Predictive Scoring Analysis for Identifying Head and Neck Squamous Cell Cancer

Christopher Lloyd[1†], My Nguyen[1†], and Christian Kauffman[2]

[1]Geisinger College of Health Sciences, Scranton, PA 18509
[2]Geisinger Medical Center, Plastic Surgery Department, Danville, PA 17822
[†]Doctor of Medicine Program
Correspondence: clloyd@som.geisinger.edu

Abstract

Background: Head and neck squamous cell carcinoma (HNSCC) accounts for an estimated 450,000 deaths annually worldwide, presenting a significant clinical burden (1). Poor survival rates are largely due to the lack of reliable screening tools for early detection, as current methods offer limited insight into tumor invasiveness and metastasis potentials. While previous studies have explored HNSCC risk factors, predicting disease's likelihood remains challenging due to lack of reliable screening tools. This retrospective study aims to creating a risk stratification scoring system for HNSCC, ultimately improves early detection and management for high-risk patients.

Methods: We developed a risk scoring system using data from the health system Geisinger, incorporating demographic and comorbidity factors such as sex, age, race, BMI, HPV history, smoking, alcohol use, diabetes, and prior diagnoses of leukoplakia or erythroplakia. The study included patients age 18 and older diagnosed with HNSCC between 2000 and 2024, excluding those with HIV/AIDS. Categorical variables were summarized with frequencies and percentages, while medians were reported for continuous variables. We used multiple logistic regression and descriptive analysis to assess the association between each risk factor and HNSCC development.

Results: Of the total 2,435 patients who were diagnosed with HNSCC, 1,607 (66.0%) are patients age 60 years or older at diagnosis. There are 620 females (25.5%) and 1,815 males (74.5%). Mean BMI is 27.5 (SD 7.0) with 707 patients (29.0%) having obesity. One thousand four hundred eighteen patients (58.2%) use tobacco, 949 patients (39.0%) use alcohol, 605 patients (24.8%) have diabetes, 62 patients (2.5%) have history of leukoplakia or erythroplakia, 14 patients (0.6%) have history of HPV, and 59 patients (2.4%) were previously diagnosed with HNSCC. Among 778 patients with known pack years, 109 (14.0%) smoked less than 10 pack years, 669 (86%) smoked at least 10 pack years, and 536 smoked at least 20 pack years (68.9%). Among 227 patients with known alcohol use frequency, 123 (54.2%) consumed less than 10 drinks per week and 104 (45.8%) consumed at least 10 drinks per week. Type/location of tumor: 621 (25.5%) tongue cancer, 358 (14.7%) tonsillar cancer, 46 (1.9%) nasopharyngeal cancer, 72 (3.0%) oropharyngeal cancer, 54 (2.2%) palatine cancer, 202 (8.3%) lip cancer, 60 (2.5%) gum cancer, 114 (4.7%) floor of mouth cancer, 270 (11.1%) glottic cancer, 259 (10.6%) supraglottic cancer, 54 (2.2%) pyriform sinus cancer, and 325 (13.3%) others. Patients who scored 6 or higher are classified as high risk, 4–6 is medium risk, and 1–3 is low risk of HNSCC.

Conclusion: Tobacco use is identified as the highest risk factor that contributes to developing HNSCC. Patients scoring 6 or more points should undergo more rigorous and proactive screening for early detection. Given the severity of HNSCC, implementing this scoring system is crucial for reducing undiagnosed cases and improving patient outcomes. Future study should focus on which factors are associated with earlier onset of HNSCC and prediction of survival rate based on risk scores.

Introduction

Preventative medicine is becoming the new gold standard of health care, especially in regard to cancer surveillance. Head and neck squamous cell cancer (HNSCC) is a commonly occurring cancer worldwide and can have poor outcomes when not detected and treated early. In order to more efficiently and quickly find at risk patients a predictive scoring system was developed, based on risk factors, social factors, and dietary conditions. This study aims to elucidate those populations at most risk for developing HNSCC.

The yearly incidence rate of head and neck squamous cell carcinoma (HNSCC) stands at an estimated 890,000 new cases with 450,000 deaths occurring annually worldwide, proving to be a burdensome disease (1). 5- and 10-year survival rates range from 50% to 70% for HNSCC depending on location (ex: higher survival rates for p16+ oropharynx 69% vs hypopharynx 51%) and tumor lineage. These troubling survival rates can be attributed to the lack of accurate and precise screening tools available to clinicians to detect earlier stages of tumor growth (2). Furthermore, quality of life deficits present for survivors of HNSCC are considerably more impactful than most other cancers (survivors of other cancers are 23.6 cases per 100,000 individuals) shown to be the second highest rate of suicide (63.4 cases per 100,000 individuals), only second to pancreatic cancer survivors, (86.4 cases per 100,000 individuals) (3). Visibility and physical exam unfortunately only relay limited information on the tumor's proclivity to invade and metastasize. More information is needed to inform better clinical management for the best prognosis possible.

This retrospective study aims to achieve a more comprehensive pathological and clinical understanding of head and neck squamous cell carcinoma (HNSCC), in order to better serve our population of potentially high-risk patients. By utilizing data from the Geisinger health system, we aimed to develop a scored system based off statewide patient histories (family history, HPV status, diabetes, obesity), demographic (age, sex, ethnicity), social factors (tobacco use, alcohol), and physical exam findings (dysplasia) that will allow a non-biased assessment that can be broadly applied to a population to most quickly and efficiently detect at-risk patients. Furthermore, this study aims to elucidate associations between risk factors such as diabetes and obesity, which have known pathophysiological mechanisms disrupted upregulating oxidative damage, yet little directed research investigating their role in the HNSCC pathogenesis.

Time is critical in oncological cases, particularly when in great proximity to adequate vasculature such as found within the head and neck areas, so it is imperative that an efficient and cost-effective tool be established to ensure that fewer patient cases go unrecognized before their outcome has suffered.

Depending on the stage of cellular degradation and particular hits made to an individual cell line (i.e., activations/inactivation of oncogenes/proto-oncogenes), particular proteins are expected to significantly increase in baseline metabolic production. This may even prove to be a useful method in determining significance of oral dysplasia found by providing secondary lab values to support physical exam findings.

This study was crafted to better elucidate the needs of screening for HNSCC, particularly for the patients within Pennsylvania. Many shared social and nutritional aspects overlap for Pennsylvania residents that leave them at high risk of developing HNSCC, particularly rural individuals lacking nearby access to affordable and timely health care screenings. Additionally, even with routine general practitioner exams, thorough inspections of the oral cavity are not routine, thus significantly increasing the risk of development of nascent leukoplakia or metaplastic growths into invasive carcinomas. Furthermore, HNSCC has proven to be a treatable condition if found early enough into disease progression, and with poor 5- and 10-year survival rates, implementation of proper screening tools, practices, and patient education should serve to greatly eliminate the burden of care for late stage HNSCC diagnoses. Below are some common risk factors of HNSCC from previous literature.

HPV

In the United States, where the greatest number of studies have been performed, approximately 70% of all oropharyngeal cancers are attributable to HPV, a much higher proportion than seen worldwide. The overall incidence of HPV-related oropharyngeal cancers was estimated to be 4.8 in 100,000 in 2013–2014, compared with the incidence of HPV-related non-oropharyngeal head and neck cancers of 0.62 per 100,000.

In the United States, the incidence of HPV-related oropharyngeal cancer differs not only by sex but also by age and race. Through 2015 in the United States, the incidence in White men of all ages increased more than any other subgroup, to greater than 18 per 100,000 people. In contrast, the incidence in Hispanic and Black men was 6 and 4 in 100,000, respectively. In White women, the incidence in 2015 was almost 4 per 100,000. In contrast, the incidence in Hispanic and Black women was approximately 2 per 100,000. In 2020, the FDA, using the evidence presented here as a surrogate to predict clinical benefit in preventing HPV associated oropharyngeal cancer, approved the use of the HPV vaccine specifically for the prevention of HPV-related head and neck cancer (4).

Tobacco

Tobacco causes most HNC, and alcohol synergistically increases the risk of HNC conferred by tobacco use. Historically, approximately 90% of patients with HNC have a history of tobacco use, with a 4-fold to 5-fold increase in risk of developing oral cavity, oropharynx, and hypopharynx cancers, and a 10-fold increased risk of laryngeal cancer, among tobacco users. The carcinogenic effects of tobacco are dose-dependent, with the risk of HNC closely related to the frequency, duration, and intensity of cigarette smoking (5).

Alcohol

Alcohol use independently increases the risk of HNC, with an estimated 1% to 4% of cases attributable to alcohol alone and a twofold increase in odds of HNC for drinkers who are never users of tobacco. In particular, alcohol use increases risk of hypopharyngeal cancer when compared with other sites.

However, the larger impact of alcohol is observed in the interaction of alcohol with tobacco use. The combined effects of alcohol and tobacco use have a greater than multiplicative impact in increasing the risk of cancer. Among individuals who smoke 2 or more packs of cigarettes and drink more than 4 alcoholic drinks per day, the risk of HNSCC is increased greater than 35-fold (5, 6).

Male Gender

HNC overall is approximately twofold to threefold more common among men than women in the United States (5).

Methods

Using data across the Geisinger health system, we developed a risk scoring system based on demographic and comorbidity factors, including sex, age, race, BMI, history of HPV, smoking, alcohol use, diabetes, and prior diagnoses of leukoplakia or erythroplakia. Risk factors were selected through existing literature, both individual studies as well as meta-analysis that looked at head and neck squamous cell development. Of note, certain risk factors were excluded, such as the use of betel chew, due to the small sample sizes and lack of documentation likely to be made on patient charts and clinical histories.

Our study included patients age 18 and older diagnosed with HNSCC diagnosis between 2000 and 2024. Those with a history of HIV/AIDS were excluded.

To describe categorical characteristics, the frequency and percentages were reported. Additional limitations on the study include lack of definitive quantities of alcohol and tobacco consumption. Patients were grouped into either low-medium alcohol/tobacco use or high alcohol/tobacco use based on self-reported data during clinical encounters. Furthermore, patients with diabetes mellitus were consolidated into one group, with no differentiation made based on length or severity of diagnosis.

Median and interquartile range (IQR) were reported for continuous characteristics. Multiple logistic regression test and descriptive analysis were conducted to examine the association between each risk factor and likelihood of developing HNSCC. Following analysis of the risk factors and their associations, scores for each risk factor were devised and assigned to correlate with the likelihood of HNSCC development following exposure.

Results

Of the total 2,435 patients who were diagnosed with HNSCC, 1,607 (66.0%) patients age 60 years or older at diagnosis, and 224 (9.2%) patients age < 50 at diagnosis (Table 1). Mean age at diagnosis is 64.8 (SD 12.1). There are 620 females (25.5%) and 1815 males (74.5%). Within the female cohort, 341 (55.0%) patients use tobacco, 201 (32.4%) patients use alcohol, and 177 (28.5%) patients have diabetes (Table 2). Within the male cohort, 1,077 (59.3%) patients use tobacco, 748 (41.2%) patients use alcohol, and 428 (23.6%) patients have diabetes.

Forty-two patients (1.7%) are Black or African American. Within this cohort, 31 (73.8%) patients are male, 12 (28.6%) patients were at least 65 years old at diagnosis, 25 (59.5%) patients were at least 60 years old at diagnosis. Average age at diagnosis is 62.0. Fourteen (33.3%) patients are obese, 25 (59.5%) patients have use tobacco, 16 (38.1%) patients have alcohol use, 11 (26.2%) patients have diabetes, 1 (2.4%) patient had leukoplakia, 1 (2.4%) patient had HPV.

Mean BMI is 27.5 (SD 7.0). 707 patients (29.0%) had obesity with BMI 30 or above. Sixty-two patients (2.5%) have history of leukoplakia or erythroplakia, 14 patients (0.6%) have history of HPV, 1418 patients (58.2%) have tobacco use, 949 patients (39.0%) have alcohol use disorder, 605 patients (24.8%) have diabetes, and 59 patients (2.4%) reported having a history of HNSCC. The highest proportion of patients

	Total (N = 2435)
Age at time of HNSCC diagnosis	
Mean	64.8 (SD 12.1)
Race, n (%)	
White	2353 (96.6%)
Black or African American	42 (1.7%)
Asian	19 (0.8%)
American Indian or Alaska Native	4 (0.2%)
Native Hawaiian or other Pacific Islander	1 (0.04%)
Unknown	16 (0.7%)
Gender, n (%)	
Male	1815 (74.5%)
Female	620 (25.5%)
Ethnicity, n (%)	
Hispanic or Latino	35 (1.4%)
Not Hispanic or Latino	2360 (97.0%)
Unknown	40 (1.6%)

Table 1. Demographics of Study Population

Age, n (%)	
≥ 60	1607 (66.0%)
< 50	224 (9.2%)
BMI, n (%)	
≥ 30 (Obesity)	707 (29.0%)
< 30	1728 (71.0%)
Gender, n (%)	
Male	1815 (74.5%)
Female	620 (25.5%)
Tobacco use, n (%)	
Yes	1418 (58.2%)
≥ 20 pack years	536
≥ 10 pack years	669
Unknown pack years	640
No	1017 (41.8%)
Alcohol use, n (%)	
Yes	949 (38.0%)
≥ 10 drinks per week	123
< 10 drinks per week	104
Unknown frequency	722
No	1486 (61.0%)
Diabetes, n (%)	
Yes	605 (24.8%)
No	1830 (75.2%)
Previous leukoplakia or erythroplakia, n (%)	
Yes	62 (2.5%)
No	2373 (97.5%)
Previous HNSCC, n (%)	
Yes	59 (2.4%)
No	2376 (97.6%)
Previous HPV, n (%)	
Yes	14 (0.6%)
No	2421 (99.4%)

Table 2. Risk Factors of HNSCC

(58.2%) who have a history of tobacco use developed HNSCC later in life. Surprisingly, only a small proportion of patients with a history of HPV (0.6%) developed HNSCC.

For patients with previous HNSCC, the odds ratio of recurrent HNSCC with tobacco use is 0.96 (95% CI 0.53 - 1.84). 1017 patients (41.8%) did not smoke, and 1,486 patients (61.0%) did not use alcohol but still developed HNSCC. Among 778 patients with known pack years, 109 (14.0%) smoked less than 10 pack years, 669 (86%) smoked at least 10 pack years, and 536 smoked at least 20 pack years (68.9%).

Among 227 patients with known alcohol use frequency, 123 (54.2%) consumed less than 10 drinks per week and 104 (45.8%) consumed at least 10 drinks per week.

Among 14 patients with HPV, 13 of them are female and only 1 male. Type/location of tumor in our study include 621 (25.5%) tongue cancer, 358 (14.7%) tonsillar cancer, 46 (1.9%) nasopharyngeal cancer, 72 (3.0%) oropharyngeal cancer, 54 (2.2%) palatine cancer, 202 (8.3%) lip cancer, 60 (2.5%) gum cancer, 114 (4.7%) floor of mouth cancer, 270 (11.1%) glottic cancer, 259 (10.6%) supraglottic cancer, 54 (2.2%) pyriform sinus cancer, and 325 (13.3%) others.

Multiple logistic regression and descriptive analysis were used to develop the risk stratification scoring system, as shown in Table 3. The predictive scoring system for head and neck squamous cell carcinoma (HNSCC) is designed to estimate the probability of an individual developing the condition based on key risk factors. This 9-point scoring system uses a logistic regression model that correlates factors like frequency, odds ratios, and clinical outcomes to predict the likelihood of the disease. Points are awarded based on the following risk factors: male sex (1 point), tobacco use (1 point for less than 10 pack years, 2 points for more than 10 pack years), alcohol use (2 points), diabetes (1 point for a history of diabetes), and previous leukoplakia, erythroplakia, or HNSCC (1 point). For females, a history of HPV

Risk Factors	Score
Male	1
Tobacco use	1 < 10 pack years 2 > 10 pack years
Alcohol use	2
Diabetes	1
Previous leukoplakia or erythroplakia	1
Previous HNSCC	1
Previous HPV (only include if female)	1
Age > 60	1

Table 3. Risk Scoring System for HNSCC

infection also contributes 1 point. Age is another critical factor, with 1 point awarded for being older than 60 years, as 66% of cases occur in this age group. The system divides patients into 3 risk categories: low risk (1–3 points), medium risk (4–6 points), and high risk (7–9 points). Smoking and alcohol consumption are highly correlated with HNSCC development, with 58.2% of diagnosed individuals having a history of smoking less than 10 pack years (earning 1 point), and 86% of diagnosed individuals having a history of smoking more than 10 pack years (earning 2 points). Although previous erythroplakia and HNSCC are highly correlated, only 2.4% of cases represent rediagnosis, with 2.5% of cases being erythroplakia-related. The scoring system aims to help identify high-risk patients, allowing for tailored prevention and treatment strategies.

A logistic regression analysis of age-related risk factors shows that age has a relatively low correlation with the outcome. With a multiple R of 0.0631, the R-squared value is only 0.00399, indicating that age alone explains only a small portion of the variance in the data. The adjusted R-squared is even lower, at 0.00357, further suggesting that age is a minimal predictor in this model. The regression statistics show that the relationship between age and the outcome is statistically significant with an F-value of 9.73 and a p-value of 0.0018.

Discussion

In this study, we identified and analyzed 10 risk factors most likely to be associated with HNSCC development. Occurrences, frequency of associated activities, and comorbidities were used to develop a novel risk stratification scoring system. The use of scoring systems has been widely used to varying degrees of success (7) but is typically field dependent in terms of its capacity to integrate seamlessly into routine clinical encounters and documentation. Of note, prediction scores have several limitations including inaccurate observations, omitted patient health information, imperfect laboratory testing with the infeasibility of 100% specific and sensitive testing, and varying degrees of patient associated comorbidities. Furthermore, after development of predictive scoring systems, there are potential ethical and economic concerns regarding administering screens or subjecting patients to potentially unnecessary exams and associated fees. Yet it is the understanding of the researchers on this paper that this should not be seen as a limiting factor in the implementation of risk stratification, as preventive medicine has proven to be effective health care. Particularly, preventative medicine and screening tools make ideal domains for integrating risk stratification scoring systems, as typically disease and economic burden is much lower when health concerns are identified earlier in disease progression.

The most significant associated risk factors were tobacco use and alcohol use, 1,418 patients (58.2%) and 949 patients (39.0%), respectively. Our findings reinforce the well-established role of tobacco and alcohol use as primary risk factors for HNSCC, consistent with (8) research, which identified cigarette smoking and alcohol consumption as the strongest predictors of head and neck cancer.

Furthermore, our study results compared appropriately against national averages and suggested likelihood ratios of development seen by other studies (9) Interestingly, when comparing light against heavy tobacco usage, among 778 patients with known pack years, 109 (14.0%) smoked less than 10 pack years and 669 (86%) smoked at least 10 pack years. This suggests potentially a proverbial tipping point in terms of HNSCC development in regard to tobacco usage. This is contrasted against alcohol usage, as a total of 227 patients with known alcohol use frequency, 123 (54.2%) consumed less than 10 drinks per week and 104 (45.8%) consumed at least 10 drinks per week, exhibiting no significant difference in risk regarding total consumption, suggesting exposure and frequency might be more correlated to HNSCC development.

The median age of our cohort was 64.8 years old and the likelihood of male HNSCC development was 3 times as likely. These figures are similarly represented in national meta-analysis data with figures of 66-year-old mean age and a male HNSCC likelihood between

2 and 4 times. Within the female cohort, 341 (55.0%) patients use tobacco, 201 (32.4%) patients use alcohol, and 177 (28.5%) patients have diabetes. Within the male cohort, 1,077 (59.3%) patients use tobacco, 748 (41.2%) patients use alcohol, and 428 (23.6%) patients have diabetes. Interestingly, these primary risk factors are comparable between genders, and do not fully explain the significant increase in male driven development. This is contrasted against Park et al. 2022 in 10-year follow-up study of 10 million healthy participants that reported the sex differences in prevalence in HNSCC to be primarily driven by social risk factors such as alcohol and tobacco (10). Contradicting findings between the studies highlight the need for further research to be accomplished on this topic.

Our cohort encompassed 42 patients that were Black or African American, and many statistics regarding diagnosis and risk factors, such as alcohol and tobacco use, compared similarly to the larger cohort. Yet interestingly, Black individuals were noted to have a substantially lower prevalence rate by age 65 (28.6%) when compared against the cohort as a whole (48.7% at 65 or older at time of diagnosis). Additionally, mean age at diagnosis of Black or African American individuals was 62 seen against 66 years old for the entire cohort. The data suggests earlier ages of onset, diagnosis, and progression the Black and African American populations, but due to the small sample size of 42 patients, further research on this topic would be necessary.

While adding to the growing body of literature available for HNSCC, we also highlight important distinctions in risk patterns, particularly regarding several non-traditionally studied risk factors such as obesity, diabetes, and HPV. Both diabetes and obesity had very strong correlations to likelihood of development of HNSCC with 29% cases being obese patients and 24.8% having diabetes.

One of the most striking findings in our study is the relatively low proportion of HNSCC patients with a history of HPV (0.6%), despite the established role of HPV in oropharyngeal cancer. HPV is typically seen with development earlier in life (1), yet the average age of this cohort's HPV positive population was 73 years of age. Of note, the HPV positive sample size was limited with 14 total individuals which may contribute to the significant difference observed. However, our dataset may underrepresent this group due to regional variations in HPV prevalence, differences in screening practices, underreporting, or data collected prior to expected age of onset for HPV positive patients.

Distinguishing the most common site of invasive growth is of particular importance. It does not only guide clinicians in performing more thorough evaluations, but proper patient education may inform patients that dysplastic and cancerous growth may be commonly found in areas difficult to view on self-assessments areas, leading to encouraging more frequent professional examinations for at-risk populations. Our study found the primary site of HNSCC to be the tongue at 25.5% of all diagnoses' locations, being a fairly visible site that may be sensitive to disruptive stimulation in addition to lip (8.3%), while still a large percentage of locations were in limited visibility areas such as tonsillar (14.7%), supraglottic (11.1%), and palatine (10.6%). Although higher levels of exposure to social risk factors, such as tobacco and alcohol use, occur within the anterior cavity, a potential explanation could be related to the inability to identify the primary dysplastic growth prior to malignant transformation due to location.

Furthermore, other sites of importance that may not be conventionally thought to be high risk area, including pyriform sinus (2.2%) and nasopharyngeal (1.9%), suggesting that individuals scoring within the high-risk category should encourage communication between dental, ENT, and general practitioners to ensure holistic inspection and screening as occurred. Typically, by the time self-assessment warning signs have been recognized, such as dysphagia (difficulty eating), odynophagia (pain when swallowing) or otalgia (ear pain), tumor progression is quite advanced, highlighting the importance of effective in office examinations and screening methods (1). Of note, 13.3% were not accurately described enough to validate a location for the study.

While our study did not directly examine educational attainment, the significant proportion of patients who developed HNSCC despite not smoking or consuming alcohol suggests that additional social determinants of health, such as occupational exposures, health care access, and nutrition as demonstrated in The INHANCE study (11) may play a role in disease development. These findings highlight the need for broader risk assessment models that incorporate social and environmental factors alongside behavioral risk factors.

Age-related trends in our study also align with previous findings. We observed that the absolute risk of HNSCC increased with age but declined after 60–70 years, potentially due to competing mortality risks from other conditions. Given the high burden of HNSCC in older populations, refining screening strategies to target high-risk individuals while balancing the risks of overdiagnosis remains a crucial challenge.

Prior studies have explored the use of several biomarkers described that were shown to be associated with HNSCC and its development as potential tools in diagnosis. Looking ahead, future research should continue focusing on refining our risk model by integrating molecular and genetic biomarkers to enhance predictive accuracy. Advances in liquid biopsy techniques, including the detection of circulating tumor DNA and protein biomarkers, hold promise for improving early detection in asymptomatic individuals.

For example, IL-8 as a biomarker for oncogenesis has been shown to have a sensitivity of 0.65-0.85 and specificities of 0.79-0.93 in vivo animal studies over the last several years. This is thought to be due to IL-8 having been shown to be elevated in the microenvironment that the cells live in, such as hypoxic or acidic conditions (i.e., smoking, chewing tobacco, poor dental hygiene, alcohol, poor diet). Therefore, IL-8 in turn induces further genetic mutations through over stimulation of pathways controlling cellular growth, division, and maintenance (i.e., NOD1, RIP2 pathways) (12–16). By pairing effective screening methods and tools, such as the novel predictive scoring system with new techniques of salivary or serum screenings, tumor cell development and viability can be properly identified, assessed for risk, and diagnosed efficiently and relatively inexpensively.

Despite the strengths of our study, several limitations should be acknowledged. First, our study is retrospective in nature, which may introduce selection bias and limit the ability to establish causal relationships between risk factors and HNSCC development. Second, our dataset relies on electronic health records from a single health care system, which may not be fully representative of the broader population and could limit the generalizability of our findings. Third, certain risk factors such as genetic predisposition, environmental exposures, and socioeconomic variables were not comprehensively accounted for, potentially underestimating their contributions to HNSCC risk. Lastly, the reliance on self-reported data for tobacco and alcohol use may introduce recall bias, which could impact the accuracy of our risk stratification model. Future studies incorporating prospective cohort designs and more comprehensive risk factor assessments will be essential to validate and refine our predictive model.

Looking ahead, future research should focus on refining our risk model by integrating molecular and genetic biomarkers to enhance predictive accuracy. Advances in liquid biopsy techniques, including the detection of circulating tumor DNA and protein biomarkers, hold promise for improving early detection in asymptomatic individuals.

Conclusion

Tobacco use is identified as the highest risk factor that contributes to developing HNSCC. Patients scoring 6 or more points should undergo more rigorous and proactive screening for early detection. Given the severity of HNSCC, implementing this scoring system is crucial for reducing undiagnosed cases and improving patient outcomes. Future study should focus on which factors are associated with earlier onset of HNSCC and prediction of survival rate based on risk scores.

Disclosures

All authors declare no conflicts of interest relevant to this study. This research was supported by funding from the Geisinger Commonwealth School of Medicine under CRF number 24-062.

Acknowledgements

We would like to express our gratitude to Dr. Lauren Blaha for her invaluable guidance and expertise in the study design and refinement of data variables. We also acknowledge Jordan Law for his essential support with data analysis. Finally, we are grateful to Geisinger Commonwealth School of Medicine for their generous financial support, which made this project possible.

References

1. Johnson DE, Burtness B, Leemans CR, Lui VWY, Bauman JE, Grandis JR. Head and neck squamous cell carcinoma. Nat Rev Dis Primers. 2020 Nov 26;6(1):92. Erratum in: *Nat Rev Dis Primers*. 2023 Jan 19;9(1):4.

2. Du E, Mazul AL, Farquhar D, Brennan P, Anantharaman D, Abedi-Ardekani B, Weissler MC, Hayes DN, Olshan AF, Zevallos JP. Long-term Survival in Head and Neck Cancer: Impact of Site, Stage, Smoking, and Human Papillomavirus Status. *Laryngoscope*. 2019 Nov;129(11):2506-2513.
3. Johnson DE, Burtness B, Leemans CR, Lui VWY, Bauman JE, Grandis JR. Head and neck squamous cell carcinoma. *Nat Rev Dis Primers*. 2020 Nov 26;6(1):92. doi:
4. Roman BR, Aragones A. Epidemiology and incidence of HPV-related cancers of the head and neck. *J Surg Oncol*. 2021 Nov;124(6):920-922.
5. Rettig EM, D'Souza G. Epidemiology of head and neck cancer. *Surg Oncol Clin N Am*. 2015 Jul;24(3):379-96.
6. Di Credico G, Polesel J, Dal Maso L, Pauli F, Torelli N, Luce D, Radoï L, Matsuo K, Serraino D, Brennan P, Holcatova I, Ahrens W, Lagiou P, Canova C, Richiardi L, Healy CM, Kjaerheim K, Conway DI, Macfarlane GJ, Thomson P, Agudo A, Znaor A, Franceschi S, Herrero R, Toporcov TN, Moyses RA, Muscat J, Negri E, Vilensky M, Fernandez L, Curado MP, Menezes A, Daudt AW, Koifman R, Wunsch-Filho V, Olshan AF, Zevallos JP, Sturgis EM, Li G, Levi F, Zhang ZF, Morgenstern H, Smith E, Lazarus P, La Vecchia C, Garavello W, Chen C, Schwartz SM, Zheng T, Vaughan TL, Kelsey K, McClean M, Benhamou S, Hayes RB, Purdue MP, Gillison M, Schantz S, Yu GP, Chuang SC, Boffetta P, Hashibe M, Yuan-Chin AL, Edefonti V. Alcohol drinking and head and neck cancer risk: the joint effect of intensity and duration. *Br J Cancer*. 2020 Oct;123(9):1456-1463.
7. Chung MS, Choi YJ, Kim SO, Lee YS, Hong JY, Lee JH, Baek JH. A Scoring System for Prediction of Cervical Lymph Node Metastasis in Patients with Head and Neck Squamous Cell Carcinoma. *AJNR Am J Neuroradiol*. 2019 Jun;40(6):1049-1054.
8. Dhull AK, Atri R, Dhankhar R, Chauhan AK, Kaushal V. Major Risk Factors in Head and Neck Cancer: A Retrospective Analysis of 12-Year Experiences. World J Oncol. 2018 Jun;9(3):80-84.
9. Gormley M, Creaney G, Schache A, Ingarfield K, Conway DI. Reviewing the epidemiology of head and neck cancer: definitions, trends and risk factors. *Br Dent J*. 2022 Nov;233(9):780-786.
10. Park JO, Nam IC, Kim CS, Park SJ, Lee DH, Kim HB, Han KD, Joo YH. Sex Differences in the Prevalence of Head and Neck Cancers: A 10-Year Follow-Up Study of 10 Million Healthy People. Cancers (Basel). 2022 May 20;14(10):2521.
11. Lee YA, Al-Temimi M, Ying J, Muscat J, Olshan AF, Zevallos JP, Winn DM, Li G, Sturgis EM, Morgenstern H, Zhang ZF, Smith E, Kelsey K, McClean M, Vaughan TL, Lazarus P, Chen C, Schwartz SM, Gillison M, Schantz S, Yu GP, D'Souza G, Gross N, Monroe M, Kim J, Boffetta P, Hashibe M. Risk Prediction Models for Head and Neck Cancer in the US Population From the INHANCE Consortium. *Am J Epidemiol*. 2020 Apr 2;189(4):330-342.
12. Chan LP, Wang LF, Chiang FY, Lee KW, Kuo PL, Liang CH. IL-8 promotes HNSCC progression on CXCR1/2-meidated NOD1/RIP2 signaling pathway. *Oncotarget*. 2016 Sep 20;7(38):61820-61831.
13. Pereira LH, Reis IM, Reategui EP, Gordon C, Saint-Victor S, Duncan R, Gomez C, Bayers S, Fisher P, Perez A, Goodwin WJ, Hu JJ, Franzmann EJ. Risk Stratification System for Oral Cancer Screening. *Cancer Prev Res* (Phila). 2016 Jun;9(6):445-55.
14. Hema Shree K, Ramani P, Sherlin H, Sukumaran G, Jeyaraj G, Don KR, Santhanam A, Ramasubramanian A, Sundar R. Saliva as a Diagnostic Tool in Oral Squamous Cell Carcinoma - a Systematic Review with Meta Analysis. *Pathol Oncol Res*. 2019 Apr;25(2):447-453.
15. Pall AH, Jakobsen KK, Grønhøj C, von Buchwald C. Circulating tumour DNA alterations as biomarkers for head and neck cancer: a systematic review. *Acta Oncol*. 2020 Jul;59(7):845-850.
16. Kaskas NM, Moore-Medlin T, McClure GB, Ekshyyan O, Vanchiere JA, Nathan CA. Serum biomarkers in head and neck squamous cell cancer. *JAMA Otolaryngol Head Neck Surg*. 2014 Jan;140(1):5-11.