Do all patients with advanced N-stage nasopharyngeal carcinoma benefit from the addition of induction chemotherapy to concurrent chemoradiotherapy?

Ji-Jin Yao iD, Ya-Nan Jin, Zhi-Gang Liu, Qiao-Dan Liu, Xiao-Feng Pei, Huai-Li Zhou, Wang-Jian Zhang, Fan Zhang, Li Lin, Wayne R. Lawrence, Si-Yang Wang, Jun Ma, Guan-Qun Zhou and Ying Sun

Abstract

Background: The aim of this study was to evaluate the benefits from the addition of induction chemotherapy (IC) to concurrent chemoradiotherapy (CCRT) in N2-3 nasopharyngeal carcinoma (NPC).

Methods: A total of 3089 patients with nonmetastatic NPC, staged as N2-3 were retrospectively reviewed. IC contained cisplatin ($80\,mg/m^2$) with 5-fluorouracil ($800\,mg/m^2$/day over 120 h), or cisplatin ($80\,mg/m^2$) with docetaxel ($80\,mg/m^2$), or cisplatin ($60\,mg/m^2$) with 5-fluorouracil ($600\,mg/m^2$ over 120 h), and docetaxel ($60\,mg/m^2$) administered at 3-week intervals for two or three cycles. Concurrent chemotherapy consisted of cisplatin (80 or $100\,mg/m^2$) given in weeks 1, 4, and 7 of radiotherapy, or cisplatin ($40\,mg/m^2$) given weekly during radiotherapy. Overall, three well-matched risk groups (low, intermediate, and high risk) were created using propensity score matching, and IC plus CCRT was compared with CCRT in each risk group. Our primary endpoint was distant metastasis-free survival (DMFS).

Results: A nomogram for DMFS was established with good prognostic accuracy (C-index, 0.69; 95% confidence interval, 0.64–0.73). The survival curves for low, intermediate, and high-risk groups stratified by the nomogram were significantly different between all three risk groups, with corresponding 5-year DMFS rates of 90.7%, 79.4%, and 64.9%, respectively ($p < 0.001$). IC plus CCRT was significantly associated with superior DMFS as compared with CCRT alone (69.5% *versus* 56.7%, $p = 0.004$) in the high-risk group. However, no significant difference between IC plus CCRT and CCRT was observed ($p = 0.831$ and 0.608, respectively) in the intermediate and low-risk groups.

Conclusions: Our findings can help accurately guide the treatment of individual patients with advanced N-stage NPC.

Keywords: advanced N-stage, benefit, induction chemotherapy, nasopharyngeal carcinoma, nomogram, propensity score matching method

Correspondence to:
Ying Sun
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 510060, Guangdong Province, China
sunying@sysucc.org.cn

Ji-Jin Yao
Guan-Qun Zhou
Department of Radiation Oncology, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong Province, China Department of Head and Neck Oncology, the Cancer Center of the Fifth Affiliated Hospital of Sun Yat-sen University, Zhuhai, Guangdong Province, China

Ya-Nan Jin
Zhi-Gang Liu
Qiao-Dan Liu
Fan Zhang
Si-Yang Wang
Department of Head and Neck Oncology, the Cancer Center of the Fifth Affiliated Hospital of Sun Yat-sen University, Zhuhai, Guangdong Province, China

Xiao-Feng Pei
Huai-Li Zhou
Department of Thoracic Oncology, the Cancer Center of the Fifth Affiliated Hospital of Sun Yat-sen University, Zhuhai, Guangdong Province, China

Wang-Jian Zhang
Wayne R. Lawrence
Department of Medical Statistics and

Introduction

Nasopharyngeal carcinoma (NPC) is endemic in Southeast Asia, where the incidence rate ranges 15–50 cases per 100,000 people.[1,2] Presently, the primary treatment modality for NPC is radiotherapy (RT). With advancements in RT technology and imaging techniques, the locoregional control rate is higher than 90%, and the main failure pattern is distant metastasis.[3,4] Patients with NPC are divided into N0–N3

Epidemiology & Health Information Research Center & Guangdong Key Laboratory of Medicine, School of Public Health, Sun Yat-sen University, Guangzhou, Guangdong Province, China
Department of Environmental Health Sciences, School of Public Health, University at Albany, State University of New York, Rensselaer, NY, USA

Li Lin
Jun Ma
Department of Radiation Oncology, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong Province, China

categories, based on the American Joint Commission on Cancer (AJCC) classification.[5] Recent evidence[6] has indicated that more than 75% of distant failures are concentrated in the advanced N-stage (stage N2–N3). Since outcomes are particularly poor for NPC patients with distant metastasis, with a median overall survival (OS) ranging from 7 to 16 months,[7–9] identifying patients with a high risk of metastasis is important for tailoring individualized therapy and reducing distant metastasis.

Currently, the standard treatment for locoregionally advanced NPC is concurrent chemoradiotherapy (CCRT).[10–12] However, CCRT alone may not be sufficient for some patients, particularly those with high-risk distant metastasis.[13] Although the addition of adjuvant chemotherapy (AC) to CCRT might reduce distant metastasis for NPC patients at an elevated distant failure risk, the concurrent–adjuvant approach has low compliance for three cycles of AC (around 60%).[14] Induction chemotherapy (IC) offers advantages for improved tolerance and early eradication of micrometastases compared with AC.[14,15] A recent phase III trial by Sun and colleagues reported that IC plus CCRT could improve distant metastasis-free survival (DMFS) in locoregionally advanced NPC compared with CCRT alone.[15] However, another trial by Fountzilas and colleagues[16] did not observe significant survival improvements. Hence, it remains unclear whether all advanced N-stage NPC patients would benefit from the addition of IC to CCRT.

To fill the current gaps in knowledge, we conceived and initiated a large-scale, real-world study to establish a nomogram model with improved prediction accuracy compared with clinical risk factors for distant metastasis in advanced N-stage NPC. We then applied this nomogram to place patients into risk groups (low, intermediate, and high risk). Additionally, individual comparisons of CCRT plus IC and CCRT alone were performed for each well-matched group based on a propensity score matching method to assess the benefit of IC in advanced N-stage NPC patients.

Materials and methods

Study population
This study utilized the NPC-specific database, which is derived from Sun Yat-sen University Cancer Center's (SYSUCC) well-established big data intelligence platform. A total of 10,126 patients were identified with histologically proven nondisseminated NPC, diagnosed from April 2009 to December 2015. Patients demographic, diagnostic, and therapeutic information were obtained from the intelligence platform using keywords such as 'diagnosis', 'histologic type', 'age at first diagnosis', 'sex', 'disease stage', 'RT technology', and 'Epstein–Barr virus (EBV) DNA'. Detailed descriptions of the SYSUCC intelligence platform have previously been published.[17] Briefly, this novel 'big data' research system, enables organizing, integrating, restructuring, and updating data in real-time from numerous clinical business systems, based on a well-designed data model and algorithm.

In this study, we identified eligible patients based on the following inclusion criteria: pathologically diagnosed NPC; no evidence of distant metastasis; N2–3 disease based on the AJCC staging system (8th edition, 2016)[5]; treated with intensity-modulated radiotherapy (IMRT); complete medical history; and baseline laboratory testing, including plasma EBV DNA, high-sensitivity C-reactive protein (hs-CRP), hemoglobin (HGB), and lactate dehydrogenase (LDH). Clinical information on the 3089 patients that met the inclusion criteria were included. The present study received approval from the Institutional Review Board (IRB-approved number, YB2018-005) of SYSUCC, Guangzhou, China. Written informed consent for the use of clinical data and collected samples for future studies (including retrospective studies) were obtained when the patients were admitted to receive treatment as a general standard procedure for patients treated in our center. All patient records were anonymous and de-identified before the analysis.

Diagnosis and treatment
All patients underwent complete evaluation before treatment, including a physical examination, fiber-optic nasopharyngoscopy, hematology and biochemistry profiling, magnetic resonance imaging (MRI) scanning of the suprasellar cistern to the collarbone, computed tomography (CT), abdominal ultrasonography, whole body bone scan (ECT), or 18F-fluorodeoxyglucose positron emission tomography (PET) and CT (PET-CT). Real-time quantitative polymerase chain reaction was used to measure plasma EBV DNA concentrations as described in prior studies.[18,19] The two radiation oncologists specializing in head and neck cancer

restaged patients.[5] Patients were treated according to treatment principles for NPC established by SYSCC (see supplementary material).

Data sharing
Key raw data were uploaded onto the Research Data Deposit public platform (RDD), with the approval number of RDDB2018000430.

Follow up and end points
Patients were examined during the first 2 years at least every 3 months, and every 6 months for 3 years thereafter. During the visits, clinical examinations, fiberoptic nasopharyngoscopy, and plasma EBV DNA were routinely performed. Patients with a clinical suspicion of metastasis were recommended for MRI, abdominal sonography, ECT, or PET-CT, followed by confirmatory cytological biopsies if possible. The study's primary endpoint was DMFS, and the secondary endpoint was OS. We calculated DMFS as the time from the date of initial treatment on day 1 to the first distant relapse date; and OS to all-cause mortality, or the last follow-up visit date, whichever occurred first.

Statistical analysis
We converted continuous variables to categorical variables according to clinical cut-off points (HGB, LDH, hs-CRP) or as determined by prior study findings (age, plasma EBV DNA).[20–23] The Kaplan–Meier method was used to calculate survival rates which were compared using log-rank tests.[24] Potential factors were analyzed with univariate and multivariate Cox proportional hazards models. Variables with $p < 0.05$ in the univariate analyses were included in the multivariable Cox regression analyses. Additionally, forest plots were generated to present adjusted hazard ratios (HRs) and 95% confidence intervals (CIs) of the validated predictors for DMFS and OS.

Based on the multivariable Cox regression analysis, nomograms were generated. The selection of the final prediction model was performed using a backward step-down selection process with the Akaike information criterion.[25] Nomogram performance was assessed using the concordance index (c-index) and evaluated through means of comparison of nomogram-predicted and observed Kaplan–Meier estimates of survival probability. Additionally, bootstraps with 1000 resamples were applied.

We scored the risk of distant metastasis for all patients with the aim of categorizing patients into groups (low, intermediate, and high risk) according to the prognostic factors in the nomogram for DMFS. We identified patients that would benefit most from CCRT with IC by performing individual comparisons of IC plus CCRT *versus* CCRT alone for each risk group. To reduce the effect of potential confounders on selection bias, the propensity score matching (PSM) method without replacement was performed for comparisons using the nearest-neighbor method with a stringent caliper of 0.05.[26] We compared the categorical variables in different groups using the Chi-square test or Fisher's exact test if indicated. The criterion for statistical significance was set at an α of 0.05 and all p values were based on two-sided tests. All statistical models were generated using R statistical software, version 3.3.2 (http://www.r-project.org/).

Results

Patient characteristics and treatment outcomes
The overall median age was 45 years (range, 18–81 years), and the male:female ratio was 3.1:1 (2337 men and 752 women). As is typical of endemic areas, histological examination revealed that 3019 (98%) patients had World Health Organization (WHO) type III disease, which is nonkeratinizing undifferentiated NPC. Table 1 shows the detailed clinicopathologic characteristics. The percentages of patients grouped as stage N2 and stage N3 were 64% and 36%, respectively. Most patients underwent IC followed by CCRT (57.7%) or CCRT alone (27.5%), 10.4% underwent IC plus RT, 2.1% underwent CCRT plus AC, and 2.3% were treated with IMRT alone. Furthermore, the number of patients with different induction, adjuvant and concurrent chemotherapy regimens are listed in Table S1. Following a median patient follow-up time of 51.2 months (range, 3.4–101.9 months), 592 (19.2%) developed distant metastases, and 533 (17.3%) died during the study period. Overall, the 3-year and 5-year DMFS, and 3-year and 5-year OS rates were 81.8% and 78.1%, and 87.4% and 76.7%, respectively.

The development of nomograms for DMFS and OS
Table 1 presents the univariate analyses results. In multivariate analysis, sex, T stage, N-stage, EBV

Table 1. Clinicopathologic characteristics and univariate analysis of DMFS and OS in the 3089 patients with advanced N-stage NPC.

Characteristic	Entire cohort no. (%)[a]	Univariate analysis			
		DMFS		OS	
Sex		HR (95% CI)	*p*	HR (95% CI)	*p*
Male	2337 (75.7)	Reference		Reference	
Female	752 (24.3)	0.72 (0.59, 0.89)	0.002	0.84 (0.68, 1.03)	0.096
Histology (WHO)					
Type I–II	70(2.3)	Reference		Reference	
Type III	3019 (97.7)	0.69 (0.44, 1.09)	0.109	0.72 (0.46, 1.12)	0.143
Age, years					
$\leqslant$30	354 (11.5)	Reference		Reference	
31–40	761 (24.6)	1.19 (0.87, 1.63)	0.274	1.15 (0.80, 1.63)	0.412
41–50	1020 (33.0)	1.41 (1.04, 1.91)	0.014	1.43 (1.03, 1.99)	0.012
51–60	652 (21.1)	1.28 (0.93, 1.77)	0.067	1.65 (1.17, 2.34)	0.006
$\geqslant$61	302 (9.8)	1.57 (1.10, 2.25)	0.070	2.19 (1.51, 3.19)	<0.001
Smoking history					
No	1877 (60.8)	Reference		Reference	
Yes	1212 (39.2)	1.24 (1.05, 1.46)	0.010	1.08 (0.91, 1.28)	0.394
Family of cancer					
No	2335 (75.6)	Reference		Reference	
Yes	754 (24.4)	1.10 (0.91, 1.32)	0.327	1.03 (0.84, 1.25)	0.793
T-stage (8th edition)					
T1	358 (11.6)	Reference		Reference	
T2	516 (16.7)	1.26 (0.88, 1.81)	0.475	1.34 (0.91, 1.96)	0.341
T3	1523 (49.3)	1.64 (1.20, 2.24)	0.016	1.66 (1.19, 2.30)	0.021
T4	692 (22.4)	1.97 (1.42, 2.73)	<0.001	2.07 (1.46, 2.93)	<0.001
N-stage (8th edition)					
N2	1981(64.1)	Reference		Reference	
N3	1108 (35.9)	1.64 (1.39, 1.92)	<0.001	1.64 (1.39, 1.95)	<0.001
EBV DNA, copy/ml[b]					
<1000	759 (24.6)	Reference		Reference	
1000–9999	807 (26.1)	1.78 (1.34, 2.36)	<0.001	1.77 (1.31, 2.39)	<0.001
10,000–99,999	992 (32.1)	2.36 (1.82, 3.07)	<0.001	2.34 (1.78, 3.09)	<0.001

Table 1. (Continued)

Characteristic	Entire cohort no. (%)ᵃ	Univariate analysis			
		DMFS		OS	
≥100,000	531 (17.2)	3.41 (2.59, 4.47)	<0.001	3.07 (2.30, 4.10)	<0.001
HGB, g/lᵇ					
<113	122 (3.9)	Reference		Reference	
113–151	2037 (65.9)	0.57 (0.41, 0.79)	0.004	0.65 (0.45, 0.93)	0.012
≥151	930 (30.1)	0.56 (0.39, 0.79)	0.001	0.55 (0.38, 0.81)	0.001
hs-CRP, g/mlᵇ					
<1.0	824 (26.7)	Reference		Reference	
1.0–3.0	1124 (36.4)	1.14 (0.91, 1.42)	0.219	1.27 (1.00, 1.61)	0.049
≥3.0	1141 (36.9)	1.56 (1.26, 1.93)	<0.001	1.69 (1.35, 2.12)	<0.001
LDH, U/lᵇ					
<245	2713 (87.8)	Reference		Reference	
≥245	376 (12.2)	1.94 (1.58, 2.38)	<0.001	1.80 (1.45, 2.24)	<0.001

CI, confidence interval; DMFS, distant metastasis-free survival; EBV, Epstein–Barr virus; HGB, hemoglobin; HR, hazard ratio; hs-CRP, high-sensitivity C-reactive protein; LDH, lactate dehydrogenase; NPC, nasopharyngeal carcinoma; OS, overall survival; WHO, World Health Organization.
ᵃPercentages may not add up to 100, due to rounding.
ᵇAll variables were measured before treatment.

DNA, HGB, and LDH had proven significant effects on DMFS [Figure 1(a)]; and age, T stage, N-stage, EBV DNA, HGB, hs-CRP, and LDH were significantly associated with OS [Figure 1(b)]. Prognostic nomograms combining all aforementioned validated predictors for 3- and 5-year DMFS and OS are presented in Figure 2(a) and (b). The prognostic nomograms provided good accuracy for predicting DMFS and OS, with corresponding c-index values of 0.69 (95% CI, 0.64–0.73) and 0.71 (95% CI, 0.66–0.75), respectively. Moreover, calibration plots for the probabilities of 3- and 5-year DMFS [Figure 2(c) and (e)] and OS [Figure 2(d) and (f)] showed ideal agreement with the nomogram predictions for DMFS [Figure 2(a)] and OS [Figure 2(b)], respectively.

Nomogram-generated risk stratification for distant metastasis

Based on the total score of nomograms for distant metastasis, all patients were categorized into risk groups: low-risk group (582 patients; total score ≤ 150 points); intermediate-risk group (1791 patients; 150 < total score ≤ 300 points); and high-risk group (716 patients; total score > 300 points). Supplementary Tables S2 and S3 present the objective hematological toxicities and liver and kidney dysfunction of each risk group before and during treatment, respectively. Survival curves showed excellent discrimination in distant failure among the low-risk group, intermediate-risk group, and high-risk group, with corresponding 5-year DMFS rates of 90.7%, 79.4%, and 64.9%, respectively [$p < 0.001$; Figure 3(a)]. In addition, the 5-year OS rates for the low-risk group (89.9%), intermediate-risk group (77.7%), and high-risk group (64.9%) also differed significantly between groups [$p < 0.001$; Figure 3(b)].

Benefits of adding IC to CCRT in each risk group

The nomogram-generated stratification based on DMFS showed patients that had an advanced N-stage receiving IC plus CCRT or CCRT alone
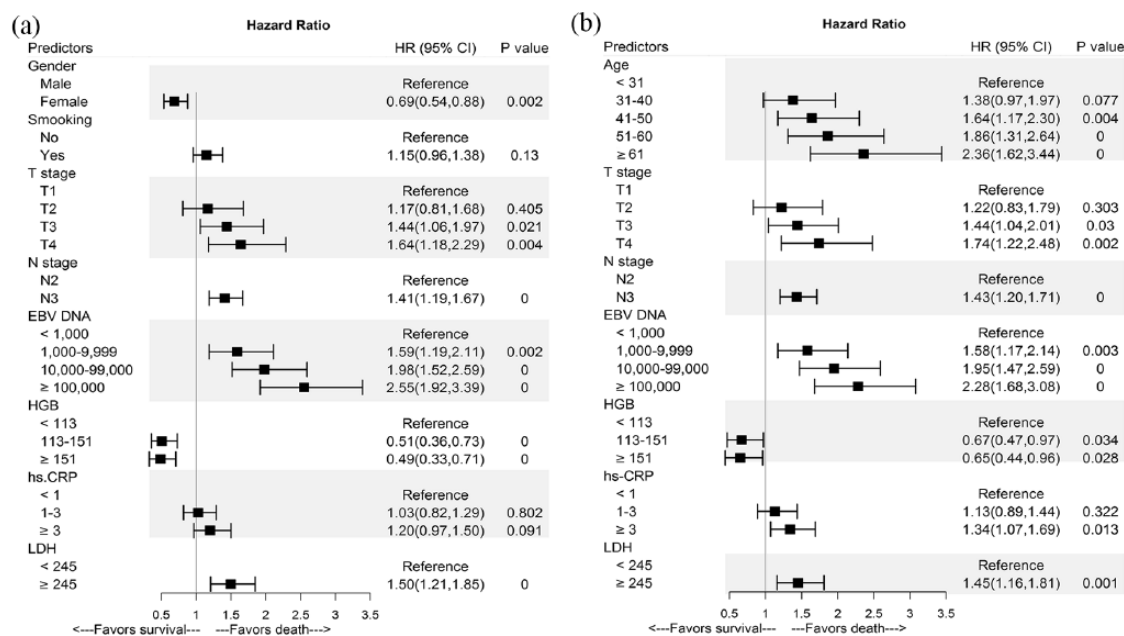
Figure 1. Forest plots depicting the multivariate association of clinicopathological characteristics with distant metastasis-free survival (a) and overall survival (b).
CI, confidence interval; EBV, Epstein–Barr virus; HGB, hemoglobin; HR, hazard ratio; hs-CRP, high-sensitivity C-reactive protein; LDH, lactate dehydrogenase.

were categorized into a low-risk group, intermediate-risk group, or high-risk group, with a total of 603, 1530, and 499 patients in each group, respectively. Based on the results of DMFS in multivariate analysis, all aforementioned validated predictors (sex, T-stage, N-stage, EBV DNA, HGB, and LDH) were selected for inclusion in the propensity score analysis in the methods section. Moreover, as serum hs-CRP has been confirmed to be an important factor of distant metastasis,[27–30] it was also incorporated into the propensity score analysis. The baseline characteristics of patients receiving either IC plus CCRT or CCRT alone were unbalanced (Table 2). Compared with the low and intermediate-risk groups, the high-risk group generally had more male patients, staged with T3–4 stage, N3 stage and IV stage disease, elevated EBV DNA, hs-CRP and LDH levels, and receiving the treatment of IC plus CCRT ($p < 0.001$ for all). Using the PSM method, three well-matched risk groups were created to compare IC plus CCRT with CCRT alone for each risk group (all $p > 0.05$). The detailed baseline characteristics for all risk groups are presented in Table 3. In addition, Supplementary Table S4 shows the number of patients with different induction, adjuvant and concurrent chemotherapy regimens for each risk group after PSM.

Figure 3 presents all direct comparisons between IC plus CCRT and CCRT alone for each risk group through the PSM method. In the high-risk group, IC plus CCRT revealed a significant improvement for DMFS compared with CCRT alone [5-year DMFS, 69.5% *versus* 56.7%, $p = 0.004$; Figure 4(a)]. In the intermediate-risk group, no significant difference was observed between IC plus CCRT and CCRT alone [$p = 0.831$; Figure 4(b)]. Additionally, IC plus CCRT tended to improve DMFS in the first 3 years after treatment compared with CCRT alone (3-year DMFS, 84.5% *versus* 83.2%), however this trend changed at year 5 (5-year DMFS, 79.8% *versus* 81.1%). In the low-risk group, patients who were treated with IC plus CCRT did not show an improvement benefit for DMFS compared with those treated with CCRT alone ($p = 0.608$). In contrast, IC plus CCRT tended to display an unfavorable DMFS in comparison with CCRT alone [3-year DMFS, 93.5% *versus* 97.0%; Figure 4(c)].

Discussion
In the current study, a prognostic nomogram based on EBV DNA, TNM staging system, and hematology and biochemistry profiling in comparison with clinical risk factors improves the ability to predict DMFS in advanced N-stage
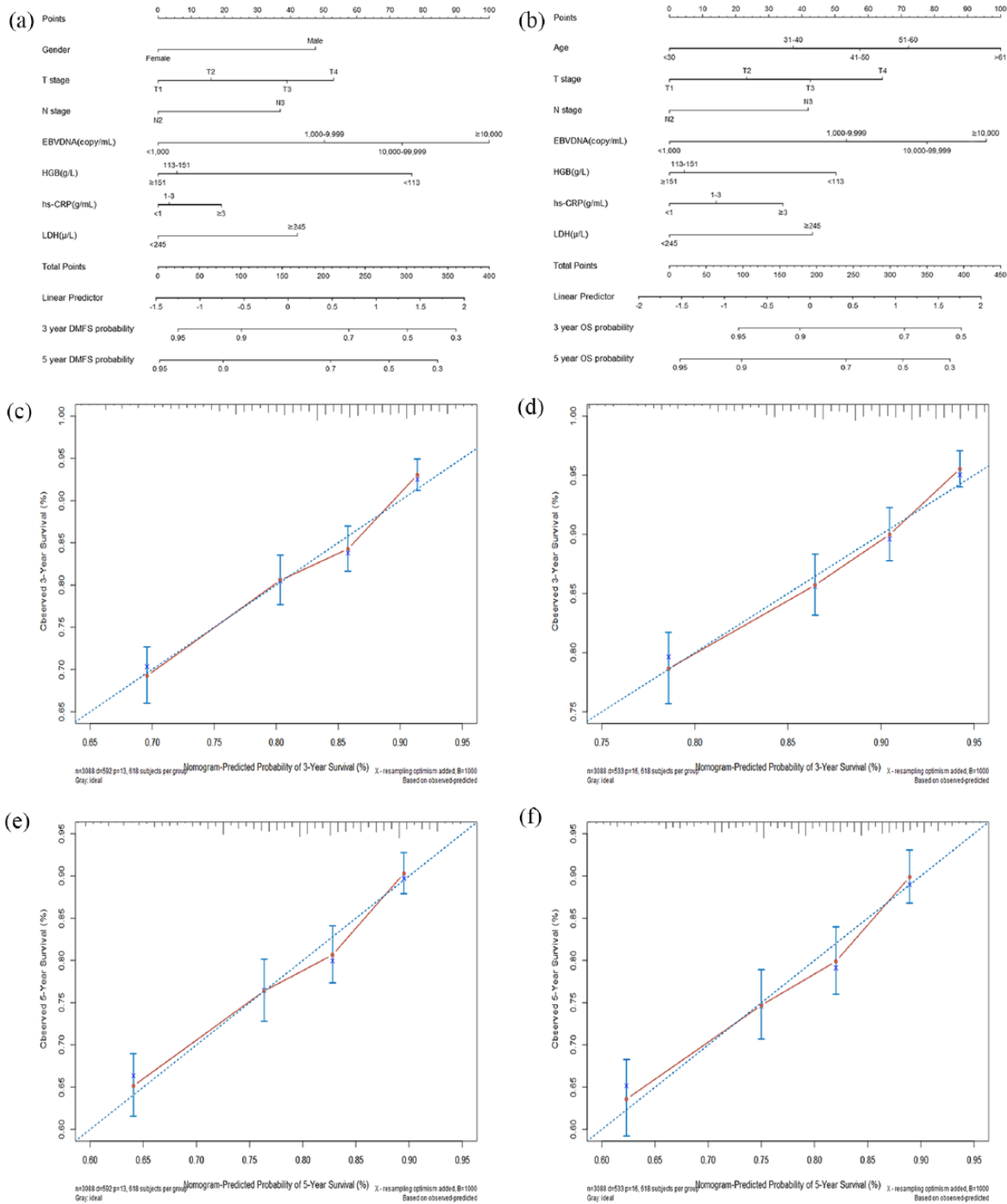
Figure 2. Prognostic nomograms (a, b) and calibration plots of survival probabilities at 3 years (c, d) and 5 years (e, f) in patients with NPC. The left panel represents the nomogram and calibration plots for distant metastasis-free survival (a, c, e); the right panel represents the nomogram and calibration plots for overall survival (b, d, f).
EBV, Epstein–Barr virus; HGB, hemoglobin; hs-CRP, high-sensitivity C-reactive protein; LDH, lactate dehydrogenase; NPC, nasopharyngeal carcinoma.

NPC patients. Our results revealed the nomograms of DMFS developed in this study categorizing patients into low, intermediate, and high-risk groups for significantly different DMFS rates. Moreover, we showed that IC plus CCRT *versus* CCRT alone provided improved survival
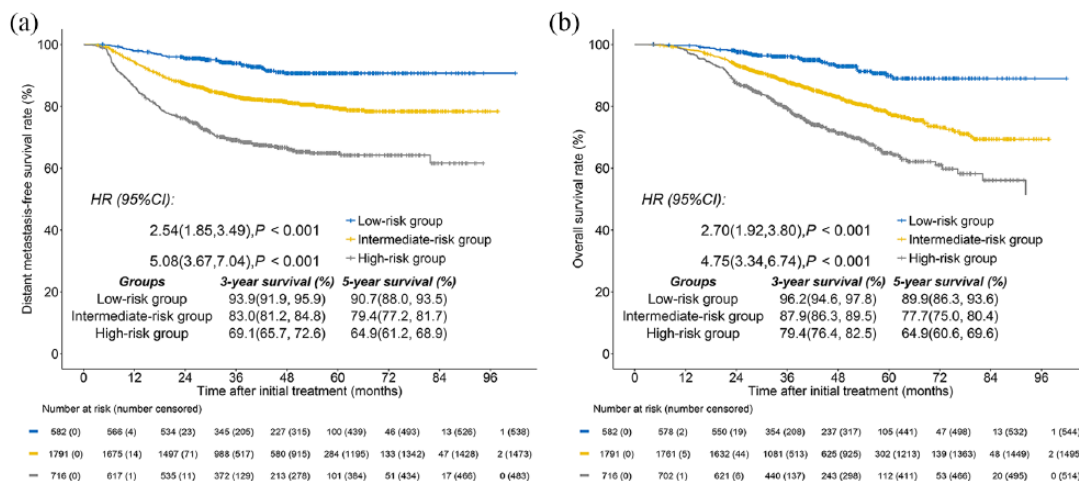
Figure 3. Kaplan–Meier survival curves are shown for (a) distant metastasis-free survival and overall survival (b) in patients with advanced N-stage NPC stratified by the total score of nomograms. Low-risk group: total score ≤ 150 points; intermediate-risk group: 150 < total score ≤ 300 points; high-risk group: total score > 300 points.
NPC, nasopharyngeal carcinoma.

Table 2. The baseline characteristics of high, intermediate, and low-risk groups in patients treated with IC plus CCRT or CCRT alone before propensity score matching.

Characteristic	High-risk group (*n* = 603)	Intermediate-risk group (*n* = 1530)	Low-risk group (*n* = 499)	*p* value
	No. (%)[a]	No. (%)[a]	No. (%)[a]	
Sex				<0.001
Male	524 (86.9)	1192 (77.9)	279 (55.9)	
Female	79 (13.1)	338 (22.1)	220 (44.1)	
Histology (WHO)				0.018
Type I–II	13 (2.2)	44 (2.9)	4 (0.8)	
Type III	590 (97.8)	1486 (97.1)	495 (99.2)	
Age, years				0.002
≤30	85 (14.1)	172 (11.2)	53 (10.6)	
31–40	148 (24.5)	382 (25.0)	159 (31.9)	
41–50	202 (33.5)	527 (34.4)	156 (31.3)	
51–60	117 (19.4)	320 (20.9)	111 (22.2)	
≥61	51 (8.5)	129 (8.4)	20 (4.0)	
Smoking history				0.018
No	336 (55.7)	885 (57.8)	362 (72.5)	
Yes	267 (44.3)	645 (42.2)	137 (27.5)	
Family of cancer				0.296
No	463 (76.8)	1162 (75.9)	364 (72.9)	
Yes	140 (23.2)	368 (24.1)	135 (27.1)	

(Continued)

Table 2. (Continued)

Characteristic	High-risk group (*n* = 603)	Intermediate-risk group (*n* = 1530)	Low-risk group (*n* = 499)	*p* value
	No. (%)[a]	No. (%)[a]	No. (%)[a]	
T-stage (8th edition)				<0.001
T1	19 (3.2)	149 (9.7)	137 (27.5)	
T2	48 (8.0)	264 (17.3)	122 (24.4)	
T3	338 (56.1)	746 (48.8)	217 (43.5)	
T4	198 (32.8)	371 (24.2)	23 (4.6)	
N-stage (8th edition)				<0.001
N2	176 (29.2)	1049 (68.6)	448 (89.8)	
N3	427 (70.8)	481 (31.4)	51 (10.2)	
Overall stage (8th edition)				<0.001
III	95 (15.8)	743 (48.6)	431 (86.4)	
IV	508 (84.2)	787 (51.4)	51 (10.2)	
EBV DNA, copy/ml[b]				<0.001
<1000	0 (0)	243 (15.9)	416 (83.4)	
1000–9999	51 (8.5)	591 (38.6)	57 (11.4)	
10,000–99,999	243 (40.3)	569 (37.2)	26 (5.2)	
≥100,000	309 (51.2)	127 (8.3)	0 (0)	
HGB, g/l[b]				<0.001
<113	58 (9.6)	45 (2.9)	2 (0.4)	
113–151	374 (62.0)	1001 (65.4)	351 (70.3)	
≥151	171 (28.4)	484 (31.6)	146 (29.3)	
hs-CRP, g/ml[b]				<0.001
<1.0	90 (14.9)	417 (27.3)	215 (43.1)	
1.0–3.0	164 (27.2)	559 (36.5)	219 (43.9)	
≥3.0	349 (57.9)	554 (36.2)	65 (13.0)	
LDH, U/l[b]				<0.001
<245	367 (60.9)	1457 (95.2)	493 (98.8)	
≥245	235 (39.0)	73 (4.8)	6 (1.2)	
Treatment				<0.001
IC plus CCRT	484 (80.3)	1027 (67.1)	272 (54.5)	
CCRT alone	119 (19.7)	503 (32.9)	227 (45.5)	

CCRT, concurrent chemoradiotherapy; EBV, Epstein–Barr virus; HGB, hemoglobin; hs-CRP, high-sensitivity C-reactive protein; IC, induction chemotherapy; LDH, lactate dehydrogenase; NPC, nasopharyngeal carcinoma; WHO, World Health Organization.
[a]Percentages may not add up to 100, due to rounding.
[b]All variables were measured before treatment.

Table 3. The baseline characteristics of the patients treated with IC plus CCRT or CCRT alone in each risk group based on the propensity score matching method.

Characteristic	High-risk group			Intermediate-risk group			Low-risk group		
	CCRT alone (*n* = 102)	IC + CCRT (*n* = 304)	*p* value	CCRT alone (*n* = 449)	IC + CCRT (*n* = 867)	*p* value	CCRT alone (*n* = 203)	IC + CCRT (*n* = 230)	*p* value
	No. (%)[a]	No. (%)[a]		No. (%)[a]	No. (%)[a]		No. (%)[a]	No. (%)[a]	
Sex			0.599			0.657			0.275
Male	14 (13.7)	35 (11.5)		88 (19.6)	161 (18.6)		81 (39.9)	80 (34.8)	
Female	88 (86.3)	269 (88.5)		361 (80.4)	706 (81.4)		122 (60.1)	150 (65.2)	
T-stage (8th edition)			0.571			0.093			0.266
T1	1 (1.0)	2 (0.7)		48 (10.7)	64 (7.4)		53 (26.1)	52 (22.6)	
T2	5 (4.9)	10 (3.3)		75 (16.7)	146 (16.8)		57 (28.1)	52 (22.6)	
T3	72 (70.6)	206 (67.8)		246 (54.8)	466 (53.7)		92 (45.3)	123 (53.5)	
T4	24 (23.5)	86 (28.3)		80 (17.8)	191 (22.0)		1 (0.5)	3 (1.3)	
N-stage (8th edition)			0.316			0.612			0.862
N2	34 (33.3)	85 (28.0)		319 (71.0)	604 (69.7)		187 (92.1)	210 (91.3)	
N3	68 (66.7)	219 (72.0)		130 (29.0)	263 (30.3)		16 (7.9)	20 (8.7)	
EBV DNA, copy/ml[b]			0.758			0.939			0.956
<1000	0 (0)	0 (0)		60 (13.4)	117 (13.5)		170 (83.7)	195 (84.8)	
1000–9999	6 (5.9)	16 (5.3)		190 (42.3)	352 (40.6)		25 (12.3)	26 (11.3)	
10,000–99,999	46 (45.1)	150 (49.3)		169 (37.6)	340 (39.2)		8 (3.9)	9 (3.9)	
≥100,000	50 (49.0)	138 (45.4)		30 (6.7)	58 (6.7)		0 (0)	0 (0)	
HGB, g/l[b]			0.249			0.597			0.051
<113	5 (4.9)	7 (2.3)		6 (1.3)	7 (0.7)		0 (0)	0 (0)	
113–151	73 (71.6)	209 (68.8)		298 (66.4)	585 (67.5)		147 (72.4)	146 (63.5)	
≥151	24 (23.5)	88 (28.9)		145 (32.3)	276 (31.8)		56 (27.6)	84 (36.5)	
hs-CRP, g/ml[b]			0.906			0.722			0.661
<1.0	14 (13.7)	42 (13.8)		118 (26.3)	217 (25.0)		87 (42.9)	93 (40.4)	
1.0–3.0	25 (24.5)	81 (26.6)		163 (36.3)	334 (38.5)		93 (45.8)	115 (50.0)	
≥3.0	63 (61.8)	181 (59.5)		168 (37.4)	316 (36.4)		23 (11.3)	22 (9.6)	
LDH, U/l[b]			0.085			0.487			0.999
<245	63 (61.8)	216 (71.1)		440 (98.0)	855 (98.6)		203 (100)	230 (100)	
≥245	39 (38.2)	88 (28.9)		9 (2.0)	12 (1.4)		0 (0)	0 (0)	

CCRT, concurrent chemoradiotherapy; EBV, Epstein–Barr virus; HGB, hemoglobin; hs-CRP, high-sensitivity C-reactive protein; IC, induction chemotherapy; LDH, lactate dehydrogenase; NPC, nasopharyngeal carcinoma.
[a]Percentages may not add up to 100, due to rounding.
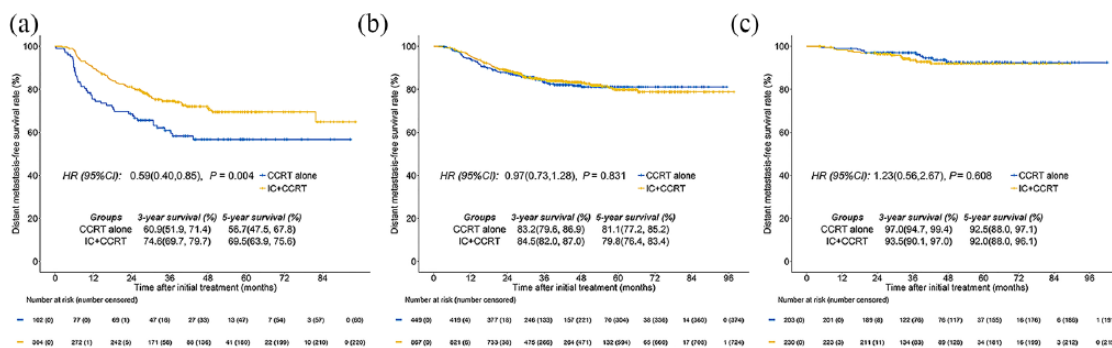[b]All variables were measured before treatment.

Figure 4. Kaplan–Meier curves of distant metastasis-free survival for the high-risk group (a), intermediate-risk group (b), and low-risk group (c) with IC plus CCRT or CCRT alone.
IC, induction chemotherapy; CCRT, concurrent chemoradiotherapy.

benefit among high-risk patients through the PSM method, but not in those classified in the intermediate-risk and low-risk groups.

Cancer is a heterogeneous disease. In addition to the anatomical extent of NPC, classic tumor-related prognostic factors, such as EBV DNA,[31] HGB,[32] hs-CRP,[27,28] and LDH,[20] were reported to independently correlate with treatment outcomes. Among these factors, we found EBV DNA was the most related prognostic factor for predicting distant failure. In fact, EBV infection is an important causative factor for NPC in endemic regions. Additionally, inherent genetic heterogeneity of NPC patients suggests inter-individual variations in EBV infection susceptibility, and EBV's ability to prompt malignant transformation.[33] Additionally, Lo and colleagues[34] found that for every 10-fold rise in plasma EBV DNA level to predict clinical events, the relative risk was 3.8 (95% CI, 1.6–9.2). For this reason, EBV DNA is potentially the ideal biomarker to complement the current TNM stage in the future.

To predict the risk of distant metastasis, we built a nomogram in the current study. Though such methods largely rely on traditional prognostic factors, such as T-stage and N-stage which reflects the anatomical information of tumors. We suggest that plasma EBV DNA reflects the tumor heterogeneity,[35] whereas blood parameters (e.g. hs-CRP and LDH) provide greater insights into the systemic inflammatory status of patients, and the increased concentrations of these are related to inferior NPC survival.[20,27] To support this, our results indicated that the established nomograms including the anatomical information of tumors, EBV DNA, and blood parameters, predicted treatment outcomes with good accuracy. Moreover,

patients grouped by high, intermediate, and low-risk groups categorized by the nomograms provided excellent discrimination in distant metastasis. Hence, our nomograms might provide a simple and precise method of predicting distant failure in advanced N-stage NPC.

In this study, the individualized risk stratification of nomograms was applied to identify subgroups where it would be advantageous for the addition of IC to CCRT to improve survival. Our results confirmed that IC for high-risk patients was advantageous for improving distant control but this was not observed in those classified as low-risk patients, despite the balanced characteristics between the two arms. The most likely reason for IC benefiting high-risk patients only in the present study is the result of bulky or extensive nodal disease, which has an elevated risk for distant metastasis.[13] Another explanation for the lower efficacy of IC in the low-risk patient group might have been the delay of RT start time for patients receiving IC before CCRT, resulting in a sub-additive effect. Although there was no difference in survival between IC and CCRT together *versus* CCRT alone for the intermediate-risk patients, IC plus CCRT achieved a higher DMFS rate than CCRT alone in the first 3 years after treatment, while this trend reversed at 5 years. Therefore, it is difficult to determine which treatment strategy is more appropriate for this group of patients, and clinical trials should be considered in this subgroup in the future.

A main strength in the present study was the head-to-head comparison among patients given CCRT with or without IC through risk stratification of nomograms and the PSM method in each group. This addressed both divergent confounders and

the selection bias associated with observational data retrospective analysis.[36] Another strength was the large-scale data derived from real-world medical records, which reflect the actual medical treatment process and the patient's health under real conditions. Nevertheless, there are some limitations that must be noted. First, although these nomograms presented excellent agreement between calibration plots and satisfactory c-indexes, external validation was not performed due to a deficiency in the available per-patient data from other hospitals within the endemic area. For this reason, the generalizability of the results to other patient populations is not certain. Another limitation was the heterogeneity of the IC regimen and dosage because of the retrospective study design. Finally, since the intelligence platform of our center failed to collect the data on acute and late toxicities, treatment-related toxicities were lacking in the current study.

In conclusion, we developed a nomogram to predict DMFS rates for advanced N-stage NPC patients in the endemic area, which provided an improved ability to predict DMFS. The low, intermediate, or high-risk groups categorized by the proposed nomograms presented excellent discrimination in DMFS rates. Although the addition of IC to CCRT provided survival benefits for high-risk patients, it failed to retain benefits in those classified as intermediate and low-risk patients. Considering the potential morbidity, cost, and inconvenience of IC, the intermediate and low-risk groups may prefer to avoid IC, given the relatively small expected benefit. The present study can serve as a catalyst for treatment discussions and facilitate informed decision-making.

Author's Contribution
Ji-Jin Yao, Ya-Nan Jin and Zhi-Gang Liu contributed equally to this work.

Conflict of interest statement
The authors declare that there is no conflict of interest.

ORCID iD
Ji-Jin Yao [iD] https://orcid.org/0000-0003-1551-5798

Supplemental material
Supplemental material for this article is available online.

References

1. Torr e LA, Bray F, Siegel RL, *et al*. Global cancer statistics, 2012. *CA Cancer J Clin* 2015; 65: 87–108.

2. Cao SM, Simons MJ and Qian CN. The prevalence and prevention of nasopharyngeal carcinoma in China. *Chin J Cancer* 2011; 30: 114–119.

3. Lai SZ, Li WF, Chen L, *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: 661–668.

4. Zhang MX, Li J, Shen GP, *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: 2587–2595.

5. Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin*. 2017; 67(2): 93–99.

6. Mao YP, Tang LL, Chen L, *et al*. Prognostic factors and failure patterns in non-metastatic nasopharyngeal carcinoma after intensity-modulated radiotherapy. *Chin J Cancer* 2016; 35: 103.

7. Foo KF, Tan EH, Leong SS, *et al.* Gemcitabine in metastatic nasopharyngeal carcinoma of the undifferentiated type. *Ann Oncol* 2002; 13: 150–156.

8. Jiang Y, Wei YQ, Luo F, *et al.* Gemcitabine and cisplatin in advanced nasopharyngeal carcinoma: a pilot study. *Cancer Invest* 2005; 23: 123–128.

9. Zhang L, Huang Y, Hong S, *et al.* Gemcitabine plus cisplatin versus fluorouracil plus cisplatin in recurrent or metastatic nasopharyngeal carcinoma: a multicentre, randomised, open-label, phase 3 trial. *Lancet* 2016; 388: 1883–1892.

10. Baujat B, Audry H, Bourhis J, *et al.* Chemotherapy in locally advanced nasopharyngeal carcinoma: an individual patient data meta-analysis of eight randomized trials and 1753 patients. *Int J Radiat Oncol Biol Phys* 2006; 64: 47–56.

11. Blanchard P, Lee A, Marguet S, *et al.* Chemotherapy and radiotherapy in nasopharyngeal carcinoma: an update of the MAC-NPC meta-analysis. *Lancet Oncol* 2015; 16: 645–655.

12. Chen YP, Wang ZX, Chen L, *et al.* A Bayesian network meta-analysis comparing concurrent chemoradiotherapy followed by adjuvant chemotherapy, concurrent chemoradiotherapy alone and radiotherapy alone in patients with locoregionally advanced nasopharyngeal carcinoma. *Ann Oncol* 2015; 26: 205–211.

13. Lin JC, Liang WM, Jan JS, *et al.* Another way to estimate outcome of advanced nasopharyngeal carcinoma–is concurrent chemoradiotherapy adequate? *Int J Radiat Oncol Biol Phys* 2004; 60: 156–164.

14. Chen L, Hu CS, Chen XZ, *et al.* Concurrent chemoradiotherapy plus adjuvant chemotherapy versus concurrent chemoradiotherapy alone in patients with locoregionally advanced nasopharyngeal carcinoma: a phase 3 multicentre randomised controlled trial. *Lancet Oncol* 2012; 13: 163–171.

15. Sun Y, Li WF, Chen NY, *et al.* Induction chemotherapy plus concurrent chemoradiotherapy versus concurrent chemoradiotherapy alone in locoregionally advanced nasopharyngeal carcinoma: a phase 3, multicentre, randomised controlled trial. *Lancet Oncol* 2016; 17: 1509–1520.

16. Fountzilas G, Ciuleanu E, Bobos M, *et al.* Induction chemotherapy followed by concomitant radiotherapy and weekly cisplatin versus the same concomitant chemoradiotherapy in patients with nasopharyngeal carcinoma: a randomized phase II study conducted by the Hellenic Cooperative Oncology Group (HeCOG) with biomarker evaluation. *Ann Oncol* 2012; 23: 427–435.

17. Lv JW, Chen YP, Huang XD, *et al.* Hepatitis B virus screening and reactivation and management of patients with nasopharyngeal carcinoma: a large-scale, big-data intelligence platform-based analysis from an endemic area. *Cancer* 2017; 123: 3540–3549.

18. Tang LQ, Chen QY, Fan W, *et al.* Prospective study of tailoring whole-body dual-modality [18F] fluorodeoxyglucose positron emission tomography/computed tomography with plasma Epstein-Barr virus DNA for detecting distant metastasis in endemic nasopharyngeal carcinoma at initial staging. *J Clin Oncol* 2013; 31: 2861–2869.

19. Shao JY, Zhang Y, Li YH, *et al.* Comparison of Epstein-Barr virus DNA level in plasma, peripheral blood cell and tumor tissue in nasopharyngeal carcinoma. *Anticancer Res* 2004; 24: 4059–4066.

20. Zhou GQ, Tang LL, Mao YP, *et al.* Baseline serum lactate dehydrogenase levels for patients treated with intensity-modulated radiotherapy for nasopharyngeal carcinoma: a predictor of poor prognosis and subsequent liver metastasis. *Int J Radiat Oncol Biol Phys* 2012; 82: e359–e365.

21. Allin KH, Bojesen SE and Nordestgaard BG. Baseline C-reactive protein is associated with incident cancer and survival in patients with cancer. *J Clin Oncol* 2009; 27: 2217–2224.

22. Chua DT, Sham J and Choy D. Prognostic impact of hemoglobin levels on treatment outcome in patients with nasopharyngeal carcinoma treated with sequential chemoradiotherapy or radiotherapy alone. *Cancer* 2004; 60: S490.

23. Tang LQ, Li CF, Li J, *et al.* Establishment and validation of prognostic nomograms for endemic nasopharyngeal carcinoma. *J Natl Cancer Inst* 2015; 108: djv291.

24. Kaplan EL and Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958; 53: 457–481.

25. Harrell FE Jr, Lee KL and Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996; 15: 361–387.

26. D'Agostino RB Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* 1998; 17: 2265–2281.

27. Xia WX, Zhang HB, Shi JL, *et al.* A prognostic model predicts the risk of distant metastasis and death for patients with nasopharyngeal carcinoma based on pre-treatment serum C-reactive protein and N-classification. *Eur J Cancer* 2013; 49: 2152–2160.

28. Xia WX, Ye YF, Lu X, *et al*. The impact of baseline serum C-reactive protein and C-reactive protein kinetics on the prognosis of metastatic nasopharyngeal carcinoma patients treated with palliative chemotherapy. *PLoS One* 2013; 8: e76958.

29. Allin KH and Nordestgaard BG. Elevated C-reactive protein in the diagnosis, prognosis, and cause of cancer. *Crit Rev Clin Lab Sci* 2011; 48: 155–170.

30. Mahmoud FA and Rivera NI. The role of C-reactive protein as a prognostic indicator in advanced cancer. *Curr Oncol Rep* 2002; 4: 250–255.

31. Lin JC, Wang WY, Chen KY, *et al*. Quantification of plasma Epstein-Barr virus DNA in patients with advanced nasopharyngeal carcinoma. *N Engl J Med* 2004; 350: 2461–2470.

32. Chua DT, Sham JS and Choy DT. Prognostic impact of hemoglobin levels on treatment outcome in patients with nasopharyngeal carcinoma treated with sequential chemoradiotherapy or radiotherapy alone. *Cancer* 2004; 101: 307–316.

33. Young LS and Rickinson AB. Epstein-Barr virus: 40 years on. *Nat Rev Cancer* 2004; 4: 757–768.

34. Lo YM, Chan AT, Chan LY, *et al*. Molecular prognostication of nasopharyngeal carcinoma by quantitative analysis of circulating Epstein-Barr virus DNA. *Cancer Res* 2000; 60: 6878–6881.

35. Leung SF, Zee B, Ma BB, *et al*. Plasma Epstein-Barr viral deoxyribonucleic acid quantitation complements tumor-node-metastasis staging prognostication in nasopharyngeal carcinoma. *J Clin Oncol* 2006; 24: 5414–5418.

36. Austin PC. The relative ability of different propensity score methods to balance measured covariates between treated and untreated subjects in observational studies. *Med Decis Making* 2009; 29: 661–677.