Racial/Ethnic Differences in Head and Neck Cancer Survival Rates

Andre Lopes Carvalho

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Public Health

University of Washington
2017

Committee:
Branko Kopjar
India J. Ornelas

Program Authorized to Offer Degree:
School of Public Health - Health Services

University of Washington

Abstract

Racial/Ethnic Differences in Head and Neck Cancer Survival Rates

Andre Lopes Carvalho

Chair of the Supervisory Committee:
Associate Professor Branko Kopjar
Department of Health Services

Background: Head and neck cancer (HNC) is the eighth most common cancer in the male population in the U.S. It is more common in men than women, and it usually affects adults ages 40 to 60. The main risk factors are tobacco and alcohol consumption and HPV infection. It is also known that race/ethnicity, social determinants of health (SDOH), and oral health influence the prevalence of HNC. Regarding its prognosis, the most important factors related to survival are the tumor site, the clinical stage at diagnosis, and proper treatment given timely. Despite all the advances in treatment for head and neck cancer in the last 30 years, survival rates have not increased at the same pace, nor are they the same in different populations. Recently, researchers have investigated how the risk factors of HNC can also affect survival. Some studies have looked for the association between race/ethnicity, SDOH, and other patient- and tumor-related factors and the occurrence of HNC but have also used those factors as predictors of patients' lives after the HNC diagnosis, producing a much more complex understanding of survival outcomes and maybe justifying why advances in treatment have not been enough to improve survival across all population. The aim of this study is to investigate the racial/ethnic differences in head and neck cancer survival rates observed in the US in recent years, and its possible association with patient and tumor characteristics, and SDOH.

Methods: This is a retrospective study based on patients registered from 2007 to 2014 in the 18-cancer registry dataset of the SEER Public-Use Data from the Surveillance, Epidemiology, and End Results (SEER) Program. Selection criteria included patients aged 18 or older, with a diagnosis of invasive

squamous cell carcinoma localized at the oral cavity, oropharynx, hypopharynx, or larynx. In addition to the individual (including clinical data) and county-level data obtain from SEER, other county-level attributes were obtained from: the National Center for Education Statistics, the U.S. Census Bureau's – American Community Survey, and the Community Health Rankings. Statistical analysis was performed on the Statistical Package SPSS version 23. Data distribution was described in tables regarding the main variables. The association of these variables with race/ethnicity was evaluated by the chi-square test. For the cancer-specific survival analysis, the Kaplan-Meier method (and the log-rank test) as well as the Cox proportional hazards model were used. For all statistical tests, the significance was 2-sided and achieved when p-values ≤ 0.05.

Results: The final sample size included 59,853 head and neck cancer patients. The majority of the patients were White (76.4%), followed by Blacks (11.2%), Hispanics (7.7%), Asian or Pacific Islanders (4.2%), and American Indians/Alaska Natives (0.5%). The tumors were localized predominantly at the oropharynx (41.6%) as is usually the case with patients at an advanced clinical stage (62.9%). The most common treatment performed was radiotherapy (with or without chemotherapy – 42.5%). For all socio-demographics characteristics evaluated (age, gender, marital status, area of residence, and health insurance), clinical characteristics (tumor site, clinical stage, and treatment), county-level SDOH attributes (age distribution, poverty, unemployment, income, GINI index, education, health literacy, nationality status, and linguistic isolation), and county scores (health outcomes, health behavior, clinical care, social and economic factors, and physical environment), statistical differences were found when comparing the distribution of all these variables between races/ethnicities, with Whites presenting, overall, a more favorable distribution with regard to patient, tumor, socioeconomic, and county-level attributes. Blacks displayed the less favorable, Asian/PI, Hispanics and AI/AN presented a mixed distribution depending on the variable being evaluated. In the univariate survival analysis, all races/ethnicities showed a higher risk of dying of cancer than Whites ($p<0.001$). Blacks, in particular, presented a much worse risk of death due to the disease, HR 1.799 (95%CI, 1.722–1.879), for Hispanics the HR was 1.231 (95%CI, 1.161 – 1.304), for Asian/PI the HR was 1.096 (95%CI, 1.011 – 1.187) and for AI/AN the HR was 1.318 (95%CI, 1.077 – 1.613). When we first adjusted for patients' characteristics (age, gender, marital status), all races/ethnicities still showed a higher risk of dying due to cancer than Whites ($p<0.001$). In the

subsequent multivariate analysis, when tumor characteristics (tumor site, clinical stage, and treatment) were added to the model, the risk of death from cancer was no longer statistically significant for AI/AN population (HR 0.954, 95%CI 0.761 – 1.195; p=0.680), compared to the Whites, but still higher for Blacks (HR 1.329, 95%CI 1.267 – 1.393; p<0.001), Hispanics (HR 1.135, 95%CI 1.068 – 1.206; p<0.001), and Asian/PI (HR 1.099; 95% CI 1.008 – 1.195; p=0.033) when compared to Whites. The next model, further adjusted by the county-level SDOH variables, showed that Hispanics (HR 1.053, 95%CI 0.991 – 1.120; p=0.097) and Asian/PI (HR 1.045, 95%CI 0.958 – 1.140; p= 0.318) no longer had a worse prognosis than Whites, but Blacks still showed a 20% higher risk of dying from cancer than Whites (HR 1.227, 95%CI 1.169 – 1.288; p<0.001). The other factors associated with survival in the model were marital status (p<0.001); health insurance (p<0.001); county-level clinical care score (p=0.024); and county-level social & economic factors score (p<0.001).

Conclusion: Our study showed that race/ethnicity plays an important role in HNC survival rates. Investments in HNC early detection, and more timely receipt of treatment would decrease the excess death risk by roughly half for all races/ethnicities (including Blacks) when compared to Whites. Further investments and policies related to socioeconomic factors, clinical care and towards decreasing social inequalities (including health care access) would further bring similar prognosis for Whites, AI/AN, Asian/PI, and Hispanics, and reduce the excess risk for Blacks by another 35%. If all this combined effort would be enough to bring comparable survival rates to Blacks, based on the complexity related to their heath access and social interactions, how much can be justified by individual's genetics or tumor biology still warrants more investigation.

Acknowledgements

I would like to thank my thesis committee members, Branko Kopjar and India Ornelas, for their guidance, critiques and support.

A special thanks to Christopher Li at Fred Hutchinson Cancer Research Center, for his insights and feedback.

I am very grateful to my lovely wife Cynthia for her understanding, support and encouragement throughout this journey and all these years.

Table of Contents

Page

List of Tables

Page

List of Figures

List of Appendixes

I. Background

Head and neck cancers (HNC) are tumors in the mucosa of the lips, oral cavity, nasopharynx, oropharynx, larynx, and hypopharynx. They also include tumors in the salivary glands, sinus cavity, and other less common sites in the head and neck region.[1] For the head and neck oncologist, there is a distinction regarding risk factors, treatment, and prognosis depending on the primary tumor location and the histology type.[2] However, in most epidemiological studies, these tumors are all grouped together, which can be confusing when specialists, generalists, and policy makers are interpreting the literature.[1,2]

The majority of the HNC in the western world occurs in the oral cavity, oropharynx, larynx, and hypopharynx and is squamous cell carcinoma, the most common histology.[1,3,4] The main risk factors are tobacco and alcohol consumption, and HPV infection (for oropharynx tumors).[1,5,6] They are more common in men than women, and they usually affect adults ages 40 to 60.[3] A higher incidence of HNC is seen in more developed countries (7.4/100,000 inhabitants vs. 5.6/100,000 inhabitants); however, a higher mortality rate is observed in less developed countries (2.5/100,000 inhabitants vs. 3.2/100,000 inhabitants).[7] The prediction is that by 2030-35, the incidence rates will be higher in less developed countries, followed by even higher mortality rates in those countries, stressing the burden of HNC in the underserved population.[8] Worldwide trends have shown a decrease in the incidence of HNC related to alcohol and tobacco consumption, but an increase of HPV related HNC.[1,4] This shift in the etiology of HNC reflects changes in the socio-demographic profile of the patients and will lead to changes in HNC prognosis, as tobacco/alcohol-induced tumors present with a worse prognosis when compared to HPV-related tumors.[9,10]

As the incidence rates increase in less developed countries, more research has been done to investigate the association of HNC incidence and prevalence with social determinants of health. Social determinants of health (SDOH) are the conditions into which people are born, grow, live, work, and age.[11] These circumstances are shaped by the distribution of money, power, and resources at global, national, and local levels. Studies have shown that SDOH are often the root causes of health inequities, which are the unfair and avoidable differences in health status seen within and between population groups.[11] Most of the times, socioeconomic status (SES) is the only SDOH investigated in the HNC literature. SDOH include, but are not exclusively related to (SES), which is defined as a measure of combined economic and social status. Broadly, this includes characteristics such as education, income, and occupation, and it is usually influenced by race/ethnicity, cultural/religious beliefs, place of residence, and other social factors.[12]

So, it is known that race/ethnicity and SDOH influence the incidence and prevalence of HNC,[13-18] but not many studies have evaluated a possible association between these variables and HNC survival rates. To date, the most important factors related to HNC survival are the cancer site, clinical stage at the time of diagnosis, and the reception of proper treatment in a timely manner.[3,13,19,20] Treatment most often consists of surgery or radiotherapy (RT) for initial clinical stage HNC and multimodality treatment (a combination of surgery, RT, and chemotherapy) for advanced patients.[1] Nowadays, oncologists can avoid surgery by using a combination of radiotherapy and chemotherapy in organ preservation strategies for selected patients, which have similar survival rates.[21,22] However, the standard up-to-date treatment needs a dedicated multidisciplinary team approach to implement advances in treatment and achieve higher survival[1]; otherwise patient outcomes can be jeopardized when the patient lives in a region where the healthcare facility has limited access to technology, skilled oncologists, multi-professional support, and rehabilitation, i.e., limited access to adequate health care.[17,23-25]

These advances in treatment for HNC in the last 30 years have resulted in an increase in the survival rates.[26,27] However, survival rates have not followed this trend at the same pace as treatment advances, nor are they the same across populations.[2,4,28] Recently, there are studies showing that the presence of comorbidities, patient’s performance status (a score that quantifies patients' general well-being and daily life activities), and other patient-related factors can influence survival rates.[13,29,30] Moreover, researchers are investigating how risk factors for HNC can also affect prognosis, for instance, gender, age at diagnosis, tobacco and alcohol consumption, HPV infection, and others are associated with survival.[3,15,23]

Regarding HNC and how its diagnosis and prognosis is related to SDOH, researchers have observed that SDOH not only influences the occurrence of HNC but also affects patients’ life after HNC diagnosis, producing a much more complex understanding of survival outcomes and possibly justifying why advances in treatment have not been enough to improve survival across the population. Studies have shown that low income and low education predict poor survival even when adjusting for covariates (age, health behaviors, medical comorbidities, and treatment modality).[13-18]

Race/Ethnicity has also been associated with HNC survival. In a previous study based on the SEER database, Blacks (compared to Whites) presented a worse prognosis for oral and oropharyngeal cancer; socioeconomic status was also associated with survival.[34] Another study, specifically for oropharynx cancer, also found that county-level SES, as well as Blacks presented with a worse overall survival.[35] This trend for a poorer prognosis among Blacks, even in multivariate models adjusting for SDOH, was observed in studies using other databases as well.[36,37] Also in studies comparing patients treated in a single institution.[38,39]

Regarding Asian/PI population, a study in British Columbia found that Asian/PI had better survival rates than other races/ethnicities, in an adjusted multivariate model.[41] However, among the Asian/PI population there is a study showing that lower neighborhood SES was associated with poorer prognosis.[40]

A study evaluating the AI/AN population with HNC and comparing it with Whites showed that they usually presented with more advanced stages of the disease in the oropharynx and were less likely to receive surgery alone for the treatment of oral cavity tumors. In a survival analysis stratified by tumor site, they observed that specifically for the oropharynx, the AI/AN population still presented with a worse prognosis in the adjusted model (adjusted for demographic characteristics, clinical stage, and treatment).[42]

For Hispanics, a case-matching analysis showed that in the univariate analysis, Hispanics presented with a worse prognosis than Whites. However, in the case-matching analysis (matched by age, gender, year of diagnosis, clinical stage, and treatment), the survival difference was no longer significant.[43]

It is important to notice that this gap in survival regarding race/ethnicity is neither new nor getting better. For instance, based on the SEER database, for oral cavity and pharynx cancers, Blacks present a worse survival rate than Whites for any clinical stage of the disease.[28] In the same publication, when comparing the trends in 5-year survival rates for the past 40 years, despite the fact that in the trend analysis, Whites and Blacks have significantly improved their survival rates for oral cavity and pharynx cancers, the 5-year survival rates for Afro-Americans in the period of 2005-2011 was lower than the rates presented for Whites in the period of 1975-1977 (45% vs. 54%, respectively).[28]

The extent to which these racial disparities reflect clinical characteristics (tumor site, stage at diagnosis, or treatment), unequal access to health care or other SDOH remains an active area of research.[44-46] It is true that age and medical comorbidities might be associated with a bias in treatment selection[47], as these factors could prevent the patient from receiving a more toxic/morbid treatment, even if the treatment would lead to a better prognosis. Yet, some authors argue that most of the association of SDOH is related to whether treatment is affordable or available for the most vulnerable populations.[17,30,47] The aim of this study is to investigate the racial/ethnic differences in head and neck cancer survival rates observed in the US in recent years, and whether these differences are related to patient and tumor characteristics, and SDOH.

II. Methods

Study Design

This is a retrospective study based on data from the Surveillance, Epidemiology, and End Results program (SEER) of the National Cancer Institute (NCI). The most recent SEER database (Nov 2016 Submission) is publicly available for researchers and includes over 8 million patients with malignant disease occurring between 1973-2014. Our study includes only head and neck carcinomas, which account for around 200,000 patients, before further selection.[48]

The SEER began collecting data on cancer cases on January 1, 1973, in the states of Connecticut, Iowa, New Mexico, Utah, and Hawaii, and the metropolitan areas of Detroit and San Francisco-Oakland. In 1974-1975, the metropolitan area of Atlanta and the 13-county Seattle-Puget Sound area were added. In 1978, 10 predominantly black rural counties in Georgia were added, followed in 1980 by the addition of American Indians residing in Arizona. Three additional geographic areas participated in the SEER program before 1990: New Orleans, Louisiana (1974-1977, rejoined 2001); New Jersey (1979-1989, rejoined 2001); and Puerto Rico (1973-1989). The NCI also funds a cancer registry that, with technical assistance from SEER, collects information on cancer cases among Alaskan Native populations. In 1992, the SEER Program was expanded to increase coverage of minority populations, especially Hispanics, by adding Los Angeles County and four counties in the San Jose-Monterey area south of San Francisco. In 2001, the SEER Program expanded coverage to include Kentucky and the remaining counties in California (Greater California); also, New Jersey and Louisiana once again became participants. In 2010, the SEER program expanded coverage to include all of the state of Georgia.[48]

The SEER currently collects and publishes cancer incidence and survival data from population-based cancer registries covering approximately 30 percent of the US population and it is the only comprehensive source of population-based information in the United States that includes the clinical stage of cancer at the time of diagnosis and patient survival data.[48]

Inclusion Criteria

Patients were included in the analysis if registered from 2007 to 2014 in the 18-cancer registry database of the SEER Public-Use Data. Patients should be age 18 years or older; have an invasive squamous cell carcinoma confirmed by histology and had the only or first primary tumor recorded during the period.

Regarding the tumor site, we selected only tumors originating from (based on ICD-10): oral cavity (C02–C05.0; C06); oropharynx (C01; C05.1,2; C09.0,1,9; C10.0,2,3); hypopharynx (C12; C13); and larynx (C32).

Exclusion Criteria

We excluded from the analysis patients with “data not available, missing, or unknown” for: race/ethnicity; county of residence; clinical stage, and health insurance status. Patients that could not be distinguished by specific site (oral cavity or pharynx), grouped as “other mouth/pharynx” in the database were also excluded as well as those with histology of carcinoma “*in situ*”.

Data Collection

The SEER Public-Use Data is a public database for research use. Data were obtained in ASCII format and transferred to SAS format (to keep all the coding for the variables) and subsequently transferred to the Statistical Package SPSS version 23.

For the purpose of the study, the SEER database offers the most accurate, reliable, comprehensive data information with the longest period of follow-up. The analysis of secondary data is already collected for a population-based study, and this is the best possible data available.

In addition to the individual (including clinical data) and county-level data obtained from SEER, other county-level attributes were obtained from: the National Center for Education Statistics[49], the U.S. Census Bureau's – American Community Survey[50], and the Community Health Rankings[51] (Appendixes I and II).

Analysis Plan

Key analysis variables on SEER database

There is detailed information on codes used for all SEER database variables and categories available online[48]. A brief description of the main variables used in the study follows below, stating how they were measured and grouped. Age was collected at the time of diagnosis as a continuous variable, and then grouped as: ≤ 50 years, 51 – 59 yrs, 60 – 59 yrs, and ≥ 70 yrs. Gender was coded in the dataset as male or female. For race/ethnicity, the original data was collected and recorded for 32 possibilities; for this analysis, we used the following classification: White, Black, Hispanic, Asian or Pacific Islander (A/PI), and American Indian/Alaska Native (AI/AN)

Regarding marital status, we grouped as: single (never married), married (including common law), and divorced/widowed (including separated). The county of residence was coded as representing the FIPS county code.

Cancer cases and tumor sites were classified according to the International Classification of Diseases for Oncology (ICD-10). For clinical stages, staging groups for analysis were defined as follows: localized (includes AJCC clinical stage I and II, or SEER stage localized), regional (includes AJCC clinical stage III, IVa and IVb, or SEER stage regional), and distant (includes AJCC clinical stage IVc or SEER stage distant). This combined criterion allowed to group the patients based on the best clinical staging data available AJCC stating system 7th ed. (2010+), AJCC staging system 6th ed. (2004+), or the SEER staging system. Sometimes, even with incomplete data from the 3 staging systems, there was enough information to stage grouping the patient. To ensure the feasibility of our methodology, we tested the staging we propose and compared the results with those of the three staging systems available, and the concordance was over 90%.

Concerning the treatment performed, there are questions regarding the completeness of the radiation or chemotherapy treatment variable in the SEER database – as patients could have it performed outside the tertiary care facility and might not be reported into the SEER Program – and the potential biases associated with the use of the radiation and chemotherapy data. In other words, there is a certainty when the treatment is said as performed in the database, but when the field is blank, there is a possibility that treatment was performed elsewhere and not recorded in the database. Thus, for the analysis, the treatment was grouped into the following categories: “surgery alone," or primary treatment based on surgery; “surgery plus radiotherapy” (pre or postoperatively); “radiotherapy," or primary treatment based on radiotherapy with or without chemotherapy; "chemotherapy alone;" "refused," which means the patient refused to accept the treatment; “recommended," which means treatment was recommended but there was no certainty that it was performed; and "untreated/unknown," this category included patients with no information on treatment in the SEER database.

County-level attribute variables

County-level attribute variables included in the analysis were originally collected from different time periods. Whenever possible, the chosen period was the one closest to, or that incorporated the midpoint of the patients' year of diagnosis (around 2010-2011).

The SEER database has some county-level attribute variables collected from different time periods. For the analysis, we selected the ones related to the 2009-2013 period[52]. They were obtained using the Census American Community Survey (ACS) 5-year files. The county-level attribute variables evaluated from the SEER database were those regarding: educational attainment, age distribution in the population, employment rate, foreign-born rate, linguistic isolation rate, median income, and poverty rate.

Details on definition, source, and years of collection regarding these county-level attribute variables are listed in Appendix I.

Moreover, as all cases have information on the individual's domicile, coded as representing the FIPS county code of residence, it was possible to crosslink the SEER database with other databases classifying the counties in variables that could reflect other SDOH: health literacy,[49,53] GINI index (income inequality).[50] Details on definition, source, and years of collection regarding these county-level attribute variables are listed in Appendix I.

Another source explored in this manner was the *County Health Rankings & Roadmaps*.[51] In this case, the rankings are based on a conceptual model of population health that includes Health Outcomes (length and quality of life) and Health Factors (social determinants of health).[51] As defined by the authors, after developing the Rankings model, weights for each component in the model and appropriate measures were established. The process for choosing weights and measures was guided by:[51] review of the literature around the impact of various factors on health outcomes; ability for factors to be modified through community action; review of America's Health Rankings methodology and indicators; availability and reliability of indicators at the county-level throughout the nation; analysis; feedback from a panel of technical experts; and standardizing measures. These outcomes and factors could be broken down into some subcomponents.[51] Briefly, the Health Outcomes Score is based on the length and quality of life, measured by morbidity and mortality rates in that county; the Health Factors Score is a combination of: 1) the Health Behaviors Score, which is based on county-level rates of tobacco, alcohol and drug use, diet and exercise, as well as sexual activity; 2) the Clinical Care Score, that is a combination of access to care and quality of care; 3) the Social & Economic Factors score, which is based on measurements of education level, employment rates, income, family and social support, and community safety; and 4) the Physical Environment Score, that is a combination of environmental quality (air and water quality) and the built environment (access to recreational facilities, access to health foods vs. fast food restaurants).[54] More details on the attributes, composition of the subcomponents, source, and year of data are listed on Appendix II.

Each measure was standardized within each state to the average of counties in that state. The measures are in a number of different scales—some are percentages, some are rates, some are averages of survey responses or other metrics. Standardizing each of these measures transforms them to the same metric—a mean (average) value of 0 and a standard deviation (measure of spread) of 1 (referred as Z-scores). Each Z-score is relative to the other counties in that state—not compared to an absolute standard—and shown in the metric of standard deviations. A positive Z-score indicates a value higher than the average of counties in that state, and usually indicates poorer health (e.g., years of potential life lost before age 75). However, for some of our measures (e.g., high school graduation) a higher score indicates better health or a more desirable value. To take this into account before computing summary scores, the authors computed the Z-score as usual but multiplied it by -1, so that higher scores indicate poorer health[51]. The formulas are listed in Appendix III. To de-emphasize the differences between individual county ranks, counties were grouped into quartiles according to their separate ranks. For each set of ranks, there are four quartiles that divide up all the units within the state into the top 25%, the second from the top 25%, the second from the bottom 25%, and the bottom 25%. The top 25% are the healthiest counties with the best ranks, the bottom 25% are the least healthy counties with the worst ranks, and the other two quartiles are in between.[54] The *County Health Rankings* dataset is publically available and was downloaded and subsequently linked with the main SEER database. The dataset used in the analysis was the one available in 2013.[51]

Statistical methods

Statistical analysis was performed using the Statistical Package SPSS version 23. Data distribution was described in tables regarding the main variables. The association of these variables with race/ethnicity was evaluated by the chi-square test.

For the cancer-specific survival analysis, the time period of follow-up was considered as the period from the date of diagnosis to the date of last contact or date of death. The outcome was defined as death due to cancer; patients were censored if alive or dead due to other causes. Five-year survival rates were estimated by the Kaplan-Meier method, and the significance of differences among survival curves was calculated by the log-rank test. Univariate and multivariable survival analyses were performed using the Cox proportional hazards model to estimate the hazard ratio (HR) with a 95% confidence interval for cancer-specific death.

For the multivariate models, the Health Outcome Score was not tested as a covariate, as it is a variable that (in the concept of the authors) represent a summary of all other scores[51]. Moreover, the scores were not tested in the same model with the individual SDOH variables (including the sub-components of the score) as it is a by-product of such variables. Moreover, as tumor site, clinical stage, and treatment performed are tightly linked, and they are the most important predictors of survival, a term of interaction for these variables was always included in the multivariate models presented.

For all statistical tests, the significance is 2-sided and achieved when *p*-values ≤ 0.05.

Ethical Considerations/Human subjects

The study is a secondary data analysis of the SEER research data (a public database). The database is considered to be anonymous. The NIH is responsible for the data and can grant access to the database for researchers. In order to have access to the SEER data, a signed SEER Research Data Agreement form is required. In signing this document, the researcher agrees that: 1) the data will not be used in any way other than for statistical reporting and analysis for research purposes; 2) all research results must be presented or published in a manner that ensures that no individual can be identified; and 3) there must be no attempt either to identify individuals from any computer file or to link with a computer file containing patient identifiers.

This agreement document was signed, and based on all this information, the study was categorized as an exempt status (as the study fits in the minimal risk and is based on public data) by the Human Subjects Division of the University of Washington; no formal approval from the IRB was necessary.

Results

The final sample size for analysis included 59,853 head and neck cancer patients. The majority of the patients were White (76.4%), followed by Blacks (11.2%), Hispanics (7.7%), Asian or Pacific Islander (4.2%) and American Indian/Alaska Native (0.5%). Ages ranged from 18 to 103 years (median 61 years; mean 61.9 years). Most patients were males (76.6%) and married (51.8%); living in metropolitan area counties (85.5%). Regarding health insurance, the majority had private insurance (78.7%), followed by Medicaid patients (16.0 %), and 5.3% of the patients were uninsured [Table 1].

Table 1 – Socio-demographics characteristics of the patients.

Variables	**Categories**	**n (%)**
Race/Ethnicity		
	White	45,712 (76.4)
	Black	6,689 (11.2)
	Hispanic	4,592 (7.7)
	Asian or Pacific Islander	2,532 (4.2)
	American Indian/Alaska Native	328 (0.5)
Age		
	≤ 50 yrs	8,826 (14.7)
	51 – 59 yrs	17,507 (29.2)
	60 – 69 yrs	18,791 (31.4)
	≥ 70 yrs	14,729 (24.6)
Gender		
	Male	45,930 (76.7)
	Female	13,923 (23.3)
Marital Status at Diagnosis	Single	11,796 (19.7)
	Married	31,020 (51.8)
	Divorced/Widowed	14,067 (23.5)
	Unknown	2,970 (5.0)
Health Insurance		
	Insured	47,118 (78.7)
	Medicaid	9,580 (16.0)
	Uninsured	3,155 (5.3)
Urban Area		
	Counties in Metropolitan Area	51,327 (85.8)
	Urban Adjacent to Metropolitan Area	4,367 (7.3)
	Urban Not Adjacent to Metropolitan Area	3,083 (5.2)
	Rural	1,013 (1.7)

The tumors were localized predominantly at the oropharynx (41.6%) followed by the oral cavity (24.9%), larynx (28.3%), and hypopharynx (5.1%). Not unexpectedly, the majority of the patients (62.9%) were diagnosed with advanced clinical stage (Regional – CS III, IVa, IVb); and just a few (3.7%) with distant metastasis. The most common treatment was based on radiotherapy (with or without chemotherapy - 42.5%) followed by surgery (alone – 19.1%; or in combination with radiotherapy – 29.0%). Less than 7% of the patients were not treated or their treatment was unknown [Table 2].

Table 2 – Clinical Characteristics of Tumors.

Variables	**Categories**	**n (%)**
Tumor Site		
	Oral Cavity	14,933 (24.9)
	Oropharynx	24,924 (41.6)
	Larynx	16,939 (28.3)
	Hypopharynx	3,057 (5.1)
Clinical Stage		
	Localized (CS I, II)	20,030 (33.5)
	Regional (CS III,IVa,IVb)	37,619 (62.9)
	Distant (CS IVc)	2,204 (3.7)
Treatment		
	Surgery Alone	11,402 (19.1)
	Surgery + RT (+/- CH)	17,338 (29.0)
	RT (+/- CH)	25,452 (42.5)
	Chemotherapy Alone	1,539 (2.6)
	Recommended*	553 (0.9)
	Refused	548 (0.9)
	Untreated/UNK	3,021 (5.0)

* Recommended, but uncertain if performed
CS — clinical stage
RT — radiotherapy
CH — chemotherapy

Regarding the socio-demographic distribution and clinical characteristics across race/ethnicity lines, all the evaluated variables showed significant changes. The Asian/PI tended to be older than other racial/ethnic groups, and Blacks were the youngest of all racial/ethnic groups (p<0.001). The gender distribution was similar (around 3:1, male:female ratio) across Whites, Blacks, and Hispanics, but not for Asian/PI (which tended to be more females – 2.5:1, male:female ratio) and AI/AN (which tended to be more males – 4:1, male:female ratio) (p<0.001). Black patients were more likely to be single, while AI/AN were more likely to be divorced (p<0.001). Regarding the site of residency, Hispanics and Asian/PI tended to live in metropolitan areas in a higher proportion (p<0.001) [Table 3].

Table 3 – Socio-demographics characteristics of the patients according to race/ethnicity.

Variables	White **n (%)**	Black **n (%)**	Hispanic **n (%)**	Asian/PI **n (%)**	AI/AN **n (%)**	**p-value**
Age						
≤ 50 yrs	6,473 (14.2)	1,031 (15.4)	823 (17.9)	447 (17.7)	52 (15.9)	<0.001
51 – 59 yrs	13,229 (28.9)	2,317 (34.6)	1,297 (28.2)	556 (22.0)	108 (32.9)	
60 – 59 yrs	14,584 (31.9)	2,141 (32.0)	1,244 (27.1)	712 (28.1)	110 (33.5)	
≥ 70 yrs	11,426 (25.0)	1,200 (17.9)	1,228 (26.7)	817 (32.3)	58 (17.7)	
Gender						
Male	35,135 (76.9)	5,132 (76.7)	3,579 (77.9)	1,821 (71.9)	263 (80.2)	<0.001
Female	10,577 (23.1)	1,557 (23.3)	1,013 (22.1)	711 (28.1)	65 (19.8)	
Marital Status						
Single	7,734 (16.9)	2,578 (38.5)	1,095 (23.8)	325 (12.8)	64 (19.5)	<0.001
Married	24,827 (54.3)	2,039 (30.5)	2,339 (50.9)	1,683 (66.5)	132 (40.2)	
Divorced/Widowed	10,911 (23.9)	1,736 (26.0)	916 (19.9)	398 (15.7)	106 (32.3)	
Unknown	2,240 (4.9)	336 (5.0)	242 (5.3)	126 (5.0)	26 (7.9)	
Health Insurance						
Insured	37,973 (83.1)	4,001 (59.8)	3,046 (66.3)	1,901 (75.1)	197 (60.1)	<0.001
Medicaid	5,647 (12.4)	2,094 (31.3)	1,203 (26.2)	515 (20.3)	121 (36.9)	
Uninsured	2,092 (4.6)	594 (8.9)	343 (7.5)	116 (4.6)	10 (3.0)	
Urban Area						
Counties in MA	38,386 (84.0)	5,931 (88.7)	4,382 (95.4)	2,427 (95.9)	201 (61.3)	<0.001
Urban adjacent MA	3,721 (8.1)	490 (7.3)	102 (2.2)	22 (0.9)	32 (9.8)	
Urban not adj. MA	2,674 (5.8)	198 (3.0)	97 (2.1)	83 (3.3)	31 (9.5)	
Rural	930 (2.0)	70 (1.0)	11 (0.2)	0 0.0)	2 (0.6)	

MA – metropolitan area
AI/AN - American Indian/Alaska Native
Asian/PI - Asian or Pacific Islander

For the clinical characteristics, oropharynx tumors were more common in Whites, Hispanics, and AI/AN, while larynx cancer was the most common site among Blacks; oral cavity was the most common among Asian/PI. Hypopharynx was less common in all races/ethnicities but was proportionally more common among AI/AN (p<0.001). All races/ethnicities presented predominantly with advanced disease, but Asian/PI presented proportionally more localized tumors compared to the others races/ethnicities. The distant disease was less common, but a higher proportion was observed among Blacks (p<0.001). Treatment based on radiotherapy was more common among Blacks and AI/AN, while treatment based on surgery (alone or with radiotherapy) was more commonly performed in Whites, Hispanics, and Asian/PI (p<0.001) [Table 4].

Table 4 – Clinical characteristics of the patients according to race/ethnicity.

Variables	White **n (%)**	Black **n (%)**	Hispanic **n (%)**	Asian/PI **n (%)**	AI/AN **n (%)**	**p-value**
Tumor site						
Oral cavity	11,353 (24.8)	1,093 (16.3)	1,263 (27.5)	1,147 (45.3)	77 (23.5)	<0.001
Oropharynx	20,000 (43.8)	2,430 (36.3)	1,684 (36.7)	676 (26.7)	134 (40.9)	
Larynx	12,296 (26.9)	2,637 (39.4)	1,389 (30.2)	536 (21.2)	81 (24.7)	
Hypopharynx	2,063 (4.5)	529 (7.9)	256 (5.6)	173 (6.8)	36 (11.0)	
Clinical Stage						
Localized (CS I, II)	15,799 (34.6)	1,656 (24.8)	1,473 (32.1)	1,028 (40.6)	74 (22.6)	<0.001
Regional (CS III, IVa,b)	28,451 (62.2)	4,595 (68.7)	2,912 (63.4)	1,423 (56.2)	238 (72.6)	
Distant (CS IVc)	1,462 (3.2)	438 (6.5)	207 (4.5)	81 (3.2)	16 (4.9)	
Treatment						
Surgery alone	9,026 (19.7)	739 (11.0)	868 (18.9)	724 (28.6)	45 (13.7)	<0.001
Surgery + RT (+/- CH)	13,424 (29.4)	1,601 (23.9)	1,384 (30.1)	831 (32.8)	98 (29.9)	
RT (+/- CH)	19,296 (42.2)	3,389 (50.7)	1,849 (40.3)	768 (30.3)	150 (45.7)	
CH alone	1,062 (2.3)	278 (4.2)	146 (3.2)	42 (1.7)	11 (3.4)	
Recommended*	376 (0.8)	95 (1,4)	54 (1.2)	24 (0.9)	4 (1.2)	
Refused	405 (0.9)	67 (1.0)	42 (0.9)	33 (1.3)	1 (0.3)	
Untreated/UNK	2,123 (4.6)	520 (7.8)	249 (5.4)	110 (4.3)	19 (5.8)	

* Recommended, but uncertain if performed
CS – clinical stage
RT – radiotherapy
CH – chemotherapy
AI/AN – American Indian/Alaska Native
Asian/PI – Asian or Pacific Islander

Regarding the county-level attributes and the association with race/ethnicity, Blacks and Hispanics lived in counties with a lower proportion of individuals ages 65 or older (p<0.001). When analyzing the economic and wealth county-level attributes, Blacks live in counties with the highest proportion of families below poverty level and persons below 150% of the poverty level, followed by Hispanics; in contrast, Asian/PI tend to live in wealthier counties (p<0.001). Blacks and Hispanics tend to live in counties with higher rates of unemployment; Whites and Asian/PI live in counties with lower rates; AI/AN seem to present a different distribution, with a fair amount residing in a low unemployment rate county and another fair amount in the high rate counties (p<0.001). The median income was higher in counties were Asian/PI and Whites live; while 43.2% of the Blacks live in the lower median household income counties (p<0.001). According to the county-level inequality index, measured by the GINI coefficient, the majority of the Blacks (51.2%) lived in counties with the highest inequality index, followed by the Hispanics and the Asian/PI. Whites presented a better distribution regarding the GINI coefficient in counties where they live; the best distribution among the groups was observed among the AI/AN (p<0.001) [Table 5].

Concerning the education attributes, AI/AN lived in counties with lower rates of people with at least bachelor's degree; Asian/PI lived in counties with the higher rates, followed by Hispanics (p<0.001). Low health literacy was observed predominantly in counties where Hispanics and Asian/PI lived, followed by the AI/AN, and Blacks (p<0.001).

Regarding place of birth, Asian/PI and Hispanics lived in counties with the highest rates of foreign-born people, which are also in the counties with higher rates of linguistic isolation (p<0.001) [Table 6].

Table 5 – County-level socioeconomic characteristics of the patients according to race/ethnicity.

Variables **County-level**	White **n (%)**	Black **n (%)**	Hispanic **n (%)**	Asian/PI **n (%)**	AI/AN **n (%)**	**p-value**
Persons Age >65yrs						
Less than 11.2%	10,760 (23.5)	2,002 (29.9)	2,098 (45.7)	705 (27.8)	113 (34.4)	<0.001
11.3 to 12.6%	11,034 (24.1)	1,451 (21.7)	1,225 (26.7)	796 (31.4)	78 (23.8)	
12.7 to 14.3%	11,310 (24.7)	2,311 (34.6)	762 (16.6)	409 (16.2)	61 (18.6)	
More than 14.4%	12,608 (27.7)	925 (13.8)	507 (11.0)	622 (24.6)	76 (23.2)	
Families Below Poverty						
Less than 8.1%	12,846 (28.1)	663 (9.9)	700 (15.2)	1,051 (41.5)	122 (37.2)	<0.001
8.2 to 11.1%	11,988 (26.2)	971 (14.5)	1,062 (23.1)	596 (23.5)	68 (20.7)	
11.2 to 14.2%	10,811 (23.7)	2,157 (32.2)	1,807 (39.4)	681 (26.9)	74 (22.6)	
More than 14.3%	10,067 (22.0)	2,898 (43.3)	1,023 (22.3)	204 (8.1)	64 (19.5)	
Persons <150% of Poverty						
Less than 19.1%	12,661 (27.7)	772 (11.5)	766 (16.7)	1,061 (41.9)	118 (36.0)	<0.001
19.2 to 25.3%	12,132 (26.5)	1,132 (16.9)	996 (21.7)	607 (24.0)	70 (21.3)	
25.4 to 29.8%	10,309 (22.6)	2,038 (30.5)	1,879 (40.9)	681 (26.9)	58 (17.7)	
More than 29.9%	10,610 (23.2)	2,747 (41.1)	951 (20.7)	183 (7.2)	82 (25.0)	
Unemployment						
Less than 8.9%	12,243 (26.8)	1,239 (18.5)	557 (12.1)	818 (32.3)	116 (35.4)	<0.001
9.0 to 10.3%	12,391 (27.1)	963 (14.4)	995 (21.7)	583 (23.0)	60 (18.3)	
10.4 to 12.0%	10,467 (22.9)	2,082 (31.1)	1,806 (39.3)	813 (32.1)	62 (18.9)	
More than 12.1%	10,611 (23.2)	2,405 (36.0)	1,234 (26.9)	318 (12.6)	90 (27.4)	
Median household income (USD)						
Less than 46,960	11,964 (26.2)	2,887 (43.2)	534 (11.6)	103 (4.1)	89 (27.1)	<0.001
46,961 to 55,910	11,115 (24.3)	1,873 (28.0)	1,996 (43.5)	729 (28.8)	58 (17.7)	
55,911 to 65,590	11,235 (24.6)	1,089 (16.3)	1,041 (22.7)	379 (15.0)	64 (19.5)	
More than 65,591	11,398 (24.9)	840 (12.6)	1,021 (22.2)	1,321 (52.2)	117 (35.7)	
GINI index						
Less than 0.43	12,638 (27.6)	830 (12.4)	395 (8.6)	660 (26.1)	145 (44.2)	<0.001
0.44 to 0.46	11,712 (25.6)	966 (14.4)	1,264 (27.5)	573 (22.6)	86 (26.2)	
0.47 to 0.48	12,348 (27.0)	1,465 (21.9)	1,299 (28.3)	586 (23.1)	69 (21.0)	
More than 0.49	9,014 (19.7)	3,428 (51.2)	1,634 (35.6)	713 (28.2)	28 (8.5)	

Table 6 – County-level education, foreign-born, and linguistic isolation of the patients according to race/ethnicity.

Variables **County-level**	White **n (%)**	Black **n (%)**	Hispanic **n (%)**	Asian/PI **n (%)**	AI/AN **n (%)**	**p-value**
Bachelor's Degree						
Less than 20.5%	11,811 (25.8)	1,620 (24.2)	887 (19.3)	172 (6.8)	85 (25.9)	<0.001
20.6 to 29.7%	11,378 (24.9)	1,988 (29.7)	684 (14.9)	363 (14.3)	151 (46.0)	
29.8 to 34.6%	11,016 (24.1)	1,694 (25.3)	1,939 (42.2)	1,066 (42.1)	60 (18.3)	
More than 34.7%	11,507 (25.2)	1,387 (20.7)	1,082 (23.6)	931 (36.8)	32 (9.8)	
Low Health Literacy						
Less than 10%	15,517 (33.9)	851 (12.7)	485 (10.6)	319 (12.6)	94 (28.7)	<0.001
10.1 to 14.1%	9,904 (21.7)	2,081 (31.1)	638 (13.9)	340 (13.4)	58 (17.7)	
14.2 to 20.4%	10,666 (23.3)	1,956 (29.2)	1,055 (23.0)	1,036 (40.9)	66 (20.1)	
More than 20.5%	9,625 (21.1)	1,801 (26.9)	2,414 (52.6)	837 (33.1)	110 (33.5)	
Foreign-born						
Less than 5.7%	12,626 (27.6)	2,040 (30.5)	212 (4.6)	61 (2.4)	68 (20.7)	<0.001
5.8 to 12.1%	12,115 (26.5)	1,911 (28.6)	600 (13.1)	253 (10.0)	133 (40.5)	
12.2 to 23.4%	11,108 (24.3)	1,305 (19.5)	1,367 (29.8)	900 (35.5)	81 (24.7)	
More than 23.5%	9,863 (21.6)	1,433 (21.4)	2,413 (52.5)	1,318 (52.1)	46 (14.0)	
Linguistic Isolation (household)						
Less than 1.9%	13,242 (29.0)	1,939 (29.0)	150 (3.3)	71 (2.8)	56 (17.1)	<0.001
2.0 to 4.1%	11,589 (25.4)	2,097 (31.3)	417 (9.1)	309 (12.2)	134 (40.9)	
4.2 to 8.1%	11,498 (25.2)	1,266 (18.9)	1,555 (33.9)	908 (35.9)	89 (27.1)	
More than 8.2%	9,383 (20.5)	1,387 (20.7)	2,470 (53.8)	1,244 (49.1)	49 (14.9)	

Asian/PI – Asian or Pacific Islander
AI/AN – American Indian/Alaska Native

When analyzing the Health Community Rankings, a higher proportion of Asian/PI live in counties belonging to the upper 25% ranked counties for the Health Outcome score followed by the Whites, while Blacks are more likely to live in counties on the lower-middle or bottom 25% (p<0.001). For the Health Behavior score rank, AI/AN are more likely to live in the bottom 25% ranked counties, followed by Blacks; Asian/PI live in the upper ranked counties, followed by the Whites and Hispanics (p<0.001). Clinical Care score rank showed a different pattern: the Hispanics tended to live in the bottom 25% ranked counties; Asian/PI, followed by Whites and Blacks were more frequently in the upper ranked counties. Concerning the Social & Economic Factors score, yet another pattern is observed: the Blacks, along with the AI/AN were more frequently living in the bottom ranked counties, while Whites and Asian/PI were in the upper ranked counties (p<0.001). Finally, for the Physical Environment score, Blacks were more likely living in the bottom ranked counties, while Asian/PI and Whites were in the upper ranked counties (p<0.0001) [Table 7].

It is interesting to notice that considering all variables analyzed so far, the Whites presented, overall, the best distribution according to the county-level attributes. Blacks were more frequently living in the worse counties. Asian/PI, Hispanics, and AI/AN have a mix of better and worse distributions depending on the county-level attribute being evaluated.

Table 7 – County-level attributes of the Community Health Rankings according to race/ethnicity.

Variables	White	Black	Hispanic	Asian/PI	AI/AN	
County-level	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**p-value**
Health Outcome Score						
Bottom 25%	7,713 (16.9)	1,811 (27.1)	683 (14.9)	200 (7.9)	74 (22.6)	<0.001
Lower-middle	8,426 (18.4)	1,356 (20.2)	739 (16.1)	268 (10.6)	120 (36.6)	
Upper-middle	14,731 (32.2)	1,979 (29.6)	2,232 (48.6)	974 (38.5)	76 (23.2)	
Upper 25%	14,842 (32.5)	1,543 (23.1)	938 (20.4)	1,090 (43.0)	58 (17.6)	
Health Behavior Score						
Bottom 25%	8,406 (18.4)	1,812 (27.1)	819 (17.8)	245 (9.7)	143 (43.6)	<0.001
Lower-middle	8,164 (17.9)	1,048 (15.7)	698 (15.2)	259 (10.2)	66 (20.1)	
Upper-middle	12,424 (27.2)	1,784 (26.7)	2,001 (43.6)	823 (32.5)	63 (19.2)	
Upper 25%	16,718 (36.5)	2,045 (30.5)	1,074 (23.4)	1,205 (47.6)	56 (17.1)	
Clinical Care Score						
Bottom 25%	9,258 (20.3)	2,031 (30.4)	1,996 (43.5)	641 (25.3)	56 (17.1)	<0.001
Lower-middle	8,122 (17.8)	929 (13.9)	847 (18.4)	329 (13.0)	126 (38.4)	
Upper-middle	12,113 (26.5)	1,146 (17.1)	880 (19.2)	518 (20.5)	77 (23.5)	
Upper 25%	16,219 (35.4)	2,583 (38.6)	869 (18.9)	1,044 (41.2)	69 (21.0)	
Social & Economic Factors Score						
Bottom 25%	7,111 (15.6)	1,978 (29.6)	825 (18.0)	233 (9.2)	111 (33.8)	<0.001
Lower-middle	12,785 (28.0)	1,969 (29.4)	2,023 (44.1)	694 (27.4)	90 (27.4)	
Upper-middle	10,179 (22.3)	1,748 (26.1)	654 (14.2)	435 (17.2)	53 (16.2)	
Upper 25%	15,637 (34.1)	994 (14.9)	1,090 (23.7)	1,170 (46.2)	74 (22.6)	
Physical Environment Score						
Bottom 25%	9,309 (20.4)	1,880 (28.1)	829 (18.1)	590 (23.3)	58 (17.7)	<0.001
Lower-middle	10,282 (22.5)	1,377 (20.6)	765 (16.7)	353 (13.9)	159 (48.5)	
Upper-middle	13,897 (30.4)	1,664 (24.9)	2,140 (46.5)	893 (35.3)	65 (19.8)	
Upper 25%	12,224 (26.7)	1,768 (26.4)	858 (18.7)	696 (27.5)	46 (14.0)	

Asian/PI - Asian or Pacific Islander
AI/AN - American Indian/Alaska Native

The observed disease-specific 5-year survival considering all patients was 66.3% (Figure 1). Disease-specific 5-year survival according to race was 68.8% for Whites; 66.9% for Asian/PI; 63.1% for Hispanics; 62.0% for AI/AN, and 51.4% for Blacks (p<0.001) [Figure 2].

Figure 1 – Disease-specific survival curve

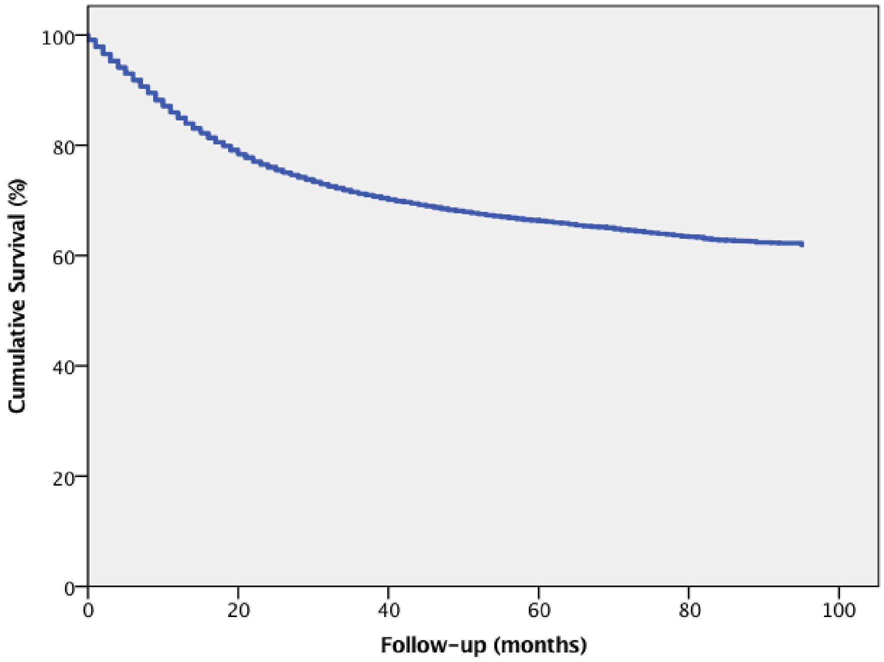
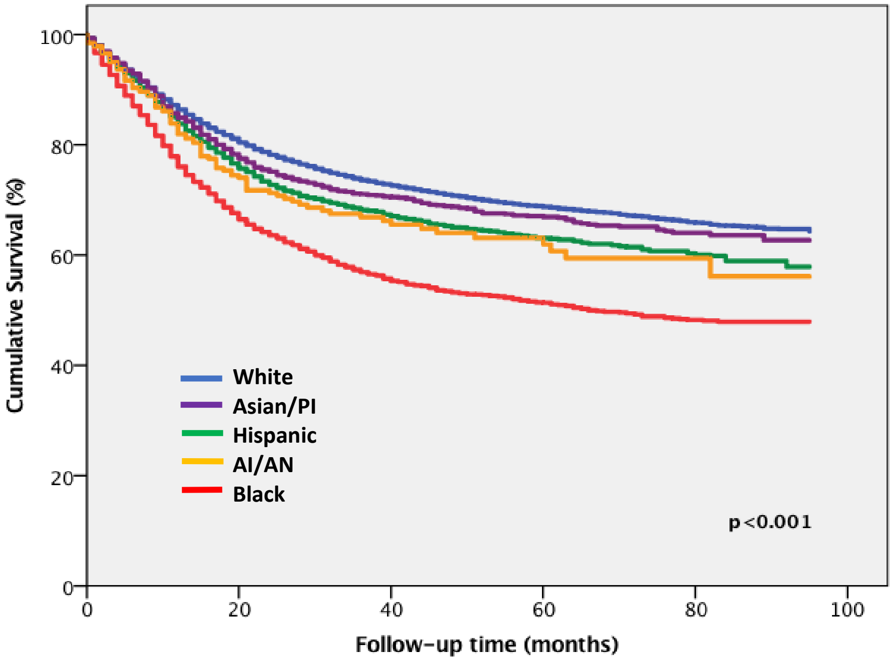


Figure 2 – Disease-specific survival curve according to the race/ethnicity



In the univariate Cox regression analysis, all variables showed a statistically significant difference in prognosis. When analyzing races/ethnicity, we could observe that not only Blacks but all other races/ethnicities presented a worse prognosis when compared to Whites (p<0.001). Blacks, in particular, presented a much worse risk of death due to the disease, HR 1.799 (95%CI, 1.722–1.879); for Hispanics, the HR was 1.231 (95%CI, 1.161 – 1.304); for Asian/PI the HR was 1.096 (95%CI, 1.011 – 1.187), and for AI/AN the HR was 1.318 (95%CI, 1.077 – 1.613). Regarding age, the patients over 70 years old presented twice the risk of dying from cancer than patients below 50 years old (HR 2.032, 95%CI, 1.926 – 2.144; p<0.001). Females presented 10% higher risk of death due to cancer (HR 1.107, 95%CI, 1.067 – 1.148; p<0.001) when compared to males and married people 50% less risk (HR 0.551, 95%CI, 0.530 – 0.574; p<0.001) than single patients. Living in the urban area (not metropolitan area) increased the risk of death due to cancer compared with those living in a metropolitan area (p<0.001), but not for those living in rural areas (p=0.239) [Table 8].

For health insurance, those patients under Medicaid presented 2 times higher risk of death due to cancer and the uninsured patients showed a 60% increased risk when compared with patients with private insurance (HR 1.967, 95%CI, 1.894 – 2.043; and HR 1.619, 95%CI, 1.520 – 1.724; p<0.001; respectively) [Table 8].

Table 8 – Univariate Cox regression survival analysis considering patients' characteristics.

Variables	HR	(95% CI)	p-value
Race/Ethnicity			
White	Ref.		
Black	1.799	(1.722 – 1.879)	<0.001
Hispanic	1.231	(1.161 – 1.304)	<0.001
Asian/PI	1.096	(1.011 – 1.187)	0.025
AI/AN	1.318	(1.077 – 1.613)	0.007
Age			
≤ 50 years	Ref.		
51 – 59 yrs	1.208	(1.143 – 1.276)	<0.001
60 – 59 yrs	1.297	(1.228 – 1.369)	<0.001
≥ 70 yrs	2.032	(1.926 – 1.144)	<0.001
Gender			
Male	Ref.		
Female	1.107	(1.067 – 1.148)	<0.001
Marital Status			
Single	Ref.		
Married	0.551	(0.530 – 0.574)	<0.001
Divorced/Widowed	1.029	(0.986 – 1.075)	0.190
Urban Area			
Counties in MA	Ref.		
Urban adjacent MA	1.082	(1.019 – 1.149)	0.010
Urban not adj. MA	1.128	(1.053 – 1.208)	0.001
Rural	1.074	(0.954 – 1.210)	0.239
Health Insurance			
Insured	Ref.		
Medicaid	1.967	(1.894 – 2.043)	<0.001
Uninsured	1.619	(1.520 – 1.724)	<0.001

MA – metropolitan area
Asian/PI - Asian or Pacific Islander
AI/AN - American Indian/Alaska Native

In the univariate analysis considering the tumors' characteristics, patients with tumors of the oropharynx (HR 0.755, 95%CI 0.725 – 0.786; p<0.001) and larynx (HR 0.845, 95%CI 0.810 – 0.882; p<0.001) presented a better prognosis than oral cavity cancer patients, while patients with hypopharynx tumors presented 2 times higher risk of dying from cancer (HR 1.956, 95%CI 1.842 – 2.078; p<0.001). Regional stage disease presented a 2.5 higher risk of death due to cancer when compared to localized tumors (p<0.001), while for the distant stage the risk was 10 times higher (p<0.001). Treatment performed also significantly influenced the survival of the patients, with the best survival being for patients undergoing surgery alone, while the worse prognosis for those treated patients was for chemotherapy alone (p<0.001) [Table 9]. It is important to notice that treatment choice depended mainly on tumor site and clinical stage at presentation (data not shown).

Table 9 – Univariate Cox regression survival analysis considering tumors' characteristics.

Variables		**HR**	**(95% CI)**	**p-value**
Tumor Site				
	Oral cavity	Ref.		
	Oropharynx	0.755	(0.725 – 0.786)	<0.001
	Larynx	0.845	(0.810 – 0.882)	<0.001
	Hypopharynx	1.956	(1.842 – 1.078)	<0.001
Clinical Stage				
	Localized (CS I, II)	Ref.		
	Regional (CS III, IVa,b)	2.555	(2.450 – 2.663)	<0.001
	Distant (CS IVc)	10.379	(9.740 – 11.059)	<0.001
Treatment				
	Surgery alone	Ref.		
	Surgery + RT (+/– CH)	1.272	(1.202 – 1.347)	<0.001
	RT (+/– CH)	1.862	(1.767 – 1.961)	<0.001
	CH alone	6.128	(5.651 – 6.646)	<0.001
	Recommended*	9.932	(8.874 – 11.116)	<0.001
	Refused	4.592	(4.021 – 5.244)	<0.001
	Untreated/UNK	7.525	(7.038 – 8.047)	<0.001

* Recommended, but uncertain if performed
CS – clinical stage
RT – radiotherapy
CH – chemotherapy

The univariate Cox regression analysis regarding the county-level attributes showed that in counties with a higher proportion of an older population, a better prognosis was observed for HNC patients (p=0.001). For the socioeconomic attributes, patients living in counties with a higher proportion of families below the poverty level or persons below 150% of the poverty level presented a 40% increased risk of dying from their cancer (p<0.001). Counties with higher rates of unemployment also showed worse survival (p<0.001). Considering the median household income, patients living in counties with a higher median income presented up to 30% decrease in the risk of death due to cancer (p<0.001). For counties with the highest inequality (GINI index) the risk of death was 30% higher (p<0.001) [Table 10].

Table 10 – Univariate Cox regression survival analysis considering county-level socioeconomic attributes.

Variables	HR	(95% CI)	p-value
Persons Age >65yrs			
Less than 11.2%	Ref.		
11.3 to 12.6%	0.926	(0.885 – 0.968)	0.001
12.7 to 14.3%	0.974	(0.932 – 1.018)	0.242
More than 14.4%	0.931	(0.891 – 0.974)	0.002
Families Below Poverty			
Less than 8.1%	Ref.		
8.2 to 11.1%	1.105	(1.054 – 1.158)	<0.001
11.2 to 14.2%	1.340	(1.281 – 1.402)	<0.001
More than 14.3%	1.451	(1.387 – 1.518)	<0.001
Persons <150% of Poverty			
Less than 19.1%	Ref.		
19.2 to 25.3%	1.095	(1.045 – 1.148)	<0.001
25.4 to 29.8%	1.297	(1.239 – 1.357)	<0.001
More than 29.9%	1.429	(1.366 – 1.495)	<0.001
Unemployed			
Less than 8.9%	Ref.		
9.0 to 10.3%	0.961	(0.918 – 0.007)	0.096
10.4 to 12.0%	1.119	(1.070 – 1.170)	<0.001
More than 12.1%	1.224	(1.170 – 1.279)	<0.001
Median Household Income (USD)			
Less than 46,960	Ref.		
46,961 to 55,910	0.894	(0.857 – 0.932)	<0.001
55,911 to 65,590	0.797	(0.763 – 0.834)	<0.001
More than 65,591	0.708	(0.676 – 0.740)	<0.001
GINI Index			
Less than 0.43	Ref.		
0.44 to 0.46	1.091	(1.042 – 1.143)	<0.001
0.47 to 0.48	1.067	(1.019 – 1.117)	0.005
More than 0.49	1.284	(1.228 – 1.343)	<0.001

Education related attributes analysis showed that higher the proportion of persons in the county with at least bachelor's degree, the lower the risk of dying from cancer (p<0.001). On the other hand, counties with a higher proportion of people with low health literacy presented with a worse prognosis (p<0.001) [Table 11].

People living in counties with a higher proportion of foreign-born individuals and higher rates of linguistic isolation presented with a better prognosis, around 5 to 10% decrease in the risk of death (p<0.001) [Table 11].

Table 11 – Univariate Cox regression survival analysis considering county-level education, foreign-born and linguistic isolation.

Variables	HR	(95% CI)	p-value
Bachelor's Degree			
Less than 20.5%	Ref.		
20.6 to 29.7%	0.940	(0.900 – 0.982)	0.006
29.8 to 34.6%	0.878	(0.841 – 0.918)	<0.001
More than 34.7%	0.782	(0.747 – 0.818)	<0.001
Low health Literacy			
Less than 10%	Ref.		
10.1 to 14.1%	1.253	(1.197 – 1.312)	<0.001
14.2 to 20.4%	1.267	(1.213 – 1.324)	<0.001
More than 20.5%	1.269	(1.215 – 1.327)	<0.001
Foreign-born			
Less than 5.7%	Ref.		
5.8 to 12.1%	0.943	(0.902 – 0.985)	0.009
12.2 to 23.4%	0.890	(0.851 – 0.931)	<0.001
More than 23.5%	0.924	(0.884 – 0.966)	<0.001
Linguistic Isolation (household)			
Less than 1.9%	Ref.		
2.0 to 4.1%	0.981	(0.938 – 1.025)	0.388
4.2 to 8.1%	0.930	(0.890 – 0.972)	0.001
More than 8.2%	0.987	(0.944 – 1.031)	0.549

For all Community Health Rankings scores evaluated, there is a gradient of worsening prognosis when comparing the upper 25% ranked counties to the bottom 25%. The highest risks of death due to the cancer were observed for the bottom 25% ranked for Health Outcomes score (HR 1.318, 95%CI 1.258 – 1.381; p<0.001) and for Social & Economic Factors score (HR 1.348, 95%CI 1.286 – 1.412; p<0.001). Followed by the Health Behavior score (HR 1.245, 95%CI 1.191 – 1.301; p<0.001); the Clinical Care score (HR 1.197, 95%CI 1.149 – 1.248; p<0.001); and the Physical Environment score (HR 1.133, 95%CI 1.082 – 1.186; p<0.001) [Table 12].

Table 12 – Univariate Cox regression survival analysis considering Community Health Rankings attributes.

Variables		**HR**	**(95% CI)**	**p-value**
Health Outcome Score				
	Upper 25%	Ref.		
	Upper-middle	1.128	(1.084 – 1.175)	<0.001
	Lower-middle	1.211	(1.156 – 1.269)	<0.001
	Bottom 25%	1.318	(1.258 – 1.381)	<0.001
Health Behavior Score				
	Upper 25%	Ref.		
	Upper-middle	1.097	(1.054 – 1.142)	<0.001
	Lower-middle	1.154	(1.102 – 1.209)	<0.001
	Bottom 25%	1.245	(1.191 – 1.301)	<0.001
Clinical Care Score				
	Upper 25%	Ref.		
	Upper-middle	1.017	(0.974 – 1.061)	0.446
	Lower-middle	1.092	(1.042 – 1.144)	<0.001
	Bottom 25%	1.197	(1.149 – 1.248)	<0.001
Social & Economic Factors Score	Upper 25%	Ref.		
	Upper-middle	1.147	(1.096 – 1.200)	<0.001
	Lower-middle	1.248	(1.198 – 1.301)	<0.001
	Bottom 25%	1.348	(1.286 – 1.412)	<0.001
Physical Environment Score				
	Upper 25%	Ref.		
	Upper-middle	1.011	(0.969 – 1.055)	0.623
	Lower-middle	1.049	(1.001 – 1.098)	0.044
	Bottom 25%	1.133	(1.082 – 1.186)	<0.001

The multivariate Cox regression was built in different models, in which a group of variables was added to each model. In model 1, which included patients' characteristics (adjusted for the SEER registry and year of diagnosis), it is possible to observe that all races/ethnicity still present with a worse prognosis when compared to Whites (p<0.001), but in this model, we can also observe that the risk has decreased most for the Blacks, from an HR of 1.799 (95%CI 1.722 – 1.879; p<0.001) in the univariate analysis to a HR of 1.588 (95%CI 1.516 – 1.663; p<0.001). The risk of death due to cancer increased for the Asian/PI population, from an HR 1.096 (95%CI 1.011 – 1.187; p=0.025) to an HR of 1.181 (95%CI 1.083 – 1.287; p<0.001). It is interesting to notice that gender changed the risk in this model when compared to the univariate analysis, where there was a higher risk of death due to cancer among females HR 1.107 (95%CI 1.067 – 1.148; p<0.001), and we can now observe a 10% decreased risk HR 0.958 (95%CI 0.922 – 0.994; p=0.024). The other independent prognostic factors in this model were age (p<0.001) and marital status (p<0.001) [Table 13].

Table 13 – Multivariate Cox regression survival analysis considering patients’ characteristics (Model 1)

Variables		**HR**	**(95% CI)**	**p-value**
Race/Ethnicity	White	Ref.		
	Black	1.588	(1.516 – 1.663)	<0.001
	Hispanic	1.208	(1.137 – 1.284)	<0.001
	Asian/PI	1.181	(1.083 – 1.287)	<0.001
	AI/AN	1.288	(1.028 – 1.613)	0.028
Age	≤ 50 yrs	Ref.		
	51 – 59 yrs	1.225	(1.159 – 1.295)	<0.001
	60 – 59 yrs	1.371	(1.298 – 1.448)	<0.001
	≥ 70 yrs	2.187	(2.070 – 2.311)	<0.001
Marital Status	Single	Ref.		
	Married	0.535	(0.513 – 0.557)	<0.001
	Divorced/Widowed	0.914	(0.874 – 0.957)	<0.001
Gender	Male	Ref.		
	Female	0.958	(0.922 – 0.994)	0.024

Model adjusted for SEER registry and year of diagnosis

Model 2 was built as model 1 plus tumor characteristics (considered the most important prognostic factors for HNC) and was adjusted for age, gender, SEER registry, year of diagnosis and an interaction term (tumor site, clinical stage, and treatment). It is possible to observe that no longer do all races/ethnicity present with a worse prognosis when compared to Whites; in this model, the AI/AN population presented with a similar risk of dying from cancer as Whites (HR 0.954, 95%CI 0.761 – 1.195; p=0.680), and Asian/PI also closed the gap, showing a much smaller statistical difference (HR 1.099; 95% CI 1.008 – 1.195; p=0.033). The Blacks and Hispanic still have significantly worse prognosis than Whites, but the excess risk for Blacks decreased by more than half (HR 1.329, 95%CI 1.267 – 1.393; p<0.001) when compared to the univariate analysis (HR 1.799, 95%CI 1.722 – 1.879; p<0.001), while the excess risk for Hispanics decreased by less than half (HR 1.135, 95%CI 1.068 – 1.206; p<0.001) when compared to the univariate analysis (HR 1.231, 95%CI 1.161 – 1.304; p<0.001) [Table 14].

The other independent prognostic factors in this model were marital status (p<0.001), tumor site (p<0.001), clinical stage (p<0.001), and treatment performed (p<0.001). All these variables kept their trend of risk seen in the univariate analysis [Table 14].

Table 14 – Multivariate Cox regression survival analysis considering patients' and tumors' characteristics (Model 2)

Variables		**HR**	**(95% CI)**	**p-value**
Race/ethnicity	White	Ref.		
	Black	1.329	(1.267 – 1.393)	<0.001
	Hispanic	1.135	(1.068 – 1.206)	<0.001
	Asian/PI	1.099	(1.008 – 1.199)	0.033
	AI/AN	0.954	(0.761 – 1.195)	0.680
Marital status	Single	Ref.		
	Married	0.664	(0.637 – 0.693)	<0.001
	Divorced/Widowed	0.963	(0.920 – 1.008)	0.104
Tumor site	Oral cavity	Ref.		
	Oropharynx	0.797	(0.731 – 0.869)	<0.001
	Larynx	0.656	(0.609 – 0.707)	<0.001
	Hypopharynx	1.448	(1.239 – 1.692)	<0.001
Clinical Stage	Localized (CS I, II)	Ref.		
	Regional (CS III, IVa,b)	3.751	(3.528 – 3.988)	<0.001
	Distant (CS IVc)	9.106	(8.007 – 10.356)	<0.001
Treatment	Surgery alone	Ref.		
	Surgery + RT (+/– CH)	1.190	(1.112 – 1.272)	<0.001
	RT (+/– CH)	1.944	(1.818 – 2.079)	<0.001
	CH alone	2.960	(2.510 – 3.490)	<0.001
	Recommended*	6.316	(5.293 – 7.536)	<0.001
	Refused	3.726	(2.995 – 4.637)	<0.001
	Untreated/UNK	5.007	(4.553 – 5.507)	<0.001

Model adjusted for age, gender, SEER registry, year or diagnosis, and an interaction term (tumor site, clinical stage and treatment).

Model 3 was built as model 2 plus county-level attributes based on the Community Health Rankings and adjusted for age, gender, SEER registry, year of diagnosis, and an interaction term (tumor site, clinical stage, and treatment). In this model, one can notice that AI/AN (HR 0.895, 95%CI 0.714 – 1.122; p=0.334); Asian/PI (HR 1.045, 95%CI 0.958 – 1.140; p= 0.318); and Hispanics (HR 1.053, 95%CI 0.991 – 1.120; p=0.097) showed a similar prognosis when compared to Whites, showing the importance of SDOH in the gap observed for survival among these races/ethnicities regarding head and neck cancer. Unfortunately, this is not the case for Blacks; they still showed a 20% higher risk of dying from cancer than their White counterparts (HR 1.227, 95%CI 1.169 – 1.288; p<0.001). Even so, it is important to notice that this excess risk of death is approximately 4 times lower than the original one (from univariate analysis – HR 1.799) and also lower when compared to the model with tumor and patient characteristics (model 2 – HR 1.329), showing that SDOH are also important determinants of increased risk of death due to cancer in the Black population [Table 15].

Also in this model, we can observe the other independent prognostic factors such as marital status; again, the married people have a lower risk of death due to cancer than the single (p<0.001), but this is not the case when comparing divorced/widowed with singles (p=0.863). Another important piece of information from this model is the association of health insurance and prognosis. The risk of Medicaid and uninsured patients had decreased from the univariate analysis, but they were still a strong predictor of survival. In particular, Medicaid patients showed a higher risk (HR 1.464, 95% CI 1.404 – 1.526; p<0.001) than uninsured patients (HR 1.365, 95% CI 1.277 – 1.460; p<0.001) [Table 15].

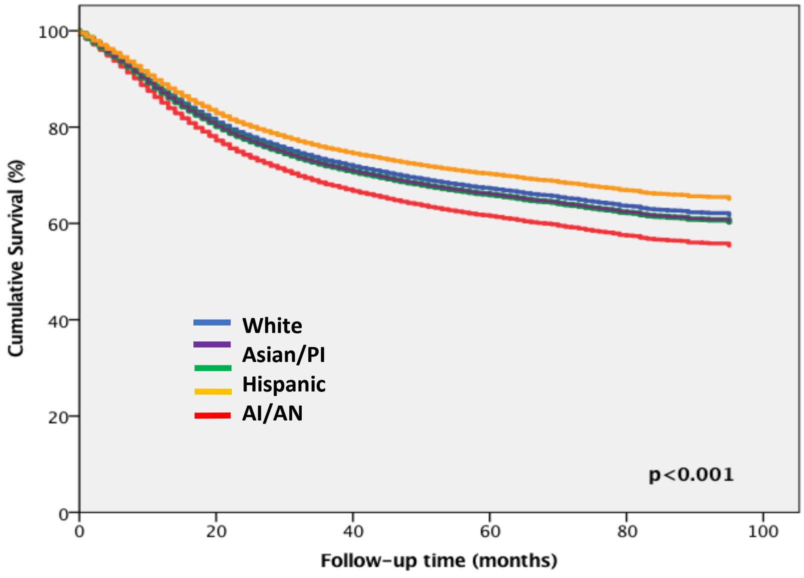
Considering the county-level measurements, only the Social Economic Factors score and the Clinical Care score are independently associated with prognosis. For the Social Economic Factors, counties with people living in the lower-middle (HR 1.121, 95%CI 1.066 – 1.179; p<0.001) and bottom 25% (HR 1.098, 95%CI 1.032 – 1.168; p=0.003) ranked counties presented a worse prognosis; patients living in the bottom 25% ranked counties considering the Clinical Care score also presented with a significantly increased risk of dying from cancer (HR 1.077, 95%CI 1.010 – 1.149; p=0.024) [Table 15]. A graph representing the survival curves according to race/ethnicity based on this model is presented in Figure 3.

Table 15 – Multivariate Cox regression analysis considering patients' and tumors' characteristics, and county-level attributes (Community Health Rankings) (Model 3)

Variables		**HR**	**(95% CI)**	**p-value**
Race/ethnicity	White	Ref.		
	Black	1.227	(1.169 – 1.288)	<0.001
	Hispanic	1.053	(0.991 – 1.120)	0.097
	Asian/PI	1.045	(0.958 – 1.140)	0.318
	AI/AN	0.895	(0.714 – 1.122)	0.334
Marital status	Single	Ref.		
	Married	0.726	(0.695 – 0.759)	<0.001
	Divorced/Widowed	0.996	(0.952 – 1.042)	0.863
Health insurance	Insured	Ref.		
	Medicaid	1.464	(1.404 – 1.526)	<0.001
	Uninsured	1.365	(1.277 – 1.460)	<0.001
Clinical Care score	Upper 25%	Ref.		
	Upper-middle	1.002	(0.956 – 1.050)	0.929
	Lower-middle	1.052	(0.995 – 1.111)	0.074
	Bottom 25%	1.077	(1.010 – 1.149)	0.024
Social & Economic Factors score	Upper 25%	Ref.		
	Upper-middle	1.036	(0.987 – 1.087)	0.153
	Lower-middle	1.121	(1.066 – 1.179)	0.000
	Bottom 25%	1.098	(1.032 – 1.168)	0.003

Model adjusted for age, gender, SEER registry, year of diagnosis and an interaction term (tumor site, clinical stage, and treatment).

Figure 3 – Survival curve according to race/ethnicity considering model 3.



IV. Discussion

In the analyzed population, the distribution of race/ethnicity among HNC patients is similar to the distribution of the US population over 45 years, for Blacks (11.2% in the dataset vs. 10.7% in the US Census), for AI/AN (0.5% vs. 0.6%, respectively), and for Asian/PI (4.2% vs. 4.8%, respectively). However, the distribution was higher for Whites (76.4% in the dataset vs. 72.3% in the US Census) and lower for Hispanics (7.7% vs. 10.7%, respectively).[55] This would not be expected considering that in a publication based on the Behavioral Risk Factor Surveillance System Survey for 2013, remarkable racial/ethnic disparities were observed in the prevalence of preventable risk factors for HNC, in which American Indians reported higher rates of binge drinking (19%) and Asian Americans reported lower (13%; p<0.001). American Indians also reported the highest rates of current smoking (28%), followed by Blacks (20%), Whites (18%), and Asians (10%; p<0.001). Moreover, the rates of HPV vaccination were lowest for American Indians (12%) compared to Whites (15%), Blacks (14%), and Asians (12%) even though the difference was not statistically significant.[56]

The decision to evaluate the association of SDOH with survival rates for HNC patients using the SEER database is based on the comprehensiveness, completeness, and representation of the US population in the database. The SEER Program is viewed as the standard for quality among cancer registries around the world. Quality improvement is an ongoing and evolving process of the SEER database to ensure validity and reliability, and each SEER Program registry has an obligation to meet specifically defined data quality goals.[48] When considering the external validity, the SEER Program currently collects and publishes cancer incidence and survival data from population-based cancer registries covering approximately 30 percent of the US population. The SEER coverage includes 26% of Blacks, 38% of Hispanics, 44% of AI/AN, 50% of Asians, and 67% of Hawaiian/Pacific Islanders.[48] The population covered by SEER is comparable to the general US population concerning measures of poverty and education. However, the SEER population tends to have a higher proportion of foreign-born persons than the general US population.[48] Despite the fact that the SEER database has strong quality control, and it is considered the gold standard for cancer registries, this is a firm of secondary data, and no new information can be added at the individual level. Thus, using a secondary data source is always a weakness of a study.

The use of county-level SDOH attributes to evaluate individual outcomes are somewhat controversial, but it is the best possible alternative when analyzing a population-based secondary data that does not provide individual-level details on such information. Combining individual- and county-level data has an intuitive appeal. The individual-level data permits identifiability of individual-level associations, and the county-level data may provide gains in the analysis, particularly if they represent large sample sizes.[57] Such associations are sometimes of interest, particularly for policy making, and can better explain a proposed model compared to limited individual-level data.[51,57]

In saying that, for all the socio-demographic individual characteristics evaluated (age, gender, marital status, area of residence, health insurance); clinical characteristics (tumor site, clinical stage, and treatment); county-level SDOH attributes (age distribution, poverty, unemployment, income, GINI index, education, health literacy, foreign-born and linguistic isolation); as well as for county-level scores (health outcome, health behavior, clinical care, social & economic factors, and physical environment) there was a statistical difference when comparing the distribution of these variables between the races/ethnicities, with the Whites presenting, overall, the best distribution according to the patient, tumor, socioeconomic and county-level attributes. Blacks more frequently presented the least favorable distribution. Asian/PI, Hispanics, and AI/AN have a mixed distribution depending on which of those variables are being evaluated. The association of the Black population with an advanced clinical stage at diagnosis, different treatment approaches (including treatment delay), lower employment rates, and lower household income has been reported by other authors when analyzing other datasets.[58] They are also more likely to be associated with increased rates of comorbidity, smoking, heavy alcohol use, and poorer performance status.[30] Also, Black patients are less likely to have insurance, more likely to reside in areas associated with lower education and income levels compared with White counterparts.[23]

So, it is important to take into consideration that higher SES tends to be positively associated with better health as SES is generally thought to influence health through three major pathways/mechanisms: 1. the ability to purchase health promoting resources and treatments (health access); 2. socialization of health habits differs by SES, but are shared among the SES strata; 3. it has been also posited that health can influence SES; such that, less healthy individuals complete fewer years of school, miss more work, and earn less, thus perpetuating the cycle.[12] This observation is also true for most SDOH.

The excess risk for death due to HNC among Blacks has been observed by other authors. In a previous study based on the SEER database, it could be observed that Blacks (compared to Whites) presented a worse prognosis for oral and oropharyngeal cancer in this study, and socioeconomic status was also associated with survival, with no mention on survival for the other races/ethnicity.[34] Another study, specifically for oropharynx cancer, could confirm that county-level SES, as well as Blacks presented with a poorer overall survival rate.[35] In a study using the LORHAN database, a detailed database of patient characteristics (including individual data on SDOH) and clinical outcomes assembled for HNC, the multivariate analysis adjusted for age, gender, smoking and alcohol consumption, education, income, health insurance, performance status, clinical stage, and treatment, showed a HR of 1.31 (95%CI 1.09 – 1.58) for Blacks when compared to Whites.[37] This trend of a worse prognosis among Blacks, even in multivariate models adjusted by SDOH, was observed in studies using other databases as well.[36,37] This was particularly true in studies comparing patients treated in a single institution.[38,39]

A few single institution studies have not observed racial/ethnic differences in HNC survival after adjusting for SDOH and clinical variables. Choi et al., observed in their study that for cancer specific survival univariate analysis that being Black, having a lower income, lower educational levels, advanced age, being unmarried, tobacco and alcohol consumption, lower BMI, advanced clinical stage, more severe comorbidity, and having chemotherapy were associated with worse prognosis. However, in the multivariate model, only education and age decreased cancer-specific survival rates, and race/ethnicity was no longer an independent prognosis factor (in a model adjusted for tumor site, clinical stage, patient's comorbidity, and treatment).[18] Ragin et al., in their study comparing survival rates between Blacks and Whites showed that in a multivariate analysis, SES, health insurance remained significantly associated with survival, but not race/ethnicity (in a model adjusted for patient's clinical characteristics).[31] Another study observed that only insurance status had a significant effect on survival in Black patients after controlling for other variables, suggesting that racial/ethnic differences in HNC outcomes are primarily related to differences in access to health care.[23]

Regarding other race/ethnicity and HNC survival rates, a study based on the British Columbia Cancer Agency population-based cancer control program could observe that Asians would present with a better prognosis than non-Asian in an analysis adjusted for (patient's characteristics and tumor's characteristics).[41] A study evaluating specifically the AI/AN population with HNC and comparing with Whites showed that they usually presented with more advanced stage disease in the oropharynx and were less likely to receive surgery alone for the treatment of oral cavity tumors. In a survival analysis stratified by a tumor site, they observed that specifically for oropharynx the AI/AN still presented with a worse prognosis in the adjusted model.[42] For Hispanics, a case-matching analysis showed that in the univariate analysis, Hispanics presented with a worse prognosis than Whites; however, in the case-matching analysis, the survival difference was no longer significant.[43]

The final model showed that county-level clinical care score and socioeconomic factors score were independently associated with prognosis in HNC patients; health insurance and marital status were also important prognosis factors. In this fully adjusted model, we observe the lowest difference in survival considering race/ethnicity, indicating that equity in clinical care (access to care and quality of care) as well as in SDOH could bring more equity in racial/ethnic differences observed for HNC survival rates.[11,18]

So, the remaining difference in survival presented only for Blacks, even in the fully adjusted model, is intriguing but had already been observed in the other study that included SDOH in the adjusted model (using a different study design, methodology, and approach than the current one).[59] The main possibility is that a much more complex scenario is in place and cannot be seen by the model built so far. This would include variables not measured or tested in this study (individual-level medical comorbidities, tobacco and alcohol consumption, HPV status, etc.), and the fact that the analyzed SDOH variables are related to the county level and not the individual level should also be taken into consideration. The other possibility discussed in the literature is the role of individuals' genetics and tumor biology.

Regarding tumor biology, it is well known that HPV-induced tumors have a different tumor biology than tobacco- and alcohol-induced tumors. Also, HPV is a risk factor for oropharynx tumors, but HPV positive tumors present a better prognosis than oropharynx tumors initiated due to tobacco and alcohol consumption.[4,5,31] The differential prevalence of HPV-induced tumors between races/ethnicities have been reported in many studies, as well as the influence of HPV on the prognosis.[1,4,5,9,10,31-33] Interestingly, the proportion of HPV-related HNC is substantially lower in Blacks when compared with Whites, and even lower among Asians.[9,31-33] However, this is related mainly to oropharynx tumors, and the analysis presented here is adjusted by tumor site.

There are few studies describing other tumor biology differences regarding race/ethnicity. In one study, it has been shown that tumors from Blacks were more likely to have loss of *CDKN2A* and gain of *SCYA3* when compared to tumors from Whites, and the multivariate modeling indicated significant differences between histopathology and tumor gene copy number alterations; both alterations can be linked with tumors presenting worse prognosis and were more frequently found among Blacks.[60] Another study evaluating epigenetic markers observed that tumors from Blacks were less likely to present with *DCC* promoter hypermethylation when compared to tumors from Whites.[61]

There are few and contradictory results in the literature trying to associate genetic ancestry and HNC survival disparities. Ramakodi et al., in an analysis evaluating ancestry-related alleles, showed results that could support the association of ancestry-related genetic factors with survival disparities in patients diagnosed with oral cavity and laryngeal cancer, mainly due to a worse treatment sensitivity and outcomes to platinum-based chemotherapy and/or radiotherapy for those with genetic alterations more commonly found in African-American descendants.[62] On the other hand, Worsham et al., in their study compared the clinical stage at presentation and survival analysis utilizing the race/ethnicity versus the genetic ancestry. Only self-reported race was associated with clinical stage at presentation; genetic ancestry was not. When a stratification within the self-denominated Black patients group was performed by West African genetic ancestry, there was no greater association with clinical stage or survival. The authors claimed that HNC disparities are more likely to be caused by SDOH rather than biological (genetic) factors.[63]

For the World Health Organization (WHO), health equity and SDOH are acknowledged as critical components of sustainable development in a global agenda and of the push toward the progressive achievement of universal health coverage (UHC). If health inequities are to be reduced, both SDOH and UHC need to be addressed in an integrated and systematic manner.[11] Furthermore, recommendations from the Institute of Medicine's Report to reduce disparities need to be implemented in treating HNC patients, as this might be the pathway to bring equity to the improvement in survival observed only by select HNC patients.[18]

V. Conclusion

In summary, our study showed that race/ethnicity plays an important role in HNC survival rates. Interestingly, the univariate analysis of all races/ethnicities showed a higher risk of dying due to cancer than Whites. When we first adjusted for patients' characteristics (age, gender, marital status), all races/ethnicities still showed a higher risk of dying due to cancer than Whites. The subsequent multivariate analysis, when tumors' characteristics (tumor site, clinical stage, and treatment) were added to the model, the risk of death from cancer was no longer statistically significant for AI/AN population, compared to the Whites, but still higher for Blacks, Hispanics, and Asian/PI when compared to Whites. The next model further adjusted for the SDOH variables, showed that Hispanics and Asian/PI no longer had a worse prognosis than Whites, but Blacks still showed a 20% higher risk of dying from cancer than Whites. Investments on HNC early detection and standard treatment performed timely could decrease the excess death risk by roughly half for all races/ethnicities (including Blacks) when compared to Whites; the investments and policies related to SDOH (including health care access) and towards decreasing social inequalities would further bring similar prognosis for Whites, AI/AN, Asian/PI, and Hispanics, and drop down the excess risk for Blacks by another 35%. If all this effort together could be enough to bring comparable survival rates to Blacks, based on the complexity related to their health access and social interactions, how much can be justified by the individual's genetics or tumor biology still warrants more investigation.

VI. References

1. Marur S, Forastiere AA. Head and neck cancer: changing epidemiology, diagnosis, and treatment. Mayo Clin Proc. 2008;83(4):489-501.
2. Carvalho AL, Nishimoto IN, Califano JA, Kowalski LP. Trends in incidence and prognosis for head and neck cancer in the United States: a site-specific analysis of the SEER database. Int J Cancer. 2005;114(5):806-16.
3. Olaleye O, Ekrikpo U, Lyne O, Wiseberg J. Incidence and survival trends of lip, intra-oral cavity and tongue base cancers in south-east England. Ann R Coll Surg Engl. 2015;97(3):229-34.
4. Hammarstedt L, Lu Y, Marklund L, Dalianis T, Munck-Wikland E, Ye W. Differential survival trends for patients with tonsillar, base of tongue and tongue cancer in Sweden. Oral Oncol. 2011;47(7):636-41.
5. Wang MB, Liu IY, Gornbein JA, Nguyen CT. HPV-Positive Oropharyngeal Carcinoma: A Systematic Review of Treatment and Prognosis. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery. 2015;153(5):758-69.
6. Hakulinen T, Tryggvadottir L, Gislum M, Storm HH, Bray F, Klint A, et al. Trends in the survival of patients diagnosed with cancers of the lip, oral cavity, and pharynx in the Nordic countries 1964-2003 followed up to the end of 2006. Acta Oncol. 2010;49(5):561-77.
7. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; . Available from: http://globocan.iarc.fr/, accessed on August/2017. 2013.
8. Bray F, Ren JS, Masuyer E, Ferlay J. Estimates of global cancer prevalence for 27 sites in the adult population in 2008. Int J Cancer. 2013;132(5):12.
9. Jiron J, Sethi S, Ali-Fehmi R, Franceschi S, Struijk L, van Doorn LJ, et al. Racial disparities in Human Papillomavirus (HPV) associated head and neck cancer. Am J Otolaryngol. 2014;35(2):147-53.
10. Chu KP, Habbous S, Kuang Q, Boyd K, Mirshams M, Liu FF, et al. Socioeconomic status, human papillomavirus, and overall survival in head and neck squamous cell carcinomas in Toronto, Canada. Cancer Epidemiol. 2016;40:102-12.
11. WHO. What are social determinants of health? 2016 [Available from: http://www.who.int/social_determinants/sdh_definition/en/.
12. Baker EH. Socioeconomic Status, Definition. The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society 2014;5.
13. Shin JY, Truong MT. Racial disparities in laryngeal cancer treatment and outcome: A population-based analysis of 24,069 patients. The Laryngoscope. 2015;125(7):1667-74.
14. McDonald JT, Johnson-Obaseki S, Hwang E, Connell C, Corsten M. The relationship between survival and socio-economic status for head and neck cancer in Canada. J Otolaryngol Head Neck Surg. 2014;43:2.
15. Chen AY, DeSantis C, Jemal A. US mortality rates for oral cavity and pharyngeal cancer by educational attainment. Arch Otolaryngol Head Neck Surg. 2011;137(11):1094-9.
16. Cosetti M, Yu GP, Schantz SP. Five-year survival rates and time trends of laryngeal cancer in the US population. Arch Otolaryngol Head Neck Surg. 2008;134(4):370-9.

17. Saini AT, Genden EM, Megwalu UC. Sociodemographic disparities in choice of therapy and survival in advanced laryngeal cancer. Am J Otolaryngol. 2016;37(2):65-9.
18. Choi SH, Terrell JE, Fowler KE, McLean SA, Ghanem T, Wolf GT, et al. Socioeconomic and Other Demographic Disparities Predicting Survival among Head and Neck Cancer Patients. PLoS One. 2016;11(3):e0149886.
19. Chen JJ, Stessin A, Christos P, Wernicke AG, Nori D, Parashar B. Differences in survival outcome between stage I and stage II glottic cancer: A SEER-based analysis. The Laryngoscope. 2015;125(9):2093-8.
20. Seoane J, Takkouche B, Varela-Centelles P, Tomas I, Seoane-Romero JM. Impact of delay in diagnosis on survival to head and neck carcinomas: a systematic review with meta-analysis. Clin Otolaryngol. 2012;37(2):99-106.
21. O'Neill CB, O'Neill JP, Atoria CL, Baxi SS, Henman MC, Ganly I, et al. Treatment complications and survival in advanced laryngeal cancer: a population-based analysis. The Laryngoscope. 2014;124(12):2707-13.
22. Newman JR, Connolly TM, Illing EA, Kilgore ML, Locher JL, Carroll WR. Survival trends in hypopharyngeal cancer: a population-based review. The Laryngoscope. 2015;125(3):624-9.
23. Gourin CG, Podolsky RH. Racial disparities in patients with head and neck squamous cell carcinoma. The Laryngoscope. 2006;116(7):1093-106.
24. Joseph LJ, Goodman M, Higgins K, Pilai R, Ramalingam SS, Magliocca K, et al. Racial disparities in squamous cell carcinoma of the oral tongue among women: a SEER data analysis. Oral Oncol. 2015;51(6):586-92.
25. Bonfante GM, Machado CJ, Souza PE, Andrade EI, Acurcio Fde A, Cherchiglia ML. [Specific 5-year oral cancer survival and associated factors in cancer outpatients in the Brazilian Unified National Health System]. Cad Saude Publica. 2014;30(5):983-97.
26. Braakhuis BJ, Leemans CR, Visser O. Incidence and survival trends of head and neck squamous cell carcinoma in the Netherlands between 1989 and 2011. Oral Oncol. 2014;50(7):670-5.
27. Znaor T, Vucemilo L, Kulis T, Znaor A. Incidence and mortality trends of head and neck cancer in Croatia in the period 1988-2008. Acta Otolaryngol. 2013;133(3):305-12.
28. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin. 2016;66(1):7-30.
29. Gimeno-Hernandez J, Iglesias-Moreno MC, Gomez-Serrano M, Carricondo F, Gil-Loyzaga P, Poch-Broto J. The impact of comorbidity on the survival of patients with laryngeal squamous cell carcinoma. Acta Otolaryngol. 2011;131(8):840-6.
30. Zakeri K, MacEwan I, Vazirnia A, Cohen EE, Spiotto MT, Haraf DJ, et al. Race and competing mortality in advanced head and neck cancer. Oral Oncol. 2014;50(1):40-4.
31. Ragin CC, Langevin SM, Marzouk M, Grandis J, Taioli E. Determinants of head and neck cancer survival by race. Head Neck. 2011;33(8):1092-8.
32. Brown LM, Check DP, Devesa SS. Oropharyngeal cancer incidence trends: diminishing racial disparities. Cancer Causes Control. 2011;22(5):753-63.
33. Ragin C, Liu JC, Jones G, Shoyele O, Sowunmi B, Kennett R, et al. Prevalence of HPV Infection in Racial-Ethnic Subgroups of Head and Neck Cancer Patients. Carcinogenesis. 2016.
34. Osazuwa-Peters N, Massa ST, Christopher KM, Walker RJ, Varvares MA. Race and sex disparities in long-term survival of oral and oropharyngeal cancer in the United States. J Cancer Res Clin Oncol. 2016;142(2):521-8.

35. Megwalu UC. Impact of County-Level Socioeconomic Status on Oropharyngeal Cancer Survival in the United States. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery. 2017;156(4):665-70.
36. Peterson CE, Khosla S, Chen LF, Joslin CE, Davis FG, Fitzgibbon ML, et al. Racial differences in head and neck squamous cell carcinomas among non-Hispanic black and white males identified through the National Cancer Database (1998-2012). J Cancer Res Clin Oncol. 2016;142(8):1715-26.
37. Hayes DN, Peng G, Pennella E, Hossain A, Carter GC, Muehlenbein C, et al. An exploratory subgroup analysis of race and gender in squamous cancer of the head and neck: inferior outcomes for African American males in the LORHAN database. Oral Oncol. 2014;50(6):605-10.
38. Murdock JM, Gluckman JL. African-American and white head and neck carcinoma patients in a university medical center setting. Are treatments provided and are outcomes similar or disparate? Cancer. 2001;91(1 Suppl):279-83.
39. Moore RJ, Doherty DA, Do KA, Chamberlain RM, Khuri FR. Racial disparity in survival of patients with squamous cell carcinoma of the oral cavity and pharynx. Ethnicity & health. 2001;6(3-4):165-77.
40. Chu KP, Shema S, Wu S, Gomez SL, Chang ET, Le QT. Head and neck cancer-specific survival based on socioeconomic status in Asians and Pacific Islanders. Cancer. 2011;117(9):1935-45.
41. Kim JD, Chang JT, Moghaddamjou A, Kornelsen EA, Ruan JY, Olson RA, et al. Asian and non-Asian disparities in outcomes of non-nasopharyngeal head and neck cancer. The Laryngoscope. 2017.
42. Dwojak SM, Sequist TD, Emerick K, Deschler DG. Survival differences among American Indians/Alaska Natives with head and neck squamous cell carcinoma. Head Neck. 2013;35(8):1114-8.
43. Schrank TP, Han Y, Weiss H, Resto VA. Case-matching analysis of head and neck squamous cell carcinoma in racial and ethnic minorities in the United States--possible role for human papillomavirus in survival disparities. Head Neck. 2011;33(1):45-53.
44. DeSantis CE, Siegel RL, Sauer AG, Miller KD, Fedewa SA, Alcaraz KI, et al. Cancer statistics for African Americans, 2016: Progress and opportunities in reducing racial disparities. CA Cancer J Clin. 2016.
45. Goodwin WJ, Thomas GR, Parker DF, Joseph D, Levis S, Franzmann E, et al. Unequal burden of head and neck cancer in the United States. Head Neck. 2008;30(3):358-71.
46. Shiboski CH, Schmidt BL, Jordan RC. Racial disparity in stage at diagnosis and survival among adults with oral cancer in the US. Community dentistry and oral epidemiology. 2007;35(3):233-40.
47. Jean RA, Kallogjeri D, Strope SA, Hardin FM, Rich JT, Piccirillo JF. Exploring SEER-Medicare for changes in the treatment of laryngeal cancer among elderly medicare beneficiaries. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery. 2014;150(3):419-27.
48. SEER. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) Research Data (1973-2013), National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch, released April 2016, based on the November 2015 submission.
49. National Center for Education Statistics. National Assessment of Adult Literacy. State & County Estimates of Low Literacy. https://ncesedgov/naal/estimates/StateEstimatesaspx. 2003.
50. Adesanya I. Income inequality: by the numbers. http://mmjvcuedu/2013/12/20/income-inequality-data/. 2013;Accessed in 2017.
51. University of Wisconsin Population Health Institute. County Health Rankings & Roadmaps 2017. http://wwwcountyhealthrankingsorg/. 2017.

52. National Cancer Institute (NCI). Surveillance, Epidemiology, and End Results (SEER) Program. County Attributes. https://seercancergov/seerstat/variables/countyattribs/ - 10.

53. Kutner M, Greenberg E, Jin Y, C P. The health literacy of America's adults: Results from the 2003 National Assessment of Adult Literacy. https://ncesedgov/pubs2006/2006483pdf. 2006.

54. University of Wisconsin Population Health Institute. County Health Rankings & Roadmaps 2017. Calculating Scores and Ranks. http://wwwcountyhealthrankingsorg/ranking-methods/calculating-scores-and-ranks. 2017.

55. Population Reference Bureau. http://wwwprborg/DataFinder/Topic/Rankingsaspx?ind=284&loc=312. 2017.

56. Dwojak S, Bhattacharyya N. Racial disparities in preventable risk factors for head and neck cancer. The Laryngoscope. 2017;127(5):1068-72.

57. Haneuse S, Bartell S. Designs for the combination of group- and individual-level data. Epidemiology (Cambridge, Mass). 2011;22(3):382-9.

58. Naghavi AO, Echevarria MI, Strom TJ, Abuodeh YA, Ahmed KA, Venkat PS, et al. Treatment delays, race, and outcomes in head and neck cancer. Cancer Epidemiol. 2016;45:18-25.

59. Molina MA, Cheung MC, Perez EA, Byrne MM, Franceschi D, Moffat FL, et al. African American and poor patients have a dramatically worse prognosis for head and neck cancer: an examination of 20,915 patients. Cancer. 2008;113(10):2797-806.

60. Worsham MJ, Stephen JK, Lu M, Chen KM, Havard S, Shah V, et al. Disparate molecular, histopathology, and clinical factors in head and neck squamous cell carcinoma racial groups. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery. 2012;147(2):281-8.

61. Carvalho AL, Chuang A, Jiang WW, Lee J, Begum S, Poeta L, et al. Deleted in colorectal cancer is a putative conditional tumor-suppressor gene inactivated by promoter hypermethylation in head and neck squamous cell carcinoma. Cancer research. 2006;66(19):9401-7.

62. Ramakodi MP, Devarajan K, Blackman E, Gibbs D, Luce D, Deloumeaux J, et al. Integrative genomic analysis identifies ancestry-related expression quantitative trait loci on DNA polymerase beta and supports the association of genetic ancestry with survival disparities in head and neck squamous cell carcinoma. Cancer. 2017;123(5):849-60.

63. Worsham MJ, Divine G, Kittles RA. Race as a Social Construct in Head and Neck Cancer Outcomes. Otolaryngology–Head and Neck Surgery. 2010;144(3):9.

Appendix I – County-level attribute variables measurement, source and year(s)

County-level attribute variables			
Attributes	**Measure**	**Source**	**Year(s)**
Educational attainment	The percentage of the population 25 years and over with at least a bachelor's degree	American Community Survey (ACS) Census, 5-year estimates	2009-2013
Age distribution	The percent of person ages 65 and over in the county	American Community Survey (ACS) Census, 5-year estimates	2009-2013
Employment rate	The percent of person ages 16 and over who are unemployed	American Community Survey (ACS) Census, 5-year estimates	2009-2013
Foreign-born rate	The percent of persons who are foreign-born: place of birth by nativity and citizenship status	American Community Survey (ACS) Census, 5-year estimates	2009-2013
Linguistic isolation rate	The percent of households that are linguistically isolated: household language by household limited English speaking status	American Community Survey (ACS) Census, 5-year estimates	2009-2013
Median income	Median household income and median family income in the past 12 months	American Community Survey (ACS) Census, 5-year estimates	2009-2013
Poverty rate	The percent of families whose incomes are below the poverty level	American Community Survey (ACS) Census, 5-year estimates	2009-2013
	The percent of persons who are below 150 percent of the poverty level is calculated using the ratio of income to the poverty level in the past 12 months.	American Community Survey (ACS) Census, 5-year estimates	2009-2013
Health literacy rate[49]	Predictive model to estimate the percentage of people who are likely to have below basic or basic skills according to the NAAL criteria (health literacy estimate of 225 or lower)	National Assessment of Adult Literacy (NAAL)*	2003
GINI index[50]	A standard economic measure of income inequality, based on Lorenz Curve. Gini scale ranges from 0 to 1. Higher the index, higher the inequality.	American Community Survey (ACS) Census, 5-year estimates	2008-2012

* for health literacy regarding AI/AN, a specific publication based on the 2003 NAAL was used[53].

For more information on this county-level attribute variables refer to:
https://seer.cancer.gov/seerstat/variables/countyattribs/#10
https://nces.ed.gov/naal/estimates/overview.aspx
http://mmj.vcu.edu/2013/12/20/income-inequality-data/

Appendix II – County-level attribute scores measurement, source and year(s) for the County Health Rankings[51]

Health Outcomes			
Focus Area	**Measure**	**Source**	**Year(s)**
Mortality	Premature death (years of potential life lost before age 75 per 100,000 pop)	National Center for Health Statistics	2008-2010
Morbidity	Poor or fair health (percent of adults reporting fair or poor health)	Behavioral Risk Factor Surveillance System	2005-2011
	Poor physical health days (average number in past 30 days)	Behavioral Risk Factor Surveillance System	2005-2011
	Poor mental health days (average number in past 30 days)	Behavioral Risk Factor Surveillance System	2005-2011
	Low birthweight (percent of live births with weight < 2500 grams)	National Center for Health Statistics	2004-2010

Health Behaviors Score			
Focus Area	**Measure**	**Source**	**Year(s)**
Tobacco use	Adult smoking (percent of adults that smoke)	Behavioral Risk Factor Surveillance System	2005-2011
Diet and exercise	Adult obesity (percent of adults that report a BMI >= 30)	National Center for Chronic Disease Prevention and Health Promotion, Division of Diabetes Translation	2009
	Physical inactivity (percent of adults that report no leisure time physical activity)	National Center for Chronic Disease Prevention and Health Promotion, Division of Diabetes Translation	2009
Alcohol use	Excessive drinking (percent of adults who report heavy or binge drinking)	Behavioral Risk Factor Surveillance System	2005-2011
	Motor vehicle crash deaths per 100,000 population	National Center for Health Statistics	2004-2010
Sexual activity	Sexually transmitted infections (chlamydia rate per 100,000 population)	National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention	2010
	Teen birth rate (per 1,000 females ages 15-19)	National Center for Health Statistics	2004-2010

Clinical Care Score			
Focus Area	**Measure**	**Source**	**Year(s)**
Access to care	Uninsured (percent of population < age 65 without health insurance)	Small Area Health Insurance Estimates	2010
	Ratio of population to primary care physicians	HRSA Area Resource File	2011-2012
	Ratio of population to dentists	HRSA Area Resource File	2011-2012
Quality of care	Preventable hospital stays (rate per 1,000 Medicare enrollees)	Dartmouth Atlas of Health Care	2010
	Diabetic screening (percent of diabetics that receive HbA1c screening)	Dartmouth Atlas of Health Care	2010
	Mammography screening (percent of females that receive screening)	Dartmouth Atlas of Health Care	2010

Social and Economic Environment Score			
Focus Area	**Measure**	**Source**	**Year(s)**
Education	High school graduation (percent of ninth grade cohort that graduates in 4 years)	State sources and the National Center for Education Statistics	Varies by state
	Some college (Percent of adults aged 25-44 years with some post-secondary education)	American Community Survey (ACS) Census, 5-year estimates	2007-2011
Employment	Unemployment rate (percent of population age 16+ unemployed)	Bureau of Labor Statistics	2011
Income	Children in poverty (percent of children under age 18 in poverty)	Small Area Income and Poverty Estimates	2011
Family and social support	Inadequate social support (percent of adults without social/emotional support)	Behavioral Risk Factor Surveillance System	2005-2010
	Percent of children that live in single-parent household	American Community Survey, 5-year estimates	2007-2011
Community safety	Violent crime rate per 100,000 population	Uniform Crime Reporting	2008-2010

Physical Environment Score			
Focus Area	**Measure**	**Source**	**Year(s)**
Environmental quality	Daily fine particulate matter (average daily measured in micrograms per cubic meter)	CDC WONDER Environmental data *Data not available for Alaska and Hawaii*	2008
	Drinking water safety (percent of population exposed to water exceeding a violation limit in the past year)	Safe Drinking Water Information System	2012
Built environment	Access to recreational facilities (rate per 100,000 population)	County Business Patterns	2010
	Limited access to health foods (percent of population who lives in poverty and more than 1 or 10 miles from a grocery store)	USDA Food Environment Atlas	2012
	Fast food restaurants (percent of all restaurants that are fast food)	County Business Patterns	2010

For more information refer to:
http://www.countyhealthrankings.org/our-approach

Appendix III – Formulas to calculate the Z-score for the County Health Rankings[54]

Standardizing each of these measures transforms them to the same metric—a mean (average) value of 0 and a standard deviation (measure of spread) of 1. We refer to these as Z-scores where:

$$Z = \frac{\text{(County Value)} - \text{(Average of Counties in State)}}{\text{(Standard Deviation of Counties in State)}}$$

For counties with a population of 20,000 or less, any z-score that is < -3.0 or > 3.0 is truncated to -3.0 or 3.0, respectively.

A weighted composite is computed by multiplying each Z-score by its weight and adding them up.

$$Composite = \sum w_i \; Z_i$$

In this formula, the Zi values are the Z-scores of the measures used to compute the summary score. The wi values are the weights applied to each Z-score. The ∑ sign simply means to add up all the Z-scores multiplied by their weights. All of the summary scores we compute use the formula above, standardized Z-scores for each measure (reverse coded when necessary), and the weights described in previous sections. It is important to notice that composite scores are computed separately by state.

---

* Some counties in the nation are too small to have reliable measurements for health outcome measures. Those counties are not ranked. If a county has data for enough measures to be ranked but is missing data for any individual measure, we assign the county the same value as the state mean for that measure.