



## OPEN ACCESS

## EDITED BY

Yong Yin,  
Shandong University, China

## REVIEWED BY

Mi Chen,  
Mianyang Third People's Hospital, China  
Chen Jian,  
The Second Affiliated Hospital of Chongqing  
Medical University, China

## \*CORRESPONDENCE

Jie Zhou  
✉ zjoncology2023@163.com

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

RECEIVED 19 June 2025

ACCEPTED 21 August 2025

PUBLISHED 12 September 2025

## CITATION

Yang Y, Shang N, Lu S, Li L, Xu P, Wang X,  
Li F, Su Y, Qin Y, Lang J and Zhou J (2025)  
Exploring the prognostic value of EBV DNA in  
advanced nasopharyngeal carcinoma treated  
with chemoradiotherapy using AI-based  
modeling.  
*Front. Oncol.* 15:1650377.  
doi: 10.3389/fonc.2025.1650377

## COPYRIGHT

© 2025 Yang, Shang, Lu, Li, Xu, Wang, Li, Su,  
Qin, Lang and Zhou. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).  
The use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Exploring the prognostic value of EBV DNA in advanced nasopharyngeal carcinoma treated with chemoradiotherapy using AI-based modeling

Yang Yang<sup>1†</sup>, Ningchuan Shang<sup>2†</sup>, Shun Lu<sup>3</sup>, Lintao Li<sup>3</sup>,  
Peng Xu<sup>3</sup>, Xianliang Wang<sup>3</sup>, Fan Li<sup>4</sup>, Yue Su<sup>3</sup>, Yuan Qin<sup>3</sup>,  
Jinyi Lang<sup>3</sup> and Jie Zhou<sup>3\*†</sup>

<sup>1</sup>Department of Oncology, The Third People's Hospital of Chengdu, Chengdu, China, <sup>2</sup>School of Clinical Medicine, Sichuan College of Traditional Chinese Medicine, Mianyang, China, <sup>3</sup>Department of Radiation Oncology, Precision Radiation in Oncology Key Laboratory of Sichuan Province, Sichuan Clinical Research Center for Cancer, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, Chengdu, China, <sup>4</sup>Department of Otorhinolaryngology Head and Neck Surgery, Chongqing General Hospital, Chongqing University, Chongqing, China

**Background:** Epstein–Barr virus (EBV) DNA is a well-established biomarker in nasopharyngeal carcinoma (NPC), but its integration into artificial intelligence (AI)–based prognostic tools remains limited. This study aimed to develop and validate AI models incorporating EBV DNA load levels to predict progression-free survival (PFS) in patients with advanced NPC treated with concurrent chemoradiotherapy (CRT).

**Methods:** A retrospective multicenter cohort of 503 patients was divided into training (n = 301) and validation (n = 202) sets. Four machine learning algorithms—Cox regression, LASSO, RSF, and GBM—were applied to predict 1- and 1.5-year PFS in patients with advanced NPC. Model performance was evaluated using the concordance index (C-index), time-dependent receiver operating characteristic (ROC), decision curve analysis (DCA), and interpretability tools such as SHAP values and partial dependence plots (PDP).

**Results:** The 1-, 3-, and 5-year PFS rates were 100.0%, 91.5%, and 88.6% in the EBV = 0 group; 99.4%, 91.2%, and 88.5% in the > 0 and < 1500 group; and 92.3%, 81.0%, and 75.7% in the ≥ 1500 group, respectively, with statistically significant differences among the three groups ( $P = 0.0024$ ). The RSF model outperformed other models with the highest C-index (0.778) and area under the ROC curve of 0.810 and 0.634 at 1 and 1.5 years, respectively. EBV DNA emerged as the most influential predictor across all interpretability analyses. Patients with EBV DNA ≥ 1500 copies/ml had the poorest predicted survival, showing a distinct threshold effect in the PDP.

**Conclusions:** High EBV DNA levels were associated with poorer PFS in advanced NPC. Among the models evaluated, the RSF model demonstrated the best predictive performance and interpretability. EBV-informed AI modeling represents a promising approach for enhancing individualized risk prediction and clinical decision-making in NPC.

#### KEYWORDS

**EBV DNA, advanced nasopharyngeal carcinoma, prognostic value, artificial intelligence, machine learning, chemoradiotherapy**

## Introduction

Nasopharyngeal carcinoma (NPC) is a malignancy of the head and neck region with high prevalence in east and Southeast Asia, and its development is strongly linked to Epstein–Barr virus (EBV) infection (1, 2). Although intensity-modulated radiotherapy (IMRT) combined with chemotherapy has become the standard treatment for advanced-stage disease, recurrence and metastasis remain major clinical challenges (3). Plasma EBV DNA has emerged as a valuable biomarker for estimating tumor burden and predicting prognosis, however, its accuracy and consistency in different clinical scenarios are still debated (4, 5).

Conventional regression-based approaches may not fully uncover the intricate relationships between EBV levels and various prognostic factors. In contrast, artificial intelligence (AI) techniques, especially machine learning models, provide a powerful framework for handling complex datasets and capturing nonlinear associations (6–8). These models have shown increasing promise in cancer prognosis and treatment optimization (9, 10). Nonetheless, few studies have systematically evaluated the integration of EBV DNA into AI-based prediction tools for NPC.

In this study, we propose to develop a robust and interpretable AI model incorporating EBV DNA levels and clinical characteristics to enhance survival prediction in patients undergoing concurrent chemoradiotherapy. This model aims to assist clinicians in tailoring personalized treatment strategies and long-term monitoring plans.

## Materials and methods

### Study design and participants

From March 2012 to May 2020, 503 patients with advanced-stage NPC who received concurrent chemoradiotherapy (CRT) were retrospectively reviewed from four hospitals in China. Inclusion criteria for this study were as follows: (1) histologically confirmed nasopharyngeal carcinoma; (2) stage III–IV disease; (3) received first-line concurrent CRT as the initial treatment; (4) complete baseline EBV DNA measurements before treatment;

(5) complete follow-up and survival data available. Exclusion criteria included: (1) history of prior malignancy or previous anti-cancer treatment; (2) with serious complications or any other serious chronic diseases; (3) unable to complete the treatment. This study was approved by the institutional review board and ethics committees (KY S2023-081-01). As this was a retrospective study, the informed consent was waived by the Ethics Committee. Participant information is confidential, and the study was conducted in accordance with the Declaration of Helsinki.

### Treatments

All enrolled patients underwent IMRT alongside synchronous chemotherapy based on platinum compounds. The delineation of gross tumor volume (GTV) included two components: the nasopharyngeal lesion itself (GTV<sub>nx</sub>) and lymph nodes confirmed as metastatic through clinical or radiologic evaluation (GTV<sub>nd</sub>), primarily based on MRI and/or PET-CT findings. To capture areas potentially harboring microscopic disease, clinical target volumes (CTV) were contoured by expanding the GTVs to incorporate adjacent high-risk tissues. Planning target volumes (PTV) were created by applying an additional 3–5 mm margin around each CTV, accounting for patient movement and setup variation during treatment sessions. A total radiation dose of 70 Gy was prescribed, aimed at ensuring that no less than 95% of the PTV received the full dose. The course was delivered over 30 to 33 fractions, typically completed within a period of 6 to 7 weeks.

### Endpoints and follow-up

Progression-free survival (PFS) was defined as the interval from the initiation of treatment to either documented disease progression, as assessed by imaging according to RECIST criteria, or death from any cause, whichever occurred first. Disease progression was evaluated based on institutional radiology reports without central imaging review, reflecting real-world clinical practice across participating centers.

TABLE 1 Baseline clinical characteristics stratified by EBV DNA levels.

| Characteristic | All         | Pre-EBV DNA |             |             | P       |
|----------------|-------------|-------------|-------------|-------------|---------|
|                |             | 0           | > 0, < 1500 | ≥ 1500      |         |
|                | 503         | 120         | 173         | 210         |         |
| Age            |             |             |             |             | 0.762   |
| < 45           | 218 (43.3%) | 54 (45.0%)  | 77 (44.5%)  | 87 (41.4%)  |         |
| ≥ 45           | 285 (56.7%) | 66 (55.0%)  | 96 (55.5%)  | 123 (58.6%) |         |
| Sex            |             |             |             |             | 0.967   |
| Female         | 139 (27.6%) | 33 (27.5%)  | 49 (28.3%)  | 57 (27.1%)  |         |
| Male           | 364 (72.4%) | 87 (72.5%)  | 124 (71.7%) | 153 (72.9%) |         |
| T Stage        |             |             |             |             | 0.004   |
| T1             | 7 (1.39%)   | 2 (1.67%)   | 2 (1.16%)   | 3 (1.43%)   |         |
| T2             | 45 (8.95%)  | 13 (10.8%)  | 15 (8.67%)  | 17 (8.10%)  |         |
| T3             | 371 (73.8%) | 93 (77.5%)  | 139 (80.3%) | 139 (66.2%) |         |
| T4             | 80 (15.9%)  | 12 (10.0%)  | 17 (9.83%)  | 51 (24.3%)  |         |
| N Stage        |             |             |             |             | < 0.001 |
| N0             | 45 (8.95%)  | 22 (18.3%)  | 17 (9.83%)  | 6 (2.86%)   |         |
| N1             | 171 (34.0%) | 46 (38.3%)  | 56 (32.4%)  | 69 (32.9%)  |         |
| N2             | 215 (42.7%) | 43 (35.8%)  | 82 (47.4%)  | 90 (42.9%)  |         |
| N3             | 72 (14.3%)  | 9 (7.50%)   | 18 (10.4%)  | 45 (21.4%)  |         |
| Stage          |             |             |             |             | < 0.001 |
| III            | 362 (72.0%) | 100 (83.3%) | 138 (79.8%) | 124 (59.0%) |         |
| IV             | 141 (28.0%) | 20 (16.7%)  | 35 (20.2%)  | 86 (41.0%)  |         |

## AI-model

A retrospective dataset comprising 503 patients was split into two groups using a 6:4 random allocation strategy, resulting in a training cohort of 301 cases and a validation cohort of 202 cases. The training set was used to construct predictive models for progression-free survival (PFS) at both 1 and 1.5 years. The following six baseline variables were incorporated as potential prognostic indicators: sex, age, T stage, N stage, clinical stage, and EBV DNA levels at diagnosis.

Four modeling algorithms were employed: Cox proportional hazards model, least absolute shrinkage and selection operator (LASSO) regression, random survival forest (RSF), and gradient boosting machine (GBM). Each model’s discriminative performance was assessed via concordance index (C-index) metrics.

Model performance was subsequently evaluated in the independent validation cohort using time-dependent receiver operating characteristic (ROC) curves and decision curve analysis (DCA). To further elucidate the underlying mechanics of the models, explainability methods such as SHAP values, partial dependence survival plots (PDP), time-dependent feature importance analysis, and Brier scores were applied in the training group.

## Statistical analyses

Categorical variables were compared using the chi-square test. The Kaplan–Meier method was used to estimate PFS, and survival differences between groups were assessed by the log-rank test. All statistical analyses related to AI-based modeling were performed using R software.

## Results

### Baseline characteristics

The cohort was divided into three groups: EBV DNA = 0 copies/ml (n = 120), > 0 and < 1500 copies/ml (n = 173), and ≥ 1500 copies/ml (n = 210). There were no statistically significant differences among the groups in terms of age (p = 0.762) or sex (p = 0.967). However, significant differences were observed in T stage distribution (p = 0.004). The proportion of T4 disease increased with rising EBV DNA levels, from 10.0% in the 0 group to 24.3% in the ≥ 1500 group. Similarly, N stage was significantly associated with EBV DNA levels (p < 0.001). The frequency of N3 involvement was markedly higher in the ≥ 1500 group (21.4%)

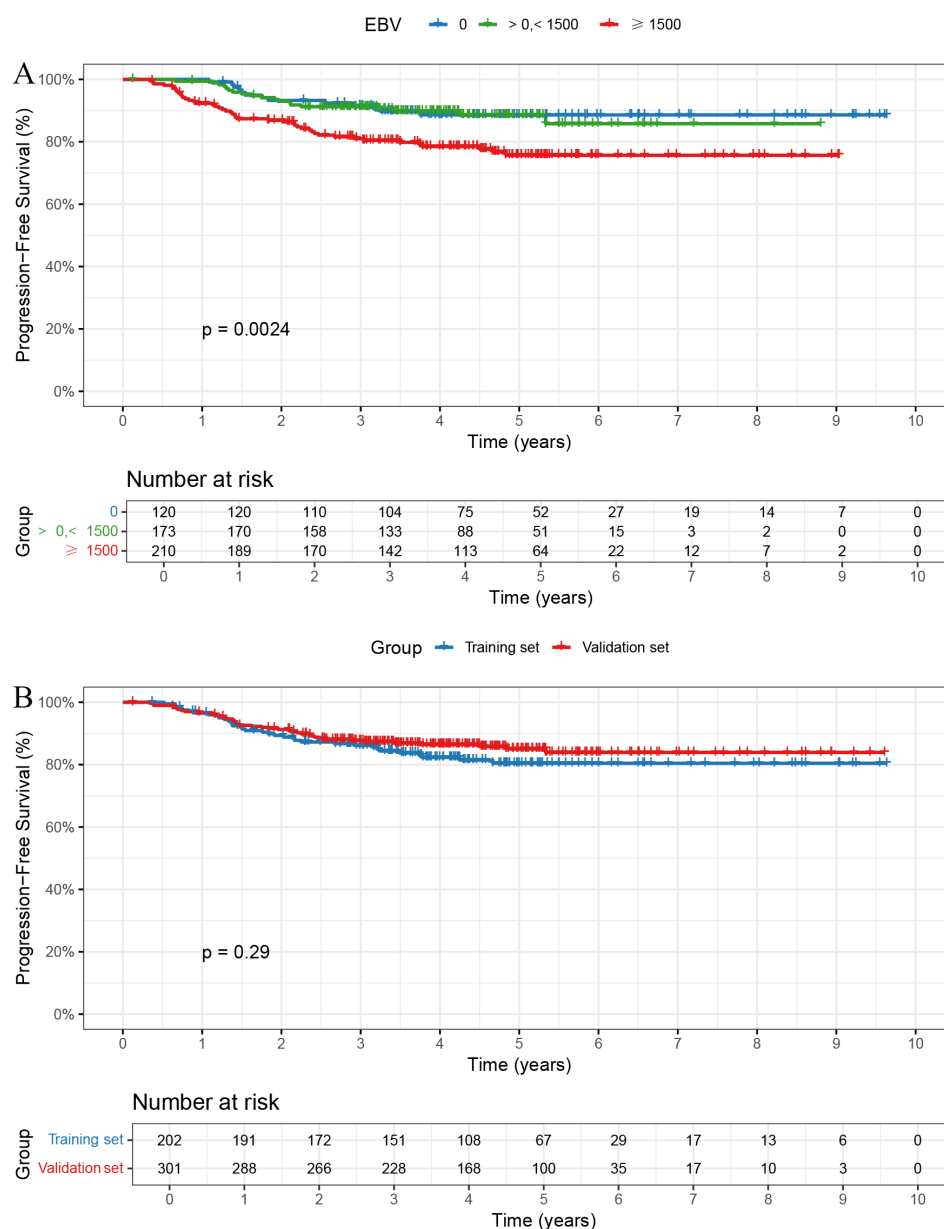


FIGURE 1

Kaplan-Meier curves for PFS. (A) Patients with EBV DNA  $\geq 1500$  copies/ml had significantly worse PFS ( $p = 0.0024$ ). (B) No significant PFS difference between training and validation sets ( $p = 0.29$ ).

compared to the 0 group (7.5%). A significant shift in clinical stage was also observed ( $p < 0.001$ ), with stage IV disease more prevalent among patients with high EBV DNA (41.0%) compared to those with undetectable levels (16.7%, Table 1).

## Survival outcomes

The follow-up duration for the cohort was consistently recorded. The median follow-up time was 4.59 years, with follow-up data collected through telephone follow-ups and hospital records. Patients were censored at the time of the last follow-up or death, whichever occurred first. In the overall cohort, the 1-, 3-,

and 5-year PFS rates were 96.60%, 87.10%, and 83.19%, respectively. The 1-, 3-, and 5-year PFS rates were 100.0%, 91.5%, and 88.6% in the EBV = 0 group; 99.4%, 91.2%, and 88.5% in the  $> 0$  and  $< 1500$  group; and 92.3%, 81.0%, and 75.7% in the  $\geq 1500$  group, respectively, with statistically significant differences among the three groups (Figure 1A,  $P = 0.0024$ ).

## AI-model construction

Baseline characteristics were well balanced between the training and validation sets, except for sex, which showed a statistically significant difference (Table 2). The 1-, 3-, and 5-year PFS rates were

TABLE 2 Comparison of baseline clinical characteristics between training and validation sets.

| Characteristic | Validation set | Training set | P     |
|----------------|----------------|--------------|-------|
|                | 202            | 301          |       |
| Age            |                |              | 0.363 |
| < 45           | 93 (46.0%)     | 125 (41.5%)  |       |
| ≥ 45           | 109 (54.0%)    | 176 (58.5%)  |       |
| Sex            |                |              | 0.003 |
| Female         | 71 (35.1%)     | 68 (22.6%)   |       |
| Male           | 131 (64.9%)    | 233 (77.4%)  |       |
| T Stage        |                |              | 0.177 |
| T1             | 0 (0.00%)      | 7 (2.33%)    |       |
| T2             | 18 (8.91%)     | 27 (8.97%)   |       |
| T3             | 151 (74.8%)    | 220 (73.1%)  |       |
| T4             | 33 (16.3%)     | 47 (15.6%)   |       |
| N Stage        |                |              | 0.965 |
| N0             | 17 (8.42%)     | 28 (9.30%)   |       |
| N1             | 71 (35.1%)     | 100 (33.2%)  |       |
| N2             | 85 (42.1%)     | 130 (43.2%)  |       |
| N3             | 29 (14.4%)     | 43 (14.3%)   |       |
| Stage          |                |              | 0.56  |
| III            | 142 (70.3%)    | 220 (73.1%)  |       |
| IV             | 60 (29.7%)     | 81 (26.9%)   |       |
| EBV DNA        |                |              | 0.757 |
| 0              | 48 (23.8%)     | 72 (23.9%)   |       |
| > 0, < 1500    | 66 (32.7%)     | 107 (35.5%)  |       |
| ≥ 1500         | 88 (43.6%)     | 122 (40.5%)  |       |

96.66%, 87.72%, and 85.03% in the training set, and 96.50%, 86.15%, and 80.43% in the validation set, respectively, with no significant differences observed between the two groups (Figure 1B,  $P = 0.29$ ).

In the training cohort, six clinical variables—sex, age, T stage, N stage, clinical stage, and EBV DNA levels—were selected to construct four prognostic models using different machine learning approaches. The C-index were 0.740 for the Cox model, 0.707 for the LASSO model, 0.778 for the RSF, and 0.732 for the GBM.

AI model validation

In the validation cohort, the 1- and 1.5-year area under the ROC curve (AUC) were 0.743 and 0.668 for the Cox model (Figure 2A), 0.801 and 0.636 for the LASSO model (Figure 2B), 0.810 and 0.634 for the RSF model (Figure 2C), and 0.785 and 0.658 for the GBM model (Figure 2D). The DCA demonstrated that the RSF model

provided favorable net clinical benefit and showed good stability across different threshold probabilities (Figures 3A, B).

Figure 4 illustrates the interpretability and performance metrics of the random survival forest (RSF) model in the training cohort. The time-dependent feature importance analysis (Figure 4A) showed that EBV DNA level and N stage were the most influential predictors throughout the entire follow-up period. The SHAP plot further demonstrated that both EBV and N stage were positively associated with increased risk (Figure 4B). Model performance was stable over time, as reflected by a consistently low Brier score and a time-dependent concordance index (C/D AUC) that remained above 0.8 for a substantial portion of the follow-up (Figure 4C). The partial dependence plot (PDP) showed that patients with EBV DNA  $\geq 1500$  copies/ml had the lowest predicted survival throughout the follow-up period (Figure 5).

Discussion

In recent years, EBV DNA has emerged as a key biomarker in the management of NPC, particularly in regions with a high disease burden such as Southeast Asia (11–13). While its prognostic relevance is well established, translating EBV DNA levels into individualized survival predictions remains a challenge. In this study, we sought to address this gap by constructing AI-based models that integrate EBV DNA with standard clinical features to predict PFS in advanced NPC patients treated with concurrent CRT.

Unlike traditional statistical methods, which typically assume linear associations and fixed hazard ratios, AI algorithms are capable of modeling more complex, nonlinear interactions (14, 15). This advantage is particularly relevant in heterogeneous cancers such as NPC, which clinical behavior is influenced by both tumor burden and host factors (16). By incorporating machine learning techniques—namely Cox regression, LASSO, RSF, and GBM—we aimed to compare the predictive capacity of conventional and data-driven approaches (17–20).

Among the four models tested, the RSF algorithm demonstrated the strongest performance (21). It achieved the highest C-index in the training cohort and showed stable results in the validation set. Beyond overall accuracy, interpretability analyses—including SHAP values, time-dependent feature importance, and partial dependence plots—consistently highlighted EBV DNA as the most significant predictor of survival outcomes. These findings reinforce the notion that EBV DNA is not merely a passive biomarker, but an active contributor to risk stratification in AI-driven models.

The partial dependence plot for EBV DNA revealed a clear inflection point: patients with baseline EBV DNA  $\geq 1500$  copies/ml consistently showed markedly lower predicted survival probabilities. Interestingly, survival curves for patients with undetectable or low-level EBV DNA copies were similar to those of patients with EBV DNA  $> 0$  and  $< 1500$  copies/ml, suggesting a threshold effect. These observations are in line with previous clinical studies and further validate the value of integration of EBV DNA into automated risk prediction frameworks.

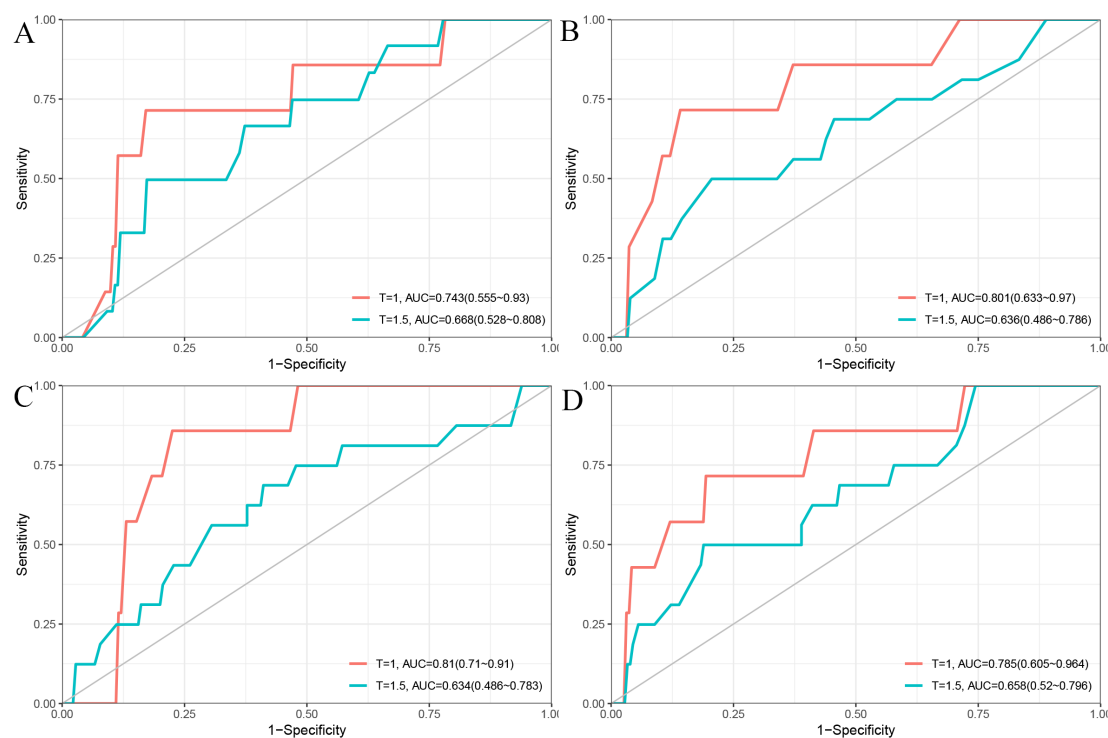


FIGURE 2

Time-dependent ROC curves for 1-year and 1.5-year PFS prediction using different models. (A) Cox model; (B) LASSO model; (C) RSF model; (D) GBM model.

Another strength of this study lies in its focus on model interpretability. One of the major criticisms of AI in healthcare is the “black box” nature of many algorithms (22, 23). By applying tools such as SHAP and PDP, we made our models more transparent and clinically meaningful (24). These visualizations allow clinicians understanding not only that EBV matters, but how and when it exerts the greatest impact on patient outcomes.

In our study, the RSF model demonstrated superior performance compared to other machine learning algorithms. RSF is an ensemble learning method that is particularly well-suited for survival analysis due to its ability to handle both censored data and complex non-linear relationships between predictors and survival outcomes (25, 26). Unlike traditional survival models, RSF does not require the assumption of

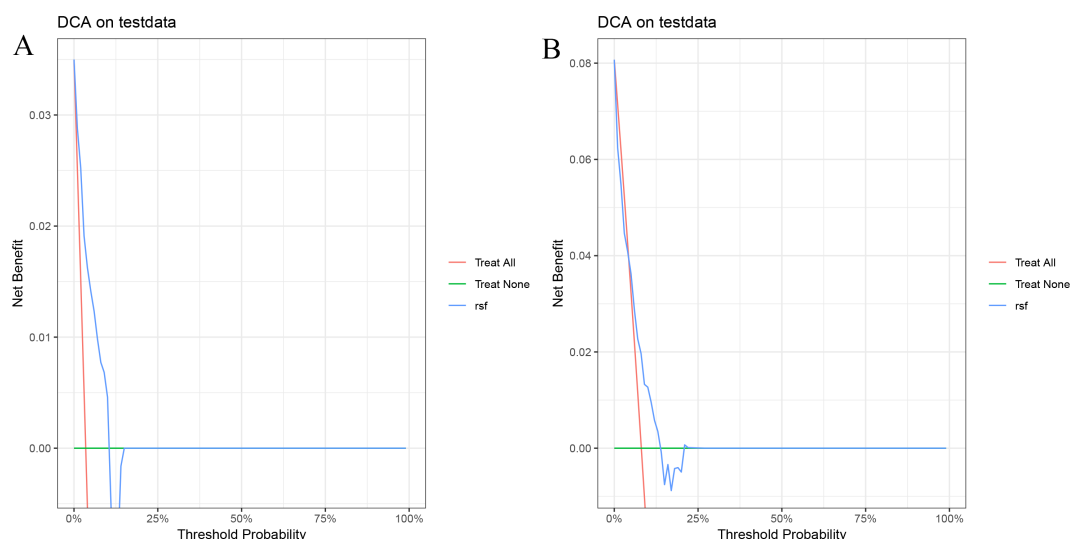
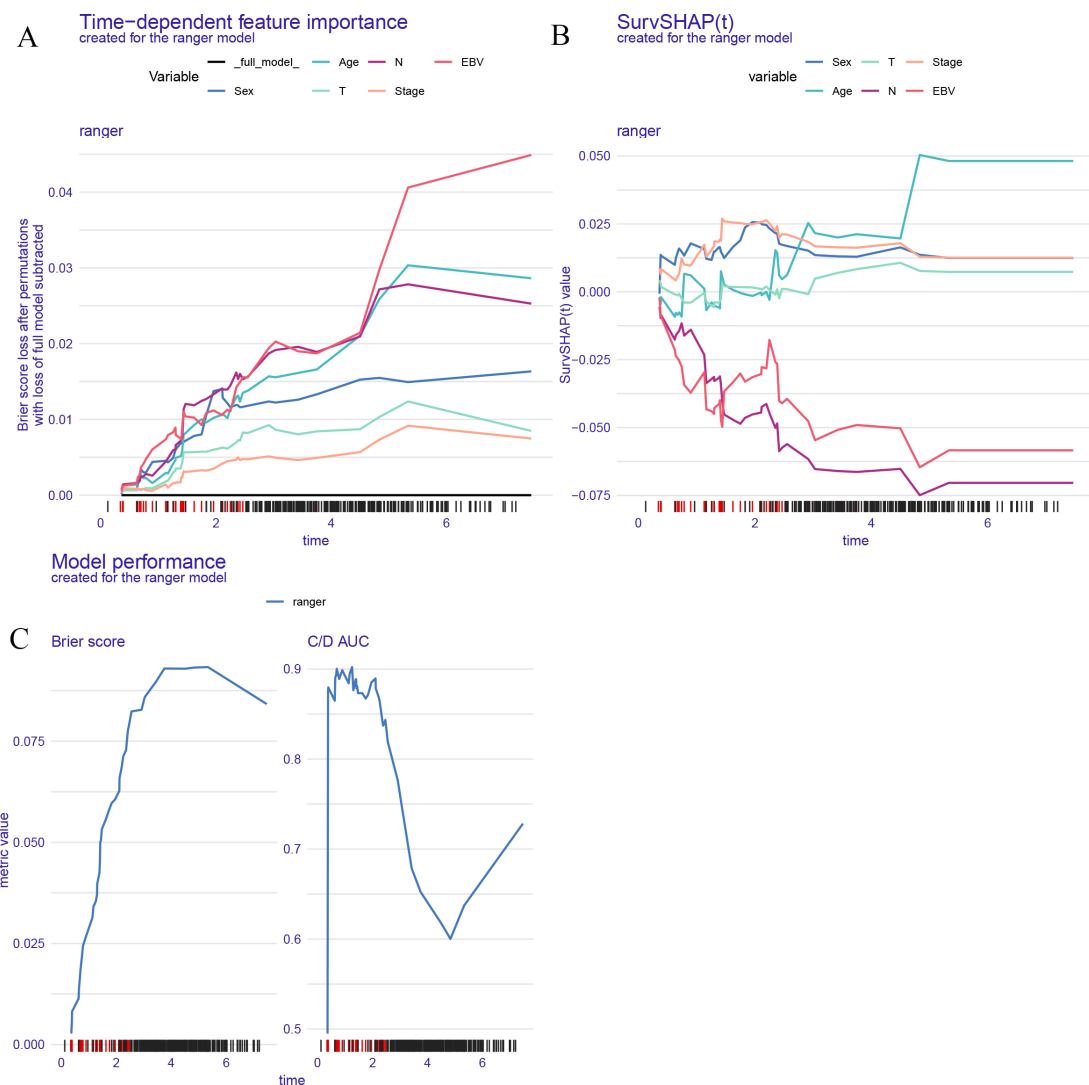


FIGURE 3

Decision curve analysis of the RSF model on the test set. (A) 1-year PFS; (B) 1.5-year PFS.



**FIGURE 4**  
Model interpretability and performance of the RSF model. **(A)** Time dependent feature importance shows EBV DNA and N stage had the greatest impact on model performance over time. **(B)** SurvSHAP values illustrate the dynamic contribution of each variable to individual risk prediction. **(C)** Brier score and time dependent C/D AUC demonstrate the model's predictive accuracy and discrimination ability over time.

proportional hazards and can model interactions between variables more effectively. Additionally, RSF can handle high-dimensional datasets with many variables and is less prone to overfitting compared to other algorithms, making it a robust choice for survival prediction in clinical datasets like ours (27).

From a clinical standpoint, integrating EBV DNA into AI-based models offers more than statistical insight—it provides a practical tool for precision medicine (28–30). By accurately identifying patients at higher risk of disease progression, clinicians can consider early treatment intensification, closer surveillance, or inclusion in clinical trials. Conversely, patients with low-risk profiles may benefit from treatment de-escalation, reducing toxicity and preserving quality of life. The model's transparency also enhances clinical confidence, making it more acceptable for integration into multidisciplinary tumor boards. As such, EBV-informed AI models may serve as decision-support systems that

personalize management strategies and improve long-term outcomes in advanced NPC (31–33).

Although the current model demonstrates promising results, integrating other modalities such as radiomics, genomic alterations, and immune-related biomarkers could further improve its prognostic performance. Radiomics could provide detailed imaging features related to tumor phenotype, allowing for more accurate risk stratification. Genomic alterations, including mutations, copy number variations, and methylation patterns, could offer insights into the underlying molecular mechanisms of tumor progression and treatment resistance (34). Additionally, immune-related biomarkers, such as tumor-infiltrating lymphocytes or checkpoint markers, may help capture the immune response to therapy, which is often a critical factor in cancer prognosis (35). Future studies incorporating these multimodal data types, in combination with EBV DNA, would



### Partial dependence survival profiles

created for the ranger model

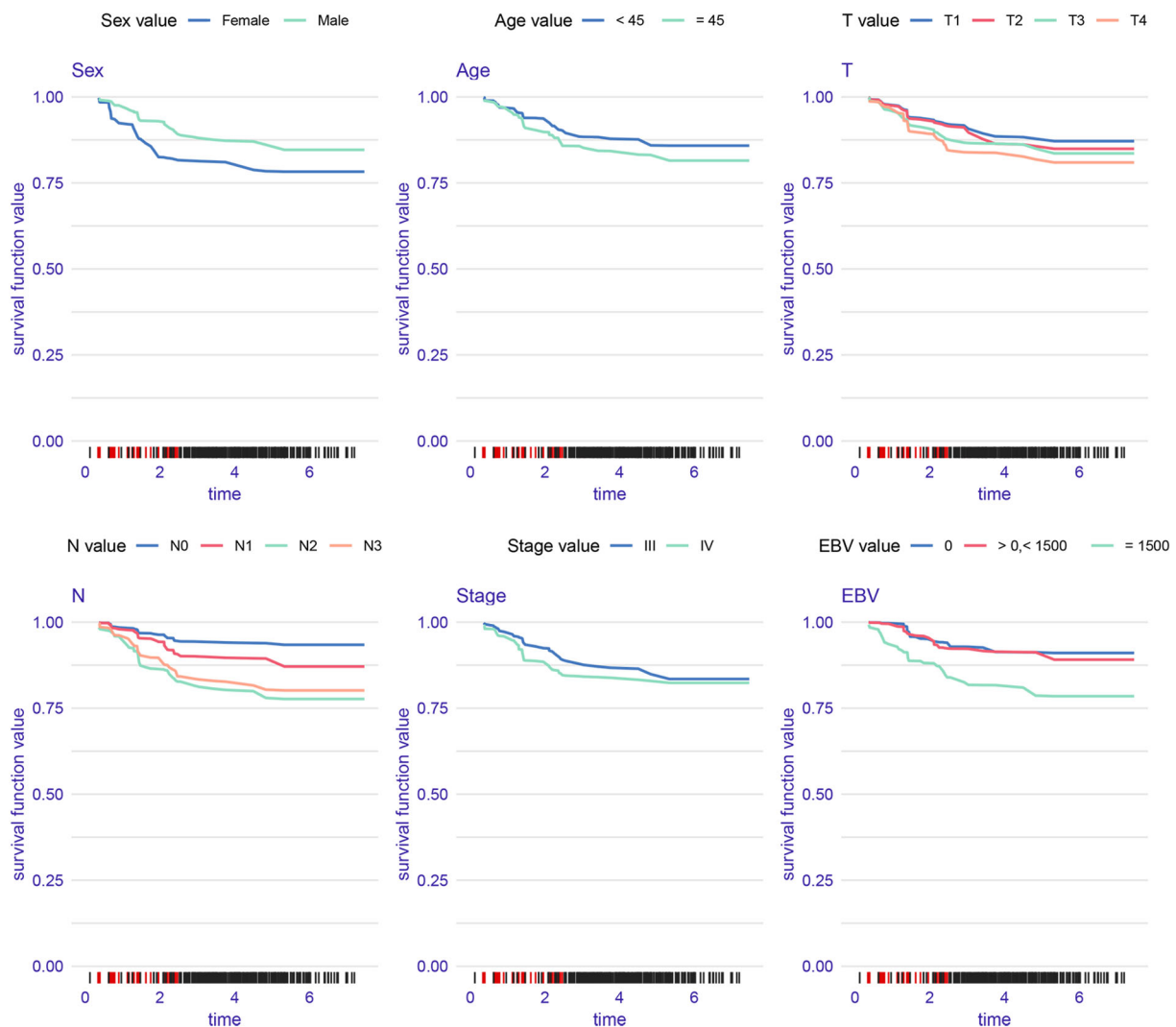


FIGURE 5  
Partial dependence survival curves of key variables in the RSF model.

likely improve the model's accuracy and utility, providing more comprehensive and personalized predictions for patients with advanced NPC.

Despite its strengths, this study has several limitations. First, the retrospective nature of the analysis, even though based on a multicenter cohort, may introduce selection and information bias. Second, while EBV DNA testing was performed uniformly within each center, inter-institutional variability in laboratory procedures cannot be completely ruled out. Third, although treatment protocols were broadly standardized (IMRT with platinum-based concurrent chemotherapy), subtle differences in chemotherapy regimens, supportive care, or radiation planning across centers may contribute to treatment heterogeneity, potentially confounding survival outcomes. Fourth, the model did not incorporate other potentially informative modalities such as radiomics, genomic alterations, or immune-related biomarkers, which could further enhance predictive accuracy. Finally, the absence

of external validation limits the generalizability of the model. Future prospective studies should include external datasets and integrate multimodal information to validate and refine AI-based prognostic models for advanced NPC.

## Conclusion

In conclusion, our study demonstrates that elevated EBV DNA levels are closely associated with poorer PFS in patients with advanced NPC. By integrating EBV DNA with advanced AI algorithms, particularly random survival forests, we were able to more effectively characterize its prognostic impact across the disease course. These findings highlight the potential of EBV-informed AI models as valuable tools for enhancing risk stratification and guiding personalized treatment decisions in clinical practice.



## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

The studies involving humans were approved by the Ethics Committee of Chongqing General Hospital (KY S2023-081-01). Participant information is confidential, and the study was conducted in accordance with the Declaration of Helsinki. The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because this was a retrospective study.

## Author contributions

YY: Methodology, Data curation, Funding acquisition, Formal Analysis, Writing – original draft. NS: Formal Analysis, Writing – original draft, Data curation, Methodology. SL: Software, Data curation, Conceptualization, Writing – review & editing, Supervision. LL: Data curation, Project administration, Writing – review & editing, Investigation. FL: Validation, Data curation, Conceptualization, Writing – original draft. PX: Writing – review & editing, Methodology, Software. XW: Software, Methodology, Data curation, Writing – review & editing. YS: Data curation, Writing – original draft. YQ: Writing – original draft, Data curation. JL: Writing – review & editing, Conceptualization, Supervision. JZ: Conceptualization, Supervision, Writing – review & editing, Methodology, Resources, Project administration, Funding acquisition.

## References

1. Hsieh HT, Zhang XY, Wang Y, Cheng XQ. Biomarkers for nasopharyngeal carcinoma. *Clin Chim Acta*. (2025) 572:120257. doi: 10.1016/j.cca.2025.120257
2. Ahmed N, Abusalah MAHA, Farzand A, Absar M, Yusof NY, Rabaa AA, et al. Updates on Epstein-Barr virus (EBV)-associated nasopharyngeal carcinoma: emphasis on the latent gene products of EBV. *Medicina-Lithuania*. (2022) 59(1):1011. doi: 10.3390/medicina59010002
3. Cao C, Fang Y, Jiang F, Jin Q, Jin T, Huang S, et al. Concurrent chemotherapy for older patients with locoregionally advanced nasopharyngeal carcinoma: A randomized clinical trial. *J Geriatr Oncol*. (2025) 16:102246. doi: 10.1016/j.jgo.2025.102246
4. Feng E, Yang Y, Yang J, Hu R, Tian L, Yang X, et al. Tumor-infiltrating CD4+ CD25+ FOXP3+ Treg is associated with plasma EBV DNA and disease progression in nasopharyngeal carcinoma. *Infect Agent Cancer*. (2025) 20:29. doi: 10.1186/s13027-025-00660-4
5. Neo J, Ong EHW, Zhang X, Chow WM, Wee JTS, Fong KW, et al. Plasma Epstein-Barr virus DNA for disease surveillance in endemic nasopharyngeal carcinoma: Analysis of a real-world database. *Eur J Cancer*. (2025) 220:115396. doi: 10.1016/j.ejca.2025.115396
6. He R, Jie P, Hou W, Long Y, Zhou G, Wu S, et al. Real-time artificial intelligence-assisted detection and segmentation of nasopharyngeal carcinoma using multimodal endoscopic data: a multi-center, prospective study. *EClinicalMedicine*. (2025) 81:103120. doi: 10.1016/j.eclinm.2025.103120
7. Shannon NB, Lye NG, Chua MLK. Leveraging artificial intelligence and radiomics for improved nasopharyngeal carcinoma prognostication. *Cancer Med*. (2025) 14:e70706. doi: 10.1002/cam4.70706
8. Yang YX, Yang X, Jiang XB, Lin L, Wang GY, Sun WZ, et al. Artificial intelligence-empowered multistep integrated radiation therapy workflow for nasopharyngeal carcinoma. *Int J Radiat Oncol Biol Phys*. (2024) 122(4):902–13. doi: 10.1016/j.ijrobp.2024.11.096
9. Alabi RO, Elmusrati M, Leivo I, Almangush A, Mäkitie AA. Machine learning explainability in nasopharyngeal cancer survival using LIME and SHAP. *Sci Rep*. (2023) 13:8984. doi: 10.1038/s41598-023-35795-0
10. Yang X, Wu J, Chen X. Application of artificial intelligence to the diagnosis and therapy of nasopharyngeal carcinoma. *J Clin Med*. (2023) 12(9):3077. doi: 10.3390/jcm12093077
11. Lu GZ, Huang YX, Guo LF, Yu YF, Lu ZZ, Lin Q, et al. Plasma Epstein-Barr virus DNA for the prediction of treatment response and disease progression in non-keratinizing differentiated nasopharyngeal carcinoma. *Infect Agent Cancer*. (2025) 20:30. doi: 10.1186/s13027-025-00661-3

## Funding

The author(s) declare financial support was received for the research and/or publication of this article. This work was supported by the Natural Science Foundation of Science and Technology Department of Sichuan Province (grant No. 2021YFH0138; No.2025ZNSFSC0787), Special Funding for Postdoctoral Research Projects in Sichuan Province (grant No. TB2022090), and Cancer Clinical Research Funding Project of Sichuan Anti-Cancer Association (grant No. XH2023 3005).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

12. Cheng H, Chen J, Li Y, Li Y, Tse C, Shen B, et al. Determining the optimal duration of oral adjuvant chemotherapy in locoregionally advanced nasopharyngeal carcinoma. *Br J Cancer*. (2025) 133(1):85–93. doi: 10.1038/s41416-025-03033-1
13. Zarkovic G, Ziegler P, Lee JH, Dresden B, Kumar A, Shuda M, et al. Disruption of ATR signaling by epstein-barr virus latent membrane protein 1 sensitizes nasopharyngeal carcinoma cells to cisplatin. *J Med Virol*. (2025) 97:e70407. doi: 10.1002/jmv.70407
14. Ng WT, But B, Choi HCW, de Bree R, Lee AWM, Lee VHF, et al. Application of artificial intelligence for nasopharyngeal carcinoma management-A systematic review. *Cancer Manag Res*. (2022) 14:339–66. doi: 10.2147/CMAR.S341583
15. Lee AW, Ng WT, Pan JJ, Poh SS, Ahn YC, AlHussain H, et al. International guideline for the delineation of the clinical target volumes (CTV) for nasopharyngeal carcinoma. *Radiother Oncol*. (2018) 126:25–36. doi: 10.1016/j.radonc.2017.10.032
16. Li B, Li H, Chen J, Xiao F, Fang X, Guo R, et al. A magnetic resonance imaging (MRI)-based deep learning radiomics model predicts recurrence-free survival in lung cancer patients after surgical resection of brain metastases. *Clin Radiol*. (2025) 85:106920. doi: 10.1016/j.crad.2025.106920
17. Hu C, Xu C, Chen J, Huang Y, Meng Q, Lin Z, et al. Deep learning MRI-based radiomic models for predicting recurrence in locally advanced nasopharyngeal carcinoma after neoadjuvant chemoradiotherapy: a multi-center study. *Clin Exp Metastasis*. (2025) 42:30. doi: 10.1007/s10585-025-10349-y
18. Mu X, Ge Z, Lu D, Li T, Liu L, Chen C, et al. Deep learning model using planar whole-body bone scintigraphy for diagnosis of skull base invasion in patients with nasopharyngeal carcinoma. *J Cancer Res Clin Oncol*. (2024) 150:449. doi: 10.1007/s00432-024-05969-y
19. Sun J, Lam SKE, Teng X, Zhang J, Lee FK, Yip CW, et al. Predicting disease progression from the rate of bodyweight change in nasopharyngeal carcinoma patient during radiotherapy. *Sci Rep*. (2025) 15:7490. doi: 10.1038/s41598-025-88810-x
20. Wang X, Song J, Qiu Q, Su Y, Wang L, Cao X. A stacked multimodality model based on functional MRI features and deep learning radiomics for predicting the early response to radiotherapy in nasopharyngeal carcinoma. *Acad Radiol*. (2024) 32:1631–44. doi: 10.1016/j.acra.2024.10.011
21. Li J, Ye Y, Cai Y, Ji H, Qin W, Luo Y, et al. Triglyceride-inflammation score established on account of random survival forest for predicting survival in patients with nasopharyngeal carcinoma: a retrospective study. *Front Immunol*. (2024) 15:1375931. doi: 10.3389/fimmu.2024.1375931
22. Wang L, Qiu T, Zhou J, Zhou Y, Sun B, Yang G, et al. A pretreatment multiparametric MRI-based radiomics-clinical machine learning model for predicting radiation-induced temporal lobe injury in patients with nasopharyngeal carcinoma. *Head Neck-J Sci Spec*. (2024) 46:2132–44. doi: 10.1002/hed.27830
23. Liao LJ, Tsai CC, Li PY, Lee CY, Lin SR, Lai WY, et al. Characterization of unique pattern of immune cell profile in patients with nasopharyngeal carcinoma through flow cytometry and machine learning. *J Cell Mol Med*. (2024) 28:e18404. doi: 10.1111/jcmm.18404
24. Zhou L, Zheng W, Huang S, Yang X. Integrated radiomics, dose-volume histogram criteria and clinical features for early prediction of saliva amount reduction after radiotherapy in nasopharyngeal cancer patients. *Discov Oncol*. (2022) 13:145. doi: 10.1007/s12672-022-00606-x
25. Sun Y, Tan J, Li C, Yu D, Chen W. Creating an interactive database for nasopharyngeal carcinoma management: applying machine learning to evaluate metastasis and survival. *Front Oncol*. (2024) 14:1456676. doi: 10.3389/fonc.2024.1456676
26. Pei W, Wang C, Liao H, Chen X, Wei Y, Huang X, et al. MRI-based random survival Forest model improves prediction of progression-free survival to induction chemotherapy plus concurrent Chemoradiotherapy in Locoregionally Advanced nasopharyngeal carcinoma. *BMC Cancer*. (2022) 22:739. doi: 10.1186/s12885-022-09832-6
27. Morgan RD, Youssi BW, Cacao R, Hernandez C, Nagy L. Random forest prognostication of survival and 6-month outcome in pediatric patients following decompressive craniectomy for traumatic brain injury. *World Neurosurg*. (2025) 193:861–7. doi: 10.1016/j.wneu.2024.10.075
28. Li Y, Huang Z, Zeng X, Pan Y, Wu L, Wang J, et al. Early recurrence as a pivotal event in nasopharyngeal carcinoma: identifying predictors and key molecular signals for survivors. *Head Face Med*. (2024) 20:55. doi: 10.1186/s13005-024-00457-7
29. Ding T, Zhang Y, Ren Z, Cong Y, Long J, Peng M, et al. EBV-associated hub genes as potential biomarkers for predicting the prognosis of nasopharyngeal carcinoma. *Viruses*. (2023) 15(9):1915. doi: 10.3390/v15091915
30. Chen X, Li Y, Li X, Cao X, Xiang Y, Xia W, et al. An interpretable machine learning prognostic system for locoregionally advanced nasopharyngeal carcinoma based on tumor burden features. *Oral Oncol*. (2021) 118:105335. doi: 10.1016/j.jor.2021.105335
31. Wan XB, Zhao Y, Fan XJ, Cai HM, Zhang Y, Chen MY, et al. Molecular prognostic prediction for locally advanced nasopharyngeal carcinoma by support vector machine integrated approach. *PloS One*. (2012) 7:e31989. doi: 10.1371/journal.pone.0031989
32. Yin X, Sha H, Cao X, Ge X, Li T, Cui Y, et al. Tumor habitat-derived radiomics features in pretreatment CT scans for predicting concurrent chemoradiotherapy responses in nasopharyngeal carcinoma: a retrospective study. *Quant Imaging Med Surg*. (2025) 15:2917–28. doi: 10.21037/qims-24-1642
33. Jian L, Sheng C, Liu H, Li H, Hu P, Ai Z, et al. Early prediction of progression-free survival of patients with locally advanced nasopharyngeal carcinoma using multiparametric MRI radiomics. *BMC Cancer*. (2025) 25:519. doi: 10.1186/s12885-025-13899-2
34. Su K, Liu X, Zeng YC, Xu J, Li H, Wang H, et al. Machine learning radiomics for predicting response to MR-guided radiotherapy in unresectable hepatocellular carcinoma: A multicenter cohort study. *J Hepatocell Carcinoma*. (2025) 12:933–47. doi: 10.2147/JHC.S521378
35. Peng G, Chi H, Gao X, Zhang J, Song G, Xie X, et al. Identification and validation of neurotrophic factor-related genes signature in HNSCC to predict survival and immune landscapes. *Front Genet*. (2022) 13:1010044. doi: 10.3389/fgene.2022.1010044