

Incorporation of grade into stage in oral cavity squamous cell carcinoma: A novel staging schema

Allen S. Ho^{a,b,*}, Michael Luu^{b,c}, Daniel Manzoor^d, Horacio Maluf^d, Cedric Bailey^d, Bonnie Balzer^d, Evan S. Walgama^{a,b}, Julie K. Jang^{b,e}, Kevin C. Scher^{b,f}, Justin T. Moyers^{b,f,g}, Jon Mallen-St. Clair^{a,b}, Joel B. Epstein^b, Zachary S. Zumsteg^{b,e}

^a Division of Otolaryngology-Head and Neck Surgery, Dept. of Surgery, Cedars-Sinai Medical Center, United States

^b Samuel Oschin Comprehensive Cancer Institute, United States

^c Dept. of Biostatistics and Bioinformatics, Cedars-Sinai Medical Center, United States

^d Department of Pathology, Cedars-Sinai Medical Center, United States

^e Department of Radiation Oncology, Cedars-Sinai Medical Center, United States

^f Division of Medical Oncology, Dept. of Medicine, Cedars-Sinai Medical Center, United States

^g The Angeles Clinic and Research Institute, Cedars-Sinai Medical Center, United States

ARTICLE INFO

Keywords:

Oral cavity carcinoma
Oral cancer
Tongue cancer
Head and neck squamous cell carcinoma
Histologic grade
Cancer staging

ABSTRACT

Background: While numerous cancer staging systems have incorporated grade into stage, the impact of grade on oral cavity carcinoma (OCC) prognosis has been conflicting. We investigated grade as a prognostic determinant in OCC staging.

Methods: Multivariable Cox regression models of OCC patients identified via U.S. cancer registry data were constructed to determine associations between grade (G1 = low-grade, G2 = intermediate-grade, G3 = high-grade) and overall survival (OS). Recursive partitioning analysis (RPA) was used to derive staging schema.

Results: Overall, 46,789 OCC cases were identified across 1,222 institutions. On univariate analysis, higher grade was associated with worse 5-yr OS (G1: 73% [95% CI 72–74%], G2: 61% [95% CI 60–61%], G3: 49% [95% CI 48–51%]) ($p < 0.001$). On multivariable analysis adjusting for other prognostic factors, these survival differences persisted. Compared to G1 tumors, both G2 (HR 1.25 [95% CI 1.19–1.30], $p < 0.001$) and G3 (HR 1.52 [95% CI 1.45–1.61], $p < 0.001$) tumors were associated with significantly worse OS. Similar results were seen when utilizing propensity score matching. RPA generated subgroups that mirrored AJCC8E, but with G3 cases performing worse for a given stage. A proposed TNM + G staging schema was created with AJCC8E G3 cases upstaged by one category. Overall, 11.5% (4,745/36,623) cases were upstaged. TNM + G performed better than AJCC8E by c-index (0.673 vs. 0.671) and Brier score (0.174 vs. 0.175).

Conclusion: Large-scale analysis supports grade as an influential predictive determinant in OCC outcomes, with hazard on par with conventional staging factors. Incorporation of grade into stage strengthens existing AJCC OCC schema and pragmatically improves its ability to convey prognosis.

Introduction

Across numerous cancers, histologic grade represents a widely accepted, foundational prognostic factor influencing outcomes and management. Staging systems have evolved to incorporate grade alongside conventional AJCC TNM factors, including breast cancer (Clinical Prognostic Stage), prostate cancer (Grade Group status), sarcoma, and esophageal malignancies.[1–4] For salivary cancers, clinical

practice guidelines and the World Health Organization (WHO) recognize grade as an important prognostic factor that drives treatment recommendations, including dedifferentiation as a high-risk, biologically aggressive state.[5,6].

As a commonly reported parameter, grade may constitute a pragmatic addition to current staging schema for oral cavity carcinoma. However, the data supporting its inclusion have been inconsistent. While some studies correlate high grade with locoregional recurrence

* Corresponding author at: Cedars-Sinai Medical Center, 8631 West Third St., Suite 915E, Los Angeles, CA 90048, United States.

E-mail address: allen.ho@cshs.org (A.S. Ho).

<https://doi.org/10.1016/j.oraloncology.2025.107847>

Received 11 November 2025; Received in revised form 24 December 2025; Accepted 30 December 2025

Available online 13 January 2026

1368-8375/© 2025 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

and risk of death [7–11], others have concluded its prognostic impact to be weak and conflicting. [12–16] As such, its importance may in fact be understated and overlooked, deserving further investigation.

Here, we conduct a large-scale analysis of grade in oral cavity carcinoma to better delineate its spectrum of behavior. We compare its impact relative to more established prognostic factors and measure grade's comparative utility in staging.

Methods

Retrospective data were abstracted from the National Cancer Database (NCDB), a tumor registry maintained by the American Cancer Society and the American College of Surgeons. The inclusion period was 2004–2019, with the final analysis date as the 2023 annual NCDB data release. The study was deemed exempt by the Cedars-Sinai institutional review board. All patients > 18 years of age with oral cavity squamous cell carcinoma, with known grade, and treated for curative intent were eligible. Reporting recommendations for tumor marker prognostic studies (REMARK) checklist guidelines are listed in **Table S1**.

NCDB selection criteria. Cases were selected from ICD to 10 codes, encompassing tongue (C02.0, C02.1, C02.2, C02.3, C02.9), gum/alveolar ridge (C03.0, C03.1, C03.9), floor of mouth (C04.0, C04.1, C04.9), hard palate (C05.0, C05.9), buccal mucosa (C06.0, C06.1), retromolar trigone (C06.2), and mouth NOS (C06.8, C06.9). Only squamous cell carcinoma ICD-O-3 morphology (8070–8078) was allowed.

Grade was transcribed from the final pathology report by a trained NCDB coder. NCDB defines well-differentiated grade (low grade or G1) to appear most like normal tissue, while moderately differentiated

(intermediate grade or G2) appears intermediate in character. We combined poorly differentiated and undifferentiated grades (high grade or G3), defined as least like normal tissue. When more than one grade was reported, the highest grade was coded. Only grade from the primary site was considered, rather than metastatic sites.

For years 2004–2017, T-classification was converted to AJCC 8th edition (AJCC 8E) by incorporating tumor depth (CS Site Specific Factor 11). Similarly, N-classification was converted to AJCC 8E by incorporating extranodal extension (CS Site Specific Factor 8). Cases in years 2018–2019 were listed natively as AJCC 8E.

Statistical methods. Baseline patient demographics stratified by major histologic variants within the NCDB cohort were compared using the Kruskal-Wallis rank sum test and the Pearson's Chi-square test for continuous and categorical variables as appropriate. The primary outcome of interest was overall survival (OS). Median patient follow-up was estimated using the reverse Kaplan-Meier method. Survival probabilities were estimated using the Kaplan-Meier method with survival curves compared using the log-rank test.

Univariate and multivariable survival analysis were performed using the Cox proportional hazard regression with pre-specified covariates included in the model. The proportional hazards assumption was assessed both visually using scaled Schoenfeld residuals, and the goodness of fit test.[17] For the multivariable analyses, missing data for covariates were imputed using the MICE (Multivariate Imputation with Chained Equation) package in R [18], under the assumption that data were missing at random. Predictive mean matching was applied for continuous variables, polytomous logistic regression for categorical variables with more than two categories, and binary logistic regression

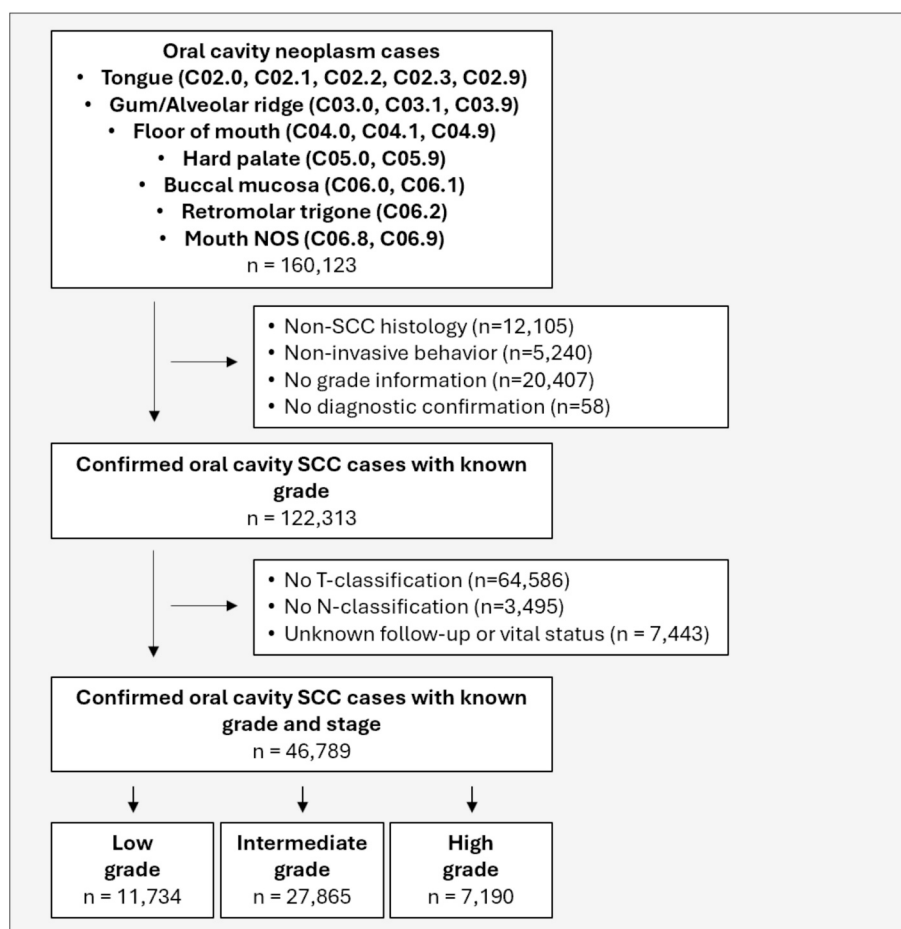


Fig. 1. Consolidated Standards of Reporting Trials (CONSORT) diagram for oral cavity squamous cell carcinoma. SCC, squamous cell carcinoma; LN, lymph node; RT, radiation therapy; chemo, chemotherapy; M1, distant metastasis.

Table 1

Patient demographics for HPV(+) oral cavity squamous cell carcinoma. Academic is defined as care received at an academic/research program classified by the Commission on Cancer Accreditation. ENE, extranodal extension.

Characteristic	Overall N = 46,789 ¹	Low Grade N = 11,734 ¹	Intermediate Grade N = 27,865 ¹	High Grade N = 7,190 ¹	p-value ²
Age					<0.001
Mean (SD)	64 (13)	65 (13)	64 (13)	63 (13)	
Median (IQR)	64 (56, 73)	66 (56, 75)	64 (55, 73)	63 (55, 72)	
Sex					<0.001
Male	27,205 (58%)	6,378 (54%)	16,587 (60%)	4,240 (59%)	
Female	19,584 (42%)	5,356 (46%)	11,278 (40%)	2,950 (41%)	
Race					<0.001
White	39,494 (85%)	10,039 (86%)	23,581 (85%)	5,874 (82%)	
Black	2,242 (4.8%)	423 (3.6%)	1,374 (5.0%)	445 (6.2%)	
Asian	2,564 (5.5%)	643 (5.5%)	1,490 (5.4%)	431 (6.0%)	
Latino	2,049 (4.4%)	511 (4.4%)	1,154 (4.2%)	384 (5.4%)	
Unknown	440	118	266	56	
Academic					0.003
Non-Academic	16,597 (37%)	4,306 (38%)	9,841 (37%)	2,450 (36%)	
Academic	28,242 (63%)	6,975 (62%)	16,854 (63%)	4,413 (64%)	
Unknown	1,950	453	1,170	327	
Region					<0.001
East	9,456 (21%)	2,286 (20%)	5,662 (21%)	1,508 (22%)	
South	15,648 (35%)	3,904 (35%)	9,461 (35%)	2,283 (33%)	
Midwest	12,532 (28%)	3,210 (28%)	7,424 (28%)	1,898 (28%)	
West	7,203 (16%)	1,881 (17%)	4,148 (16%)	1,174 (17%)	
Unknown	1,950	453	1,170	327	
Insurance					<0.001
Private	17,841 (38%)	4,471 (38%)	10,624 (38%)	2,746 (38%)	
Uninsured	1,396 (3.0%)	294 (2.5%)	869 (3.1%)	233 (3.2%)	
Medicaid	4,076 (8.7%)	754 (6.4%)	2,530 (9.1%)	792 (11%)	
Medicare	22,191 (47%)	5,913 (50%)	13,046 (47%)	3,232 (45%)	
Other	1,285 (2.7%)	302 (2.6%)	796 (2.9%)	187 (2.6%)	
Income					<0.001
<38 k	6,197 (16%)	1,391 (14%)	3,793 (16%)	1,013 (16%)	
38 k-63 k	20,013 (50%)	4,933 (49%)	11,987 (51%)	3,093 (50%)	
63 k+	13,520 (34%)	3,645 (37%)	7,773 (33%)	2,102 (34%)	
Unknown	7,059	1,765	4,312	982	
Education					<0.001
Low	6,015 (15%)	1,382 (14%)	3,644 (15%)	989 (16%)	
Medium	23,543 (59%)	5,870 (59%)	13,987 (59%)	3,686 (59%)	
High	10,188 (26%)	2,725 (27%)	5,927 (25%)	1,536 (25%)	
Unknown	7,043	1,757	4,307	979	
Urban/Rural					<0.001
Metro	36,769 (83%)	9,379 (84%)	21,769 (82%)	5,621 (83%)	
Urban	6,955 (16%)	1,600 (14%)	4,293 (16%)	1,062 (16%)	
Rural	802 (1.8%)	210 (1.9%)	463 (1.7%)	129 (1.9%)	
Unknown	2,263	545	1,340	378	
Crowfly					<0.001
0-10mi	19,771 (42%)	4,936 (42%)	11,805 (42%)	3,030 (42%)	
10-20mi	8,009 (17%)	2,133 (18%)	4,665 (17%)	1,211 (17%)	
21-50mi	9,589 (20%)	2,411 (21%)	5,766 (21%)	1,412 (20%)	
51-100mi	5,547 (12%)	1,343 (11%)	3,348 (12%)	856 (12%)	
100mi+	3,873 (8.3%)	911 (7.8%)	2,281 (8.2%)	681 (9.5%)	
Charlson Deyo					0.208
0	33,923 (73%)	8,588 (73%)	20,117 (72%)	5,218 (73%)	
1	8,653 (18%)	2,140 (18%)	5,215 (19%)	1,298 (18%)	
2	2,485 (5.3%)	611 (5.2%)	1,478 (5.3%)	396 (5.5%)	
3+	1,728 (3.7%)	395 (3.4%)	1,055 (3.8%)	278 (3.9%)	
Subsite					<0.001
Tongue	24,518 (52%)	6,179 (53%)	14,506 (52%)	3,833 (53%)	
Upper/Lower Gum	6,327 (14%)	1,983 (17%)	3,604 (13%)	740 (10%)	
FOM	7,103 (15%)	1,215 (10%)	4,630 (17%)	1,258 (17%)	
Hard Palate	1,106 (2.4%)	354 (3.0%)	616 (2.2%)	136 (1.9%)	
Buccal Mucosa	3,718 (7.9%)	1,169 (10.0%)	2,025 (7.3%)	524 (7.3%)	
RMT	2,362 (5.0%)	453 (3.9%)	1,475 (5.3%)	434 (6.0%)	
Mouth NOS	1,655 (3.5%)	381 (3.2%)	1,009 (3.6%)	265 (3.7%)	
T Classification					<0.001
T1	17,760 (38%)	6,333 (54%)	9,682 (35%)	1,745 (24%)	
T2	13,025 (28%)	2,775 (24%)	8,053 (29%)	2,197 (31%)	
T3	5,868 (13%)	925 (7.9%)	3,693 (13%)	1,250 (17%)	
T4	10,136 (22%)	1,701 (14%)	6,437 (23%)	1,998 (28%)	
N Classification					<0.001
N0	25,342 (64%)	7,518 (83%)	14,897 (61%)	2,927 (45%)	
N1	4,290 (11%)	599 (6.6%)	2,847 (12%)	844 (13%)	
N2	4,315 (11%)	439 (4.9%)	2,859 (12%)	1,017 (16%)	
N3	5,948 (15%)	480 (5.3%)	3,728 (15%)	1,740 (27%)	

(continued on next page)

Table 1 (continued)

Characteristic	Overall N = 46,789 ¹	Low Grade N = 11,734 ¹	Intermediate Grade N = 27,865 ¹	High Grade N = 7,190 ¹	p-value ²
M Classification					
Unknown	6,894	2,698	3,534	662	
M0	46,483 (99%)	11,699 (100%)	27,681 (99%)	7,103 (99%)	<0.001
M1	306 (0.7%)	35 (0.3%)	184 (0.7%)	87 (1.2%)	
LVI					
No	35,837 (83%)	9,996 (94%)	21,161 (82%)	4,680 (72%)	<0.001
Yes	7,119 (17%)	689 (6.4%)	4,597 (18%)	1,833 (28%)	
Unknown	3,833	1,049	2,107	677	
ENE					
No	33,474 (84%)	8,478 (94%)	20,324 (84%)	4,672 (72%)	<0.001
Yes	6,313 (16%)	538 (6.0%)	3,940 (16%)	1,835 (28%)	
Unknown	7,002	2,718	3,601	683	
Margin					
Negative	41,490 (90%)	10,795 (93%)	24,629 (89%)	6,066 (86%)	<0.001
Positive	4,690 (10%)	769 (6.6%)	2,904 (11%)	1,017 (14%)	
Unknown	609	170	332	107	
Surgery					
No	4,668 (10.0%)	1,708 (15%)	2,431 (8.7%)	529 (7.4%)	<0.001
Yes	42,121 (90%)	10,026 (85%)	25,434 (91%)	6,661 (93%)	
Postop RT					
No	29,095 (62%)	9,138 (78%)	16,539 (59%)	3,418 (48%)	<0.001
Yes	17,694 (38%)	2,596 (22%)	11,326 (41%)	3,772 (52%)	
Postop Chemo					
No	38,589 (83%)	10,798 (93%)	22,634 (82%)	5,157 (72%)	<0.001
Yes	7,924 (17%)	871 (7.5%)	5,069 (18%)	1,984 (28%)	
Unknown	276	65	162	49	

¹n (%)²Kruskal-Wallis rank sum test; Pearson's Chi-squared test

for binary variables. Five imputed datasets were generated with five iterations each, incorporating all model covariates and relevant auxiliary variables in the imputation models. Analyses were conducted separately on each imputed dataset, and results were pooled using Rubin's rules.

To further mitigate the potential for bias, propensity score adjusted survival curves were constructed stratified by grade. The propensity score was estimated using a multivariable logistic regression adjusting for covariates as listed in the multivariable Cox model. Balanced cohorts were constructed using 1-to-1 matching with the nearest neighbor algorithm, with balance diagnostics assessed using the standardized mean differences. [19].

Recursive partitioning analysis (RPA) using conditional inference trees was performed to develop a novel TNM + G staging system. The conditional inference tree was constructed with OS as the primary outcome with T-classification, N-classification, M-classification, and histologic grade as covariates. For practicality, low and intermediate grade cases were consolidated (G1 + G2) relative to high grade cases (G3). The implementation of this algorithm assessed a regression relationship by binary recursive partitioning among the covariates with a stop criterion based on a Bonferroni adjusted p-value of $p < 0.00000001$ thresholds for each branch. [20,21] Performance of the RPA-derived TNM + G staging system was compared with the AJCC 8th edition staging system using c-indices in patients with determinable AJCC stage. Internal validation was conducted using 10-fold cross-validation to adjust for potential optimism in performance estimates, evaluating discrimination via the c-index and calibration via the Brier score at 60 months. [22].

All statistical analyses were performed with R [23] (version 4.5.1; R Foundation, Vienna, Austria) and packages including tidyverse [24], rms [25], and partykit [26] with 2-sided tests and a significance level of 0.05.

Results

Overall, 46,789 oral cavity cases were identified across 1,222 institutions (Table 1, Fig. 1) (median follow-up of 53.0 [95% CI 53–54]

months). A broad spectrum of oral cavity subsites was identified, with tongue cancer (52.4%) comprising the majority of cases followed by FOM (15.2%). Relative to low-grade cases, high-grade cases harbored a higher incidence of negative prognostic features: this included T4 classification (14.5% vs. 27.8%, $p < 0.001$), N3 classification (5.3% vs. 27.0%, $p < 0.001$), M1 classification (0.3% vs. 1.2%, $p < 0.001$), LVI (6.4% vs. 28.1%, $p < 0.001$), ENE (6.0% vs. 28.2%, $p < 0.001$), and positive margins (6.6% vs. 14.0%, $p < 0.001$).

Anatomic subsite.

On univariate analysis, all subsites relative to tongue correlated with significantly worsening overall survival (OS) (Table 2). Estimated 5-year OS was significantly different between subtypes ($p < 0.001$) (tongue: 66% [95% CI 65%–67%]; gum/alveolar ridge: 61% [95% CI 60%–63%]; FOM: 56% [95% CI 55%–58%]; hard palate: 57% [95% CI 54%–61%]; buccal mucosa: 57% [95% CI 56%–59%]; RMT: 54% [95% CI 52%–57%]; mouth NOS: 53% [95% CI 52%–57%]). Notably, in multivariable analysis that adjusted for differences in stage, grade, and other prognostic factors, these survival differences substantially changed. Compared to tongue cases, gum/alveolar ridge cases (HR 0.82 [95% CI 0.78–0.86], $p < 0.0001$), hard palate cases (HR 0.91 [95% CI 0.82–1.00], $p = 0.057$), RMT cases (HR 0.99 [95% CI 0.92–1.05], $p = 0.69$), and mouth NOS (HR 0.94 [95% CI 0.86–1.02], $p = 0.11$) performed better. In contrast, FOM (HR 1.05 [95% CI 1.00–1.10], $p = 0.029$) and buccal mucosa (HR 1.07 [95% CI 1.01–1.13], $p = 0.022$) performed worse.

Grade.

On univariate analysis, higher grade was associated with worse overall survival. Estimated 5-year OS was significantly different between grade subtypes ($p < 0.001$) (G1: 73% [95% CI 72–74%], G2: 61% [95% CI 60–61], G3: 49% [95% CI 48–51] (Fig. 2a). On MVA, these survival differences persisted (Table 2). Compared to low-grade tumors, both intermediate-grade (HR 1.25 [95% CI 1.19–1.30], $p < 0.001$) and high-grade (HR 1.52 [95% CI 1.45–1.61], $p < 0.001$) tumors were associated with significantly worse overall survival.

To account for potential imbalances in prognostic factors in higher-grade versus lower-grade tumors, propensity score matching was performed. High-grade tumors continued to be associated with significant worse survival than lower-grade tumors even after balancing for other

Table 2

Univariate and multivariable analyses of overall survival in oral cavity carcinoma. HR, hazard ratio; CI, confidence interval; NOS, not otherwise specified; FOM, floor of mouth;; RMT, retromolar trigone; LVI, lymphovascular invasion; ENE, extranodal extension; RT, radiation therapy.

Characteristic	Univariate HR [†]	95% CI [†]	p-value	Multivariable HR [†]	95% CI [†]	p-value
Age	1.03	1.03–1.03	<0.001	1.03	1.03–1.03	<0.001
Sex						
Male	—	—	—	—	—	—
Female	0.90	0.87–0.92	<0.001	0.89	0.86–0.92	<0.001
Race						
White	—	—	—	—	—	—
Black	1.29	1.21–1.38	<0.001	1.07	1.00–1.15	0.055
Asian	0.76	0.71–0.82	<0.001	0.83	0.77–0.89	<0.001
Latino	0.90	0.83–0.97	0.006	0.83	0.77–0.91	<0.001
Academic						
Non-Academic	—	—	—	—	—	—
Academic	0.97	0.94–1.00	0.04	0.97	0.93–1.00	0.042
Region						
East	—	—	—	—	—	—
South	1.07	1.03–1.12	0.001	1.00	0.96–1.05	0.868
Midwest	1.08	1.03–1.13	<0.001	1.01	0.96–1.06	0.705
West	1.01	0.96–1.06	0.8	1.01	0.95–1.06	0.795
Insurance						
Private	—	—	—	—	—	—
Uninsured	1.55	1.42–1.70	<0.001	1.37	1.25–1.50	<0.001
Medicaid	1.84	1.74–1.95	<0.001	1.43	1.34–1.51	<0.001
Medicare	1.97	1.90–2.04	<0.001	1.20	1.15–1.26	<0.001
Other	1.61	1.46–1.77	<0.001	1.17	1.06–1.29	0.001
Income						
<38 k	—	—	—	—	—	—
38 k–63 k	0.89	0.86–0.93	<0.001	0.94	0.89–0.99	0.021
63 k+	0.72	0.69–0.76	<0.001	0.84	0.78–0.90	<0.001
Education						
Low	—	—	—	—	—	—
Medium	0.95	0.91–1.00	0.046	1.07	1.01–1.13	0.031
High	0.80	0.76–0.84	<0.001	1.03	0.96–1.10	0.46
Urban/Rural						
Metro	—	—	—	—	—	—
Urban	1.16	1.12–1.21	<0.001	1.04	0.99–1.09	0.116
Rural	1.05	0.94–1.18	0.4	0.95	0.84–1.07	0.399
Crowfly						
0–10mi	—	—	—	—	—	—
10–20mi	0.94	0.90–0.98	0.005	1.01	0.96–1.05	0.789
21–50mi	0.96	0.92–1.00	0.056	0.99	0.94–1.03	0.519
51–100mi	1.06	1.01–1.11	0.027	0.97	0.92–1.02	0.22
100mi+	1.01	0.96–1.07	0.7	0.90	0.85–0.96	0.001
Charlson Deyo						
0	—	—	—	—	—	—
1	1.32	1.27–1.37	<0.001	1.18	1.14–1.23	<0.001
2	1.55	1.46–1.65	<0.001	1.27	1.19–1.35	<0.001
3+	1.94	1.81–2.08	<0.001	1.55	1.44–1.66	<0.001
Subsite						
Tongue	—	—	—	—	—	—
Upper/Lower Gum	1.22	1.17–1.28	<0.001	0.82	0.78–0.86	<0.001
FOM	1.36	1.31–1.42	<0.001	1.05	1.00–1.10	0.029
Hard Palate	1.36	1.24–1.50	<0.001	0.91	0.82–1.00	0.057
Buccal Mucosa	1.37	1.30–1.45	<0.001	1.07	1.01–1.13	0.022
RMT	1.45	1.36–1.55	<0.001	0.99	0.92–1.05	0.686
Mouth NOS	1.54	1.42–1.66	<0.001	0.94	0.86–1.02	0.114
T-classification						
T1	—	—	—	—	—	—
T2	1.68	1.61–1.75	<0.001	1.48	1.42–1.55	<0.001
T3	2.50	2.38–2.62	<0.001	2.04	1.93–2.15	<0.001
T4	2.97	2.85–3.09	<0.001	2.39	2.27–2.51	<0.001
N-classification						
N0	—	—	—	—	—	—
N1	1.49	1.41–1.58	<0.001	1.37	1.29–1.46	<0.001
N2	2.21	2.11–2.32	<0.001	1.88	1.78–1.99	<0.001
N3	3.18	3.05–3.31	<0.001	2.00	1.73–2.30	<0.001
M-classification						
M0	—	—	—	—	—	—
M1	3.45	2.89–4.11	<0.001	1.93	1.62–2.31	<0.001
LVI						
No	—	—	—	—	—	—
Yes	2.01	1.94–2.09	<0.001	1.24	1.18–1.30	<0.001
ENE						
No	—	—	—	—	—	—

(continued on next page)

Table 2 (continued)

Characteristic	Univariate HR [†]	95% CI [†]	p-value	Multivariable HR [†]	95% CI [†]	p-value
Yes	2.74	2.64–2.84	<0.001	1.33	1.16–1.53	<0.001
Margin						
Negative	—	—	—	—	—	—
Positive	2.02	1.94–2.11	<0.001	1.48	1.41–1.55	<0.001
Surgery						
No	—	—	—	—	—	—
Yes	1.20	1.14–1.26	<0.001	0.95	0.90–1.00	0.05
Post-op RT						
No	—	—	—	—	—	—
Yes	1.36	1.32–1.40	<0.001	0.71	0.68–0.74	<0.001
Post-op Chemo						
No	—	—	—	—	—	—
Yes	1.71	1.65–1.77	<0.001	0.96	0.91–1.00	0.066
Grade						
Low (G1)	—	—	—	—	—	—
Intermediate (G2)	1.55	1.49–1.61	<0.001	1.25	1.19–1.30	<0.001
High (G3)	2.24	2.14–2.35	<0.001	1.52	1.45–1.61	<0.001

[†] HR = Hazard Ratio, CI = Confidence Interval

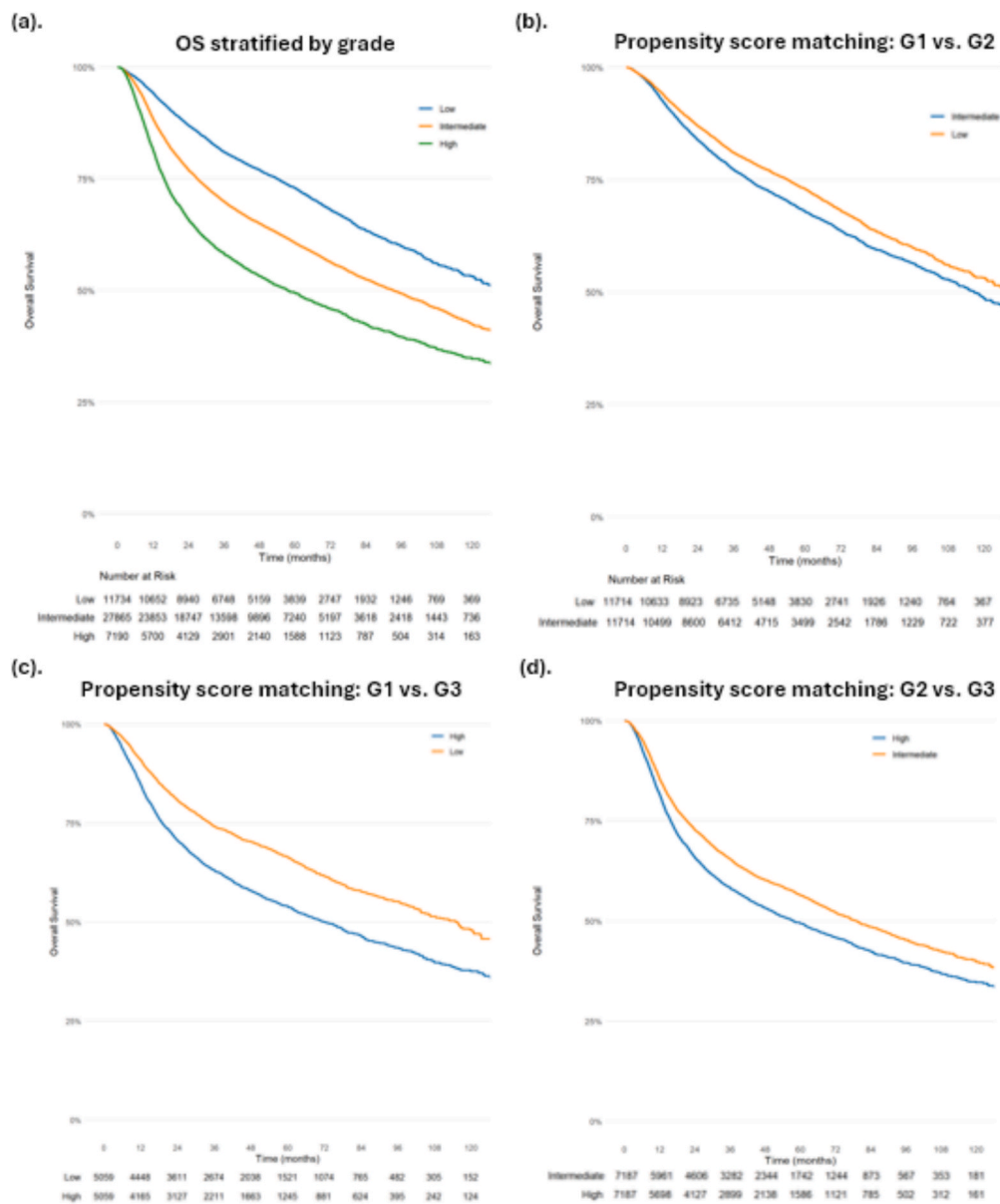


Fig. 2. Kaplan Meier curves illustrating overall survival for oral cavity squamous cell carcinoma (OCC) patients undergoing surgery ± adjuvant therapy. (a). All OCC patients stratified by histologic grade. (b). Propensity score-matched (PSM) comparison of low-grade (G1) vs intermediate-grade (G2) OCC. (c). PSM comparison of low-grade (G1) vs. high-grade (G3) OCC. (d). PSM comparison of intermediate-grade (G2) vs. high-grade (G3) OCC.

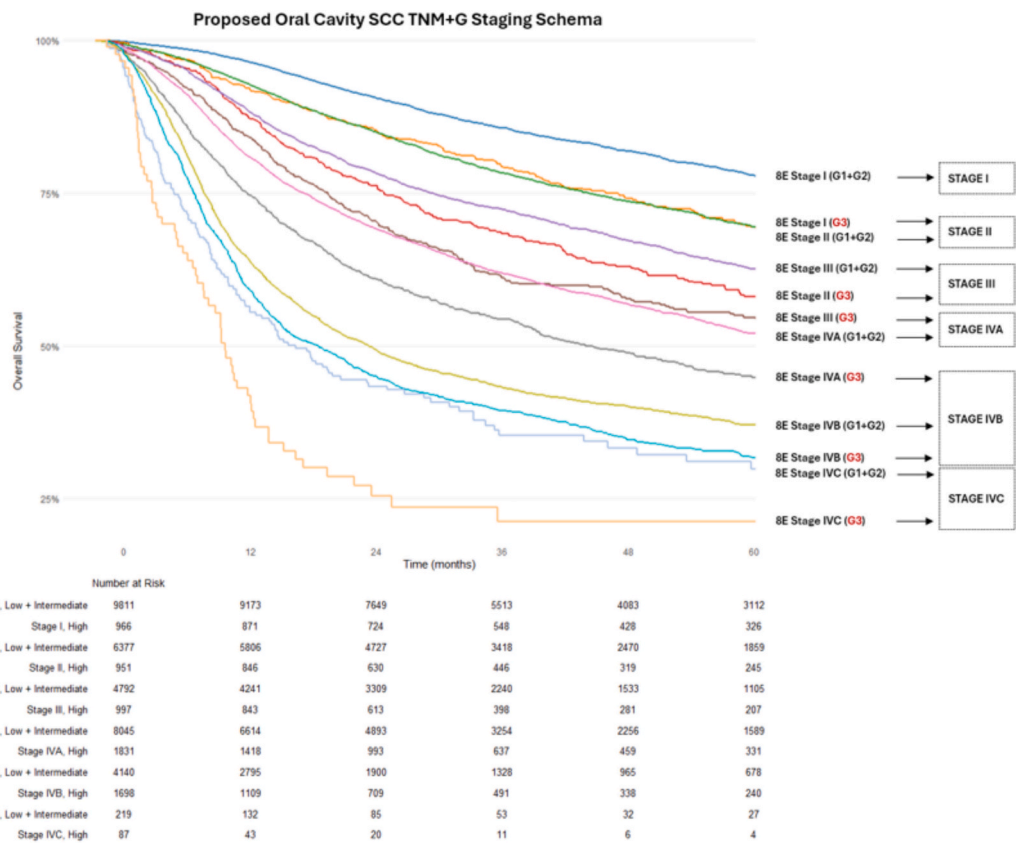


Fig. 3. Kaplan-Meier overall survival curves for oral cavity squamous cell carcinoma based upon branches determined by recursive partitioning analysis, focusing on T-classification, N-classification, M-classification, and histologic grade. Low and intermediate grade cases were collapsed into one category (G1 + G2), separate from high grade cases (G3). 8E refers to the American Joint Committee on Cancer (AJCC) 8th Edition staging system.

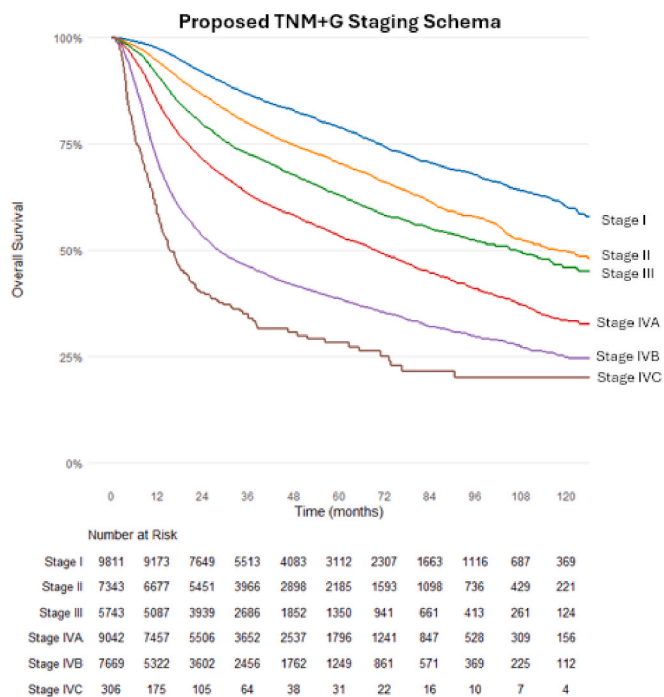


Fig. 4. Kaplan-Meier overall survival curves based on proposed TNM + G staging schema.

prognostic factors in all histologic subtypes ($p < 0.001$) (Fig. 2b-2d).
Proposed TNM + G staging system.
Given the fact that grade appeared to be a primary histologic driver

of prognosis, we investigated the utility of incorporating grade into oral cavity cancer staging. RPA was performed based on T-classification, N-classification, M-classification, and grade to generate an unbiased, agnostic subgroups (Fig. 3, Fig. S1). For simplicity, G1 and G2 were combined. Kaplan-Meier survival estimates of these empirically derived subgroups were visualized and then grouped (Fig. 4). The 5-yr OS was 79% [95% CI 78%-80%] (Stage I), 70% [95% CI 69%-72%] (Stage II), 63% [95% CI 61%-65%] (Stage III), 53% [95% CI 52%-55%] (Stage IVA), 39% [95% CI 37%-40%] (Stage IVB), and 28% [95% CI 23%-35%] (Stage IVC). Similar results were observed when using surgical patients only (Fig. S3).

It was noted that the AJCC8E schema remained largely intact (Stage I-IVC), but with high grade (G3) cases performing worse for each AJCC8E stage. As such, an augmented AJCC8E schema was produced, whereby G3 conferred upstaging by one category (e.g., high-grade AJCC8 Stage I to proposed Stage II, or high-grade AJCC8 Stage IVA to proposed Stage IVB) (Fig. 5). Stage IVB cases that were G3 were not upstaged, since Stage IVC was reserved for distant metastasis. Overall, 11.5% (4,745/36,623) of cases were upstaged in the new schema.

The optimism-corrected c-index showed improvement in predictive ability for the proposed system over AJCC8E (0.673 vs. 0.671) (higher score better). Similarly, the Brier score was also superior (0.174 vs. 0.175) (lower score better).

Discussion

In this large-scale analysis, we demonstrate that grade is a key predictor of mortality risk in oral cavity carcinoma. Even after adjusting for other covariates, the hazard of worsening grade approximates or exceeds that of conventional factors such as positive margins or extranodal extension (ENE) (Table 2). Propensity score matching shows stepwise

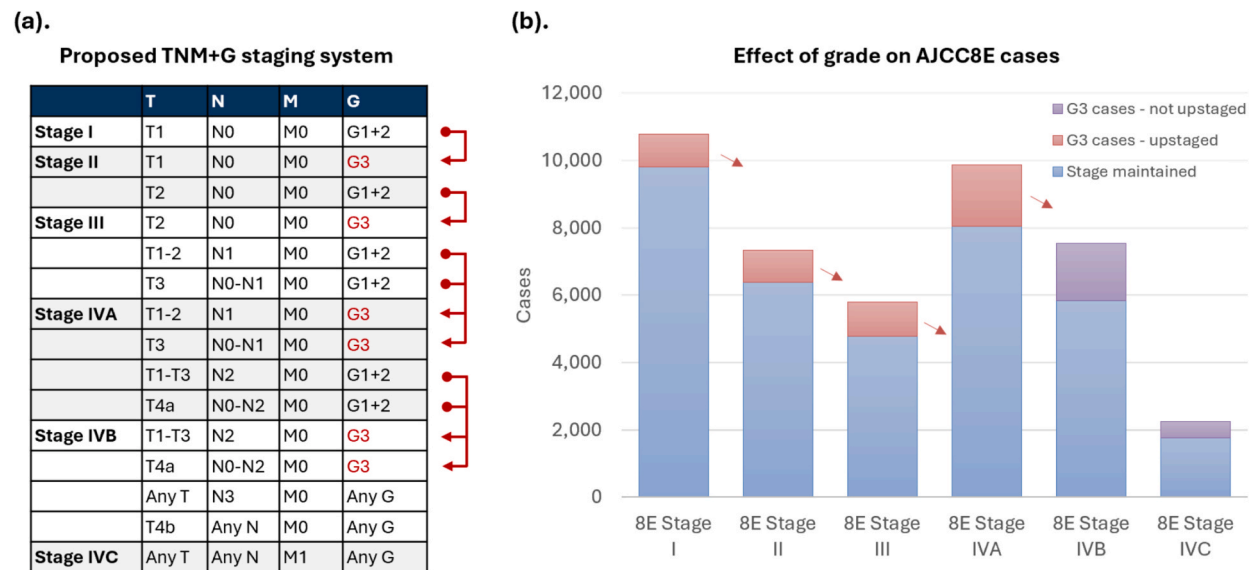


Fig. 5. Upstaging of AJCC8E oral cavity staging schema based on grade. High-grade (G3) cases are upstaged by 1 level with the exception of Stage IVB, as Stage IVC is reserved for M1 cases. (a). Proposed TNM + G staging system, with G3 upstaged cases illustrated by red arrows. (b). Impact of incorporating grade on AJCC8E, with red boxes illustrating cases that would be upstaged from AJCC8E stage grouping.

worsening of survival in parallel with worsening differentiation. Using recursive partitioning analysis, we moreover integrate grade into the existing TNM system to reflect its conspicuously adverse outcomes. This novel schema outperforms AJCC8E, highlights the highest-risk cases for upstaging, and pragmatically aligns oral cavity cancers with other disease sites. [1–4,27,28].

Our analysis challenges current views of grade in the context of oral cavity SCC. Like many head and neck malignancies, treatment morbidity remains problematic and improved prognostication is needed to best calibrate therapeutic intensity. High grade can act as a determinant that parses out cases with worse prognosis than would be expected based on traditional staging (11.5% in our cohort), which may warrant consideration of management changes including elective neck dissection, adjuvant radiation, chemotherapy, or closer surveillance. Our staging schema augments AJCC8E via an approach that is both meaningful and measurable, without adding undue complexity. [29] By staging criteria, it also exhibits robust separation of curves, predictive ability, hazard consistency, and balance between groups. [30].

The effect of grade has been decidedly mixed in the oral cavity literature, fueling controversy and limiting adoption. While some investigators have shown grade to impact nodal metastasis, perineural invasion, ENE, margin status, and overall survival [7–11], others have shown poor correlation with outcome. [12–16] More recent proposals have attempted to combine grade with related pathologic factors, such as tumor budding (presence of small clusters of < 5 tumor cells), worst pattern of invasion (WPOI), and tumor stromal ratio.[31–33] For now, the WHO, College of American Pathologists (CAP), and the International Collaboration on Cancer Reporting (ICCR) recommend recording grade based on a 4-tier system from Broders et al that relies on degree of keratinization, cytonuclear atypia, and mitotic features.[6,34–36] However, the WHO has nonetheless discounted its utility, cautioning that grade alone does not correlate well with prognosis.

Several factors could explain these conflicting results. The absence of effect could be due to the subjective nature in assigning grade, especially in small single-institution series. Its role as an independent prognostic factor could be blunted by related histologic factors such as tumor budding, TSR, PNI, LVI, and DOI. Conversely, many pathologists may inherently incorporate these factors into grade (rather than consider them as separate variables, a recent phenomenon) – this would artificially amplify its impact. Finally, grade’s magnitude of effect could be

too small for underpowered studies to detect, especially for advanced stage cases. This may explain why some studies only observe an impact of grade on early-stage oral cavity cancers. [7,16] In contrast, our analysis detects a clear signal of strong independent hazard.

The presumed subjectivity in assigning grade is also worthwhile to consider. It is plausible that the lack of uniformity in grading, changes in grading over time, and wide variance between institutions may weaken enthusiasm for adoption. However, our large-scale data suggests the opposite: by consolidating national-level data from a spectrum of institutional types, the results indicate that grade’s impact is generalizable and not dependent on specific criteria, specialized pathology expertise, or particular institutional workflows. The clear delineation of low, intermediate, and high grade cohorts despite heterogeneity in reporting across hundreds of institutions (n = 1,222) in fact underpins its importance. We hypothesize that with more uniform or standardized grading and centralized review, the observed effect of grade could be even stronger.

A number of caveats deserve mention. Registry data is retrospective and at risk of coding errors, though audits are routinely performed. NCDB tracks overall survival but lacks the nuance that may be conferred with disease-specific survival. Pathology did not undergo centralized review, which may mitigate risk of misdiagnosis and streamline borderline cases. Nonsurgical cases were included to increase generalizability (i.e., definitive radiation cases or distant metastasis cases), but had grade determined by biopsy confirmation rather than surgical resection. Though classic anatomic factors such as T-category, ENE, LVI, and margins were corrected for in MVA, we acknowledge the absence of salient factors such as perineural invasion and DOI. We furthermore could not compare related histologic determinants such as tumor budding or WPOI, which are not recorded in NCDB. Finally, histologic determinants such as grade may in the future give way to molecular ones that offer greater precision. Nonetheless, our analysis strongly supports grade to be an intuitive, independent prognostic factor that is a biomarker of biologic aggressiveness.

Altogether, we describe and quantify the impact of grade on oral cavity carcinoma mortality risk. While grade should not be considered in isolation, the same is true for any determinant, making staging a complex exercise to consolidate prognosis. Our proposed schema is intended to strengthen existing staging and provoke discussion on improving ways to convey mortality risk. As treatment paradigms evolve in head

and neck cancer [37], it is crucial to consider such factors to better enhance decision-making and ultimately patient outcomes.

CRedit authorship contribution statement

Allen S. Ho: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Michael Luu:** Writing – review & editing, Methodology, Formal analysis. **Daniel Manzoor:** Writing – review & editing, Conceptualization. **Horacio Maluf:** Writing – review & editing, Conceptualization. **Cedric Bailey:** Writing – review & editing, Conceptualization. **Bonnie Balzer:** Writing – review & editing, Data curation, Conceptualization. **Evan S. Walgama:** Writing – review & editing, Formal analysis, Conceptualization. **Julie K. Jang:** Writing – review & editing, Conceptualization. **Kevin C. Scher:** Writing – review & editing, Conceptualization. **Justin T. Moyers:** Writing – review & editing, Conceptualization. **Jon Mallen-St. Clair:** Writing – review & editing, Formal analysis. **Joel B. Epstein:** Writing – review & editing, Conceptualization. **Zachary S. Zumsteg:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.oraloncology.2025.107847>.

Data availability

Data will be made available on request.

References

- [1] Rakha EA, El-Sayed ME, Lee AH, et al. Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. *J Clin Oncol* 2008;26(19):3153–8.
- [2] Tanaka K, Ozaki T. New TNM classification (AJCC eighth edition) of bone and soft tissue sarcomas: JCOG Bone and Soft Tissue Tumor Study Group. *Japanese journal of clinical oncology*. 2019;49(2):103–107.
- [3] Epstein JI, Amin MB, Reuter VE, Humphrey PA. Contemporary Gleason Grading of Prostatic Carcinoma: An Update With Discussion on Practical Issues to Implement the 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. *The American journal of surgical pathology*. 2017;41(4):e1–e7.
- [4] Amin MB, Edge SB, Greene FL, et al. *AJCC cancer staging manual*. Vol 1024: Springer; 2017.
- [5] van Herpen C, Vander Poorten V, Skalova A, et al. Salivary gland cancer: ESMO–European Reference Network on Rare Adult Solid Cancers (EURACAN) clinical practice guideline for diagnosis, treatment and follow-up. *ESMO Open* 2022;7(6):100602.
- [6] Head and Neck Tumours. World Health Organization (WHO) Classification of Tumours. 5th Edition, Volume 9. 2024.
- [7] Thomas B, Stedman M, Davies L. Grade as a prognostic factor in oral squamous cell carcinoma: a population-based analysis of the data. *Laryngoscope* 2014;124(3):688–94.
- [8] Brandwein-Gensler M, Teixeira MS, Lewis CM, et al. Oral squamous cell carcinoma: histologic risk assessment, but not margin status, is strongly predictive of local disease-free and overall survival. *Am J Surg Pathol* 2005;29(2):167–78.
- [9] Piffko J, Bankfalvi A, Öfner D, et al. Prognostic value of histobiological factors (malignancy grading and AgNOR content) assessed at the invasive tumour front of oral squamous cell carcinomas. *Br J Cancer* 1997;75(10):1543–6.
- [10] Roland NJ, Caslin AW, Nash J, Stell PM. Value of grading squamous cell carcinoma of the head and neck. *Head Neck* 1992;14(3):224–9.
- [11] Anderson EM, Luu M, Balzer BL, et al. Variations in the association of grade with survival across the head and neck cancer landscape. *Head Neck* 2021;43(4):1105–15.
- [12] Silva LAB, Lopes MLDS, Sá MC, et al. Histopathologic grading and its relationship with outcome in oral tongue squamous cell carcinoma. *J Oral Pathol Med* 2021;50(2):183–90.
- [13] Xu B, Salama AM, Valero C, et al. The prognostic role of histologic grade, worst pattern of invasion, and tumor budding in early stage oral tongue squamous cell carcinoma: a comparative study. *Virchows Archiv : an international journal of pathology* 2021;479(3):597–606.
- [14] Almangush A, Mäkitie AA, Triantafyllou A, et al. Staging and grading of oral squamous cell carcinoma: an update. *Oral Oncol* 2020;107:104799.
- [15] Dik EA, Ipenburg NA, Kessler PA, van Es RJJ, Willems SM. The value of histological grading of biopsy and resection specimens in early stage oral squamous cell carcinomas. *Journal of cranio-maxillo-facial surgery : official publication of the European Association for Cranio-Maxillo-Facial Surgery* 2018;46(6):1001–6.
- [16] Xu QS, Wang C, Li B, et al. Prognostic value of pathologic grade for patients with oral squamous cell carcinoma. *Oral Dis* 2018;24(3):335–46.
- [17] Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 1994;81(3):515–26.
- [18] van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011;45(3):1–67.
- [19] Ho D, Imai K, King G, Stuart EA. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *J Stat Softw* 2011;42(8):1–28.
- [20] Hothorn T, Hornik K, Zeileis A. Unbiased Recursive Partitioning: a Conditional Inference Framework. *J Comput Graph Stat* 2006;15(3):651–74.
- [21] HaW S. Christian. On the Asymptotic Theory of Permutation Statistics 1999;8: 220–50.
- [22] Harrell FE. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Springer Cham; 2015.
- [23] Team RC. R: A Language and Environment for Statistical Computing. 2021; <https://www.R-project.org/>.
- [24] Wickham HAM, Bryan J, et al. Welcome to the Tidyverse. *Journal of Open Source Software* 2019;4(43):1686.
- [25] Harrell Jr FE. Rms: Regression Modeling Strategies. 2022; <https://CRAN.R-project.org/package=rms>.
- [26] Zeileis A, Hothorn T, Hornik K. Model-based Recursive Partitioning. *J Comput Graph Stat* 2008;17(2):492–514.
- [27] Ho AS, Luu M, Balzer BL, et al. Comparative impact of grade on mortality across salivary cancers: a novel, unifying staging system. *Head Neck* 2023;45(8):2028–39.
- [28] Ho AS, Luu M, Barrios L, et al. Prognostic Impact of Histologic Grade for Papillary Thyroid Carcinoma. *Ann Surg Oncol* 2021;28(3):1731–9.
- [29] Lydiatt WM. Cancer staging: the Goldilocks principle. *Cancer* 2024;130(3):342–3.
- [30] Groome PA, Schulze K, Boysen M, Hall SF, Mackillop WJ. A comparison of published head and neck stage groupings in carcinomas of the oral cavity. *Head Neck* 2001;23(8):613–24.
- [31] Arora A, Husain N, Bansal A, et al. Development of a New Outcome Prediction Model in Early-stage Squamous Cell Carcinoma of the Oral Cavity based on Histopathologic Parameters with Multivariate Analysis: the Aditi-Nuzhat Lymph-node Prediction score (ANLPS) system. *Am J Surg Pathol* 2017;41(7):950–60.
- [32] Elseragy A, Bello IO, Wahab A, et al. Emerging histopathologic markers in early-stage oral tongue cancer: a systematic review and meta-analysis. *Head Neck* 2022;44(6):1481–91.
- [33] Bastien AJ, Manzoor D, Maluf H, et al. A review of histopathologic assessment for head and neck oncologists. *Oral Oncol* 2025;165:107286.
- [34] Ac B. Carcinoma of the mouth: types and degree of malignancy. *Ann J Roentgenol Rad Ther Nucl Med* 1927;17:90–3.
- [35] Müller S, Boy SC, Day TA, et al. Data set for the reporting of oral cavity carcinomas: explanations and recommendations of the guidelines from the international collaboration of cancer reporting. *Arch Pathol Lab Med* 2019;143(4):439–46.
- [36] Seethala RRB, Faquin WC, et al. Protocol for the Examination of Specimens from patients with Cancers of the Oral Cavity. Version 4.2.0.0. College of American Pathologists 2023.
- [37] Uppaluri R, Haddad RI, Tao Y, et al. Neoadjuvant and Adjuvant Pembrolizumab in locally Advanced Head and Neck Cancer. *N Engl J Med* 2025;393(1):37–50.