Oral Oncology 99 (2019) 104438

Contents lists available at ScienceDirect

Oral Oncology

journal homepage: www.elsevier.com/locate/oraloncology

Prognostic value of radiologic extranodal extension and its potential role in future N classification for nasopharyngeal carcinoma

Tianzhu Lu[a,1], Yujun Hu[a,1], Youping Xiao[b], Qiaojuan Guo[c,d], Shao Hui Huang[e], Brian O'Sullivan[e], Yanhong Fang[b], Jingfeng Zong[c], Ying Chen[b], Shaojun Lin[c], Yunbin Chen[b,*], Jianji Pan[c,f]

[a] Department of Radiation Oncology, Fujian Medical University Cancer Hospital, Fuzhou, China
[b] Department of Radiology, Fujian Cancer Hospital & Fujian Medical University Cancer Hospital, Fuzhou, China
[c] Department of Radiation Oncology, Fujian Cancer Hospital & Fujian Medical University Cancer Hospital, Fuzhou, China
[d] Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai, China
[e] Department of Radiation Oncology, Princess Margaret Cancer Centre/University of Toronto, Toronto, Canada
[f] Fujian Provincial Key Laboratory of Translational Cancer Medicine, Fujian Medical University Cancer Hospital, Fuzhou, China

ARTICLE INFO

Keywords:
Head and neck cancers
Nasopharyngeal carcinoma
Extranodal extension
Prognostication
N classification

ABSTRACT

Purpose: We evaluated the prognostic value of various grades of radiologic extranodal extension (rENE) and their potential roles in N-classification refinement for nasopharyngeal carcinoma (NPC).

Methods and Materials: All NPC patients treated with IMRT in our institution between 2005 and 2011 were included. Pre-treatment MR of cN+ cases were reviewed and rENE was recorded asG0: lymph nodes (LNs) without rENE; G1: tumor infiltrating beyond individual nodal capsule(s) into the surrounding fat plane; G2: coalescent nodal mass with unequivocal evidence of rENE; G3: tumor infiltrating beyond nodal capsule into adjacent structures. Multivariable analysis (MVA) assessed prognostic value of rENE for distant metastasis (DM) and death adjusted for age, gender, LDH, T-classification, N-classification, and chemotherapy cycles.

Results: A total of 1390 of 1616 (86%) NPC were cN+, and rENE was detected in 826/1390 (59%) patients: 256 (18.4%) G1-rENE, 487 (35%) G2-rENE, and 83 (6%) G3-rENE. MVA confirmed that G2-/G3-rENE had increased risk of DM (HR: 2.05/3.18, both $p < 0.001$) and death (HR: 1.62/2.39, $p = 0.002/p < 0.001$), while G1-rENE was non-prognostic (DM: $p = 0.172$; death: $p = 0.320$). We propose a refined N: *New-N1*: N1/N2 without G2-/G3-rENE; *New-N2*: N1_G2-rENE; *New-N3*: N2_G2-rENE, N1/N2_G3-rENE, or N3. The *New-N* classification had a lower AIC and higher c-index for DM (AIC: 3809.6 vs 3830.9; c-index: 0.700 vs. 0.677) and death (AIC: 3693.8 vs. 3705.9; c-index: 0.735 vs. 0.725) versus TNM-8 N.

Conclusions: G2- and G3-rENE are independently prognostic for DM and death in NPC. Compared to the TNM8 N-classification, a refined N-classification incorporating G2- and G3-rENE improves prognostication of DM and mortality risk.

Introduction

Clinical extranodal extension (cENE) has been introduced as a new parameter in the cN classification in the UICC/AJCC 8th edition TNM (TNM-8) for non-viral related head and neck cancer (HNC) [1]. However, the cENE definition is primarily based on stringent clinically unambiguous evidences (e.g. dermal involvement or soft tissue invasion with deep fixation/tethering to underlying structures or clinical signs of nerve involvement), radiologic ENE (rENE) is not the essential criterion for cENE. This was because there was paucity data on the reliability of rENE assessment, and hence uncertainty about its prognostic value. Emerging evidence has shown that rENE based on unequivocal CT/MR evidence can be reliably assessed with excellent specificity for pENE [2–4]. Besides, the presence of rENE has been shown to be a strong adverse prognostic factor in both non-viral related HNC [4] and HPV-positive oropharyngeal cancer [5,6].

The prognostic importance of rENE in NPC was reported as early as 2008 [7]. However, two subsequent studies showed that rENE was not an independent prognostic factor in NPC [8,9], likely related to the rENE assessment criteria where overcall is a possibility if non-stringent

* Corresponding authors at: Department of Radiology, Fujian Cancer Hospital & Fujian Medical University Cancer Hospital, Fuzhou 350014, China (Y. Chen). Department of Radiation Oncology, Fujian Medical University Cancer Hospital, No. 420 Fuma Road, Fuzhou 350014, China (J. Pan).
E-mail addresses: yunbinchen@126.com (Y. Chen), panjianji1956@fjmu.edu.cn (J. Pan).
[1] Both Tianzhu Lu and Yujun Hu and contributed equally to this study.

https://doi.org/10.1016/j.oraloncology.2019.09.030
Received 30 July 2019; Received in revised form 21 September 2019; Accepted 29 September 2019
Available online 22 October 2019
1368-8375/

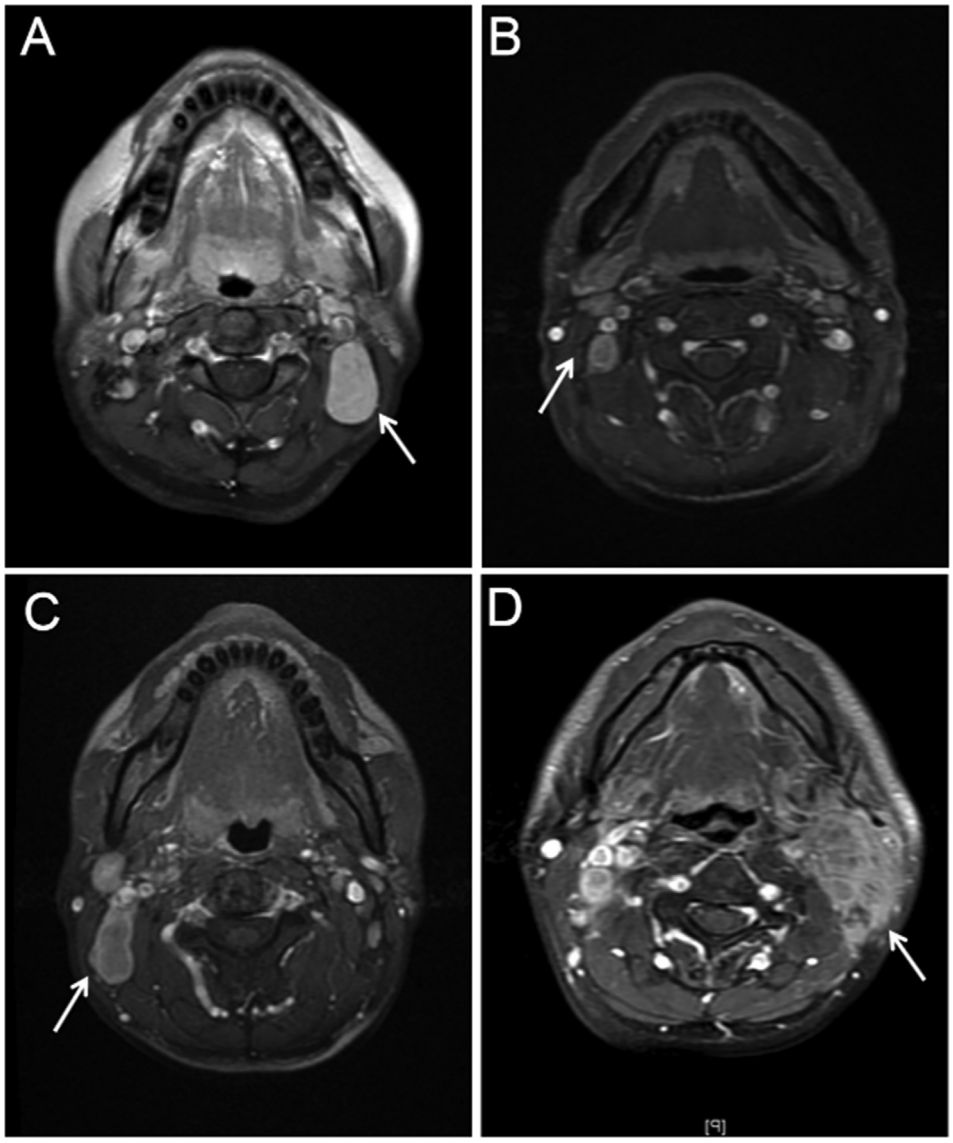

Fig. 1. Axial T1-weighted fat-suppressed contrast MRI of NPC patients with metastatic neck nodes along internal jugular chain. (A) a metastatic node without rENE; (B) a metastatic node with rENE infiltrating surrounding fat (G1-rENE); (C) a metastatic node with coalescent LNs (G2-rENE); (D) a metastatic node with rENE infiltrating adjacent muscles (G3-rENE).

criteria or low-quality images are used. The incidence or rENE varied among different centers and different radiologists (range 33.6–75.6%) for N positive (cN+) patients [7–9]. Another potential inconsistency about the prognostic value of rENE in the literature is that less extensive rENE might not be reliably assessed and may not be as prognostically important in the IMRT era compared to overt ("high grade") rENE. A recent study [10] showed that only rENE with adjacent structure infiltration was independently prognostic for distant metastasis-free survival (DMFS) and overall survival (OS), while low-grade rENE, manifested by infiltration into surrounding fat only, did not show significant impact on prognosis. Our recent study [11] focusing on stage II patients also confirmed these findings. In addition, we also found that a coalescent nodal mass comprising 2 or more inseparable lymph nodes (LNs) could represent a special form of rENE and also carries prognostic significance in NPC. Our study also showed that the prognostic value of rENE increased with both the grade and certainty of rENE determined by a radiologist.

Based on these findings, we conducted this retrospective study to comprehensively evaluate the role of various grades of rENE in NPC patients. We hypothesize that high-grade rENE, including either a coalescent nodal mass or invasion of adjacent muscles/vessels, is prognostically important in the IMRT era and could refine the current N-classification to improve outcome prediction in cN+ NPC patients. In addition, we also evaluated that reliability (inter-rater concordance) and reproducibility (intra-rater concordance) of various grades of rENE.

Methods and materials

Patients

Following Hospital Review Board approval, we conducted this retrospective review of all histologically confirmed non-metastatic (M0) NPC patients treated with definitive IMRT with or without chemotherapy in our institution between June 2005 and December 2011. All patients were reclassified according to the TNM-8 [12].

Evaluation of rENE on magnetic resonance imaging (MRI)

Pre-treatment gadolinium-enhanced MRI was performed using 1.5 T.MR (Signa Excite 1.5 T HD Twin Speed, GE Healthcare, WI, USA)

Table 1
Clinical Characteristics of 1616 Nasopharyngeal Carcinoma Patients.

Covariate	cN0 (n = 226, 14.0%)	cN+ patients (n = 1390, 86.0%)		
		No rENE (n = 564, 39.9%)	rENE+ (n = 826, 60.1%)	*p-value*
Age (year)	48	45	46	*0.116*
Median (range)	(19–78)	(12–84.)	(11–82)	
LDH (IU/ml)	133	137	146	*< 0.001*
Median (range)	(71–458)	(86–411)	(88–586)	
Sex				*0.013*
Female	57 (25.2)	161 (28.5)	187 (22.6)	
Male	169 (74.8)	403 (71.5)	639 (77.4)	
Pathology				*0.707*
Keratinizing squamous cell	4 (1.8)	6 (1.1)	7 (0.8)	
Nonkeratinizing, differentiated	16 (7.1)	38 (6.7)	62 (7.5)	
Nonkeratinizing, undifferentiated	206 (91.2)	520 (92.2)	757 (91.6)	
T-category				*0.001*
T1	77 (34.1)	167 (29.6)	169 (20.5)	
T2	32 (14.2)	90 (16.0)	166 (20.1)	
T3	74 (32.7)	199 (35.3)	318 (38.5)	
T4	43 (19.0)	108 (19.1)	173 (20.9)	
N-category				*< 0.001*
N0	226 (100)			
N1		464 (8.8)	420 (51.5)	
N2		91 (16.6)	255 (30.4)	
N3		9 (1.6)	151 (18.1)	
Clinical stage				*< 0.001*
I	77 (34.1)			
II	32 (14.2)	207 (36.7)	158 (19.1)	
III	74 (32.7)	240 (42.6)	365 (44.1)	
IV	43 (19.0)	117 (20.7)	304 (36.8)	
Chemotherapy (cycles)				*< 0.001*
≤3	174 (77.0)	339 (60.1)	388 (47.0)	
> 3	52 (23.0%)	225 (39.9)	438 (53.0)	
Median FU (range) (month)	55 (10–105)	55 (2–105)	51 (2–105)	
5-year LRC (95% CI)	94.1% (91.0–97.2)	96.4% (89.3–94.3)	88.5% (96.2–90.8)	*< 0.001*
5-year DMFS (95% CI)	93.7% (90.6–86.8)	89.9% (87.2–92.6)	73.8% (70.5–77.1)	*< 0.001*
5-year OS (95% CI)	90.1.% (85.8–94.4)	87.6% (84.7–90.5)	77.3% (77.3–80.4)	*< 0.001*

Abbreviation: rENE: radiologic extra-nodal extension. FU: fellow-up time; RRFS: regional relapse-free survival; DMFS: distant metastasis-free survival; OS; overall survival; 5y: 5 years.

or 3.0 T MR (Achieva 3.0 T, Philips Healthcare, Best, Netherlands) with a slice thickness of 5 mm encompassing from the lower temporal lobe to the supraclavicular region. An axial fat-suppressed T2-weighted sequence, axial T1-weighted spin-echo sequence, and post-contrast-enhanced T1-weighted spin-echo sequence with fat-suppression following a bolus injection of gadolinium-based agents was obtained. The detailed MR scanning protocol were described previously [13]. Pre-treatment MRIs were reviewed by a specialized neuroradiologist (YX). rENE was defined as an involved LN that had an unequivocal ill-defined nodal border, i.e., clearly discernible loss of sharp plane between the nodal capsule and the surrounding fat [2,4,5]. Equivocal/uncertain cases were classified as rENE−, while rENE+ was further classified into three grades: G1: overt LN with infiltration into surrounding fat plane only; G2: coalescent LNs (comprised of ≥2 LNs) with clear evidence of rENE); and G3: tumor invading beyond LN capsule into adjacent structures (i.e., muscles, nerves, parotid glands, etc.). The LN with the highest grade of ENE was recorded. Fig. 1 depicts the typical imaging appearance of rENE− and G1-3 rENE. To assess the inter/intra-rater reliability of rENE, MR scans of a subset of randomly selected cases were re-reviewed by the same radiologist (YX) after a three-month interval and were also independently reviewed by a second neuroradiologist (YF) [10].

Treatment and follow-up

All patients received IMRT according to our institutional protocols, described previously [14]. Generally, stage I diseases were treated by radiotherapy alone, while stages II-IVdisease were treated by a combination of chemo-radiotherapy. The most commonly used chemotherapy regimen was platinum plus paclitaxel or gemcitabine. After treatment completion, Follow-up intervals were 3 months in the 2 years, 3–6 months for the next 3–5 years and then annually.

Statistical analyses

Baseline characteristics were compared between the rENE+ vs. rENE− cohorts using Chi-square, Fisher's exact, or Wilcoxon-Mann-Whitney test for categorical variables. Kaplan–Meier survival analyses were used to estimate the locoregional control (LRC), DMFS, and OS with log-rank test for comparison of survival curves. Multivariable analyses (MVA) with Cox proportional hazard methods were used to estimate the risk of DM, LRC or death. The MVA models were constructed using a backward step. The criterion for the backward step is for significance at the 0.05 level before a variable could stay in the model, while parameters with a significance of greater than 0.10 were removed from the model.

Inter-rater and intra-rater concordance of rENE assessment were calculated using the Cohen Kappa coefficient (k) [5]. The prognostic performance of refined N-classification by incorporating rENE grades was compared with TNM8-N using Akaike information criterion (AIC) and Harrell's concordance index (c-Index) [15] where lower AIC or high c-index indicated more robust MVA models. All statistical tests were two-sided, p values < 0.05 were considered statistically significant. Statistical analyses were performed using SPSS (IBM version 18.0) and R-Studio (https://www.rstudio.com/).

Results

Patient population

A total of 1616 consecutive M0 NPC were identified, of which 1390 (86%) were cN+. The clinical characteristics of the 1616 eligible patients are listed in Table 1. rENE were identified in 826 (59%) cN+ patients, of whom 256 (18.4%), 487 (35%) and 83(6%) had G1-rENE, G2-rENE and G3-rENE, respectively. The frequency of rENE+ increased with higher N-category: N1 (n = 884): 420 (48%); N2 (n = 346): 255 (74%), N3 (n = 160): 151 (94%). Similarly, the frequency of higher grade rENE (G1/G2/G3) increased with higher N-category: N1: 152 (17%)/239 (27%)/29 (3%), N2:75 (22%)/159 (46%)/21 (6%), N3:29 (18%)/89 (56%)/33 (21%).

Prognostic values of rENE and its grades

Median follow-up for the entire cohort was 53 months. Locoregional failure, DM, and deaths were detected in 140, 273, and 271, respectively. Five-year LRC, DMFS, and OS were 90.6%, 82.2%, and 82.7%, respectively. When compared to patients with rENE−, the rENE+ cohort had a significantly inferior 5-years DMFS (73.8% vs 88.4%, p < 0.001]), OS (77.3% vs 87.6%, p < 0.001) and LRC (88.5% vs. 91.8%, p = 0.016) [Fig. 2]. Notably, DMFS were similar for patients with a coalescent node comprised 2 LNs vs ≥ 3 LNs (5 year-DMFS: 73.2% vs. 74.5%, p = 0.921; 5y-OS: 79.4% vs. 77.2%, p = 0.483) [Supplementary Fig. 1].

The prognostic value of rENE+ was confirmed in MVA for DM (HR 2.26, 95% CI: 1.65–3.09, p < 0.001) and death (HR 1.57 [1.22–2.11], p < 0.001), but not LRC (HR 1.28 [0.86–1.92], p = 0.228) after

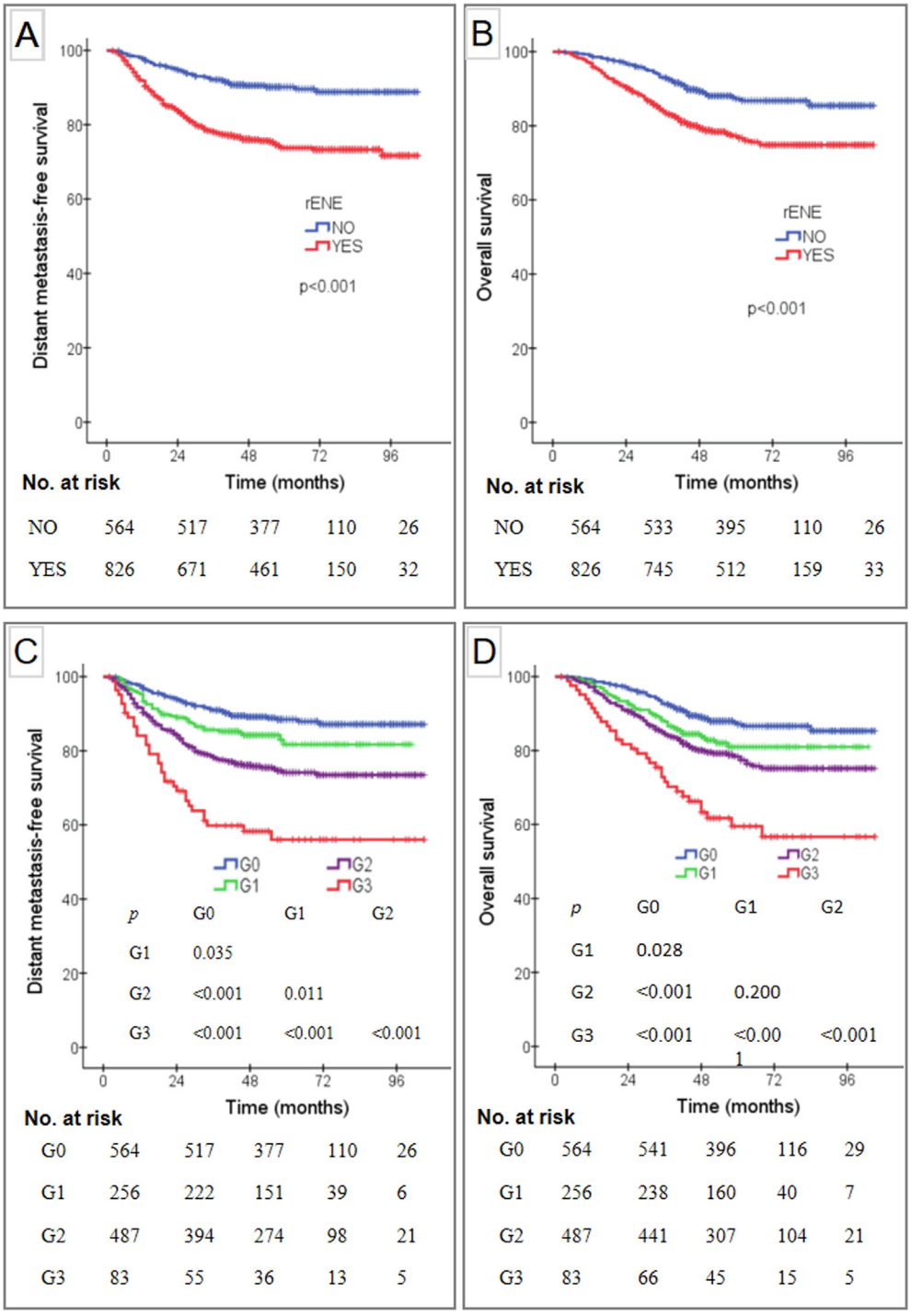

Fig. 2. The Kaplan-Meier curves of DMFS and OS in cN+ NPC patients with different grades of rENE: DMFS (A) and OS (B) for cN+ patients with or without rENE; DMFS (C) and OS (D) for cN+ patients with various grades of rENE.

adjustment for age, gender, LDH, T-classification, N-classification, and chemotherapy cycles [Table 2].

Further exploration of the prognostic value in various grades of rENE showed that the higher the grade of rENE, the lower the 5-year DMFS (rENE−: 88.7%, G1/G2/G3-rENE+: 82.7%/74.1%/56.0%; $p < 0.001$) and OS (rENE−: 87.5%, G1/G2/G3-rENE+: 81.0%/77.9%/59.5%, $p < 0.001$). After adjustment for age, gender, T, N, chemotherapy and LDH, MVA confirmed that G2-rENE and G3-rENE had increased HRs for DM (HR: 2.05 and 3.18, $p < 0.001$ and $p < 0.001$) and death (HR: 1.62 and 2.39, $p = 0.003$ and < 0.001), while G1-rENE was not significantly prognostic for DM (HR = 1.32, $p = 0.172$) nor death (HR = 1.22, $p = 0.321$) [Table 2].

Refinement of the N classification by incorporation of rENE

In viewing of prognostic importance of G2 and G3-rENE, we proposed a New N-classification incorporating G2/G3-rENE as a modifier to refine TNM-8 N-classification. Thus, 1390 cN+ patients were divided into six subgroups: N1 without G2/G3-rENE (n = 616), N2 without G2/G3-rENE (n = 166), N1 with G2-rENE (n = 239), N2 with G2-rENE (n = 159), N3 without G3-rENE (n = 127) and all N with G3-rENE (n = 83). Five-year DMFS (87.4% vs. 85.1%, $p = 0.522$) and OS (85.9%

Table 2
Multivariate Analysis of DMFS, OS and LRC by rENE Adjusting for Other Potential Predictors in 1390 cN+ Nasopharyngeal Carcinoma Patients.

Co-variate	DMFS		OS		LRC	
	HR (95%CI)	p-value	HR (95%CI)	p-value	HR (95%CI)	p-value
rENE		< 0.001		< 0.001		0.228
no	1.00		1.00		1.00	
yes	2.26 (1.65–3.09)		1.57 (1.22–2.11)		1.28 (0.86–1.92)	
Age (years)	1.01 (1.00–1.02)	0.007	1.04 (1.03–1.06)	< 0.001	1.02 (1.00–1.03)	0.031
Gender		0.030		0.029		
male	1.00		1.00			
female	0.70 (0.51–0.97)		0.67 (0.49–0.96)			
N classification	1.33 (1.13–1.60)	0.023	1.40 (1.15–1.69)	0.001	1.46 (1.22–1.75)	< 0.001
T Classification	1.26 (1.13–1.41)	< 0.001	1.38 (1.17–2.11)	< 0.001	1.45 (1.15–1.83)	0.002
LDH	1.00 (1.00–1.01)	0.019	1.00 (1.00–1.01)	< 0.001		
Co-variate	DMFS		OS		LRC	
	HR (95%CI)	p-value	HR (95%CI)	p-value	HR (95%CI)	p-value
rENE		< 0.001		< 0.001		0.608
G0	1.00		1.00		1.00	
G1	1.32 (0.89–1.95)	0.172	1.22 (0.83–1.78)	0.321	1.37 (0.84–2.25)	0.210
G2	2.05 (1.49–2.82)	< 0.001	1.62 (1.17–2.24)	< 0.001	1.28 (0.82–2.00)	0.283
G3	3.18 (2.02–4.99)	< 0.001	2.39 (1.52–3.76)	< 0.001	1.23 (0.57–2.69)	0.600
Age	1.01 (1.00–1.03)	0.005	1.04 (1.03–1.06)	< 0.001	1.02 (1.00–1.03)	0.031
gender		0.032		0.033		
male	1.00		1.00			
female	0.70 (0.51–0.97)		0.69 (1.03–1.06)			
N classification	1.26 (1.13–1.41)	< 0.001	1.32 (1.08–1.60)	0.006	1.46 (1.22–1.75)	< 0.001
T Classification	1.26 (1.05–1.53)	0.016	1.37 (1.22–1.54)	< 0.001	1.45 (1.15–1.83)	0.002
LDH	1.00 (1.00–1.01)	0.035	1.00 (1.00–1.01)	0.033		

Abbreviation: DMFS = distant metastasis-free survival; OS = overall survival; LRC: locoregional control; rENE: radiologic extra-nodal extension.

vs. 83.2%, p = 0.630) were similar between N1 and N2 without G2/G3-rENE subgroups, and both were higher compared to N1 with G2-rENE (87.4% vs. 80.9%, p = 0.015). N2 with G2-rENE had lower 5-year DMFS, (70.9% vs. 80.9%, p = 0.028) and OS (73.1% vs. 82.8%, p = 0.032) compared to N1 with G2-rENE, but similar to N3 without G3-rENE (DMFS: 70.9% vs. 66.8%, p = 0.662; OS: 73.1% vs. 75.4%, p = 0.672); G3-rENE patients had the lowest DMFS (56.0%) regardless of TNM-8 N (all p < 0.05), excepting similar DMFS with N3 with G0/G1/G2-rENE (56.0% vs. 66.8%, p = 0.073) [Supplement-Fig. 2].

Based on DMFS endpoint, we propose a refined *New_N* classification that included G2-rENE and G3-rENE: *New_N1*: N1/N2 without G2/G3-rENE (n = 781); *New_N2*: N1 with G2-rENE (n = 239); *New_N3*: N2 with G2-rENE, or N1-N2 with G3-rENE, or any N3 (n = 370) with the 5-year DMFS of 86.9%, 79.9%, and 66.8%, respectively [Fig. 3]. The MVA model using *New_N* variable instead of TNM-8 N appeared to have a better performance with a lower AIC and higher c-index for DM (AIC: 3809.6 vs. 3830.9; c-index: 0.700 vs. 0.677) and death (AIC: 3693.8 vs. 3705.9; c-index: 0.735 vs. 0.725). Besides, the *New_N* classification seemed to have a more balanced sample size distribution compared to TNM-8 N (Table 3).

Inter-rater and Intra-rater variation of rENE assessment

Inter-rater and intra-rater concordance of rENE assessments were performed in 100 randomly selected cases which would provide ≥80% power to detect significant discordance in rENE. The inter-rater and intra-rater kappa for rENE was 0.793 and 0.896, respectively. The kappa coefficients for G1-rENE, G2-rENE, and G3-rENE were 0.48 (0.26–0.71), 0.77 (0.64–0.91), and 0.94 (0.81–1.00) for the inter-rater, and 0.83 (0.68–0.97), 0.96 (0.90–1.00), and 1.0 for the intra-rater assessment respectively (Supplement-Table).

Discussion

This large population-based cohort study confirms the differential prognostic importance of rENE grades in cN+ NPC. G3-rENE manifested as invading adjacent structures, has the highest risk for DM; G2-rENE, defined as a coalescent nodal mass (a.k.a. “matted nodes”) is also prognostic. Although G1-rENE, manifested as an irregular nodal border only, was also associated with poor outcome in UVA but was not prognostic when adjusted for other covariates in MVA. Based on our data, we propose a *New_N* classification that incorporates G2-rENE and G3-rENE to refine TNM8 N as follows: *New_N1*: N1/N2 without G2-/G3 rENE (n = 781); *New_N2*: N1 with G2-rENE (n = 239); *New_N3*: N1/N2 with G3-rENE, N2 with G2-rENE or any N3 (n = 370). The 5-year DMFS and OS for the *New_N1*, *New_N2*, and *New_N3* separate distinctly with an orderly reduction in the outcome. Compared to the TNM-8 N-classification, the proposed *New_N* improved prognostication for the risk of DM and death with a lower AIC and higher c-index. Finally, we were able to show excellent intra-rater and inter-rater concordance for G2- and G3-rENE, supporting consideration to include this variable in clinical N classification. However, the inter-rater concordance for G1-rENE is unsatisfactory, which partially explains the lack of prognostication for G1-rENE. Nonetheless, this comprehensive analysis of the various pattern of rENE in a large cohort of NPC patients provides opportunity for future refinement of N classification based on DM risk, one of the major adverse outcomes of the disease.

Our study found that rENE is a strong prognostic factor for DM but not LRC. Presence of rENE could identify a subgroup of patients with a significantly higher DM risk (5-year risk: 26.2% vs. 11.6%) in a population that otherwise should have an excellent prognosis. Besides NPC, the strong association of rENE with higher DM risk has also been observed in several other head and neck cancer sites, such as oral cavity and HPV-positive oropharyngeal cancer [5,16,17]. The prognostic importance of rENE on DM is not surprising. When tumor invades through the nodal capsule, it is conceivable that it increases the risk of tumor cells entering the bloodstream increases.

Different grades of rENE have different prognostic impacts. We found that the HR for DM is highest in G3-rENE. It is possible that when tumor foci invade adjacent vasculature and muscle (i.e., G3 rENE), it is

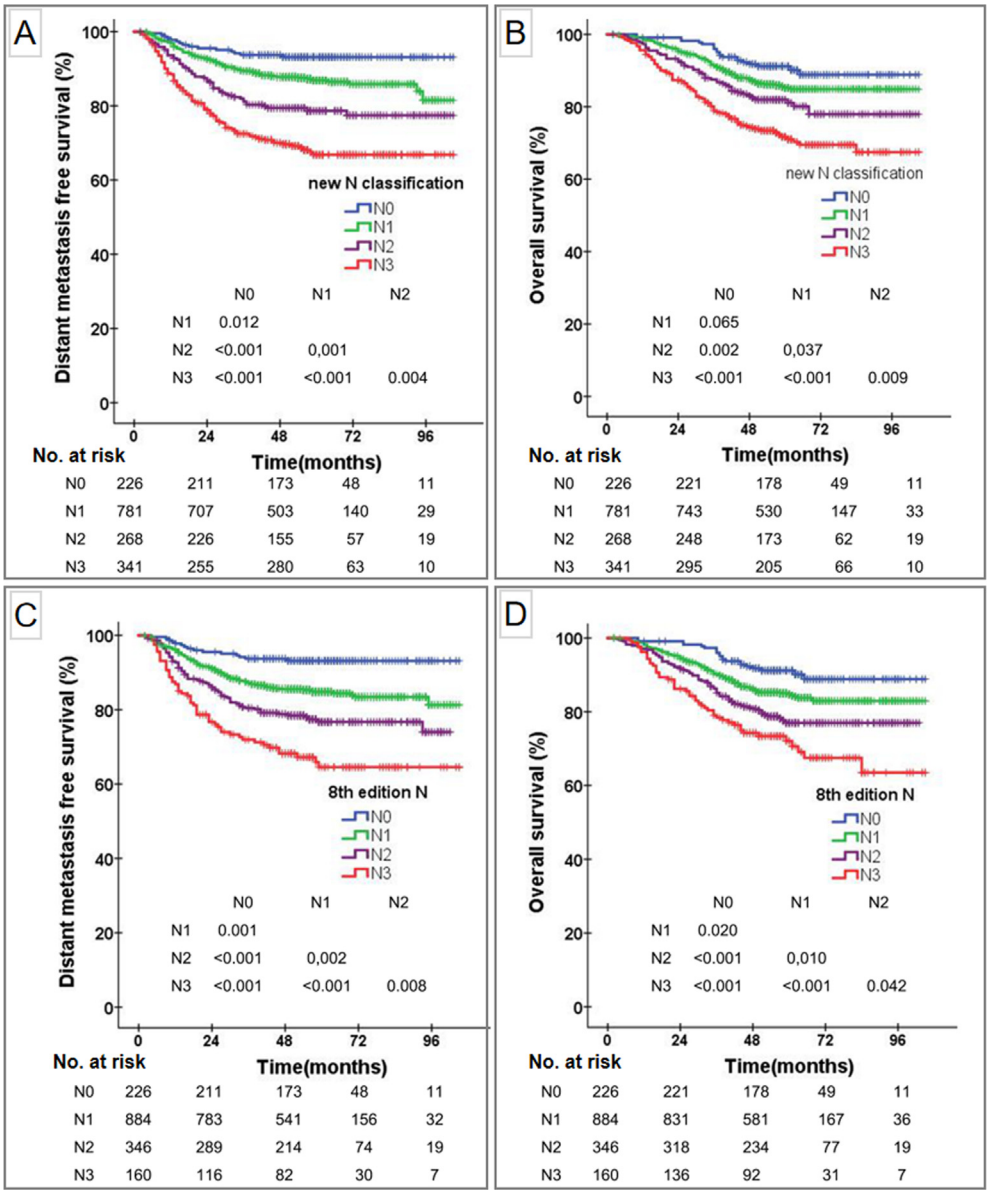

Fig. 3. Kaplan-Meier curves for DMFS or OS in NPC patients: (A) DMFS of the proposed new N classification, (B) OS of the proposed new N classification, (C) DMFS of the 8 edition N classification, and (D) OS of the 8 edition N classification.

easier to gain access to the blood circulation, thereby resulting in the highest DM risk. The strong detrimental impact of G3-rENE demonstrated in our study and by others [10] supports our proposal to reclassify any cN+ cases with G3-rENE to the highest N (N3) category.

Besides G3-rENE, we found that G2-rENE (coalescent nodal mass) is also prognostic although the HR is not as strong as G3-rENE. We propose a step-wise reclassification for cases with G2-rENE: N1 with G2-rENE to be *New_N2* while N2 with G2-rENE to be *New_N3*. Our study defined G2-rENE (coalescent LN) as 2 or more LNs merging together (assumed due to rENE), which was similar to the "matted nodes" phenomenon observed in previous studies [18–20], although the latter was arbitrarily defined as 3 or more LNs merged together (assumed to represent rENE). In their studies, there was no true explanation of why three LNs was chosen as the cutoff. It might be more reasonable for coalescent LNs to be defined as two or more LNs merging due to rENE. Our previous study [11] of stage II NPC showed that the 5 year DMFS of patients with coalescent nodes comprising 2 LNs was similar to that represented by 3 or more LNs (i.e., "matted nodes") (81.3% vs. 80.2%, $p = 0.943$). Data from this large study also yielded consistent results.

Similar to previous reports from our group [11] and others [10], we found that, while G1-rENE is significantly associated with inferior DMFS and OS in UVA, it lost its prognostic importance in MVA. It is possible that G1-rENE represents nodal capsule invasion confined to the fat plane without reaching blood vessels, making tumor cell less likely to shed into the bloodstream. Lack of prognostication of G1-rENE might also be attributable to poor inter-rater concordance (0.48) which leads to over-call of G1-rENE, thereby masking its prognostic value.

We have observed similar DMFS and OS between ipsilateral (N1) and bilateral (N2) neck disease without G2/3-rENE, and proposed to combine them as New_N1. Lack of distinction in prognosis between N1 and N2 neck disease in the absence of G2-/G3-rENE is not surprising. The prognostic difference of N2 and N1 disease observed previously might be mainly attributable to the rENE variable where the N2 subset has a higher proportion of G2-/G3-rENE (52.0%) compared to N1 cases (30.3%). Once the rENE, potentially the principal culprits as a cause of DM, is removed, the difference in DM risk disappears (14.9% vs. 12.9%, $p = 0.552$) [Supplementary Fig. 3].

DM has become the leading cause of failure in non-metastatic (M0)

Table 3
Performance of New N and TNM-8 N Classification in Predicting Risk of Distant Metastasis and Overall Survival.

Co-variate	TNM-8 N	New N
N Classification		
N0	N0	N0
N1	N1	N1/N2 without G2-/G3-rENE
N2	N2	N1 with G2-rENE
N3	N3	G2-rENE, N1-2_G3-rENE, N3
Distribution		
N0	226 (14.0%)	226 (14.0%)
N1	884 (54.7%)	781 (48.3%)
N2	346 (21.4%)	239 (14.8%)
N3	160 (9.9%)	370 (22.8%)
5-year DMFS		
N0	93.1%	93.1%
N1	84.6%	86.9%
N2	76.7%	79.9%
N3	64.5%	66.8%
c-index of DMFS (95%CI)	0.677 (0.646–0.708)	0.700 (0.671–0.729)
AIC of DMFS	3830.9	3809.6
5-year OS		
N0	90.1%	90.1%
N1	84.7%	85.2%
N2	77.0%	82.8%
N3	71.3%	70.9%
c-index of OS (95%CI)	0.725 (0.696–0.754)	0.735 (0.706–0.764)
AIC of OS	3705.9	3693.8

Abbreviation: DMFS = distant metastasis-free survival; OS = overall survival; AIC = Akaike information criterion.

NPC [21,22]. It is crucial to improve the prediction of DM risk in the treatment and design of clinical trials for NPC. Although many studies showed that elevated plasma EBV-DNA can as a DM predictor and should be included in clinical stage classification [23,24], the testing methodology of EBV-DNA has not been standardized nor is widely used in a routine clinical care setting so far [25]. Currently, the N classification is a significant and widely-used factor in the assessment of metastatic risk in NPC. Consideration of G2-/G3-rENE with good inter-rater and intra-rater concordance and strong prognostic factor of DM, it may permit refinement of a new N classification with DMFS as a reference. Our data suggest that such a new N classification, which has a lower AIC and a higher C-index, could strongly predict the risk of DM and death compare with TNM-8 N classification. Based on robust data with full statistical justification, rENE may have a potential role in future N classification for NPC. Further external studies are needed to validate our findings.

Several limitations should be noted when interpreting these results. Firstly, this is a retrospective study of a relatively unselected large NPC population that includes the entire disease. Although we have adjusted major confounders for DMFS, OS, and LRC, other factors might still influence results. Secondly, the incidence rate of rENE varies significantly in the literature, raising concern about the reliability of rENE assessment. Improved consistency in evaluating rENE parameter among centers is of paramount importance. Thirdly, the data were derived from an academic cancer center in an endemic area. Whether the findings can be reproduced and be generalizable to other cancer central remains to be determined. Further external validation in independent cohorts, especially multicenter studies, is warranted.

Conclusions

Radiologic ENE is an independent factor for DMFS and OS for cN+ patients, but is not a prognostic factor for LRC. Subdivision of the degree of rENE shows that G2- and G3-rENE can independently predict poor DMFS and OS for NPC. We propose refined N classification based DMFS endpoint where N1_G2-rENE+ is reclassified to N2, N2_G2-rENE + to N3, and any N with G3-rENE to N3. Compared to the TNM8 N classification, an N classification incorporating G2- and G3-rENE improves prognostication of risk of DM and death. If our findings are validated in other external studies, it will facilitate the future clinical trial design and TNM refinement in NPC.

Declaration of Competing Interest

None declared.

Acknowledgement

This work was sponsored by National Clinical Key Specialty Construction Program and Key Clinical Specialty Discipline Construction Program of Fujian, People's Republic of China. This research is also supported by grants from the National Natural Science Foundation of China (No. U1405221), grant from the Fujian Provincial Natural Science Foundation of China (No. 2016J01438, 2018J01275, and 2019J01194), grant from Science and Technology Program of Fujian Province, China (No. 2018Y2003), and grant from Startup Fund for scientific research, Fujian Medical University (No: 2017XQ2050). We acknowledge the O. Warwick Prize of the Canadian Cancer Society for supporting BO'S academic activities.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.oraloncology.2019.09.030.

References

[1] Compton C, Hess K, Halabi S, Balis U, Gershenwald J, Gimotty P, et al. Risk models for individualized prognosis in the practice of precision oncology. AJCC Cancer Staging Manual 8ed. Springer International Publishing; 2017. p. 47–52.
[2] Prabhu RS, Magliocca KR, Hanasoge S, Aiken AH, Hudgins PA, Hall WA, et al. Accuracy of computed tomography for predicting pathologic nodal extracapsular extension in patients with head-and-neck cancer undergoing initial surgical resection. Int J Radiat Oncol Biol Phys 2014;88:122–9.
[3] Geltzeiler M, Clayburgh D, Gleysteen J, Gross ND, Hamilton B, Andersen P, et al. Predictors of extracapsular extension in HPV-associated oropharyngeal cancer treated surgically. Oral Oncol 2017;65:89–93.
[4] Almulla A, Noel CW, Lu L, Xu W, O'Sullivan B, Goldstein DP, et al. Radiologic-pathologic correlation of extranodal extension in patients with squamous cell carcinoma of the oral cavity: implications for future editions of the TNM classification. Int J Radiat Oncol Biol Phys 2018;102:698–708.
[5] Billfalk-Kelly A, Yu E, Su J, O'Sullivan B, Waldron J, Ringash J, et al. Radiologic extranodal extension portends worse outcome in cN+ TNM-8 stage I human papillomavirus-mediated oropharyngeal cancer. Int J Radiat Oncol Biol Phys 2019.
[6] Kaka AS, Kumar B, Kumar P, Wakely Jr. PE, Kirsch CM, Old MO, et al. Highly aggressive human papillomavirus-related oropharyngeal cancer: clinical, radiologic, and pathologic characteristics. Oral Surg, Oral Med, Oral Pathol Oral Radiol 2013;116:327–35.
[7] Mao YP, Liang SB, Liu LZ, Chen Y, Sun Y, Tang LL, et al. The N staging system in nasopharyngeal carcinoma with radiation therapy oncology group guidelines for lymph node levels based on magnetic resonance imaging. Clin Cancer Res: An Off J Am Assoc Cancer Res 2008;14:7497–503.
[8] Guo Q, Pan J, Zong J, Zheng W, Zhang C, Tang L, et al. Suggestions for lymph node classification of UICC/AJCC staging system: a retrospective study based on 1197 nasopharyngeal carcinoma patients treated with intensity-modulated radiation therapy. Medicine 2015;94:e808.
[9] Li WF, Sun Y, Mao YP, Chen L, Chen YY, Chen M, et al. Proposed lymph node staging system using the International Consensus Guidelines for lymph node levels is predictive for nasopharyngeal carcinoma patients from endemic areas treated with intensity modulated radiation therapy. Int J Radiat Oncol Biol Phys 2013;86:249–56.
[10] Ai QY, King AD, Poon DMC, Mo FKF, Hui EP, Tong M, et al. Extranodal extension is a criterion for poor outcome in patients with metastatic nodes from cancer of the nasopharynx. Oral Oncol 2019;88:124–30.
[11] Hu Y, Lu T, Huang SH, Lin S, Chen Y, Fang Y, et al. High-grade radiologic extranodal extension predicts distant metastasis in stage II nasopharyngeal carcinoma. Head Neck 2019;41:3317–27.
[12] Pan JJ, Ng WT, Zong JF, Chan LL, O'Sullivan B, Lin SJ, et al. Proposal for the 8th edition of the AJCC/UICC staging system for nasopharyngeal cancer in the era of intensity-modulated radiotherapy. Cancer 2016;122:546–58.
[13] Xiao Y, Pan J, Chen Y, Lin S, Chen Y, Zong J, et al. Prognostic value of MRI-derived masticator space involvement in IMRT-treated nasopharyngeal carcinoma patients. Rad Oncol (London, England) 2015;10:204.

[14] Lin S, Pan J, Han L, Zhang X, Liao X, Lu JJ. Nasopharyngeal carcinoma treated with reduced-volume intensity-modulated radiation therapy: report on the 3-year outcome of a prospective series. Int J Rad Oncol Biol Phys 2009;75:1071–8.
[15] Akaike H. A new look at the statistical model identification. IEEE Trans Auto Control 1974;19:716–23.
[16] Kann BH, Buckstein M, Carpenter TJ, Bakst R, Misiukiewicz K, Genden E, et al. Radiographic extracapsular extension and treatment outcomes in locally advanced oropharyngeal carcinoma. Head Neck 2014;36:1689–94.
[17] Liu JT, Kann BH, De B, Buckstein M, Bakst RL, Genden EM, et al. Prognostic value of radiographic extracapsular extension in locally advanced head and neck squamous cell cancers. Oral Oncol 2016;52:52–7.
[18] Spector ME, Gallagher KK, Light E, Ibrahim M, Chanowski EJ, Moyer JS, et al. Matted nodes: poor prognostic marker in oropharyngeal squamous cell carcinoma independent of HPV and EGFR status. Head Neck 2012;34:1727–33.
[19] de Bree R, Ljumanovic R, Hazewinkel MJ, Witte BI, Castelijns JA. Radiologic extranodal spread and matted nodes: Important predictive factors for development of distant metastases in patients with high-risk head and neck cancer. Head Neck 2016;38(Suppl 1):E1452–8.
[20] Spector ME, Chinn SB, Bellile E, Gallagher KK, Ibrahim M, Vainshtein J, et al. Matted nodes as a predictor of distant metastasis in advanced-stage III/IV oropharyngeal squamous cell carcinoma. Head Neck 2016;38:184–90.
[21] Lee AW, Ma BB, Ng WT, Chan AT. Management of nasopharyngeal carcinoma: current practice and future perspective. J Clin Oncol: Off J Am Soc Clin Oncol 2015;33:3356–64.
[22] Lee AW, Lin JC, Ng WT. Current management of nasopharyngeal cancer. Sem Radiat Oncol 2012;22:233–44.
[23] Lee VH, Kwong DL, Leung TW, Choi CW, O'Sullivan B, Lam KO, et al. The addition of pretreatment plasma Epstein-Barr virus DNA into the 8th edition of nasopharyngeal cancer TNM stage classification. Int J Cancer 2018.
[24] Guo R, Tang LL, Mao YP, Du XJ, Chen L, Zhang ZC, et al. Proposed modifications and incorporation of plasma Epstein-Barr virus DNA improve the TNM staging system for Epstein-Barr virus-related nasopharyngeal carcinoma. Cancer 2019;125:79–89.
[25] Kim KY, Le QT, Yom SS, Ng RHW, Chan KCA, Bratman SV, et al. Clinical utility of Epstein-Barr Virus DNA testing in the treatment of nasopharyngeal carcinoma patients. Int J Radiat Oncol Biol Phys 2017;98:996–1001.