

## ORIGINAL ARTICLE

# Development and validation of a web-based calculator to predict individualized conditional risk of site-specific recurrence in nasopharyngeal carcinoma: Analysis of 10,058 endemic cases

Chen-Fei Wu<sup>1,†</sup> | Jia-Wei Lv<sup>1,†</sup> | Li Lin<sup>1,†</sup> | Yan-Ping Mao<sup>1</sup>  | Bin Deng<sup>2</sup> | Wei-Hong Zheng<sup>1</sup> | Dan-Wan Wen<sup>1</sup> | Yue Chen<sup>1</sup> | Jia Kou<sup>1</sup> | Fo-Ping Chen<sup>1</sup> | Xing-Li Yang<sup>1</sup> | Zi-Qi Zheng<sup>1</sup> | Zhi-Xuan Li<sup>1</sup> | Si-Si Xu<sup>1</sup> | Jun Ma<sup>1</sup> | Ying Sun<sup>1</sup> 

<sup>1</sup> 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, Guangdong 510060, P. R. China

<sup>2</sup> Department of Radiation Oncology, Wuzhou Red Cross Hospital, Wuzhou, Guangxi 543002, P. R. China

## Correspondence

Ying Sun, Ph.D., Department of Radiation Oncology, Sun Yat-sen University Cancer Center, 651 Dongfeng Road East, Guangzhou 510060, Guangdong, P. R. China.

Email: [sunying@sysucc.org.cn](mailto:sunying@sysucc.org.cn)

<sup>†</sup>These authors contributed equally to this work.

## Funding information

National Natural Science Foundation of China, Grant/Award Numbers: 81872463, 81930072; Special Support Program of Sun Yat-sen University, Grant/Award Number: 16zxtzlc06; Key-Area Research and Development Program of Guangdong Province, Grant/Award Numbers: 2019A1515012045,

## Abstract

**Background:** Conditional survival (CS) provides dynamic prognostic estimates by considering the patients existing survival time. Since CS for endemic nasopharyngeal carcinoma (NPC) is lacking, we aimed to assess the CS of endemic NPC and establish a web-based calculator to predict individualized, conditional site-specific recurrence risk.

**Methods:** Using an NPC-specific database with a big-data intelligence platform, 10,058 endemic patients with non-metastatic stage I–IVA NPC receiving intensity-modulated radiotherapy with or without chemotherapy between April 2009 and December 2015 were investigated. Crude CS estimates of conditional overall survival (COS), conditional disease-free survival (CDFS), conditional locoregional relapse-free survival (CLRRFS), conditional distant metastasis-free survival (CDMFS), and conditional NPC-specific survival (CNPC-SS) were calculated. Covariate-adjusted CS estimates were generated using inverse probability

**Abbreviations:** ASD, absolute standardized difference; Bone-M, bone metastasis; CCRT, concurrent chemoradiotherapy; CDFS, conditional disease-free survival; CDMFS, conditional distant metastasis-free survival; CLRRFS, conditional locoregional relapse-free survival; CNPC-SS, conditional nasopharyngeal carcinoma-specific survival; COS, conditional overall survival; CS, conditional survival; DFS, disease-free survival; DMFS, distant metastasis-free survival; EBV, Epstein-Barr virus; IMRT, intensity-modulated radiotherapy; IPW, inverse probability weighting; LR, local relapse; LRRFS, locoregional relapse-free survival; Liver-M, liver metastasis; Lung-M, lung metastasis; NCCN, National Comprehensive Cancer Network; NPC, nasopharyngeal carcinoma; NPC-SD, nasopharyngeal carcinoma-specific death; NPC-SS, nasopharyngeal carcinoma-specific survival; OC-SD, other cause-specific death; OS, overall survival; RDD, research data deposit; RR, regional relapse; WHO, World Health Organization.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *Cancer Communications* published by John Wiley & Sons Australia, Ltd. on behalf of Sun Yat-sen University Cancer Center

2019B020230002; Natural Science Foundation of Guangdong Province, Grant/Award Number: 2017A030312003; Health & Medical Collaborative Innovation Project of Guangzhou City, China, Grant/Award Number: 201803040003; Innovation Team Development Plan of the Ministry of Education, Grant/Award Number: IRT\_17R110; Overseas Expertise Introduction Project for Discipline Innovation, Grant/Award Numbers: 111 Project, B14035

weighting. A prediction model was established using competing risk models and was externally validated with an independent, non-metastatic stage I–IVA NPC cohort undergoing intensity-modulated radiotherapy with or without chemotherapy ( $n = 601$ ) at another institution.

**Results:** The median follow-up of the primary cohort was 67.2 months. The 5-year COS, CDFS, CLRRFS, CDMFS, and CNPC-SS increased from 86.2%, 78.1%, 89.8%, 87.3%, and 87.6% at diagnosis to 87.3%, 87.7%, 94.4%, 96.0%, and 90.1%, respectively, for an existing survival time of 3 years since diagnosis. Differences in CS estimates between prognostic factor subgroups of each endpoint were noticeable at diagnosis but diminished with time, whereas an ever-increasing disparity in CS between different age subgroups was observed over time. Notably, the prognoses of patients that were poor at diagnosis improved greatly as patients survived longer. For individualized CS predictions, we developed a web-based model to estimate the conditional risk of local (C-index, 0.656), regional (0.667), bone (0.742), lung (0.681), and liver (0.711) recurrence, which significantly outperformed the current staging system ( $P < 0.001$ ). The performance of this web-based model was further validated using an external validation cohort (median follow-up, 61.3 months), with C-indices of 0.672, 0.736, 0.754, 0.663, and 0.721, respectively.

**Conclusions:** We characterized the CS of endemic NPC in the largest cohort to date. Moreover, we established a web-based calculator to predict the CS of site-specific recurrence, which may help to tailor individualized, risk-based, time-adapted follow-up strategies.

#### KEYWORDS

endemic nasopharyngeal carcinoma, conditional survival, big data, web-based, individualized prediction model, overall survival, disease-free survival, locoregional relapse-free survival, distant metastasis-free survival, NPC-specific survival

## 1 | BACKGROUND

In clinical practice, clinicians estimate a patient's survival at diagnosis and consider it to be static over time [1]. However, this estimation neglects the dynamic nature of prognosis and might be inaccurate once the patients have survived for certain periods [1–5]. Therefore, conditional survival (CS), which represents a dynamically changing prognosis with time-based on a certain survival period, could be clinically more meaningful than traditional static survival estimates. CS denotes the survival probability given that a patient has already survived for a defined period of time, and it has been applied in several types of malignancies [2,4–9]. CS provides more precise prognostic information by considering the patient's existing survival time, which could assist patients to make better-informed decisions concerning their health and treatment and help clinicians make subsequent treatment and follow-up related medical decisions.

Owing to new treatment modalities, i.e., intensity-modulated radiotherapy (IMRT) and concurrent chemoradiotherapy (CCRT), the clinical outcomes of nasopharyngeal carcinoma (NPC) have remarkably improved, contributing to a longer duration of event-free follow-up [10,11]. However, a higher risk of treatment failure was observed within the first 3 posttreatment years, suggesting nonconstant hazards of treatment failure over time [12]. Therefore, it would be more clinically helpful to estimate the CS of NPC patients. Our previous research reported the CS in 7713 NPC patients diagnosed between 1973 and 2007 from the Surveillance, Epidemiology, and End Results database [13], which mainly focused on non-endemic NPC patients. As endemic and non-endemic NPC have distinct clinicopathological features, it is crucial to calculate the CS of endemic NPC patients based on large-scale and long-term follow-up data [10]. Moreover, the historical populations may not represent the current treatment modalities in the IMRT era and restrict the evaluation of

relationships between CS and emerging prognostic factors such as plasma Epstein-Barr virus (EBV) DNA load.

Based on a large NPC-specific database from our big-data intelligence platform, we performed a contemporary evaluation of CS among 10,058 endemic NPC patients with long-term follow-up treated by radical chemoradiotherapy. First, we calculated the crude CS estimates of multiple endpoints. Then, we evaluated how prognostic factors of each endpoint impacted CS after adjusting for other covariates using inverse probability weighting (IPW). Finally, we developed and externally validated a user-friendly, web-based model to predict the conditional risk of site-specific recurrence. These data could provide dynamic prognostic information to patients and guide individualized risk-based therapies and time-adapted follow-up strategies.

## 2 | MATERIALS AND METHODS

### 2.1 | Data source and study population

We conducted a retrospective cohort study using the NPC-specific database from the Big-data Intelligence Platform of Sun Yat-Sen University Cancer Center (Guangzhou, China). This database is a patient-level research system enabling real-time organization, integration, and updating of medical records automatically from several clinical systems based on well-designed data model and algorithms. It has been applied to various clinical studies [14–17]. Details of this platform were as previously described [15].

We retrieved the data of 10,058 patients with non-metastatic, biopsy-confirmed World Health Organization (WHO) type II/III NPC diagnosed cases between April 2009 and December 2015 from the NPC-specific database. All patients received routine pretreatment evaluations and were restaged according to the Union for International Cancer Control/American Joint Committee on Cancer (UICC/AJCC) Eighth edition staging system [18]. Details on the clinical workup and restaging are presented in the Supplementary Methods. The institutional ethics committee approved the study protocol and waived the requirement for informed consent given the study's retrospective nature. The study's authenticity has been validated by uploading the key raw data onto the research data deposit (RDD) public platform (<http://www.researchdata.org.cn>, approval RDD number: RDDA2018000934).

### 2.2 | Treatment

All patients received radical IMRT as primary treatment. During the study period, our institutional treatment guidelines recommended IMRT alone for stage I

patients and platinum-based CCRT ± induction/adjuvant chemotherapy for stage II-IVA patients, based on the recommendations of the updated National Comprehensive Cancer Network (NCCN) guideline for head and neck cancer version 1.2020 [19]. Reasons for deviation of patient treatment from our institutional treatment guidelines mentioned above included age, patient's refusal of treatment, or organ dysfunction suggestive of intolerance to treatment. Salvage treatments, including re-radiation, chemotherapy, and surgery, were provided during recurrence or persistent disease despite the above-mentioned treatments. The chemoradiotherapy protocols are detailed in the Supplementary Methods.

### 2.3 | Follow-up and outcomes

The patients were regularly followed-up every 3 months during the first 2 years, every 6 months during the first 3–5 years, and annually thereafter. At each visit, physical examinations, plasma EBV DNA load, and fiberoptic nasopharyngoscopy were routinely performed. MRI of the nasopharynx and neck, chest radiography, abdominal sonography, bone scan, or PET/CT were repeated annually or when clinically suspected recurrence occurred. The follow-up duration was calculated from the date of diagnosis to either date of death or last follow-up. The date of last follow-up was October 2019.

The primary endpoint was overall survival (OS), defined as the time from diagnosis to death from any cause. The secondary endpoints included disease-free survival (DFS), the time from diagnosis to tumor recurrence at any site or death from any cause, whichever came first; locoregional relapse-free survival (LRRFS), the time to locoregional relapse; distant metastasis-free survival (DMFS), the time to metastasis; and NPC-specific survival (NPC-SS), the time to death from NPC. Patients without evidence of any events of interest were censored at the time of the last follow-up.

### 2.4 | Statistical analysis

Continuous variables were transformed into categorical variables. Plasma EBV DNA load was categorized by every 10-fold increase as previously described [20–22]. The cut-off values for other laboratory variables were determined using maximally selected rank statistics based on OS, a widely used method that generates cut-off values with the most significant log-rank statistics [23–25]. The handling of cut-off values is detailed in the Supplementary Methods.

## 2.5 | Conditional survival

CS denotes the probability of survival for an additional  $y$  years given an existing survival time of  $x$  years, which is calculated as  $CS(y|x) = S(x + y)/S(x)$ . The concept of CS can extend to multiple endpoints, deriving conditional OS (COS), conditional DFS (CDFS), conditional LRRFS (CLRFS), conditional DMFS (CDMFS) and conditional NPC-SS (CNPC-SS) [1,3,26]. For example, the 3-year CDFS at 5 years demonstrates the probability of being disease-free for an additional 3 years given an existing disease-free time of 5 years.

We evaluated CS estimates for all endpoints mentioned above. Crude CS estimates were assessed using the Kaplan-Meier method. Adjusted 3-year conditional risk (1 minus CS estimates) stratified by prognostic factors of each endpoint was calculated after adjusting for other covariates using IPW, an algorithm used to balance covariates without sacrificing sample size and statistical power [27–29]; clinicopathological variables (i.e., tumor [T] stage, node [N] stage, EBV DNA, age, sex, histology, smoking, alcohol consumption, family history of NPC, lactate dehydrogenase [LDH], albumin [ALB], C-reactive protein [CRP] and hemoglobin [HGB]) that were statistically significant in IPW-adjusted log-rank tests ( $P < 0.050$ ) were considered to be the prognostic factors of the endpoints. Considering that the median follow-up time of the primary cohort was 67.2 months, we chose the 3-year conditional risk because it permitted the estimation of conditional risk given existing survival times of 0–5 years, which has been widely used in studies of CS in other malignancies with similar median follow-up times [26,30–33]. Absolute standardized difference (ASD) and chi-square tests were employed to determine the balance of baseline covariates after IPW adjustment. ASD values less than 0.1 indicated negligible imbalance [29,34]. Then, patients were categorized in a competing risk framework according to their first events (i.e., local relapse [LR], regional relapse [RR], bone metastasis [Bone-M], liver metastasis [Liver-M], lung metastasis [Lung-M], NPC-specific death [NPC-SD] and other cause-specific death [OC-SD]). The conditional risk of each competing event was calculated using the cumulative incidence function.

## 2.6 | Development and validation of prediction models

To reduce overfitting, our primary cohort was split into a derivation cohort and an internal validation cohort according to the diagnosis time. Models were constructed using cause-specific hazard models. Variables with  $P < 0.100$  in univariate analysis were subjected to multivariate analy-

sis to identify independent predictors for the models using backward selection with the Akaike information criterion. The proportional hazards assumption was verified based on the Schoenfeld residuals [35].

The models were validated internally and externally using the Harrell's C-index and calibration plots, and they were compared with the current 8<sup>th</sup> edition UICC/AJCC staging system. The bias-corrected C-index was obtained using 1000 bootstraps resamples in the derivation cohort. The models were externally validated in an independent cohort of 601 non-metastatic, biopsy-confirmed WHO type II/III NPC patients diagnosed between February 2012 and July 2015 from Wuzhou Red Cross Hospital receiving radical IMRT with or without chemotherapy to evaluate the models' general applicability. The date of last follow-up was October 2019, and other information related to this cohort is detailed in the Supplementary Methods.

## 2.7 | Handling of missing data

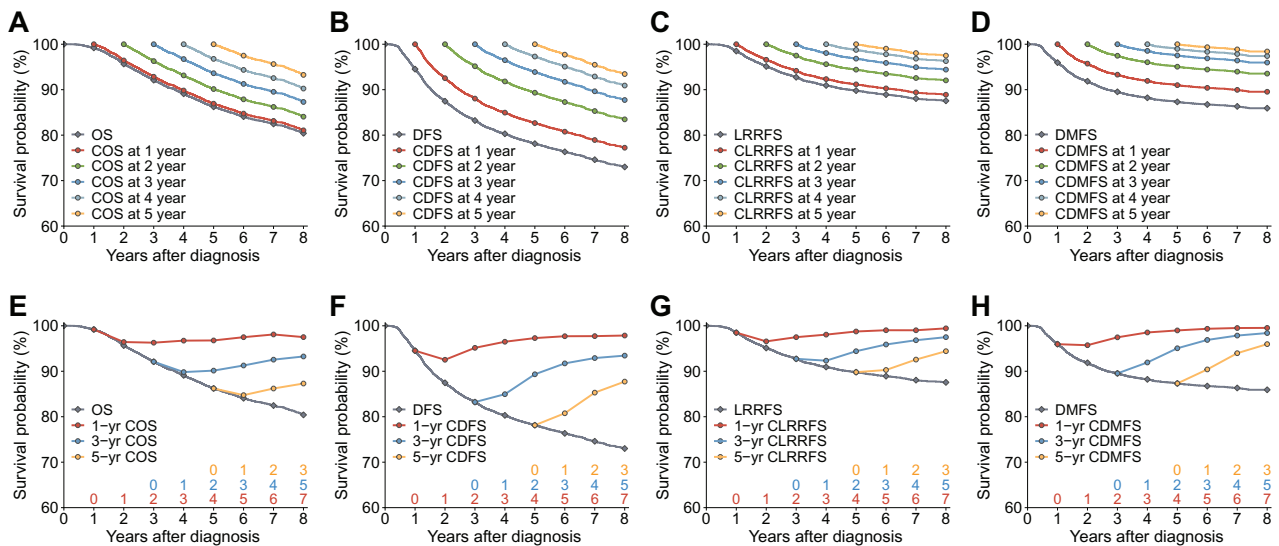
For the primary cohort of 10,058 patients, including the derivation cohort and the internal validation cohort, missing data were considered missing completely at random. Therefore, the complete-case analysis was performed, resulting in a complete-case cohort of 9302 patients. The patient characteristics were well balanced between the original cohort and the complete-case cohort (Supplementary Table S1). It is worth noting that the analyses for the crude CS estimates and the conditional risk of competing events used the primary cohort of 10,058 patients because no missing values were observed during these analyses, while other analyses utilized the complete-case cohort.

For the external validation cohort, we utilized the multiple imputations by chained equations (MICE) with the predictive mean matching method to imputed missing values [36–39]. To guarantee the robustness of the validation, 1000 imputations with 30 iterations were carried out to yield the pooled Harrell's C-index and calibration plots based on Rubin's rules [36,40,41]. All tests were two-sided, and  $P < 0.050$  was defined as statistically significant. Analyses were performed using the R software, version 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria). Detailed statistical considerations are presented in the Supplementary Methods.

## 3 | RESULTS

### 3.1 | Patient characteristics

A total of 10,058 eligible patients were investigated. All patients received IMRT, and 8839 (87.9%) of the patients



**FIGURE 1** Conditional survival of multiple endpoints in 10,058 endemic patients with non-metastatic nasopharyngeal carcinoma. Traditional Kaplan-Meier estimates of (A) OS overlaid by COS; (B) DFS overlaid by CDFS; (C) LRRFS overlaid by CLRRFS; and (D) DMFS overlaid by CDMFS. The conditional survival probabilities were estimated given the existing survival time from 1 to 5 years after diagnosis, denoted by 5 lines with different colors. Traditional Kaplan-Meier estimates of (E) OS overlaid by the 1-, 3- and 5-year COS; (F) DFS overlaid by the 1-, 3- and 5-year CDFS; (G) LRRFS overlaid by the 1-, 3- and 5-year CLRRFS; and (H) DMFS overlaid by the 1-, 3- and 5-year CDMFS. Conditional survival probabilities were estimated as a function of existing survival time after diagnosis, denoted by the numbers in the same colors above the x-axis.

Abbreviations: CDFS, conditional disease-free survival; CDMFS, conditional distant metastasis-free survival; CLRRFS, conditional locoregional relapse-free survival; COS, conditional overall survival; DFS, disease-free survival; DMFS, distant metastasis-free survival; LRRFS, locoregional relapse-free survival; OS, overall survival

received additional chemotherapy (stage I: 0 [0%], stage II: 1414 [79.7%], and stage III–IV: 7425 [96.0%]). The median follow-up time was 67.2 months (interquartile range, 53.9–83.0 months). Among them, 1059 (10.5%) patients developed locoregional relapse, 1306 (13.0%) patients developed distant metastasis, and 1581 (15.7%) patients died (1365 [13.6%] were NPC-SD). Patient characteristics are presented in [Supplementary Table S2](#).

### 3.2 | Crude CS estimates of multiple endpoints

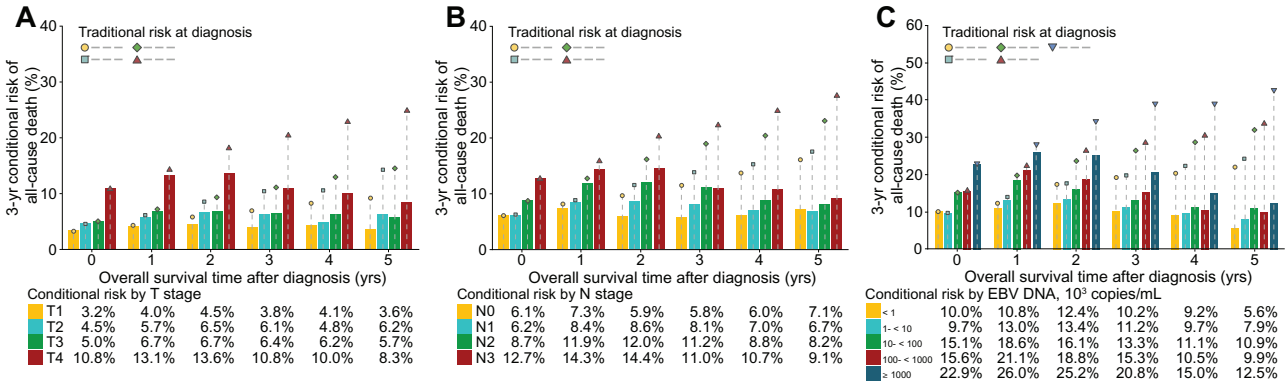
We first investigated how crude CS estimates of COS, CDFS, CLRRFS, CDMFS, and CNPC-SS changed with time. The CS probabilities improved with each additional year of existing survival time for the abovementioned endpoints (Figure 1A–D, [Supplementary Figure S1A](#)). Contrary to the ever-decreasing traditional survival estimates, the 1-, 3- and 5-year CS probabilities gradually increased as a function of the existing survival time (Figure 1E–H, [Supplementary Figure S1B](#)). Specifically, the 5-year COS, CDFS, CLRRFS, CDMFS and CNPC-SS increased from 86.2%, 78.1%, 89.8%, 87.3%, and 87.6% at diagnosis (i.e., an existing survival time of 0 year) to 87.3%, 87.7%, 94.4%,

96.0%, and 90.1%, respectively, given an existing survival time of 3 years since diagnosis ([Supplementary Table S3](#)). In addition, the 3-year CDFS, CLRRFS, and CDMFS increased considerably in the first 3 years but plateaued thereafter, indicating that the patients' prognosis improved primarily in the first 3 years (Figure 1F–H). In contrast, the 3-year COS and CNPC-SS improved continuously without plateauing until beyond the fifth year (Figure 1E, [Supplementary Figure S1B](#)). Interestingly, the almost overlapping curves of “COS at 1 year” and “OS” in Figure 1A denoted that the COS with a 1-year existing survival time barely improved compared with the baseline survival estimates at diagnosis. Likewise, CNPC-SS at 1 year also increased minimally against the baseline estimates ([Supplementary Figure S1A](#)).

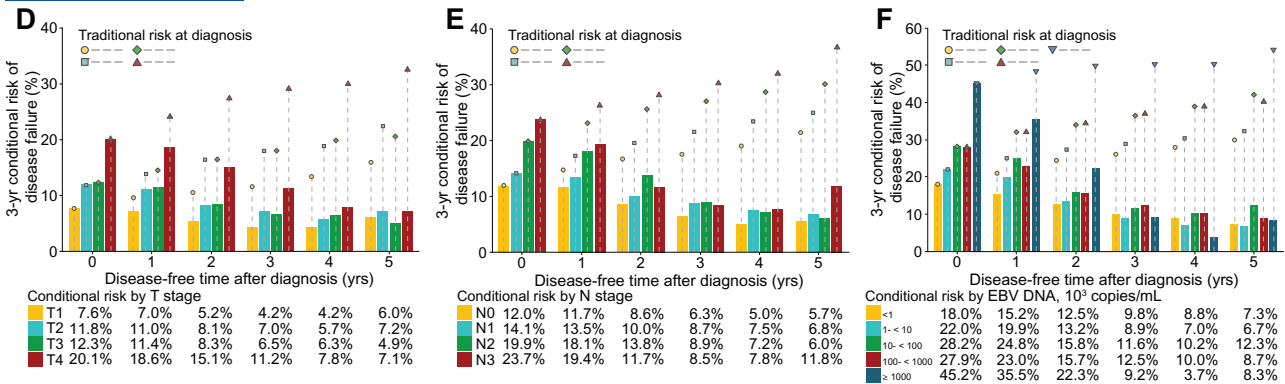
### 3.3 | IPW-adjusted conditional risk stratified by prognostic variables

To further understand how prognostic factors of NPC influenced conditional risks, we calculated the 3-year conditional risk separately stratified by the prognostic factors of each endpoint after adjusting for other covariates using IPW. The IPW algorithm-generated well-balanced

**All-cause death**



**Disease failure**



**FIGURE 2** The IPW-adjusted 3-year conditional risk for all-cause of death stratified by (A) T stage, (B) N stage and (C) plasma EBV DNA load; and disease failure (including all-cause death, LRR, and DM) stratified by (D) T stage, (E) N stage and (F) plasma EBV DNA load in 9,302 endemic patients of the complete-case cohort with non-metastatic nasopharyngeal carcinoma. The bars depict the conditional risk, and the tables below the graphs represent the value of the conditional risk. Points connected with dashed lines indicate the corresponding traditional estimates of risk calculated at diagnosis. For example, if a patient with T4 has not experienced recurrence and on follow-up is found to be disease-free, his/her 3-year conditional risk for having disease failure after 5 years from diagnosis would be 8.3%, while the traditional estimate of risk for having disease failure after 8 years from diagnosis (5 years + 3 years = 8 years) is approximately 37%. Note that all the risks were calculated using IPW-adjusted Kaplan-Meier estimates that adjusted for other covariates.

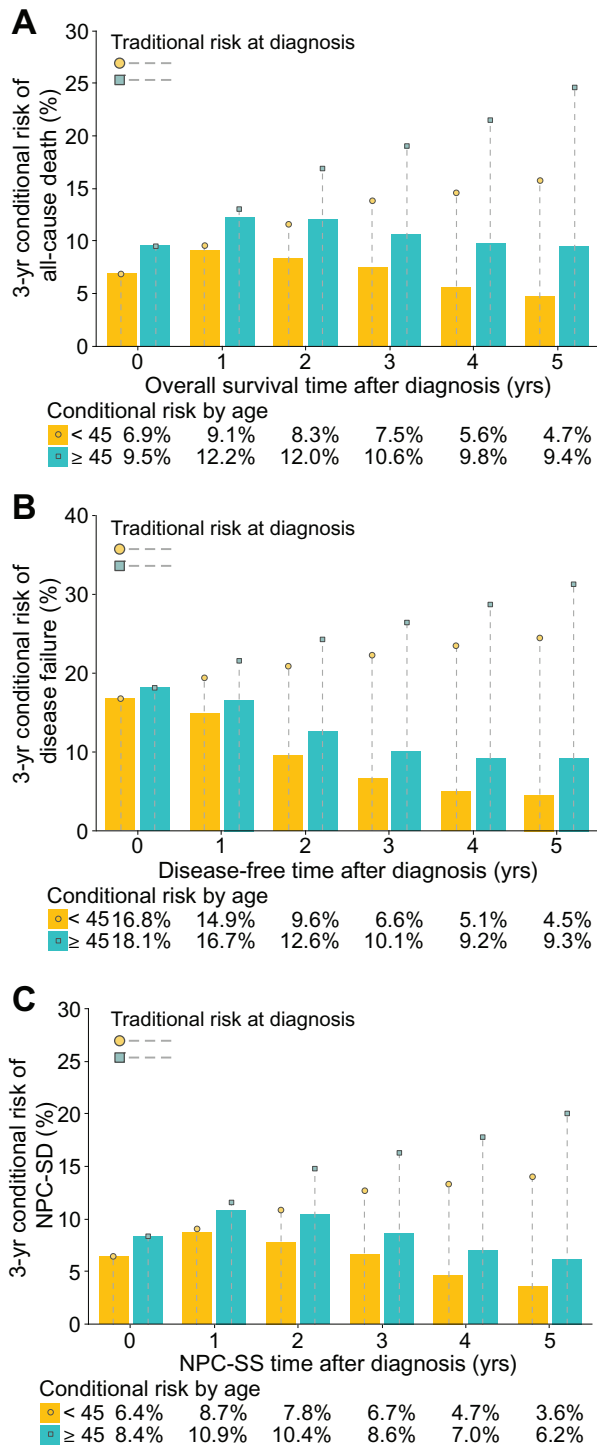
Abbreviations: EBV DNA, Epstein-Barr virus DNA; IPW, inverse probability weighting; N, node; T, tumor

cohorts with comparable baseline covariates between subgroups (Supplementary Figure S2), and the ASDs of covariates considerably decreased after IPW adjustment, with values all below or near the thresholds indicating negligible imbalance (Supplementary Table S4). The prognostic factors independently associated with each endpoint in IPW-adjusted log-rank tests are presented in Supplementary Table S5. The number at risk of each subgroup is listed in Supplementary Table S6.

For T stage, N stage, and EBV DNA, the disparities in the 3-year conditional risk of all-cause of death and disease failure among each subgroup were notable at diagnosis, but they gradually tended to level with time (Figure 2A–F), and the same trend was observed for the stratified conditional risk of locoregional relapse, distant metastasis, and NPC-SD (Supplementary Figure S3A–H). This trend occurred due to the conditional risk for patients

with an initially poorer prognosis (e.g., T4, N3, and EBV DNA  $\geq 1000 \times 10^3$  copies/mL), which decreased more substantially than for patients with an initially better prognosis. For instance, the 3-year conditional risk of disease failure for T4 decreased from 20.1% at diagnosis to 7.1% by the fifth year of disease-free survival, while those with T1 only decreased from 7.6% to 6.0% (Figure 2D).

Other prognostic factors for each endpoint (i.e. sex, histology, LDH and HGB for all-cause death, disease failure, distant metastasis and NPC-SD; smoking for disease failure and locoregional relapse; and ALB and CRP for all-cause death) showed similar patterns as mentioned above (Supplementary Figure S4A–T). However, age demonstrated enlarged disparities over time in the 3-year conditional risk of all-cause death, disease failure, and NPC-SD between subgroups (Figure 3A–C). Interestingly, this was because the conditional risk of recurrence or death for



**FIGURE 3** The IPW-adjusted 3-year conditional risk for (A) all-cause of death, (B) disease failure (including all-cause death, LRR, and DM) and (C) NPS-SD stratified by age in 9302 endemic patients of the complete-case cohort with non-metastatic nasopharyngeal carcinoma. The bars depict the conditional risk, and the tables below the graphs represent the value of the conditional risk. Points connected with dashed lines indicate the corresponding traditional estimates of risk calculated at diagnosis. For example, if a patient with an age of <45 years has not experienced recurrence and on follow-up is found to be disease-free, his 3-year conditional risk for having disease failure after 5 years from diagnosis would be 4.5%, while the traditional

the older group (initially poorer prognosis) reduced less than that for the younger group (initially better prognosis), which was opposite to other factors. A possible reason was that older patients might experience a higher prolonged risk of recurrence or death due to increased comorbidity burdens, decreased performance status, and reduced tolerance to treatments [17,42–45].

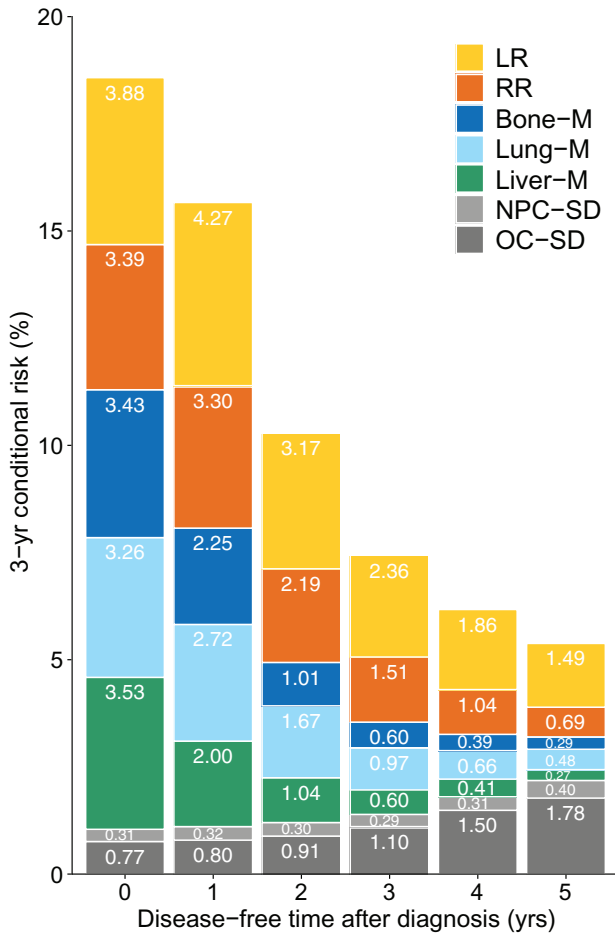
### 3.4 | Establishment of prediction models for the conditional risk of site-specific recurrence

To clarify the conditional risk of specific events, the patients were further categorized according to their first events in a competing risk framework, and the cumulative risk of each competing event was estimated (Supplementary Figure S5). We then assessed how the conditional risk of each competing event changed with the existing disease-free time. The 3-year conditional risk of LR, RR, bone-M, lung-M, and liver-M decreased along with the existing disease-free time after diagnosis, and the decrease in metastases was more remarkable than that in LR or RR (Figure 4). In contrast, the 3-year conditional risk of NPC-SD remained nearly constant, while that of OC-SD gradually increased. The 1-year and 5-year conditional risk of specific events displayed similar trends to those mentioned above (Supplementary Figure S6A and B). Notably, the 3-year conditional risk for OC-SD surpassed the risk of bone-M, lung-M, and liver-M, given a 3-year disease-free time, while that for OC-SD exceeded the risk of RR and LR at 4-year and 5-year disease-free times, respectively (Figure 4).

Given that different sites showed unique conditional risk patterns, we established competing risk models for LR, RR, Bone-M, Lung-M, and Liver-M, respectively. Patient characteristics of the derivation cohort are listed in Supplementary Table S7. Variables with  $P < 0.100$  in univariate analyses (Supplementary Table S8) were entered into multivariate analyses to select the independent predictors. The final prediction models are detailed in Table 1. To facilitate user-friendly access, the models were then integrated into a web-based application available at <https://cr-npc.yiducloud.com.cn>, which allows individualized

estimate of risk for having disease failure after 8 years from diagnosis (5 years + 3 years = 8 years) is approximately 24%. Note that all the risks were calculated using IPW-adjusted Kaplan-Meier estimates that adjusted for other covariates.

Abbreviations: DM, distant metastasis; IPW, inverse probability weighting; LRR, locoregional relapse; NPS-SD, nasopharyngeal carcinoma-specific death; NPS-SS, nasopharyngeal carcinoma-specific survival



**FIGURE 4** Three-year conditional risk of site-specific recurrence given the existing disease-free time from 0 to 5 years after diagnosis in 10,058 endemic patients with non-metastatic nasopharyngeal carcinoma. The stacked bars with different colors annotated with numbers indicate the absolute probability of each event. For example, if a patient has not experienced recurrence and on follow-up is found to be disease-free, his 3-year conditional risk for having LR after 5 years from diagnosis would be 1.49%.

Abbreviations: Bone-M, bone metastasis; Liver-M, liver metastasis; LR, local relapse; Lung-M, lung metastasis; NPC-SD, nasopharyngeal carcinoma-specific death; OC-SD, other cause-specific death; RR, regional relapse

conditional risk estimation of site-specific recurrence by entering patients' clinicopathological information and existing disease-free time.

### 3.5 | Performance of the prediction models and external validation

In the derivation and internal validation cohorts, the C-indices for the models to predict LR, RR, Bone-M, Lung-M and Liver-M were 0.656 and 0.656, 0.667 and 0.676, 0.742 and 0.709, 0.681 and 0.669, and 0.711 and 0.708, respec-

tively, which were all significantly higher than those for the current staging system ( $P < 0.001$ ), with respective values of 0.613 and 0.628, 0.581 and 0.578, 0.634 and 0.637, 0.647 and 0.616, and 0.625 and 0.615 (Table 1). The calibration plots indicated good agreement between the model-predicted and observed survival estimates in both cohorts (Supplementary Figure S7A, B, D, E, G, H, J, K, M, and N).

To investigate the general applicability of the models, they were further validated externally with an independent cohort of 601 NPC patients from Wuzhou Red Cross Hospital (median follow-up time = 61.3 months [interquartile range, 49.9-67.9 months]), with patient characteristics presented in Supplementary Table S7. Intriguingly, the C-indices for LR, RR, Bone-M, Lung-M and Liver-M were 0.672, 0.736, 0.754, 0.663, and 0.721, respectively, which significantly outperformed the current staging system ( $P < 0.050$ ), with C-indices of 0.642, 0.673, 0.659, 0.618 and 0.612 (Table 1). The calibration plots also demonstrated good agreement between the predictions and observations (Supplementary Figure S7C, F, I, L, and O).

## 4 | DISCUSSION

To our knowledge, this is the first study on CS for endemic NPC patients based on the largest cohort to date of 10,058 patients receiving IMRT with long-term follow-up. We found that the CS probabilities of endemic NPC patients improved remarkably with time, especially for those with an initially poor prognosis. Therefore, CS could provide more accurate and clinically meaningful prognostic estimates for cancer survivors than traditional survival estimates. Accordingly, we established and externally validated a web-based calculator to predict the individualized conditional risk of site-specific recurrence for the first time. Overall, our research provides dynamic prognostic information for endemic NPC patients and facilitates the establishment of individualized, risk-based, time-adapted surveillance strategies.

Previous studies on CS in NPC only investigated COS and CNPC-SS [9,13]. In this study, we extended the concept of CS to the CDFS, CLRRFS, and CDMFS, and revealed distinct patterns compared with those of COS and CNPC-SS. For COS and CNPC-SS, the CS probabilities improved continuously without plateauing until after the fifth year, which agreed with the report of Lv et al. [13] that the probabilities plateaued at 9 years after diagnosis. However, the probabilities of CDFS, CLRRFS, and CDMFS increased rapidly in the first 3 years and plateaued thereafter, indicating that patients outliving a 3-year disease-free milestone might obtain excellent long-term survival free from recurrence. This time point could be considered a potential endpoint in clinical trials



TABLE 1 Multivariate analysis and performance of the site-specific competing risk models

Variables	Competing risk model <sup>†</sup>		Model performance		
	HR (95% CI)	P	Cohorts	C-index (95% CI)	P <sup>‡</sup>
<b>Local relapse</b>					
T stage			Competing risk model		
T1	1 [Ref]		Derivation cohort	0.656 (0.628-0.684)	
T2	2.526 (1.517-4.206)	< 0.001	Internal validation cohort	0.656 (0.624-0.688)	
T3	2.892 (1.833-4.563)	< 0.001	External validation cohort	0.672 (0.578-0.766)	
T4	6.057 (3.814-9.620)	< 0.001	TNM staging system <sup>§</sup>		
LDH, IU/L			Derivation cohort	0.613 (0.587-0.639)	< 0.001
<210	1 [Ref]		Internal validation cohort	0.628 (0.597-0.659)	< 0.001
≥210	1.452 (1.153-1.829)	0.002	External validation cohort	0.642 (0.562-0.721)	0.034
ALB, g/L					
<42	1 [Ref]				
≥42	0.768 (0.603-0.978)	0.032			
<b>Regional relapse</b>					
Histology, WHO type			Competing risk model		
II	1 [Ref]		Derivation cohort	0.667 (0.635-0.698)	
III	0.546 (0.312-0.9533)	0.033	Internal validation cohort	0.676 (0.640-0.711)	
N stage			External validation cohort	0.736 (0.637-0.836)	
N0	1 [Ref]		TNM staging system <sup>§</sup>		
N1	2.131 (1.234-3.681)	0.007	Derivation cohort	0.581 (0.550-0.611)	< 0.001
N2	3.682 (2.086-6.501)	< 0.001	Internal validation cohort	0.578 (0.541-0.616)	< 0.001
N3	4.494 (2.465-8.193)	< 0.001	External validation cohort	0.673 (0.588-0.758)	0.017
EBV DNA, × 10 <sup>3</sup> copies/mL					
<1	1 [Ref]				
1 – <10	1.569 (1.120-2.197)	0.009			
10 – <100	1.997 (1.454-2.743)	< 0.001			
100 – <1000	1.612 (1.053-2.468)	0.028			
≥1000	3.571 (1.862-6.850)	< 0.001			
<b>Bone metastasis</b>					
Sex			Competing risk model		
Male	1 [Ref]		Derivation cohort	0.742 (0.711-0.772)	
Female	0.424 (0.281-0.639)	< 0.001	Internal validation cohort	0.709 (0.675-0.743)	
N stage			External validation cohort	0.754 (0.668-0.839)	
N0	1 [Ref]		TNM staging system <sup>§</sup>		
N1	1.859 (0.920-3.755)	0.084	Derivation cohort	0.634 (0.602-0.666)	< 0.001
N2	3.446 (1.681-7.065)	< 0.001	Internal validation cohort	0.637 (0.603-0.671)	< 0.001
N3	3.609 (1.700-7.664)	< 0.001	External validation cohort	0.659 (0.556-0.761)	0.014
EBV DNA, × 10 <sup>3</sup> copies/mL					
<1	1 [Ref]				
1 – <10	2.753 (1.765-4.294)	< 0.001			
10 – <100	2.948 (1.905-4.563)	< 0.001			
100 – <1000	4.510 (2.804-7.254)	< 0.001			
≥1000	7.584 (3.780-15.214)	< 0.001			
LDH, IU/L					
<210	1 [Ref]				
≥210	1.425 (1.059-1.918)	0.020			

(Continues)

TABLE 1 (Continued)

Variables	Competing risk model <sup>†</sup>		Model performance		
	HR (95% CI)	<i>P</i>	Cohorts	C-index (95% CI)	<i>P</i> <sup>‡</sup>
<b>Liver metastasis</b>					
Sex			Competing risk model		
Male	1 [Ref]		Derivation cohort	0.681 (0.645-0.717)	
Female	0.626 (0.434-0.902)	0.012	Internal validation cohort	0.669 (0.632-0.705)	
T stage			External validation cohort	0.663 (0.550-0.775)	
T1	1 [Ref]		TNM staging system <sup>§</sup>		
T2	1.759 (0.970-3.187)	0.063	Derivation cohort	0.647 (0.614-0.680)	< 0.001
T3	1.799 (1.057-3.063)	0.031	Internal validation cohort	0.616 (0.580-0.652)	< 0.001
T4	2.115 (1.201-3.723)	0.010	External validation cohort	0.618 (0.516-0.720)	0.040
N stage					
N0	1 [Ref]				
N1	2.026 (1.005-4.084)	0.048			
N2	3.241 (1.571-6.683)	0.002			
N3	5.787 (2.773-12.077)	< 0.001			
EBV DNA, × 10 <sup>3</sup> copies/mL					
<1	1 [Ref]				
1 – <10	1.337 (0.861-2.077)	0.196			
10 – <100	2.542 (1.742-3.709)	< 0.001			
100 – <1000	2.182 (1.368-3.481)	0.001			
≥1000	4.296 (2.206-8.367)	< 0.001			
HGB, g/L					
<125	1 [Ref]				
≥125	0.565 (0.373-0.856)	0.007			
<b>Lung metastasis</b>					
Sex			Competing risk model		
Male	1 [Ref]		Derivation cohort	0.711 (0.675-0.746)	
Female	0.596 (0.424-0.839)	0.003	Internal validation cohort	0.708 (0.673-0.742)	
T stage			External validation cohort	0.721 (0.597-0.845)	
T1	1 [Ref]		TNM staging system <sup>§</sup>		
T2	1.012 (0.558-1.836)	0.968	Derivation cohort	0.625 (0.591-0.658)	< 0.001
T3	1.452 (0.899-2.346)	0.128	Internal validation cohort	0.615 (0.582-0.649)	< 0.001
T4	3.222 (1.978-5.247)	< 0.001	External validation cohort	0.612 (0.508-0.715)	0.005
N stage					
N0	1 [Ref]				
N1	1.776 (1.039-3.036)	0.036			
N2	2.204 (1.241-3.915)	0.007			
N3	2.660 (1.426-4.960)	0.002			
EBV DNA, × 10 <sup>3</sup> copies/mL					
<1	1 [Ref]				
1 – <10	1.366 (0.944-1.976)	0.098			
10 – <100	1.565 (1.099-2.230)	0.013			
100 – <1000	1.685 (1.083-2.623)	0.021			
≥1000	3.441 (1.679-7.052)	< 0.001			

Boldfaced *P*-value indicates statistical significance.

Abbreviations: ALB, Albumin; CI, confidence interval; EBV DNA, Epstein-Barr virus DNA; HGB, hemoglobin; HR, hazard ratio; LDH, lactate dehydrogenase; N, node; Ref, reference; T, tumor.

<sup>†</sup>Competing risk model was constructed based on the cause-specific hazard model.

<sup>‡</sup>*P*-value refers to the comparison of the Harrell's C-index between the competing risk model and the TNM staging system.

<sup>§</sup>According to the 8th edition of the AJCC/UICC Staging System.

because it symbolizes the necessary observation period for patients to obtain long-term curability [2,9]. Furthermore, as surveillance protocols are primarily formulated based on recurrence data, our results suggest that the follow-up for NPC patients can be significantly less intensive after a 3-year disease-free interval. Considering that the surveillance strategies for NPC patients recommended by the current NCCN guidelines for head and neck cancer version 1.2020 are hardly justified based on time-adapted prognostic information [19], the dynamic recurrence data we provided could serve as valuable evidence in developing optimal follow-up frequency and duration.

Given the limited data on the relationships between prognostic factors of multiple endpoints and CS, we systematically investigated the individual contribution of these factors to CS. We found that the prognosis for patients with an initially poorer prognosis improved more prominently over time, which is consistent with findings in other malignancies [2–4,26,45]. For NPC patients, particularly those with advanced disease, understanding their improvement in prognosis could reduce their fear of recurrence and improve their quality of life. Moreover, since the current NCCN guidelines version 1.2020 still recommends a uniform follow-up strategy for all NPC patients [19], the substantial heterogeneity in CS among patients with different clinical characteristics further advocates the rationale and necessity of instituting individualized follow-up schemes. Additionally, the follow-up strategies for NPC were derived from non-nasopharyngeal head and neck cancer. Given the unique biological behaviors and recurrence patterns of NPC, the follow-up strategies should be customized for NPC patients, and our results contribute substantial evidence that could be useful to that practice.

To facilitate the utilization of conditional risk in clinical practice, we established a novel, web-based model to predict the conditional risk of site-specific recurrence, including LR, RR, Bone-M, Lung-M, and Liver-M. This model features the following key characteristics. First, since the CS estimates must be interpreted clinically for individual patients with diverse clinicopathological characteristics, we simultaneously incorporated the calculation of CS and the prognostication of significant risk factors. Therefore, during ongoing surveillance, clinicians can easily update the real-time prognosis of patients with their unique clinicopathological profiles carefully weighted, and thus modify subsequent follow-up strategies. Second, in addition to the optimization of when to visit, clinicians may also be concerned about what examinations to perform but usually lack risk-based references. Hence, we introduced a prediction model that can generate site-specific recurrence risk. Accordingly, clinicians can select cost-effective diagnostic procedures to facilitate diagnostic efficiency while avoiding unnecessary expenditures for the patients. Third,

considering that a first recurrence may influence the risk and observation of subsequent events [4,8,46], we implemented the competing risk models that only investigated the first event, which was better in predicting site-specific recurrence [46,47]. Last but not most importantly, the excellent performance observed during the internal and external validation confirmed the reliability and generalizability of our model to be implemented in clinical practice for endemic NPC patients. Collectively, our easily accessible model can help physicians establish personalized, risk-based, time-adapted surveillance strategies that are updated throughout the patient's treatment and surveillance course, including when to follow up with patients and what examinations to perform. It can also facilitate patients' informed rights and help reduce patient anxiety and economic burden. Additionally, incorporating this model into clinical trial designs or cost-effective analyses could seem promising.

Several limitations need to be highlighted. First, the study's retrospective nature may have introduced potential bias. Although the large cohort and IPW adjustment might reduce bias, prospective studies are still warranted to confirm this study findings. Second, we did not include the effect of treatment in our model. Previous studies have suggested that treatment (e.g., radiotherapy alone versus chemoradiotherapy) should not be included in prediction models as a covariate unless the treatment data are derived from randomized clinical trials; otherwise, it may introduce bias (e.g., multicollinearity), as treatments are highly dependent on patient clinical characteristics in real-world clinical practice [17,48–52]. Given this information, we excluded treatment in our models. Third, the study was based on endemic NPC patients in China. The applicability of our findings to non-endemic or non-Asian populations needs further validation.

## 5 | CONCLUSIONS

This study represents the largest cohort to date investigating 10,058 patients to comprehensively characterize the CS of endemic NPC patients receiving IMRT. The impacts of well-known prognostic factors on CS were investigated, and an individualized, web-based, site-specific recurrence prediction model was established and validated. These data provide dynamic prognostic information for NPC patients. Moreover, this information could help clinicians formulate individualized, risk-based, time-adapted surveillance strategies.

## ACKNOWLEDGMENTS

The authors thank Yiducloud (Beijing) Technology Ltd. for the construction of the Big-data Intelligence Platform at

Sun Yat-sen University Cancer Center and their assistance during the establishment of the website for the web-based prediction model.

### AUTHOR CONTRIBUTIONS

Study concept and design: CFW, JW, YS. Administrative support: YS, JM. Data acquisition: CFW, JW, LL, YPM, BD, YS. Data analysis and interpretation: CFW, JW, LL, YPM, BD, WHZ, DWW, YC, JK, FPC, XLY, ZQZ, ZXL, SSX. Manuscript writing: All authors. Critical revision: All authors. Final approval of manuscript: All authors.

### ETHICS APPROVAL

The study protocol was approved by the institutional ethics committee of Sun Yat-sen University Cancer Center (B2020-119), and the requirement for informed consent was waived given the retrospective nature of the study.

### CONSENT FOR PUBLICATION

Not applicable.

### COMPETING INTERESTS

The authors declare no competing interest.

### DATA AVAILABILITY STATEMENT

The data regarding patient baseline characteristics, therapeutic information, and survival outcomes have been deposited in the Research Data Deposit public platform ([www.researchdata.org.cn](http://www.researchdata.org.cn)) with the accession code DDA2018000934. All the other data of this study are available within the article and the supplementary materials and from the corresponding author upon reasonable request.

### FUNDING

This work was supported by the National Natural Science Foundation of China (81872463 and 81930072), Special Support Program of Sun Yat-sen University (16zxtzlc06), Key-Area Research and Development Program of Guangdong Province (2019A1515012045 and 2019B020230002), Natural Science Foundation of Guangdong Province (2017A030312003), Health & Medical Collaborative Innovation Project of Guangzhou City, China (201803040003), Innovation Team Development Plan of the Ministry of Education (No. IRT\_17R110) and Overseas Expertise Introduction Project for Discipline Innovation (111 Project, B14035).

### ORCID

Yan-Ping Mao  <https://orcid.org/0000-0003-1027-071X>

Ying Sun  <https://orcid.org/0000-0002-5888-2929>

### REFERENCES

- Zabor EC, Gonen M, Chapman PB, Panageas KS. Dynamic prognostication using conditional survival estimates. *Cancer*. 2013;119(20):3589–92.
- Chang GJ, Hu C-Y, Eng C, Skibber JM, Rodriguez-Bigas MA. Practical application of a calculator for conditional survival in colon cancer. *J Clin Oncol*. 2009;27(35):5938–43.
- Liu X, Wu T, Zhu S-Y, Shi M, Su H, Wang Y, et al. Risk-dependent conditional survival and failure hazard after radiotherapy for early-stage extranodal natural killer/T-cell lymphoma. *JAMA Netw Open*. 2019;2(3):e190194.
- Ploussard G, Shariat SF, Dragomir A, Kluth LA, Xylinas E, Masson-Lecomte A, et al. Conditional survival after radical cystectomy for bladder cancer: evidence for a patient changing risk profile over time. *Eur Urol*. 2014;66(2):361–70.
- Polley M-YC, Lamborn KR, Chang SM, Butowski N, Clarke JL, Prados M. Conditional probability of survival in patients with newly diagnosed glioblastoma. *J Clin Oncol*. 2011;29(31):4175–80.
- Carbognin L, Sperduti I, Ciccarese M, Fabi A, Petrucelli L, Vari S, et al. Prognostic model for advanced breast carcinoma with luminal subtype and impact of hormonal maintenance: implications for post-progression and conditional survival. *Breast*. 2016;29:24–30.
- Skuladottir H, Olsen JH. Conditional survival of patients with the four major histologic subgroups of lung cancer in Denmark. *J Clin Oncol*. 2003;21(16):3035–40.
- Hippisley-Cox J, Coupland C. Development and validation of risk prediction equations to estimate survival in patients with colorectal cancer: cohort study. *BMJ*. 2017;357:j2497.
- Fuller CD, Wang SJ, Thomas CR, Hoffman HT, Weber RS, Rosenthal DI. Conditional survival in head and neck squamous cell carcinoma: results from the SEER dataset 1973-1998. *Cancer*. 2007;109(7):1331–43.
- Chen Y-P, Chan ATC, Le Q-T, Blanchard P, Sun Y, Ma J. Nasopharyngeal carcinoma. *Lancet*. 2019;394(19):64–80.
- Liao J-F, Zhang Q, Du X-J, Lan M, Liu S, Xia Y-F, et al. Concurrent chemoradiotherapy with weekly docetaxel versus cisplatin in the treatment of locoregionally advanced nasopharyngeal carcinoma: a propensity score-matched analysis. *Cancer Commun (Lond)*. 2019;39(1):40.
- Liu X, Tang L-L, Du X-J, Li W-F, Chen L, Zhou G-Q, et al. Changes in disease failure risk of nasopharyngeal carcinoma over time: analysis of 749 patients with long-term follow-up. *J Cancer*. 2017;8(3):455–9.
- Lv J-W, Huang X-D, Chen Y-P, Zhou G-Q, Tang L-L, Mao Y-P, et al. A national study of survival trends and conditional survival in nasopharyngeal carcinoma: analysis of the national population-based Surveillance Epidemiology and End Results registry. *Cancer Res Treat*. 2018;50(2):324–34.
- Chen F-P, Wen D-W, Li F, Lin L, Kou J, Zheng W-H, et al. The role of post-neoadjuvant chemotherapy tumor volume for prognostication and treatment guidance in loco-regionally advanced nasopharyngeal carcinoma. *Cancers*. 2019;11(11):1632.
- Lv J-W, Chen Y-P, Huang X-D, Zhou G-Q, Chen L, Li W-F, 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(18):3540–9.

16. Lv J, Chen Y, Zhou G, Qi Z, Tan KRL, Wang H, et al. Liquid biopsy tracking during sequential chemo-radiotherapy identifies distinct prognostic phenotypes in nasopharyngeal carcinoma. *Nat Commun.* 2019;10(1):3941.
17. Huang X-D, Zhou G-Q, Lv J-W, Zhou H-Q, Zhong C-W, Wu C-F, 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(2):389–95.
18. Amin MB, Edge SB, Greene FL, Byrd DR, Brookland RK, Washington MK, et al. *AJCC Cancer Staging Manual*. 8th ed. New York, NY: Springer International Publishing; 2017:1032.
19. Pfister DG, Spencer S, Adelstein D, Adkins D. *NCCN Clinical Practice Guidelines in Oncology: Head and Neck Cancers*. Version 1.2020. Accessed February 12, 2020. To view the most recent version, visit [NCCN.org](http://NCCN.org).
20. Zou X, You R, Liu H, He Y-X, Xie G-F, Xie Z-H, et al. Establishment and validation of M1 stage subdivisions for de novo metastatic nasopharyngeal carcinoma to better predict prognosis and guide treatment. *Eur J Cancer.* 2017;77:117–26.
21. Tang L-Q, Li C-F, Li J, Chen W-H, Chen Q-Y, Yuan L-X, et al. Establishment and validation of prognostic nomograms for endemic nasopharyngeal carcinoma. *J Natl Cancer Inst.* 2016;108(1):djv291.
22. Lo YM, Chan AT, Chan LY, Leung SF, Lam CW, Huang DP, et al. Molecular prognostication of nasopharyngeal carcinoma by quantitative analysis of circulating Epstein-Barr virus DNA. *Cancer Res.* 2000;60(24):6878–81.
23. Zapatka M, Borozan I, Brewer DS, Iskar M, Grundhoff A, Alawi M, et al. The landscape of viral associations in human cancers. *Nat Genet.* 2020;52(3):320–30.
24. Hillengass J, Fechtner K, Weber M-A, Bäuerle T, Ayyaz S, Heiss C, et al. Prognostic significance of focal lesions in whole-body magnetic resonance imaging in patients with asymptomatic multiple myeloma. *J Clin Oncol.* 2010;28(9):1606–10.
25. Lausen B, Schumacher M. Maximally selected rank statistics. *Biometrics.* 1992;48(1):73–85.
26. Kurta ML, Edwards RP, Moysich KB, Mcdonough K, Bertollet M, Weissfeld JL, et al. Prognosis and conditional disease-free survival among patients with ovarian cancer. *J Clin Oncol.* 2014;32(36):4102–12.
27. Halpern EF. Behind the numbers: inverse probability weighting. *Radiology.* 2014;271(3):625–8.
28. Buchanan AL, Hudgens MG, Cole SR, Lau B, Adimora AA. Worth the weight: using inverse probability weighted Cox models in AIDS research. *AIDS Res Hum Retroviruses.* 2014;30(12):1170–7.
29. Bertelsen CA, Neuenschwander AU, Jansen JE, Tenma JR, Wilhelmssen M, Kirkegaard-Klitbo A, et al. 5-year outcome after complete mesocolic excision for right-sided colon cancer: a population-based cohort study. *Lancet Oncol.* 2019;20(11):1556–65.
30. São Julião GP, Karagkounis G, Fernandez LM, Habr-Gama A, Vailati BB, Dattani M, et al. Conditional survival in patients with rectal cancer and complete clinical response managed by watch and wait after chemoradiation: recurrence risk over time. *Ann Surg.* 2020;272(1):138–44.
31. Margonis GA, Buettner S, Andreatos N, Wagner D, Sasaki K, Barbon C, et al. Prognostic factors change over time after hepatectomy for colorectal liver metastases: a multi-institutional, international analysis of 1099 patients. *Ann Surg.* 2019;269(6):1129–37.
32. Jamil ML, Keeley J, Sood A, Dalela D, Arora S, Peabody JO, et al. Long-term risk of recurrence in surgically treated renal cell carcinoma: a post hoc analysis of the Eastern Cooperative Oncology Group-American College of Radiology Imaging Network E2805 Trial Cohort. *Eur Urol.* 2020;77(2):277–81.
33. Bischof DA, Kim Y, Dodson R, Jimenez MC, Behman R, Cocieru A, et al. Conditional disease-free survival after surgical resection of gastrointestinal stromal tumors: a multi-institutional analysis of 502 patients. *JAMA Surg.* 2015;150(4):299–306.
34. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med.* 2015;34(28):3661–79.
35. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika.* 1982;69(1):239–41.
36. Rubin DB *Multiple Imputation for Nonresponse in Surveys*. New York, NY: John Wiley & Sons; 2004:258.
37. Li P, Stuart EA, Allison DB. Multiple imputation: a flexible tool for handling missing data. *JAMA.* 2015;314(18):1966–7.
38. Van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *J Stat Softw.* 2011;45(3):1–67.
39. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med.* 2011;30(4):377–99.
40. Carlin JB, Li N, Greenwood P, Coffey C. Tools for analyzing multiple imputed datasets. *Stat Med.* 2003;3(3):226–44.
41. Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prev Sci.* 2007;8(3):206–13.
42. Moller MB, Pedersen NT, Christensen BE. Conditional survival of patients with diffuse large B-cell lymphoma. *Cancer.* 2006;106(10):2165–70.
43. He Q, Huang Y, Wan G, Feng M, Zeng H, Liu M, et al. A novel prognostic marker based on risk stratification with prognostic nutritional index and age for nasopharyngeal carcinoma patients who received neoadjuvant chemotherapy. *Biomark Med.* 2019;13(12):1013–23.
44. Huang S-J, Tang Y-Y, Liu H-M, Tan G-X, Wang X, Zhang H, et al. Impact of age on survival of locoregional nasopharyngeal carcinoma: an analysis of the Surveillance, Epidemiology, and End Results program database, 2004–2013. *Clin Otolaryngol.* 2018;43(5):1209–18.
45. Margonis GA, Buettner S, Andreatos N, Wagner D, Sasaki K, Barbon C, Beer A, et al. Prognostic factors change over time after hepatectomy for colorectal liver metastases: a multi-institutional, international analysis of 1099 patients. *Ann Surg.* 2019;269(6):1129–37.
46. Zhang Y, Zheng D, Xie J, Li Y, Wang Y, Li C, et al. Development and validation of web-based nomograms to precisely predict conditional risk of site-specific recurrence for patients with completely resected non-small cell lung cancer: a multiinstitutional study. *Chest.* 2018;154(3):501–11.
47. Rasmussen JH, Grønhoj C, Håkansson K, Friborg J, Andersen E, Lelkaitis G, et al. Risk profiling based on p16 and HPV DNA more accurately predicts location of disease relapse in

- patients with oropharyngeal squamous cell carcinoma. *Ann Oncol.* 2019;30(4):629–36.
48. Geuskens GA, Burdorf A, Barendregt PJ, Hazes JMW. A high physical workload and high job demands hamper the good prognosis in physical functioning in persons with early inflammatory joint conditions. *Rheumatology.* 2011;50(4):789–98.
  49. Labidi SI, Ménétrier-Caux C, Chabaud S, Chassagne C, Sebban C, Gargi T, et al. Serum cytokines in follicular lymphoma. Correlation of TGF-beta and VEGF with survival. *Ann Hematol.* 2010;89(1):25–33.
  50. Elwood JM, Tawfiq E, Tintin S, Marshall RJ, Phung TM, Campbell I, et al. Development and validation of a new predictive model for breast cancer survival in New Zealand and comparison to the Nottingham prognostic index. *BMC Cancer.* 2018;18(1):897.
  51. Balachandran VP, Gonen M, Smith JJ, Dematteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol.* 2015;16(4):e173–e180.
  52. Kane GC, Maradit-Kremers H, Slusser JP, Scott CG, Frantz RP, Mcgoon MD. Integration of clinical and hemodynamic param-

eters in the prediction of long-term survival in patients with pulmonary arterial hypertension. *Chest.* 2011;139(6):1285–93.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Wu C-F, Lv J-W, Lin L, et al. Development and validation of a web-based calculator to predict individualized conditional risk of site-specific recurrence in nasopharyngeal carcinoma: Analysis of 10,058 endemic cases. *Cancer Commun.* 2020;1-14.  
<https://doi.org/10.1002/cac2.12113>