European Journal of Nuclear Medicine and Molecular Imaging
https://doi.org/10.1007/s00259-022-05793-x

ORIGINAL ARTICLE

Deep learning signatures reveal multiscale intratumor heterogeneity associated with biological functions and survival in recurrent nasopharyngeal carcinoma

Xun Zhao[1,2] · Yu-Jing Liang[3,4] · Xu Zhang[4,5] · Dong-Xiang Wen[3,4] · Wei Fan[4,5] · Lin-Quan Tang[3,4] · Di Dong[1,2] · Jie Tian[1,2] · Hai-Qiang Mai[3,4]

Received: 14 January 2022 / Accepted: 3 April 2022

Abstract

Purpose How to discriminate different risks of recurrent nasopharyngeal carcinoma (rNPC) patients and guide individual treatment has become of great importance. This study aimed to explore the associations between deep learning signatures and biological functions as well as survival in (rNPC) patients.

Methods A total of 420 rNPC patients with PET/CT imaging and follow-up of overall survival (OS) were retrospectively enrolled. All patients were randomly divided into a training set ($n=269$) and test set ($n=151$) with a 6:4 ratio. We constructed multi-modality deep learning signatures from PET and CT images with a light-weighted deep convolutional neural network EfficienetNet-lite0 and survival loss DeepSurvLoss. An integrated nomogram was constructed incorporating clinical factors and deep learning signatures from PET/CT. Clinical nomogram and single-modality deep learning nomograms were also built for comparison. Furthermore, the association between biological functions and survival risks generated from an integrated nomogram was analyzed by RNA sequencing (RNA-seq).

Results The C-index of the integrated nomogram incorporating age, rT-stage, and deep learning PET/CT signature was 0.741 (95% CI: 0.688–0.794) in the training set and 0.732 (95% CI: 0.679–0.785) in the test set. The nomogram stratified patients into two groups with high risk and low risk in both the training set and test set with hazard ratios (HR) of 4.56 (95% CI: 2.80–7.42, $p<0.001$) and 4.05 (95% CI: 2.21–7.43, $p<0.001$), respectively. The C-index of the integrated nomogram was significantly higher than the clinical nomogram and single-modality nomograms. When stratified by sex, N-stage, or EBV DNA, risk prediction of our integrated nomogram was valid in all patient subgroups. Further subgroup analysis showed that patients with a low-risk could benefit from surgery and re-irradiation, while there was no difference in survival rates between patients treated by chemotherapy in the high-risk and low-risk groups. RNA sequencing (RNA-seq) of data further explored the mechanism of high- and low-risk patients from the genetic and molecular level.

Conclusion Our study demonstrated that PET/CT-based deep learning signatures showed satisfactory prognostic predictive performance in rNPC patients. The nomogram incorporating deep learning signatures successfully divided patients into different risks and had great potential to guide individual treatment: patients with a low-risk were supposed to be treated with surgery and re-irradiation, while for high-risk patients, the application of palliative chemotherapy may be sufficient.

Keywords Recurrent nasopharyngeal carcinoma · Survival analysis · Radiomics · Deep learning

Introduction

Nasopharyngeal carcinoma (NPC) is an endemic disease in southeastern Asia, southern China, and north Africa, accounting for 0.7% (29,079) of new cases in 2018 [1]. A relatively satisfactory locoregional control rate was obtained in the intensity-modulated radiation therapy (IMRT) era, with local failure rates of 5–14%. However, the local control for T4 disease was still challenging with

This article is part of the Topical Collection on Oncology—Head and Neck.

Xun Zhao, Yu-Jing Liang, Xu Zhang, Dong-Xiang Wen and Wei Fan contributed equally to this work.

Extended author information available on the last page of the article

Published online: 26 April 2022

rates ranging from 74 to 80% [2]. Re-irradiation, surgery, or systemic chemotherapy has been used as salvage treatment methods in recurrent NPC patients (rNPC). But neither local salvage nor the patient's quality of life obtained satisfactory results because of the delicate balance between risk and benefits in these post-radiotherapy patients [3]. The efficacy and adverse effects of treatment modalities above have not been compared in high evidence randomized controlled trials. In a case–control study involving patients with rT3–4 N0–1, the 5-year OS rates were similar between patients treated by chemotherapy alone or re-irradiation [4]. Additionally, a previous study demonstrated that patients with low and intermediate-risk instead of high-risk could benefit from re-irradiation or surgery [5]. The tradeoff decisions between the chance of salvage and the risk of life-threatening adverse effects is a daunting dilemma for the oncologist and recurrent patients since severe re-irradiation-related adverse effects including massive hemorrhage or necrosis are frequent, accounting for a high proportion of mortality at 34.7% [6]. Therefore, how to discriminate high-risk from low-risk recurrent patients and guide individual treatment becomes of great importance.

Sun et al. established a nomogram incorporating age, hypertension, rT-stage, and EBV DNA level to stratify different risk groups based on the total score derived from the nomogram model [5]. The model symbolized the first step to predict OS and guide individualized treatment, but the concordance index (C-index) of 0.687 was not satisfactory enough.

Recently, deep learning which integrates parameters derived from medical imaging like computed tomography (CT), magnetic resonance imaging (MRI) or positron emission computer tomography (PET), and genomic data has become a promising field in oncology, providing important supplement information on cancer physiology [7–12]. Peng et al. reported that PET/CT-based deep learning signatures could be used as a reliable and useful prognostic prediction tool that may help guide individualized induction chemotherapy in NPC patients [13]. Zhong and colleagues demonstrated that a nomogram incorporating deep learning based on pre-treatment MRI images showed satisfactory predictive performance in the prognosis of patients with T3N1M0 treated with different regimens and promoted personalized treatment of NPC [14]. Advances in deep learning provide a simple, effective, and reliable means to identify certain characteristics as well as molecular subgroups and divide different risk groups [7]. Clinical data incorporating deep learning signatures have consistently predicted excellent performance in risk stratification and prognosis in patients with lung cancer, glioblastoma, gastric cancer, and breast cancer [15–19]. However, there is a lack of research on the associations between deep learning signatures and prognosis in rNPC patients.

In this study, we explored the associations among deep learning signatures of multi-modality CT images and PET images, biological functions revealed by RNA-seq, as well as survival in rNPC patients.

Materials and methods

Patients

This study was approved by the Clinical Research Ethics Committee in Sun Yat-sen University Cancer Center (SYSUCC), and due to the observational nature of this retrospective study, the requirement for informed consent was waived.

Between October 2010 and May 2019, the medical records of 478 rNPC patients undergoing PET/CT without distant metastases were screened. The eligibility criteria were as follows: (1) pathology confirmed or evidence by PET/CT with a consistent clinical process of recurrence; (2) without distant metastasis; (3) adequate organ function; and (4) available clinical data and PET/CT imaging data. Patients without treatment after diagnosis were excluded. Finally, 420 eligible patients were identified (Fig. 1).

Diagnosis and treatment

A series of evaluations were applied to patients including head and neck physical examination, nasopharyngoscopy with biopsy, nasopharynx and neck MRI, PET/CT, and blood tests.

Salvage treatment included re-irradiation, surgery, or systemic chemotherapy. Surgical approaches for endoscopic nasopharyngectomy were chosen based on the anatomical structure and extension of the tumor lesion by experienced otolaryngologists [20]. Re-irradiation was performed by intensity-modulated radiotherapy (IMRT) technology, with a total prescription dosage of 60–70 Gy, divided into 30–35 fractions, 5 times a week on working days. Common palliative chemotherapy regimens included GP: gemcitabine (1 g/m^2, days 1 and 8) and cisplatin (75–80 mg/m^2, day 1); TPF: docetaxel (60–75 mg/m^2) or paclitaxel (135 mg/m^2) in combination with cisplatin (60 mg/m^2, day 1) and 5-fluorouracil (3–3.75 g/m^2 civ120h); PF: cisplatin (80 mg/m^2, day 1) plus 5-fluorouracil (4 g/m^2 civ120h); and TP: docetaxel (75–80 mg/m^2, day 1) combined with cisplatin (75–80 mg/m^2, day 1). The intravenous chemotherapy

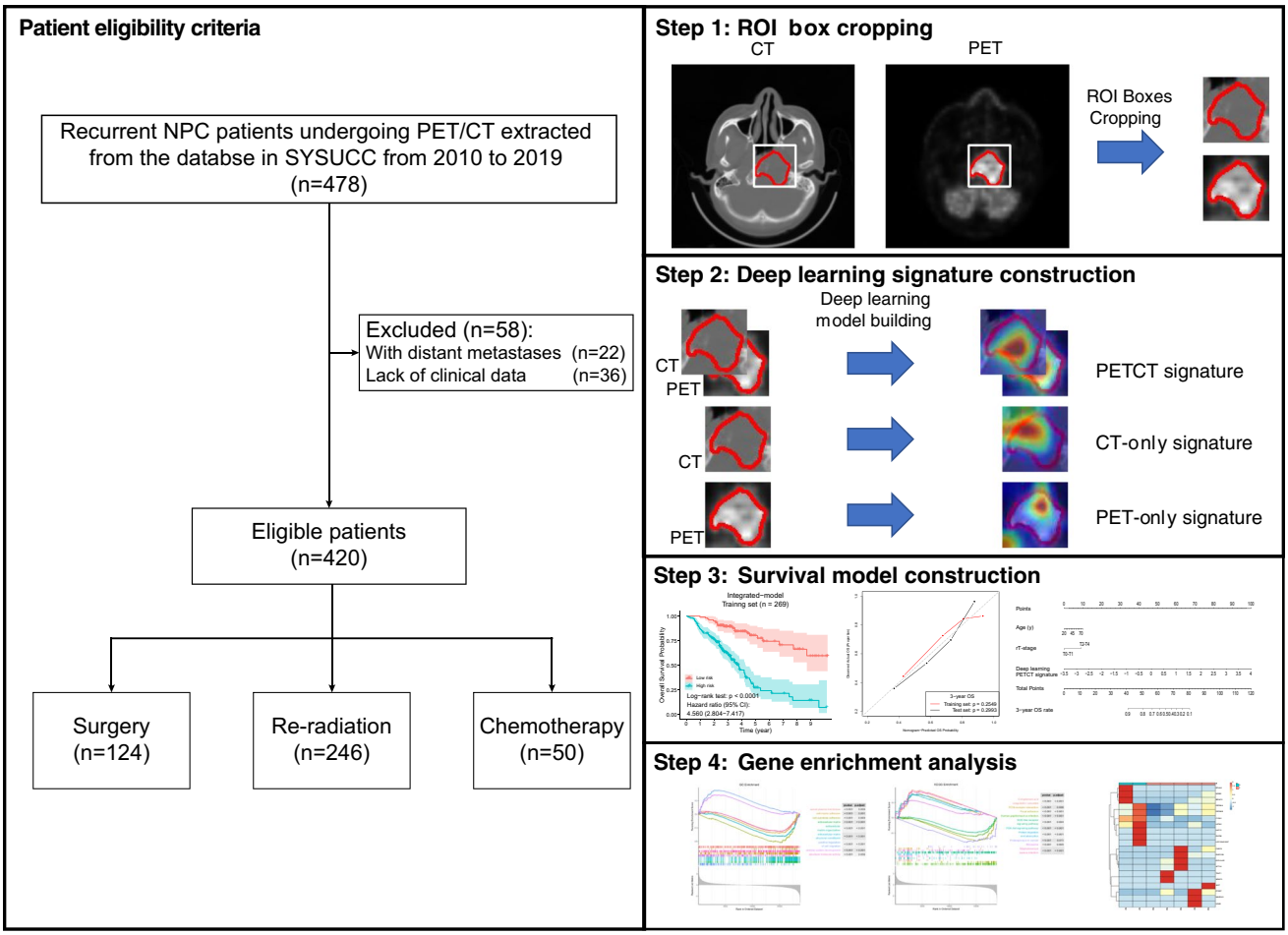

Fig. 1 Patients eligibility criteria and deep learning signature construction workflow. (Left) Patients eligibility criteria. (Right Step 1) Cropping CT and PET images according to region of interest annotated by clinicians. (Right Step 2) Building single-modality and multi-modality deep learning signatures from CT and PET cropped images using deep convolutional networks. (Right Step 3) Constructing and validating clinical nomogram, deep learning nomogram, and integrated nomogram. (Right Step 4) Analyzing gene enrichment between high-risk group and low-risk group

regimens were applied through a peripherally inserted central catheter every 3 weeks. The oral chemotherapy regimen was the administration of capecitabine (1000 mg/m^2) twice daily on days 1–14.

Clinical outcome and follow-up

The primary outcome of our study was overall survival (OS), which was defined as the time from diagnosis to the date of death from any cause. After treatment, patients underwent a series of assessments every 3–6 months thereafter until death.

Deep learning model construction

We interpolated all PET/CT images and cropped regions of interest (ROIs) of each cross-section slice into square boxes. Deep convolutional neural networks (DCNNs) were constructed using these cropped images.

CT images, PET images, and ROIs were concatenated to build a multi-modality model. In DCNN training, we compared 5 classical light-weight structures and selected EfficientNet-lite0 (Supplementary Table S4, Supplementary Table S5) [21]. The DCNN was trained with a survival loss, DeepSurv loss [22]. The neural network structure and loss function are provided in Supplementary Methods. For comparison, we also built two single-modality models, CT-only model and PET-only model, using CT images and PET images, respectively. The details of data preprocessing and deep learning model training are provided in Supplementary Methods.

The output values of deep learning models were denoted as deep learning signatures.

Nomogram construction

A deep learning multi-modality nomogram was built using a PET/CT deep learning signature. We also built a clinical nomogram with clinical factors selected from a univariate analysis and multi-variate analysis (Supplementary Methods).

To integrate deep learning information and clinical information, an integrated nomogram was built using a PET/CT deep learning signature and selected clinical factors.

Patients were divided into low-risk and high-risk groups, according to their nomogram outputs stratified by a threshold learned in the training set (Supplementary Methods). Subgroup analysis was performed based on the results of the high-risk and low-risk split, showing the robustness of nomogram on different patient groups.

RNA extraction, qualification, and sequencing

The association between integrated nomogram predictions and biological functions was analyzed by RNA-seq. TRIzol reagent (Invitrogen, Carlsbad, CA, USA) was used to isolate the total RNA from 7 samples (high-risk, $n = 5$; low-risk, $n = 2$) of rNPC patients after diagnosis. The extraction of total cell RNA was performed with the help of the Novogene Bioinformatics Institute (Beijing, China). The integrity and purity of RNA was assessed with agarose gel electrophoresis and the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA). These libraries were sequenced on the Illumina HiSeq 4000 platform, and 150 bp paired-end reads were generated after the index-coded samples were clustered. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses with a p value of < 0.05 were considered significantly enriched and the top 10 most significantly enriched pathways of differentially expressed genes associated with different risk groups stratified by the integrated nomogram in rNPC patients were presented.

Statistical analysis

The categorical variables were displayed as the number of cases (%) and were compared with the chi-square test or Fisher's exact test. Kaplan–Meier curves were performed to assess the cumulative survival rates compared by the log-rank test. Concordance index (C-index) was used to assess the performance of nomograms. We used calibration curves with the Hosmer–Lemeshow test to evaluate the agreement between prediction survival probability and observed OS proportion. Area under the receiver operating characteristic (ROC) curve (AUC) was used to assess the classification performance of nomograms at given time points.

All analyses were performed by R (http://www.R-project.org, 4.0.2), and a two-tailed p value of < 0.05 was considered statistically significant.

Results

Patient characteristics and clinical nomogram

Five clinical factors associated with OS were provided, including sex, age, rT-stage, N-stage, and EBV DNA. All of them were balanced and distributed between the training set and test set ($p > 0.05$), except rT-stage ($p = 0.010$). The median follow-up time was 36.4 months (IQR 26.77–50.67) in the training set and 35.13 (IQR 22.86–48.33) in the test set. At the end of the follow-up, the overall survival ratio was 59.9% (161/269) in the training set and 53.6% (81/151) in the test set, and there was no significant difference of death events distribution between the two sets ($p = 0.120$) (Table 1).

Using univariable analysis and multi-variable analysis with Akaike information criterion (AIC), we selected age and rT-stage, which were significantly related to OS (Table 2). We used Cox proportional hazard regression (CPH) to build a clinical nomogram with these two clinical factors, and its C-index was 0.673 (95% CI: 0.621–0.726, $p < 0.001$) in the training set and 0.667 (95% CI: 0.607–0.728, $p < 0.001$) in the test set. For additional information from metabolic parameters, 6 standardized uptake values (SUV) were provided, including mean, peak, and maximum SUV of tumor and lymph node. Tumor SUV_{max} and tumor SUV_{peak} were selected by univariable and multi-variable analyses. The clinical nomogram with metabolic parameters had a 0.693 C-index in the training set and 0.682 in the test set, which were better than the clinical nomogram without metabolic parameters (Supplementary Table S2).

Deep learning signature

We built 3 deep learning signatures, including PET/CT signature, CT-only signature, and PET-only signature. All of these three deep learning signatures were significantly related to OS (Table 2). The C-index of PET/CT signature was 0.722 (95% CI: 0.666–0.777, $p < 0.001$) in the training set and 0.722 (95% CI: 0.669–0.775, $p < 0.001$) in the test set. CT-only signature and PET-only signature yielded C-index of 0.660 and 0.706 in the training set and 0.679 and 0.692 in the test set,

Table 1 Baseline clinical information of the training set and test set

Characteristics	Training set ($n=269$)	Test set ($n=151$)	p value
Age (years), median (range)	47 (22–75)	50 (16–69)	0.231
Sex, no. (%)			0.561
Male	198 (73.6)	108 (71.5)	
Female	71 (26.4)	43 (28.5)	
rT-stage, no. (%)			0.010*
T0	34 (12.6)	24 (15.9)	
T1	27 (10.0)	20 (13.2)	
T2	28 (10.4)	8 (5.3)	
T3	108 (40.1)	72 (47.7)	
T4	72 (26.8)	27 (17.9)	
N-stage, no. (%)			0.315
N0	170 (63.2)	88 (58.3)	
N1	78 (29.0)	51 (33.8)	
N2	11 (4.1)	4 (2.6)	
N3	10 (3.7)	8 (5.3)	
EBV-DNA, no. (%)			0.990
Undetectable	121 (45.0)	68 (45.0)	
Detectable	148 (55.0)	83 (55.0)	
Treatment, no. (%)			0.428
Surgery	79 (29.4)	45 (29.8)	
Re-radiation	161 (59.9)	85 (56.3)	
Chemotherapy	29 (10.8)	21 (13.9)	
Tumor SUV, median (range)[1]			
SUV_{max}	8.88 (0.00–66.28)	8.88 (0.00–44.75)	0.838
SUV_{peak}	6.10 (0.00–48.17)	6.10 (0.00–34.92)	0.916
SUV_{mean}	3.99 (0.00–16.47)	3.99 (0.00–11.07)	0.795
Lymph node SUV, median (range)[1]			
SUV_{max}	0.00 (0.00–32.92)	0.00 (0.00–20.62)	0.268
SUV_{peak}	0.00 (0.00–17.21)	0.00 (0.00–16.24)	0.128
SUV_{mean}	0.00 (0.00–19.88)	0.00 (0.00–6.56)	0.550

EBV DNA, Epstein-Barr virus DNA; $SUV_{max/peak/mean}$, maximum/peak/mean standardized uptake value

*denotes p value less than 0.05

[1]For SUV, missing values were filled with median of non-missing values of training set

respectively. The prognostic performance of multi-modality PET/CT signature was significantly better than both single-modality CT-only signature and PET-only signature in the test set, with a higher C-index and better prognosis (Table 3, Supplementary Figure S1).

Integrated nomogram

To integrate clinical and deep learning information, we constructed an integrated nomogram with two selected clinical factors and PET/CT deep learning signatures (Fig. 2A). The C-index of the nomogram was 0.741 (95% CI: 0.688–0.794, $p<0.001$) in the training set and 0.732 (95% CI: 0.679–0.785, $p<0.001$) in the test set. The integrated nomogram brought a significant improvement to the clinical nomogram in C-indices (Supplementary Table S1). The integrated nomogram was slightly better than the deep learning signature (Table 3, Fig. 3). However, after adding metabolic parameters, there were no significant changes in C-indices, resulting in 0.742 and 0.729 in the training set and test set (Supplementary Table S2).

When accessing the classification performance of our nomograms, the integrated nomogram also yielded 0.744 AUC of 3-year OS in the training set and 0.776 AUC in the test set, which was the best compared with the clinical nomogram and deep learning nomogram (Supplementary Figure S2). Moreover, the calibration curve showed that the prediction OS of the integrated nomogram was highly accurately matched with the observed OS with a p value >0.05 in both the training set and test set (Fig. 2B).

Table 2 Univariate hazard ratios of clinical factors, metabolic parameters, and deep learning signatures on the training set and test set

Characteristics		Training set ($n=269$)		Test set ($n=151$)	
		Hazard ratio (95% CI)	p value	Hazard ratio (95% CI)	p value
Sex		0.728 (0.455–1.163)	0.182	0.841 (0.492–1.438)	0.527
Age (years)		1.020 (1.002–1.039)	0.026*	1.024 (1.000–1.049)	0.045*
rT-stage (T0–2 vs. T3–4)		3.609 (2.189–5.950)	< 0.001*	3.533 (1.893–6.592)	< 0.001*
N-stage (N0 vs. other)		0.881 (0.673–1.154)	0.358	1.008 (0.749–1.357)	0.956
EBV-DNA (detectable vs. undetectable)		1.297 (0.884–1.904)	0.183	0.941 (0.589–1.505)	0.800
Tumor SUV	SUV_{max}	1.043 (1.020–1.067)	< 0.001*	1.046 (1.021–1.072)	< 0.001*
	SUV_{peak}	1.068 (1.037–1.100)	< 0.001*	1.066 (1.032–1.101)	< 0.001*
	SUV_{mean}	1.184 (1.084–1.293)	< 0.001*	1.247 (1.120–1.388)	< 0.001*
Lymph node SUV	SUV_{max}	0.972 (0.935–1.010)	0.148	0.970 (0.926–1.015)	0.187
	SUV_{peak}	0.964 (0.907–1.025)	0.240	0.962 (0.900–1.028)	0.246
	SUV_{mean}	0.921 (0.840–1.011)	0.085	0.942 (0.840–1.057)	0.308
Deep learning signature	CT-only	3.185 (2.500–4.058)	< 0.001*	2.223 (1.714–2.883)	< 0.001*
	PET-only	5.494 (3.644–8.283)	< 0.001*	2.237 (1.471–3.401)	< 0.001*
	PETCT	2.751 (2.021–3.744)	< 0.001*	2.458 (1.698–3.559)	< 0.001*

*denotes p value less than 0.05

Table 3 C-indices of clinical, deep learning, and integrated nomograms on the training set and test set

Nomograms		Training set ($n=269$)		Test set ($n=151$)	
		C-index (95% CI)	p value	C-index (95% CI)	p value
Clinical		0.673 (0.621–0.726)	< 0.001*	0.667 (0.607–0.728)	< 0.001*
Deep learning	CT-only	0.660 (0.601–0.719)	< 0.001*	0.679 (0.613–0.745)	< 0.001*
	PET-only	0.706 (0.649–0.762)	< 0.001*	0.692 (0.633–0.752)	< 0.001*
	PETCT	0.722 (0.666–0.777)	< 0.001*	0.722 (0.669–0.775)	< 0.001*
Integrated		0.741 (0.688–0.794)	< 0.001*	0.732 (0.679–0.785)	< 0.001*

*denotes p value less than 0.05

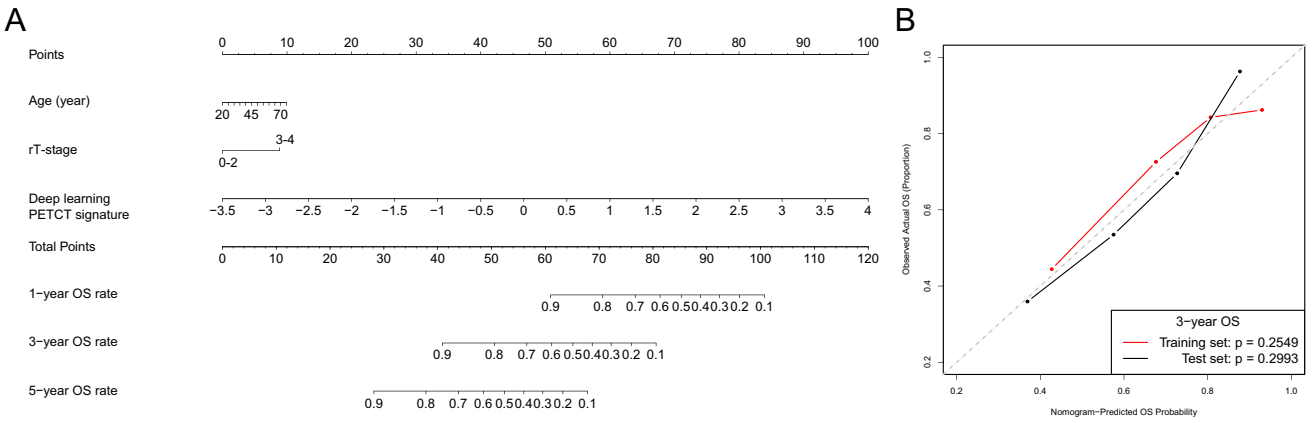

Fig. 2 Integrated nomogram of clinical factors and PET/CT deep learning signatures and its calibration curves. **A** An integrated nomogram was built from clinical factors and PET/CT deep learning signatures to predict overall survival. **B** The 3-year overall survival rate calibration curves of integrated nomogram in training set and test set. The p value of calibration was calculated by Hosmer–Lemeshow goodness-of-fit tests

A cutoff value of 0.304 was learned from nomogram outputs in the training set, corresponding to a 50.76 total score in the integrated nomogram. This threshold could divide patients into two groups as high-risk and low-risk. In the training set, the high-risk group had a higher 3-year OS rate compared with the low-risk group (89.2%

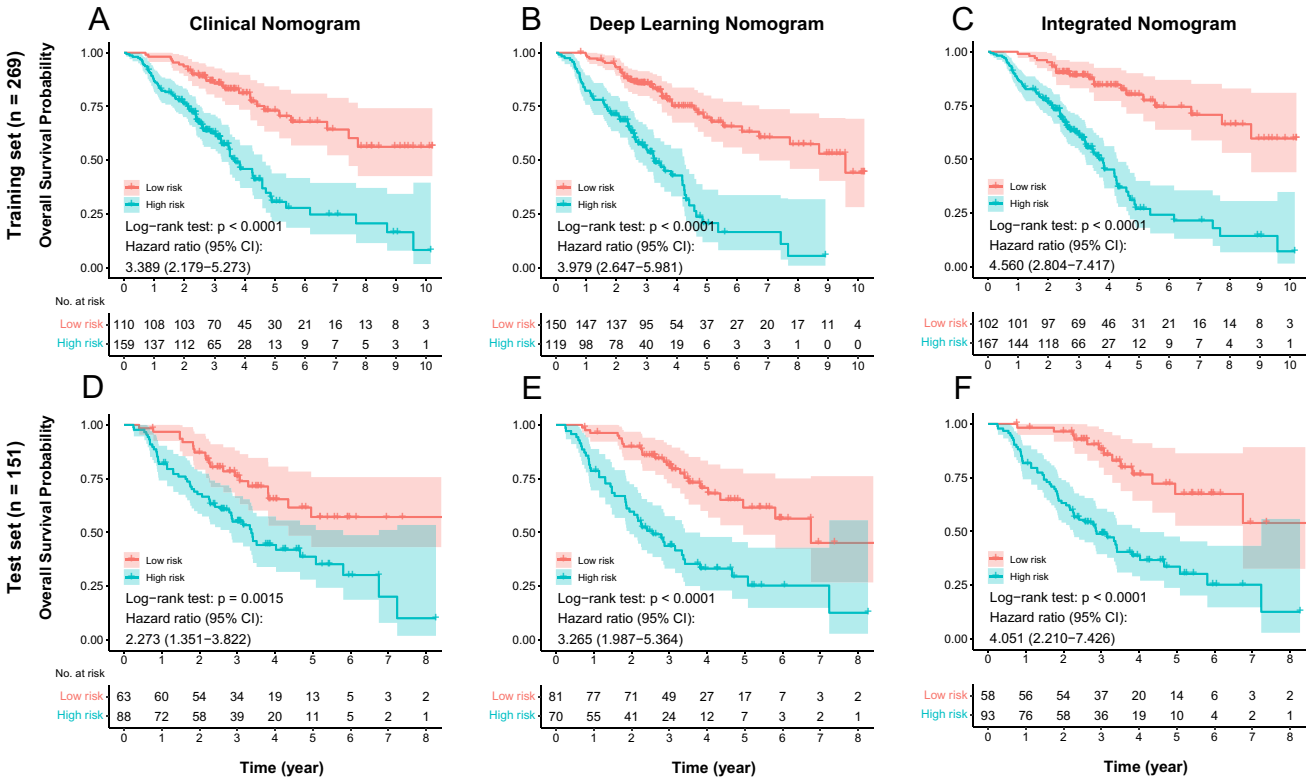

Fig. 3 Risk stratification of recurrent nasopharyngeal carcinoma using clinical, deep learning, and integrated nomograms in training set and test set. **A** Kaplan–Meier curves of clinical nomogram in training set. **B** Kaplan–Meier curves of deep learning nomogram in training set. **C** Kaplan–Meier curves of integrated nomogram in training set. **D** Kaplan–Meier curves of clinical nomogram in test set. **E** Kaplan–Meier curves of deep learning nomogram in test set. **F** Kaplan–Meier curves of integrated nomogram in test set. Curves were stratified by a cutoff value with best log-rank test in training set using each nomogram

vs. 64.7%), with a hazard ratio (HR) of 4.56 (95% CI: 2.80–7.42, $p < 0.001$). The threshold value remained valid in the test set with an HR of 4.05 (95% CI: 2.21–7.43, $p < 0.001$, OS rate: 89.7% vs. 50.5%) (Fig. 3). The prognostic ability of the integrated nomogram was kept in patient subgroups when stratified with sex (male or female), EBV DNA (undetectable or detectable), N-stage (N0 or N1–N3), or treatment type (surgery or re-irradiation) (Fig. 4). Detailed clinical information differences between the high-risk group and low-risk group are shown in Supplementary Table S3.

Association between high- and low-risk deep learning characteristics and biological processes

Seven patients (4 from the training set and 3 from the test set) with RNA-seq data were stratified into the high-risk group and low-risk group by the integrated nomogram. Heatmaps of the gene expression revealed the difference between the two risk groups (Supplementary Figure S3). The enriched pathways of differentially expressed genes were implemented by a GO enrichment analysis and identified pathways including apical plasma membrane, cell–matrix adhesion, cell-substrate adhesion extracellular matrix, extracellular matrix organization, extracellular matrix structural constituent, positive regulation of cell migration, skeletal system development, and structural molecule activity. The analysis of KEGG pathway enrichment showed that the differentially expressed genes (DEGs) were involved in pathways including complement and coagulation cascades, ECM-receptor interaction, focal adhesion, human papillomavirus infection, NOD-like receptor signaling pathway, PI3K-Akt signaling pathway, protein digestion and absorption, proteoglycans in cancer, ribosomes, and staphylococcus aureus infection (Fig. 5).

Discussion

To the best of our knowledge, this study presents the first attempt to investigate the multiscale intratumor heterogeneity associated with biological functions and survival in rNPC by deep learning signatures.

In this work, we constructed a multi-modality DCNN to generate deep learning signatures from joint CT and PET images. Our multi-modality deep learning signatures had

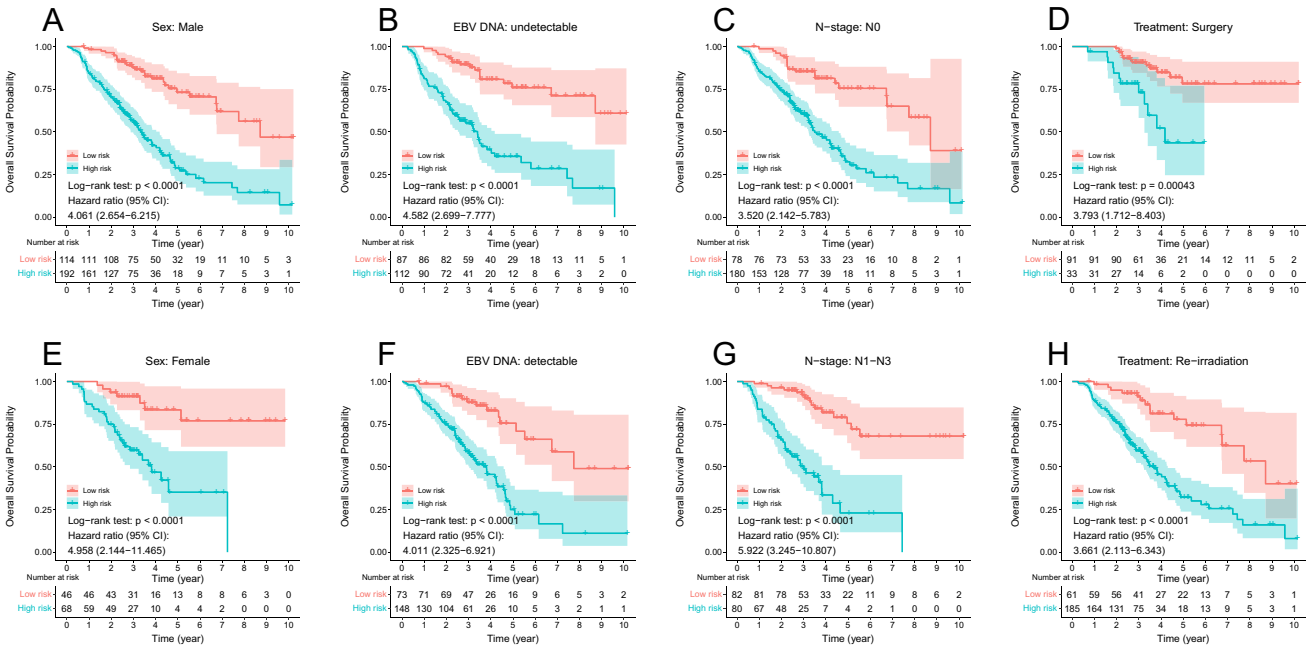

Fig. 4 Subgroup risk stratification of recurrent nasopharyngeal carcinoma using integrated nomogram. Kaplan–Meier curves of integrated nomogram in patient subgroups, which were grouped by sex: male (**A**) vs. female (**E**), EBV-DNA: undetectable (**B**) vs. detectable (**F**), N-stage: N0 (**C**) vs. N1–N3 (**G**), and treatment surgery (**D**) vs. re-irradiation (**H**). Curves were stratified by a cutoff value with best log-rank test in training set using integrated nomogram

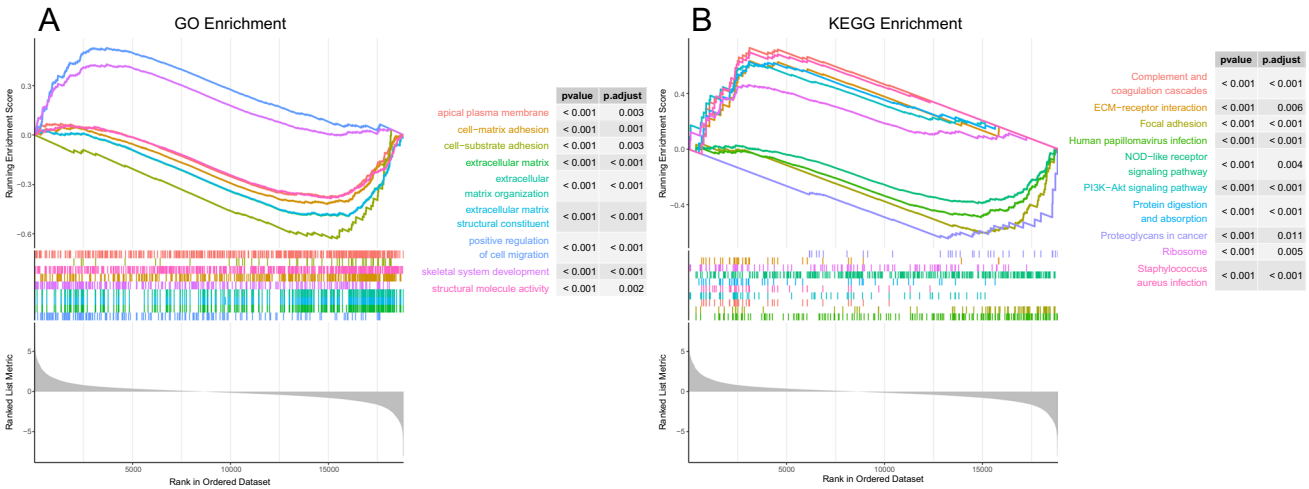

Fig. 5 Biological functions associated with risk group stratification of integrated nomogram. **A** Top 10 Gene ontology (GO) pathways associated with integrated nomogram risk stratification. **B** Top 10 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways associated with integrated nomogram risk stratification

better prognostic ability than single-modality deep learning signatures and clinical factors. To further improve our deep learning signatures, our integrated nomogram with deep learning information and clinical information can divide patients with newly diagnosed rNPC into low-risk group and high-risk group. Patients in the low-risk group were characterized by better OS and better survival with surgery and re-irradiation than high-risk group. When using RNA-seq, the prediction of our integrated nomogram was found to be associated with biological functions with enrichment of downregulated cell growth and development related pathways in the low-risk group.

Currently, the rTNM staging system has limitations in predicting the survival rate of rNPC patients. As shown

by our results, the integrated nomogram performed better than the clinical system in the risk stratification in both the training set and test set, (C-index: 0.741 and 0.732 vs. 0.673 and 0.667 in the training and test set, respectively), indicating the potential advantages of deep learning signatures. The tumor TNM system was developed based on the extent of tumor invasion. However, even patients with the same tumor stage may have a different prognosis due to tumor heterogeneity [23, 24]. Therefore, traditional clinical tumor staging cannot provide enough information for prognosis. Moreover, the deep learning signature carries information about intratumoral heterogeneity, which was reported to be a significant prognostic factor [25].

Plasma EBV DNA has been widely recognized as a promising biomarker in clinical practice for NPC [26–28]. In our study, EBV DNA did not have statistical significance in univariable analysis and multi-variable analysis. This phenomenon could be explained as the sensitivity of EBV DNA was significantly higher in treatment-naive patients than those with locoregional recurrent tumors. As a consequence, the level of EBV DNA in rNPC patients was often divided into detectable EBV DNA levels and undetectable ones, which makes it less meaningful [29].

Standardized uptake value (SUV) was shown to be supplementary to the PET/CT-based radiomics nomogram in lymphovascular invasion prediction [30]. In our study, tumor SUV_{max} and tumor SUV_{peak} were significantly associated with OS, but could not improve the performance of the integrated nomogram. This could be explained by the correlation and collinearity between the SUV and deep learning signature (Pearson's $r > 0.4$).

Subgroup analysis showed that low-risk patients could benefit from the locoregional treatment including surgery and re-irradiation, while there was no difference in survival rates between patients treated by chemotherapy in two groups. The balance between risks and benefits was very delicate among them; therefore, it was difficult to benefit from local treatments of re-irradiation and endoscopic nasopharyngectomy. For re-irradiation, given the risks of late toxicities from the previous course of RT, it was difficult for patients to tolerate another course of RT. Especially the high incidence of fatal adverse events including severe temporal lobe necrosis, mucosal necrosis, and massive hemorrhage after re-irradiation, which were reported to be 13%, 16%, and 16%, respectively [3]. For patients with bone, nerve, or vascular invasion, the difficulty of salvage surgical resection was greatly increased. Therefore, patients undergoing salvage surgical resection need to be strictly selected. Given the clue that the therapeutic value of salvage surgery and re-irradiation would be weakened in high-risk rNPC patients, the management of high-risk patients should be a top priority regarding palliative care and the improvement in the quality of life, and the application of palliative chemotherapy may be sufficient.

Recently, the emergence of deep learning has resulted in significant improvements in the predictive performance of prognosis in malignancy. Our research innovatively provided spatial transcriptome sequencing data to further explore the mechanism on high- and low-risk patients from the genetic and molecular levels. Through an enrichment analysis, we found that pathway association with the cytoskeleton and cell migration were significantly upregulated in high-risk patients, which further supported the results that high-risk patients tended to have a poor prognosis.

Limitations were inevitable in our study and were acknowledged. Firstly, the quality of life, acute and late adverse effects during and after treatment were not documented and analyzed due to the retrospective nature of the study. Secondly, tissues for sequencing were obtained from only a few patients. Thirdly, the data were collected from a single center and external verification may be required. Therefore, results of our research should be further validated in well-designed prospective studies in the future.

Conclusion

In conclusion, our study demonstrated that PET/CT-based deep learning signatures showed satisfactory prognostic predictive performance in rNPC patients. The nomogram incorporating deep learning signatures successfully divided patients into different risks and had great potential to guide individual treatment: low-risk patients were supposed to be treated with surgery and re-irradiation while the management of high-risk patients should pay more attention to palliative care and the improvement in the quality of life, hence the application of palliative chemotherapy may be sufficient.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00259-022-05793-x.

Acknowledgements The authors thank Prof. Xuebin Xie from KiangWu Hospital, Macau, China for his great help about the clinical part in this paper.

Funding This work was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (XDB38040200), National Key R&D Program of China (2017YFA0205200, 2017YFC0908500, 2017YFC1309003), National Natural Science Foundation of China (82022036, 91959130, 81971776, 81771924, 62027901, 81930053, 81425018, 81672868, 81802775, 82073003, 82002852, 82003267, 81803105), the Beijing Natural Science Foundation (L182061), the Youth Innovation Promotion Association CAS (Y2021049, 2017175), the Sci-Tech Project Foundation of Guangzhou City (201707020039), the Sun Yat-sen University Clinical Research 5010 Program

(2016010, 201315, 2015021, 2017010, 2016013, 2019023), Innovative research team of high-level local universities in Shanghai (SSMU-ZLCX20180500), the Natural Science Foundation of Guangdong Province (2017A030312003, 2018A0303131004), the Natural Science Foundation of Guangdong Province for Distinguished Young Scholar (2018B030306001), the Health & Medical Collaborative Innovation Project of Guangzhou City (No. 201803040003), Pearl River S&T Nova Program of Guangzhou (201806010135), the Planned Science and Technology Project of Guangdong Province (2019B020230002), Natural Science Foundation of Guangdong Province (2017A030312003), the Key Youth Teacher Cultivating Program of Sun Yat-sen University (20ykzd24), and the Fundamental Research Funds for the Central Universities.

Declarations

Ethics approval All procedures performed in the study and involving human participants were carried out in accordance with the ethical standards of the institutional and/or national research committee and with the principles of the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Consent to participate The ethical review board of each participating center approved this retrospective study and waived the requirement of informed consent.

Conflict of interest The authors declare no competing interests.

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;68:394–424.
2. Lee AWM, Ma BBY, Ng WT, Chan ATC. Management of nasopharyngeal carcinoma: current practice and future perspective. J Clin Oncol. 2015;33:3356–64.
3. Ng WT, Soong YL, Ahn YC, AlHussain H, Choi HCW, Corry J, et al. International recommendations on reirradiation by intensity modulated radiation therapy for locally recurrent nasopharyngeal carcinoma. Int J Radiat Oncol Biol Phys. 2021;110:682–95.
4. Liu L-T, Chen Q-Y, Tang L-Q, Zhang L, Guo S-S, Guo L, et al. With or without reirradiation in advanced local recurrent nasopharyngeal carcinoma: a case-control study. BMC Cancer. 2016;16:774.
5. Sun X-S, Liang Y-J, Jia G-D, Liu S-L, Liu L-T, Guo S-S, et al. Establishment of a prognostic nomogram to identify optimal candidates for local treatment among patients with local recurrent nasopharyngeal carcinoma. Oral Oncol. 2020;106:104711.
6. Han F, Zhao C, Huang S-M, Lu L-X, Huang Y, Deng X-W, et al. Long-term outcomes and prognostic factors of re-irradiation for locally recurrent nasopharyngeal carcinoma using intensity-modulated radiotherapy. Clin Oncol. 2012;24:569–76.
7. Bruixola G, Remacha E, Jiménez-Pastor A, Dualde D, Viala A, Montón JV, et al. Radiomics and radiogenomics in head and neck squamous cell carcinoma: potential contribution to patient management and challenges. Cancer Treat Rev. 2021;99:102263.
8. Beig N, Singh S, Bera K, Prasanna P, Singh G, Chen J, et al. Sexually dimorphic radiogenomic models identify distinct imaging and biological pathways that are prognostic of overall survival in glioblastoma. Neuro-Oncol. 2021;23:251–63.
9. Dong D, Zhang F, Zhong L-Z, Fang M-J, Huang C-L, Yao J-J, et al. Development and validation of a novel MR imaging predictor of response to induction chemotherapy in locoregionally advanced nasopharyngeal cancer: a randomized controlled trial substudy (NCT01245959). BMC Med. 2019;17:190.
10. Zhang B, Tian J, Dong D, Gu D, Dong Y, Zhang L, et al. Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma. Clin Cancer Res. 2017;23:4259–69.
11. Zhang F, Zhong L-Z, Zhao X, Dong D, Yao J-J, Wang S-Y, et al. A deep-learning-based prognostic nomogram integrating microscopic digital pathology and macroscopic magnetic resonance images in nasopharyngeal carcinoma: a multi-cohort study. Ther Adv Med Oncol. 2020;12:1758835920971416.
12. Zhang L, Zhou H, Gu D, Tian J, Zhang B, Dong D, et al. Radiomic nomogram: pretreatment evaluation of local recurrence in nasopharyngeal carcinoma based on MR imaging. J Cancer. 2019;10:4217–25.
13. Peng H, Dong D, Fang M-J, Li L, Tang L-L, Chen L, et al. Prognostic value of deep learning PET/CT-based radiomics: potential role for future individual induction chemotherapy in advanced nasopharyngeal carcinoma. Clin Cancer Res. 2019;25:4271–9.
14. Zhong L, Dong D, Fang X, Zhang F, Zhang N, Zhang L, et al. A deep learning-based radiomic nomogram for prognosis and treatment decision in advanced nasopharyngeal carcinoma: a multicentre study. EBioMedicine. 2021;70:103522.
15. Kirienko M, Sollini M, Corbetta M, Voulaz E, Gozzi N, Interlenghi M, et al. Radiomics and gene expression profile to characterise the disease and predict outcome in patients with lung cancer. Eur J Nucl Med Mol Imaging. 2021;48:3643–55.
16. Zinn PO, Singh SK, Kotrotsou A, Hassan I, Thomas G, Luedi MM, et al. A coclinical radiogenomic validation study: conserved magnetic resonance radiomic appearance of periostin-expressing glioblastoma in patients and xenograft models. Clin Cancer Res. 2018;24:6288–99.
17. Fan M, Xia P, Clarke R, Wang Y, Li L. Radiogenomic signatures reveal multiscale intratumour heterogeneity associated with biological functions and survival in breast cancer. Nat Commun. 2020;11:4861.
18. Dong D, Fang M-J, Tang L, Shan X-H, Gao J-B, Giganti F, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study. Ann Oncol. 2020;31:912–20.
19. Dong D, Tang L, Li Z-Y, Fang M-J, Gao J-B, Shan X-H, et al. Development and validation of an individualized nomogram to identify occult peritoneal metastasis in patients with advanced gastric cancer. Ann Oncol. 2019;30:431–8.
20. Wen Y-H, Wen W-P, Chen H-X, Li J, Zeng Y-H, Xu G. Endoscopic nasopharyngectomy for salvage in nasopharyngeal carcinoma: a novel anatomic orientation. Laryngoscope. 2010;120:1298–302.
21. Tan M, Le QV. Efficientnet: rethinking model scaling for convolutional neural networks. In: ICML. 2019. pp. 6105–14.
22. Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. BMC Med Res Methodol. 2018;18:24.
23. Friemel J, Rechsteiner M, Frick L, Böhm F, Struckmann K, Egger M, et al. Intratumor heterogeneity in hepatocellular carcinoma. Clin Cancer Res. 2015;21:1951–61.
24. Yang F, Wang Y, Li Q, Cao L, Sun Z, Jin J, et al. Intratumor heterogeneity predicts metastasis of triple-negative breast cancer. Carcinogenesis. 2017;38:900–9.

25. Huang Y-Q, Liang C-H, He L, Tian J, Liang C-S, Chen X, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. J Clin Oncol. 2016;34:2157–64.
26. Leung S, Zee B, Ma BB, Hui EP, Mo F, Lai M, et al. Plasma Epstein-Barr viral deoxyribonucleic acid quantitation complements tumor-node-metastasis staging prognostication in nasopharyngeal carcinoma. J Clin Oncol. 2006;24:5414–8.
27. Tang L-Q, Chen Q-Y, Fan W, Liu H, Zhang L, Guo L, et al. Prospective study of tailoring whole-body dual-modality [18F] fluorodeoxyglucose positron emission tomography/computed tomography with plasma Epstein-Barr virus DNA for detecting distant metastasis in endemic nasopharyngeal carcinoma at initial staging. J Clin Oncol. 2013;31:2861–9.
28. Guo R, Tang L-L, Mao Y-P, Du X-J, Chen L, Zhang Z-C, et al. Proposed modifications and incorporation of plasma Epstein-Barr virus DNA improve the TNM staging system for Epstein-Barr virus-related nasopharyngeal carcinoma. Cancer. 2019;125:79–89.
29. Leung S-F, Lo YMD, Chan ATC, To K-F, To E, Chan LYS, et al. Disparity of sensitivities in detection of radiation-naïve and postirradiation recurrent nasopharyngeal carcinoma of the undifferentiated type by quantitative analysis of circulating Epstein-Barr virus DNA1,2. Clin Cancer Res. 2003;9:3431–4.
30. Nie P, Yang G, Wang N, Yan L, Miao W, Duan Y, et al. Additional value of metabolic parameters to PET/CT-based radiomics nomogram in predicting lymphovascular invasion and outcome in lung adenocarcinoma. Eur J Nucl Med Mol Imaging. 2021;48:217–30.

Authors and Affiliations

Xun Zhao[1,2] · Yu-Jing Liang[3,4] · Xu Zhang[4,5] · Dong-Xiang Wen[3,4] · Wei Fan[4,5] · Lin-Quan Tang[3,4] · Di Dong[1,2] · Jie Tian[1,2] · Hai-Qiang Mai[3,4]

✉ Lin-Quan Tang
tanglq@sysucc.org.cn

✉ Di Dong
di.dong@ia.ac.cn

✉ Jie Tian
jie.tian@ia.ac.cn

✉ Hai-Qiang Mai
maihq@sysucc.org.cn

[1] School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, People's Republic of China

[2] CAS Key Laboratory of Molecular Imaging, Beijing Key Laboratory of Molecular Imaging, The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, No. 95 Zhongguancun East Road, Hai Dian District, Beijing, People's Republic of China

[3] Sun Yat-Sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Guangzhou 510060, People's Republic of China

[4] Department of Nasopharyngeal Carcinoma, Sun Yat-Sen University Cancer Center, 651 Dongfeng Road East, Guangzhou 510060, People's Republic of China

[5] Department of Nuclear Medicine, Sun Yat-Sen University Cancer Center, Guangzhou 510060, People's Republic of China