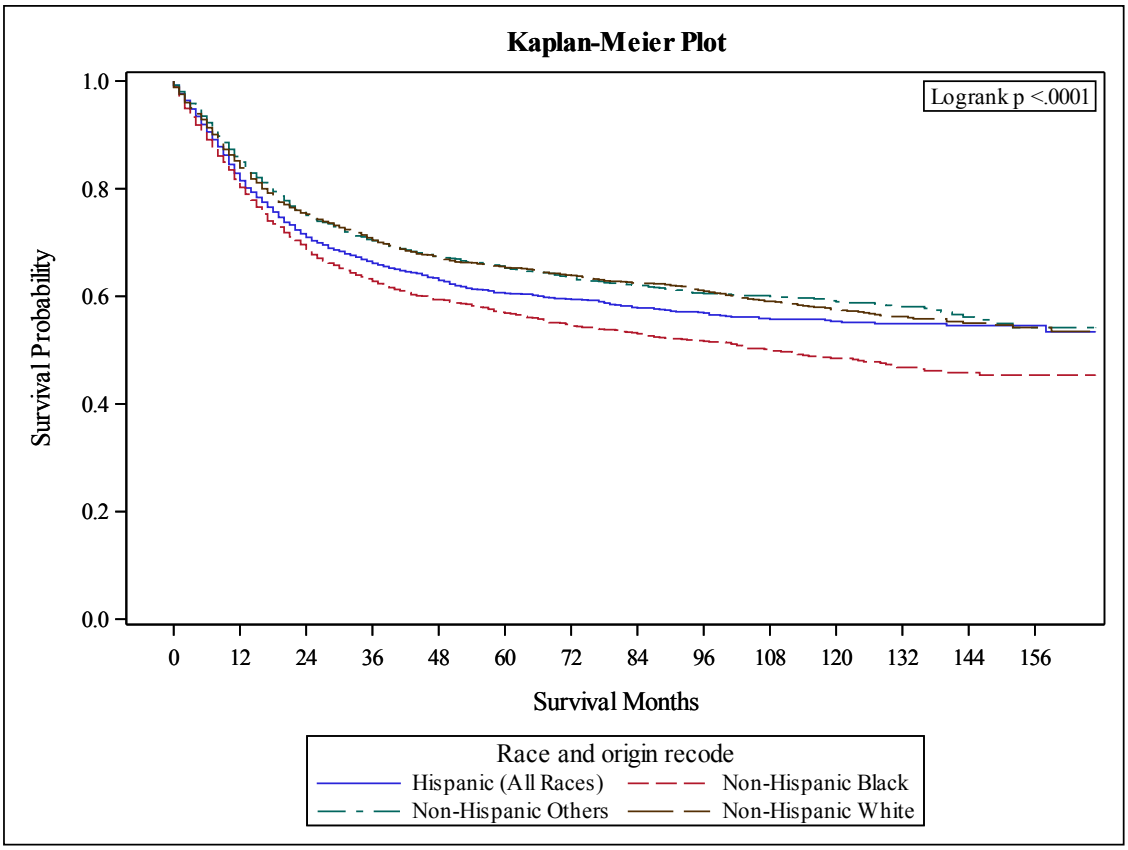

Figure 3.3 Kaplan-Meier Overall Survival Estimate of Oral Cavity Cancer Patients by Race in Matched Sample

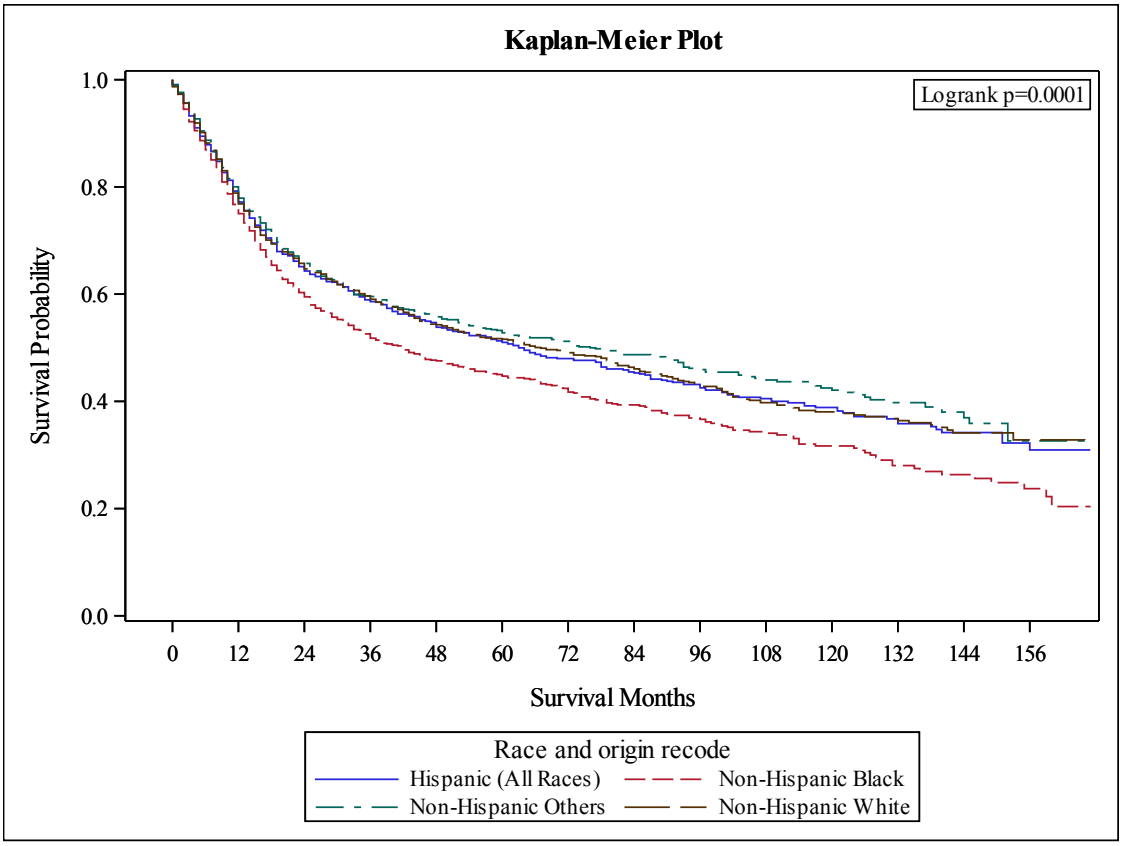

Figure 3.4 Kaplan-Meier Overall Survival Estimate of Pharyngeal Cancer Patients by Race in Matched Sample

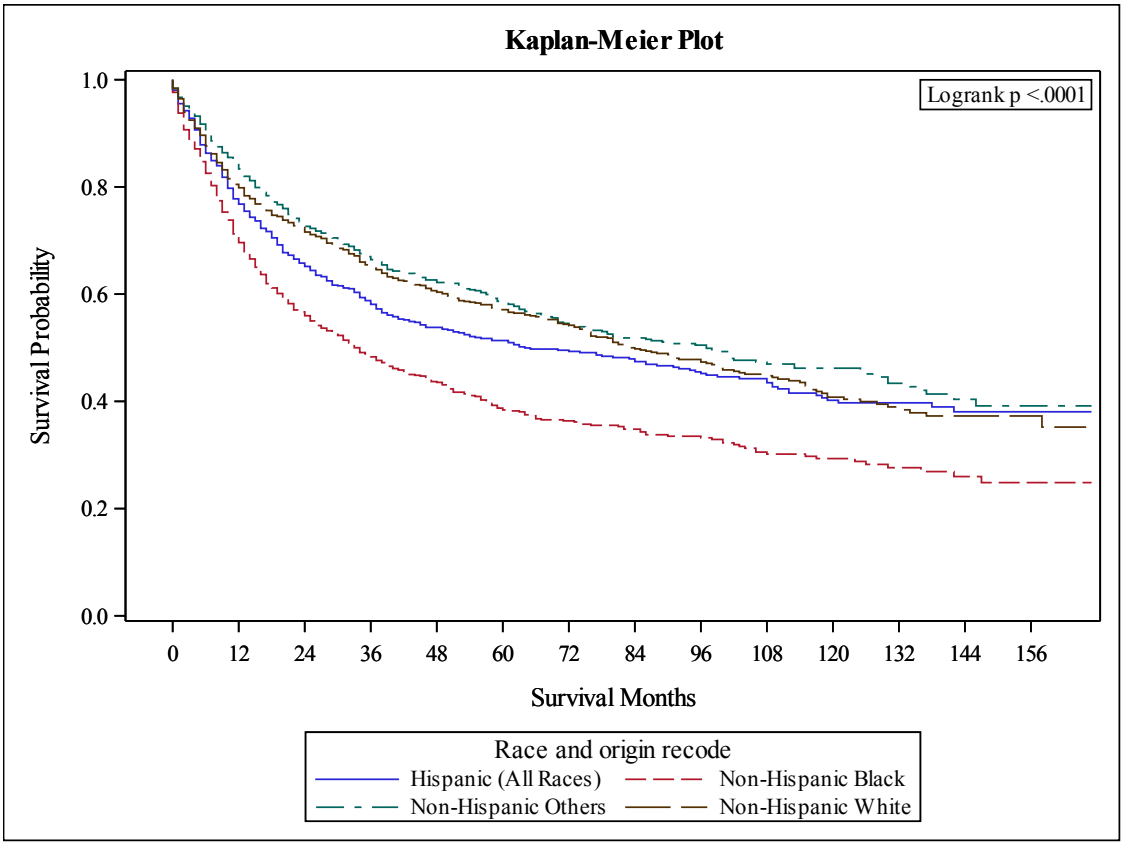

Figure 3.5 Kaplan-Meier Overall Survival Estimate of Laryngeal Cancer Patients by Race in Matched Sample