Boosting Generalizability in NPC ART Prediction via Multi-Omics Feature Mapping

Jiabao Sheng[1], Zhe Li[2,3] , Jiang Zhang[1], Saikit LAM[4], Zhi Chen[1], Lei Xing[5], and Jing Cai(✉)[1]

[1] Dept. of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong SAR
jing.cai@polyu.edu.hk
[2] Dept. of Electrical and Electronic Engineering, The Hong Kong Polytechnic University, Hong Kong SAR
[3] Microsoft Research Asia, Beijing, China
[4] Dept. of Biomedical Engineering, The Hong Kong Polytechnic University, Hong Kong SAR
[5] Dept. of Radiation Oncology, Stanford University, USA

Abstract. Adaptive radiotherapy (ART) improves treatment precision by adapting to anatomical changes, but its clinical adoption is limited by high costs, patient burden, and institutional variability. To address this, we propose a robust multi-omics nomogram for predicting ART eligibility in nasopharyngeal carcinoma (NPC) patients by integrating multi-modality Genomap signatures with clinical factors. Using retrospective data from 311 patients at Queen Elizabeth Hospital (training set) and 192 patients at Queen Mary Hospital (external test set), we extracted 7,956 radiomics features from six regions-of-interest (ROIs) across contrast-enhanced computed tomography (CECT), magnetic resonance imaging (MRI), and dose modalities, alongside 132 geometric features capturing spatial relationships between ROIs. Feature selection via LASSO identified 35 radiomic, 8 dosiomic, and 4 geometric features for analysis. The Genomap model achieved an accuracy of 80% and an AUC of 90% across modalities, while the integrated nomogram demonstrated superior performance with 88% accuracy and 96% AUC. Our results show that Genomap ensures generalizability and robustness, providing a reliable tool for personalized ART planning in NPC patients.

Keywords: Adaptive radiotherapy (ART) · Multi-Omics Integration · Nasopharyngeal Carcinoma (NPC) · Genomap Signatures · Clinical Nomogram

1 Introduction

Nasopharyngeal carcinoma (NPC) is a distinctive type of cancer originating in the nasopharynx, with high incidence rates in East and Southeast Asia, particularly southern China [3,13]. This geographic disparity is attributed to genetic predispositions [1], environmental factors [18], and the prevalent Epstein-Barr

virus (EBV), which are closely associated with NPC pathogenesis [11,16]. The standard treatment for NPC involves a multidisciplinary approach, including radiation therapy, chemotherapy, and occasionally surgery [9,14].

Radiation therapy is crucial for managing NPC due to its radiosensitivity. However, traditional radiation plans, based solely on pre-treatment imaging, often overlook intra-treatment anatomical changes such as tumor shrinkage and weight loss, which can significantly alter tumor positioning and morphology [5]. ART marks a significant advancement in NPC treatment, featuring dynamic adjustments to radiation plans in response to observed changes during therapy [10]. In clinical practice, ART relies on regular imaging updates (e.g., daily cone beam computed tomography or weekly magnetic resonance imaging), but these imaging technologies have limitations such as high cost, long processing time, and high burden on patients. By accurately predicting which patients tend to respond favorably to adaptive strategies, healthcare providers can tailor treatments more effectively, improving therapeutic outcomes and reducing unnecessary radiation exposure.

Radiomics, an emerging field in medical imaging, extracts quantitative features from radiological images to reveal disease characteristics imperceptible to the naked eye, playing a crucial role in predicting ART tasks by providing detailed textural, shape, and intensity data [2,4]. Recent studies have shown promising results using radiomics for predicting ART in NPC [8]. More recently, multi-omics—which integrates radiomics, dosiomics, and contouromics from multi-modality radiotherapy data—has gained attention for its potential to enhance ART prediction models. For instance, Zhang *et al.* [17] explored the impact of multi-omics on ART prediction models, improving the AUC to 70%-80% by addressing data imbalance issues. However, the application of multi-omics in ART prediction faces significant challenges. Variability in scanners, imaging protocols, dose calculation algorithms, and structure delineations across different institutions can undermine the consistency and reproducibility of multi-omics features, resulting in prediction models that are not generally applicable across hospitals [7,15]. This highlights the need for specialized modeling techniques to enhance the inter-institutional generalizability of omics-based ART prediction models.

Genomap [6] is a novel modeling technique that can enhance model generalizability by re-representing the multi-omics data in a feature map manner. In multi-omics analysis, it can control variables through the normalized partial correlation coefficient, explore stronger dependencies between features, and obtain the optimal order of mapping features to a graph structure based on the Gromov-Wasserstein calculation. Each sample is then represented in such a map, which reduces the structural differences in the representation of data from different institutions.

The contributions of this paper are summarized as follows:

- This study aims to develop a highly generalizable multi-omics model for predicting the efficacy of ART in NPC patients, utilizing the map-constructing

module of Genomap [6]. The model was validated both internally and externally using data from two distinct institutions.
- In this study, the map-constructing module of Genomap was employed to transform multi-omics data, including CECT, multi-parameter MRI, dose map, and contour structure, into a unified feature map. Subsequently, a convolutional neural network (CNN) was used to recalibrate the features, generating predictive factors that integrate different modalities with clinical factors for nomogram construction.
- The generalizability of the transformed multi-omics data, as output by the Genomap module, was evaluated and compared with that of conventionally adopted machine learning models.

2 Methodology

2.1 Overview

This work proposes a novel approach that integrates the construction of a feature map and the probability for prediction. As illustrated in Fig. 1, the process begins with the construction of an interaction matrix of features after feature extraction and selection, which is then using the Gromov-Wasserstein (GW) to calculate discrepancy. Subsequently, a CNN model learns a discriminative feature from this feature map. The feature embeddings extracted from the final fully connected (FC) layer are then fed into a logistic regression model as input features for prediction. To enhance interpretability, prediction probabilities are visualized using a Nomogram.

2.2 Normalized Partial Correlation Coefficient Matrix

Let $\boldsymbol{X} \in \mathbb{R}^{m \times n}$ represent the feature matrix, where m denotes the number of patients and n denotes the number of imaging features. Each element x_{pi} corresponds to the i-th imaging feature of the p-th patient. We represent the i-th feature across all patients as a column vector $\boldsymbol{X}_i = (\boldsymbol{x}_{1i}, \ldots, \boldsymbol{x}_{mi})^\top \in \mathbb{R}^m$. Our objective is to learn feature interactions through entropy maximization constrained by pairwise statistical dependencies.

The covariance matrix $\boldsymbol{\Sigma}$ is derived from the feature matrix $\boldsymbol{X}$. Specifically, the covariance matrix describes the linear relationships between features. For the feature matrix $\boldsymbol{X}$, the element $\boldsymbol{\sigma}_{ij}$ of the covariance matrix represents the covariance between the i-th and j-th features, calculated by the formula:

$$\boldsymbol{\sigma}_{ij} = \frac{1}{m-1} \sum_{p=1}^{m} (\boldsymbol{x}_{pi} - \boldsymbol{\mu}_i)(\boldsymbol{x}_{pj} - \boldsymbol{\mu}_j) \tag{1}$$

where $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ are the mean values of the i-th and j-th features, respectively.

The normalized partial correlation coefficient matrix is derived from the precision matrix $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}$, which encodes the conditional dependencies between features. Specifically, we compute the normalized partial correlations:

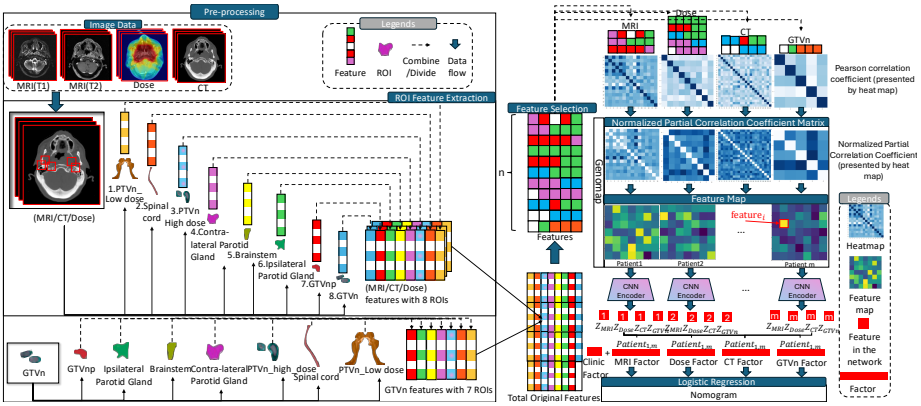

Fig. 1. Model Workflow Overview. This model includes three key steps: 1) Preprocessing: radiomic features are extracted from image data within ROIs, and geometric relationships between GTVn and other ROIs are calculated. Feature selection is subsequently performed, resulting in the retention of 6 ROIs from the initial set of 8. 2) Feature transformation: selected features are processed using Genomap's modules for normalization and mapping. 3) Nomogram construction: predictive factors from CNN feature maps and clinical data are combined to train and generate a nomogram for prediction.

$$\rho_{ij} = \begin{cases} -\dfrac{\boldsymbol{\Omega}_{ij}}{\sqrt{\boldsymbol{\Omega}_{ii}\boldsymbol{\Omega}_{jj}}} & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \tag{2}$$

These normalized partial correlation coefficients reveal the conditional dependencies between features when other variables are controlled.

2.3 Genomap Construction

We followed the Genomap construction procedure proposed by Islam *et al.* [6] to compute the Gromov-Wasserstein discrepancy between the normalized partial correlation coefficient matrix $\boldsymbol{C} \in \mathbb{R}^{n\times n}$ and the distance matrix $\overline{\boldsymbol{C}} \in \mathbb{R}^{n\times n}$, n denotes the embedding dimension. This calculation enables us to determine the optimal arrangement of features, thereby facilitating the creation of the final feature map. We define the Gromov-Wasserstein discrepancy between matrices $\boldsymbol{C}$ and $\overline{\boldsymbol{C}}$ as follows:

$$\begin{aligned} \mathcal{E}_{\boldsymbol{C},\overline{\boldsymbol{C}}}(\boldsymbol{T}) &\overset{\text{def.}}{=} \sum_{i,j,k,\ell} L\left(\boldsymbol{C}_{i,k}, \overline{\boldsymbol{C}}_{j,\ell}\right) \boldsymbol{T}_{i,j}\boldsymbol{T}_{k,\ell}, \\ GW(\boldsymbol{C}, \overline{\boldsymbol{C}}, \boldsymbol{u}, \boldsymbol{v}) &\overset{\text{def.}}{=} \min_{\boldsymbol{T}\in\mathcal{C}_{u\times v}} \mathcal{E}_{\boldsymbol{C},\overline{\boldsymbol{C}}}(\boldsymbol{T}), \end{aligned} \tag{3}$$

where $\boldsymbol{T}$ denotes the coupling between $\boldsymbol{C}$ and $\overline{\boldsymbol{C}}$, $\boldsymbol{u}$ and $\boldsymbol{v}$ encapsulate the relative importance of positions in the feature and feature maps.

To quantify the discrepancy between two distributions, we adopt the Kullback–Leibler (KL) divergence as the loss function $\mathcal{L}$. For inputs a and b, the KL divergence is defined as

$$\mathcal{L}(a,b) = \mathrm{KL}(a \,\|\, b) \overset{\text{def.}}{=} a \log\left(\frac{a}{b}\right) - a + b. \tag{4}$$

In this study, we introduced $\boldsymbol{C}$ and $\overline{\boldsymbol{C}}$ to this loss:

$$\mathcal{L}(\boldsymbol{C}, \overline{\boldsymbol{C}}) \overset{\text{def.}}{=} \left(\mathcal{L}\left(\boldsymbol{C}_{i,k}, \overline{\boldsymbol{C}}_{j,\ell}\right)\right)_{i,j,k,\ell}. \tag{5}$$

We have $\mathcal{E}_{\boldsymbol{C},\overline{\boldsymbol{C}}}(\boldsymbol{T}) = \langle \mathcal{L}(\boldsymbol{C}, \overline{\boldsymbol{C}}) \otimes \boldsymbol{T}, \boldsymbol{T} \rangle$, where $\otimes$ denotes the tensor-matrix multiplication as follows:

$$\mathcal{L} \otimes \boldsymbol{T} \overset{\text{def.}}{=} \left(\sum_{k,\ell} \mathcal{L}_{i,j,k,\ell} \boldsymbol{T}_{k,\ell}\right)_{i,j}. \tag{6}$$

2.4 CNN Feature Extraction and Nomogram Construction

Based on the optimal arrangement of the feature scheme derived from matrix $\boldsymbol{T}$, a feature vector $\boldsymbol{F}_i \in \mathbb{R}^n$ is generated for each sample, where n denotes the dimension of the feature vector. By stacking all samples, we obtain the feature matrix $\boldsymbol{F} \in \mathbb{R}^{m \times n}$, where m is the number of samples. The matrix $\boldsymbol{F}$ is subsequently fed into the CNN model. The weight matrix $\boldsymbol{W} \in \mathbb{R}^{n \times k}$ represents the weights of the FC layer in a CNN, where k is the number of neurons in the layer.

To quantify the contribution of each feature across all samples to the output of the FC layer, matrix multiplication is performed as $\boldsymbol{Z} = \boldsymbol{F} \cdot \boldsymbol{W}$, where $\boldsymbol{Z} \in \mathbb{R}^{m \times k}$ represents the transformed feature representations for all samples, and k denotes the number of neurons in the FC layer.

We propose a nomogram derived from a logistic regression model to predict binary outcomes:

$$\mathrm{logit}(p) = \ln\left(\frac{p}{1-p}\right), \tag{7}$$

where p denotes the probability of the positive class. This probability is computed using the logistic regression function:

$$p = \frac{1}{1 + \exp\left(-(\boldsymbol{Z} \cdot \boldsymbol{\beta} + \boldsymbol{\beta}_0)\right)}, \tag{8}$$

where $\boldsymbol{\beta}$ is the weight vector, and $\boldsymbol{\beta}_0$ is the bias.

In the constructed nomogram, each input factor $\boldsymbol{Z}_j$ is represented by a corresponding axis that displays its value range. The total score is calculated by aggregating the weighted contributions of all features, with each contribution determined by its respective weight $\boldsymbol{\beta}_j$. The nomogram thus encapsulates the

Table 1. The number of selected features from each data modality and ROI.

Data Modality	ROI	Feature Number
MRI (T1)	Ipsilateral Parotid	1
	GTVnp	2
	Contra Parotid	1
MRI (T2)	Ipsilateral Parotid	3
	GTVnp	4
	Contra Parotid	3
CT	Ipsilateral Parotid	7
	GTVnp	3
	GTVn	4
	Contra Parotid	7
Dose	Ipsilateral Parotid	2
	GTVnp	1
	GTVn	5
Contouromics	Spinal Cord	3
	Ipsilateral Parotid	1

contributions of each feature in $\boldsymbol{Z}$ to the final prediction probability. Each feature's influence is scaled by its respective weight $\boldsymbol{\beta}$, providing a visual and interpretable summary of the prediction model. This intuitive representation aids in clinical decision-making by translating the logistic model's outputs into a straightforward, visual format.

3 Experiments and Results

3.1 Datasets

Data source institutions: This retrospective study included 503 NPC patients treated with radiotherapy from Queen Elizabeth Hospital (QEH) for model training and Queen Mary Hospital (QMH) for external validation. A total of 311 patients treated at QEH from 2012 to 2015 and 192 patients treated at QMH from 2012 to 2020 were recruited. Radiation oncologists reviewed clinical records to determine whether each patient received ART. Among the 311 and 192 patients from QEH and QMH, 111 and 27 received ART, respectively.

Multi-omics: Data collection and feature extraction were performed as follows: 1) CECT images in DICOM format and MR images were collected from QEH and QMH. ROIs including GTVp, GTVn, parotid glands, brainstem, and spinal cord were delineated based on NCCN guidelines [12]. 2) Radiomic features were extracted from CECT and MR images for 6 ROIs, resulting in 6,348 features per image, including 84 morphologic features. 3) Dosiomic features were extracted based on DVH and spatial dose distribution within 6 VOIs, yielding 1,608 features. 4) Contouromics features(132 in total) were calculated to represent the geometric relationships between four pairs of OIs based on OVH and POV descriptors.

Selected features: After feature selection, 31 radiomic features from four ROIs were retained (IpsiParotid, ContraParotid, GTVp, and GTVn), 8 dosiomic features from three ROIs were selected (IpsiParotid, GTVp, and GTVn), and 4

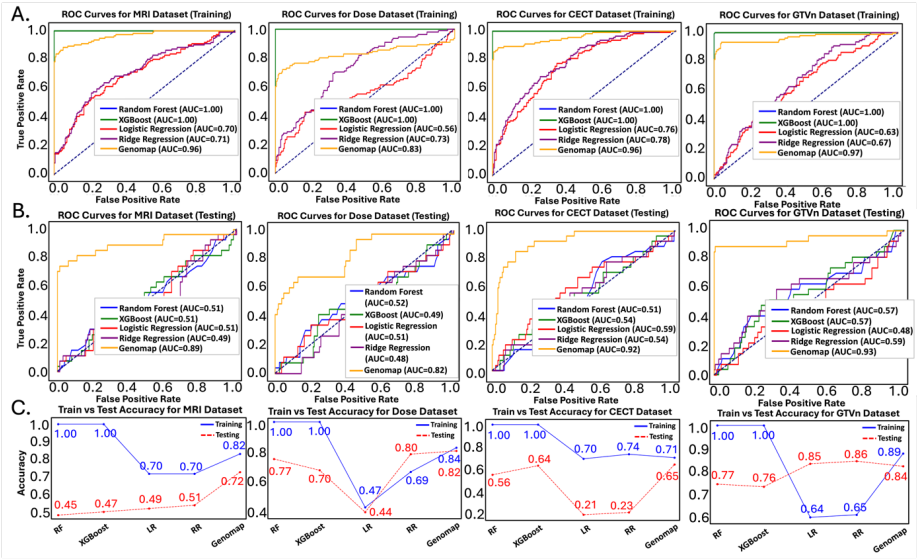

Fig. 2. Model Generalizability and Performance. (A) AUC(%) results for QEH data(training set). (B) AUC(%) results for QMH data (test set). (C) ART prediction accuracy of single-modality features and four baseline models on QEH(training) and QMH(test) data.

geometric features from GTVn-SpinalCord and GTVn-IpsiParotid remained, as detailed in Table 1.

3.2 Comparison Results

During training (Fig. 2A), the random forest and XGBoost models overfitted the dataset. The remaining two baseline models, though avoiding overfitting, achieved AUC values of 56%–78%. Conversely, Genomap averted overfitting and maintained stable AUC scores of 83%–97% across all datasets. In testing (Fig. 2B), Genomap's AUC scores remained robust at 82%–93%, while other models exhibited AUC values around 50%. Consistent trends emerged in the accuracy (ACC) results in Fig. 2C. Although Genomap did not consistently outperform all models on every dataset, it achieved the highest training and testing accuracy on MRI and dose data, with accuracy on the remaining datasets closely matching the best-performing models.

3.3 Ablation Study

The training data were sourced from QEH, while the testing data were obtained from QMH. Genomap demonstrated remarkable performance and generalizability across different data sources. During training Fig. 2A, Genomap

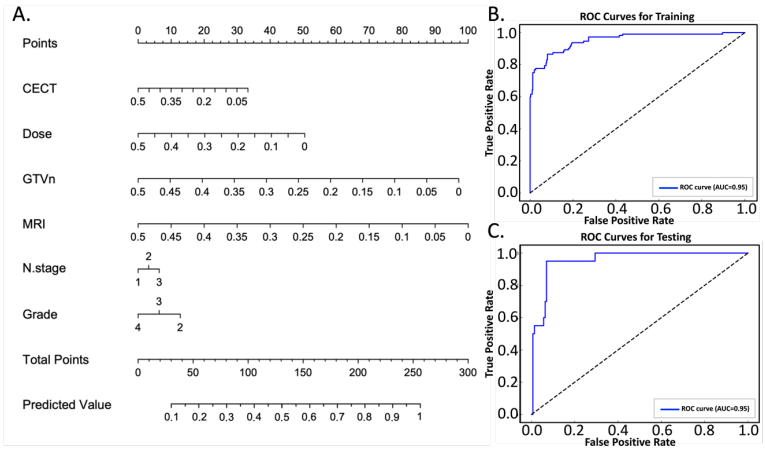

Fig. 3. Nomogram and Model Validation. (A) The nomogram shows the contribution of six predictors (CECT, Dose, GTVn, MRI, N. stage, and grade) to the total points, with each feature mapped to a probability score (0–1) indicating ART likelihood. Model trained on QEH data and tested on QMH data. (B)AUC(%) results for QEH data (training set). (C) AUC (%) results for QMH data (test set).

achieved an AUC of approximately 96% and maintained high accuracy with values ranging from 71% on CECT to 80%–90% on other data modalities. In the testing Fig. 2B, where data were from a different institution, Genomap consistently achieved approximately 90% AUC across all four data modalities and maintained high accuracy, peaking at 89% on GTVn. The minimal differences between training and testing performance indicate that Genomap effectively bridged the gap between data from different institutions. Genomap's consistent performance across both training and testing phases, particularly with diverse data sources, underscores its exceptional generalization ability. The map-constructing module of Genomap effectively mitigates the impact of differences arising from distinct data sources, making it a robust solution for predictive modeling in this context.

3.4 Nomogram Results

Fig. 3 presents the nomogram developed for predicting the probability of requiring ART in patients, based on Genomap factors and clinical factors. Fig. 3B presents the training results obtained with the QEH dataset, yielding an AUC of 95% and an accuracy of 97%. Similarly, Fig. 3C illustrates the testing results using the QMH dataset, achieving an AUC of 96% and an accuracy of 88%.

4 Conclusion

Model generalizability across diverse clinical settings is crucial, as variations in imaging protocols, scanner types, and patient characteristics can undermine prediction accuracy. This study addresses these challenges using a normalized partial correlation coefficient matrix and Genomap's map-constructing module, outperforming benchmark methods with superior generalizability and predictive accuracy. By integrating Genomap-derived multi-modality factors with clinical factors, we developed a nomogram providing individualized ART probability estimates for patients. This tool aids clinicians in personalized treatment planning and optimized radiation strategies. Future research will expand the dataset to enhance model adaptability across different clinical environments.

Acknowledgments. This work was in part supported by the International Collaborative Research Fellowship and the Research Student Attachment Programme of HKPolyU.

Disclosure of Interests. The authors have no competing interests to declare.

References

1. Abubakar, M.B., Gan, S.H.: Genomic and epigenetic landscapes of nasopharyngeal carcinoma. An Evidence-Based Approach to the Management of Nasopharyngeal Cancer: From Basic Science to Clinical Presentation and Treatment p. 31 (2020)
2. Aerts, H.J., Velazquez, E.R., Leijenaar, R.T., Parmar, C., Grossmann, P., Carvalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., et al.: Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nature communications **5**(1), 4006 (2014)
3. Chua, M.L., Wee, J.T., Hui, E.P., Chan, A.T.: Nasopharyngeal carcinoma. The Lancet **387**(10022), 1012–1024 (2016)
4. Gillies, R.J., Kinahan, P.E., Hricak, H.: Radiomics: images are more than pictures, they are data. Radiology **278**(2), 563–577 (2016)
5. Glide-Hurst, C.K., Lee, P., Yock, A.D., Olsen, J.R., Cao, M., Siddiqui, F., Parker, W., Doemer, A., Rong, Y., Kishan, A.U., et al.: Adaptive radiation therapy (art) strategies and technical considerations: a state of the art review from nrg oncology. International Journal of Radiation Oncology* Biology* Physics **109**(4), 1054–1075 (2021)
6. Islam, M.T., Xing, L.: Cartography of genomic interactions enables deep analysis of single-cell expression data. Nature Communications **14**(1), 679 (2023)
7. Kumar, V., Gu, Y., Basu, S., Berglund, A., Eschrich, S.A., Schabath, M.B., Forster, K., Aerts, H.J., Dekker, A., Fenstermacher, D., et al.: Radiomics: the process and the challenges. Magnetic resonance imaging **30**(9), 1234–1248 (2012)
8. Lam, S.K., Zhang, Y., Zhang, J., Li, B., Sun, J.C., Liu, C.Y.T., Chou, P.H., Teng, X., Ma, Z.R., Ni, R.Y., et al.: Multi-organ omics-based prediction for adaptive radiation therapy eligibility in nasopharyngeal carcinoma patients undergoing concurrent chemoradiotherapy. Frontiers in oncology **11**, 792024 (2022)

9. Lee, A.W., Ng, W., Chan, Y., Sze, H., Chan, C., Lam, T.: The battle against nasopharyngeal cancer. Radiotherapy and Oncology **104**(3), 272–278 (2012)
10. Li, J., Xu, Z., Pilar, A., O'Sullivan, B., Huang, S.H.: Adaptive radiotherapy for nasopharyngeal carcinoma. Annals of Nasopharynx Cancer **4** (2020)
11. Lin, J.C., Wang, W.Y., Chen, K.Y., Wei, Y.H., Liang, W.M., Jan, J.S., Jiang, R.S.: Quantification of plasma epstein–barr virus dna in patients with advanced nasopharyngeal carcinoma. New England Journal of Medicine **350**(24), 2461–2470 (2004)
12. Pfister, D.G., Spencer, S., Adelstein, D., Adkins, D., Anzai, Y., Brizel, D.M., Bruce, J.Y., Busse, P.M., Caudell, J.J., Cmelak, A.J., et al.: Head and neck cancers, version 2.2020, nccn clinical practice guidelines in oncology. Journal of the National Comprehensive Cancer Network **18**(7), 873–898 (2020)
13. Su, Z.Y., Siak, P.Y., Leong, C.O., Cheah, S.C.: The role of epstein–barr virus in nasopharyngeal carcinoma. Frontiers in microbiology **14**, 1116143 (2023)
14. Tham, I.W.K., Lu, J.J.: Controversies and challenges in the current management of nasopharyngeal cancer. Expert Review of Anticancer Therapy **10**(9), 1439–1450 (2010)
15. Van Timmeren, J.E., Cester, D., Tanadini-Lang, S., Alkadhi, H., Baessler, B.: Radiomics in medical imaging—"how-to" guide and critical reflection. Insights into imaging **11**(1), 91 (2020)
16. Young, L.S., Murray, P.G.: Epstein–barr virus and oncogenesis: from latent genes to tumours. Oncogene **22**(33), 5108–5121 (2003)
17. Zhang, J., Lam, S.K., Teng, X., Ma, Z., Han, X., Zhang, Y., Cheung, A.L.Y., Chau, T.C., Ng, S.C.Y., Lee, F.K.H., et al.: Radiomic feature repeatability and its impact on prognostic model generalizability: A multi-institutional study on nasopharyngeal carcinoma patients. Radiotherapy and Oncology **183**, 109578 (2023)
18. Zheng, Y., Tuppin, P., Hubert, A., Jeannel, D., Pan, Y., Zeng, Y., De Thé, G.: Environmental and dietary risk factors for nasopharyngeal carcinoma: a case-control study in zangwu county, guangxi, china. British journal of cancer **69**(3), 508–514 (1994)