Original Article

Development and Validation of Web-Based Nomograms to Precisely Predict Survival Outcomes of Non-metastatic Nasopharyngeal Carcinoma in an Endemic Area

Ji-Jin Yao[1,2], **Li Lin**[1], **Tian-Sheng Gao**[3], **Wang-Jian Zhang**[4], **Wayne R. Lawrence**[4], **Jun Ma**[1], **Ying Sun**[1]

[1]Department of Radiation Oncology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Guangzhou, [2]Department of Head and Neck Oncology, The Cancer Center of the Fifth Affiliated Hospital of Sun Yat-sen University, Zhuhai, [3]Department of Radiation Oncology, Wuzhou Red Cross Hospital, Wuzhou, China, [4]Department of Environmental Health Sciences, School of Public Health, University at Albany, State University of New York, Rensselaer, NY, USA

Purpose This study aimed to develop web-based nomograms to precisely predict survival outcomes in patients with non-metastatic nasopharyngeal carcinoma (NPC) in an endemic area.

Materials and Methods A total of 10,126 patients who underwent radical intensity-modulated radiotherapy at Sun Yat-sen University Cancer Center (SYSUCC) from 2009 to 2015 were analyzed. We assigned patients into a training cohort (SYSUCC-A, n=6,751) and an internal validation cohort (SYSUCC-B, n=3,375) based on computer-generated random numbers. Patients collected from Wuzhou Red Cross Hospital (WZRCH) between 2012 and 2015 were used as the independent external validation cohort (WZRCH, n=450). Concordance index (C-index) was used to determine predictive accuracy and discriminative ability for the nomogram. The web-based clinicopathologic prediction models for predicting survival were based on Cox regression.

Results The C-indexes for SYSUCC-A, SYSUCC-B, and WZRCH cohorts for the established nomograms to predict 3-year overall survival (OS) was 0.736, 0.715, and 0.691. Additionally, C-indexes to predict 3-year distant metastasis-free survival (DMFS) was 0.717, 0.706, and 0.686, disease-free survival (DFS) was 0.713, 0.697, and 0.656, local relapse-free survival was 0.695, 0.684, and 0.652, and regional relapse-free survival was 0.672, 0.650, and 0.616. The calibration plots showed great agreement between nomogram-predicted 3-year survival outcomes and actual 3-year survival outcomes. Moreover, C-indexes of the nomograms for OS, DMFS, and DFS were significantly superior than TNM stage (p < 0.001 for all).

Conclusion These user-friendly nomograms can precisely predict survival endpoints in patients with non-metastatic NPC. They may serve as a useful tool for providing patient counseling and help physicians to make individual follow-up plans.

Key words Nasopharyngeal carcinoma, Survival outcomes, Web-based nomograms, Validation, Intensity-modulated radiotherapy

Introduction

Nasopharyngeal carcinoma (NPC) is highly prevalent in Southeast Asia, accounting for over 50% of new cases worldwide [1]. Primary treatment for NPC is radiotherapy (RT). The American Joint Committee on Cancer (AJCC) TNM staging system is generally used for developing treatment strategies and predicting clinical outcomes for NPC patients. However, clinical differences still occur among patients with the same TNM stage that have undergone similar treatment strategies. These differences indicate that the current staging system is inadequate for predicting prognosis [2].

There has been an emergence of studies utilizing nomogram to graphically depict predictive statistical models for patients diagnosed with various cancers [3-5]. Numerous studies have confirmed the feasibility of nomogram for predicting NPC survival [6,7], where they consistently demonstrated this model to be better than conventional TNM staging system. Though nomograms have better predictive power, there are also limitations such as obsolete RT techniques [8], lack of validation [7,9], using current obsolete editions of the AJCC staging system [6], and absence of publicly available online nomograms.

To fill current gaps in knowledge, we conceived and initiated a large-scale real-world study to comprehensively investigate the survival outcomes and prognostic nomograms in patients with NPC under modern RT. Moreover, user-friendly web-based nomograms were developed to predict survival

outcomes, provide patient counseling, and help physicians make individual follow-up plans.

Materials and Methods

1. Patient cohort

Ten thousand one hundred and twenty-six patients with histologically proven NPC diagnosed between April 2009 and December 2015 were consecutively extracted from the well-established big-data intelligence platform at Sun Yat-sen University Cancer Centre (SYSUCC). The training (SYSUCC-A, n=6,751) cohort and internal validation (SYSUCC-B, n=3,375) cohort were determined by assigned computer-generated random numbers. Patients from Wuzhou Red Cross Hospital (WZRCH) between February 2012 and July 2015 were the independent external validation cohort (WZRCH, n=450). Patients were enrolled in this study if they met the following inclusion criteria: (1) histologic diagnosis of NPC; (2) no distant metastasis ascertained by whole body bone scan, computed tomography (CT) or 18F-fluorodeoxyglucose positron emission tomography (PET)-CT; and (3) treated with intensity-modulated RT (IMRT). Patients were excluded if medical record indicated synchronous malignancies, cardiac disease, or uncontrolled renal disease that required treatment.

Patients were restaged based on the 8th edition of the AJCC staging system. Two radiologists with at least 10 years of experience were selected to independently review all imaging data to minimize heterogeneity in restaging.

2. Treatment protocol

The planning protocol for IMRT for both institutions were previously described [10]. Doses prescribed were 66 to 72 Gy/28-33 fractions to the planning target volume (PTV) of the primary gross tumor volume (GTVnx); 64 to 70 Gy/28-33 fractions to the PTV of the GTV for the involved lymph nodes (GTVnd); 60 to 63 Gy/28-33 fractions to the PTV of the high-risk clinical target volume (CTV1); and 54 to 56 Gy/28-33 fractions for the PTV of the low-risk clinical target volume (CTV2). Dose constraints to the critical organs at risk were as described by the Radiation Therapy Oncology Group (RTOG)-0225 trial. Patients adhered to the routine scheduled established for treatment (one fraction daily for 5 days each week).

During the study period, institutional guidelines recommended no chemotherapy for stage I, concurrent chemotherapy for stage II, and concurrent chemotherapy +/− induction/adjuvant chemotherapy stage III to IVB NPC, as defined by the 7th edition of the AJCC staging system. Concurrent chemotherapy involved cisplatin (80 or 100 mg/

m²) provided at week 1, 4, and 7 of RT, or cisplatin (40 mg/m²) provided weekly during RT starting on day 1 of RT. The induction or adjuvant chemotherapy included cisplatin (60 mg/m²), docetaxel (60 mg/m²), and 5-fluorouracil (600 mg/m²/day over 120 hours), or cisplatin (80 mg/m²) plus 5-fluorouracil (800 mg/m²/day over 120 hours), or cisplatin (80 mg/m²) plus docetaxel (80 mg/m²), or gemcitabine (1 g/m², on day 1, 8) plus cisplatin (80 mg/m², on day 1) every third week for three cycles.

3. Data sharing

The authenticity of this study has been validated by uploading the key raw data onto the Research Data Deposit (RDD) public platform http://www.researchdata.org.cn, and the approval RDD number is RDDA2019001315.

4. Statistical analysis

Clinical and demographic characteristics were obtained for each study cohort reported. Continuous variables (age, hemoglobin [HGB, g/L], and albumin [ALB, g/L] levels) were converted to categorical variables, determined by routine cutoff points in clinical application. Chi-square test and Fisher exact test were used for comparing differences between groups for clinical and demographic characteristics. Our primary endpoint was overall survival (OS), which was calculated from the date of initial treatment to date of death from any cause. Secondary endpoints were distant metastasis-free survival (DMFS), disease-free survival (DFS), local relapse-free survival (LRFS), and regional relapse-free survival (RRFS). DMFS was calculated from the date of first distant relapse; LRFS was calculated from date of first regional relapse; RRFS was calculated from the date of first regional relapse; and DFS was calculated from the first relapse at any site, death from any cause, or date of last follow-up visit, whichever first occurred.

Kaplan-Meier and log-rank test were used to depict survival curves. Regarding the development of prediction models, we first developed univariable Cox regression models to identify the potentially important predictors (p < 0.1), then developed a multivariable Cox regression model for each outcome (i.e., the endpoint and the time to the event) based on the selected predictors. Following the methods described in previous studies [11], the proportional hazards assumption for the multivariable Cox regression models was examined. Time-dependent variables or sensitivity analysis were conducted in the case where the assumption was not satisfied. Finally, we used the nomogram function in the R package *rms* [12] to display the multivariable models as nomograms. Concordance index (C-index) was used to assess the performance of the nomogram. Moreover, C-index and calibration curves derived from regression analysis. In

Table 1. Patient demographics and clinical characteristics

Characteristic	Entire cohort (n=10,126)	Training cohort (n=6,751)	Internal validation cohort (n=3,375)	External validation cohort (n=450)	p-value[a]
Age (yr)					
≤ 29	848 (9.3)	571 (9.5)	277 (9.1)	30 (6.7)	0.121
30-39	2,349 (25.9)	1,549 (25.7)	800 (26.2)	95 (21.1)	
40-49	3,492 (38.4)	2,322 (38.5)	1,170 (38.3)	195 (43.3)	
50-59	1,255 (13.8)	823 (13.6)	432 (14.2)	65 (14.4)	
≥ 60	1,143 (12.6)	770 (12.8)	373 (12.2)	65 (14.4)	
Sex					
Male	7,440 (73.5)	4,992 (73.9)	2,448 (72.5)	322 (71.6)	0.208
Female	2,686 (26.5)	1,759 (26.1)	927 (27.5)	128 (28.4)	
Histology (WHO)					
Type I	66 (0.7)	43 (0.6)	23 (0.7)	5 (1.1)	0.005
Type II	205 (2.0)	144 (2.1)	61 (1.8)	21 (4.7)	
Type III	9,855 (97.3)	6,564 (97.2)	3,291 (97.5)	424 (94.2)	
T category[b]					
T1	1,674 (16.5)	1,103 (16.3)	571 (16.9)	77 (17.1)	< 0.001
T2	1,633 (16.1)	1,084 (16.1)	549 (16.3)	96 (21.3)	
T3	4,702 (46.4)	3,145 (46.6)	1,557 (46.1)	147 (32.7)	
T4	2,117 (20.9)	1,419 (21.0)	698 (20.7)	130 (28.9)	
N category[b]					
N0	1,586 (15.7)	1,066 (15.8)	520 (15.4)	40 (8.9)	< 0.001
N1	5,133 (50.7)	3,412 (50.5)	1,721 (51.0)	220 (48.9)	
N2	2,171 (21.4)	1,449 (21.5)	722 (21.4)	124 (27.6)	
N3	1,236 (12.2)	824 (12.2)	412 (12.2)	66 (14.7)	
Clinical stage[b]					
I	558 (5.5)	382 (5.7)	176 (5.2)	18 (4.0)	0.001
II	1,788 (17.7)	1,161 (17.2)	627 (18.6)	89 (19.8)	
III	4,672 (46.1)	3,132 (46.4)	1,540 (45.6)	168 (37.3)	
IVA	3,108 (30.7)	2,076 (30.8)	1,032 (30.6)	175 (38.9)	
Treatment					
IMRT alone	1,035 (10.2)	684 (10.1)	351 (10.4)	34 (7.6)	0.121
CCRT	3,752 (37.1)	2,463 (36.5)	1,289 (38.2)	154 (34.2)	
IC+CCRT	3,857 (38.1)	2,597 (38.5)	1,260 (37.3)	198 (44.0)	
CCRT+AC	463 (4.6)	315 (4.7)	148 (4.4)	24 (5.3)	
IC+IMRT	1,019 (10.1)	692 (10.3)	327 (9.7)	40 (8.9)	
HGB (g/L)					
< 113	334 (3.4)	215 (3.3)	119 (3.7)	29 (6.4)	< 0.001
113-151	6,290 (64.6)	4,222 (65.0)	2,068 (63.7)	303 (67.7)	
≥ 151	3,115 (32.0)	2,055 (31.7)	1,060 (32.6)	118 (26.2)	
Unknown	387 (3.8)	259 (3.8)	128 (3.8)	0	
ALB (g/L)					
< 40	856 (8.6)	589 (8.9)	267 (8.0)	59 (13.1)	< 0.001
≥ 40	9,084 (91.4)	6,027 (91.1)	3,057 (92.0)	391 (86.9)	
Unknown	186 (1.8)	135 (2.0)	51 (1.5)	0	
Family of cancer					
0	7,444 (73.5)	4,979 (73.8)	2,465 (73.0)	357 (79.3)	0.015
1	2,682 (26.5)	1,772 (26.2)	910 (27.0)	93 (20.7)	

(Continued to the next page)

Table 1. Continued

Characteristic	Entire cohort (n=10,126)	Training cohort (n=6,751)	Internal validation cohort (n=3,375)	External validation cohort (n=450)	p-value[a)]
Smoking history					
No	6,575 (64.9)	4,383 (64.9)	2,192 (64.9)	273 (60.7)	0.179
Yes	3,551 (35.1)	2,368 (35.1)	1,183 (35.1)	177 (39.3)	
Drinking history					
No	8,732 (86.2)	5,807 (86.0)	2,925 (86.7)	388 (86.2)	0.675
Yes	1,394 (13.8)	944 (14.0)	450 (13.3)	62 (13.8)	

Values are presented as number (%). AC, adjuvant chemotherapy; ALB, albumin; CCRT, concurrent chemoradiotherapy; HGB, hemoglobin; IC, induction chemotherapy; IMRT, intensity-modulated radiotherapy; WHO, World Health Organization. [a)]p-values were derived from comparisons among training, internal and external validation cohorts, [b)]According to the 8th edition of the American Joint Committee on Cancer staging system.

Table 2. The number of events (e.g., death, distant metastasis, local recurrence, regional recurrence, and disease recurrence) in the training, internal, and external validation cohort

Characteristic	Training	Internal testing	External testing
Death	899 (13.3)	425 (12.6)	44 (13.1)
Distant metastasis	858 (12.7)	398 (11.8)	44 (12.4)
Local recurrence	406 (6.0)	221 (6.5)	16 (6.2)
Regional recurrence	344 (5.1)	182 (5.4)	11 (5.2)
Disease recurrence	1,443 (21.4)	712 (21.1)	77 (21.3)

Values are presented as number (%).

addition, we analyzed two data cohorts (SYSUCC-B and WZRCH cohorts) within the same study period allowing for the nomogram based on SYSUCC-A to be validated by data from another institute for generalizability and robustness. Comparison of the C-index between the 8th AJCC staging system and nomograms was conducted using the compareC package in R [13]. All statistical tests were two-sided and p < 0.05 was considered statistically significant. Statistical models were performed using the rms package in R ver. 3.4.2 software (The R Foundation for Statistical Computing, Vienna, Austria; https://www.rproject. org).

Results

1. Study population and follow-up

Table 1 presents patient characteristics from both the validation and training cohorts. Generally, WZRCH had more N2-3 category, T4 and IVA stage disease, World Health Organization (WHO) type II disease, less family history of cancer, and reduced HGB and ALB levels than SYSUCC-A and SYSUCC-B. No statistically significant difference was observed for age (p=0.121), sex (p=0.208), smoking history (p=0.179), history of alcohol consumption (p=0.675), and

treatment strategy (p=0.121) among SYSUCC-A, SYSUCC-B, and WZRCH cohorts (Table 1). Univariate analysis for the prognostic factors of OS, DMFS, DFS, LRFS, and RRFS are listed in S1 Table. The median follow-up time for all patients was 54.3 months (interquartile range [IQR], 41.8 to 70.0), with 54.0 months (IQR, 41.4 to 70.0) for the training set and 54.4 months (IQR, 42.0 to 70.0) for the testing set. The median follow-up time for the external validation cohort was 46.5 months (IQR, 37.9 to 52.4). Additionally, the number of events (e.g., death, distant metastasis, local recurrence, regional recurrence, and disease recurrence) in the training, internal and external validation cohort were showed in Table 2.

2. Overall survival

Overall, 1,324 patients (1,324/10,126, 13%) died in the Guangzhou cohort, and the 1-, 3-, and 5-year OS rates were 99.2%, 92.3%, and 86.5%, respectively (Fig. 1A). Mortality risk significantly rose in accordance with clinical stage increase from stage I to stage IVA (p < 0.001 for all), with corresponding 5-year OS rates of 99.5%, 92.5%, 88.6%, and 77.2%, respectively (Fig. 2A). Although patients with T4 disease (77.8%) had a significantly lower 5-year OS rate than T1 (94.4%), T2 (88.8%), and T3 (86.7%) disease (p < 0.001 for all), the OS curves for T2 and T3 disease almost overlapped in the
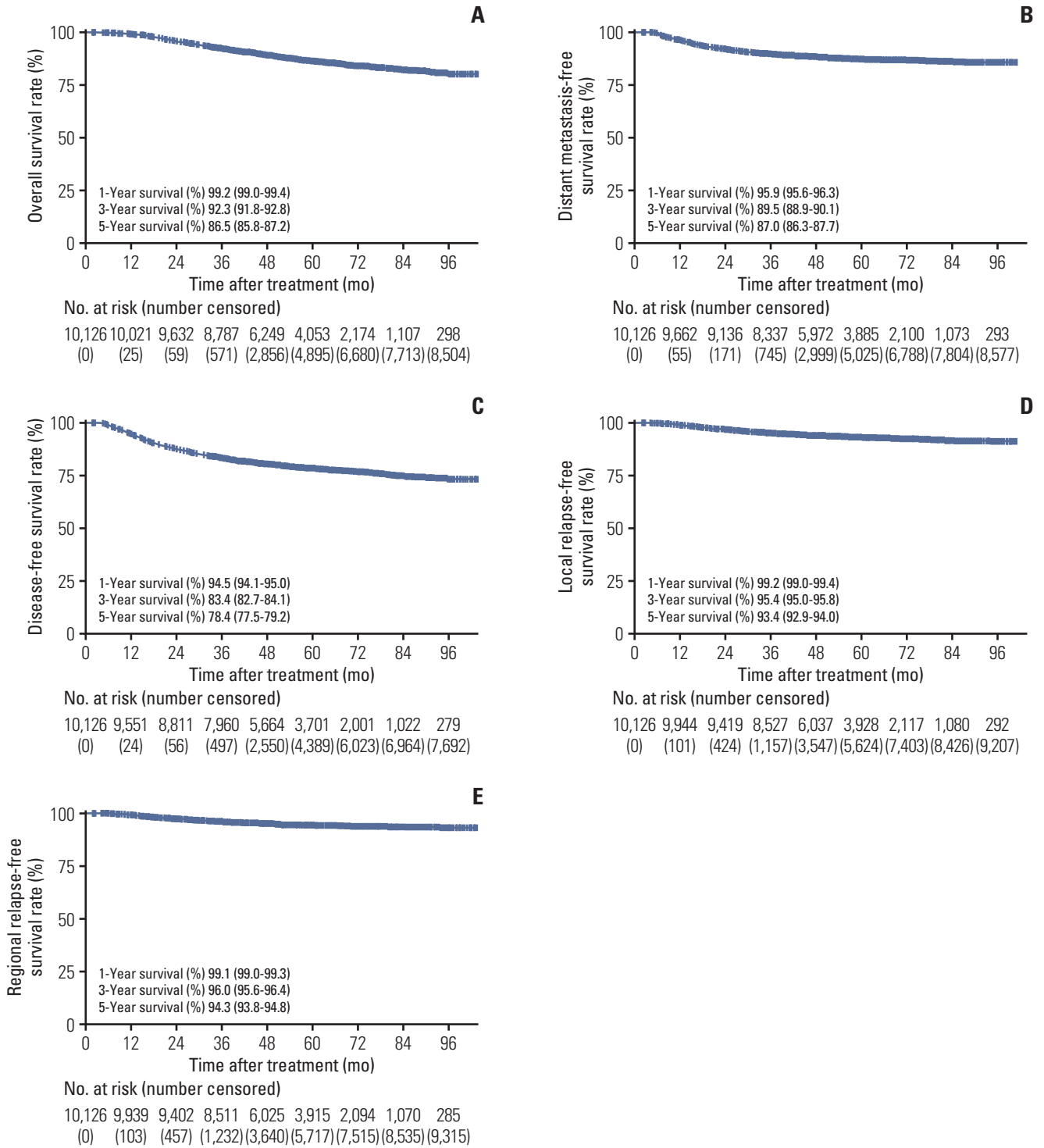
Fig. 1. Kaplan-Meier curves of 10,126 nasopharyngeal carcinoma patients for 1-, 3-, and 5-year overall survival (A), distant metastasis-free survival (B), disease-free survival (C), local relapse-free survival (D), and regional relapse-free survival (E).
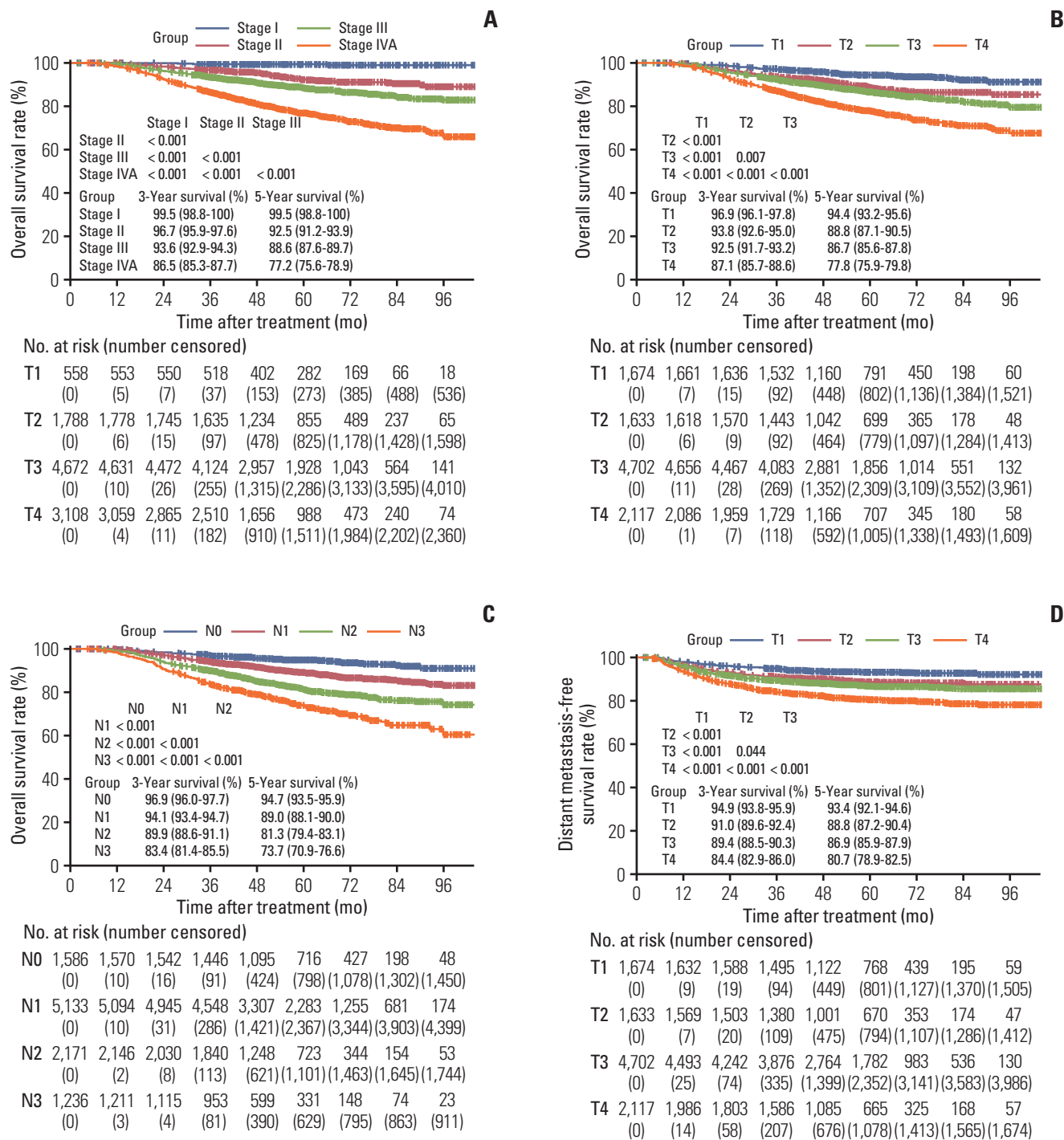
A

Overall survival rate (%) vs Time after treatment (mo)

Group — Stage I — Stage III — Stage II — Stage IVA

	Stage I	Stage II	Stage III
Stage II	< 0.001		
Stage III	< 0.001	< 0.001	
Stage IVA	< 0.001	< 0.001	< 0.001

Group	3-Year survival (%)	5-Year survival (%)
Stage I	99.5 (98.8-100)	99.5 (98.8-100)
Stage II	96.7 (95.9-97.6)	92.5 (91.2-93.9)
Stage III	93.6 (92.9-94.3)	88.6 (87.6-89.7)
Stage IVA	86.5 (85.3-87.7)	77.2 (75.6-78.9)

No. at risk (number censored)

T1	558	553	550	518	402	282	169	66	18
	(0)	(5)	(7)	(37)	(153)	(273)	(385)	(488)	(536)
T2	1,788	1,778	1,745	1,635	1,234	855	489	237	65
	(0)	(6)	(15)	(97)	(478)	(825)	(1,178)	(1,428)	(1,598)
T3	4,672	4,631	4,472	4,124	2,957	1,928	1,043	564	141
	(0)	(10)	(26)	(255)	(1,315)	(2,286)	(3,133)	(3,595)	(4,010)
T4	3,108	3,059	2,865	2,510	1,656	988	473	240	74
	(0)	(4)	(11)	(182)	(910)	(1,511)	(1,984)	(2,202)	(2,360)

B

Overall survival rate (%) vs Time after treatment (mo)

Group — T1 — T2 — T3 — T4

	T1	T2	T3
T2	< 0.001		
T3	< 0.001	0.007	
T4	< 0.001	< 0.001	< 0.001

Group	3-Year survival (%)	5-Year survival (%)
T1	96.9 (96.1-97.8)	94.4 (93.2-95.6)
T2	93.8 (92.6-95.0)	88.8 (87.1-90.5)
T3	92.5 (91.7-93.2)	86.7 (85.6-87.8)
T4	87.1 (85.7-88.6)	77.8 (75.9-79.8)

No. at risk (number censored)

T1	1,674	1,661	1,636	1,532	1,160	791	450	198	60
	(0)	(7)	(15)	(92)	(448)	(802)	(1,136)	(1,384)	(1,521)
T2	1,633	1,618	1,570	1,443	1,042	699	365	178	48
	(0)	(6)	(9)	(92)	(464)	(779)	(1,097)	(1,284)	(1,413)
T3	4,702	4,656	4,467	4,083	2,881	1,856	1,014	551	132
	(0)	(11)	(28)	(269)	(1,352)	(2,309)	(3,109)	(3,552)	(3,961)
T4	2,117	2,086	1,959	1,729	1,166	707	345	180	58
	(0)	(1)	(7)	(118)	(592)	(1,005)	(1,338)	(1,493)	(1,609)

C

Overall survival rate (%) vs Time after treatment (mo)

Group — N0 — N1 — N2 — N3

	N0	N1	N2
N1	< 0.001		
N2	< 0.001	< 0.001	
N3	< 0.001	< 0.001	< 0.001

Group	3-Year survival (%)	5-Year survival (%)
N0	96.9 (96.0-97.7)	94.7 (93.5-95.9)
N1	94.1 (93.4-94.7)	89.0 (88.1-90.0)
N2	89.9 (88.6-91.1)	81.3 (79.4-83.1)
N3	83.4 (81.4-85.5)	73.7 (70.9-76.6)

No. at risk (number censored)

N0	1,586	1,570	1,542	1,446	1,095	716	427	198	48
	(0)	(10)	(16)	(91)	(424)	(798)	(1,078)	(1,302)	(1,450)
N1	5,133	5,094	4,945	4,548	3,307	2,283	1,255	681	174
	(0)	(10)	(31)	(286)	(1,421)	(2,367)	(3,344)	(3,903)	(4,399)
N2	2,171	2,146	2,030	1,840	1,248	723	344	154	53
	(0)	(2)	(8)	(113)	(621)	(1,101)	(1,463)	(1,645)	(1,744)
N3	1,236	1,211	1,115	953	599	331	148	74	23
	(0)	(3)	(4)	(81)	(390)	(629)	(795)	(863)	(911)

D

Distant metastasis-free survival rate (%) vs Time after treatment (mo)

Group — T1 — T2 — T3 — T4

	T1	T2	T3
T2	< 0.001		
T3	< 0.001	0.044	
T4	< 0.001	< 0.001	< 0.001

Group	3-Year survival (%)	5-Year survival (%)
T1	94.9 (93.8-95.9)	93.4 (92.1-94.6)
T2	91.0 (89.6-92.4)	88.8 (87.2-90.4)
T3	89.4 (88.5-90.3)	86.9 (85.9-87.9)
T4	84.4 (82.9-86.0)	80.7 (78.9-82.5)

No. at risk (number censored)

T1	1,674	1,632	1,588	1,495	1,122	768	439	195	59
	(0)	(9)	(19)	(94)	(449)	(801)	(1,127)	(1,370)	(1,505)
T2	1,633	1,569	1,503	1,380	1,001	670	353	174	47
	(0)	(7)	(20)	(109)	(475)	(794)	(1,107)	(1,286)	(1,412)
T3	4,702	4,493	4,242	3,876	2,764	1,782	983	536	130
	(0)	(25)	(74)	(335)	(1,399)	(2,352)	(3,141)	(3,583)	(3,986)
T4	2,117	1,986	1,803	1,586	1,085	665	325	168	57
	(0)	(14)	(58)	(207)	(676)	(1,078)	(1,413)	(1,565)	(1,674)

Fig. 2. Kaplan-Meier curves for 10,126 nasopharyngeal carcinoma patients receiving intensity-modulated radiotherapy: overall survival stratified by clinical stage (A), T category (B), and N category (C); distant metastasis-free survival T category (D), clinical stage (E), and N category (F); disease-free survival stratified by T category (G), clinical stage (H), and N category (I). *(Continued to the next page)*

first 6 years after treatment (p=0.007) (Fig. 2B). In comparison with T category and clinical stage, we found that N category presented better discrimination in OS, where the correspond-ing 5-year OS rates for N0, N1, N2, and N3 categories were 94.7%, 89.0%, 81.3%, and 73.7%, respectively (p < 0.001) (Fig. 2C). Significant predictors for OS in the multivariable analy-
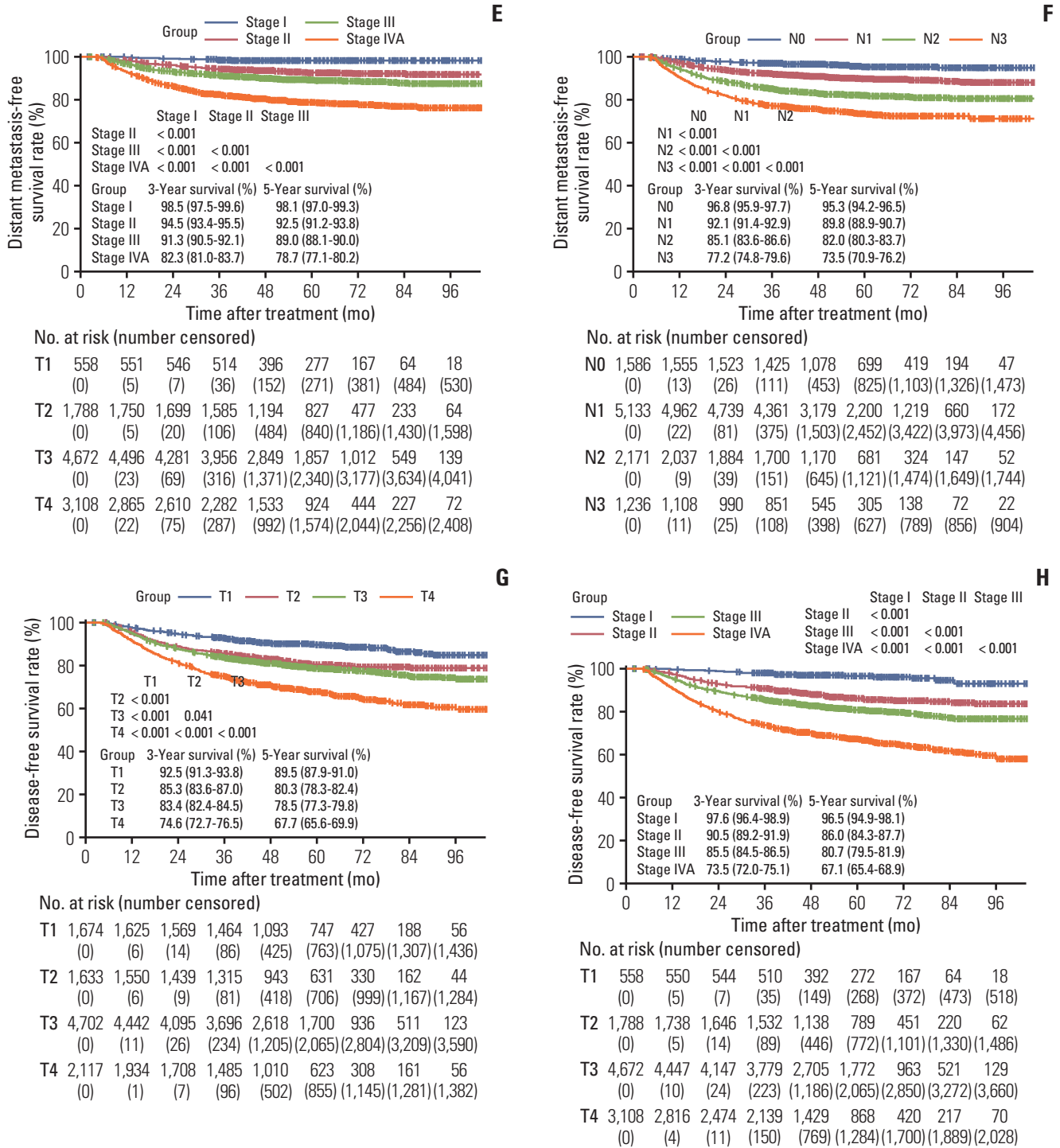
E

Distant metastasis-free survival rate (%)

Group — Stage I — Stage III — Stage II — Stage IVA

	Stage I	Stage II	Stage III
Stage II	< 0.001		
Stage III	< 0.001	< 0.001	
Stage IVA	< 0.001	< 0.001	< 0.001

Group	3-Year survival (%)	5-Year survival (%)
Stage I	98.5 (97.5-99.6)	98.1 (97.0-99.3)
Stage II	94.5 (93.4-95.5)	92.5 (91.2-93.8)
Stage III	91.3 (90.5-92.1)	89.0 (88.1-90.0)
Stage IVA	82.3 (81.0-83.7)	78.7 (77.1-80.2)

Time after treatment (mo)

No. at risk (number censored)

	0	12	24	36	48	60	72	84	96
T1	558	551	546	514	396	277	167	64	18
	(0)	(5)	(7)	(36)	(152)	(271)	(381)	(484)	(530)
T2	1,788	1,750	1,699	1,585	1,194	827	477	233	64
	(0)	(5)	(20)	(106)	(484)	(840)	(1,186)	(1,430)	(1,598)
T3	4,672	4,496	4,281	3,956	2,849	1,857	1,012	549	139
	(0)	(23)	(69)	(316)	(1,371)	(2,340)	(3,177)	(3,634)	(4,041)
T4	3,108	2,865	2,610	2,282	1,533	924	444	227	72
	(0)	(22)	(75)	(287)	(992)	(1,574)	(2,044)	(2,256)	(2,408)

F

Distant metastasis-free survival rate (%)

Group — N0 — N1 — N2 — N3

	N0	N1	N2
N1	< 0.001		
N2	< 0.001	< 0.001	
N3	< 0.001	< 0.001	< 0.001

Group	3-Year survival (%)	5-Year survival (%)
N0	96.8 (95.9-97.7)	95.3 (94.2-96.5)
N1	92.1 (91.4-92.9)	89.8 (88.9-90.7)
N2	85.1 (83.6-86.6)	82.0 (80.3-83.7)
N3	77.2 (74.8-79.6)	73.5 (70.9-76.2)

Time after treatment (mo)

No. at risk (number censored)

	0	12	24	36	48	60	72	84	96
N0	1,586	1,555	1,523	1,425	1,078	699	419	194	47
	(0)	(13)	(26)	(111)	(453)	(825)	(1,103)	(1,326)	(1,473)
N1	5,133	4,962	4,739	4,361	3,179	2,200	1,219	660	172
	(0)	(22)	(81)	(375)	(1,503)	(2,452)	(3,422)	(3,973)	(4,456)
N2	2,171	2,037	1,884	1,700	1,170	681	324	147	52
	(0)	(9)	(39)	(151)	(645)	(1,121)	(1,474)	(1,649)	(1,744)
N3	1,236	1,108	990	851	545	305	138	72	22
	(0)	(11)	(25)	(108)	(398)	(627)	(789)	(856)	(904)

G

Disease-free survival rate (%)

Group — T1 — T2 — T3 — T4

	T1	T2	T3
T2	< 0.001		
T3	< 0.001	0.041	
T4	< 0.001	< 0.001	< 0.001

Group	3-Year survival (%)	5-Year survival (%)
T1	92.5 (91.3-93.8)	89.5 (87.9-91.0)
T2	85.3 (83.6-87.0)	80.3 (78.3-82.4)
T3	83.4 (82.4-84.5)	78.5 (77.3-79.8)
T4	74.6 (72.7-76.5)	67.7 (65.6-69.9)

Time after treatment (mo)

No. at risk (number censored)

	0	12	24	36	48	60	72	84	96
T1	1,674	1,625	1,569	1,464	1,093	747	427	188	56
	(0)	(6)	(14)	(86)	(425)	(763)	(1,075)	(1,307)	(1,436)
T2	1,633	1,550	1,439	1,315	943	631	330	162	44
	(0)	(6)	(9)	(81)	(418)	(706)	(999)	(1,167)	(1,284)
T3	4,702	4,442	4,095	3,696	2,618	1,700	936	511	123
	(0)	(11)	(26)	(234)	(1,205)	(2,065)	(2,804)	(3,209)	(3,590)
T4	2,117	1,934	1,708	1,485	1,010	623	308	161	56
	(0)	(1)	(7)	(96)	(502)	(855)	(1,145)	(1,281)	(1,382)

H

Disease-free survival rate (%)

Group — Stage I — Stage III — Stage II — Stage IVA

	Stage I	Stage II	Stage III
Stage II	< 0.001		
Stage III	< 0.001	< 0.001	
Stage IVA	< 0.001	< 0.001	< 0.001

Group	3-Year survival (%)	5-Year survival (%)
Stage I	97.6 (96.4-98.9)	96.5 (94.9-98.1)
Stage II	90.5 (89.2-91.9)	86.0 (84.3-87.7)
Stage III	85.5 (84.5-86.5)	80.7 (79.5-81.9)
Stage IVA	73.5 (72.0-75.1)	67.1 (65.4-68.9)

Time after treatment (mo)

No. at risk (number censored)

	0	12	24	36	48	60	72	84	96
T1	558	550	544	510	392	272	167	64	18
	(0)	(5)	(7)	(35)	(149)	(268)	(372)	(473)	(518)
T2	1,788	1,738	1,646	1,532	1,138	789	451	220	62
	(0)	(5)	(14)	(89)	(446)	(772)	(1,101)	(1,330)	(1,486)
T3	4,672	4,447	4,147	3,779	2,705	1,772	963	521	129
	(0)	(10)	(24)	(223)	(1,186)	(2,065)	(2,850)	(3,272)	(3,660)
T4	3,108	2,816	2,474	2,139	1,429	868	420	217	70
	(0)	(4)	(11)	(150)	(769)	(1,284)	(1,700)	(1,889)	(2,028)

Fig. 2. *(Continued from the previous page) (Continued to the next page)*

ses for the SYSUCC-A cohort were age, sex, T category, and N category (Table 3). A nomogram was constructed based on the weighting of four significant covariates for the SYSUCC-A cohort) (S2A Fig.), which generated a Harrell's C-index of 0.736 (95% confidence interval [CI], 0.690 to 0.783) for OS (S2B Fig.). The calibration plots for 3-year OS were predicted in the SYSUCC-B (C-index, 0.715; 95% CI, 0.676 to 0.751) (S2C Fig.) and WZRCH (C-index, 0.691; 95% CI, 0.557 to 0.836)
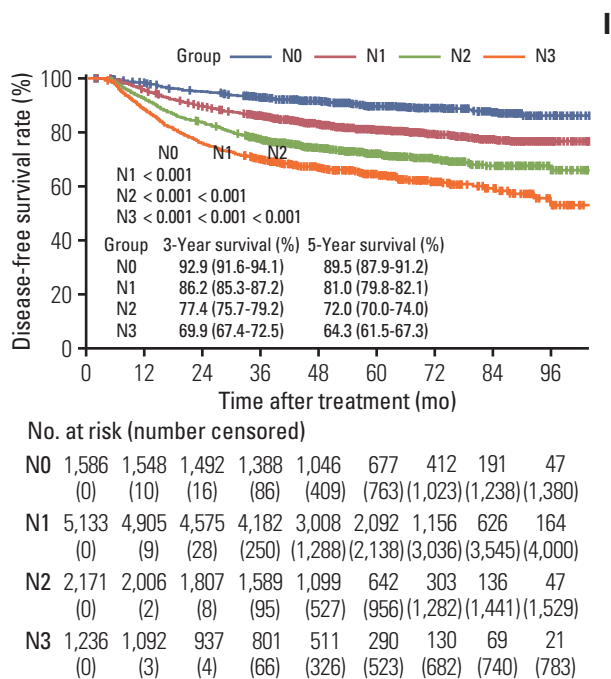
Fig. 2. *(Continued from the previous page)*

(S2D Fig.) cohorts. The nomogram for OS had a significantly higher C-index (0.736; 95% CI, 0.690 to 0.783) than T category (C-index, 0.606; 95% CI, 0.567 to 0.644; p < 0.001), N category (C-index, 0.626; 95% CI, 0.586 to 0.666; p < 0.001), and clinical stage (C-index=0.644; 95% CI, 0.608 to 0.680; p < 0.001).

3. Distant metastasis-free survival

One thousand two hundred fifty-six patients (1,256/10,126, 12%) suffered distant metastasis from the Guangzhou cohort. The 1-, 3-, and 5-year DMFS rates were 95.9%, 89.5%, and 87.0%, respectively (Fig. 1B). Although all T category, N category, and clinical stage presented an associated discrimination in DMFS (p < 0.05 for all) (Fig. 2D-F), the survival curves for different N category were separated better than T category and clinical stage. Among patients with N0, N1, N2, and N3 category, the corresponding 3- and 5-year DMFS rates were 96.8% and 95.3%, 92.1% and 89.8%, 85.1% and 82.0%, and 77.2% and 73.5%, respectively. In the SYSUCC-A cohort, independent risk factors for DMFS were sex, age, T category, and N category (Table 3). We then constructed a nomogram for DMFS that was based on the above four significant covariates in the SYSUCC-A cohort (S3A Fig.) yielding a C-index of 0.717 (95% CI, 0.665 to 0.768) (S3B Fig.). The calibration plots for 3-year DMFS predicted well in the SYSUCC-B (C-index, 0.706; 95% CI, 0.669 to 0.744) (S3C Fig.) and WZRCH (C-index, 0.686; 95% CI, 0.535 to 0.827) (S3D Fig.) cohorts. In addition, the nomogram for DMFS had a significantly higher

C-index (0.717; 95% CI, 0.665 to 0.768) than the T category (C-index, 0.584; 95% CI, 0.548 to 0.621; p < 0.001), N category (C-index, 0.651; 95% CI, 0.614 to 0.688; p < 0.001), and clinical stage (C-index, 0.630; 95% CI, 0.594 to 0.665; p < 0.001).

4. Disease-free survival

Among the Guangzhou cohort, 2,155 of 10,126 patients (21%) suffered disease failure, and the 1-, 3-, and 5-year RRFS rates were 94.5%, 83.4%, and 78.4%, respectively (Fig. 1C). The 1-, 3-, and 5-year rates for DFS stratified by T category, N category, and clinical stage are detailed in Fig. 2G-I. Although T4 disease (67.7%) had a significantly lower 5-year DFS rate in comparison with T1 (89.5%), T2 (80.3%), and T3 (78.5%) disease (p < 0.001), the DFS curves for T2 and T3 category almost overlapped in the first 6 years after treatment (p=0.041) (Fig. 2G). The risk of suffering disease failure significantly increased with clinical stage escalating from stage I to stage IVA (p < 0.001 for all) (Fig. 2H). When N category was analyzed as an ordinal variable (from N0 to N3), the cumulative survival curves presented an excellent discrimination in DFS (p < 0.001 for all), with corresponding 5-year DFS rates of 89.5%, 81.0%, 72.0%, and 64.3%, respectively (Fig. 2I). Additionally, we performed multivariable analysis to generate a nomogram to predict DFS in the SYSUCC-A cohort. Age, sex, smoking history, T category, and N category were associated with DFS (p < 0.05 for all) (Table 3). A nomogram based on the above five significant covariates was further constructed in the SYSUCC-A cohort (S4A Fig.). The calibration plots for the three-year DFS predicted well in SYSUCC-A (C-index, 0.713; 95% CI, 0.677 to 0.752) (S4B Fig.), SYSUCC-B (C-index, 0.695; 95% CI, 0.667 to 0.727) (S4C Fig.), and WZRCH (C-index, 0.656; 95% CI, 0.532 to 0.774) (S4D Fig.) cohorts. In addition, the nomogram for DFS had a significantly higher C-index (0.713; 95% CI, 0.677 to 0.752) than the T category (C-index, 0.587; 95% CI, 0.557 to 0.616; p < 0.001), N category (C-index, 0.616; 95% CI, 0.586 to 0.646; p < 0.001), and clinical stage (C-index, 0.618; 95% CI, 0.590 to 0.647; p < 0.001).

5. Local relapse-free survival

Six hundred twenty-seven patients (6%) experienced a local failure event in the Guangzhou cohort. The 1-, 3-, and 5-year LRFS rates were 99.2%, 95.4%, and 93.4%, respectively (Fig. 1D). Five-year LRFS rates for T1, T2, and T3 disease were 98.7%, 94.9%, and 93.7% (p < 0.05), respectively, and all significantly greater when compared with T4 disease (87.4%, p < 0.001 for all) (Fig. 3A). Most likely due to the difficultly of differentiating T2 and T3 disease, the LRFS curves for T2 and T3 disease almost overlapped in the first 6 years after treatment (p=0.047) (Fig. 3A). Additionally, in the SYSUCC-A cohort we performed multivariable analysis for LRFS to gen-

Table 3. Multivariable analysis of prognostic factors for NPC patients in the training cohort (n=6,751)

Endpoints	HR (95% CI)	p-value[a]
Overall survival		
Sex		
Female	Reference	
Male	1.33 (1.01-1.75)	0.040
Age (yr)		
≤ 30	Reference	
31-40	1.85 (1.10-3.10)	0.020
41-50	2.21 (1.35-3.63)	0.002
51-60	2.73 (1.61-4.64)	< 0.001
≥ 61	5.52 (3.33-9.16)	< 0.001
T category (8th edition)[b]		
T1	Reference	
T2	1.54 (0.96-2.48)	0.072
T3	2.08 (1.38-3.12)	< 0.001
T4	3.65 (2.40-5.56)	< 0.001
N category (8th edition)[b]		
N0	Reference	
N1	2.41 (1.53-3.80)	< 0.001
N2	3.65 (2.28-5.85)	< 0.001
N3	5.43 (3.35-8.80)	< 0.001
Distant metastasis-free survival		
Sex		
Female	Reference	
Male	1.53 (1.16-2.03)	0.003
Age (yr)		
≤ 30	Reference	
31-40	1.70 (1.06-2.70)	0.026
41-50	1.64 (1.04-2.58)	0.033
51-60	1.96 (1.19-3.20)	0.008
≥ 61	2.40 (1.47-3.92)	< 0.001
T ctaegory (8th edition)[b]		
T1	Reference	
T2	1.27 (0.82-1.95)	0.283
T3	1.62 (1.13-2.34)	0.009
T4	2.04 (1.38-3.01)	< 0.001
N category (8th edition)[b]		
N0	Reference	
N1	2.71 (1.66-4.43)	< 0.001
N2	4.28 (2.59-7.08)	< 0.001
N3	6.34 (3.80-10.60)	< 0.001
Disease-free survival		
Sex		
Female	Reference	
Male	1.32 (1.06-1.63)	0.011

(Continued)

Table 3. Continued

Endpoints	HR (95% CI)	p-value[a]
Age (yr)		
≤ 30	Reference	
31-40	1.40 (0.99-1.98)	0.060
41-50	1.51 (1.08-2.12)	0.015
51-60	1.79 (1.24-2.59)	0.002
≥ 61	2.56 (1.79-3.67)	< 0.001
Smoking history		
No	Reference	
Yes	1.24 (1.04-1.49)	0.017
T category (8th edition)[b]		
T1	Reference	
T2	1.62 (1.15-2.27)	0.005
T3	1.79 (1.33-2.40)	< 0.001
T4	2.64 (1.94-3.60)	< 0.001
N category (8th edition)[b]		
N0	Reference	
N1	2.44 (1.73-3.44)	< 0.001
N2	3.34 (2.33-4.78)	< 0.001
N3	4.55 (3.14-6.59)	< 0.001
Local relapse-free survival		
Smoking history		
No	Reference	
Yes	1.46 (1.10-1.93)	0.009
T category (8th edition)[b]		
T1	Reference	
T2	3.06 (1.38-6.80)	0.006
T3	3.73 (1.80-7.73)	< 0.001
T4	7.80 (3.75-16.23)	< 0.001
N category (8th edition)[b]		
N0	Reference	
N1	1.87 (1.07-3.28)	0.028
N2	2.28 (1.26-4.12)	0.006
N3	1.98 (1.02-3.84)	0.043
Regional relapse-free survival		
T category (8th edition)[b]		
T1	Reference	
T2	2.17 (1.24-3.77)	0.006
T3	1.37 (0.82-2.30)	0.232
T4	1.38 (0.78-2.46)	0.269
N category (8th edition)[b]		
N0	Reference	
N1	3.40 (1.56-7.39)	0.002
N2	5.22 (2.35-11.56)	< 0.001
N3	7.43 (3.31-16.69)	< 0.001

CI, confidence interval; HR, hazard ratio; NPC, nasopharyngeal carcinoma; WHO, World Health Organization. [a]p-values were calculated using an adjusted Cox proportional hazards model, [b]According to the 8th edition of the American Joint Committee on Cancer staging system.
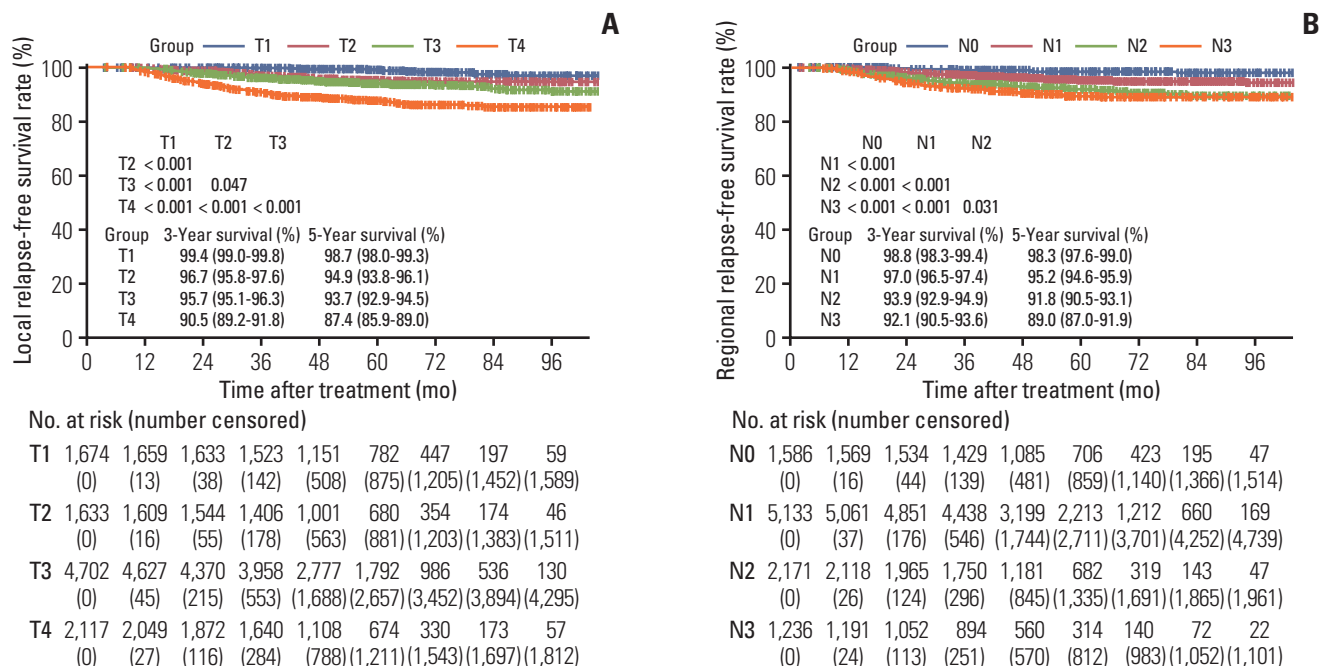
Fig. 3. Kaplan-Meier curves for 10,126 nasopharyngeal carcinoma patients receiving intensity-modulated radiotherapy: local relapse-free survival (A) stratified by T category and regional relapse-free survival (B) stratified by N category.

erate a nomogram to predict LRFS. The predictors included smoking history, T category, and N category (Table. 3). We further built a nomogram (S5A Fig.) based on weighting of the three significant covariates in the SYSUCC-A cohort, yielding for LRFS a Harrell's C-index of 0.697 (95% CI, 0.647 to 0.746) (S5B Fig.). The calibration plots for the 3-year LRFS predicted well in the SYSUCC-B (C-index, 0.684; 95% CI, 0.634 to 0.735) (S5C Fig.) and the WZRCH (C-index, 0.652; 95% CI, 0.501 to 0.831) (S5D Fig.) cohorts. In addition, the nomogram for LRFS had a higher C-index (0.697; 95% CI, 0.647 to 0.746) than T category (C-index, 0.650; 95% CI, 0.597 to 0.703), though narrowly statistically significant (p=0.053).

6. Regional relapse-free survival

Five hundred twenty-six of the 10,126 patients (5%) suffered regional recurrence in the overall Guangzhou cohort, and the 1-, 3-, and 5-year RRFS rates were 99.1%, 96.0%, and 94.3%, respectively (Fig. 1E). The 1-, 3-, and 5-year rates for RRFS stratified by N category was detailed in Fig. 3B. As the N category increased from N0 to N3, risk of suffering regional recurrence was also associated with an increase accordingly (p < 0.05 for all) (Fig. 3B). This is most likely due to the difficultly in differentiating between N2 and N3 disease. The RRFS curves for N2 and N3 disease were not clearly separated following 6 years after treatment (p=0.031) (Fig. 3B). Among the SYSUCC-A cohort, multivariable analysis

for RRFS was conducted to generate a nomogram to predict RRFS. Significant predictors of RRFS were T category and N category (Table 3). A nomogram that was based on weighting of the two significant covariates were constructed in the SYSUCC-B (S6A Fig.). In the SYSUCC-A cohort, calibration plots for 3-year RRFS were well predicted (C-index, 0.672; 95% CI, 0.593 to 0.751) (S6B Fig.). This was also observed in the SYSUCC-B cohort (C-index, 0.650; 95% CI, 0.594 to 0.705) (S6C Fig.) and WZRCH cohort (C-index, 0.616; 95% CI, 0.431 to 0.785) (S6D Fig.). In addition, the nomogram for RRFS had a higher C-index (0.672; 95% CI, 0.593 to 0.751) than N category (C-index, 0.643; 95% CI, 0.581 to 0.705), with a potential trend towards statistical significance (p=0.054).

7. The development of web-based nomograms

To create user-friendly access, the underlying statistical formulas were implemented in the web-based nomograms. Web-based nomogram users will be able to estimate the individual 1-, 3-, and 5-year OS, DMFS, DFS, LRFS, and RRFS by entering the sex, age, history of smoking status, T category, and N category. S7 Fig. presents a screenshot of the web-based nomograms which are available on https://modelstore.yiducloud.com.cn/#/?lang=en. For example, a male patient that was 46 years of age with T3N2M0 NPC enters the clinic. This patient had no prior history of smoking. The estimated 1-, 3-, and 5-year for OS were 99.14%, 90.61%, and

86.35%, DMFS were 94.07%, 84.05%, and 82.81%, DFS were 92.61%, 77.97%, and 72.48%, LRFS were 99.02%, 95.27%, and 92.95%, and RRFS were 98.55%, 93.19%, and 90.32%.

Discussion

Although various studies have built nomograms for predicting survival outcomes after RT for NPC, a web-based prediction model is still needed for clinical assessment and patient counseling. To the best of our knowledge, this is the first study to include external validation to characterize patterns of disease recurrence, and establish web-based nomograms derived from large series of patients with NPC who underwent radical IMRT. These user-friendly web-based models may serve as a useful tool for providing patient counseling and helping physicians to make individual follow-up plans.

The current study showed that patients with stage I disease had an excellent survival rate of 99.5%. In contrast, stage IVA disease patients were most challenging to treat, with 25% of patients dying within 5 years. A similar trend was also observed in patients with N3 category. With the use of IMRT, local relapse rate in NPC improved greatly and more than 70% of the deaths are attributed to distant metastasis [14]. For this reason, prognostic value for T category has potentially become weaker compared with N category because of great local control. Consistent with the conclusion, N category has comparable prognostic accuracy in OS compared with clinical stage in our study. This observation was the result of weakening impact for the T category on clinical stage, which would reduce the prediction efficiency of clinical stage for OS [15]. Although prognostic models could provide a more accurate and precise prediction than the TNM staging system, there are few prognostic models related to OS for NPC. Recently, Liang et al. [16] developed a nomogram for predicting OS based on 1520 non-metastatic patients diagnosed with NPC that underwent curative IMRT with the Harrel's C-index of 0.69 (95% CI, 0.67 to 0.71). In the current series, a prognostic nomogram to predict OS was built and our nomogram provided a higher C-index 0.736 (95% CI, 0.690 to 0.783) than the study reported by Liang et al. [16]. Moreover, the nomogram performance was verified in all validation cohorts. Therefore, our nomogram potentially provides a more accurate method for predicting OS among NPC patients.

With the improvement of locoregional control in the modern era, distant metastasis is now the primary failure pattern for NPC. The result from the present study was found to be in line with previous studies [17]. Our study showed that the actual 5-year DMFS was 86.9%, which is comparable to

results by Sun et al. [18] and Lai et al. [19], though slightly higher than a study by Wu et al. (5-year DMFS, 81.8%) [20]. We postulate that the utilization rate of chemotherapy might partly be attributed to the inconsistency. It is well recognized that distant control could be improved by using chemotherapy, especially among advanced-stage patients [21]. In the study by Wu et al. [20], nearly 20% of stage III-IV disease did not receive chemotherapy. In contrast, less than 5% of stage III-IV disease received RT only. Although patients with N0 disease achieved an excellent distant control rate (5-year DMFS, 95%), 20%-30% of patients with N2-3 disease suffered distant failure. Efforts to improve distant control for NPC will likely require greater attention to patients with N2-3 disease. Since the primary advantage of induction chemotherapy (IC) is improving distant control [22], IC of adequate intensity, such as four cycles is potentially a reasonable approach to lessen distant metastasis of N2-3 disease [23]. However, this hypothesis must be further examined in prospective clinical trials. Additionally, we built a nomogram to predict risk of distant metastasis, yielding a satisfactory C-index (0.717; 95% CI, 0.665 to 0.768), and in all validation cohorts the nomogram's performance was verified. Therefore, the nomogram in the current study potentially provides a simple and accurate method for predicting distant metastasis in the clinical management of NPC.

The 5-year DFS shown in this study is slightly higher than that by Lai et al. [19], but to some extent lower than that by Zhang et al. [24]. We found that the 3- and 5-year DFS rates were 83.4% and 78.4%, respectively. Interestingly, only a small decrease was found between 3- and 5-year DFS (5.0%), comparable with findings by Lee et al. [25]. Since 10% to 20% of NPC patients with local or systemic recurrence might be cured with supplementary treatment [17], it is essential to identify individuals early that are at high risk of recurrence. In the present study, we developed and validated a prognostic nomogram based on clinicopathologic factors, and proved that the prognostic nomogram was effective for DFS prediction (C-index, 0.713; 95% CI, 0.677 to 0.752). Recently, an MRI-based radiomics nomogram was reported to have better prognostic ability for advanced NPC (C-index, 0.776; 95% CI, 0.678 to 0.873) [26]. Although it provided a better predictive efficacy, there is still a long way to go before clinical application because of constraints in the instability of MRI-based radiomics. Compared with previous prognostic models [26,27], a major strength of our model is that prognostic factors were easy to obtain, and the nomogram yielded a satisfactory C-index as well.

Recent improvements in local control for NPC patients is partly attributable to IMRT [28]. Early studies [14,18] by independent groups reported 5-year local control rates of 85%-95% in the IMRT era for NPC aligning with our find-

ings. Largely, advanced T category and poor local control are associated, but the recent study by Au et al. [29] found no associated difference in local recurrence between T2 and T3. Though in the present study our survival curves for T2 and T3 nearly overlapped, they were still slightly significant. This inconsistency might be due to the large sample size in the present study. Since local control rate achieved using IMRT is higher than 90%, the prognostic power of T categories decreased. Hence, it seems reasonable to suggest T categories could be further optimized, for example through merging T2 and T3 or incorporating other distinguishing factors [15]. Although the local control for patients treated by IMRT significantly improved compared with those receiving two-dimensional RT (2DRT), it remains unsatisfactory for patients with T4 category having a local recurrence rate of nearly 15% suffered in the current study. This suggests that better patient stratification of T4 disease for treatment intensification is needed [7].

Based on previous research, both IMRT and 2DRT were equally effective in achieving regional control in the lymph region [19,24]. Our 5-year RRFS rate was approximately 94%, analogous to regional control rates achieved in previous IMRT studies [15,29]. To date, it remains unclear the prophylactic irradiation of all neck node levels in NPC patients without lymph node metastasis. Most prior clinical trials on NPC treatment encouraged routine bilateral entire neck irradiation (from retropharyngeal region to level IV/V [15]. However, some investigators have confirmed that elective upper neck lymph drainage region (level II, III, and VA) is appropriate for NPC negative cervical lymph nodes [30]. In the current study, patients with N0 disease only received prophylactic irradiation to the upper neck lymph drainage region (level II and III), not including the level IV to VB, and supraclavicular lower neck lymph node drainage areas. The RRFS rate for these group of patients was up to 99%. This potentially suggest that prophylactic irradiation of level II and III provides sufficient regional control for patients with N0 disease, and prophylactic irradiation excluding the level IV-VB and supraclavicular region may not increase the risk of regional recurrence in N0 disease [31]. However, comparative randomized clinical trials are still needed.

In the current study, the main strength is the large-scale data that utilizes objective information such as medical records, which provides accurate medical treatment outcomes for NPC. However, there are several limitations that should be mentioned. First, our nomograms did not include plasma Epstein-Barr virus (EBV) DNA. There are increasing data suggesting plasma EBV DNA can be a beneficial molecular marker for prognosis in NPC patients [32]. However, the testing methodology has not been standardized globally. Additionally, the test is expensive and the value for

optimal cutoff has yet to be defined. Second, nomograms were determined based on data collected from endemic areas in the current study. Therefore, uncertainty remains whether these nomograms are applicable to patients in nonendemic areas. Third, as the median follow-up for the validation cohort was about 4 years, close monitoring and 5-year follow-up data are still required for these patients. Fourth, not all of advanced-stage patients could receive PET-CT due to the issue of health insurance in our country, which may cause inaccurate in tumor staging. However, patients with advanced-stage NPC who have no conditions to pay PET-CT fees were required to complete chest and abdomen enhanced CT and ECT in our study, which to a certain extent ensures the accuracy of tumor staging. Finally, adding possible clinically significant interactions is excellent. However, it is beyond the scope of the current study which focused on the prediction model rather than association assessment. But we will consider it in future studies.

In conclusion, our study provides comprehensive insight on clinical features, patterns of recurrence, and survival outcomes for non-metastatic NPC in an endemic area. The training and validation of nomograms based on existing prognostic factors provides a satisfactory prediction efficiency. To encourage widespread clinical use, we published our nomograms as a publicly available online tool. The user-friendly web-based models may serve as a useful tool for helping physicians make individual follow-up plans and providing patient counseling. Future research is needed to test these nomograms on multiple datasets across nonendemic areas to further strengthen accuracy of prediction.

Electronic Supplementary Material
Supplementary materials are available at Cancer Research and Treatment website (https://www.e-crt.org).

Ethical Statement
This study was reviewed and approved by the Institutional Review Board and Ethics Committee of SYSUCC (approval number: YB2019-76). Written informed consent for the use of clinical data was obtained when the patients were admitted to receive treatment as a general standard procedure for those treated at SYSUCC. Patients' clinical and demographic information were anonymized prior to analysis.

Author Contributions
Conceived and designed the analysis: Yao JJ, Lin L, Gao TS, Lawrence WR, Ma J, Sun Y.
Collected the data: Yao JJ, Lin L, Gao TS, Lawrence WR, Ma J, Sun Y.
Contributed data or analysis tools: Yao JJ, Gao TS, Zhang WJ, Sun Y.
Performed the analysis: Yao JJ, Zhang WJ, Sun Y.
Wrote the paper: Yao JJ, Lin L, Sun Y.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68:394-424.

2. 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.

3. Tendulkar RD, Agrawal S, Gao T, Efstathiou JA, Pisansky TM, Michalski JM, et al. Contemporary update of a multi-institutional predictive nomogram for salvage radiotherapy after radical prostatectomy. J Clin Oncol. 2016;34:3648-54.

4. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. J Clin Oncol. 2016;34:2157-64.

5. Liang W, Zhang L, Jiang G, Wang Q, Liu L, Liu D, et al. Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. J Clin Oncol. 2015;33:861-9.

6. Tang LQ, Li CF, Li J, Chen WH, Chen QY, Yuan LX, et al. Establishment and validation of prognostic nomograms for endemic nasopharyngeal carcinoma. J Natl Cancer Inst. 2016;108:djv291.

7. Zhang LL, Li YY, Hu J, Zhou GQ, Chen L, Li WF, et al. Proposal of a pretreatment nomogram for predicting local recurrence after intensity-modulated radiation therapy in T4 nasopharyngeal carcinoma: a retrospective review of 415 Chinese patients. Cancer Res Treat. 2018;50:1084-95.

8. Zeng Q, Hong MH, Shen LJ, Meng XQ, Guo X, Qian CN, et al. Nomograms for predicting long-term survival in patients with non-metastatic nasopharyngeal carcinoma in an endemic area. Oncotarget. 2016;7:29708-19.

9. Huang XD, Zhou GQ, Lv JW, Zhou HQ, Zhong CW, Wu CF, et al. Competing risk nomograms for nasopharyngeal carcinoma in the intensity-modulated radiotherapy era: a big-data, intelligence platform-based analysis. Radiother Oncol. 2018;129:389-95.

10. Jin YN, Zhang WJ, Cai XY, Li MS, Lawrence WR, Wang SY, et al. The characteristics and survival outcomes in patients aged 70 years and older with nasopharyngeal carcinoma in the intensity-modulated radiotherapy Era. Cancer Res Treat. 2019;51:34-42.

11. Therneau T, Crowson C, Atkinson E. Using time dependent covariates and time dependent coefficients in the Cox model [Internet]. Vienna: R Foundation for Statistical Computing; 2014 [cited 2020 Dec 4]. Available from: https://cran.r-project.org/web/packages/survival/vignettes/timedep.pdf.

12. rms: Regression Modeling Strategies [Internet]. Vienna: R Foundation for Statistical Computing; 2014 [cited 2020 Dec 4]. Available from: https://CRAN.R-project.org/package=rms.

13. Package 'compareC' [Internet]. Vienna: R Foundation for Statistical Computing; 2015 [cited 2020 Dec 4]. Available from: https://cran.r-project.org/web/packages/compareC/compareC.pdf.

14. Mao YP, Tang LL, Chen L, Sun Y, Qi ZY, Zhou GQ, et al. Prognostic factors and failure patterns in non-metastatic nasopharyngeal carcinoma after intensity-modulated radiotherapy. Chin J Cancer. 2016;35:103.

15. Tang LL, Chen YP, Mao YP, Wang ZX, Guo R, Chen L, et al. Validation of the 8th Edition of the UICC/AJCC staging system for nasopharyngeal carcinoma from endemic areas in the intensity-modulated radiotherapy era. J Natl Compr Canc Netw. 2017;15:913-9.

16. Liang W, Shen G, Zhang Y, Chen G, Wu X, Li Y, et al. Development and validation of a nomogram for predicting the survival of patients with non-metastatic nasopharyngeal carcinoma after curative treatment. Chin J Cancer. 2016;35:98.

17. Chen YP, Chan ATC, Le QT, Blanchard P, Sun Y, Ma J. Nasopharyngeal carcinoma. Lancet. 2019;394:64-80.

18. Sun X, Su S, Chen C, Han F, Zhao C, Xiao W, et al. Long-term outcomes of intensity-modulated radiotherapy for 868 patients with nasopharyngeal carcinoma: an analysis of survival and treatment toxicities. Radiother Oncol. 2014;110:398-403.

19. Lai SZ, Li WF, Chen L, Luo W, Chen YY, Liu LZ, et al. How does intensity-modulated radiotherapy versus conventional two-dimensional radiotherapy influence the treatment results in nasopharyngeal carcinoma patients? Int J Radiat Oncol Biol Phys. 2011;80:661-8.

20. Wu LR, Liu YT, Jiang N, Fan YX, Wen J, Huang SF, et al. Ten-year survival outcomes for patients with nasopharyn-

geal carcinoma receiving intensity-modulated radiotherapy: an analysis of 614 patients from a single center. Oral Oncol. 2017;69:26-32.

21. Yao JJ, Jin YN, Liu ZG, Liu QD, Pei XF, Zhou HL, et al. Do all patients with advanced N-stage nasopharyngeal carcinoma benefit from the addition of induction chemotherapy to concurrent chemoradiotherapy? Ther Adv Med Oncol. 2019; 11:1758835919833863.

22. Chua DT, Ma J, Sham JS, Mai HQ, Choy DT, Hong MH, et al. Long-term survival after cisplatin-based induction chemotherapy and radiotherapy for nasopharyngeal carcinoma: a pooled data analysis of two phase III trials. J Clin Oncol. 2005;23:1118-24.

23. Chang H, Peng L, Tao YL, Chen C, Xiao WW, Hu YH, et al. Necessity of concurrent chemotherapy in N2-3 nasopharyngeal carcinoma treated with neoadjuvant chemotherapy of ≥ 3 cycles followed by intensity-modulated radiotherapy. Cancer Med. 2019;8:2823-31.

24. Zhang MX, Li J, Shen GP, Zou X, Xu JJ, Jiang R, et al. Intensity-modulated radiotherapy prolongs the survival of patients with nasopharyngeal carcinoma compared with conventional two-dimensional radiotherapy: A 10-year experience with a large cohort and long follow-up. Eur J Cancer. 2015;51:2587-95.

25. Lee AW, Fee WE Jr, Ng WT, Chan LK. Nasopharyngeal carcinoma: salvage of local recurrence. Oral Oncol. 2012;48:768-74.

26. Zhang B, Tian J, Dong D, Gu D, Dong Y, Zhang L, et al. Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma. Clin Cancer Res. 2017;23:4259-69.

27. Gu B, Zhang J, Ma G, Song S, Shi L, Zhang Y, et al. Establishment and validation of a nomogram with intratumoral heterogeneity derived from (18)F-FDG PET/CT for predicting individual conditional risk of 5-year recurrence before initial treatment of nasopharyngeal carcinoma. BMC Cancer. 2020;20:37.

28. Peng G, Wang T, Yang KY, Zhang S, Zhang T, Li Q, et al. A prospective, randomized study comparing outcomes and toxicities of intensity-modulated radiotherapy vs. conventional two-dimensional radiotherapy for the treatment of nasopharyngeal carcinoma. Radiother Oncol. 2012;104:286-93.

29. Au KH, Ngan RK, Ng AW, Poon DM, Ng WT, Yuen KT, et al. Treatment outcomes of nasopharyngeal carcinoma in modern era after intensity modulated radiotherapy (IMRT) in Hong Kong: a report of 3328 patients (HKNPCSG 1301 study). Oral Oncol. 2018;77:16-21.

30. Sun Y, Tang LL, Chen L, Li WF, Mao YP, Liu LZ, et al. Promising treatment outcomes of intensity-modulated radiation therapy for nasopharyngeal carcinoma patients with N0 disease according to the seventh edition of the AJCC staging system. BMC Cancer. 2012;12:68.

31. Cho WK, Oh D, Lee E, Kim TG, Lee H, Nam H, et al. Feasibility of selective neck irradiation with lower elective radiation dose in treating nasopharynx cancer patients. Cancer Res Treat. 2019;51:603-10.

32. Jin YN, Yao JJ, Zhang F, Wang SY, Zhang WJ, Zhou GQ, et al. Is pretreatment Epstein-Barr virus DNA still associated with 6-year survival outcomes in locoregionally advanced nasopharyngeal carcinoma? J Cancer. 2017;8:976-82.