

Comparison of the American Joint Committee on Cancer N1 versus N2a nodal categories for predicting survival and recurrence in patients with oral cancer: Time to acknowledge an arbitrary distinction and modify the system

Ardalan Ebrahimi, MBBS, MPH,^{1,2,3*} Ziv Gil, MD, PhD,^{4,5} Moran Amit, MD, PhD,^{4,5} Tzu-Chen Yen, MD,⁶ Chun-Ta Liao, MD, PhD,⁶ Pankaj Chatturvedi, MBBS,⁷ Jaiprakash Agarwal, MBBS, PhD,⁷ Luiz Kowalski, MD, PhD,⁸ Matthias Kreppel, MD, PhD,⁹ Claudio Cernea, MD,¹⁰ Jose Brandao, MD,¹⁰ Gideon Bachar, MD,¹¹ Andrea Bolzoni Villaret, MD,¹² Dan Fliss, MD,¹³ Eran Fridman, MD,¹³ K. Thomas Robbins, MD,¹⁴ Jatin Shah, MD,¹⁵ Snehal Patel, MD,¹⁵ Jonathan Clark, MBBS, BSc, MD,^{1,2} The International Consortium for Outcome Research (ICOR) in Head and Neck Cancer

¹Sydney Head and Neck Cancer Institute, Royal Prince Alfred Hospital, Sydney, New South Wales, Australia, ²Department of Head and Neck Surgery, Liverpool Hospital, Sydney, New South Wales, Australia, ³University of New South Wales, Sydney, New South Wales, Australia, ⁴The Laboratory for Applied Cancer Research, the Technion, Israel Institute of Technology, Haifa, Israel, ⁵Department of Otolaryngology, Rambam Medical Center, Rappaport School of Medicine, the Technion, Israel Institute of Technology, Haifa, Israel, ⁶Chang Gung Memorial Hospital, Taoyuan, Taiwan, ⁷Tata Memorial Hospital, Mumbai, India, ⁸Hospital A.C. Camargo, São Paulo, Brazil, ⁹Department of Oral and Cranio-Maxillo and Facial Plastic Surgery, University of Cologne, Cologne, Germany, ¹⁰Department of Head and Neck Surgery, University of São Paulo Medical School, São Paulo, Brazil, ¹¹Department of Otolaryngology Head and Neck Surgery, Rabin Medical Center, Petach Tikva, Israel, ¹²Department of ENT, University of Brescia, Brescia, Italy, ¹³Department of Otolaryngology Head and Neck Surgery, Tel Aviv Medical Center, Tel Aviv, Israel, ¹⁴Southern Illinois University School of Medicine, Carbondale, Illinois, ¹⁵Head and Neck Surgery Service, Memorial Sloan-Kettering Cancer Center, New York, New York.

Accepted 12 September 2014

Published online 18 June 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/hed.23871

ABSTRACT: *Background.* We hypothesized that pathological N1 (pN1) and N2a (pN2a) nodal disease portend a similar prognosis in patients with oral cancer.

Methods. An international multicenter study of 739 oral squamous cell carcinoma (SCC) patients with pN1 or pN2a stage disease was conducted. Multivariable analyses were performed using Cox proportional hazard models to compare locoregional failure, disease-specific survival (DSS), and overall survival (OS). Institutional heterogeneity was assessed using 2-stage random effects meta-analysis techniques.

Results. Univariate analysis revealed no difference in locoregional failure ($p = .184$), DSS ($p = .761$), or OS ($p = .475$). Similar results were

obtained in adjusted multivariable models and no evidence of institutional heterogeneity was demonstrated.

Conclusion. The prognosis of pN2a and pN1 disease is similar in oral SCC suggesting these categories could be combined in future revisions of the nodal staging system to enhance prognostic accuracy. However, these results may reflect more aggressive treatment of N2a disease; hence, we caution against using these data to deintensify treatment. © 2015 Wiley Periodicals, Inc. *Head Neck* 38: 135–139, 2016

KEY WORDS: head and neck neoplasms, oral squamous cell carcinoma, lymph node metastases, cancer staging, prognosis

INTRODUCTION

In 1977, the first edition of the Manual for Staging of Cancer was published by the American Joint Committee on Cancer (AJCC).¹ The AJCC made a distinction in nodal (N) staging for patients with oral squamous cell carcinoma (SCC) and a single ipsilateral positive node based on size: with those ≤ 3 cm in greatest dimension classified as N1, whereas those > 3 cm but ≤ 6 cm in greatest dimension were classified as N2a.² They qualified this by suggesting that most cervical nodal masses over 3 cm in diameter are not single nodes, but rather confluent nodes or soft tissue deposits in the neck.² From a staging perspective, this distinction takes on an even greater significance given it up-stages the TNM classification

for patients with T1 to 3 primaries from stage III to IVa.³

Although the simplicity and consistency of the current AJCC staging system for head and neck cancer promotes clinical utility, it is widely acknowledged that the prognostic performance is suboptimal in selected subgroups.^{4–7} Specific to the current study, there is some evidence from studies of multiple head and neck cancer sites to suggest that the prognosis for N1 and N2a disease is similar,^{2,8,9} and that perhaps these 2 categories should be combined for staging purposes.^{2,8} However, this observation has been made in a number of single institution analyses in which the number of patients with N2a disease tends to be very small, limiting statistical power and inference.

We hypothesized that patients with oral SCC and pathological N2a (pN2a) neck disease have a similar outcome to those with pN1 disease. The primary purpose of this study was to compare the prognosis of pN1 and pN2a neck disease in a multicenter cohort of patients with oral SCC treated by primary surgery.

*Corresponding author: A. Ebrahimi, Sydney Head and Neck Cancer Institute, Royal Prince Alfred Hospital, Level 5, 119-143 Missenden Road, Chris O'Brien Lifehouse, Camperdown, New South Wales, Sydney, Australia. E-mail: are472@mail.harvard.edu

MATERIALS AND METHODS

Study population

This multicenter study included pooled data on patients with oral SCC undergoing surgical resection of the primary tumor and neck dissection from 11 participating cancer centers worldwide. We identified 739 patients staged as pN1 or pN2a as candidates for inclusion in this study. After excluding cases with neoadjuvant therapy, perioperative mortality, lack of follow-up data, and age <20 years, the final study population consisted of 729 patients treated between 1970 and 2011. Ethics approval was obtained from local institutional review board committees of participating centers.

Histopathological analysis

In general, procedures at participating centers were in accordance with guidelines for the histopathological assessment of head and neck carcinoma, with assessments performed by pathologists experienced in the examination of head and neck tumors at each center. We assumed some heterogeneity in specimen dissection and tissue handling in view of the extended time period of the study as well as the number of involved institutions, surgeons, and pathologists. We sought to determine if significant between-center heterogeneity exists in regard to the prognosis of pN1 versus pN2a disease, which could reflect center-specific practice.

Statistical analysis

Statistical analysis was performed using Stata version 12.0 SE (StataCorp LP, College Station, TX). All statistics were 2-sided and a value of $p < .05$ was considered statistically significant. Categorical data were compared using the chi-square test or Fisher's exact test when appropriate. Overall survival (OS) was calculated from the date of surgery to the date of death or last follow-up. For disease-specific survival (DSS), patients who died from causes other than oral SCC were censored at the time of death. Locoregional failure was defined as pathologically proven tumor relapse in the primary site or neck. Differences in survival and locoregional failure rates between pN1 and pN2a patients were determined using univariate Cox regression analysis, and survival curves generated using the Kaplan–Meier method when appropriate. Other covariates of interest included age at diagnosis in years (continuous), pT category (T1–2, T3–4), surgical margin status (clear, involved), extracapsular spread (ECS; absent, microscopic, macroscopic), time period of primary treatment (1970–1989, 1990–2011), and postoperative radiotherapy (PORT). Multivariable analyses were performed using Cox proportional hazards regression, stratified by study center, to determine if the prognosis associated with pN1 and pN2a disease differs after adjusting for these covariates. Model diagnostics were performed to check for linearity of continuous predictors and the proportional hazards assumption.

Investigation for the presence of between-center heterogeneity was performed using a 2-stage random effects model.¹⁰ At the first stage of analysis, the difference in prognosis between pN2a and pN1 patients was deter-

mined for each center. We used univariable analyses in view of the small number of patients and events in most institutions. In the second stage, the center-specific effect estimates were introduced into the random effects model of DerSimonian and Laird, which allows for unexplained sources of heterogeneity between centers.¹¹ Heterogeneity across centers was assessed using Cochran's Q test ($p < .1$ was considered statistically significant given 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).¹²

RESULTS

Patient demographics

The study cohort consisted of 729 patients with oral SCC, including 452 men and 277 women, with a median age of 54 years (range, 22–91 years) and median follow-up of 35 months. Relevant demographic and clinicopathological details are summarized in Table 1. There were 661 pN1 and 68 pN2a stage patients. The rate of ECS was slightly higher in the pN2a group but the difference was not statistically significant (56.1% vs 48.9%; $p = .391$). This may represent a type II error because of inadequate power. Patients with pN2a disease were more likely to receive PORT (94.1% vs 86.1%; $p < .001$) and more likely to be treated with concurrent chemotherapy or cetuximab (33.9% vs 18.4%; $p < .001$). As expected, the pN2a group was more likely to have clinically evident disease (80.9% vs 61.8%; $p = .002$) and undergo comprehensive neck dissection (55.8% vs 32.3%; $p = .001$).

Study centers

Heterogeneity between study centers was noted with statistically significant differences in mean age ($p < .001$), sex distribution ($p < .001$), nodal yield ($p < .001$), pT category ($p < .001$), pN category ($p < .001$), presence of ECS ($p < .001$), margin status of the primary ($p < .001$), adjuvant therapy use ($p < .001$), and distribution of year of primary treatment ($p < .001$).

Survival and locoregional failure analyses

The median OS for the study population was 84 months. There were 264 deaths, 171 of which were due to oral SCC. Locoregional failure occurred in 177 patients. As shown in Table 2 and Figure 1, univariate analysis revealed no significant difference in locoregional failure (hazard ratio [HR] = 1.37; 95% confidence interval [CI] = 0.86–2.19; $p = .184$), DSS (HR = 0.92; 95% CI = 0.53–1.59; $p = .761$), or OS (HR = 0.84; 95% CI = 0.53–1.35; $p = .475$) between patients with pN2a versus pN1 disease. Similar results were obtained after adjusting for age, pT category, surgical margin status, ECS, time period of primary treatment, and PORT in Cox proportional hazards regression models (Table 2). A sensitivity analysis was performed restricting the study population to patients receiving PORT ($n = 633$). Similar results were obtained for locoregional failure, DSS, and OS in both univariable and multivariable analyses. Finally, 2-stage random effects meta-analyses were used to investigate for between-center heterogeneity in the difference pN2a versus pN1 disease. We confirmed the

TABLE 1. Comparison of baseline clinicopathological characteristics in pN1 versus pN2a staged disease (*N* = 729).

Variables	pN1 classification		pN2a classification		<i>p</i> value
	No. of patients	%	No. of patients	%	
Age, y					.861
≤45	181	27.4	19	27.9	
46–55	182	27.5	17	25.0	
56–65	169	25.6	16	23.5	
≥66	129	19.5	16	23.5	
Sex					.407
Male	413	62.5	39	57.4	
Female	248	37.5	29	42.6	
Pathological T classification					.213
T1–2	289	43.7	24	35.8	
T3–4	372	56.3	43	64.2	
TNM classification					<.001
III	387	58.6	0	0.0	
IV	274	41.5	68	100.0	
Extracapsular nodal spread*					.391
Absent	297	51.1	29	43.9	
Microscopic	214	36.8	30	45.5	
Macroscopic	70	12.1	7	10.6	
Surgical excision margin					.338
Clear	605	91.7	60	88.2	
Involved	55	8.3	8	11.8	
Adjuvant therapy					<.001
No	93	14.1	3	4.4	
Radiotherapy	446	67.5	42	61.8	
Chemoradiotherapy	94	14.2	8	11.8	
Radiotherapy + cetuximab	28	4.2	15	22.1	
Time period of primary treatment					.126
1970–1989	80	12.1	4	5.9	
1990–2011	581	87.9	64	94.1	
Type of neck dissection					.002
Elective	252	38.2	13	19.1	
Therapeutic	407	61.8	55	80.9	
Extent of neck dissection					.001
Selective (I–III or I–IV)	255	51.4	13	25.0	
Comprehensive (I–V)	160	32.3	29	55.8	
Bilateral	81	16.3	10	19.2	
Study center					
Brescia, Italy	11	1.7	2	2.9	
Camargo, São Paulo, Brazil	93	14.1	7	10.3	
Cologne, Germany	30	4.5	15	22.1	
MSKCC, USA	70	10.6	2	2.9	
Petach Tikva, Israel	9	1.4	0	0.0	
São Paulo, Brazil	33	5.0	10	14.7	
SHNCI, Australia	48	7.3	6	8.8	
Southern Illinois, USA	15	2.3	3	4.4	
CGMH-Taoyuan, Taiwan	171	25.9	6	8.8	
Tata Memorial Hospital, India	165	25.0	16	23.5	
Tel Aviv, Israel	16	2.4	1	1.5	

Abbreviations: pN1, pathological stage N1; pN2, pathological stage N2a; MSKCC, Memorial Sloan-Kettering Cancer Center; SHNCI, Sydney Head & Neck Cancer Institute; CGMH, Chang Gung Memorial Hospital.

*There were 82 patients with missing data on extracapsular spread.

absence of significant institutional heterogeneity for locoregional failure ($I^2 = 0\%$; $p = .773$), DSS ($I^2 = 0\%$; $p = .926$), and OS ($I^2 = 0\%$; $p = .974$).

DISCUSSION

Since the first edition of the AJCC staging system for oral SCC was published in 1977, patients with single ipsilateral positive node ≤ 3 cm in greatest dimension have

been classified as N1, whereas those >3 cm but ≤ 6 cm in greatest dimension are classified as N2a.¹ Furthermore, the N2a designation upstages patients with pT1-3 primary tumors and a single ipsilateral node to TNM stage IVa, rather than stage III, as in the case of N1 disease.³ Although there is some evidence to suggest that N1 and N2a disease share a similar prognosis,^{2,8,9} the number of patients with N2a disease in single institution analyses tends to be too small to draw reliable conclusions. For

TABLE 2. Univariate and multivariable analyses of the association difference in survival and locoregional failure between N1 and N2a neck disease.

pN2a vs pN1	Locoregional failure		DSS		OS	
	HR (95% CI)	<i>p</i> value	HR (95% CI)	<i>p</i> value	HR (95% CI)	<i>p</i> value
Univariate	1.37 (0.86–2.19)	.184	0.84 (0.53–1.35)	.475	0.84 (0.53–1.35)	.475
Multivariable*	1.30 (0.74–2.30)	.361	0.75 (0.43–1.31)	.308	0.75 (0.43–1.31)	.308

Abbreviations: pN1, pathological stage N1; pN2, pathological stage N2a; HR, hazard ratio; CI, confidence interval.

* Multivariable analyses performed with Cox regression models adjusting for age in years, pathological T classification (T1–2, T3–4), surgical margin status (clear, involved), extracapsular nodal spread (absent, microscopic, macroscopic), time period of primary treatment (1970–1989, 1990–2011), and adjuvant radiotherapy (yes, no).

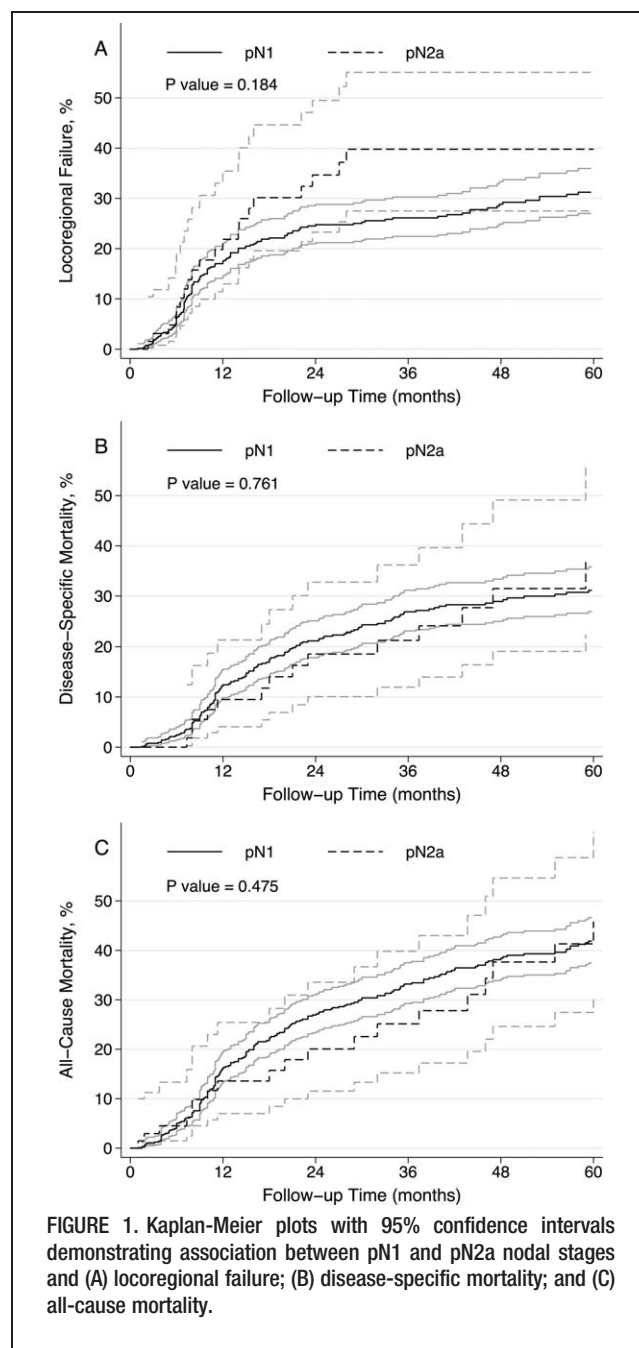
such a small proportion of patients to be given a distinct staging category only makes clinical sense if their prognosis is markedly different from other patients in adjacent N categories. This prompted the current multicenter international study, which demonstrated no difference in survival or recurrence between pN1 and pN2a stage disease, suggesting a modification of the current AJCC nodal staging to combine these categories may be appropriate.

We found no significant difference in prognosis for pN2a disease compared with pN1 for OS, DSS, or locoregional failure on either univariable or multivariable adjusted models. Given that we noted significant variability between study centers in a range of relevant clinicodemographic characteristics, and the likelihood of variation in factors, such as preoperative staging and imaging, as well as surgical and pathological techniques, we sought to assess whether the uniform prognosis seen in pN2a and pN1 disease was institution dependent. However, a 2-stage random effects model failed to show any evidence for heterogeneity between centers in this regard, suggesting the finding is robust and widely applicable.

We hypothesize that the similar prognosis observed in pN2a and pN1 disease may be related to more frequent and/or aggressive administration of adjuvant therapy in the former group. Unfortunately, we were unable to formally test this hypothesis because, in our cohort, only 3 patients (4.4%) with pN2a disease were treated by surgery alone. However, we noted that pN2a patients were more likely to receive PORT and this was more likely to be accompanied by concurrent chemotherapy or cetuximab compared with pN1 staged patients. When we restricted our analysis to only patients who received PORT, the results were unchanged.

The key issue arising from our findings is the potential implications for staging. Certainly, it is widely acknowledged that the prognostic performance of the AJCC staging system for head and neck cancer is suboptimal in some subgroups.^{4–7} Specific to the current study, some authors have proposed a revision of nodal staging that combines the N1 and N2a categories,^{2,8} although this has not been adopted into clinical practice. The integration of our results depends critically on the balance between the goals of the TNM staging system. On the one hand, an ideal staging system should be useful for guiding treatment decisions.⁷ In this respect, we cannot exclude the possibility that, in the absence of adjuvant therapy, pN2a disease has a worse prognosis than pN1 disease and the current distinction is reasonable on these grounds. However, arguably the most important goal of any staging system is the ability to accurately predict prognosis.⁴ In this

respect, combining the pN1 and pN2a categories seems appropriate, particularly if it is common practice for both groups to receive adjuvant therapy, as supported by our



study population. If the staging was modified and the pN2a category absorbed into pN1, 39 of 68 patients (57.4%) in our study population would be downstaged from TNM class IVa to III, because their primary tumors were staged as pT1-3.

This study had several limitations. First, it was a retrospective analysis and treatment was not assigned in a randomized fashion, which makes direct comparison of pN1 and pN2a disease problematic given the latter may be treated more aggressively. We have attempted to restrict this concern by subgroup analysis of only patients who received adjuvant therapy. Second, despite using data from 11 comprehensive cancer centers worldwide over an extended period of time, the number of pN2a patients was still relatively small at 68, which may have limited the power of the study. Finally, our sample size was inadequate to perform exploratory analyses to investigate for prognostic differences in subgroups based on the presence of ECS or the administration of PORT.

In conclusion, our results indicate that the overall prognosis associated with pN2a nodal disease is comparable to that of pN1 disease in patients with oral SCC. We noted that pN2a disease was treated more aggressively in regard to adjuvant therapy, which may account for the observed similar prognosis. In view of our findings and the fact that these patients generally do receive postoperative radiotherapy, it may be reasonable to downstage pN2a disease by absorbing it into the pN1 category. Our results suggest that such a modification to the AJCC stag-

ing system may enhance prognostic accuracy but we caution against using these data to deintensify treatment of patients.

REFERENCES

1. American Joint Committee for Cancer Staging and End Results Reporting: Manual for Staging of Cancer. Chicago, IL: American Joint Committee; 1977.
2. Hall SF, Groome PA, Dixon PF. Does N stage predict survival? *J Otolaryngol* 1996;25:296-299.
3. Edge SB, Byrd DR, Compton C, Fritz AG, Greene FL, Trotti A III, editors. AJCC Cancer Staging Manual. 7th ed. New York, NY: Springer; 2010.
4. 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;23:613-624.
5. 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;31:1-8.
6. Kreppel M, Drebber U, Rothamel D, et al. Prognostic impact of different TNM-based stage groupings for oral squamous cell carcinoma. *Head Neck* 2011;33:1467-1475.
7. Manikantan K, Sayed SI, Syrigos KN, et al. Challenges for the future modifications of the TNM staging system for head and neck cancer: case for a new computational model? *Cancer Treat Rev* 2009;35:639-644.
8. Hall SF, Groome PA, Rothwell D, Dixon PF. Using TNM staging to predict survival in patients with squamous cell carcinoma of head & neck. *Head Neck* 1999;21:30-38.
9. Bataini JP, Bernier J, Jaulerry C, Brunin F, Pontvert D. Impact of cervical disease and its definitive radiotherapeutic management on survival: experience in 1013 patients with squamous cell carcinomas of the oropharynx and pharyngolarynx. *Laryngoscope* 1990;100:716-723.
10. Stukel TA, Demidenko E, Dykes J, Karagas MR. Two-stage methods for the analysis of pooled data. *Stat Med* 2001;20:2115-2130.
11. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177-188.
12. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21:1539-1558.