ORIGINAL RESEARCH

RuleFit-Based Nomogram Using Inflammatory Indicators for Predicting Survival in Nasopharyngeal Carcinoma, a Bi-Center Study

Chao Luo[1,*], Shuqi Li[1,*], Qin Zhao[1,*], Qiaowen Ou[2], Wenjie Huang[1], Guangying Ruan[1], Shaobo Liang[3], Lizhi Liu[1,4], Yu Zhang[5], Haojiang Li [ID][1]

[1]Department of Radiology, 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, People's Republic of China; [2]Department of Clinical Nutrition, The First Affiliated Hospital of Guangdong Pharmaceutical University, Guangzhou, Guangdong, People's Republic of China; [3]Department of Radiotherapy, Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, Guangdong, People's Republic of China; [4]Department of Radiology, The Third People's Hospital of Shenzhen, Shenzhen, Guangdong, People's Republic of China; [5]Department of Pathology, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong, People's Republic of China

*These authors contributed equally to this work

Correspondence: Haojiang Li, Department of Radiology, 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, 651 Dongfeng Road East, Guangzhou, Guangdong, 510060, People's Republic of China, Tel +86-20-87342135, Fax +86-20-87342125, Email lihaoj@sysucc.org.cn; Yu Zhang, Department of Pathology, Sun Yat-sen University Cancer Center, 651 Dongfeng Road East, Guangzhou, Guangdong, 510060, People's Republic of China, Email zhangyu@sysucc.org.cn

Purpose: Traditional prognostic studies utilized different cut-off values, without evaluating potential information contained in inflammation-related hematological indicators. Using the interpretable machine-learning algorithm RuleFit, this study aimed to explore valuable inflammatory rules reflecting prognosis in nasopharyngeal carcinoma (NPC) patients.

Patients and Methods: In total, 1706 biopsy-proven NPC patients treated in two independent hospitals (1320 and 386) between January 2010 and March 2014 were included. RuleFit was used to develop risk-predictive rules using hematological indicators with no distributive difference between the two centers. Time-event-dependent hematological rules were further selected by stepwise multivariate Cox analysis. Combining high-efficiency hematological rules and clinical predictors, a final model was established. Models based on other algorithms (AutoML, Lasso) and clinical predictors were built for comparison, as well as a reported nomogram. Area under the receiver operating characteristic curve (AUROC) and concordance index (C-index) were used to verify the predictive precision of different models. A site-based app was established for convenience.

Results: RuleFit identified 22 combined baseline hematological rules, achieving AUROCs of 0.69 and 0.64 in the training and validation cohorts, respectively. By contrast, the AUROCs of the optimal contrast model based on AutoML were 1.00 and 0.58. For overall survival, the final model had a much higher C-index than the base model using TN staging in two cohorts (0.769 vs 0.717, *P*<0.001; 0.752 vs 0.688, *P*<0.001), and showing great generalizability in training and validation cohorts. The two models based on RuleFit rules performed best, compared with other models. As for other endpoints, the final model showed a similar trend. Kaplan–Meier curve exhibited 22.9% (390/1706) patients were "misclassified" by AJCC staging, but the final model could assess risk classification accurately.

Conclusion: The proposed final models based on inflammation-related rules based on RuleFit showed significantly elevated predictive performance.

Keywords: machine learning, nomograms, nasopharyngeal carcinoma, prognosis, survival analysis

Plain Language Summary

Nasopharyngeal carcinoma (NPC) is a malignant head and neck cancer with highest mortality rate in China. The American Joint Cancer Committee (AJCC) TNM staging system is the most commonly used standard for treatment decision-making and prediction of

prognosis. However, with the heterogeneities of tumors and individual physical conditions, TNM staging remains far from sufficient. Characterized by potentially wide application, interlaboratory stability, and cost efficiency in clinical settings, inflammatory hematological parameters may gain widespread clinical use as prognostic markers. Traditional prognostic studies utilized different cut-off values, ignoring potential information contained in inflammation-related hematological indicators. Instead of common cut-off values alone, this study dug out inflammatory rules of significant prognostic value from original hematological data, using an interpretable algorithm RuleFit. Different from other algorithms, RuleFit provides users with a relatively small rule ensemble that can be easily interpreted and applied by creating a large initial ensemble of prediction rules and further selecting rules to improve the predictive accuracy. Finally, we successfully established a cost-efficient prognostic nomogram in nasopharyngeal carcinoma patients, dividing patients into two groups with a significant survival difference. For convenience of clinicians and patients, a site-based app was established (https://bloodscore4npc.shinyapps.io/shinyblood/) and the model based on R language with detail instruction for replicated blood scores calculation was uploaded here (https://github.com/trackse/bloodscore). In the long run, this study provides us with an effective model of prognostic risk modeling, which can be applied to more prognostic studies.

Introduction

In malignant head and neck cancer, nasopharyngeal carcinoma (NPC) is the one with highest mortality rate in China, and approximately 130,000 new cases were reported worldwide in 2018.[1,2] The geographical distribution of NPC is extremely unbalanced: more than 70% newly diagnosed cases are in east and southeast Asia, with an age-standardized rate of 3.0 per 100,000 in China.[1,2] The Union Internationale Contre le Cancer/American Joint Cancer Committee (UICC/AJCC) TNM staging system is the most commonly used standard for treatment decision-making and prediction of prognosis. However, with the heterogeneities of tumors and individual physical conditions, TNM staging remains far from sufficient.[2,3] Clinical factors, molecular and imaging biomarkers have been widely studied to improve predictive performance.[4–8]

Emerging evidence shows that systemic immunity plays a crucial role in tumor elimination, regression, and metastasis.[9–11] Compared with other reported markers, hematological indicators are characterized by potentially wide application, interlaboratory stability, and cost efficiency in clinical settings. Hence, hematological parameters may gain widespread clinical use as prognostic markers.[12,13] Nonetheless, limitations of currently available studies are as follows: 1) traditional epidemiological researchers focused on a specified cutoff value for categorizing patients with different hematological immune-related levels into high-risk and low-risk groups,[13,14] potentially ignoring the survival outcomes of patients with extreme cell levels with this dichotomous-level evaluation system[15] and the information contained a large number of hematological indicators; 2) the majority of studies are single-center without an independent external validation cohort,[6,8] whereby the results are non-comparable; 3) current combined indices are calculated to acquire a significant effect instead of revealing actual relationships, but their value is restricted because of the concern with certain indices and inability to reflect the intrinsic relationship among immune-related variables; and 4) some results of these studies were frequently contradictory.[7] In other words, few results from earlier research can be applied to predict individual outcomes or guide clinical treatments. On the basis of a great deal of original hematological data, this study aims to explore the prognostic value of hematological indicators in NPC patients.

Clinicians are familiar with the common machine-learning (ML) algorithms (linear and tree-based) that are easy to interpret because those algorithms only consider the correlation between independent and dependent variables. However, those ML algorithms are unsuitable for processing complex information, as they prevent the extraction of useful information from multiple correlated or unrelated variables in large datasets. Complex models represented by deep neural networks can be inferred at multiple levels to facilitate the analysis of complex relationships among variables and to achieve a high predictive performance. However, such complexity makes the model a black box that lacks clinical interpretability, which poses a great challenge for the transformation of information from scientific research to clinical application. RuleFit,[16] a flexible algorithm, combines tree ensembles and linear models to benefit from the accuracy of the former and the interpretability of the latter. This provides users with a relatively small rule ensemble that can be easily interpreted and applied by creating a large initial ensemble of prediction rules and further selecting rules to improve the predictive accuracy.[17] Clinicians can further optimize the rules of time-event-related prognosis using traditional statistical methods, such as Cox regression analyses. At present, RuleFit has been widely used in commercial fields and studied in medical research for higher sensitivity and specificity, such as in psychological practice,[18] coronary

artery calcification prediction,[19] and prediction of the risk of type 1 diabetes.[20] It appears to be a promising method for facilitating clinical interpretability and creating decision-making tools that are less time-consuming, with accuracy that is comparable to traditional actuarial methods.[17] To date, however, this algorithm has not been applied to the prognostication of cancer patients.

Using the RuleFit algorithm, the primary objectives of this study were to explore the hematological rules of prognostication, and establish a cost-effective model to refine individualized risk prediction in NPC patients.

Materials and Methods

Patients and Treatment

This large-scale bi-center, retrospective study enrolled 1706 Chinese patients with biopsy-proven NPC, including 1320 patients who were treated at Sun Yat-sen University Cancer Center (SYUCC; training cohort) and 386 patients treated at First People's Hospital of Foshan (validation cohort) between January 2010 and March 2014. The inclusion criteria were as follows: 1) newly confirmed as nasopharyngeal carcinoma by histological analysis, 2) availability of complete clinical data and information on hematological investigations, 3) performing MR examinations of the nasopharynx and cervical regions before treatment, and 4) treated with radiation therapy. The exclusion criteria were 1) with second primary tumor, 2) with distant metastasis at first diagnosis, and 3) failure to receive complete radiation therapy plan for NPC. The ethics committee of SYUCC approved this study (approval number: B2019-222), and the need for informed consent was waived due to the retrospective nature of study. To cover the privacy of all participants, this study maintained their information confidential and anonymous, and was conducted in accordance with the Declaration of Helsinki.

Patients were followed up every 3 months for the first 2 years and every 6 months thereafter for subsequent 5 years. The primary endpoint was overall survival (OS), which was calculated as the interval (months) from the day of first treatment to death from any cause or the last follow-up visit. Distant metastasis-free survival (DMFS), locoregional relapse-free survival (LRFS), and progression-free survival (PFS) were set as the secondary endpoints. Detailed treatments and data acquisition are described in the Supplementary Materials.

Data Collection

According to the 8th AJCC TNM staging system, we retrospectively restage enrolled patients' clinical TNM stage based on MR imaging. All restaging work were evaluated by senior radiologists with at least five years of experience diagnosing head and neck tumors. Pretreatment original data of hematological indicators were collected, including albumin, absolute lymphocyte count (ALC), absolute neutrophil count, C-reactive protein (CRP), hemoglobin, lactate dehydrogenase, monocyte (MON), platelet count (PLT), prognostic nutritional index (PNI), white blood cell (WBC). Besides, clinical information including age, sex, histological type, EBV DNA copies and treatment strategy were recorded as well. The cut-off value of EBV DNA copies was 1000, and 10,000 copies/mL, based on a previous study.[6]

Model Building

The main workflow of the prognostic prediction system is illustrated in Figure 1. The chi-square test and Mann–Whitney *U*-test were used to identify baseline hematological indices without distribution differences, which represent relative stability and high repeatability in different medical centers. In the training cohort, potential hematological rules identified by RuleFit were established as a prognostic model named RuleFit scores, with the parameter of "max_rule_length" set as 3 to simplify the interpretation of the rules for clinicians. Similarly, we used a wrapper function, AutoML, which is able to automatically select the optimal algorithm and establish a contrast model, by integrating multiple ML algorithms including not only deep learning, Light GBM, and Stack Ensembles but also common models, such as Lasso and Ridge Regression, Random Forest, and Naivebayes. AutoML scores were established in the default setting (5-fold cross-validation, 3600 seconds for training, and 100 training models). The Lasso regression model was included in AutoML, but the R glmnet package is more commonly used to train the Lasso model. Thus, a prognostic model based on Lasso was established separately (Lasso scores).

Based on rules discovered by RuleFit, stepwise multivariable analyses with the Cox proportional hazards model were used to identify time-event-dependent prognostic variables adjusted with clinical prognostic factors. Then blood scores based on
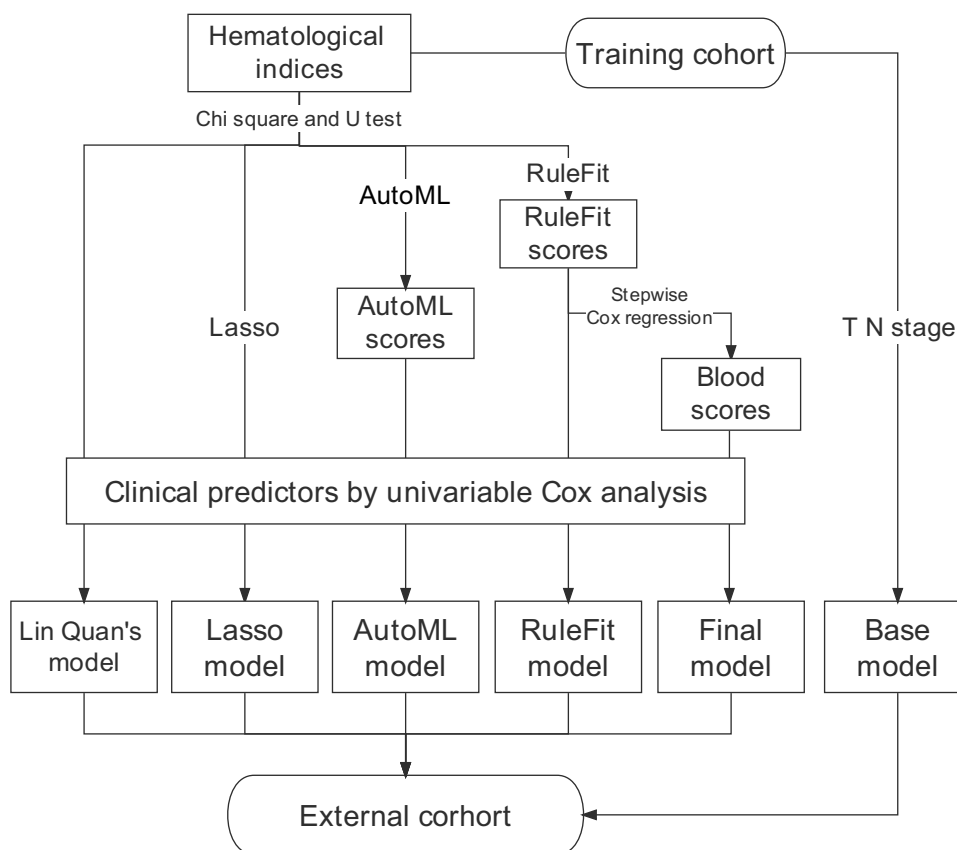
Figure 1 Flowchart of the study procedures.

those prognostic variables were constructed. Finally, the base TN stage model, and combined models incorporating the above-described scores models with clinical prognostic factors were established. Furthermore, Lin-Quan's model was established by integrating the hematological indices reported by a previous study[6] and the clinical risk predictors in our study.

Statistical Analysis

The discrimination performance of the scores models was quantified by the area under the receiver operating characteristic (AUROC) value in the primary training set and validated in the independent validation set. We evaluated the performance of the established models for predicting survival using Harrell's concordance index (C-index) and calibration curves. Decision curve analysis (DCA) was conducted to determine the clinical usefulness of the models by calculating the net benefits at different threshold probability values. Kaplan–Meier survival analysis and Log rank tests were performed to evaluate final models and AJCC stage. In order to serve clinicians and patients better, an online application based on the final risk stratification model concerning patient-specific survival prediction was constructed by shiny app (https://npc2science.shinyapps.io/shiny/).

All statistical analyses were performed using R software (version 3.2.5, https://www.r-project.org/), including the packages stats, survival, Hmisc, RuleFit, AutoML, glmnet, ggplot2, survminer, etc. RuleFit and AutoML are both included in the H_20 package (version 3.32.0.2, source code will be provided if necessary). $P<0.05$ was considered statistical significant.

Results
Patient Characteristics

The characteristics of the patients are presented in Table 1. In this study cohort of 1706 patients, the median age was 46 years (interquartile range [IQR]: 38–54 years), and the median follow-up was 61.8 months (range: 3.3–99.1 months). There were no statistically significant differences in sex, T stage, N stage, and clinical stages, indicating a similar distribution of patients between the two cohorts. During the follow-up period, 12.1% (161/1320) and 19.1% (74/386) of

Table 1 Clinical Characteristics of Patients in the Primary and Validation Cohorts

Characteristics	Total (N=1706)	Primary (N=1320)	Validation (N=386)	P-value
Age (years)				<0.001*
Median (IQR)	46 (38–54)	45 (38–54)	47 (40–57)	
Sex				0.385
Male	1261 (73.9%)	969 (73.4%)	292 (75.6%)	
Female	445 (26.1%)	351 (26.6%)	94 (24.4%)	
Histologic type				<0.001*
WHO type 1/2	63 (3.8%)	63 (4.8%)	0 (0.0%)	
WHO type 3	1643 (96.2%)	1257 (95.2%)	386 (100.0%)	
EBV DNA, 10^3 copies/mL				<0.001*
<1	806 (47.2%)	583 (44.2%)	223 (57.8%)	
<10	493 (28.9%)	339 (25.7%)	154 (39.9%)	
≥10	407 (23.9%)	398 (30.2%)	9 (2.3%)	
T classification				0.204
T1	434 (25.4%)	331 (25.1%)	103 (26.7%)	
T2	206 (12.1%)	153 (11.6%)	53 (13.7%)	
T3	665 (39.0%)	532 (40.3%)	133 (34.5%)	
T4	401 (23.5%)	304 (23.0%)	97 (25.1%)	
N classification				0.064
N0	341 (20.0%)	278 (21.1%)	63 (16.3%)	
N1	969 (56.8%)	752 (57.0%)	217 (56.2%)	
N2	281 (16.5%)	205 (15.5%)	76 (19.7%)	
N3	115 (6.7%)	85 (6.4%)	30 (7.8%)	
Clinical stage				0.297
I	144 (8.4%)	115 (8.7%)	29 (7.5%)	
II	377 (22.1%)	288 (21.8%)	89 (23.1%)	
III	691 (40.5%)	547 (41.4%)	144 (37.3%)	
IV	494 (29.0%)	370 (28.0%)	124 (32.1%)	
Treatment				0.020*
RT alone	211 (12.4%)	155 (11.7%)	56 (14.5%)	
CCRT	622 (36.5%)	504 (38.2%)	118 (30.6%)	
IC + CCRT	872 (51.1%)	660 (50.0%)	212 (54.9%)	

Notes: Data are presented as numbers with percentages or medians with interquartile ranges. Histologic type was based on basing on 2005 World Health Organization classification of tumors. T, N and clinical stage was based on the 8th American Joint Cancer Committee TNM staging system. *P-value <0.05, P values were calculated by Fisher's exact test or chi-square test.
Abbreviations: IQR, interquartile range; WHO, World Health Organization; EBV, Epstein-Barr virus; RT, radiotherapy; CCRT, concurrent chemoradiotherapy; IC, induction chemotherapy.

patients died in the training and validation cohorts, respectively, within 5 years. The 5-year OS, DMFS, LRFS, and PFS were 85.5%, 86.3%, 89.9%, and 76.8%, respectively.

Hematological Models Based on RuleFit and AutoML

Nineteen hematological indices exhibited no statistically significant difference between the two cohorts, representing relative stability in different laboratories, and the associated details are listed in Table S1. These stable hematological parameters were used to establish scores models for prognosis prediction in the training cohort by different algorithms. Ensemble ML used multiple learning algorithms to obtain better predictive performance than can be obtained from any of the constituent learning algorithms alone. Stack Ensembles in AutoML achieved the top three AUROC values but may overfit without adjusting parameters. Through 5-fold cross-validation, AutoML automatically selected a random forest (DRF)-based model that showed the best performance among many models as output. Finally, RuleFit used approximately 20 seconds to study and build the RuleFit scores model that includes 23 rules, achieving an AUROC value of 0.695 and 0.645 in the training and validation cohorts, respectively. In contrast, after 1 hour of training, AutoML scores had a high AUROC value of 1.000 in the training set but only 0.580 in the validation set, suggesting overfitting.

Prognostic Models Combining Clinical Predictors

According to the univariate Cox analysis, TN stage was significantly related to all endpoints; age was related to OS and PFS; and Epstein-Barr virus DNA (EBV DNA) was significantly related to OS, DMFS, and PFS. Chemotherapy was related to DMFS and PFS (Table S2), but treatment options were largely determined by clinical stage. Based on the 23 rules identified by the RuleFit algorithm, multivariable Cox regression with stepwise analysis was performed to explore independent prognostic indices adjusted by corresponding clinical prognosticator. Four rules were selected and used to build the blood scores model: 1) 2.20×10^9/L $\leq$ ALC $< 3.02 \times 10^9$/L and MON count $> 0.62 \times 10^9$/L; 2) CRP < 6.69mg/L, MON $< 0.75 \times 10^9$/L, and platelet count $< 366.86 \times 10^9$/L; 3) white blood cell count $< 8.26 \times 10^9$/L, MON $< 0.79 \times 10^9$/L, and neutrophil-to-lymphocyte ratio < 1.27; and 4) MON $< 0.62 \times 10^9$/L, absolute lymphocyte count $< 3.44 \times 10^9$/L, and white blood cell count $\geq 5.25 \times 10^9$/L.

In the final analysis, the blood scores model was an independent predictor of prognosis per multivariable analysis (Table S3). The C-index of different models are listed in Table 2. In the validation set, the final model achieved a much higher C-index than the base model and models based on other algorithms. The AutoML model had the lowest C-index at 0.605. There was no statistical difference in the performance between the final model and RuleFit model. Interestingly, better performance of the final model was observed in other endpoints, except for PFS. Compared with Lin-Quan's model, the base model had a similar C-index without a statistically significant difference, and the RuleFit model showed better predictive value in terms of OS, LRFS, and PFS. Similarly, the blood scores model showed better performance than the Lasso model with regard to OS (0.752 vs 0.686, $P<0.001$), DMFS (0.725 vs 0.689, $P<0.001$), and PFS (0.701 vs 0.662, $P<0.001$). A nomogram was further built based on the combined prognostic model integrating clinical predictors with blood scores (Figure 2). We used zero as linear predictor cutoff value to divide NPC patients into two groups with different risk ratio, and further validated with K-M survival curves (Figure S1). Correspondingly, the cutoff value of total points in nomogram equals to 170.5 (Figure 1). In Figure S2, the calibration plot for the probability of 5-year OS showed optimal agreement between the prediction by the final model and the actual OS in both the primary and training cohorts. Net predictive benefit of the final nomogram was further validated using the DCA diagram in the validated set (Figure 3). As shown in the graph, the decision curve of final model is far above extreme lines and higher than the base and RuleFit models only. Moreover, NPC patients have a good prognosis, with a mortality rate of less than 30% even in advanced T4 patients, suggesting the final model has great discrimination between high- and low-risk stratification in patients with NPC. In Figure 4, patients were classified into two risk stratifications based on AJCC (early stage: I, II vs advanced stage: III, IV) and final model, respectively. A and D groups represent patients with same risk stratifications using two models. For patients in group B and C, they showed opposite risk in the two models. Patients in group B (low risk in final model but high in AJCC) showed no survival difference with group A, and group C (high risk in our model but low in AJCC) have similar survival with D group (both $P>0.05$). For B and C group patients, survival difference is of statistical significance ($P<0.01$). For convenience of clinicians and patients, we transferred the final model into a network-based

Table 2 C-Index of Different Modules in the Primary and Validation Cohorts

Models	Primary Cohort		Validation Cohort	
	C-Index	**P-value**	**C-Index**	**P-value**
OS:				
Final model	0.769 (0.735~0.802)	Reference	0.752 (0.697~0.808)	Reference
RuleFit model	0.776 (0.742~0.810)	<0.001	0.734 (0.679~0.790)	0.124
AutoML model	0.963 (0.956~0.970)	<0.001	0.605 (0.542~0.668)	<0.001
Lin Quan's model	0.746 (0.711~0.782)	<0.001	0.702 (0.642~0.762)	<0.001
Lasso model	0.754 (0.719~0.789)	0.041	0.686 (0.623~0.750)	<0.001
Clinical model	0.747 (0.711~0.782)	<0.001	0.715 (0.655~0.775)	<0.001
Base model	0.717 (0.682~0.753)	<0.001	0.688 (0.624~0.752)	<0.001
DMFS:				
Final model	0.734 (0.698~0.770)	Reference	0.725 (0.666~0.784)	Reference
RuleFit model	0.741 (0.705~0.777)	0.015	0.703 (0.641~0.765)	0.169
AutoML model	0.827 (0.795~0.860)	<0.001	0.619 (0.549~0.69)	<0.001
Lin Quan's model	0.715 (0.679~0.751)	0.021	0.697 (0.633~0.762)	0.145
Lasso model	0.716 (0.679~0.752)	0.002	0.689 (0.623~0.754)	0.010
Clinical model	0.713 (0.676~0.749)	0.001	0.695 (0.631~0.759)	0.008
Base model	0.687 (0.649~0.726)	<0.001	0.704 (0.639~0.770)	0.006
LRFS:				
Final model	0.649 (0.599~0.700)	Reference	0.633 (0.543~0.724)	Reference
RuleFit model	0.647 (0.595~0.698)	0.178	0.622 (0.535~0.709)	0.865
AutoML model	0.733 (0.683~0.782)	<0.001	0.566 (0.475~0.658)	0.008
Lin Quan's model	0.647 (0.598~0.696)	0.696	0.587 (0.495~0.678)	0.031
Lasso model	0.649 (0.599~0.698)	0.982	0.614 (0.529~0.700)	0.128
Clinical model	0.645 (0.596~0.694)	0.224	0.596 (0.507~0.686)	0.017
Base model	0.645 (0.596~0.694)	0.224	0.596 (0.507~0.686)	0.017
PFS:				
Final model	0.691 (0.660~0.722)	Reference	0.701 (0.647~0.754)	Reference
RuleFit model	0.692 (0.661~0.723)	0.005	0.683 (0.629~0.737)	0.154
AutoML model	0.798 (0.770~0.826)	<0.001	0.587 (0.53~0.644)	<0.001
Lin Quan's model	0.678 (0.647~0.709)	0.005	0.677 (0.622~0.733)	0.003
Lasso model	0.679 (0.648~0.710)	0.002	0.662 (0.605~0.719)	0.001
Clinical model	0.677 (0.646~0.708)	<0.001	0.674 (0.618~0.729)	0.001
Base model	0.665 (0.634~0.695)	<0.001	0.663 (0.606~0.720)	<0.001

Note: Clinical models were contributed by clinical predictors determined by univariate analysis (Table S2).
Abbreviations: OS, overall survival; DMFS, distant metastasis-free survival; LRFS, locoregional relapse-free survival; PFS, progression-free survival.
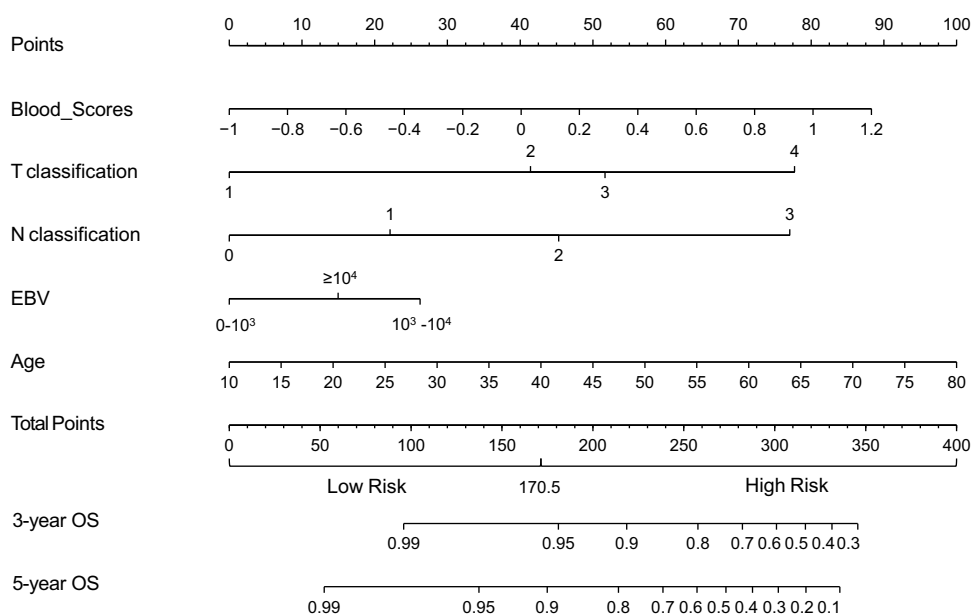
Figure 2 Nomogram of final model.
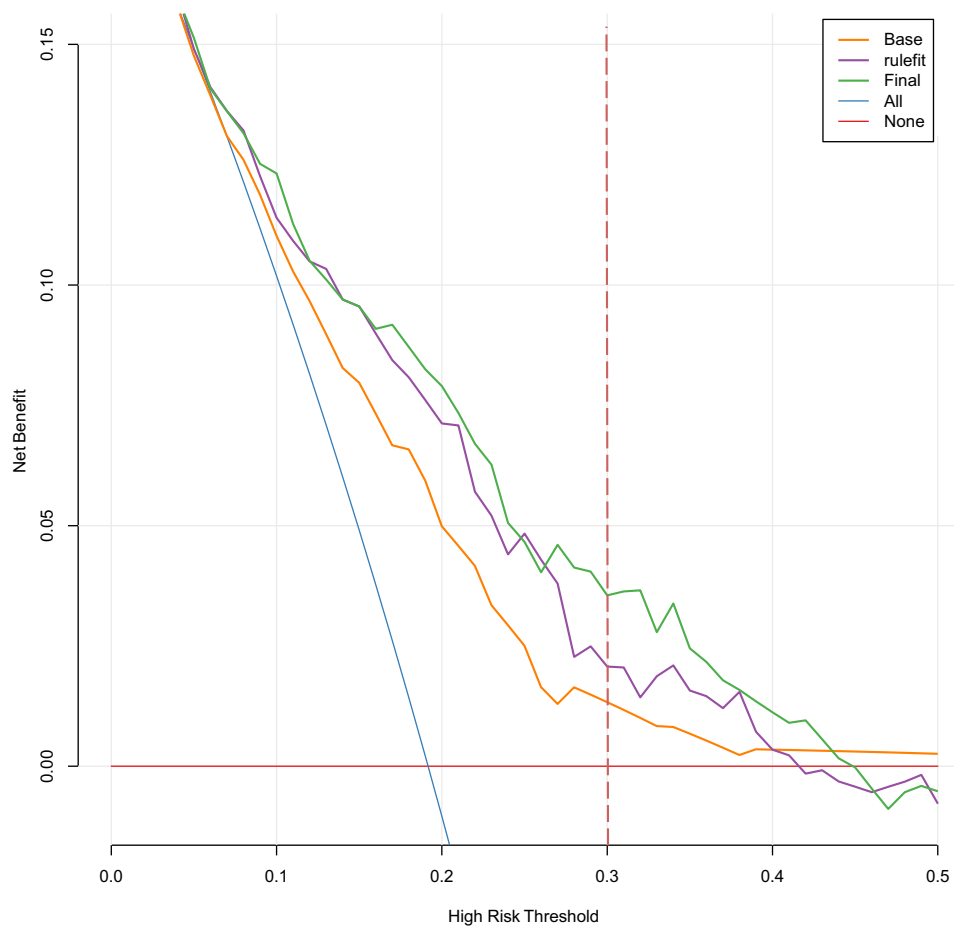**Abbreviations**: OS, overall survival; EBV, Epstein-Barr virus.

Figure 3 Decision curve analysis of the 5-year overall survival predictions.
Notes: Horizontal and axis represent the mortality risk and net benefit, respectively. The curves of models are above on lines of extreme assumptions, indicating the net benefit.
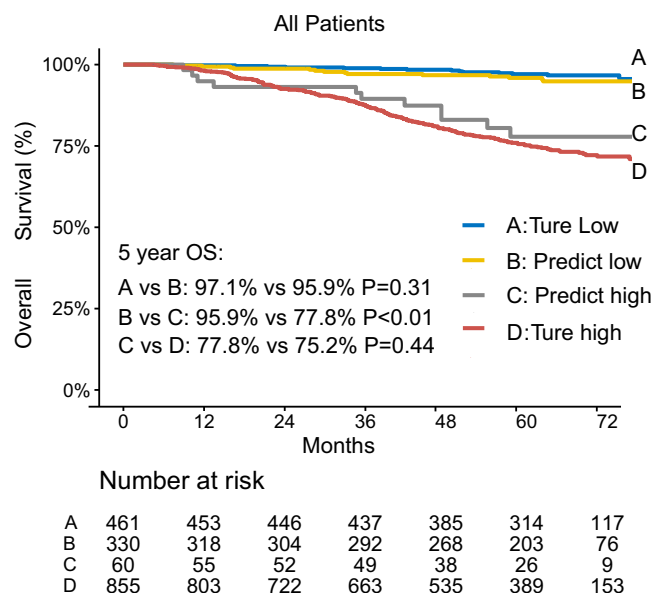
Figure 4 Risk stratification in final model and AJCC stage.
Notes: AJCC stage I, II were thought low-risk and III, IV stage were high-risk. (**A** and **D**) groups represent patients with same high- or low-risk stratification in AJCC stage and final model. (**B**) group was predicted low risk by final model but high risk by AJCC, and C group was predicted high risk by final model but low risk by AJCC.
Abbreviations: OS, overall survival; AJCC, American Joint Cancer Committee.

app (https://bloodscore4npc.shinyapps.io/shinyblood/). By choosing or entering variables in the app, a corresponding survival curve will be generated, as well as reference curves based on AJCC stage (Figure 5).

Discussion

At present, the research on hematological indicators often uses different cut-off values, and few researchers mined the large amount of information behind raw hematological data. ML is widely used to deal with large amounts of information. This study explored and built inflammatory rules based on RuleFit, showing great prognostic value in patients with NPC. As compared with the AutoML algorithm, RuleFit shows higher robustness: the AUROC values of the model directly built by RuleFit achieved 0.695 and 0.645 in the training and validation cohorts, respectively, whereas AutoML scores were 1.000 and 0.580. The final model combining blood scores and clinical predictors showed first-class performance in terms of the C-index for OS, DMFS, LRFS, and PFS, compared with the base model and other contrast models. To the best of our knowledge, this is the first study to apply the RuleFit algorithm to establish a tumor prognosis model.

To date, there are quite a few predictive models incorporating hematological biomarkers and clinical characteristics.[6,8,21] In this study, we found hematological predictors from large-scale hematological data of NPC patients in two independent cohorts. Instead of traditional cut-off values, primary continuous hematological variables were studied, allowing evaluation of the information contained in a large number of hematological indicators. According to the final 4 hematological rules, there were no indices related to a common cut-off value, although we included some hematological categorical variables base on cut-off values of clinically normal and statistically significant receiver operating characteristic curve into original study, which indicate the irrationality of traditional cut-off values in prognostic prediction to some extent.

Compared with the base model alone, the final model achieved much higher C-index, confirming that hematological indices do elevate the prognostic predictive value in NPC patients. In addition, the final model has stable predictive value in both training and validation cohorts, and performs well in all endpoints. As compared with Lin-Quan's[6] model based on traditional hematological biomarkers, the final model also showed better performance. Differing from our and Li et al[8] studies, sex was shown to be an independent predictor in Lin-Quan's study,[6] which may be attributed to selection bias. In addition, body mass index was not included in this study. Owing to different cohorts and endpoints, C-indices in the two studies are not comparable, but the final model achieved statistically higher C-index than Lin-Quan's model. A previous
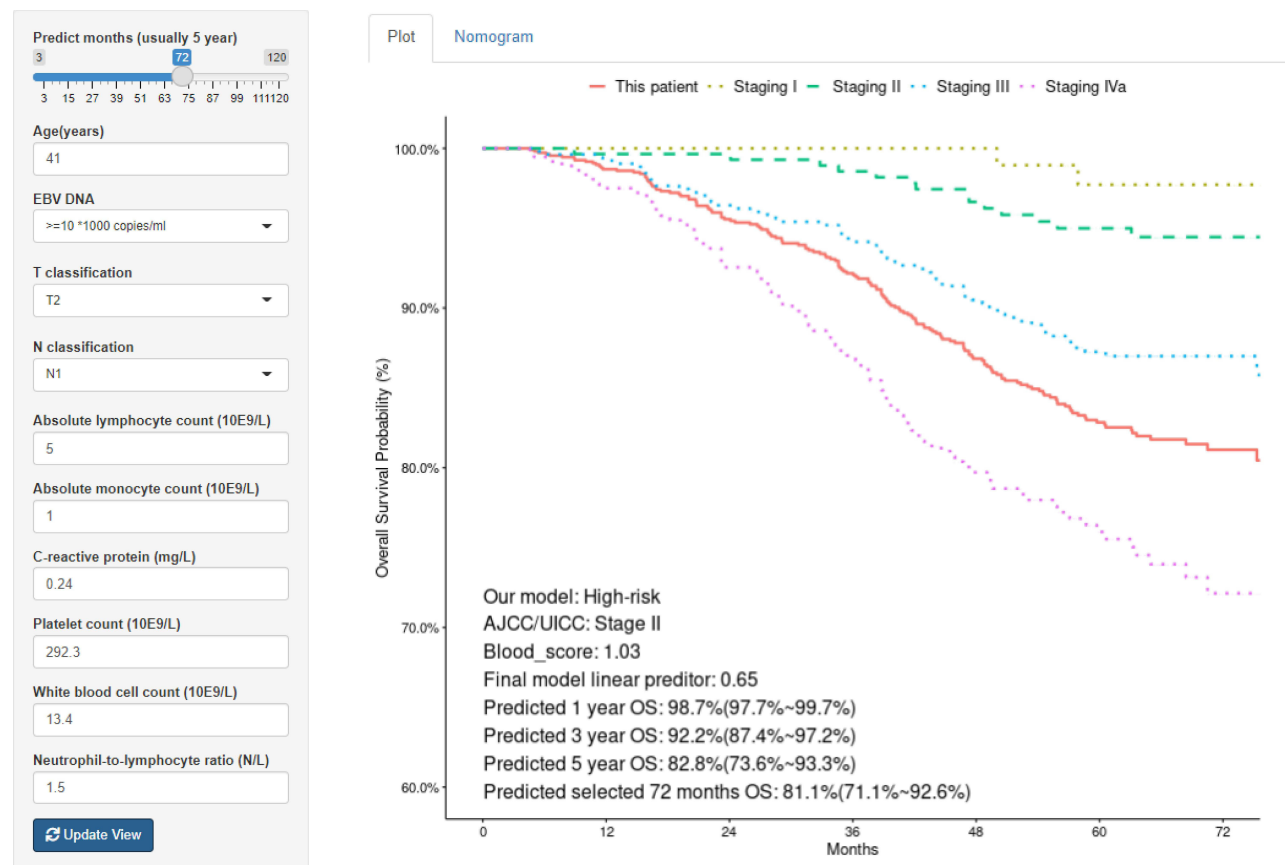
Prognosis Predictor for Nasopharyngeal Carcinoma

Figure 5 A case of NPC patients using the network-based predictor for prognosis predicting.
Notes: Overall survival was set as the endpoint in app. Detailed instruction for replicated blood scores calculation was uploaded here to facilitate clinical and patient usage (*https://github.com/trackse/bloodscore*).
Abbreviations: OS, overall survival; EBV, Epstein-Barr virus.

nomogram[22] incorporated lactate dehydrogenase-to albumin ratio (LAR) and platelet-to-lymphocyte ratio (PLR) for predicting survival in NPC, improving the predictive performance of TNM stage. In this study, OS and PFS were set as end points, and we share a similar C-index for TN stage (validation cohort: OS 0.688 vs 0.688; PFS 0.666 vs 0.663). This nomogram elevated C-index from 0.688 (TNM stage) to 0.747 in the validation cohort (8.6%). In addition, their study only classified EBV DNA as detectable and indetectable, causing loss of effective information. It is fair to say our model increased predictive performance in NPC using cost-effective hematological indicators.

As the final four inflammation-related hematological rules show, a total of 6 hematological indicators were used. The absolute MON and lymphocyte count appear most frequently, and the absolute lymphocyte count was included in each rule, implying large contributions to related model. This confirmed the significance of monocytes and lymphocytes in the prognostic prediction of NPC patients, which is consistent with findings from previous studies.[23,24] The exact relationship between hematological parameters and malignancy prognosis has not been fully reported yet, but this association may be illustrated by immune cells and inflammatory proteins: lymphocytes play a crucial part in immunologic antitumor responses by inhibiting tumor cell proliferation and inducing cell death. A high pretreatment lymphocyte count has been reported to be associated with a good prognosis of NPC,[25] and tumor-infiltrating lymphocytes are positive prognostic predictors in various cancers.[26,27] Our previous study[28] also confirmed absolute lymphocyte count as an important marker in predicting NPC survival. Platelets and CRP are also important factors affecting the inflammatory response, thereby affecting the prognosis of patients. Platelets may protect tumor cells from immune elimination, promote tumor

metastasis,[29] and mediate tumor cell growth by producing various growth factors.[30] CRP, an acute-phase response protein, is mainly synthesized by hepatocytes induced by proinflammatory cytokines. Elevated CRP levels are related to inflammation accompanied by hypoalbuminemia, resulting in dystrophia and further leading to a poor prognosis.[31] The final four inflammatory rules indicated a potential relationship among these hematological indices, which has rarely been reported and warrants further exploration.

ML is a preferred option for resolving challenges in the fields of statistical analysis, data mining, and model optimization. However, most ML involves extensive parameters and takes a long time to learn, which makes it a huge challenge for clinicians to deal with a large number of algorithm parameters and to interpret the established model from a clinical perspective. AutoML, a wrapper function algorithm that integrates multiple ML algorithms, can automatically select the optimal algorithm and establish a model with high performance (at least in the primary cohort), which seems to relieve the burden of different types of algorithms for clinicians. However, in our study, the very high C-index of the final model in the training cohort may indicate overfitting, which may be validated by the very low C-index in the validation set. Moreover, models obtained by the optimal ML algorithms are usually difficult to interpret clinically. RuleFit, the main algorithm we used, requires a much shorter training time than AutoML but showed better performance in the validation cohort. Most importantly, rules obtained through the RuleFit algorithm are directly related to clinical variables, which greatly improve its clinical interpretability. Using the Cox regression algorithm, we optimized four independent time-event prognosticators that can effectively stratify patients into high- and low-prognostic-risk groups. Incorporating the four rules and clinical risk predictors, we further established a blood model with a C-index that is equivalent to that of the RuleFit model using the same clinical predictors. That is to say, without tuning, RuleFit outperforms AutoML in generalization performance, calculation cost, and interpretability in prognosis of NPC using hematological indices. Furthermore, the nature of interpretability makes further statistical processing like Cox analysis possible, thus obtaining concise and high-efficiency hematological rules.

A large number of patients from two independent clinical centers were enrolled, and a rigorous statistical methodology was developed. However, several limitations of this study need to be addressed. First, although rules obtained by the RuleFit algorithm are interpretable, they seem to have no correlation with the median, cutoff value, or mean values of the hematological data. Thus, clinicians may find it challenging to further explain these rules. Second, distribution differences of some hematological indices were found between the two clinical centers, resulting in the exclusion of these indices from further investigation. Third, the blood parameters may fluctuate over time, which may hide the fittest cutoff and its measurement timing. Absence of more comprehensive SII, SIRI, and/or pan-immune-inflammation value (PIV), as well as chemokines and cytokines is another limitation. Besides, inevitable differences in salvage therapies during the progressive disease state between the patients' groups may have unintentionally aggregated in one group to favor it over the others. Finally, our models were validated in one relatively small cohort of 386 patients, causing a less ideal effect in the AutoML model. Larger and multicenter cohorts can be used to resolve these limitations in future studies.

To the best of our knowledge, our work is the first hematological risk model ever developed for tumors using RuleFit. By identifying a comprehensive set of risk-predictive rules from hematological variables, risk estimation can be performed by examining individual profiles of hematological data. The statistical framework we adopted has the following advantages. First, it deals with a mix of categorical and continuous variables showing no statistical significance in two sets, thereby representing good consistency in different laboratories. Second, it can combine variables of varying biological characteristics without difficulties in interpretation, as rules can provide a clear representation of complex combination data. Third, the dual-center design makes our conclusions more reliable and reproducible. Finally, the introduction of an app makes the application of the model easy for both clinicians and patients.

In conclusion, inflammatory rules based on RuleFit are promising prognostic predictors and closely related with the survival outcomes of NPC patients. Moreover, this study provides us with an effective model of prognostic risk modeling, which can be applied to more prognostic studies.

Acknowledgment

Funding

Disclosure

The authors report no conflicts of interest in this work.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424. doi:10.3322/caac.21492
2. Chen YP, Chan ATC, Le QT, Blanchard P, Sun Y, Ma J. Nasopharyngeal carcinoma. *Lancet*. 2019;394(10192):64–80. doi:10.1016/s0140-6736(19)30956-0
3. Tang LL, Chen YP, Mao YP, et al. Validation of the 8th Edition of the UICC/AJCC staging system for nasopharyngeal carcinoma from endemic areas in the intensity-modulated radiotherapy era. *J Natl Compr Canc Netw*. 2017;15(7):913–919. doi:10.6004/jnccn.2017.0121
4. Fang W, Zhang J, Hong S, et al. EBV-driven LMP1 and IFN-γ up-regulate PD-L1 in nasopharyngeal carcinoma: implications for oncotargeted therapy. *Oncotarget*. 2014;5(23):12189–12202. doi:10.18632/oncotarget.2608
5. Mo X, Wu X, Dong D, et al. Prognostic value of the radiomics-based model in progression-free survival of hypopharyngeal cancer treated with chemoradiation. *Eur Radiol*. 2020;30(2):833–843. doi:10.1007/s00330-019-06452-w
6. Tang LQ, Li CF, Li J, et al. Establishment and validation of prognostic nomograms for endemic nasopharyngeal carcinoma. *J Natl Cancer Inst*. 2016;108(1):djv291. doi:10.1093/jnci/djv291
7. Tu X, Ren J, Zhao Y. Prognostic value of prognostic nutritional index in nasopharyngeal carcinoma: a meta-analysis containing 4511 patients. *Oral Oncol*. 2020;110:104991. doi:10.1016/j.oraloncology.2020.104991
8. Li J, Chen S, Peng S, et al. Prognostic nomogram for patients with Nasopharyngeal Carcinoma incorporating hematological biomarkers and clinical characteristics. *Int J Biol Sci*. 2018;14(5):549–556. doi:10.7150/ijbs.24374
9. Forget P, Echeverria G, Giglioli S, et al. Biomarkers in immunonutrition programme, is there still a need for new ones? A brief review. *Ecancermedicalscience*. 2015;9:546. doi:10.3332/ecancer.2015.546
10. Ohno Y. Role of systemic inflammatory response markers in urological malignancy. *Int J Urol*. 2019;26(1):31–47. doi:10.1111/iju.13801
11. Balkwill F, Mantovani A. Inflammation and cancer: back to Virchow? *Lancet*. 2001;357(9255):539–545. doi:10.1016/s0140-6736(00)04046-0
12. Wilcox RA, Ristow K, Habermann TM, et al. The absolute monocyte and lymphocyte prognostic score predicts survival and identifies high-risk patients in diffuse large-B-cell lymphoma. *Leukemia*. 2011;25(9):1502–1509. doi:10.1038/leu.2011.112
13. Zhang Y, Zhou GQ, Liu X, et al. Exploration and validation of C-reactive protein/albumin ratio as a novel inflammation-based prognostic marker in nasopharyngeal carcinoma. *J Cancer*. 2016;7(11):1406–1412. doi:10.7150/jca.15401
14. Li X, Han Z, Cheng Z, Yu J, Yu X, Liang P. Prognostic value of preoperative absolute lymphocyte count in recurrent hepatocellular carcinoma following thermal ablation: a retrospective analysis. *Onco Targets Ther*. 2014;7:1829–1835. doi:10.2147/ott.S69227
15. Su L, Zhang M, Zhang W, Cai C, Hong J. Pretreatment hematologic markers as prognostic factors in patients with nasopharyngeal carcinoma: a systematic review and meta-analysis. *Medicine*. 2017;96(11):e6364. doi:10.1097/md.0000000000006364
16. Friedman JH, Popescu BE. Predictive learning via rule ensembles. *Ann Appl Stat*. 2008;2(3):916–954. doi:10.1214/07-aoas148
17. Fokkema M, Smits N, Kelderman H, Penninx B. Connecting clinical and actuarial prediction with rule-based methods. *Psychol Assess*. 2015;27(2):636–644. doi:10.1037/pas0000072
18. Lin Y, Huang S, Simon GE, Liu S. Data-based decision rules to personalize depression follow-up. *Sci Rep*. 2018;8(1):5064. doi:10.1038/s41598-018-23326-1
19. Sun YV, Bielak LF, Peyser PA, et al. Application of machine learning algorithms to predict coronary artery calcification with a sibship-based design. *Genet Epidemiol*. 2008;32(4):350–360. doi:10.1002/gepi.20309
20. Lin Y, Qian X, Krischer J, Vehik K, Lee HS, Huang S. A rule-based prognostic model for type 1 diabetes by identifying and synthesizing baseline profile patterns. *PLoS One*. 2014;9(6):e91095. doi:10.1371/journal.pone.0091095
21. Yang L, Hong S, Wang Y, et al. Development and external validation of nomograms for predicting survival in nasopharyngeal carcinoma patients after definitive radiotherapy. *Sci Rep*. 2015;5:15638. doi:10.1038/srep15638
22. Peng RR, Liang ZG, Chen KH, Li L, Qu S, Zhu XD. Nomogram based on lactate dehydrogenase-to-albumin ratio (LAR) and platelet-to-lymphocyte ratio (PLR) for predicting survival in nasopharyngeal carcinoma. *J Inflamm Res*. 2021;14:4019–4033. doi:10.2147/jir.S322475
23. Liu LT, Chen QY, Tang LQ, et al. The prognostic value of treatment-related lymphopenia in nasopharyngeal carcinoma patients. *Cancer Res Treat*. 2018;50(1):19–29. doi:10.4143/crt.2016.595
24. Jiang R, Cai XY, Yang ZH, et al. Elevated peripheral blood lymphocyte-to-monocyte ratio predicts a favorable prognosis in the patients with metastatic nasopharyngeal carcinoma. *Chin J Cancer*. 2015;34(6):237–246. doi:10.1186/s40880-015-0025-7
25. He J, Shen G, Ren Z, et al. Pretreatment levels of peripheral neutrophils and lymphocytes as independent prognostic factors in patients with nasopharyngeal carcinoma. *Head Neck*. 2012;34(12):1769–1776. doi:10.1002/hed.22008

26. Gooden MJ, de Bock GH, Leffers N, Daemen T, Nijman HW. The prognostic influence of tumour-infiltrating lymphocytes in cancer: a systematic review with meta-analysis. *Br J Cancer*. 2011;105(1):93–103. doi:10.1038/bjc.2011.189

27. So YK, Byeon SJ, Ku BM, et al. An increase of CD8(+) T cell infiltration following recurrence is a good prognosticator in HNSCC. *Sci Rep*. 2020;10(1):20059. doi:10.1038/s41598-020-77036-8

28. Lin W, Cao D, Dong A, et al. Systematic construction and external validation of an immune-related prognostic model for nasopharyngeal carcinoma. *Head Neck*. 2022;44:1086–1098. doi:10.1002/hed.26996

29. Gay LJ, Felding-Habermann B. Contribution of platelets to tumour metastasis. *Nat Rev Cancer*. 2011;11(2):123–134. doi:10.1038/nrc3004

30. Haemmerle M, Stone RL, Menter DG, Afshar-Kharghan V, Sood AK. The platelet lifeline to cancer: challenges and opportunities. *Cancer Cell*. 2018;33(6):965–983. doi:10.1016/j.ccell.2018.03.002

31. Yamagata K, Fukuzawa S, Ishibashi-Kanno N, Uchida F, Bukawa H. Association between the C-reactive protein/albumin ratio and prognosis in patients with oral squamous cell carcinoma. *Sci Rep*. 2021;11(1):5446. doi:10.1038/s41598-021-83362-2