Precision Medicine and Imaging

Prognostic Value of Deep Learning PET/CT-Based Radiomics: Potential Role for Future Individual Induction Chemotherapy in Advanced Nasopharyngeal Carcinoma

Hao Peng[1], Di Dong[2,3], Meng-Jie Fang[2,3], Lu Li[4,5], Ling-Long Tang[1], Lei Chen[1], Wen-Fei Li[1], Yan-Ping Mao[1], Wei Fan[5], Li-Zhi Liu[6], Li Tian[6], Ai-Hua Lin[7], Ying Sun[1], Jie Tian[2,8,9], and Jun Ma[1]

Abstract

Purpose: We aimed to evaluate the value of deep learning on positron emission tomography with computed tomography (PET/CT)–based radiomics for individual induction chemotherapy (IC) in advanced nasopharyngeal carcinoma (NPC).

Experimental Design: We constructed radiomics signatures and nomogram for predicting disease-free survival (DFS) based on the extracted features from PET and CT images in a training set ($n = 470$), and then validated it on a test set ($n = 237$). Harrell's concordance indices (C-index) and time-independent receiver operating characteristic (ROC) analysis were applied to evaluate the discriminatory ability of radiomics nomogram, and compare radiomics signatures with plasma Epstein–Barr virus (EBV) DNA.

Results: A total of 18 features were selected to construct CT-based and PET-based signatures, which were significantly associated with DFS ($P < 0.001$). Using these sig-

natures, we proposed a radiomics nomogram with a C-index of 0.754 [95% confidence interval (95% CI), 0.709–0.800] in the training set and 0.722 (95% CI, 0.652–0.792) in the test set. Consequently, 206 (29.1%) patients were stratified as high-risk group and the other 501 (70.9%) as low-risk group by the radiomics nomogram, and the corresponding 5-year DFS rates were 50.1% and 87.6%, respectively ($P < 0.0001$). High-risk patients could benefit from IC while the low-risk could not. Moreover, radiomics nomogram performed significantly better than the EBV DNA-based model (C-index: 0.754 vs. 0.675 in the training set and 0.722 vs. 0.671 in the test set) in risk stratification and guiding IC.

Conclusions: Deep learning PET/CT-based radiomics could serve as a reliable and powerful tool for prognosis prediction and may act as a potential indicator for individual IC in advanced NPC.

Introduction

Nasopharyngeal carcinoma (NPC) is a special kind of head and neck cancers which is mainly endemic in South Asia (1). Although the advance in radiotherapy technique and chemotherapy strategies has improved the prognosis of NPC, outcomes of patients with advanced disease still remain unsatisfactory, with nearly 30% of cases suffering treatment failure (2, 3). Unfortunately,

more than 70% of the patients present with locoregionally advanced disease at initial diagnosis (4, 5). Management of advanced disease remains a challenge for clinicians.

Induction chemotherapy (IC), given before radical radiotherapy, has widely been proven a feasible neoadjuvant treatment with satisfactory efficacy and low toxicities in advanced NPC during the past decade (6–9). Consequently, IC has been

[1]Department of Radiation Oncology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in Southern China, Collaborative Innovation Center for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Guangdong, P. R. China. [2]CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, P. R. China. [3]University of Chinese Academy of Sciences, Beijing, P. R. China. [4]Department of Radiation Oncology, Nanfang Hospital, Southern Medical University, Guangzhou, P. R. China. [5]Department of Nuclear Medicine, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, Guangdong, P. R. China. [6]Imaging Diagnosis and Interventional Center, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in Southern China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, Guangdong, P. R. China. [7]Department of Medical Statistics and Epidemiology, School of Public Health, Sun Yat-sen University, P. R. China. [8]Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine, Beihang University, Beijing, P. R. China. [9]Engineering Research Center of Molecular and Neuro Imaging of Ministry of Education,

School of Life Science and Technology, Xidian University, Xi'an, Shaanxi, P. R. China.

Corresponding Authors: Jun Ma, Sun Yat-sen University Cancer Center, 651 Dongfeng Road East, Guangzhou 510060, People's Republic of China. Phone: 86-20-87343469; Fax: 86-20-87343392; E-mail: majun2@mail.sysu.edu.cn; and Jie Tian, CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Beijing 100190, P. R. China; E-mail: jie.tian@ia.ac.cn

Peng et al.

Translational Relevance

Induction chemotherapy (IC) plus concurrent chemoradiotherapy (CCRT) has emerged as the standard care for locoregionally advanced nasopharyngeal carcinoma (LA-NPC). However, most of the patients could not benefit from additional IC, and we still lack effective markers to perform individualized IC. Our current study developed and validated that deep learning-aided PET/CT radiomics could serve as a powerful prognostication and help to individual IC. Our findings would provide important evidence for clinical treatment of LA-NPC.

routinely recommended for advanced NPC. However, it should be pointed out that the advanced disease consisted of many subgroups and not all of them could benefit from additional IC (10, 11). Thus, identifying the high-risk subgroups who could benefit from IC is the key to improve management of advanced NPC. Although a few retrospective studies have found that pretreatment plasma Epstein–Barr virus DNA (pre-DNA) could act as an indicator for IC (12, 13), these pieces of evidence were not strong. Most importantly, the assay standardization of plasma EBV DNA has constrained its wide application because different labs used different polymerase chain reaction assays and therefore produced inconsistent results (14). Thus, it is worth identifying novel and powerful factors to guide IC.

Radiomics has recently emerged as a promising field in oncology, and is based on the premise that medical imaging can provide important information on tumor physiology (15, 16). By translating medical imaging into mineable, high-dimension, and quantitative imaging features via high-throughput extraction of data-characterization algorithms, radiomics offers an easy, effective, and reliable method of stratifying patients into risk groups

and aids decision-making (15–17). Meanwhile, the novel deep learning techniques have shown the promising capabilities to extract correlative quantitative representation in many medical applications (18, 19). Specially, the patch-based strategy makes it possible to implement the training process on a relatively small data set (20–22). Given this, we conducted this study to evaluate the role of deep learning positron emission tomography with computed tomography (PET/CT)-based radiomics in risk stratification and guiding individual IC for patients with advanced NPC undergoing intensity-modulated radiotherapy (IMRT).
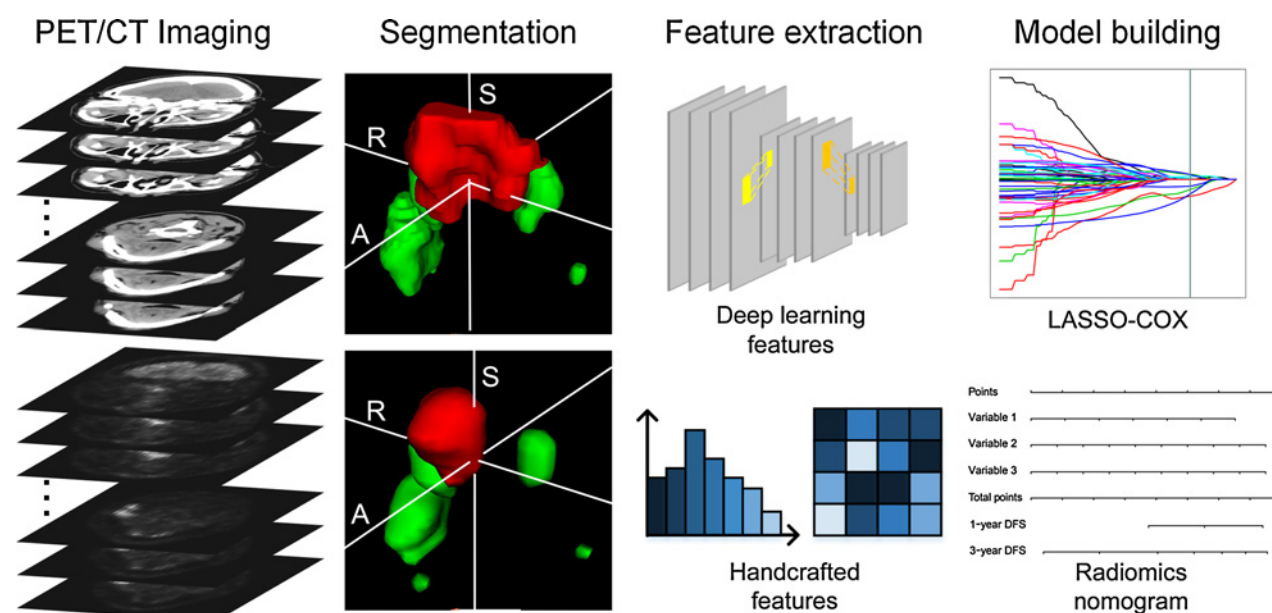
Materials and Methods

Participant inclusion

Patients treated at our center between December 2009 and December 2014 were reviewed and included for this study if they: (i) received pretreatment 18F-FDG PET/CT test; (ii) had newly diagnosed stage III-IVA disease; (iii) treated by concurrent chemoradiotherapy (CCRT) with or without IC; (iv) received IMRT; (v) did not have other malignancies. Flow chart of patient inclusion was presented in Supplementary Fig. S1. This study was approved by the Research Ethics Committee of our Center, and written informed consent was obtained from all patients before treatment. Also, our study was carried out in accordance with the Declaration of Helsinki. The study data underlying the findings of current work were deposited at the Research Data Deposit platform (RDDA2018000721, available at http://www.researchdata.org.cn/).

PET/CT imaging protocol

18F-FDG PET/CT scans were performed using a dedicated PET/CT system (Discovery ST16; GE Medical Systems). Imaging was performed using a combination PET/CT scanner according to PET/CT tumor imaging guidelines (23). Detailed information on PET/CT protocol was described in Supplementary Methods.

Figure 1.
Radiomics workflow in this study.

Imaging segmentation

PET/CT images were retrieved from the picture archiving and communication system and then loaded into ITK-SNAP software (version 2.2.0; www.itksnap.org) for manual segmentation. A radiation oncologist (L. Chen) with 13 years of experience outlined the regions of interest (ROI), which included volumes of the tumor and lymph nodes, on the PET and CT images, respectively. Therefore, there were four different ROIs being segmented for each patient in this study (Supplementary Fig. S2). After 3 months, 50 patients in the training set were selected randomly and segmented again by him and another radiation oncologist (L.-L. Tang) with 15 years of experience to assess intra/interreader agreement of the radiomics analysis.

Radiomics feature extraction

Both deep learning features and handcrafted features were extracted based on the PET/CT images to quantify the tumor phenotype (Fig. 1). For each ROI, 136 deep learning features and 133 handcrafted features were extracted. We constructed and trained four deep convolutional neural networks (DCNNs with 12 or 8 weighted layers) to extract deep learning features on the four groups of ROIs, respectively (Supplementary Fig. S3). A set of handcrafted features, which was defined by experiential algorithms, was also extracted. The features could be divided into four groups: shape features, histogram features, gray-level co-occurrence matrix features, and gray-level run-length matrix features.

The architecture and implementation of our DCNNs and the feature extraction pipeline were detailed in Supplementary Methods. Our DCNNs were implemented based on the Python Keras package (https://github.com/fchollet/keras) with the TensorFlow library (https://www.tensorflow.org) as the backend. The handcrafted feature extraction was performed in MATLAB 2017a (MathWorks) using an in-house developed tool box.

Feature selection and radiomics signature building

We built two radiomics signatures reflecting the phenotypic characteristics of the primary tumor and the lymph nodes in CT and PET images, respectively, as independent predictors of disease-free survival (DFS), i.e., the CT-based signature and the PET-based signature. The least absolute shrinkage and selection operator (LASSO) Cox regression method was used to select the most useful prognostic combination of features. Then, the radiomics score (Rad-score) was computed for each patient through a linear combination of selected features weighted by their respective coefficients. Both feature selection and the following radiomics signature construction were performed in the training set. Supplementary Methods detailed the feature selection and radiomics signature construction. Furthermore, signatures combining either the handcrafted features or the deep learning features were also developed using the same methods for comparison.

Staging workup and treatment

All patients were staged by PET/CT and MRI. Two radiologists (L.-Z. Liu and L. Tian) reviewed the MRI scans independently, and discrepancy was solved by consensus. Tumor stage was grouped according to the 8th edition of the *International Union against Cancer/American Joint Committee on Cancer* manual.

All the patients received radical IMRT. The cumulative radiation doses were 66 Gy or greater to the primary tumor and 60 to 70 Gy to the involved neck area. All potential sites of local infiltration and bilateral cervical lymphatics were irradiated to 50 Gy or greater. All patients were treated with 30 to 35 fractions with five daily fractions per week for 6 to 7 weeks. IC included cisplatin-based regimens every three weeks for 2 to 4 cycles. Concurrent chemotherapy was weekly or triple-weekly cisplatin.

Clinical endpoints and follow-up

To allow earlier individual treatment (24), we set DFS (time from diagnosis to disease progression or death from any cause) as the main endpoint and nomograms were built based on it. Other endpoints included OS (time from diagnosis to death from any cause), distant metastasis-free survival (DMFS, time from

Table 1. Baseline information of the training and internal validation sets.

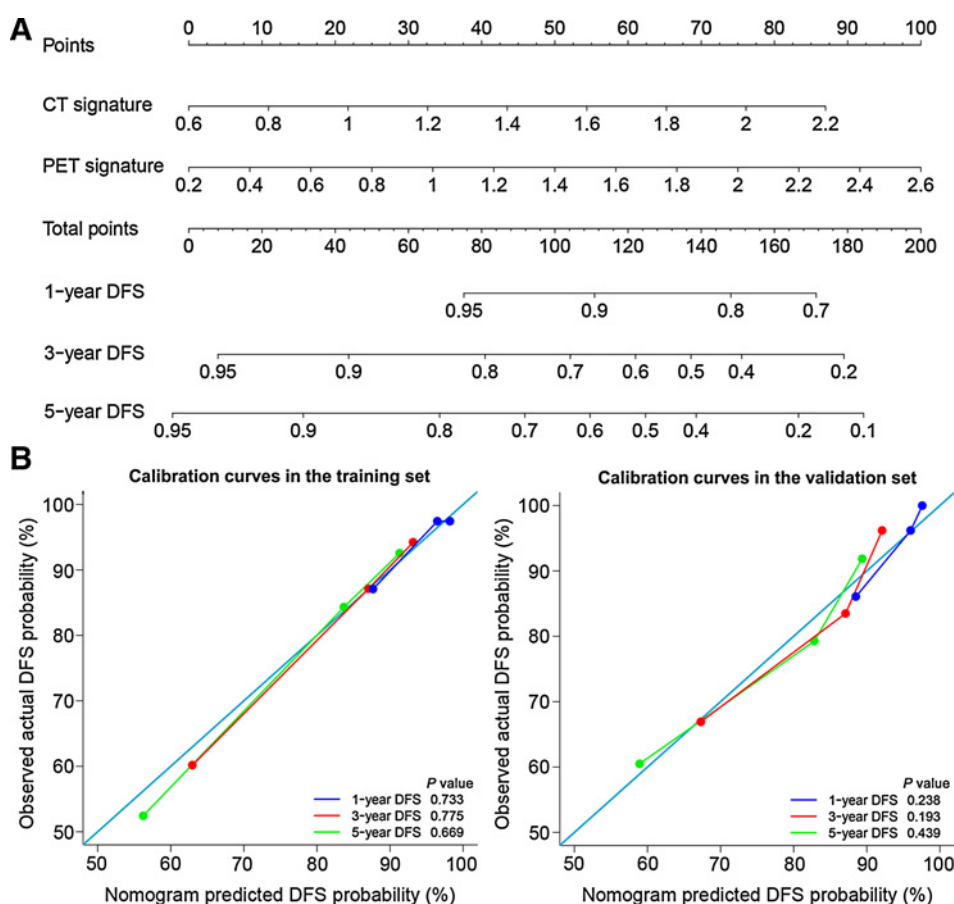| Characteristics	Training set ($n = 470$) No. (%)	Test set ($n = 237$) No. (%)	P^a
Age (y)			0.759
Median (range)	45 (9–76)	44 (10–76)	
Gender			0.458
Female	111 (23.6)	62 (26.2)	
Male	359 (76.4)	175 (73.8)	
Smoking			0.155
Yes	157 (33.4)	92 (38.8)	
No	313 (66.6)	145 (61.2)	
Drinking			0.454
Yes	58 (12.3)	34 (14.3)	
No	412 (87.7)	203 (85.7)	
WHO pathology type			0.710
I	3 (0.6)	1 (0.4)	
II–III	467 (99.4)	236 (99.6)	
Family history of cancer			0.231
Yes	143 (30.4)	62 (26.2)	
No	327 (69.6)	175 (73.8)	
LDH (U/L)			0.120
Median (range)	177 (100–658)	174 (118–626)	
HGB (g/L)			0.092
Median (range)	146 (79–178)	144 (91–176)	
ALB (g/L)			0.484
Median (range)	44.2 (31–53)	44 (25–54)	
CRP (mg/L)			0.549
Median (range)	2 (0–127.2)	2.1 (0–126.6)	
Pre-DNA (copies/mL)[b]			0.376
Median (range)	5,385 (0–68,700,000)	4,855 (0–1,840,000)	
T category[c]			0.118
T1	24 (5.1)	10 (4.2)	
T2	50 (10.6)	13 (5.5)	
T3	287 (61.1)	151 (63.7)	
T4	109 (23.2)	63 (26.6)	
N category[c]			0.694
N0	46 (9.8)	24 (10.1)	
N1	206 (43.8)	111 (46.8)	
N2	135 (28.7)	58 (24.5)	
N3	83 (17.7)	44 (18.6)	
Overall stage[c]			0.664
III	292 (62.1)	143 (60.3)	
IVA	178 (37.9)	94 (39.7)	
Treatment			0.085
IC + CCRT	322 (68.5)	147 (62.0)	
CCRT alone	148 (31.5)	90 (38.0)	

Abbreviations: WHO, World Health Organization; LDH, lactate dehydrogenase; HGB, hemoglobin; ALB, albumin; CRP, C-reaction protein; Pre-DNA, pretreatment plasma Epstein–Barr Virus DNA; IC, induction chemotherapy; CCRT, concurrent chemoradiotherapy.
[a]P values were calculated by the χ^2 test for categorical variables and nonparametric test for continuous variables.
[b]Twenty-three patients lost these data.
[c]According to the 8th edition of the *International Union against Cancer/American Joint Committee on Cancer* (*UICC/AJCC*) staging manual.

Figure 2.
A, Radiomics nomogram; **B,** Radiomics nomogram calibration curves. PET, positron emission tomography; CT, computed tomography.

diagnosis to first distant metastasis) and locoregional relapse-free survival (LRRFS, time from diagnosis to first local or regional recurrent or both).

Patients were followed by routine imaging methods every 3 months during the first 2 years, every 6 months at the 3rd to 5th years, and annually thereafter. Follow-up duration was measured from the day of diagnosis to last visit or death. All local and regional recurrence was confirmed by pathology. Distant metastasis was diagnosed mainly based on imaging methods such as MRI, CT, or PET/CT.

Statistical analysis

To compare radiomics signatures with pre-DNA, we also developed two clinical nomograms, one using only clinical factors (age, gender, smoking, drinking, family history of cancer, lactate dehydrogenase, hemoglobin, albumin, C-reaction protein, T category, N category, and overall stage) without pre-DNA (nomogram A), and another using these clinical factors with pre-DNA (nomogram B). The radiomics nomogram was defined as nomogram C.

To evaluate the reproducibility of our model's prognostic performance and the stability of the feature selection, we repeated the randomized assignment of training/test sets 10 times. Subsequently, the model was retrained and validated repeatedly.

Statistical analysis was conducted with R software (version 3.4.4; http://www.Rproject.org) and MATLAB. A two-sided P value < 0.05 was used as the criterion to indicate a statistically significant difference. Detailed information on statistical methods was shown in Supplementary Methods.
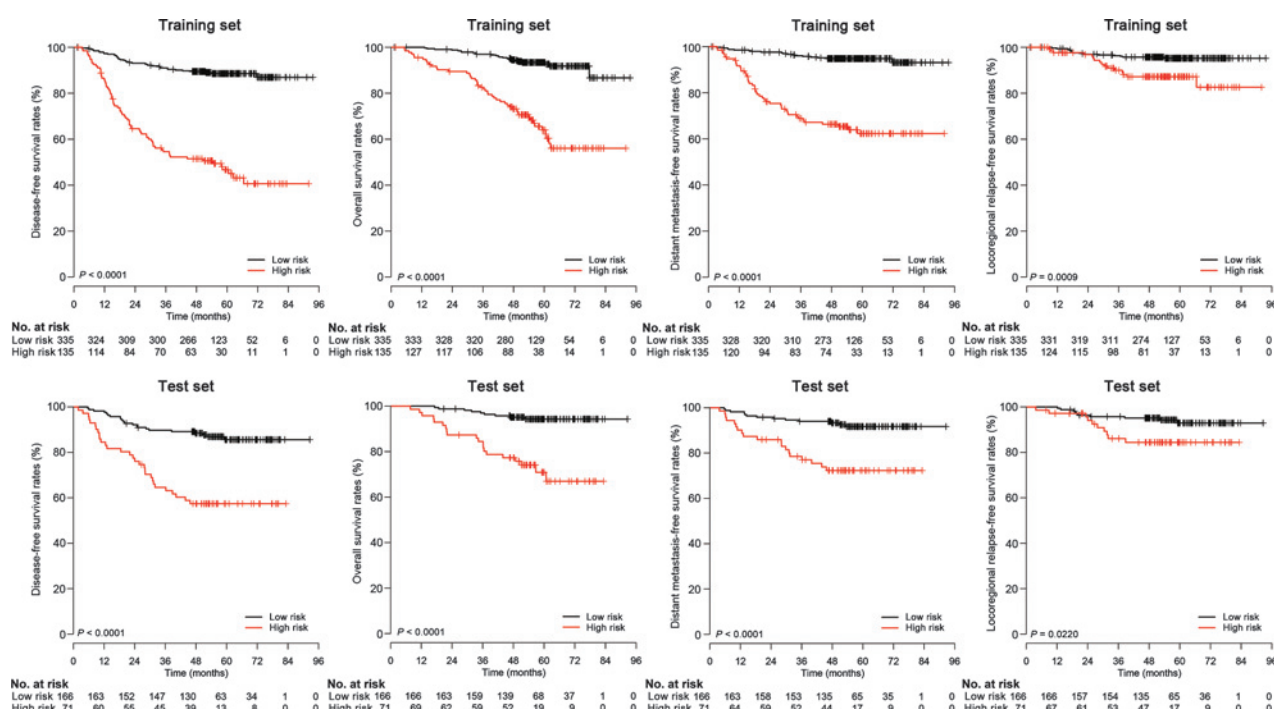
Results

Baseline information of participants

In total, 707 patients were recruited for this study, among them 436 (61.7%) and 271 (38.3%) patients had stage III and IVA disease, respectively. Additionally, 469 (66.3%) received IC plus CCRT and 238 (33.7%) received CCRT alone. We then used computer-generated random numbers to divide patients into a training set ($n = 470$) and a test set ($n = 237$; Table 1). The median follow-up duration of the whole cohort was 55.7 months (range, 1.3–93.6 months). Upon the last follow-up, 109 (23.2%) in the training set and 52 (21.9%) patients in the test set experienced a confirmed disease progression ($P = 0.708$).

Radiomics signature building and validation

Five and 13 radiomics features were selected from the CT-based and PET-based feature sets, respectively, and the detailed selection process was presented in Supplementary Tables S1 and S2. The selected features and corresponding coefficients in the formula of each Rad-score were listed in Supplementary Table S3. In the training set, the CT-based and PET-based Rad-score yielded C-indexes of 0.738 (95% CI, 0.690–0.786) and 0.730 (95% CI, 0.683–0.776), respectively. Good prognostic performances were validated with the corresponding C-indexes of 0.707 (95% CI, 0.635–0.779) and 0.683 (95% CI, 0.610–0.755) in the test set. Furthermore, as presented in Supplementary Table S4, the radiomics signature achieved the best discriminatory ability when it combined both handcrafted features and deep learning features.

Figure 3.
DFS, overall survival, DMFS, and locoregional relapse-free survival Kaplan–Meier curves between the radiomics nomogram–defined high-risk and low-risk groups in the training and test sets.

Development of an individualized prediction model

For univariate analysis, clinical factors including pre-DNA, N stage, and overall stage were found significantly associated with DFS (Supplementary Table S5). When multivariate Cox proportional hazard model was performed, the two radiomics signatures (CT-based signature [per 1 increase]: HR, 2.99; 95% CI, 1.84–4.86; $P < 0.001$; PET-based signature [per 1 increase]: HR, 2.32; 95% CI, 1.55–3.46; $P < 0.001$) remained significant for DFS after adjustment for various cofactors (Supplementary Table S6). Then, a radiomics nomogram for individualized DFS estimation was built using the above regression coefficients (Fig. 2A).

Performance and validation of the radiomics nomogram

The radiomics nomogram was significantly associated with DFS (all $P < 0.001$), with C-indexes of 0.754 (95% CI, 0.709–0.800) in the training set and 0.722 (95% CI, 0.652–0.792) in the test set. The calibration curves of nomogram for DFS are shown in Fig. 2B, which showed better agreement between the estimated outcomes and the observed outcomes (all $P > 0.05$). Moreover, the prognostic accuracy of the radiomics nomogram at 1, 3, and 5 years was also satisfactory (all $P < 0.01$; Supplementary Fig. S4).

We identified the cutoff score of radiomics nomogram as 0.311 corresponding to the total point of 79 in Fig. 2A. Consequently, 135 (28.7%) in the training set and 71 (30.0%) in the test set with scores ≥ 0.311 were classified as high-risk group, and 335 (71.3%) and 166 (70.0%) in the training and test sets with scores < 0.311 as low-risk group (Supplementary Fig. S5). Baseline information of the high-risk and low-risk groups was presented in Supplementary Table S7. For high-risk versus low-risk group, the 5-year DFS rate
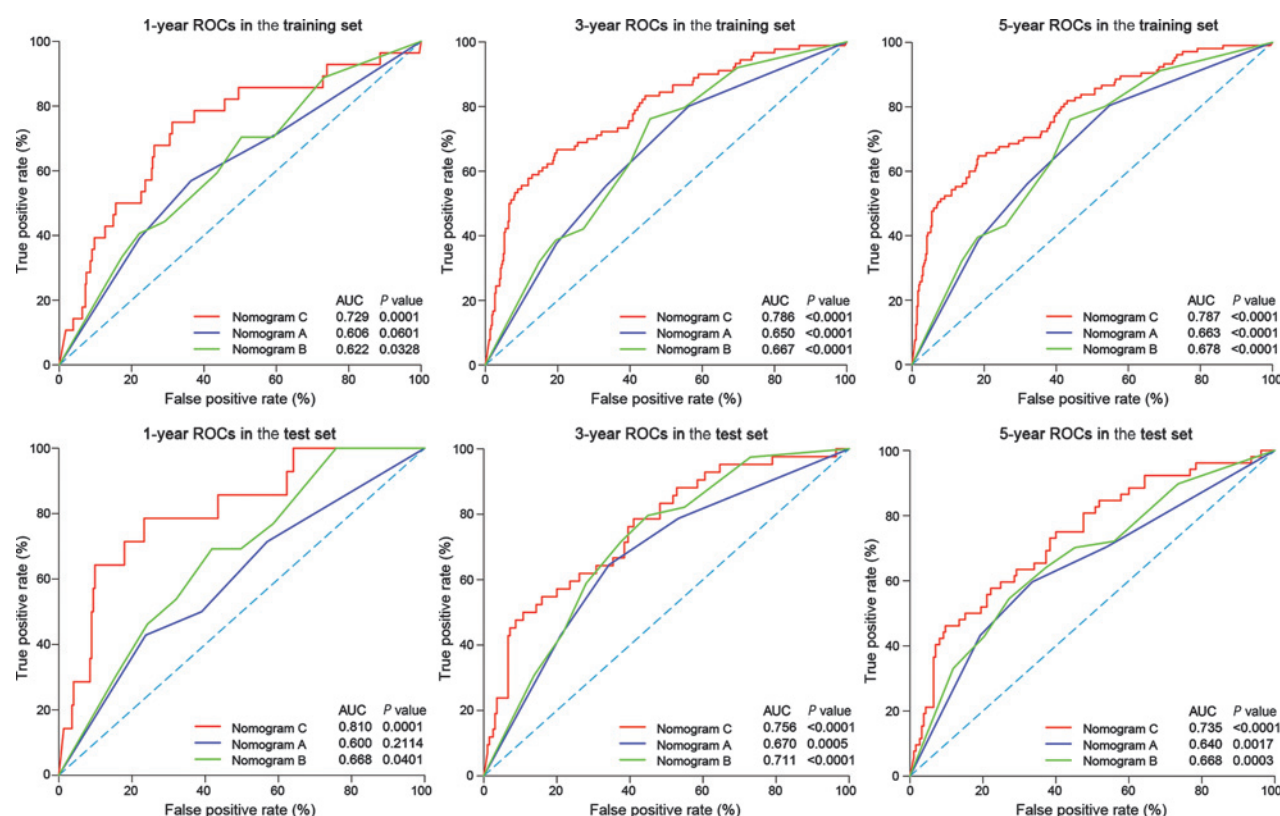
was 46.7% versus 88.6% (HR, 6.29; 95% CI, 4.24–9.35; $P < 0.001$) in the training set, and 57.4% vs. 85.6% (HR, 3.90; 95% CI, 2.24–6.76; $P < 0.001$) in the test set (Fig. 3). Similarly, patients in the low-risk group also achieved better OS, DMFS, and LRRFS (all $P < 0.01$; Fig. 3; Supplementary Table S8). When stratified by age (> 45 or ≤ 45 years), gender (female or male), and pre-DNA (>4,000 copies/mL or ≤ 4,000 copies/mL), the radiomics nomogram remained a clinically and statistically significant prognostic model (Supplementary Fig. S6). The KM curves of the low/high-risk groups crossed approximately at 2 years for DFS on the patients with overall stage IVA, which suggested that a finer-grained model constructed based on a larger-scale training set was needed. Meanwhile, our model successfully split the patients for different OS in all stratification cases, except for the female group in which only four patients died during the follow-up (Supplementary Fig. S7).

Furthermore, we split the whole data set into paired training (70%) and test (30%) sets 10 times, followed by the repeating construction and validation of the predictive model. In this experiment, the features involved into the new models yielded a very high possibility (144/165) to be highly correlated with the 18 selected features (i.e., with the Pearson correlation coefficients > 0.8). Moreover, no significant difference was found between the resulted C-indexes ranging from 0.703 to 0.749 in the holdout test sets.

Comparing radiomics signature with pre-DNA

Overall, data on pre-DNA were available for 456 patients in the training set and 228 patients in the test set. Independent factors and their coefficients for nomograms A and B were shown in Supplementary Results. In the training set, nomogram C (C-index,

Figure 4.
ROC curves comparing the predictive power of three nomograms for DFS in the training and test sets. ROC, receiver operator characteristic; AUC, area under the curve.

0.754; 95% CI, 0.709–0.800) achieved stronger prognostic ability for DFS than nomogram A (C-index, 0.684; 95% CI, 0.621–0.747) and nomogram B (C-index, 0.675; 95% CI, 0.619–0.731). This finding was also validated in the test set (nomogram C: C-index, 0.722; 95% CI, 0.652–0.792; nomogram A: C-index, 0.661; 95% CI 0.565–0.758; nomogram B: C-index, 0.671; 95% CI 0.590–0.752). Furthermore, time-independent receiver operating curve (ROC) analysis also validated that nomogram C had the best prognostic power (Fig. 4).

Benefit of induction chemotherapy

For the whole cohort, survival outcomes were comparable between IC + CCRT and CCRT-alone groups (Supplementary Fig. S8; Supplementary Table S9). Then, we applied our radiomics nomogram to predict if patients could benefit from IC. Within the high-risk group, patients receiving IC plus CCRT ($n = 173$) achieved significantly better 5-year DFS (53.5% vs. 32.5%, $P = 0.001$), OS (71.8% vs. 38.1%, $P < 0.001$), and DMFS (70.6% vs. 40.0%, $P < 0.001$; Fig. 5; Supplementary Table S10) rates than those receiving CCRT alone ($n = 33$). However, for the 501 patients with low risk, 5-year DFS (88.9% vs. 85.7%, $P = 0.505$), OS (93.5% vs. 94.0%, $P = 0.611$), DMFS (93.6% vs. 93.9%, $P = 0.815$), and LRRFS (94.3% vs. 90.0%, $P = 0.162$; Supplementary Fig. S9; Supplementary Table S11) rates did not signifirantly differ between IC plus CCRT ($n = 296$) and CCRT alone ($n = 205$). When applying nomograms A and B to predict the benefit of IC, they

either failed or had less power than nomogram C (Supplementary Results).

Discussion

We undertook this study to develop and validate the prognostic value of multiparametric PET/CT-based radiomics in advanced NPC, and our findings suggested that the radiomics nomogram was powerful in risk stratification and guiding the individual IC. Moreover, the radiomics signatures performed better than current TNM staging system and prognostic biomarker plasma EBV DNA, indicating that it could act as a novel and useful tool for the future management of advanced NPC. The prediction models built in this study are available on our website (www.radiomics.net.cn/platform.html).

One of the main challenges of our study is the extraction and selection of radiomics features, which were mostly associated with DFS, to develop radiomics signatures. Initially, 136 deep learning and 133 handcrafted features from each ROI were extracted. For deep learning feature extraction, we constructed 4 DCNNs and trained the weighted parameters through a patch-based strategy. After data augmentation, the size of training samples reached the order of ten thousand. Moreover, instead of using the DCNNs as the predictive tools directly or collecting the outputs of some layers as the features, we quantified the characteristics of the feature maps (Supplementary Fig. S10) from many aspects using the statistical algorithms to extract more comprehensive features,

Figure 5.

A–D, Kaplan–Meier survival curves between IC + CCRT and CCRT alone within the radiomics nomogram–defined high-risk group. IC, induction chemotherapy; CCRT, concurrent chemoradiotherapy.

as well as to improve stability and generalization. By using LASSO, 18 features were finally selected. It should be noted that LASSO is suitable for handling a mass of radiomics features with a relatively small sample size and avoid overfitting (25, 26). The radiomics features selected by LASSO are usually accurate, and the regression coefficients of extracted features are shrunk to zero during the process of model fitting, allowing the selection of features that are most strongly associated with DFS and making the model easier to interpret (27). Most importantly, LASSO allows radiomics signatures to be constructed by combining the selected features. In our study, the identified features were highly associated with DFS in both training and test sets.

As shown by our results, radiomics nomogram performed better than the clinical TNM staging system in risk stratification (C-index: 0.754 in the training set and 0.722 in the test set). There may be two major reasons for this: First, the TNM system was developed based on tumor size, lymph node status, and metastasis status, which only reflect anatomic information. Patients even with the same tumor stage could have different prognosis (28). Second, our signature features carry information on intratumor heterogeneity, which is an established prognostic factor (29, 30). Radiomics extracts the tumor imaging characteristics on medical images, providing a powerful means of interpreting intratumor heterogeneity; traditional clinical tumor stages cannot provide this information. This may be the main reason why the radiomics signatures and proposed nomogram performed better than TNM classification in predicting prognosis and stratifying risk.

As plasma EBV DNA has been widely identified as a reliable and useful biomarker at clinical practice (31–34), we then compared radiomics signatures with it. Intriguingly, the C-index of nomogram C is higher than that of nomogram B (0.785 vs. 0.683 in the training set, 0.771 vs. 0.671 in the test set), indicating that the prognostic ability of radiomics signatures was better than that of pre-DNA. Moreover, this conclusion was further supported by the results of time-independent ROC analysis (Fig. 5). When using these nomograms to predict the benefit of IC, nomogram C significantly performed better than both nomogram A and B. Taken these together, radiomics signatures were more powerful than plasma EBV DNA in prognosis prediction. Notably, we did not include EBV DNA into the radiomics nomogram initially because a few patients lost the data.

Currently, distant metastasis after radical radiotherapy has emerged as the predominant failure pattern for advanced NPC as IMRT has improved local and regional control greatly (3, 35). IC, given before radiotherapy, has been proven as a robust tool against this treatment failure (6–9). Although the most effective triplet IC regimen of docetaxel plus cisplatin with fluorouracil was delivered, the absolute benefit was observed only in 8% of the patients (8), meaning that more than 70% of patients could not benefit from IC. Meanwhile, these patients have to suffer from the severe toxicities and economic burden brought by IC. Given these, it is of great importance to identify patients with non-IC benefit. Although previous studies found that pre-DNA may play this role (12, 13), these studies were retrospective and had small sample size, making the results inconclusive. In our study, we

established PET/CT-based radiomics as a strong indicator for IC, i.e., high-risk patients could benefit from IC while low-risk could not. These findings provided a new insight into the future delivery of IC.

Compared with previous studies regarding radiomics (36–39), there were mainly four advantages in our study. First, the sample size was larger, thus improving the test power and the predictive ability of the model. Second, all patients were staged by PET/CT, which achieved higher diagnostic accuracy than conventional staging workup in NPC (40, 41). Undoubtedly, this accurate stage classification enables the robust prognostic prediction by radiomics signatures. Third, all patients received the standard care of CCRT with or without IC, which could reduce treatment-related bias on our conclusion. Finally, a deep learning method, named convolutional neural network (42), was applied for feature extraction. Deep learning radiomics method could learn features included in neural nets' hidden layers automatically from imaging data, and thus do not need object segmentation and hard-coded feature extraction (43). Limitations of our study should also be acknowledged. Follow-up duration may not be long enough; therefore, we constructed nomograms based on DFS. Study data were collected from a single center, and external validation may be warranted in the future. Moreover, potential patient selection biases confounded with radiomics signatures and outcomes may exist because IC treatment was not randomly assigned to participants as a result of retrospective nature, indicating that our results should be further validated in prospective and well-designed studies.

In summary, this study identified PET/CT-based radiomics as a powerful approach for predicting prognosis in patients with advanced NPC. The radiomics nomogram successfully stratified patients into high-risk and low-risk groups for all endpoints, and thereby may act as a potential tool for individualized treatment strategies: high-risk patients should receive more intensity treatment like IC plus CCRT; for low-risk patients, CCRT may be enough. Future prospective studies with external validation are needed to validate our findings.

References

1. Wei KR, Zheng RS, Zhang SW, Liang ZH, Li ZM, Chen WQ. Nasopharyngeal carcinoma incidence and mortality in China, 2013. Chin J Cancer 2017;36:90.
2. Yi JL, Gao L, Huang XD, Li SY, Luo JW, Cai WM, et al. Nasopharyngeal carcinoma treated by radical radiotherapy alone: ten-year experience of a single institution. Int J Radiat Oncol Biol Phys 2006;65:161–8.
3. Sun X, Su S, Chen C, Han F, Zhao C, Xiao W, et al. Long-term outcomes of intensity-modulated radiotherapy for 868 patients with nasopharyngeal carcinoma: an analysis of survival and treatment toxicities. Radiother Oncol 2014;110:398–403.
4. Mao YP, Xie FY, Liu LZ, Sun Y, Li L, Tang LL, et al. Re-evaluation of 6th edition of AJCC staging system for nasopharyngeal carcinoma and proposed improvement based on magnetic resonance imaging. Int J Radiat Oncol Biol Phys 2009;73:1326–34.
5. Pan JJ, Ng WT, Zong JF, Chan LL, O'Sullivan B, Lin SJ, et al. Proposal for the 8th edition of the AJCC/UICC staging system for nasopharyngeal cancer in the era of intensity-modulated radiotherapy. Cancer 2016;122:546–58.
6. Cao SM, Yang Q, Guo L, Mai HQ, Mo HY, Cao KJ, et al. Neoadjuvant chemotherapy followed by concurrent chemoradiotherapy versus concurrent chemoradiotherapy alone in locoregionally advanced nasopharyngeal carcinoma: a phase III multicentre randomised controlled trial. Eur J Cancer 2017;75:14–23.
7. Hui EP, Ma BB, Leung SF, King AD, Mo F, Kam MK, et al. Randomized phase II trial of concurrent cisplatin-radiotherapy with or without neoadjuvant docetaxel and cisplatin in advanced nasopharyngeal carcinoma. J Clin Oncol 2009;27:242–9.
8. Sun Y, Li WF, Chen NY, Zhang N, Hu GQ, Xie FY, et al. Induction chemotherapy plus concurrent chemoradiotherapy versus concurrent chemoradiotherapy alone in locoregionally advanced nasopharyngeal carcinoma: a phase 3, multicentre, randomised controlled trial. Lancet Oncol 2016;17:1509–20.
9. Frikha M, Auperin A, Tao Y, Elloumi F, Toumi N, Blanchard P, et al. A randomized trial of induction docetaxel-cisplatin-5FU followed by concomitant cisplatin-RT versus concomitant cisplatin-RT in nasopharyngeal carcinoma (GORTEC 2006–02). Ann Oncol 2017;29:731–36.
10. Lan XW, Xiao Y, Zou XB, Zhang XM, OuYang PY, Xie FY. Outcomes of adding induction chemotherapy to concurrent chemoradiotherapy for stage T3N0–1 nasopharyngeal carcinoma: a propensity-matched study. OncoTargets Ther 2017;10:3853–60.
11. Wu LR, Yu HL, Jiang N, Jiang XS, Zong D, Wen J, et al. Prognostic value of chemotherapy in addition to concurrent chemoradiotherapy in T3–4N0-1 nasopharyngeal carcinoma: a propensity score matching study. Oncotarget 2017;8:76807–15.
12. Guo SS, Tang LQ, Chen QY, Zhang L, Liu LT, Guo L, et al. Induction chemotherapy followed by concurrent chemoradiotherapy versus concurrent chemoradiotherapy alone in stage III-IVb nasopharyngeal carcinoma patients with Epstein-Barr virus DNA ≥ 4000 copies/ml: a matched study. Oncotarget 2016;7:29739–48.

13. Peng H, Chen L, Li WF, Guo R, Zhang Y, Zhang F, et al. Prognostic value of neoadjuvant chemotherapy in locoregionally advanced nasopharyngeal carcinoma with low pre-treatment Epstein-Barr virus DNA: a propensity-matched analysis. J Cancer 2016;7:1465–71.

14. Le QT, Jones CD, Yau TK, Shirazi HA, Wong PH, Thomas EN, et al. A comparison study of different PCR assays in measuring circulating plasma Epstein-Barr virus DNA levels in patients with nasopharyngeal carcinoma. Clin Cancer Res 2005;11:5700–7.

15. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.

16. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441–6.

17. Dong D, Tang L, Li ZY, Fang MJ, Gao JB, Shan XH, 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.

18. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60–88.

19. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annu Rev Biomed Eng 2017;19:221–48.

20. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans Med Imaging 2016;35:1207–16.

21. Cordier N, Delingette H, Ayache N. A patch-based approach for the segmentation of pathologies: application to glioma labelling. IEEE Trans Med Imaging 2016;35:1066–76.

22. Ning Z, Luo J, Li Y, Han S, Feng Q, Xu Y, et al. Pattern classification for gastrointestinal stromal tumors by integration of radiomics and deep convolutional features. IEEE J Biomed Health Inform 2018;23:1181–91.

23. Delbeke D, Coleman RE, Guiberteau MJ, Brown ML, Royal HD, Siegel BA, et al. Procedure guideline for tumor imaging with 18F-FDG PET/CT 1.0. J Nucl Med 2006;47:885–95.

24. Rotolo F, Pignon JP, Bourhis J, Marguet S, Leclercq J, Tong Ng W, et al. Surrogate end points for overall survival in loco-regionally advanced nasopharyngeal carcinoma: an individual patient data meta-analysis. J Natl Cancer Inst 2017;109. pii: djw239.

25. Gui J, Li H. Penalized Cox regression analysis in the high-dimensional and low-sample size settings, with applications to microarray gene expression data. Bioinformatics 2005;21:3001–8.

26. Hepp T, Schmid M, Gefeller O, Waldmann E, Mayr A. Approaches to regularized regression – a comparison between gradient boosting and the lasso. Methods Inf Med 2016;55:422–30.

27. Ndhlovu ZM, Chibnik LB, Proudfoot J, Vine S, McMullen A, Cesa K, et al. High-dimensional immunomonitoring models of HIV-1-specific CD8 T-cell responses accurately identify subjects achieving spontaneous viral control. Blood 2013;121:801–11.

28. Mao YP, Liang SB, Liu LZ, Chen Y, Sun Y, Tang LL, et al. The N staging system in nasopharyngeal carcinoma with radiation therapy oncology group

29. Friemel J, Rechsteiner M, Frick L, Bohm F, Struckmann K, Egger M, et al. Intratumor heterogeneity in hepatocellular carcinoma. Clin Cancer Res 2015;21:1951–61.

30. 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.

31. Chan AT, Lo YM, Zee B, Chan LY, Ma BB, Leung SF, et al. Plasma Epstein-Barr virus DNA and residual disease after radiotherapy for undifferentiated nasopharyngeal carcinoma. J Natl Cancer Inst 2002;94:1614–9.

32. Lin JC, Wang WY, Chen KY, Wei YH, Liang WM, Jan JS, et al. Quantification of plasma Epstein-Barr virus DNA in patients with advanced nasopharyngeal carcinoma. N Engl J Med 2004;350:2461–70.

33. Lo YM, Chan AT, Chan LY, Leung SF, Lam CW, Huang DP, et al. Molecular prognostication of nasopharyngeal carcinoma by quantitative analysis of circulating Epstein-Barr virus DNA. Cancer Res 2000;60:6878–81.

34. Peng H, Guo R, Chen L, Zhang Y, Li WF, Mao YP, et al. Prognostic impact of plasma Epstein-Barr virus DNA in patients with nasopharyngeal carcinoma treated using intensity-modulated radiation therapy. Sci Rep 2016;6:22000.

35. Lai SZ, Li WF, Chen L, Luo W, Chen YY, Liu LZ, et al. How does intensity-modulated radiotherapy versus conventional two-dimensional radiotherapy influence the treatment results in nasopharyngeal carcinoma patients? Int J Radiat Oncol Biol Phys 2011;80:661–8.

36. Huang Y, Liu Z, He L, Chen X, Pan D, Ma Z, et al. Radiomics signature: a potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer. Radiology 2016;281:947–57.

37. Huang YQ, Liang CH, He L, Tian J, Liang CS, 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.

38. Li H, Zhu Y, Burnside ES, Drukker K, Hoadley KA, Fan C, et al. MR imaging radiomics signatures for predicting the risk of breast cancer recurrence as given by research versions of MammaPrint, Oncotype DX, and PAM50 Gene Assays. Radiology 2016;281:382–91.

39. 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.

40. Gordin A, Golz A, Daitzchman M, Keidar Z, Bar-Shalom R, Kuten A, et al. Fluorine-18 fluorodeoxyglucose positron emission tomography/computed tomography imaging in patients with carcinoma of the nasopharynx: diagnostic accuracy and impact on clinical management. Int J Radiat Oncol Biol Phys 2007;68:370–6.

41. Peng H, Chen L, Tang LL, Li WF, Mao YP, Guo R, et al. Significant value of (18)F-FDG-PET/CT in diagnosing small cervical lymph node metastases in patients with nasopharyngeal carcinoma treated with intensity-modulated radiotherapy. Chin J Cancer 2017;36:95.

42. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44.

43. Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, et al. Convolutional neural networks for medical image analysis: full training or fine tuning? IEEE Trans Med Imaging 2016;35:1299–312.

guidelines for lymph node levels based on magnetic resonance imaging. Clin Cancer Res 2008;14:7497–503.

Clinical Cancer Research

Prognostic Value of Deep Learning PET/CT-Based Radiomics: Potential Role for Future Individual Induction Chemotherapy in Advanced Nasopharyngeal Carcinoma

Hao Peng, Di Dong, Meng-Jie Fang, et al.

Updated version	Access the most recent version of this article at: doi:10.1158/1078-0432.CCR-18-3065
Supplementary Material	Access the most recent supplemental material at: http://clincancerres.aacrjournals.org/content/suppl/2019/04/11/1078-0432.CCR-18-3065.DC1

Cited articles	This article cites 42 articles, 9 of which you can access for free at: http://clincancerres.aacrjournals.org/content/25/14/4271.full#ref-list-1

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, use this link http://clincancerres.aacrjournals.org/content/25/14/4271. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.