Survival in Head and Neck Squamous Cell Carcinoma by Clinical Stage and Tumor Site

By: Stephanie Chu & Thanh Mai

INTRODUCTION

Cancer remains one of the leading causes of death worldwide and is the second leading cause of death in the United States, accounting for approximately one in six deaths each year.[1] Among the many cancer types, head and neck cancers represent a significant global health burden and were estimated to be the seventh most common cancer worldwide in 2018.[2] Head and neck squamous cell carcinoma (HNSCC) refers to a group of malignant tumors that arise from the squamous cells lining the mucosal surfaces of the upper aerodigestive tract.[3] These surfaces include anatomical structures involved in breathing, speaking, and swallowing, particularly within the oral cavity, pharynx, and larynx. The disease disproportionately affects men, with incidence rates in males approximately two to four times higher than in females. This disparity has been largely attributed to historically higher rates of behavioral risk factors such as tobacco use and alcohol consumption among men.[4]

The prognosis for patients with HNSCC can vary widely. Clinical stage at the time of diagnosis is one of the most well-established predictors of patient survival, with earlier stages generally associated with more favorable outcomes.[5] However, patients classified within the same clinical stage may still experience different survival outcomes depending on the primary site of the tumor. These differences may reflect variations in tumor biology, patterns of disease progression, and differences in treatment approaches across anatomical sites.

Understanding how survival outcomes vary by both clinical stage and tumor location is important for improving prognostic assessment and guiding treatment strategies. Therefore, the aim of this study is to examine whether overall survival differs based on clinical stage and primary tumor site among patients diagnosed with HNSCC. By evaluating these factors together, this analysis seeks to provide further insight into the relationship between tumor location, stage at diagnosis, and patient prognosis.

DATA

This analysis used the publicly available University of Michigan Head and Neck Specialized Program of Research Excellence (SPORE) dataset,[6] which includes data from a cohort of 1,843 patients diagnosed with head and neck squamous cell carcinoma (HNSCC) between 2003 and 2014. The dataset contains patient demographics, tumor characteristics, self-reported behavioral factors, and clinical outcomes.

METHODS

For this analysis, we examined variables relevant to survival outcomes, including age at diagnosis, gender, race, smoking status, drinking status, clinical stage at diagnosis, and primary tumor site. Age was modeled as a continuous variable. Gender was categorized as male or female, and race as White, Black, Asian, American Indian, or Other. Smoking and drinking status were classified as never, current, or former (quit >12 months). Clinical stage was defined according to the American Joint Committee on Cancer, using stage groups from 0 to 4. The primary tumor site was initially recorded in ten categories (Glottic, Larynx, Oral Cavity, Oropharynx, Hypopharynx, Skull Base, Unknown Primary, Nasopharynx, Salivary Gland, and Other) and subsequently collapsed into four clinically meaningful groups to ensure adequate sample size: Larynx (Glottic and Larynx), Oral Cavity, Oropharynx, and Other (Hypopharynx, Skull Base, Unknown Primary, Nasopharynx, Salivary Gland, and other rare sites). No patients were excluded based on these variables.

Baseline characteristics were summarized by clinical stage at diagnosis. Age was described using the median and interquartile range and compared across stage groups using one-way ANOVA. Categorical variables (gender, race, smoking status, drinking status, and collapsed tumor site) were summarized using counts and percentages, with differences assessed using chi-square tests.

Overall survival was defined as the time from diagnosis to death or last follow-up. Survival functions by clinical stage and tumor site were estimated using Kaplan–Meier methods and compared using log-rank tests. Cox proportional hazards regression was used to estimate hazard ratios. Two separate unadjusted models were fit: one including clinical stage as the primary predictor and one including collapsed tumor site as the primary predictor. Multivariable models were then constructed adjusting for age, gender, race, smoking status, and drinking status. The proportional hazards assumption was assessed using standard diagnostic methods, including evaluation of Schoenfeld residuals. Evidence of a potential violation of the proportional hazards assumption was observed for the tumor site variable. To address this issue, an additional Cox model stratified by tumor site was fit. All statistical tests were two-sided and conducted using a significance level of $\alpha = 0.05$.

RESULTS

A. Baseline Characteristics by Clinical Stage at Diagnosis

Baseline characteristics of the study population stratified by clinical stage are presented in Table 1. Tumor sites differed significantly across stages, with laryngeal cancers more frequently observed in earlier stages and oropharyngeal cancers appearing more often in later stages. Gender distribution also varied significantly by stage. Although the majority of patients were male across all stages, there was a slight increase in the proportion of female patients in stage 2. In contrast, race, smoking status, drinking status, and age were similarly distributed across clinical stages, with no statistically significant differences observed.

Table 1. Summary of Baseline Characteristics by Clinical Stage at Diagnosis

Characteristic	**Stage 0** N = 31[1]	**Stage 1** N = 220[1]	**Stage 2** N = 196[1]	**Stage 3** N = 249[1]	**Stage 4** N = 1,147[1]	**p-value**[2]
Tumor Site						<0.001
Larynx	12 (39%)	80 (36%)	55 (28%)	69 (28%)	157 (14%)	
Oral Cavity	17 (55%)	123 (56%)	101 (52%)	76 (31%)	252 (22%)	
Oropharynx	2 (6.5%)	12 (5.5%)	31 (16%)	78 (31%)	548 (48%)	
Other	0 (0%)	5 (2.3%)	9 (4.6%)	26 (10%)	190 (17%)	
Gender						<0.001
Female	10 (32%)	66 (30%)	78 (40%)	75 (30%)	232 (20%)	
Male	21 (68%)	154 (70%)	118 (60%)	174 (70%)	915 (80%)	
Race						0.15
White	30 (97%)	192 (89%)	179 (94%)	233 (95%)	1,032 (91%)	
Black	0 (0%)	15 (6.9%)	7 (3.7%)	5 (2.0%)	70 (6.2%)	
Asian	0 (0%)	2 (0.9%)	0 (0%)	1 (0.4%)	2 (0.2%)	
American Indian	1 (3.2%)	3 (1.4%)	2 (1.1%)	5 (2.0%)	19 (1.7%)	
Other	0 (0%)	4 (1.9%)	2 (1.1%)	1 (0.4%)	5 (0.4%)	
Unknown	0	4	6	4	19	
Smoking Status						0.5

Characteristic	Stage 0 N = 31[1]	Stage 1 N = 220[1]	Stage 2 N = 196[1]	Stage 3 N = 249[1]	Stage 4 N = 1,147[1]	p-value[2]
Never	4 (13%)	48 (22%)	46 (24%)	47 (19%)	284 (25%)	
Current	17 (55%)	100 (47%)	86 (45%)	116 (48%)	490 (44%)	
Former	10 (32%)	67 (31%)	58 (31%)	81 (33%)	352 (31%)	
Unknown	0	5	6	5	21	
Drinking Status						0.4
Never	4 (13%)	17 (7.9%)	26 (14%)	25 (10%)	104 (9.3%)	
Current	20 (65%)	143 (67%)	111 (58%)	149 (61%)	740 (66%)	
Former	7 (23%)	55 (26%)	53 (28%)	71 (29%)	278 (25%)	
Unknown	0	5	6	4	25	
Age	61 (49, 69)	61 (53, 71)	62 (52, 69)	59 (53, 68)	58 (52, 67)	0.2
Unknown	0	3	2	2	5	

[1]n (%); Median (Q1, Q3)

[2]Pearson's Chi-squared test with simulated p-value (based on 2000 replicates); One-way analysis of means (not assuming equal variances)

B. Kaplan-Meier Survival Analysis

To evaluate whether differences in baseline characteristics across clinical stages translated into differences in survival outcomes, Kaplan–Meier survival curves were generated to examine overall survival by clinical stage, and differences between groups were assessed using the log-rank test. As shown in Figure 1, survival differed significantly across clinical stages (log-rank $p < 0.001$). Patients diagnosed at earlier stages (0 and 1) experienced fewer deaths than expected, while patients diagnosed at later stages (3 and 4) experienced more deaths than expected. These findings indicate that stage at diagnosis is strongly associated with overall survival.

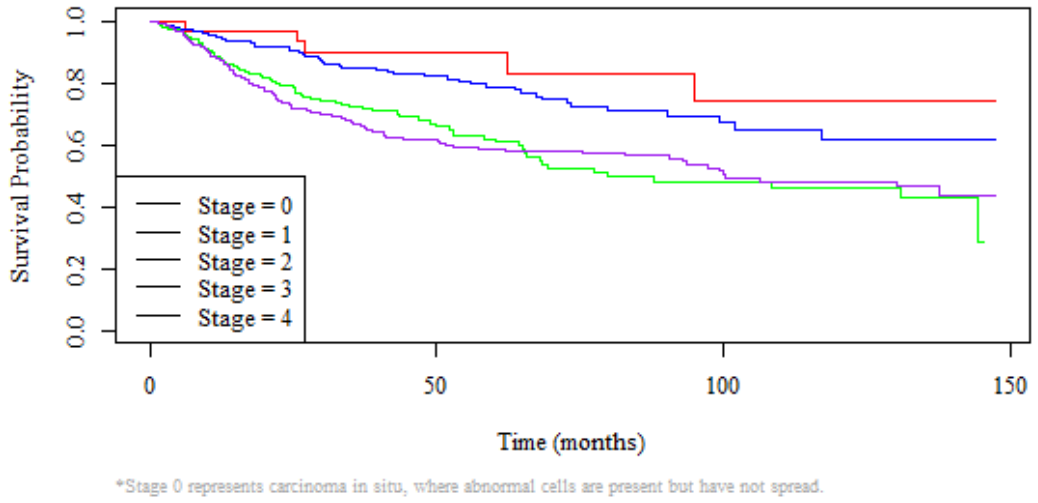

Figure 1. Kaplan-Meier Survival Curve by Clinical Stage at Diagnosis

Given the observed differences in survival by stage, a separate Kaplan–Meier analysis was conducted to evaluate overall survival by tumor site. The resulting survival curves are presented in Figure 2. The log-rank test indicated a statistically significant difference in survival across tumor sites ($p < 0.001$).

Laryngeal and oral cavity tumors generally demonstrated better survival outcomes, whereas oropharyngeal tumors showed fewer deaths early in follow-up but higher mortality later in the observation period. The remaining tumor sites exhibited intermediate survival patterns.

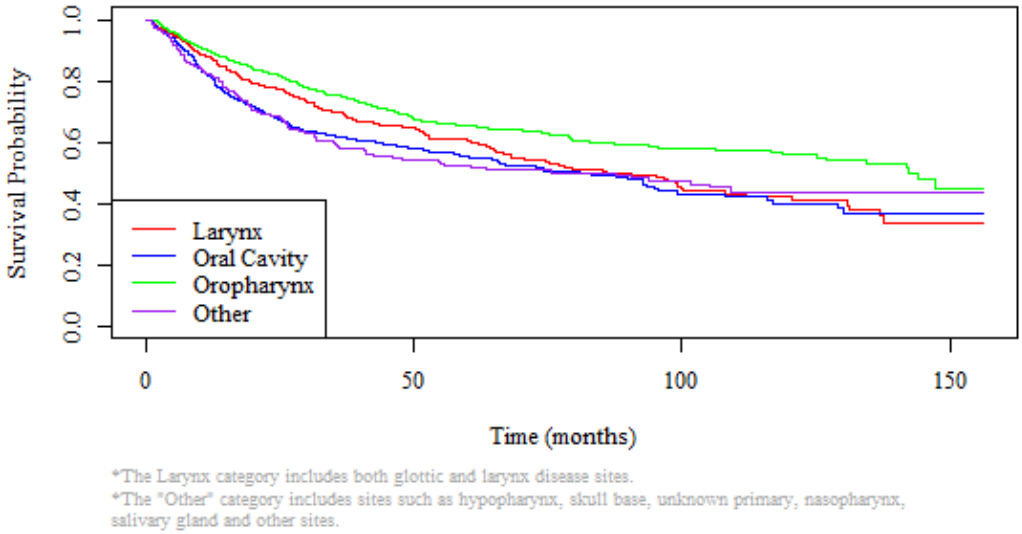

Figure 2. Kaplan-Meier Survival Curve by Tumor Site at Diagnosis

Together, these findings suggest that both clinical stage and tumor site may be important predictors of overall survival in patients with head and neck squamous cell carcinoma. Because significant differences were observed in the Kaplan–Meier analyses, Cox proportional hazards regression models were subsequently used to quantify these associations while adjusting for relevant covariates. Visual inspection of the Kaplan–Meier curves indicated some crossing of survival curves, highlighting the importance of assessing the proportional hazards assumption prior to interpreting the Cox model results.

C. Cox Proportional Hazards Models

Cox proportional hazards regression models were used to quantify the association between clinical stage, tumor site, and overall survival.

In the unadjusted Cox model examining clinical stage, later stages were associated with progressively higher hazards of death compared with the reference group (Stage 0). Patients diagnosed with Stage 2, Stage 3, and Stage 4 disease had significantly higher mortality risk, with hazard ratios of 3.10, 3.22, and 3.76, respectively. Table 2 contains the confidence intervals and p-values for the given hazard ratio estimates. Stage 1 was not significantly associated with mortality compared with Stage 0. These findings are consistent with the separation observed in the Kaplan–Meier curves, indicating substantially worse survival among patients diagnosed at more advanced stages.

Table 2. Summary of Unadjusted Cox-PH Models

Unadjusted Cox-PH Model - Clinical Stage				
Variable	**Hazard Ratio**	**Lower 95% CI**	**Upper 95% CI**	**p-value**
Stage 1	1.598	0.637	4.007	0.318
Stage 2	3.101	1.256	7.661	0.014
Stage 3	3.216	1.312	7.883	0.011

Unadjusted Cox-PH Model - Clinical Stage				
Variable	**Hazard Ratio**	**Lower 95% CI**	**Upper 95% CI**	**p-value**
Stage 4	3.762	1.559	9.074	0.003

Unadjusted Cox-PH Model - Tumor Site				
Variable	**Hazard Ratio**	**Lower 95% CI**	**Upper 95% CI**	**p-value**
Oral Cavity	1.162	0.958	1.410	0.127
Oropharynx	0.748	0.614	0.910	0.004
Other	1.132	0.892	1.436	0.309

In the unadjusted model examining tumor site, survival also differed across anatomical locations. Relative to the reference group (laryngeal tumors), oral cavity tumors were associated with a higher hazard of death, while oropharyngeal tumors demonstrated a lower hazard of death. Tumors classified as other sites showed intermediate survival patterns. These findings suggest that tumor location may influence survival outcomes among patients with head and neck squamous cell carcinoma. Because both clinical stage and tumor site were associated with survival in unadjusted analyses, a multivariable Cox model was subsequently fit to adjust for potential confounders.

A multivariable Cox model was fitted adjusting for age, gender, race, smoking status, and drinking status. After adjustment, clinical stage remained a strong predictor of mortality, with increasing stage associated with higher hazards of death. Compared with patients diagnosed at Stage 0, those diagnosed at Stage 2, Stage 3, and Stage 4 had significantly higher mortality risk, with hazard ratios of 3.40, 3.87, and 4.92, respectively. Age was also positively associated with mortality risk, with each additional year of age corresponding to a 4% increase in the hazard of death. Male patients experienced higher mortality compared with female patients (HR = 1.24). Smoking status was strongly associated with survival, as both current smokers (HR = 2.36) and former smokers (HR = 1.60) had significantly higher hazards of death compared with individuals that never smoked. In contrast, drinking status was not significantly associated with mortality after adjustment for other covariates. Survival differences were also observed across tumor sites, with oral cavity tumors associated with a higher hazard of death compared with laryngeal tumors (HR = 1.37), while other site comparisons were not statistically significant. Table 3 contains the confidence intervals and p-values for the given hazard ratio estimates.

Table 3. Summary of Adjusted Cox-PH Model

Adjusted Cox-PH Model				
Variable	**Hazard Ratio**	**Lower 95% CI**	**Upper 95% CI**	**p-value**
Clinical Stage				
Stage 1	1.459	0.580	3.673	0.422
Stage 2	3.404	1.375	8.422	0.008
Stage 3	3.875	1.577	9.522	0.003
Stage 4	4.918	2.029	11.918	<0.001
Tumor Site				
Oral Cavity	1.367	1.115	1.676	0.003
Oropharynx	0.807	0.648	1.005	0.056
Other	1.111	0.863	1.431	0.414

Adjusted Cox-PH Model				
Variable	**Hazard Ratio**	**Lower 95% CI**	**Upper 95% CI**	**p-value**
Age	1.043	1.035	1.051	<0.001
Male	1.237	1.033	1.479	0.020
Race				
Black	1.384	1.053	1.819	0.020
Asian	1.350	0.188	9.69	0.765
American Indian	0.727	0.400	1.325	0.298
Other	2.703	1.113	6.565	0.028
Smoking Status				
Current	2.356	1.879	2.955	<0.001
Former	1.603	1.270	2.023	<0.001
Drinking Status				
Current	0.991	0.759	1.295	0.949
Former	1.020	0.769	1.352	0.891

Tumor site was initially included as a covariate in the multivariable model. However, diagnostic testing using Schoenfeld residuals suggested a potential violation of the proportional hazards assumption for this variable. This finding was consistent with the Kaplan–Meier curves, which showed evidence of crossing survival curves across tumor site groups.

To address this violation, a stratified Cox model was fitted with tumor site included as a stratification factor. This approach allows the baseline hazard function to vary across tumor site strata while estimating common hazard ratios for the remaining covariates. In the stratified model, the proportional hazards assumption was satisfied for all variables. Clinical stage remained strongly associated with mortality, with higher stages associated with increased hazards of death after adjustment for demographic and behavioral factors. The estimated hazards of death were very similar to those observed in the adjusted Cox model that included tumor site as a covariate, suggesting that the overall association between clinical stage and survival was robust to the modeling approach. Other covariates also showed consistent associations with mortality across the two models. Age remained positively associated with mortality risk and male patients experienced higher mortality compared with female patients. Both current smokers and former smokers had significantly elevated hazards of death compared with individuals who never smoked. Race was also associated with mortality in some categories, while drinking status was not significantly associated with survival. As shown in Figure 3, the magnitude and direction of the hazard ratios in the stratified model closely mirror those from the adjusted model, indicating that accounting for non-proportional hazards by stratifying on tumor site did not materially change the estimated effects of the remaining covariates.

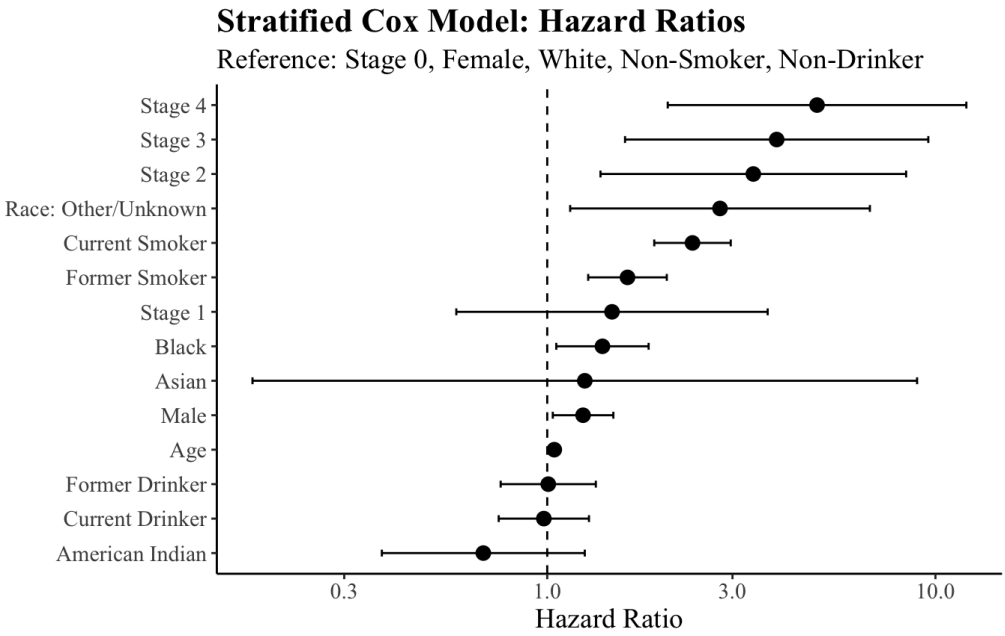

Figure 3. Forest Plot of Hazard Ratios from the Stratified Cox Model by Tumor Site

DISCUSSION

A. Key Points

This study examined whether overall survival among patients with head and neck squamous cell carcinoma differed according to clinical stage at diagnosis and primary tumor site. Kaplan–Meier analyses demonstrated significant differences in survival across both clinical stage and tumor site. Patients diagnosed at earlier stages experienced substantially better survival outcomes than those diagnosed at later stages, highlighting the importance of stage at diagnosis in determining prognosis. Cox proportional hazards models further quantified these associations. In both unadjusted and multivariable analyses, increasing clinical stage was strongly associated with higher hazards of death. Patients diagnosed with Stage 2, Stage 3, and Stage 4 disease experienced significantly greater mortality risk compared with those diagnosed at Stage 0. Age and smoking status were also important predictors of survival, with older patients and individuals with a history of smoking experiencing worse outcomes.

Initial multivariable models suggested that survival also differed across tumor sites. However, diagnostic testing indicated a violation of the proportional hazards assumption for tumor site. After fitting a stratified Cox model that allowed the baseline hazard to vary by tumor site, the estimated effects of clinical stage and other covariates remained largely unchanged. This finding indicates that the association between stage and survival is robust across tumor site strata, while also accounting for potential differences in baseline hazard functions across anatomical locations.

Together, these results emphasize that clinical stage at diagnosis remains a dominant predictor of survival in patients with head and neck squamous cell carcinoma, even after adjusting for demographic and behavioral factors.

B. Limitations & Future Research

Several limitations should be considered. First, the dataset represents patients enrolled in a single institutional research cohort, which may limit generalizability to broader patient populations. Second,

behavioral variables such as smoking and alcohol use were self-reported and may be subject to misclassification. Third, several rare tumor sites were combined into a single “Other” category to ensure adequate sample sizes, potentially obscuring differences among less common anatomical sites. Additionally, the dataset did not include certain clinical factors that may influence survival, such as treatment modality, HPV status, tumor grade, or comorbid conditions. Incorporating these variables in future studies could improve understanding of prognostic factors in head and neck cancer.

Future research using larger, multi-institutional datasets could further investigate differences across tumor sites and explore how treatment strategies or biological characteristics influence survival outcomes. In this analysis, tumor site was ultimately included as a stratification variable to address violations of the proportional hazards assumption, which prevented direct estimation of its effect on survival within the final Cox model. Future work could explore interaction models to examine whether the association between clinical stage and mortality differs across tumor sites. Such analyses could help clarify whether the impact of disease stage on survival varies by anatomical location, providing additional insight into potential biological or clinical differences across tumor subtypes.

C. Conclusion

This study confirms that clinical stage at diagnosis remains the primary determinant of survival in patients with head and neck squamous cell carcinoma, reinforcing the critical role of early detection and accurate staging in clinical practice. Although tumor site may influence survival patterns, stratified analyses indicate that the prognostic impact of stage is consistent across anatomical locations, suggesting that stage captures the dominant clinical risk regardless of tumor location.

These findings have important implications for patient counseling, risk stratification, and treatment planning, emphasizing that patients diagnosed at advanced stages require closer monitoring and potentially more aggressive management. Furthermore, the study highlights the need to consider patient characteristics such as age and smoking history, which independently contribute to mortality risk.

Future research leveraging larger, multi-institutional datasets and incorporating additional clinical and biological factors such as HPV status, tumor grade, comorbidities, and treatment modalities, could further elucidate how tumor location and disease stage interact to influence survival. Understanding these interactions may support more individualized prognostic assessments and guide targeted therapeutic strategies, ultimately improving outcomes for patients with head and neck cancer.

REFERENCES

1. *Cancer Fact Sheet*. (2025, February 3). World Health Organization. https://www.who.int/news-room/fact-sheets/detail/cancer
2. Vilches, L. A., Anantharaman, D., Brennan, P., & Leemans, C. R. (2020). Head and neck cancers. In *World Cancer Report: Cancer research for cancer prevention*. International Agency for Research on Cancer. https://www.ncbi.nlm.nih.gov/books/NBK606499/
3. Johnson, D. E., Burtness, B., Leemans, C. R., Lui, V. W. Y., Bauman, J. E., & Grandis, J. R. (2020). Head and neck squamous cell carcinoma. *Nature Reviews. Disease Primers*, *6*(1), 92. https://doi.org/10.1038/s41572-020-00224-3
4. Zhang, C. (2024, April 24). Cancer Types | Head and Neck Cancer—National Foundation for Cancer Research. *NFCR*. https://www.nfcr.org/cancer-types/cancer-types-head-neck-cancers/
5. Worsham, M. J. (2011). Identifying the risk factors for late stage head and neck cancer. *Expert Review of Anticancer Therapy*, *11*(9), 1321–1325. https://doi.org/10.1586/era.11.135
6. Bellile, E. L., Taylor, J. M., & Wolf, G. T. (2023, June 13). *Head and Neck Cancer University of Michigan SPORE clinical outcomes analytic dataset*.