

Original Investigation

Primary Tumor Staging for Oral Cancer and a Proposed Modification Incorporating Depth of Invasion

An International Multicenter Retrospective Study

The International Consortium for Outcome Research (ICOR) in Head and Neck Cancer

IMPORTANCE The current American Joint Committee on Cancer (AJCC) staging system for oral cancer demonstrates wide prognostic variability within each primary tumor stage and provides suboptimal staging and prognostic information for some patients.

OBJECTIVE To determine if a modified staging system for oral cancer that integrates depth of invasion (DOI) into the T categories improves prognostic performance compared with the current AJCC T staging.

DESIGN, SETTING, AND PARTICIPANTS Retrospective analysis of 3149 patients with oral squamous cell carcinoma treated with curative intent at 11 comprehensive cancer centers worldwide between 1990 and 2011 with surgery \pm adjuvant therapy, with a median follow-up of 40 months.

MAIN OUTCOMES AND MEASURES We assessed the impact of DOI on disease-specific and overall survival in multivariable Cox proportional hazard models and investigated for institutional heterogeneity using 2-stage random effects meta-analyses. Candidate staging systems were developed after identification of optimal DOI cutpoints within each AJCC T category using the Akaike information criterion (AIC) and likelihood ratio tests. Staging systems were evaluated using the Harrel concordance index (C-index), AIC, and visual inspection for stratification into distinct prognostic categories, with internal validation using bootstrapping techniques.

RESULTS The mean and median DOI were 12.9 mm and 10.0 mm, respectively. On multivariable analysis, DOI was significantly associated with disease-specific survival ($P < .001$), demonstrated no institutional prognostic heterogeneity ($I^2 = 6.3\%$; $P = .38$), and resulted in improved model fit compared with T category alone (lower AIC, $P < .001$). Optimal cutpoints of 5 mm in T1 and 10 mm in T2-4 category disease were used to develop a modified T staging system that was preferred to the AJCC system on the basis of lower AIC, visual inspection of Kaplan-Meier curves, and significant improvement in bootstrapped C-index.

CONCLUSIONS AND RELEVANCE We propose an improved oral cancer T staging system based on incorporation of DOI that should be considered in future versions of the AJCC staging system after external validation.

JAMA Otolaryngol Head Neck Surg. 2014;140(12):1138-1148. doi:10.1001/jamaoto.2014.1548
Published online July 30, 2014.

The authors and members of The International Consortium for Outcome Research (ICOR) in Head and Neck Cancer and their affiliations are listed at the end of the article.

Corresponding Author: Ardalan Ebrahimi, MBBS, FRACS, MPH, Sydney Head and Neck Cancer Institute, Royal Prince Alfred Hospital, Missenden Road, Camperdown, New South Wales 2050, Australia (ebrahim.ardalan@post.harvard.edu).

The first edition of the *Manual for Staging of Cancer* was published by the American Joint Committee on Cancer (AJCC) in 1977.¹ Since then, the primary tumor staging for oral squamous cell carcinoma (SCC) has remained unchanged, with the exception of refinements to the T4 classification. Although the simplicity of the current system promotes clinical utility, it is widely acknowledged that the prognostic performance is suboptimal in some patients with head and neck cancer.²⁻⁴ This may reflect the substantial changes in management of oral SCC in the last 4 decades, based on advances in cross-sectional imaging, the introduction of positron emission tomography, the widespread use of microvascular free flap reconstruction, and significant progress in adjuvant therapy protocols. In conjunction with this evolution in management paradigms, a vast body of literature has accumulated detailing important prognostic factors in oral cancer. One such factor is the depth of invasion (DOI) of the primary tumor, which is well established as an independent predictor of recurrence and survival⁵⁻¹⁷ and may provide information that should be integrated into the existing AJCC staging system to improve prognostic performance.

Depth of invasion or thickness is already a feature in the AJCC staging of other cancers such as melanoma, cutaneous SCC, and the uterine cervix.¹⁸ In addition, the degree of invasion based on anatomical layers, which can be considered analogous to tumor DOI, is central to staging of cancers of the esophagus, stomach, colon, and rectum.¹⁸ Although several groups have suggested how DOI might be incorporated into the staging of oral SCC,^{9,10} these proposals have not been adopted by the AJCC.

The primary aim of this study was to develop a modified staging system for oral SCC that integrates DOI into the T categories. Our secondary aims were 2-fold: we aimed (1) to determine if there is significant heterogeneity between institutions in terms of the prognostic impact of DOI in view of the variable definitions used by previous authors¹⁹ and (2) to compare the prognostic performance of our modified system with the current AJCC T categories as well as previously proposed modifications based on DOI or thickness.

Methods

Study Population

This international multicenter retrospective study included pooled individual patient data from 11 participating comprehensive cancer centers worldwide, treated between 1990 and 2011. Ethical approval was obtained from local institutional review board committees of the 11 participating centers. Patients with histologically confirmed oral SCC undergoing surgical resection of the primary tumor and neck dissection with curative intent were candidates for inclusion. We identified 3781 patients as candidates for inclusion in the study. We excluded cases if they had received neoadjuvant therapy (n = 22), were younger than 20 years (n = 5), experienced perioperative mortality (n = 16), or had inadequate information to determine primary DOI (n = 581), pathological T (pT) category (n = 0), or pathological N (pN) category (n = 8). The final study population consisted of 3149 patients (Table 1).

Table 1. Patient Clinicopathological Data

Variable	Patients, No. (%)
Age, y	
≤45	875 (27.8)
46-55	909 (28.9)
56-65	761 (24.2)
≥66	604 (19.2)
Sex	
Male	2074 (65.9)
Female	1075 (34.1)
Pathological T category	
T1	454 (14.4)
T2	955 (30.3)
T3	413 (13.1)
T4	1327 (42.1)
Pathological N category	
N0	1702 (54.1)
N1	488 (15.5)
N2a	54 (1.7)
N2b	728 (23.1)
N2c	169 (5.4)
N3	8 (0.3)
TNM stage	
I	350 (11.1)
II	568 (18.0)
III	457 (14.5)
IV	1774 (56.3)
Extracapsular nodal spread	
Absent	2192 (69.6)
Present	921 (29.2)
Unknown	36 (1.1)
Surgical excision margin	
Clear	2515 (79.9)
Close, <5 mm	429 (13.6)
Involved	202 (6.4)
Unknown	3 (0.1)
Adjuvant therapy	
No	884 (28.1)
Radiotherapy	1610 (51.1)
Chemoradiotherapy	500 (15.9)
Radiotherapy + cetuximab	153 (4.9)
Unknown	2 (0.1)
Decade of primary treatment	
1990-1999	615 (19.5)
2000-2011	2534 (80.5)

Histopathological Analysis

Although many authors use the terms *thickness* and *depth of invasion* synonymously, they are not the same, and a distinction should be made.^{19,20} Tumor thickness was defined according to the definition proposed by Moore et al,²⁰ which extends from the level of adjacent normal mucosa to the deepest point of tumor invasion. In contrast, DOI is considered to be the extent of invasion below the epithelial basement

membrane.¹⁹ Rather than restricting the analysis to a homogeneous group in this respect, we assumed some heterogeneity in view of the extended study period as well as the number of involved institutions and pathologists. Hence, we elected to determine if significant between-center heterogeneity exists in regard to the clinical impact of DOI, which may reflect center-specific definitions and measurement techniques.

Statistical Analysis

Statistical analysis was performed using Stata version 12.0 SE (StataCorp LP, College Station, Texas). All statistics were 2-sided, and $P < .05$ was considered statistically significant. The clinical end points of interest were overall survival (OS) and disease-specific survival (DSS). Overall survival was calculated from the date of surgery to the date of death or last follow-up visit. For DSS, patients who died from causes other than oral SCC were censored at the time of death.

We first sought to determine if DOI provides significant prognostic information beyond pT and pN categories in multivariable-adjusted models. Depth of invasion was analyzed as a continuous variable after logarithmic transformation because the distribution was right skewed. Multivariable analyses were performed using Cox proportional hazard models, stratified by study center and adjusted for pT category (T1, T2, T3, T4), pN category (N0, N1, N2a, N2b, N2c, N3), time period of primary treatment (1990-1999, 2000-2011), and adjuvant radiotherapy (minimally adjusted model). A sensitivity analysis was performed to determine whether additional adjustment for age, sex, extracapsular spread, and margin status affected the results (fully adjusted model).

Next, we used a 2-stage random effects modeling approach to investigate for the presence of heterogeneity between institutions in terms of the prognostic impact of DOI.²¹ This was performed for both minimally and fully adjusted models. At the first stage, the effect of DOI on DSS was determined for each center using adjusted Cox regression models. In the second stage of analysis, the center-specific estimates were introduced into the random effects model of DerSimonian and Laird, which allows for unexplained sources of heterogeneity.²² Heterogeneity across centers was assessed using the Cochran Q test ($P < .10$ was considered statistically significant, given that the test has limited power) and quantified using the I^2 measure (the percentage of total variation across centers attributable to heterogeneity rather than chance).²³

We then performed exploratory analyses to identify optimal cutpoints for DOI with an a priori decision to dichotomize the variable in the interests of developing a parsimonious staging system. This analysis excluded 112 patients with pT4 disease who had primary tumors less than 5 mm thick, assuming these were largely advanced primary tumors with bone invasion, in which the DOI obtained from the soft-tissue component may provide a misleading underestimate of thickness. Although these patients were excluded in the model development stage, they were included in later analyses evaluating and comparing the prognostic performance of staging systems. Optimal cutpoints were identified for each T category in multivariable models stratified by study center. The best-fitting model including DOI was identified using the

Akaike Information Criterion (AIC), which takes into account how well the model fits the data with penalties for model complexity.²⁴ This was then compared with the baseline model using a likelihood ratio test to determine whether the addition of DOI improved model fit.

On the basis of the results of exploratory analyses, 5 candidate primary tumor staging systems were evaluated using the AIC, the Harrel concordance index (C-index), and visual inspection for stratification into distinct prognostic categories. The C-index provides a measure of model discrimination, with a value of 1 indicating perfect prediction, while 0.5 is equivalent to the toss of a coin.²⁵ Discrimination is the ability of a model to distinguish individuals who experience the outcome from those who remain event free. For a prognostic model, the C-index is the chance that given 2 individuals, one who will develop the event of interest and one who will remain event free, the prediction model will assign a higher probability of an event to the former.

Because predictive models perform better in the data from which they were derived than on external data, we performed bootstrap resampling to obtain a bias-corrected C statistic, hence providing a more accurate estimate of model performance in other populations.²⁶ This method of internal validation was chosen over others, such as split-sample modeling, because bootstrap resampling techniques have been shown to produce stable and nearly unbiased estimates of predictive accuracy with better efficiency than other methods, particularly in large samples.²⁶ For this purpose, 200 random bootstrap samples with replacement, and of the same size as the original sample, were drawn from the original data set consisting of all patients. Sensitivity analyses were performed repeating the bootstrap taking into account clustering by study institution to ensure consistent results.²⁷ The most precise prognostic staging system was chosen for comparison with the current AJCC staging system, as well as previously proposed modifications of T staging based on DOI by Howaldt et al¹⁰ and Yuen et al.⁹ To establish whether observed differences in the C-index were statistically significant a bootstrapped 95% CI, and P value was generated for the difference in C-index using the SOMERSD package in Stata.

Results

Patient Demographics

The study population consisted of 3149 patients with oral SCC, treated at 11 participating tertiary cancer centers from 8 countries. There were 2074 men and 1075 women, with a median age of 53 years (range, 20-93 years) and median follow-up of 40 months. A summary of relevant demographic and clinicopathological details is provided in Table 1.

Primary DOI

The mean and median DOI were 12.9 mm and 10.0 mm, respectively. As given in Table 2, there was a statistically significant association between increasing DOI and more advanced disease, including higher pT category ($P < .001$) and pN category ($P < .001$), extracapsular spread ($P < .001$), and involved margins ($P < .001$). A wide range of median DOI was

noted based on pT categories: 5 mm in pT1; 9 mm in pT2; 13.5 mm in pT3; and 15 mm in pT4 tumors. A significant association was also noted with the use and type of adjuvant therapy ($P < .001$). Finally, the median DOI differed significantly by treating institution, ranging from 7 to 12 mm ($P < .001$).

Model Development

The 5-year DSS for the cohort was 76.0% (95% CI, 74.1%-77.8%), with 569 deaths due to oral SCC. Multivariable analysis showed a statistically significant association between the log-transformed DOI (hazard ratio [HR], 1.31; 95% CI, 1.13-1.51) ($P < .001$) and DSS, after adjusting for pT category, pN category, adjuvant therapy, and time of treatment (Table 3). A sensitivity analysis confirmed that the results were robust to additional adjustment for age, sex, extracapsular spread, and margin status (Table 3). Importantly, pT category remained an important predictor of DSS in the presence of information on DOI ($P < .001$), suggesting that these 2 variables provide complementary prognostic information. In support of this hypothesis, we compared the minimally adjusted model with and without DOI and found significant improvement in model fit with the inclusion of DOI ($P < .001$). Finally, a 2-stage random effects meta-analysis confirmed that there was no evidence of between-center heterogeneity in the prognostic impact of DOI in both minimally adjusted ($I^2 = 16.2\%$; $P = .30$) and fully adjusted models ($I^2 = 6.3\%$; $P = .38$).

We proceeded to identify optimal cutpoints for dichotomizing DOI separately for each T category, using adjusted Cox regression models stratified by study center. Based on the AIC, the optimal cutpoints were less than 5 mm vs 5 mm or greater in pT category I disease and less than 10 mm vs 10 mm or greater in pT category II, III, and IV disease. The inclusion of the category-specific DOI variable significantly improved model fit for pT1 ($P = .01$), pT2 ($P = .001$), and pT3 ($P = .004$) but not pT4 ($P = .11$). Figure 1 demonstrates Kaplan-Meier plots of cumulative risk of disease-specific failure according to DOI and pT category.

On the basis of the results of these exploratory analyses, 5 candidate staging systems were devised for assessment as described in Table 4 and shown in Figure 2. Model 1 was a simple model using DOI alone to define T1-3 categories and performed poorly compared with the other candidate models (Figure 2B). This was consistent with our earlier findings suggesting that AJCC T staging provides additional complementary information and cannot be replaced by DOI. Model 2 performed poorly at differentiating between T1 and T2 categories (Figure 2C). This resulted from the increased hazard associated with combining all thin (<5 mm) tumors smaller than 4 cm into T1. Models 3 and 5 were designed to further stratify early tumors; however, there was considerable overlap in 95% CIs for T1a, T1b, and T2 categories in these models as shown in Figure 2D and F. Of the candidate staging systems, model 4 (Figure 2E) was preferred based on a combination of the AIC, C-index, stratification of patients into distinct prognostic groups with fairly evenly separation, minimal overlap of 95% CIs, and parsimony.

As shown in Figure 2A, the current AJCC staging system performed poorly in regard to discrimination and stratifica-

Table 2. Summary Data for Primary Depth of Invasion

Variable	Patients, No. (%)	Depth of Invasion, mm	
		Median (IQR)	P Value
Overall	3149 (100)	10 (11)	
Study center			
Brescia, Italy	22 (0.7)	11.5 (6)	<.001
São Paulo, Brazil	236 (7.5)	12 (14.5)	
Petach Tikva, Israel	58 (1.8)	9 (10)	
SHNCI, Australia	200 (6.4)	10 (7.5)	
Camargo, São Paulo, Brazil	97 (3.1)	10 (11)	
Cologne, Germany	237 (7.5)	11.4 (7.9)	
MSKCC, United States	199 (6.3)	8 (9)	
Southern Illinois, United States	26 (0.8)	7 (7)	
CGMH, Taoyuan, Taiwan	1219 (38.7)	10 (10)	
Tel Aviv, Israel	80 (2.5)	8.5 (9.5)	
Tata Memorial Hospital, India	775 (24.6)	12 (12)	
Pathological T category			
T1	454 (14.4)	5 (4.1)	<.001
T2	955 (30.3)	9 (6)	
T3	413 (13.1)	13.5 (9.4)	
T4	1327 (42.1)	15 (12)	
Pathological N category			
N0	1702 (54.1)	10 (10)	<.001
N1	488 (15.5)	10 (9.5)	
N2a	54 (1.7)	14.6 (10)	
N2b	728 (23.1)	12 (12)	
N2c	169 (5.4)	16 (15)	
N3	8 (0.3)	21.4 (18.8)	
Extracapsular nodal spread			
Absent	2192 (70.4)	10 (9)	<.001
Present	921 (29.6)	13 (12)	
Surgical excision margin			
Clear	2515 (80.0)	10 (10)	<.001
Close, <5 mm	429 (13.6)	11 (11.2)	
Involved	202 (6.4)	14.7 (11.5)	
Adjuvant therapy			
No	884 (28.1)	7 (7)	<.001
Radiotherapy	1610 (51.2)	12 (12)	
Chemoradiotherapy	500 (15.9)	12.4 (11.2)	
Radiotherapy + cetuximab	153 (4.9)	14.5 (6.6)	

Abbreviations: CGMH, Chang Gung Memorial Hospital; IQR, interquartile range; MSKCC, Memorial Sloan-Kettering Cancer Center; SHNCI, Sydney Head & Neck Cancer Institute.

tion of T3 and T4 disease. On the basis of the lower AIC, visual inspection of Kaplan-Meier curves, and statistically significant improvement in bootstrapped comparisons of the C-index ($P = .007$), model 4 outperformed the AJCC stage. The 5-year disease-specific mortality rate based on model 4 was 4% for T1, 13% for T2, 24% for T3, and 37% for T4 disease. In contrast, the 5-year disease-specific mortality based on the AJCC was 8% for T1, 18% for T2, 35% for T3, and 34% for T4 disease.

Table 3. Multivariable Analyses for Clinicopathological Factors Associated With Disease-Specific Survival^a

Factor	Minimally Adjusted Model		Fully Adjusted Model	
	HR (95% CI)	P Value	HR (95% CI)	P Value
Primary depth of invasion ^b	1.31 (1.13-1.51)	<.001	1.31 (1.13-1.52)	<.001
Age, 10-y increment	NA	NA	1.05 (0.97-1.13)	.23
Sex				
Male	NA	NA	1 [Reference]	
Female	NA	NA	0.94 (0.75-1.19)	.63
Pathological T category				
T1	1 [Reference]		1 [Reference]	
T2	1.95 (1.302-92)	.001	1.89 (1.27-2.83)	.002
T3	2.94 (1.91-4.52)	<.001	2.79 (1.81-4.30)	<.001
T4	3.82 (2.50-5.84)	<.001	3.24 (2.11-4.97)	<.001
Pathological N category				
N0	1 [Reference]		1 [Reference]	
N1	2.26 (1.73-2.95)	<.001	1.95 (1.47-2.59)	<.001
N2a	1.44 (0.72-2.90)	.30	1.09 (0.54-2.21)	.82
N2b	4.11 (3.25-5.21)	<.001	2.93 (2.21-3.89)	<.001
N2c	4.30 (3.12-5.91)	<.001	2.81 (1.96-4.03)	<.001
N3	6.71 (2.68-16.82)	<.001	4.00 (1.54-10.35)	.004
ECS				
Absent	NA	NA	1 [Reference]	
Present	NA	NA	1.72 (1.37-2.15)	<.001
Excision margin				
Clear	NA	NA	1 [Reference]	
Close, <5 mm	NA	NA	1.32 (1.01-1.71)	.04
Involved	NA	NA	1.98 (1.52-2.59)	<.001
Adjuvant therapy				
None	1 [Reference]		1 [Reference]	
Adjuvant RT	0.97 (0.74-1.27)	.83	0.96 (0.74-1.26)	.78
Decade of primary treatment				
1990-1999	1 [Reference]		1 [Reference]	
2000-2011	0.69 (0.55-0.87)	.002	0.68 (0.54-0.87)	.002

Abbreviations: ECS, extracapsular spread; HR, hazard ratio; NA, not applicable; RT, radiotherapy.

^a Cox proportional hazards regression models stratified by study center. Note that the minimally adjusted model represents the main analysis model. The fully adjusted model provided a sensitivity analysis demonstrating that the results for depth of invasion were stable with additional adjustment for age, sex, ECS, and margin status.

^b Depth of invasion was analyzed as a continuous variable after logarithmic transformation.

Next, we compared model 4 with the T staging systems proposed by Howaldt et al¹⁰ (Howaldt stage) and Yuen et al⁹ (Yuen stage). Model 4 was more informative based on the lower AIC and stratification into more distinct prognostic categories. In addition, model 4 had the highest C-index, with the difference statistically significant for both Howaldt¹⁰ ($P = .01$) and Yuen⁹ ($P < .001$) stages on bootstrapped comparison. As shown in Figure 3, our proposed model was also superior when the end point was OS. Again, the bootstrapped C-index was significantly higher than for AJCC stage ($P < .001$), Howaldt stage ($P < .001$), and Yuen stage ($P < .001$).

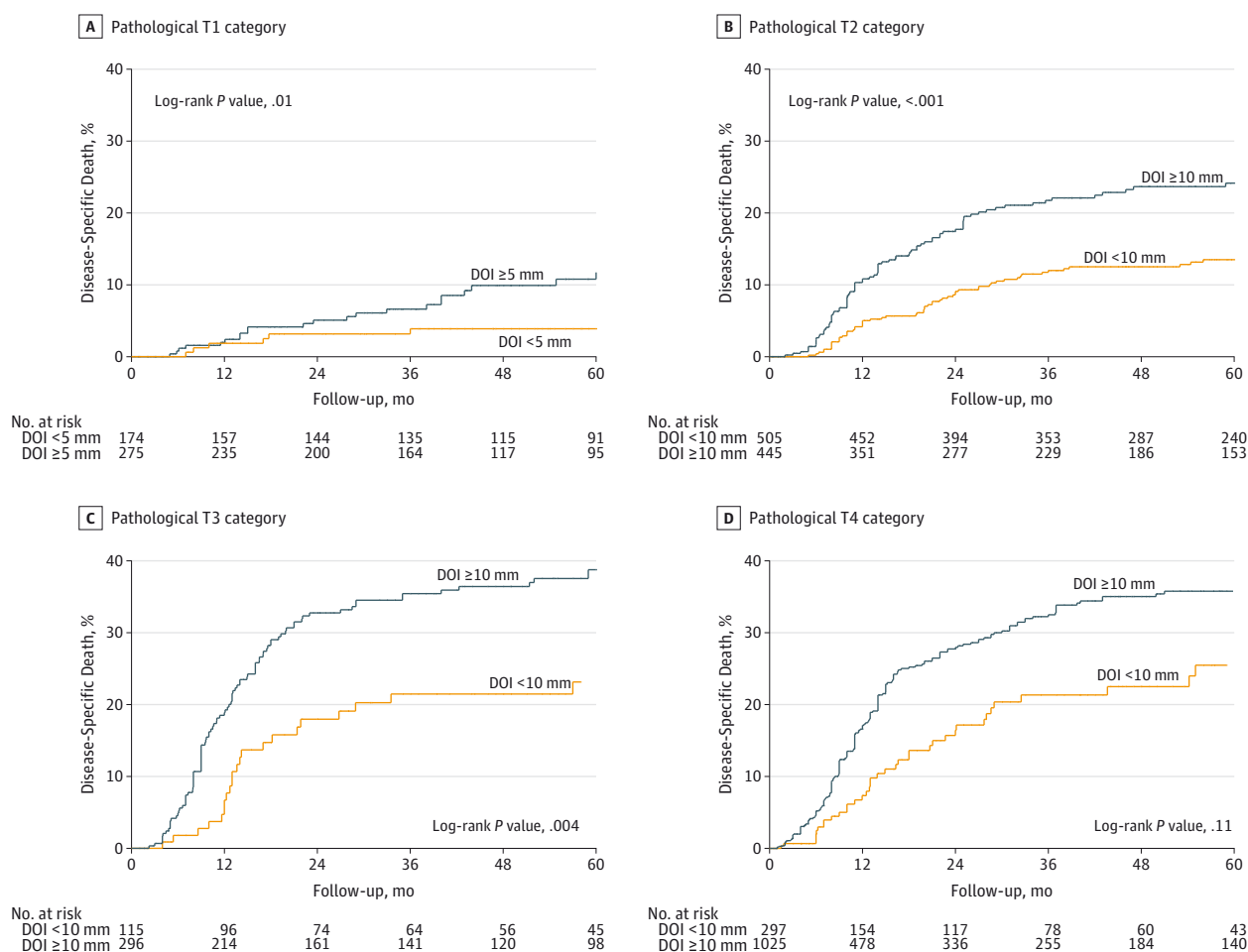
Table 5 summarizes changes in the T category comparing the AJCC system vs our proposed system based on model 4. The T category was unchanged in 58.1% of patients, upstaged in 32.5%, and downstaged in 9.4%. Specifically, 62.8% of patients with T1 disease were upstaged to T2 based on a DOI of 5 mm or greater, 46.8% of patients with T2 disease were upstaged to T3 based on a DOI of 10 mm or greater, and 71.9% of patients with T3 disease were upstaged to T4 based on a DOI of 10 mm or greater. Finally, 22.4% of patients with T4 disease were downstaged based on a DOI less than 10 mm.

Discussion

The AJCC primary tumor staging for oral SCC has remained essentially unchanged since it was first published in 1977.¹ Although the simplicity and consistency across subsites of the existing AJCC system promotes clinical utility, the prognostic performance is suboptimal in many patients, which may reflect the substantial changes in management of oral SCC in the last 4 decades. This has led to numerous proposed modifications of the staging system based on a variety of clinicopathological factors including DOI,^{9,10} which is now well established as an independent predictor of recurrence and survival in oral SCC.⁵⁻¹⁵ In the present study, when patients were restaged with a modification of the current AJCC staging system that incorporates DOI, we observed improved discrimination between T categories with respect to both DSS and OS.

In agreement with other studies, we found that DOI is an independent predictor of DSS in multivariable analyses.⁵⁻¹⁵ Importantly, the association between pathological T cat-

Figure 1. Kaplan-Meier Plots With Risk Tables Demonstrating the Association Between DOI and Disease-Specific Survival Within Each Pathological T Category



A, T1 category; B, T2 category; C, T3 category; and D, T4 category. Log-rank test is provided for each comparison. DOI indicates depth of invasion.

category and DSS remained significant after controlling for DOI, suggesting that they provide complementary information. On the basis of the results from exploratory analyses to identify optimal cutpoints for DOI, we generated 5 candidate staging systems that modify the existing AJCC T category by incorporating DOI. Because of the lower AIC, higher C-index, excellent stratification into distinct prognostic categories with regard to DSS, and relative simplicity, we selected model 4 (Table 4). Comparisons with the current AJCC primary tumor staging and the proposed staging systems of Howaldt et al¹⁰ and Yuen et al⁹ suggested that our model provides improved prognostic stratification and discrimination.

The tremendous variability and frequent lack of clarity in the literature regarding the exact definitions of tumor thickness vs DOI is an important issue.^{19,28} However, the difference between the line of the epithelial basement membrane (used to measure DOI) and the surrounding healthy mucosa (most commonly used to measure tumor thickness) is often more theoretical than practical because

of the limited thickness of healthy epithelium.¹⁹ Hence, rather than restricting the analysis, we assumed that heterogeneity between centers was unavoidable in view of the extended period of the study as well as the number of involved institutions and pathologists. We hypothesized that these differences would not translate into clinically relevant differences in the prognostic impact of DOI or tumor thickness. Our hypothesis was confirmed using a 2-staged random effects meta-analysis demonstrating minimal between-center heterogeneity in this respect. This, along with the large sample size of individual patient data from 11 international cancer centers, demonstrates that the proposed staging system should be generalizable to clinical practice irrespective of institutional variability.

However, if DOI is to be introduced into the staging system, it is important to reach consensus among institutions regarding a precise definition and measurement technique because the difference may be important in selected circumstances and individual patients. For example, if thickness is measured from the tumor surface, rather than adja-

Table 4. AJCC T Categories and Candidate Models for Primary Tumor Staging of Oral Squamous Cell Carcinoma^a

T Staging System	Description
AJCC T categories	
T1	Tumor ≤2 cm in greatest dimension
T2	Tumor >2 cm but ≤4 cm in greatest dimension
T3	Tumor >4 cm in greatest dimension
T4a	Tumor invades through cortical bone, deep extrinsic tongue musculature, maxillary sinus, or facial skin
T4b	Tumor invades masticator space, pterygoid plates, skull base, or encases internal carotid artery
Model 1	
T1	Maximum DOI <5 mm
T2	Maximum DOI ≥5 mm but <10 mm
T3	Maximum DOI ≥10 mm
T4	AJCC T4, any DOI
Model 2	
T1	AJCC T1-2, maximum DOI <5 mm
T2	AJCC T1, maximum DOI ≥5 mm or AJCC T2, maximum DOI ≥5 mm but <10 mm
T3	AJCC T2, maximum DOI ≥10 mm or AJCC T3-4, maximum DOI <10 mm
T4	AJCC T3-4, maximum DOI ≥10 mm
Model 3	
T1a	AJCC T1, maximum DOI <5 mm
T1b	AJCC T1, maximum DOI ≥5 mm
T2	AJCC T2, maximum DOI <10 mm
T3	AJCC T2, maximum DOI ≥10 mm or AJCC T3-4, maximum DOI <10 mm
T4	AJCC T3-4, maximum DOI ≥10 mm
Model 4	
T1	AJCC T1, maximum DOI <5 mm
T2	AJCC T1, maximum DOI ≥5 mm or AJCC T2, maximum DOI <10 mm
T3	AJCC T2, maximum DOI ≥10 mm or AJCC T3-4, maximum DOI <10 mm
T4	AJCC T3-4, maximum DOI ≥10 mm
Model 5	
T1a	AJCC T1, maximum DOI <5 mm
T1b	AJCC T1, maximum DOI ≥5 mm
T2	AJCC T2, maximum DOI <10 mm
T3	AJCC T2, maximum DOI ≥10 mm or AJCC T3, maximum DOI <10 mm
T4	AJCC T3, maximum DOI ≥10 mm or AJCC T4
T stages proposed by Howaldt et al ¹⁰	
T1	Tumor ≤20 mm in greatest dimension, DOI ≤5 mm
T2	Tumor ≤20 mm in greatest dimension, DOI >5 mm but ≤20 mm or Tumor >20 mm in greatest dimension, DOI ≤5 mm
T3	Tumor ≤20 mm in greatest dimension, DOI >20 mm or tumor >20 mm but ≤40 mm in greatest dimension, DOI >5 mm but ≤20 mm or tumor >40 mm in greatest dimension, DOI >5 mm but ≤10 mm
T4	Tumor >20 mm in greatest dimension, DOI >20 mm or tumor >40 mm in greatest dimension, DOI >10 mm
T stages proposed by Yuen et al ⁹	
T1	DOI ≤3 mm
T2	DOI >3 mm and ≤9 mm
T3	DOI >9 mm
T4	AJCC T4

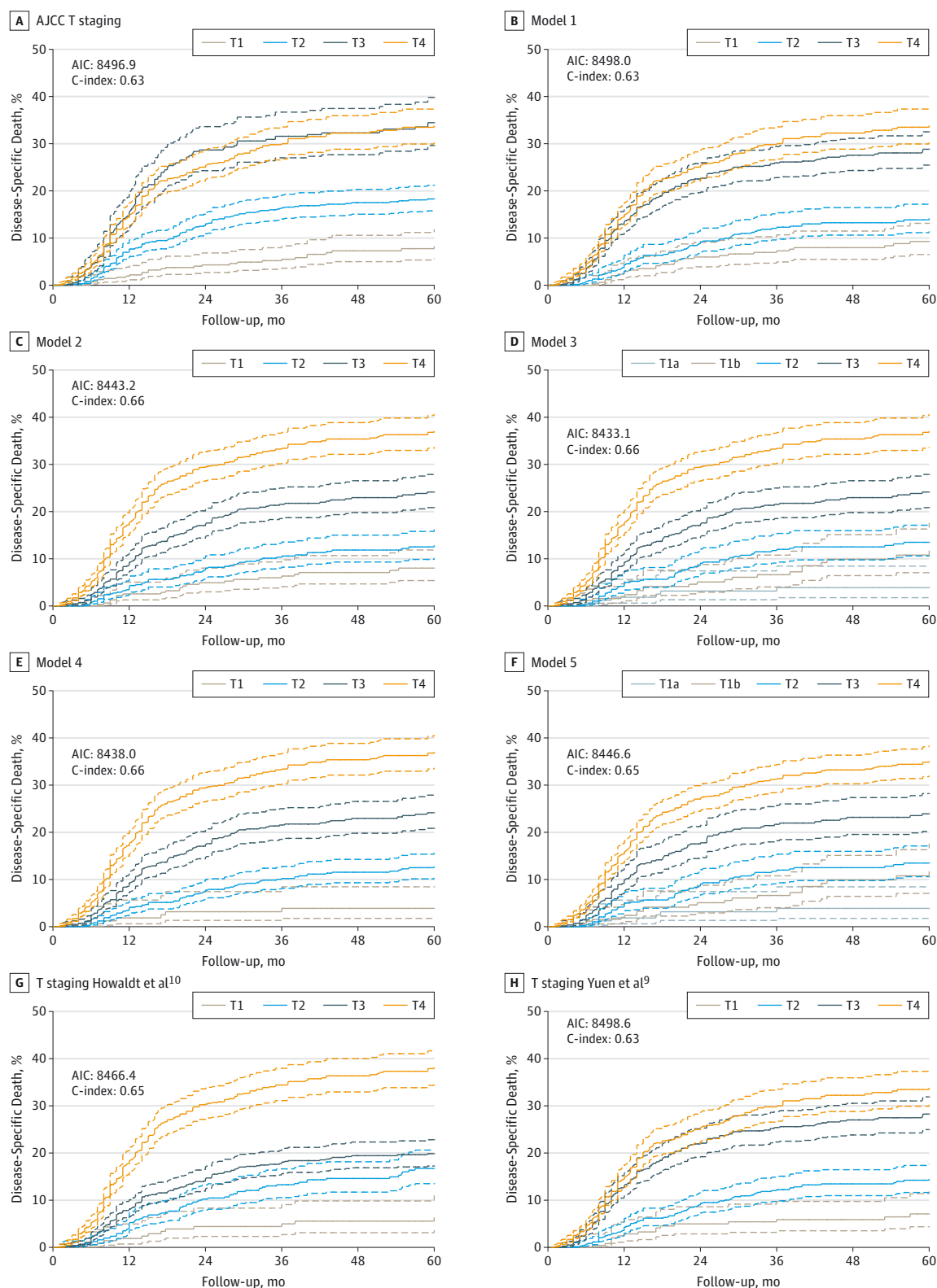
Abbreviations: AJCC, American Joint Committee on Cancer; DOI, depth of invasion.

^a Adapted from *AJCC Cancer Staging Manual* (7th ed).¹⁸ Note the designations Tx (primary tumor cannot be assessed), T0 (no evidence of primary tumor), and Tis (carcinoma in situ) are not included for brevity.

cent normal mucosal level as suggested by Moore et al,²⁰ then the thickness for large exophytic tumors may differ substantially compared with DOI, depending on technique.¹⁹ Similarly, measurement from the base of

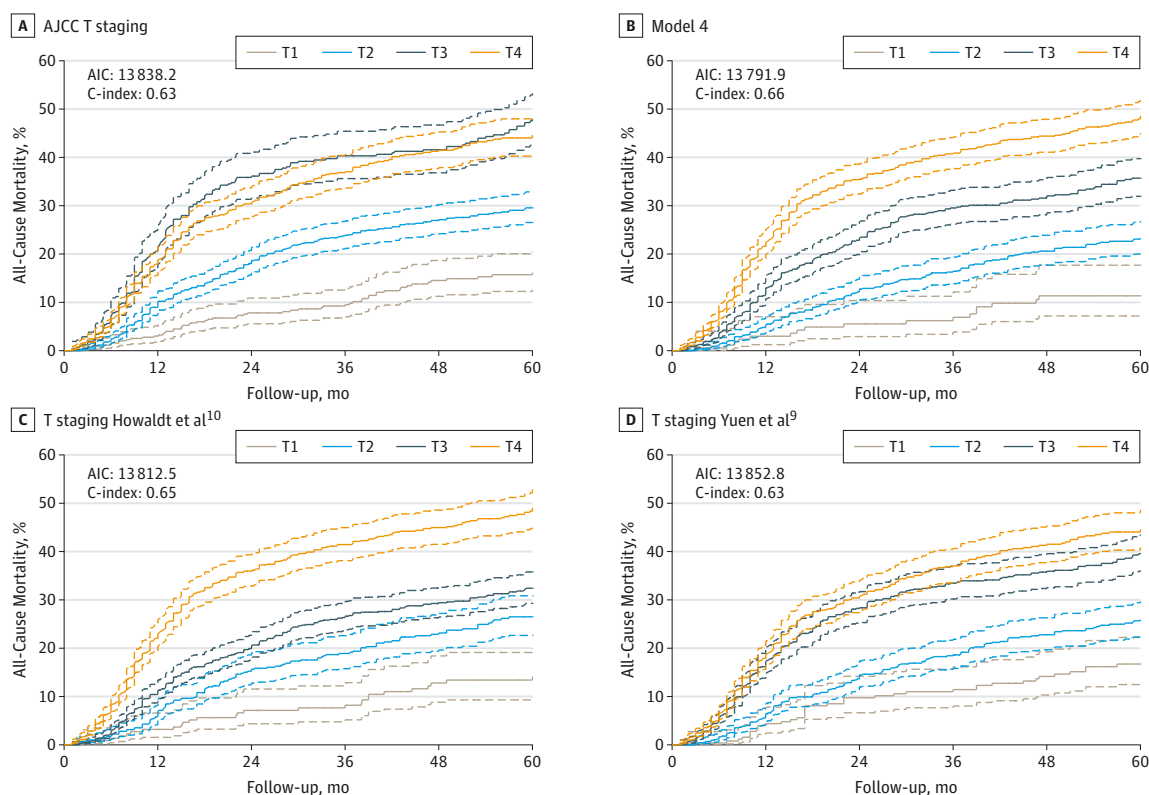
deeply ulcerative lesions may underestimate DOI.¹⁹ While an argument can be made for including other histopathological features such as perineural or lymphovascular invasion in a modified staging system, there is considerable

Figure 2. Kaplan-Meier Plots With 95% CIs Demonstrating the Association Between Different Staging Systems Describing the Primary Tumor Characteristics and Disease-Specific Survival



A, Current American Joint Committee on Cancer (AJCC) T staging; B, model 1; C, model 2; D, model 3; E, model 4; F, model 5; G, T stages proposed by Howaldt et al¹⁰; and H, T stages proposed by Yuen et al.⁹ Dashed lines indicate 95% CI. For details of candidate staging models please refer to Table 4.

Figure 3. Kaplan-Meier Plots With 95% CIs Demonstrating the Association Between Overall Survival and T Staging Systems



A, American Joint Committee on Cancer (AJCC) T categories; B, our proposed T categories based on model 4; C, T stages proposed by Howaldt et al¹⁰; and D, T stages proposed by Yuen et al.⁹ Dashed lines indicate 95% CI.

Table 5. T Category Changes From AJCC to Proposed System

AJCC T Category	T Category in Proposed System (Model 4), No. of Patients			
	T1	T2	T3	T4
T1	175	279	0	0
T2	0	508	447	0
T3	0	0	116	297
T4	0	0	297	1030

Abbreviation: AJCC, American Joint Committee on Cancer.

correlation between DOI and other adverse features, making their inclusion overly complicated without introducing sufficient additional prognostic information. Furthermore, thickness can be assessed clinically with the use imaging modalities such as ultrasonography and magnetic resonance imaging.²⁹⁻³¹

Another issue that requires further study is the use of DOI information for staging in tumors of the alveolar ridge and hard palate with underlying bone. We were unable to study this further owing to a lack of data on bone invasion and hence an inability to determine which patients were staged pT4 on this basis vs other features such as invasion of extrinsic tongue musculature or facial skin. There is some evidence to suggest that the prognosis of T4 tumors with bone invasion varies substan-

tially according to tumor size, and small primary tumors with bone invasion are associated with a relatively favorable prognosis that does not warrant T4 classification.^{32,33} Further study is required to confirm whether incorporation of DOI data is necessary for tumors in these subsites, particularly after reclassification of tumors with bone invasion based on consideration of tumor size.

This study has several limitations. First, it was performed using a combination of prospectively and retrospectively collected data, and treatment was not assigned in a randomized fashion. Second, despite internal validation by bootstrapping, validation using an external cohort remains a critical step before considering implementation in clinical practice. Third, there is likely to be variability among institutions over the extended period of the study in terms of pathology protocols used to assess DOI and interobserver variability among pathologists. Fourth, we lacked tumor subsite data and cannot exclude the possibility that critical values of thickness may differ between subsites.³⁴ Finally, our study population was limited to patients undergoing neck dissection, and therefore patients undergoing resection of the oral cavity primary tumor alone could not be assessed. However, the majority of these cases will be thin (<5 mm) T1 primary tumors, and²⁸ we anticipate that their inclusion will improve our proposed model's performance.

Conclusions

Our results show that DOI is an independent predictor of DSS in oral SCC and provides complementary prognostic information to the AJCC T category. We propose a modification that incorporates DOI to improve discrimination among patient sub-

groups with respect to DSS compared with the current AJCC staging system. This staging system can be easily implemented in clinical practice after external validation. However, for DOI to be integrated, the definition and methods of pathological assessment should also be described in the staging manual to ensure reproducibility and facilitate accurate comparisons between institutions.

ARTICLE INFORMATION

Submitted for Publication: May 4, 2014; final revision received June 30, 2014; accepted July 7, 2014.

Published Online: July 30, 2014.
doi:10.1001/jamaoto.2014.1548.

Authors and Members of The International Consortium for Outcome Research (ICOR) in Head and Neck Cancer: Ardalan Ebrahimi, MBBS, MPH, FRACS; Ziv Gil, MD, PhD; Moran Amit, MD, MSc; Tzu-Chen Yen, MD; Chun-Ta Liao, MD; Pankaj Chaturvedi, MS, DNB, FICS, MNAMS; Jai Prakash Agarwal, MD, DMRT; Luiz P. Kowalski, PhD; Matthias Kreppel, MD; Claudio R. Cernea, PhD; Jose Brandao, MD; Gideon Bachar, MD; Andrea Bolzoni Villaret, MD; Dan Fliss, MD; Eran Fridman, MD; K. Thomas Robbins, MD; Jatin P. Shah, MD; Snehal G. Patel, MD; Jonathan R. Clark, FRACS.

Affiliations of Authors and Members of The International Consortium for Outcome Research (ICOR) in Head and Neck Cancer: Sydney Head and Neck Cancer Institute, Royal Prince Alfred Hospital, Sydney, New South Wales, Australia (Ebrahimi, Clark); Department of Head and Neck Surgery, Liverpool Hospital, Sydney, New South Wales, Australia (Ebrahimi, Clark); Australian School of Advanced Medicine, Macquarie University, Sydney, New South Wales, Australia (Ebrahimi); South Western Sydney Clinical School, University of New South Wales, Sydney, New South Wales, Australia (Ebrahimi); The Laboratory for Applied Cancer Research, Haifa, Israel (Gil, Amit); Department of Otolaryngology Rambam Medical Center, Rappaport School of Medicine, the Technion, Israel Institute of Technology, Haifa, Israel (Gil, Amit); Chang Gung Memorial Hospital, Taoyuan, Taiwan (Yen, Liao); Tata Memorial Hospital, Mumbai, India (Chaturvedi, Agarwal); Hospital A. C. Camargo, São Paulo, Brazil (Kowalski); Department of Oral and Cranio-Maxillo and Facial Plastic Surgery University of Cologne, Cologne, Germany (Kreppel); Department of Head and Neck Surgery, University of São Paulo Medical School, São Paulo, Brazil (Cernea, Brandao); Department of Otolaryngology Head and Neck Surgery, Rabin Medical Center, Petach Tikva, Israel (Bachar); Ear, Nose, and Throat Department, University of Brescia, Brescia, Italy (Bolzoni Villaret); Department of Otolaryngology Head and Neck Surgery, Tel Aviv Medical Center, Tel Aviv, Israel (Fliss, Fridman); Southern Illinois University School of Medicine, Springfield, Illinois (Robbins); Head and Neck Surgery Service, Memorial Sloan Kettering Cancer Center, New York, New York (Shah, Patel).

Author Contributions: Drs Ebrahimi and Clark had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Ebrahimi, Kreppel, Bolzoni Villaret, Fliss, Shah, Patel.

Acquisition, analysis, or interpretation of data: Ebrahimi, Gil, Amit, Yen, Liao, Chaturvedi, Agarwal, Kowalski, Kreppel, Cernea, Brandao, Bachar, Fridman, Robbins, Clark.
Drafting of the manuscript: Ebrahimi, Kreppel, Bachar, Fliss, Robbins.
Critical revision of the manuscript for important intellectual content: Ebrahimi, Gil, Amit, Yen, Liao, Chaturvedi, Agarwal, Kowalski, Cernea, Brandao, Bolzoni Villaret, Fliss, Fridman, Shah, Patel, Clark.
Statistical analysis: Ebrahimi, Gil, Amit, Yen, Liao, Kreppel.
Administrative, technical, or material support: Ebrahimi, Yen, Liao, Brandao, Bachar, Fliss, Patel.
Study supervision: Chaturvedi, Kreppel, Fliss.

Conflict of Interest Disclosures: None reported.

Previous Presentation: This study was presented at the Fifth World Congress of the International Federation of Head and Neck Oncologic Societies and the Annual Meeting of the American Head & Neck Society; July 30, 2014; New York, New York.

REFERENCES

1. American Joint Committee for Cancer Staging and End-Results Reporting. *Manual for Staging of Cancer*. Chicago, IL: American Joint Committee; 1977.
2. 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):613-624.
3. Hall SF, Groome PA, Irish J, O'Sullivan B. TNM-based stage groupings in head and neck cancer: application in cancer of the hypopharynx. *Head Neck*. 2009;31(1):1-8.
4. Kreppel M, Drebber U, Rothamel D, et al. Prognostic impact of different TNM-based stage groupings for oral squamous cell carcinoma. *Head Neck*. 2011;33(10):1467-1475.
5. O'Brien CJ, Lauer CS, Fredricks S, et al. Tumor thickness influences prognosis of T1 and T2 oral cavity cancer—but what thickness? *Head Neck*. 2003;25(11):937-945.
6. Ganly I, Goldstein D, Carlson DL, et al. Long-term regional control and survival in patients with “low-risk,” early stage oral tongue cancer managed by partial glossectomy and neck dissection without postoperative radiation: the importance of tumor thickness. *Cancer*. 2013;119(6):1168-1176.
7. Ling W, Mijiti A, Moming A. Survival pattern and prognostic factors of patients with squamous cell carcinoma of the tongue: a retrospective analysis of 210 cases. *J Oral Maxillofac Surg*. 2013;71(4):775-785.
8. Tan WJ, Chia CS, Tan HK, Soo KC, Iyer NG. Prognostic significance of invasion depth in oral tongue squamous cell carcinoma. *ORL J Otorhinolaryngol Relat Spec*. 2012;74(5):264-270.
9. Yuen AP, Lam KY, Wei WI, et al. A comparison of the prognostic significance of tumor diameter, length, width, thickness, area, volume, and clinicopathological features of oral tongue carcinoma. *Am J Surg*. 2000;180(2):139-143.
10. Howaldt HP, Kainz M, Euler B, Vorst H. Proposal for modification of the TNM staging classification for cancer of the oral cavity: DOSAK. *J Craniomaxillofac Surg*. 1999;27(5):275-288.
11. Al-Rajhi N, Khafaga Y, El-Husseiny J, et al. Early stage carcinoma of oral tongue: prognostic factors for local control and survival. *Oral Oncol*. 2000;36(6):508-514.
12. Mishra RC, Parida G, Mishra TK, Mohanty S. Tumour thickness and relationship to locoregional failure in cancer of the buccal mucosa. *Eur J Surg Oncol*. 1999;25(2):186-189.
13. Gonzalez-Moles MA, Esteban F, Rodriguez-Archilla A, Ruiz-Avila I, Gonzalez-Moles S. Importance of tumour thickness measurement in prognosis of tongue cancer. *Oral Oncol*. 2002;38(4):394-397.
14. Urist MM, O'Brien CJ, Soong SJ, Visscher DW, Maddox WA. Squamous cell carcinoma of the buccal mucosa: analysis of prognostic factors. *Am J Surg*. 1987;154(4):411-414.
15. Clark JR, Naranjo N, Franklin JH, de Almeida J, Gullane PJ. Established prognostic variables in NO oral carcinoma. *Otolaryngol Head Neck Surg*. 2006;135(5):748-753.
16. Kang CJ, Lin CY, Wang HM, et al. The number of pathologically positive lymph nodes and pathological tumor depth predicts prognosis in patients with poorly differentiated squamous cell carcinoma of the oral cavity. *Int J Radiat Oncol Biol Phys*. 2011;81(4):e223-e230.
17. Liao CT, Lin CY, Fan KH, et al. Identification of a high-risk group among patients with oral cavity squamous cell carcinoma and pT1-2N0 disease. *Int J Radiat Oncol Biol Phys*. 2012;82(1):284-290.
18. Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A, eds. *AJCC Cancer Staging Manual*. 7th ed. New York, NY: Springer; 2010.
19. Pentenero M, Gandolfo S, Carrozzo M. Importance of tumor thickness and depth of invasion in nodal involvement and prognosis of oral squamous cell carcinoma: a review of the literature. *Head Neck*. 2005;27(12):1080-1091.
20. Moore C, Kuhns JG, Greenberg RA. Thickness as prognostic aid in upper aerodigestive tract cancer. *Arch Surg*. 1986;121(12):1410-1414.
21. Stukel TA, Demidenko E, Dykes J, Karagas MR. Two-stage methods for the analysis of pooled data. *Stat Med*. 2001;20(14):2115-2130.
22. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177-188.
23. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539-1558.

24. Akaike H. New look at statistical-model identification. *IEEE Trans Automat Contr*. 1982;19:716-723.
25. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *JAMA*. 1982;247(18):2543-2546.
26. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001;54(8):774-781.
27. Bouwmeester W, Moons KG, Kappen TH, et al. Internal validation of risk models in clustered data: a comparison of bootstrap schemes. *Am J Epidemiol*. 2013;177(11):1209-1217.
28. Huang SH, Hwang D, Lockwood G, Goldstein DP, O'Sullivan B. Predictive value of tumor thickness for cervical lymph-node involvement in squamous cell carcinoma of the oral cavity: a meta-analysis of reported studies. *Cancer*. 2009;115(7):1489-1497.
29. Lam P, Au-Yeung KM, Cheng PW, et al. Correlating MRI and histologic tumor thickness in the assessment of oral tongue cancer. *AJR Am J Roentgenol*. 2004;182(3):803-808.
30. Park JO, Jung SL, Joo YH, Jung CK, Cho KJ, Kim MS. Diagnostic accuracy of magnetic resonance imaging (MRI) in the assessment of tumor invasion depth in oral/oropharyngeal cancer. *Oral Oncol*. 2011;47(5):381-386.
31. Mark Taylor S, Drover C, Maceachern R, et al. Is preoperative ultrasonography accurate in measuring tumor thickness and predicting the incidence of cervical metastasis in oral cancer? *Oral Oncol*. 2010;46(1):38-41.
32. Ebrahimi A, Murali R, Gao K, Elliott MS, Clark JR. The prognostic and staging implications of bone invasion in oral squamous cell carcinoma. *Cancer*. 2011;117(19):4460-4467.
33. Fried D, Mullins B, Weissler M, et al. Prognostic significance of bone invasion for oral cavity squamous cell carcinoma considered T1/T2 by AJCC size criteria. *Head Neck*. 2014;36(6):776-781.
34. Woolgar JA, Scott J. Prediction of cervical lymph node metastasis in squamous cell carcinoma of the tongue/floor of mouth. *Head Neck*. 1995;17(6):463-472.