A radiomics prognostic scoring system for predicting progression-free survival in patients with stage IV non-small cell lung cancer treated with platinum-based chemotherapy

Lan He[1*], Zhenhui Li[1,2*], Xin Chen[3*], Yanqi Huang[1,4], Lixu Yan[5], Changhong Liang[1], Zaiyi Liu[1]

[1]Department of Radiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, China; [2]Department of Radiology, the Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Yunnan Cancer Center, Kunming 650118, China; [3]Department of Radiology, Guangzhou First People's Hospital, School of Medicine, South China University of Technology, Guangzhou 510120, China; [4]The Second School of Clinical Medicine, Southern Medical University, Guangzhou 510515, China; [5]Department of Pathology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, China
*These authors contributed equally to this work.
Correspondence to: Zaiyi Liu, PhD. Department of Radiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, No. 106 Zhongshan Er Road, Guangzhou 510080, China. Email: zyliu@163.com; Changhong Liang, PhD. Department of Radiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, No. 106 Zhongshan Er Road, Guangzhou 510080, China. Email: liangchanghong@gdph.org.cn.

Abstract

Objective: To develop and validate a radiomics prognostic scoring system (RPSS) for prediction of progression-free survival (PFS) in patients with stage IV non-small cell lung cancer (NSCLC) treated with platinum-based chemotherapy.

Methods: In this retrospective study, four independent cohorts of stage IV NSCLC patients treated with platinum-based chemotherapy were included for model construction and validation (Discovery: n=159; Internal validation: n=156; External validation: n=81, Mutation validation: n=64). First, a total of 1,182 three-dimensional radiomics features were extracted from pre-treatment computed tomography (CT) images of each patient. Then, a radiomics signature was constructed using the least absolute shrinkage and selection operator method (LASSO) penalized Cox regression analysis. Finally, an individualized prognostic scoring system incorporating radiomics signature and clinicopathologic risk factors was proposed for PFS prediction.

Results: The established radiomics signature consisting of 16 features showed good discrimination for classifying patients with high-risk and low-risk progression to chemotherapy in all cohorts (All P<0.05). On the multivariable analysis, independent factors for PFS were radiomics signature, performance status (PS), and N stage, which were all selected into construction of RPSS. The RPSS showed significant prognostic performance for predicting PFS in discovery [C-index: 0.772, 95% confidence interval (95% CI): 0.765−0.779], internal validation (C-index: 0.738, 95% CI: 0.730−0.746), external validation (C-index: 0.750, 95% CI: 0.734−0.765), and mutation validation (C-index: 0.739, 95% CI: 0.720−0.758). Decision curve analysis revealed that RPSS significantly outperformed the clinicopathologic-based model in terms of clinical usefulness (All P<0.05).

Conclusions: This study established a radiomics prognostic scoring system as RPSS that can be conveniently used to achieve individualized prediction of PFS probability for stage IV NSCLC patients treated with platinum-based chemotherapy, which holds promise for guiding personalized pre-therapy of stage IV NSCLC.

Keywords: Non-small cell lung cancer; radiomics; prognostic scoring system; progression-free survival; platinum-based chemotherapy

www.cjcrcn.org   *Chin J Cancer Res* 2021;33(5):592-605

Introduction

Lung cancer remains the most commonly diagnosed cancer, with high incidence and mortality rates in both males and females (1,2). Non-small cell lung cancer (NSCLC) accounts for 85% of all lung cancer cases. Nearly two-thirds of NSCLC patients have unresectable metastatic disease upon diagnosis (3). For patients without positive markers (e.g., EGFR/ALK/ROS1), or patients with low programmed death ligand 1 (PD-L1) expression, first-line treatment of platinum-based chemotherapy is recommended for stage IV NSCLC according to the American Society of Clinical Oncology guideline (4,5). Even when a curative treatment approach is feasible, the survival outcomes among patients are various (6). Thus, a method for individually predicting the progression probability of patients can determine whether to continue, escalate, discontinue, or change a patient's therapeutic schedule so as to improve therapeutic outcomes and potentially reduce therapeutic toxicity (7). However, how to assess the individual patient's potential progression probability to platinum-based chemotherapy remains very challenging.

To date, clinical and genetic factors have been extensively investigated to select patients who may benefit from specific chemotherapies (8). It is increasingly suggested that tailored individual chemotherapy, based on molecular biomarkers, represents a novel avenue for NSCLC treatment (9). However, these methods are invasive and limited by the fact that they are acquired at a single timepoint and from a single anatomical location (10,11). To address this topic, radiomics is an emerging field that involves throughput conversion of quantitative automated imaging features into mineable data (12-14). With its capability of simultaneous modelling of both multiple image phenotypic features and clinical factor effects, radiomics analyses aim to provide comprehensive quantification of tumor phenotypes, which can be incorporated into prediction and prognostic models (15-19). With advances in radiomics, tumor characterization is not just limited to anatomy, but it also reveals cellular and genomic level information that is quantifiable as an imaging phenotype (14). However, this type of radiomics

prognostic scoring system (RPSS) for predicting disease progression probability to platinum-based chemotherapy of stage IV NSCLC is currently lacking (20).

Hence, in this study we aimed to establish an effective RPSS to individually predict the progression probability to platinum-based chemotherapy for stage IV NSCLC patients with the aim of providing valuable information for personalized medicine.

Materials and methods

Patients

This retrospective study had ethical approval, and the informed consent requirement was waived. In this multi-cohort study, we reviewed 159 patients with stage IV NSCLC treated with first-line platinum-based chemo-therapy at Guangdong Provincial People's Hospital between November 2007 and March 2012 as the discovery cohort, and 156 patients treated from April 2012 to June 2016 in the same institution as the internal validation cohort. Another independent cohort of 81 patients was included from November 2007 to June 2016 in the Yunnan Cancer Hospital as the external validation cohort for this study (*Figure 1*).

All patients were treated according to the American Society of Clinical Oncology guideline (5). The inclusion criteria were as follows: 1) Patients who were histological or cytological confirmed NSCLC; 2) Patients diagnosed with stage IV according to the TNM classification system of the American Joint Committee on Cancer (21); 3) Patients who were not harboring EGFR-sensitizing mutations or *ALK* gene rearrangements; 4) Patients with low PD-L1 expression; 5) Patients underwent at least 4 cycles of platinum-based chemotherapy as first-line treatment; 6) Patients received no radiotherapy and anticancer therapy before chemotherapy; and 7) Patients underwent a full follow-up. Patients with a history of other malignancies or surgery resection were excluded from this study. For each patient, the clinical variables that might correlate with the prognostic outcomes were also recorded, including age, gender, stage at diagnosis, smoking status, performance status (PS), histological type, and the administered therapeutic regimen.
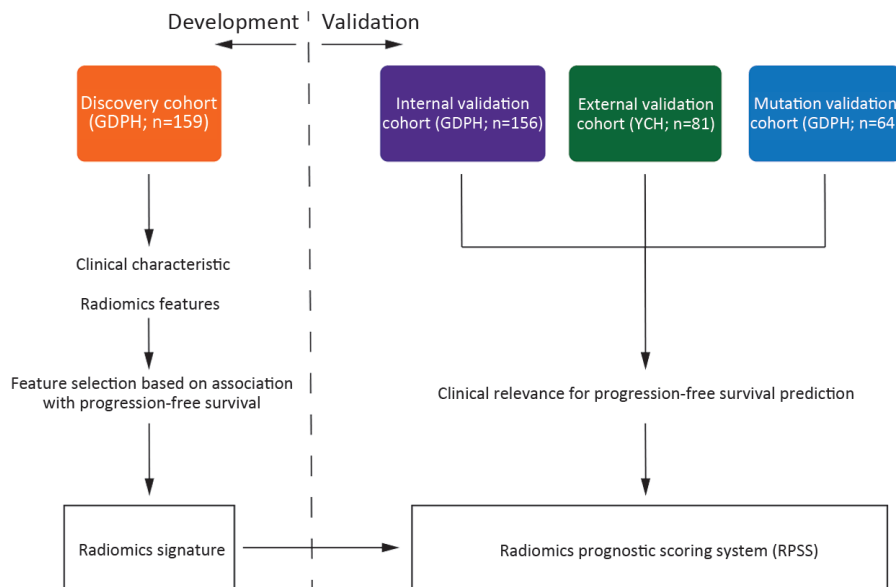
Figure 1 Study design flowchart. GDPH, Guangdong Provincial People's Hospital; YCH, Yunnan Cancer Hospital; RPSS, Radiomics prognostic scoring system.

Furthermore, with the aim of extending the prognostic value of RPSS we established for predicting progression-free survival (PFS), we included in the analysis a mutation validation cohort of stage IV NSCLC patients with EGFR-sensitizing mutation who only received platinum-based chemotherapy as first-line treatment from November 2007 to June 2016 in Guangdong Provincial People's Hospital.

Treatment procedure

All patients only received cisplatin or carboplatin based double chemotherapy as first-line treatment (5). The other administrated regimens included pemetrexed, gemcitabine, paclitaxel, and docetaxel. The details of drug treatment are provided in *Supplementary materials*. During treatment, routine blood tests were conducted, and the liver and renal functions were monitored.

Follow-up and assessment of PFS

Each patient was followed at 4-week intervals with a CT of the chest/abdomen, physical examination, and routine laboratory tests. Additional CT or magnetic resonance imaging was routinely performed if extrapulmonary metastasis was suspected. PFS was defined as the time interval between the date of initiation of chemotherapy and either disease progression or death, which was censored at the date of death from other cause or the date of the last follow-up visit for the progression-free patients.

CT image acquisition, interpretation, and radiomics feature extraction

All patients performed contrast-enhanced CT scans within 2 weeks before chemotherapy start. The details of CT image acquisition are described in *Supplementary materials*. Then, all acquired CT images were gathered for tumor segmentation and radiomics feature extraction. The primary tumors from each patient were manually segmented by an experienced radiologist, with more than 10 years of clinical experience in chest CT interpretation. If multiple disconnected tumor volumes were found, the largest by volume was chosen for features extraction.

To ensure the reproducibility and accuracy, 80 patients were randomly selected for a reproducibility analysis using the inter- and intra-class correlation coefficients (ICCs). The details of ICCs for the reproducibility analysis of radiomics features extraction are described in the *Supplementary materials*. An ICC greater than 0.75 indicated a good agreement.

Finally, we programmed the radiomics analysis algorithms to automatically extract the radiomics features from manually segmented tumor region. All medical images were resampled to the same voxel of 1 mm^3 × 1 mm^3 × 1 mm^3 by using cubic spline interpolation in our study. The voxel intensity within the region of interest was discretized to the same range of 64 intensity values. The details of the radiomics analysis algorithms are summarized

in *Supplementary materials*. In total, 1,182 quantitative radiomics features were extracted from each patient's contrast-enhanced chest CT images, which included tumor intensity, texture features, wavelet features, and Gabor features. All features were compliant with the Image Biomarker Standardization Initiative (IBSI) (22).

Feature selection and radiomics signature construction

The radiomics features data were firstly normalized with z-score normalization (23), and were filtered based on their independence from other features, as determined by the Pearson's correlation coefficient among the features (cut-off value of 0.9) (24). Then, the feature selection was done in the discovery cohort using the least absolute shrinkage and selection operator method (LASSO) penalized Cox proportional regression (25,26). Finally, the radiomics signature was built through the combination of selected features weighted by their respective coefficients.

All patients were stratified into high-risk and low-risk progression groups by the established signature, which was achieved by using time-dependent receiver operating characteristic (ROC) curves (27) based on Kaplan-Meier survival analyses and log-rank test (28). The significant difference in PFS was analyzed to investigate the clinical benefits across different progression subgroups by the Kaplan-Meier survival analysis and the Cox proportional hazards regression model.

Development and validation of RPSS

First, the established radiomics signature and potential clinicopathologic characteristics, including age, gender, smoking history, histological type, PS, T stage, N stage, the site and volume of metastases, and chemotherapy regimens, were involved in multivariable Cox regression analysis. Then, a final model selection was performed by the backward step-down selection process with Akaike information criterion (29). Subsequently, RPSS was calculated based on the result of multivariate analysis for individualized probability prediction of PFS. The performance of RPSS was measured by the concordance index (C-index), the integrated area under the ROC curve (iAUC), and the integrated Brier score (iBS) (30), and also was assessed by comparing the predicted versus observed survival probability using the calibration curve (31,32). In addition, the prognostic predictive power of the RPSS was further validated in the internal validation, external validation, and mutation validation cohort.

Clinical usefulness

To determine the clinical benefits of this radiomics signature, we established another clinicopathologic-based model with only clinicopathological characteristics. Then, comparisons between RPSS and clinicopathologic-based model were evaluated by C-index, iAUC, iBS, net reclassification improvement (NRI), and integrated discrimination improvement (IDI) (33,34). The larger C-index which resulted in the positive value for NRI and IDI, the more accurate was the prognostic prediction for predicting PFS (35). For iAUC, the higher value indicated better performance; while for iBS, the lower value indicated better performance. Finally, a decision curve analysis was performed for determining the clinical usefulness of RPSS by quantifying the net benefits at different threshold probabilities (36-38).

Statistical analysis

The Kaplan-Meier method was used to calculate PFS, and the differences in PFS were compared using a log-rank test. The data were analyzed using R software (Version 3.2.3; R Foundation for Statistical Computing, Vienna, Austria), and the details of all R packages used in this study are described in *Supplementary materials*. The results are considered statistically significant at $P<0.05$ using a two-tailed test.

Results

Characteristics of study population

The demographics and clinical characteristics of all patient cohorts are shown in *Table 1*. The median follow-up time was 5.73 [interquartile range (IQR): 3.53−8.00] months in discovery cohort, 5.47 (IQR: 3.33−7.93) months in internal validation cohort, 7.07 (IQR: 4.03−12.02) months in external validation cohort, and 6.04 (IQR: 3.49−8.83) months in mutation cohort, respectively.

Radiomics analysis profiling of PFS

Based on the reproducibility and the accuracy of radiomics features extraction, we found that the intra- and inter-observer reproducibility of features extraction was satisfactory. The inter-observer ICCs calculated ranged from 0.751 to 0.957 based on Reader 1's and Reader 2's first extraction features, and the intra-observer ICCs calculated ranged from 0.763 to 0.911 based on Reader 1's

Table 1 Demographic and pathological characteristics of all patient cohorts

Characteristics	n (%)			
	Discovery cohort (N=159)	Internal validation cohort (N=156)	External validation cohort (N=81)	Mutation validation cohort (N=64)
Age (year) [median (IQR)]	58 (52, 66)	59 (49, 65)	55 (49, 63)	57 (46, 65)
Gender				
Male	120 (75.5)	119 (76.3)	48 (59.3)	40 (62.5)
Female	39 (24.5)	37 (23.7)	33 (40.7)	24 (37.5)
Smoking status				
Yes	72 (45.3)	76 (48.7)	30 (37.0)	37 (57.8)
No	87 (54.7)	80 (51.3)	51 (63.0)	27 (42.2)
Histological type				
Squamous cell carcinoma	33 (20.8)	24 (15.4)	10 (12.3)	4 (6.3)
Adenocarcinoma	120 (75.5)	115 (73.7)	71 (87.7)	59 (92.2)
Others	6 (3.8)	17 (10.9)	0 (0)	1 (1.6)
PS				
<2	124 (78.0)	126 (80.8)	78 (96.3)	60 (93.8)
≥2	35 (22.0)	30 (19.2)	3 (3.7)	4 (6.3)
T stage				
T1	7 (4.4)	20 (12.8)	6 (7.4)	5 (7.8)
T2	58 (36.5)	43 (27.6)	31 (38.3)	23 (35.9)
T3	34 (21.4)	44 (28.2)	12 (14.8)	12 (18.8)
T4	60 (37.7)	49 (31.4)	32 (39.5)	24 (37.5)
N stage				
N0	20 (12.6)	15 (9.6)	12 (14.8)	9 (14.1)
N1	14 (8.8)	12 (7.7)	12 (14.8)	8 (12.5)
N2	61 (38.4)	64 (41.0)	38 (46.9)	22 (34.4)
N3	64 (40.3)	65 (41.7)	19 (23.5)	25 (39.1)
Chemotherapy regimens				
DDP/CBP+GEM/PEM	128 (80.5)	117 (75.0)	54 (66.7)	50 (78.1)
DDP/CBP+TAX/TXT/DOC	31 (19.5)	39 (25.0)	27 (33.3)	14 (21.9)

IQR, interquartile range; PS, performance status; DDP, cisplatin; CBP, carboplatin; GEM, gemcitabine; PEM, pemetrexed; TAX, paclitaxel; TXT, taxinol; DOC, docetaxel.

twice feature extractions. A total of 1,182 radiomics features were extracted from the pre-treatment CT images of each patient. After filtering by correlation analysis, 100 radiomics features were used for the subsequent analysis (*Supplementary Figure S1*). Then, 16 key features with non-zero coefficients in the LASSO Cox regression model were selected (*Supplementary Figure S2*). The radiomics signature was constructed based on the regression analysis with a corresponding value calculated for each patient (*Supplementary materials*). The cut-off value of radiomics signature was −0.734, which was generated by a time-dependent ROC curve (*Supplementary Figure S3*), and used

for dividing patients into high-risk and low-risk progression groups. The Kaplan-Meier curves clearly showed different prognostic strata in PFS between the high-risk and low-risk progression subgroups in all cohorts, with a high statistical significance (log-rank P<0.01 in all cases, *Figure 2*). Lower value of radiomics signature was associated with improved PFS in discovery (HR: 5.829, 95% CI: 3.532−9.618), internal validation (HR: 3.701, 95% CI: 2.262−6.054), external validation (HR: 3.077, 95% CI: 1.617−5.857), and mutation validation cohort (HR: 3.050, 95% CI: 1.390−6.693). The PFS and disease progression rate in the high-risk progression and low-risk progression
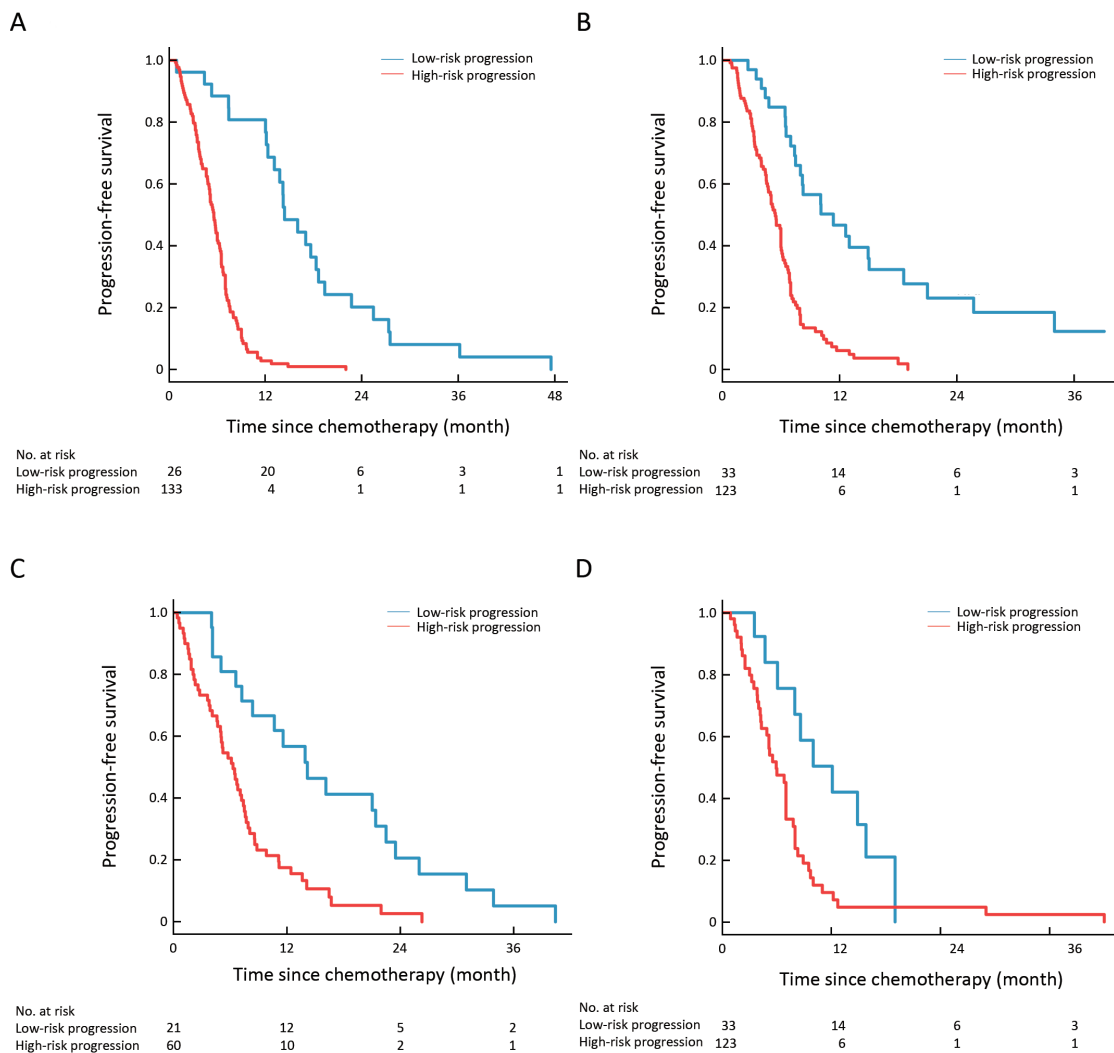
Figure 2 Kaplan-Meier survival curves for PFS in (A) discovery cohort (P<0.001, HR: 5.829, 95% CI: 3.532−9.618); (B) internal validation cohort (P<0.001, HR: 3.701, 95% CI: 2.262−6.054); (C) external validation cohort (P=0.001, HR: 3.077, 95% CI: 1.617−5.857); and (D) mutation validation cohort (P=0.008, HR: 3.050, 95% CI: 1.390−6.693) according to radiomics signature to classify stage IV NSCLC patients with high-risk and low-risk progression treated by platinum-based chemotherapy. PFS, progression-free survival; HR, hazard ratio; 95% CI, 95% confidence interval; NSCLC, non-small cell lung cancer.

group in all cohort are listed in *Table 2*.

RPSS for predicting of PFS

The multivariate analysis demonstrated that PS, N stage, and radiomics signature were independent risk factors for prediction of PFS (*Table 3*). RPSS was computed per patient according to the multivariate analyses, using a linear combination of selected independent risk factors weighted by their respective coefficients (*Table 3*). The C-index of RPSS for prediction of PFS were 0.772 (95% CI: 0.765−0.779) in discovery, 0.738 (95% CI: 0.730−0.746) in

internal validation, 0.750 (95% CI: 0.734−0.765) in external validation, and 0.739 (95% CI: 0.720−0.758) in mutation validation. The iAUC was 0.760 in discovery, 0.700 in internal validation, 0.741 in external validation, and 0.705 in mutation validation cohort. AUC at 1-year PFS was 0.956 (95% CI: 0.914−0.998) in discovery, 0.825 (95% CI: 0.725−0.925) in internal validation, 0.819 (95% CI: 0.715−0.923) in external validation, and 0.818 (95% CI: 0.649−0.986) in mutation validation cohort. And the iBS was 0.038 in discovery, 0.072 in internal validation, 0.070 in external validation, and 0.054 in mutation validation cohort. The calibration plot for the probability of survival

598

He et al. A RPSS for patients with stage IV NSCLC

Table 2 PFS and disease progression rate in high-risk progression and low-risk progression group

Parameter	Discovery cohort			Internal validation cohort			External validation cohort			Mutation validation cohort		
	High-risk progression	Low-risk progression	Total	High-risk progression	Low-risk progression	Total	High-risk progression	Low-risk progression	Total	High-risk progression	Low-risk progression	Total
No. of patients [n (%)]	133 (83.6)	26 (16.4)	159	123 (78.8)	33 (21.2)	156	60 (74.1)	21 (25.9)	81	51 (79.7)	13 (20.3)	64
1-year PFS [median (IQR)] (month)	5.1 (3.4, 7.0)	14.3 (11.8, 20.2)	5.7 (3.5, 8.0)	4.8 (3.2, 6.8)	10.1 (6.5, 16.9)	5.5 (3.3, 7.9)	6.2 (2.6, 8.6)	13.9 (6.9, 23.0)	6.0 (3.5, 8.8)	5.2 (2.8, 8.0)	10.0 (5.4, 15.3)	6.0 (3.5, 8.8)
No. of progression patients [n (%)]												
At 10 months	117 (88.0)	12 (46.2)	129 (81.1)	95 (77.2)	15 (45.5)	110 (70.5)	46 (76.7)	8 (38.1)	54 (66.7)	41 (80.4)	6 (46.2)	47 (73.4)
At 1 year	120 (90.2)	25 (96.2)	145 (91.2)	103 (83.7)	25 (75.8)	128 (82.1)	55 (91.7)	20 (95.2)	75 (92.6)	45 (88.2)	11 (84.6)	56 (87.5)

PFS, progression-free survival; IQR, interquartile range.

Table 3 Multivariate Cox regression analysis of models in discovery cohort

Variable	RPSS			Clinicopathologic-based model		
	Coefficient	HR (95% CI)	P	Coefficient	HR (95% CI)	P
PS	0.474	1.606 (1.058, 2.438)	0.026	0.632	1.881 (1.260, 2.810)	0.002
N stage	0.214	1.239 (1.033, 1.485)	0.021	0.337	1.401 (1.169, 1.679)	<0.001
Radiomics signature	1.296	3.655 (2.767, 4.827)	<0.001	–	–	–

PS, performance status; RPSS, radiomics prognostic scoring system; HR, hazard ratio; 95% CI, 95% confidence interval. The score value of RPSS was calculated as follow: Score = 0.474 × performance status + 0.213 × N stage + 1.296 × radiomcis signature.

at the 10-month or 1-year PFS showed an optimal agreement between the prediction by RPSS and the actual observation (P<0.05, *Figure 3*).

Comparison of predictive accuracy for PFS between RPSS and clinicopathologic-based model

As shown in *Table 3*, hazard ratios of the radiomics signature in the RPSS for survival were higher than that for other factors (PS and N stage). By removing signature from RPSS, a clinicopathologic-based model, with only the clinicopathologic characteristics (PS and N stage), was constructed. The predictive power for prognosis of PFS between RPSS and clinicopathologic-based model was compared. The C-index values for PFS prediction by clinicopathologic-based model were 0.686 (95% CI: 0.678−0.694) in discovery, 0.649 (95% CI: 0.639−0.659) in internal validation, 0.519 (95% CI: 0.502−0.536) in external validation, and 0.674 (95% CI: 0.655−0.693) in mutation validation cohort, respectively, which were significantly lower than the C-index by RPSS (All P<0.01). The iAUC was 0.616 in discovery, 0.638 in internal validation, 0.544 in external validation, and 0.659 in mutation validation cohort, which were significantly lower than the iAUC by RPSS. And the iBS was 0.055 in

discovery, 0.084 in internal validation, 0.101 in external validation, and 0.056 in mutation validation cohort, which were significantly higher than the iBS by RPSS (of note: for iAUC and C-index, the higher the better, while for iBS, the lower the better) (*Figure 4,5*). Consequently, these results suggested that the RPSS was a more accurate and useful tool for the prediction of PFS regarding NRI of 0.732 (95% CI: 0.497−0.834) in discovery, 0.480 (95% CI: 0.203−0.676) in internal validation, 0.463 (95% CI: 0.131−0.674) in external validation, and 0.287 (95% CI: 0.232−0.759) in mutation validation cohort (All P<0.05), and IDI of 0.335 (95% CI: 0.210−0.459) in discovery, 0.113 (95% CI: 0.036−0.184) in internal validation, 0.263 (95% CI: 0.076−0.397) in external validation, and 0.119 (95% CI: 0.002−0.248) in mutation validation cohort (All P<0.05).

Clinical usefulness of RPSS

RPSS displayed as a nomogram is shown in *Figure 6A*, which used as an easy-to-use tool for probability scoring of 10-month and 1-year PFS. The decision curve analysis for RPSS provided a range of risk thresholds that yielded a positive net benefit compared with the clinicopathologic-based model, which is shown in *Figure 6B*. The RPSS provided the largest overall net benefit for predicting PFS
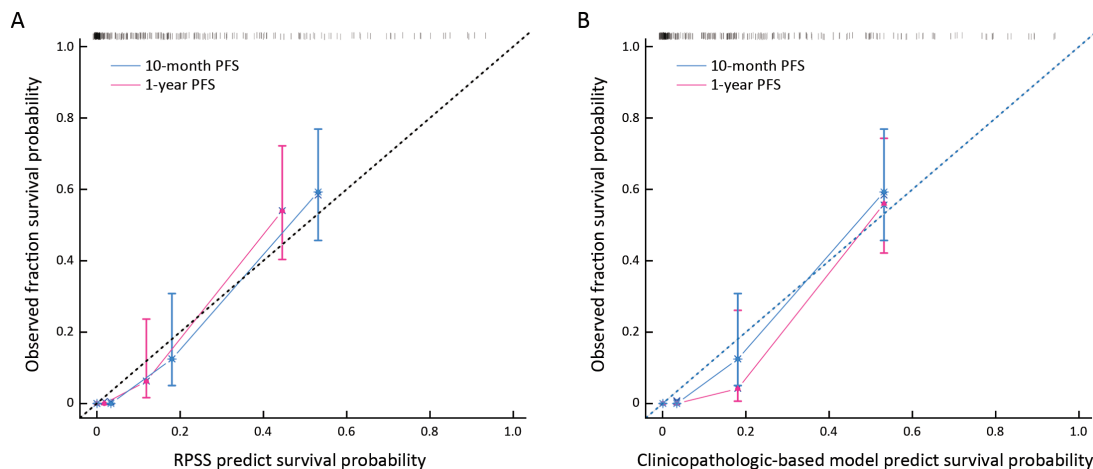
Figure 3 Calibration curves of RPSS and clinicopathologic-based model in terms of agreement between predicted and observed 1-year PFS and 10-month PFS. Calibration curve for RPSS (A) and clinicopathologic-based model (B) to predict survival probability. RPSS, radiomics prognostic scoring system; PFS, progression-free survival.
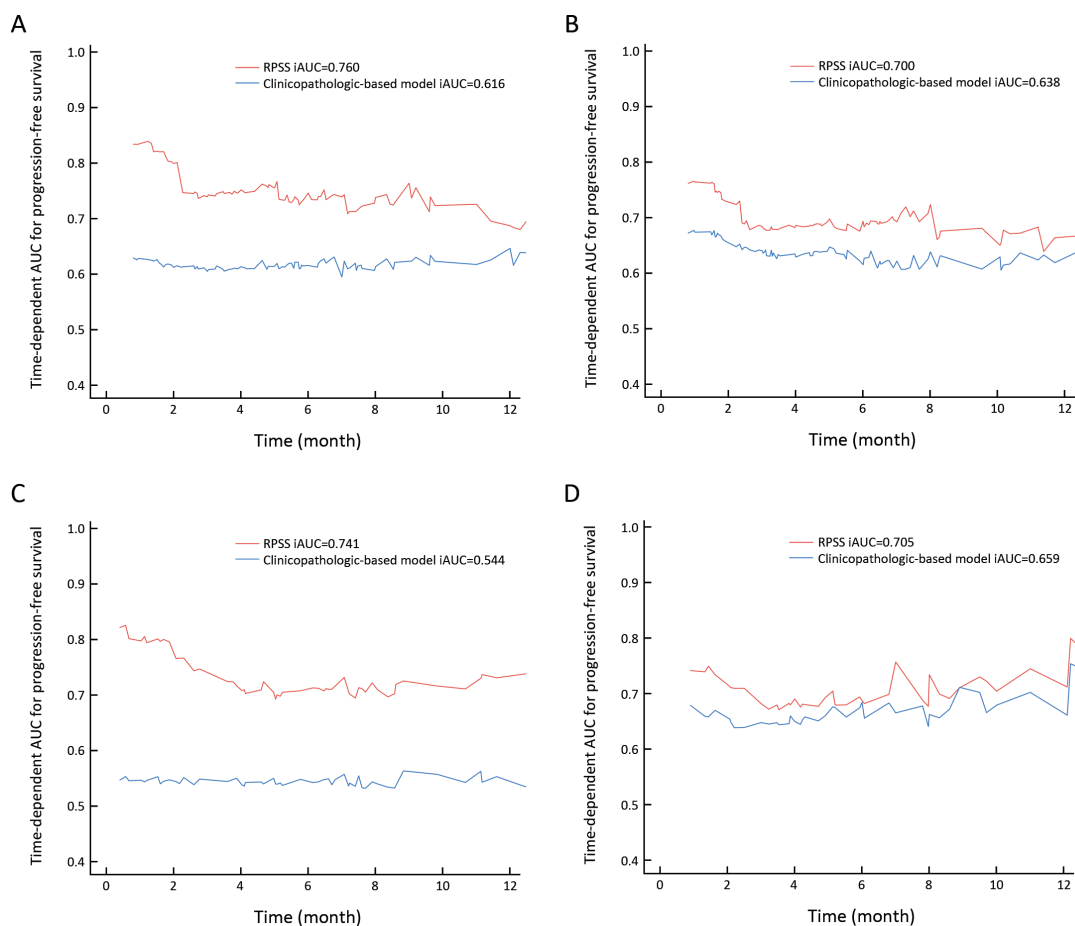

Figure 4 Prognostic value of RPSS and clinicopathologic-based model in terms of iAUC. (A) Discovery cohort; (B) internal validation cohort; (C) external validation cohort; and (D) mutation validation cohort. The higher value of iAUC indicated the better prognostic performance of clinical model. RPSS, radiomics prognostic scoring system; iAUC, integrated area under the curve.
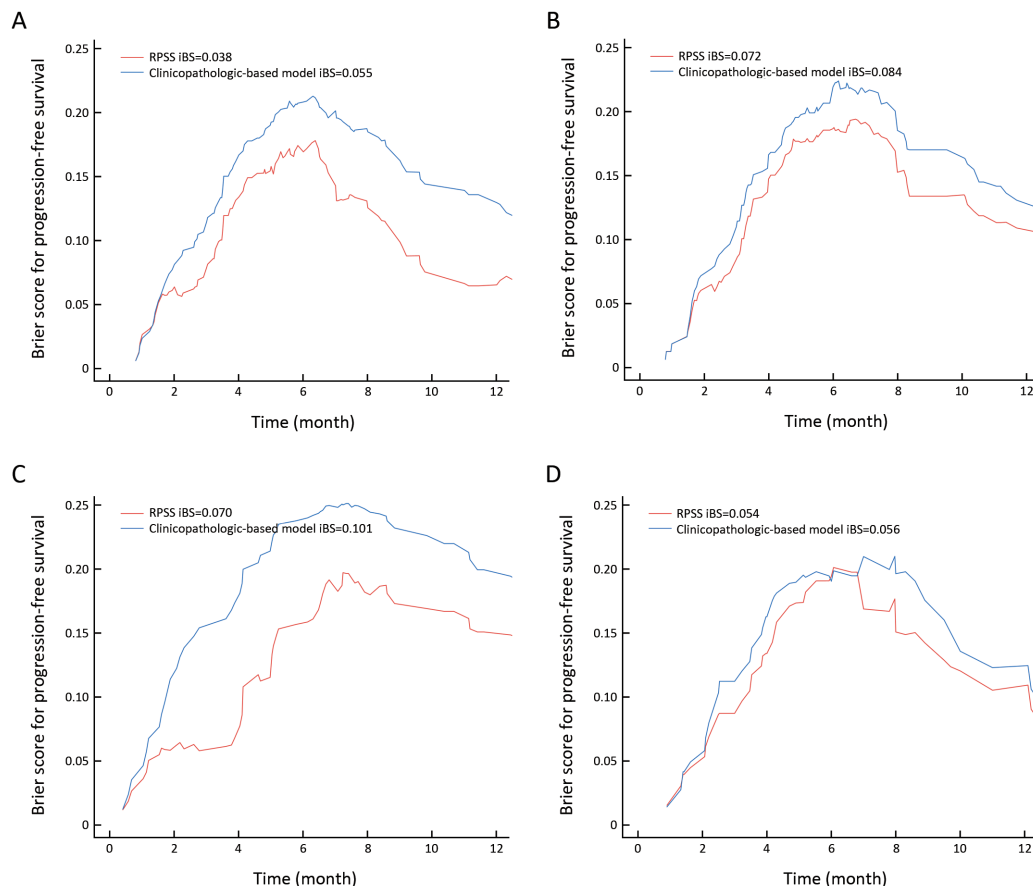
Figure 5 Prognostic value of RPSS and clinicopathologic-based model in terms of iBS. (A) Discovery cohort; (B) internal validation cohort; (C) external validation cohort; and (D) mutation validation cohort. The lower value of iBS indicated the better prognostic performance of clinical model. RPSS, radiomics prognostic scoring system; iBS, integrated Brier score.

compared with the clinicopathologic-based model.

Discussion

Although target therapy and immunotherapy have been recently developed to treat lung cancer, platinum-based chemotherapy is still first-line treatment for most stage IV NSCLC patients without gene mutations or for those with low PD-L1 expression (4,5). Disease progression is the common reason to stop first-line cytotoxic chemotherapy according to the American Society of Clinical Oncology guideline (4). First-line cytotoxic chemotherapy should be stopped at disease progression or after four cycles in patients whose disease is stable but not responding to treatment. Because NSCLC is remarkably heterogeneous, with regard to survival of the individual patients (6), the rarity of the biomarkers as a clinical entity means that any attempts to create predictive models to give an indication of prognosis are extremely challenging. In the present

study, we developed a noninvasive approach as RPSS to solve this problem. This scoring system aimed to estimate the probability of PFS based on a multivariate Cox proportional hazards model that included a radiomics signature and two clinical variables (PS and N stage). Based on three independent validation cohorts, RPSS was validated as a reliable tool to predict disease progression in these patients, and was superior to the clinicopathologic-based model, with only clinicopathological characteristics (PS and N stage). Furthermore, deriving of decision curve analysis and displaying as a nomogram for RPSS enhanced its practical utility.

A large number of studies have tried to identify genetic variations or clinical factors that can be used to individually predict the outcomes of platinum-based chemotherapy (7). However, due to their inherent limitations, which are achieved invasively or inconsistently, it is impractical to noninvasively conduct the prediction with univariate
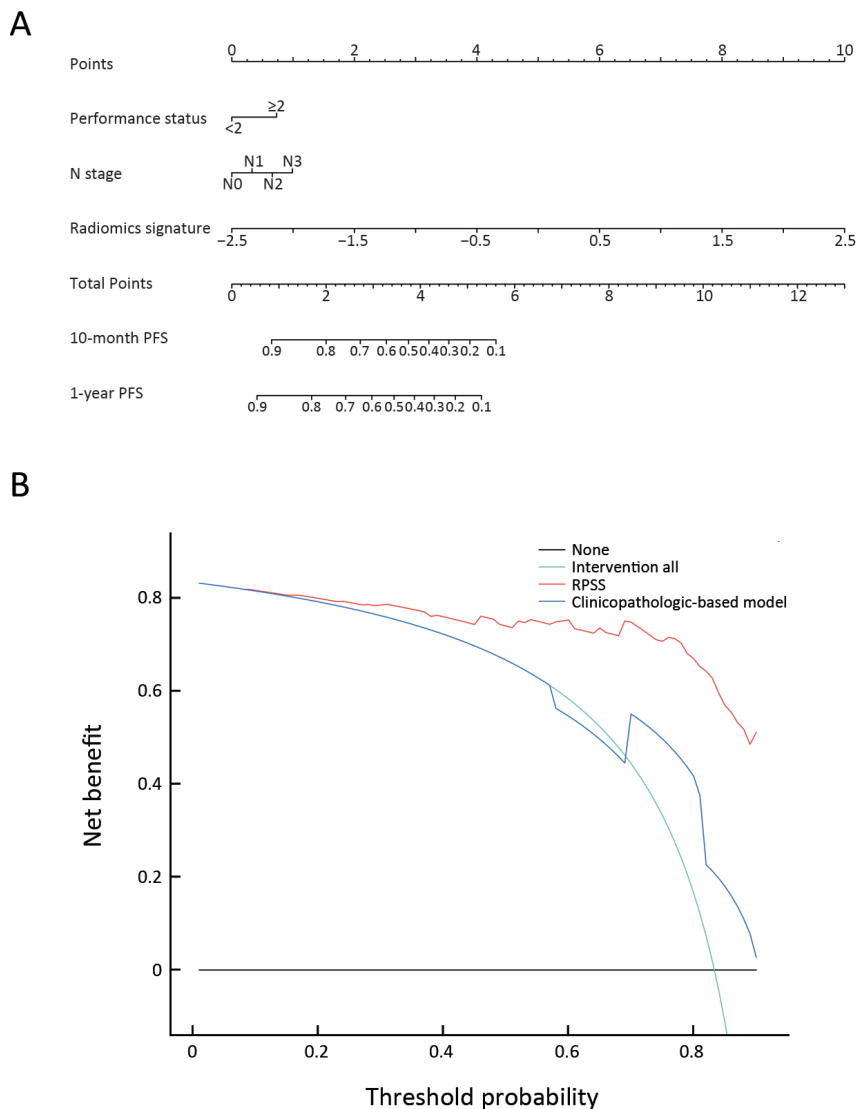
Figure 6 Clinical usefulness of RPSS. (A) Proposed nomogram to estimate the risk of disease progression for stage IV NSCLC treated with platinum-based chemotherapy. An individual patient's value is located on each variable axis, and a line is drawn upward to determine the number of points received for each variable value. The sum of these numbers is located on the Total Points axis, and a line is drawn downward to the survival axes to determine the likelihood of 10-month and 1-year PFS; (B) Decision curve analysis for comparison of the proposed RPSS (red line) and the clinicopathologic-based model (blue line) in terms of clinical usefulness. The y-axis measures the net benefit. The net benefit was calculated by subtracting the proportion of all patients who are false positive from the proportion who are true positive, and weighting by the relative harm of forgoing treatment compared with the negative consequences of an unnecessary treatment. Accordingly, the proposed RPSS had the higher net benefit compared with the clinicopathologic-based model for prediction of PFS in stage IV NSCLC patients treated with platinum-based chemotherapy. RPSS, radiomics prognostic scoring system; NSCLC, non-small cell lung cancer.

genetic or clinical factors. We thus extracted 1,182 image features from the pre-therapy CT images, and then reduced these to 16 potential predictors for integration into a radiomics signature by using LASSO Cox regression. All radiomics features selected for construction of radiomics signature were compliant with Image Biomarker Standardization Initiative (IBSI) (22). Our study showed that the established radiomics signature, which can potentially capture biologic properties like intra- and inter-tumor heterogeneities (39), can successfully stratify patients

602

He et al. A RPSS for patients with stage IV NSCLC

into high-risk and low-risk progression and demonstrate that approximately 70%−80% of the patients were predicted as high-risk progression of chemotherapy. Our results showed that higher value of radiomics signature was associated with higher risk of disease progression for stage IV NSCLC patients treated with platinum-based chemotherapy (All P<0.05). Through the multivariate analysis, we identified that PS and N stage were independent prognostic factors. These findings were in high concordance with previous studies for stage IV NSCLC patients treated with platinum-based chemotherapy (7,40). Additionally, we found that there were significant differences in T stage between the discovery cohort and validation cohorts. The possible reasons for the inconsistency might be that the eligible patients were enrolled from different periods, which is a stronger approach for the validation of RPSS (32). Therefore, consistent with a previous study on the prediction of EGFR-TKI treatment outcome in stage IV NSCLC (15), T stage was not suitable to be included as an independent factor into the final model in this study.

Even though overall survival (OS) is as the gold standard endpoint in clinical trials of chemotherapy for lung cancer, some of the disadvantages of this endpoint are the need for long-term follow-up and larger number of patients. Serial studies had found a high level of evidence that PFS can be considered a surrogate endpoint for OS in advanced NSCLC treated with chemotherapy (41). Disease progression also is the common reason to stop first-line cytotoxic chemotherapy according to the American Society of Clinical Oncology guideline (4,5). Thus, we kept PFS as the primary endpoint for this study. Furthermore, we have validated this radiomics signature we established for OS prediction on discovery and internal validation cohorts. The results showed that radiomics signature also successfully achieved patient stratification with respect to OS on discovery and internal validation cohorts, and remained as an independent predictor adjusted for clinicopathological parameters (All P<0.05, *Supplementary materials*).

Here, we also investigate whether the prognostic value of RPSS for predicting PFS could be extended in the mutation validation cohort of stage IV NSCLC patients with gene mutation who only received platinum-based chemotherapy as the first-line treatment. Surprisingly, we found that the RPSS also displayed a good discrimination in predicting PFS in mutation validation cohort (C-index=0.739; iAUC=0.705; and AUC at 1-year PFS=0.818).

This came to the robustness and effectiveness of our proposed RPSS for prognostic prediction. Thus, our established RPSS showed a good accuracy for prediction of PFS in all stage IV NSCLC patients treated with platinum-based chemotherapy.

To further investigate that how much extra clinical benefits we can obtain for individualized prediction on PFS by incorporating the radiomics signature, we also developed and compared a clinicopathologic-based model incorporated clinicopathologic characteristics (PS and N stage) without signature. Finally, results showed that RPSS showed better discrimination performance than the clinicopathologic-based model in all cohorts (C-index, RPSS *vs.* clinicopathologic-based model, discovery cohort: 0.772 *vs.* 0.686; internal validation cohort: 0.738 *vs.* 0.649; external validation cohort: 0.750 *vs.* 0.519; mutation validation cohort: 0.739 *vs.* 0.674; all comparisons P<0.05). And the decision curve analysis proved that RPSS offered significant improvement for individualized PFS prediction comparing with clinicopathologic-based model.

Although the RPSS we established demonstrated good levels of accuracy for the prediction of PFS, there are some limitations to our study. That is, the prognostic factors we used were restricted to the pre-therapy CT image phenotypic features and common clinical variables, and some potential molecular or immune biomarkers, such as excision repair complementing group 1 (ERCC1) protein (42), were not included as variables in the model analysis. The refinement of RPSS, with the identification of additional clinical, pathological or molecular predictors, will permit the optimization of this model. However, results of immunohistochemically staining are often inconsistent, assays for many biological or molecular markers are not widely available, and it may be difficult to standardize such results across the clinical practice.

Another limitation was that the primary tumors' features were extracted in our study. For IV carcinoma, the metastases might be more important than primary tumor to predict the chemotherapy response. However, stage IV NSCLC patients might have more than one tumor including the metastases, which might affect the uniformity of extracted radiomics features. Most of radiomics studies built radiomics models based on primary tumor to predict the development of distant metastases from NSCLC (43). Lack of radiomics studies had shown the reproducibility of extracted radiomics features from the metastases tumor. Thus, we selected the primary tumors of each eligible patient for features extraction. If multiple disconnected

tumor volumes were found, the largest by volume was chosen for features extraction. Further investigations are warranted to explore the potential usefulness of radiomics features extracted from metastases tumor for the prediction of chemotherapy response in patients with stage IV NSCLC. To the best of our knowledge, our study provided the first prognostic scoring system integrating both pre-therapy image biomarkers and clinical factors for stage IV NSCLC patients treated with platinum-based chemo-therapy.

Conclusions

We developed and externally validated a RPSS that was conveniently used to achieve an individualized prediction of PFS probability for stage IV NSCLC patients treated with platinum-based chemotherapy, which holds the promise of guiding the personalized pre-therapy of stage IV NSCLC.

Acknowledgements

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA Cancer J Clin 2017;67:7-30.
2. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71:209-49.
3. Lancia A, Merizzoli E, Filippi AR. The 8th UICC/AJCC TNM edition for non-small cell lung cancer staging: getting off to a flying start? Ann Transl Med 2019;7(suppl 6):S205.
4. Hanna N, Johnson D, Temin S, et al. Systemic therapy for stage IV non-small-cell lung cancer: American Society of Clinical Oncology clinical

practice guideline update. J Clin Oncol 2017;35: 3484-515.
5. Masters GA, Temin S, Azzoli CG, et al. Systemic therapy for stage IV non-small-cell lung cancer: American Society of Clinical Oncology clinical practice guideline update. J Clin Oncol 2015;33: 3488-515.
6. Pilkington G, Boland A, Brown T, et al. A systematic review of the clinical effectiveness of first-line chemo-therapy for adult patients with locally advanced or metastatic non-small cell lung cancer. Thorax 2015;70:359-67.
7. Yin JY, Li X, Li XP, et al. Prediction models for platinum-based chemotherapy response and toxicity in advanced NSCLC patients. Cancer Lett 2016; 377:65-73.
8. Giovannetti E, Toffalorio F, De Pas T, et al. Pharmacogenetics of conventional chemotherapy in non-small-cell lung cancer: a changing landscape? Pharmacogenomics 2012;13:1073-86.
9. Zhang Q, Zhu X, Zhang L, et al. A prospective study of biomarker-guided chemotherapy in patients with non-small cell lung cancer. Cancer Chemother Pharmacol 2014;74:839-46.
10. Fisher R, Pusztai L, Swanton C. Cancer hetero-geneity: implications for targeted therapeutics. Br J Cancer 2013;108:479-85.
11. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 2012;366: 883-92.
12. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441-6.
13. Verma V, Simone CB 2nd, Krishnan S, et al. The rise of radiomics and implications for oncologic management. J Natl Cancer Inst 2017:109.
14. Caudell JJ, Torres-Roca JF, Gillies RJ, et al. The future of personalised radiotherapy for head and neck cancer. Lancet Oncol 2017;18:e266-73.
15. Song J, Shi J, Dong D, et al. A new approach to predict progression-free survival in stage IV EGFR-mutant NSCLC patients with EGFR-TKI therapy. Clin Cancer Res 2018;24:3583-92.
16. de Jong EE, van Elmpt W, Leijenaar RT, et al.

[18F]FDG PET/CT-based response assessment of stage IV non-small cell lung cancer treated with paclitaxel-carboplatin-bevacizumab with or without nitroglycerin patches. Eur J Nucl Med Mol Imaging 2017;44:8-16.

17. Bortolotto C, Lancia A, Stelitano C, et al. Radiomics features as predictive and prognostic biomarkers in NSCLC. Expert Rev Anticancer Ther 2021;21:257-66.

18. Huang YQ, Liang CH, He L, 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.

19. Huang Y, Liu Z, He L, 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.

20. Khorrami M, Khunger M, Zagouras A, et al. Combination of peri- and intratumoral radiomic features on baseline CT scans predicts response to chemotherapy in lung adenocarcinoma. Radiol Artif Intell 2019;1:e180012.

21. Edge BE, Byrd-Holt D, Compton C, et al. editors. AJCC Cancer Staging Manual. 7th edition. New York: Springer, 2010.

22. Zwanenburg A, Vallières M, Abdalah MA, et al. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology 2020;295:328-38.

23. Acharya UR, Dua S, Du X, et al. Automated diagnosis of glaucoma using texture and higher order spectra features. IEEE Trans Inf Technol Biomed 2011;15:449-55.

24. Segal E, Sirlin CB, Ooi C, et al. Decoding global gene expression programs in liver cancer by noninvasive imaging. Nat Biotechnol 2007;25:675-80.

25. Friedman JH, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Software 2010;33:1-22.

26. Tibshirani R. Regression shrinkage and selection via the LASSO. J Royal Stat Society 1996;58:267-88.

27. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. Biometrics 2005;61:92-105.

28. Buyske S, Fagerstrom R, Ying Z. A class of weighted log-rank tests for survival data when the event is rare. Public Am Stat Assoc 2000;95:249-58.

29. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361-87.

30. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128-38.

31. Harrell FE Jr. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer, 2015.

32. Moons KG, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162:W1-73.

33. Pepe MS, Janes H, Li CI. Net risk reclassification p values: valid or misleading? J Natl Cancer Inst 2014;106:dju041.

34. Tangri N, Stevens LA, Griffith J, et al. A predictive model for progression of chronic kidney disease to kidney failure. JAMA 2011;305:1553-9.

35. Huitzil-Melendez FD, Capanu M, O'Reilly EM, et al. Advanced hepatocellular carcinoma: which staging systems best predict prognosis? J Clin Oncol 2010;28:2889-95.

36. Kerr KF, Brown MD, Zhu K, et al. Assessing the clinical impact of risk prediction models with decision curves: Guidance for correct interpretation and appropriate use. J Clin Oncol 2016;34:2534-40.

37. Fitzgerald M, Saville BR, Lewis RJ. Decision curve analysis. JAMA 2015;313:409-10.

38. Ten Haaf K, Jeon J, Tammemägi MC, et al. Risk prediction models for selection of lung cancer screening candidates: A retrospective validation study. PLoS Med 2017;14:e1002277.

39. Marusyk A, Almendro V, Polyak K. Intra-tumour heterogeneity: a looking glass for cancer? Nat Rev Cancer 2012;12:323-34.

40. Papadaki C, Sfakianaki M, Lagoudaki E, et al. PKM2 as a biomarker for chemosensitivity to front-line platinum-based chemotherapy in patients with metastatic non-small-cell lung cancer. Br J Cancer 2014;111:1757-64.

41. Mauguen A, Pignon JP, Burdett S, et al. Surrogate endpoints for overall survival in chemotherapy and radiotherapy trials in operable and locally advanced lung cancer: a re-analysis of meta-analyses of individual patients' data. Lancet Oncol 2013;14: 619-26.

42. Lee SM, Falzon M, Blackhall F, et al. Randomized prospective biomarker trial of ERCC1 for comparing platinum and nonplatinum therapy in advanced non-small-cell lung cancer: ERCC1 trial (ET). J Clin Oncol 2017;35:402-11.

43. Thawani R, McLane M, Beig N, et al. Radiomics and radiogenomics in lung cancer: A review for the clinician. Lung Cancer 2018;115:34-41.

Supplementary materials

Treatment of patients

All patients only received cisplatin- or carboplatin-based double chemotherapy as the first-line treatment. The other administrated regimens included pemetrexed, gemcitabine, paclitaxel, and docetaxel. The regimens were administered as follow: 1) cisplatin/carboplatin + pemetrexed: pemetrexed 500 mg/m^2 on d 1 plus cisplatin 75 mg/m^2 (PP) or carboplatin area under the curve (AUC) 5 mg/mL/min (PC) on d 1; 2) cisplatin/carboplatin + gemcitabine: gemcitabine 1.00 g/m^2 on d 1 plus cisplatin 75 mg/m^2 (GP) or carboplatin AUC 5 mg/mL/min (GC) on d 1; 3) cisplatin/carboplatin + paclitaxel: paclitaxel 175 mg/m^2 on d 1 plus cisplatin 75 mg/m^2 (TP) or carboplatin AUC 5 mg/mL/min (TC) on day 1; and 4) cisplatin/carboplatin + docetaxel: docetaxel 75 mg/m^2 on d 1 plus cisplatin 75 mg/m^2 (DP) or carboplatin AUC 5 mg/mL/min (DC) on day 1. All drugs were administered intravenously every three weeks, and patients were treated for 2−6 cycles until disease progression, unacceptable toxicity, or patient's refusal, as the therapeutic regimen.

Computed tomography (CT) image acquisition parameters

The acquisition parameters were as follows: 120 kV; 160 mAs; 0.5- or 0.4-s rotation time; detector collimation: 8 mm × 2.5 mm or 64 mm × 0.625 mm; field of view, 350 mm × 350 mm; and matrix, 512×512. The contrast-enhanced CT image was taken after a 25-s delay following the intravenous administration of 85 mL of iodinated contrast material (Ultravist 370, Bayer Schering Pharma, Berlin, Germany) at a rate of 3 mL/s with a pump injector (Ulrich CT Plus 150, Ulrich Medical, Ulm, Germany) after routine non-enhanced CT. Images in DICOM format images were retrieved from the picture archiving and communication system (PACS; Carestream, Canada).

Inter- and intra-class correlation coefficients (ICCs) for reproducibility analysis of radiomics features extraction

For the assessment of the inter-observer agreement of radiomics feature extraction, two radiologists with 15 years (Reader 1) and 12 years (Reader 2) of experience in chest CT interpretation performed the region of interest-based radiomics feature extraction procedure, in a blind fashion, respectively. Then, Reader 1 repeated the procedure with an interval of 1 week for assessment of the intra-observer agreement of radiomics feature extraction.

Algorithm for radiomics features calculation

The radiomics features analysis was applied to the pretreatment CT using in-house radiomics analysis software, with algorithms implemented in MATLAB 2014a (Mathworks, Natick, USA). The contrast-enhanced CT image data were retrieved from the institution archive and were loaded to a personal laptop for further analysis. All contrast-enhanced CT images were gathered for tumor segmentation. The primary tumors of all eligible patients were manually segmented by an experienced radiologist with more than 10 years of clinical experience in chest CT interpretation, thus a volume of interest (VOI) was delineated initially around the tumor outline for the whole volume area. The VOI was further refined by excluding the air area with a thresholding procedure that removed any pixels from the analysis with an attenuation values below −50 HU and beyond 300 HU.

In total, 1,182 quantitative radiomics features were extracted from each chest contrast-enhanced chest CT images, which included the tumor intensity, texture features, wavelet features, and Gabor features. The features of the tumor intensity and texture features were extracted without/after a filtration of the Laplacian of Gaussian filter (filter parameter = 1.0, 1.5, 2.0, 2.5, respectively) from the CT image. The wavelet features and Gabor features focused the features in different frequency ranges within the tumor pixels.

Laplacian of Gaussian filtration for gray-level histogram features and texture features
The Laplacian of the Gaussian filter ($\nabla^2 G$) distribution is given by

$$\nabla^2 G(x, y) = \frac{-1}{\pi\sigma^4}(1 - \frac{x^2 + y^2}{2\sigma^2})e^{-\left(\frac{x^2 + y^2}{2\sigma^2}\right)}$$

x, y denote the spatial coordinates of the pixel, and σ is the value of the filter parameter.

Gray-level histogram features

$X(i)$ indicates the intensity of the gray level i, and N denotes the sum of the pixels in the image. β indicates the top percentage of the histogram curve, which could be 50%, 25%, and 10%, M denotes the number of pixels in the histogram on the percentage of $(1-\beta)$.

  1) Mean

$$mean = \frac{1}{N}\sum_{i=1}^{N} X(i)$$

  2) SD

$$SD = \frac{1}{N}\sum_{i=1}^{N} (X(i) - \bar{X})^2$$

  3) Percentile mean and Percentile SD

$$\beta\_mean = \frac{1}{N - M}\sum_{i=M}^{N} X(i)$$

$$\beta\_SD = \frac{1}{N - M}\sum_{i=M}^{N} (X(i) - \bar{X})^2$$

  4) Kurtosis

$$kurtosis = \frac{\frac{1}{N}\sum_{i=1}^{N}(X(i) - \bar{X})^4}{\left(\sqrt{\frac{1}{N}\sum_{i=1}^{N}(X(i) - \bar{X})^2}\right)^4}$$

  5) Skewness

$$skewness = \frac{\frac{1}{N}\sum_{i=1}^{N}(X(i) - \bar{X})^3}{\left(\sqrt{\frac{1}{N}\sum_{i=1}^{N}(X(i) - \bar{X})^2}\right)^3}$$

Gray-level co-occurrence matrix features

A matrix $P(i, j)$ to indicate the relative frequency with the intensity values of two pixels (i and j) at the three distances (δ=1,2,3) and in four directions (0°, 45°, 90°, 135°). N_g is the number of discrete intensity levels in the image; x, y denote the spatial coordinates of the pixel. μ, $\mu_x(i)$, $\mu_y(j)$ is the mean of $P(i, j)$, $P_x(i)$, $P_y(j)$, and $\sigma_x(i)$, $\sigma_y(j)$ is the standard deviation of $P_x(i)$, $P_y(j)$, respectively.

  1) Contrast

$$contrast = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} |i - j|^2 P(i,j)$$

  2) Correlation

$$correlation = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} ijP(i,j) - \mu_i(i)\mu_j(j)}{\sigma_x(i)\sigma_y(j)}$$

3) Entropy

$$entropy = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} P(i,j)\log[P(i,j)]$$

4) Energy

$$energy = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} [P(i,j)]^2$$

5) Homogeneity

$$homogeneity = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{P(i,j)}{1+|i-j|^2}$$

Gray-level run-length matrix features

A gray level run was defined as the length in the number of pixels, of the consecutive pixels, that had the same gray level value. In a gray level run length matrix $p(i,j|\theta)$, the (i,j)th element describes the number of times j a gray level i appears consecutively in the direction specified by θ (0º, 45º, 90º, 135º), and N_g is the number of discrete gray level intensities. N_r is the number of different run lengths, and N_p is the number of pixels in the image.

1) Short Run Emphasis (SRE)

$$SRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\left[\frac{p(i,j|\theta)}{j^2}\right]}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j|\theta)}$$

2) Long Run Emphasis (LRE)

$$LRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} j^2 p(i,j|\theta)}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j|\theta)}$$

3) Gray Level Non-Uniformity (GLN)

$$GLN = \frac{\sum_{i=1}^{N_g}\left[\sum_{j=1}^{N_r} p(i,j|\theta)\right]^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j|\theta)}$$

4) Run Length Non-Uniformity (RLN)

$$RLN = \frac{\sum_{j=1}^{N_r}\left[\sum_{i=1}^{N_g} p(i,j|\theta)\right]^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j|\theta)}$$

5) Run Percentage (RP)

$$RP = \sum_{i=1}^{N_g}\sum_{j=1}^{N_r} \frac{p(i,j|\theta)}{N_p}$$

6) Low Gray Level Run Emphasis (LGLRE)

$$LGLRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\left[\frac{p(i,j|\theta)}{i^2}\right]}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j|\theta)}$$

7) High Gray Level Run Emphasis (HGLRE)

$$HGLRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} i^2 p(i,j|\theta)}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j|\theta)}$$

8) Short Run Low Gray Level Emphasis (SRLGLE)

$$SRLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \left[\frac{p(i,j|\theta)}{i^2 j^2} \right]}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta)}$$

9) Short Run High Gray Level Emphasis (SRHGLE)

$$SRHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \left[\frac{p(i,j|\theta)i^2}{j^2} \right]}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta)}$$

10) Long Run Low Gray Level Emphasis (LRLGLE)

$$LRLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \left[\frac{p(i,j|\theta)j^2}{i^2} \right]}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta)}$$

11) Long Run High Gay Level Emphasis (LRHGLE)

$$LRHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta)i^2 j^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta)}$$

Wavelet features

Wavelet transform effectively decouples textural information by decomposing the original image, in a similar manner as a Fourier analysis, in low- and high-frequencies. In this study, three-level discrete wavelet transform was applied to each CT image, which decomposed the original image X into 9 decompositions.

Gabor features

Gabor filter, named after Dennis Gabor, is a linear filter used for edge detection, which is usually used in the field of face recognition. It selects valuable image information in different directions and different scales. In this study, we used four directions and three scales to extract Gabor features. The mean was used to construct the Gabor feature group.

$$Mean\,(Gabor) = \frac{1}{N} \sum_{i=1}^{N} X\,(i)$$

N indicates the sum of the image pixels and the $X\,(i)$ presents the intensity i of on the Gabor image.

R packages used

The least absolute shrinkage and selection operator method (LASSO) Cox regression model was done using the "glmnet" package. The time-dependent receiver operating characteristic (ROC) analyses were done using the "survivalROC" package. The multivariable Cox regression analysis and calibration plots were done with the "rms" package. The calculation of the C-index was performed with the "Hmisc" package, and the calculation of net reclassification index (NRI) and integrated discrimination improvement (IDI) were performed with the "survIDINRI" package. The calculation of the integrated area under the ROC curve (iAUC) and the integrated Brier score (iBS) were performed with the "risksetROC" package. The internal validation of the C-index was performed with the "rms" package. The acquisition of the decision curve analysis was performed with the "stdca.R" package.

Calculation of radiomics signature based on LASSO Cox regression analysis

$$\begin{aligned}
\textit{The value of radiomics signature} \\
= 0.08454937 \times 2\text{D\_skewness\_0} + 0.42527225 \times 2\text{D\_kurtosis\_0} \\
- 0.34077344 \times 2\text{D\_contrast\_0\_1\_0} + 0.46372619 \\
\times 2\text{D\_contrast\_45\_1\_0} + 0.00708416 \times 2\text{D\_contrast\_90\_1\_0} \\
+ 0.01030149 \times \text{c\_correlation\_135\_2\_0} - 0.35094852 \\
\times \text{c\_correlation\_0\_3\_0} - 0.06798770 \times \text{c\_correlation\_45\_3\_0} \\
+ 0.37192199 \times \text{c\_energy\_0\_1\_0} + 0.11469022 \\
\times \text{c\_his\_25\_mean\_0} + 0.07791898 \times \text{c\_his\_10\_SD\_1.0} \\
- 0.04199315 \times 3\text{D\_correlation\_45\_1\_0} + 0.12571617 \\
\times 3\text{D\_correlation\_45\_3\_0} + 0.21898221 \times 3\text{D\_his\_10\_SD\_1.0} \\
- 0.04686967 \times 3D\text{\_kurtosis\_2.5} + 0.12042190 \times 3D\text{\_wavelet\_1}
\end{aligned}$$

Validation of radiomics signature for predicting overall survival (OS) in discovery and internal validation cohorts

Methods

First, all patients were stratified into high-risk and low-risk groups by the established signature using the cutoff=−0.734. The significant difference in OS was analyzed to investigate the clinical benefits across different subgroups by the Kaplan-Meier survival analysis and the Cox proportional hazards regression model on both discovery and internal validation cohorts. Then, the established radiomics signature and potential clinicopathologic characteristics, including age, gender, smoking history, histological type, performance status (PS), T stage, N stage, the site and volume of metastases, and chemotherapy regimens, were involved in multivariable Cox regression analysis for OS prediction.

Results

The Kaplan-Meier curves clearly showed different prognostic strata in OS between the high-risk and low-risk subgroups in both discovery and internal validation cohorts, with a high statistical significance (log-rank P<0.01 in all cases). Lower value of radiomics signature was associated with improved OS in discovery (HR: 3.112, 95% CI: 1.689−5.734) and internal validation cohorts (HR: 2.231, 95% CI: 1.246−3.994). After the multivariate analyses, the radiomics signature remained as an independent risk factor for prediction of OS (HR: 1.913, 95% CI: 1.336−2.741; P=0.0004).
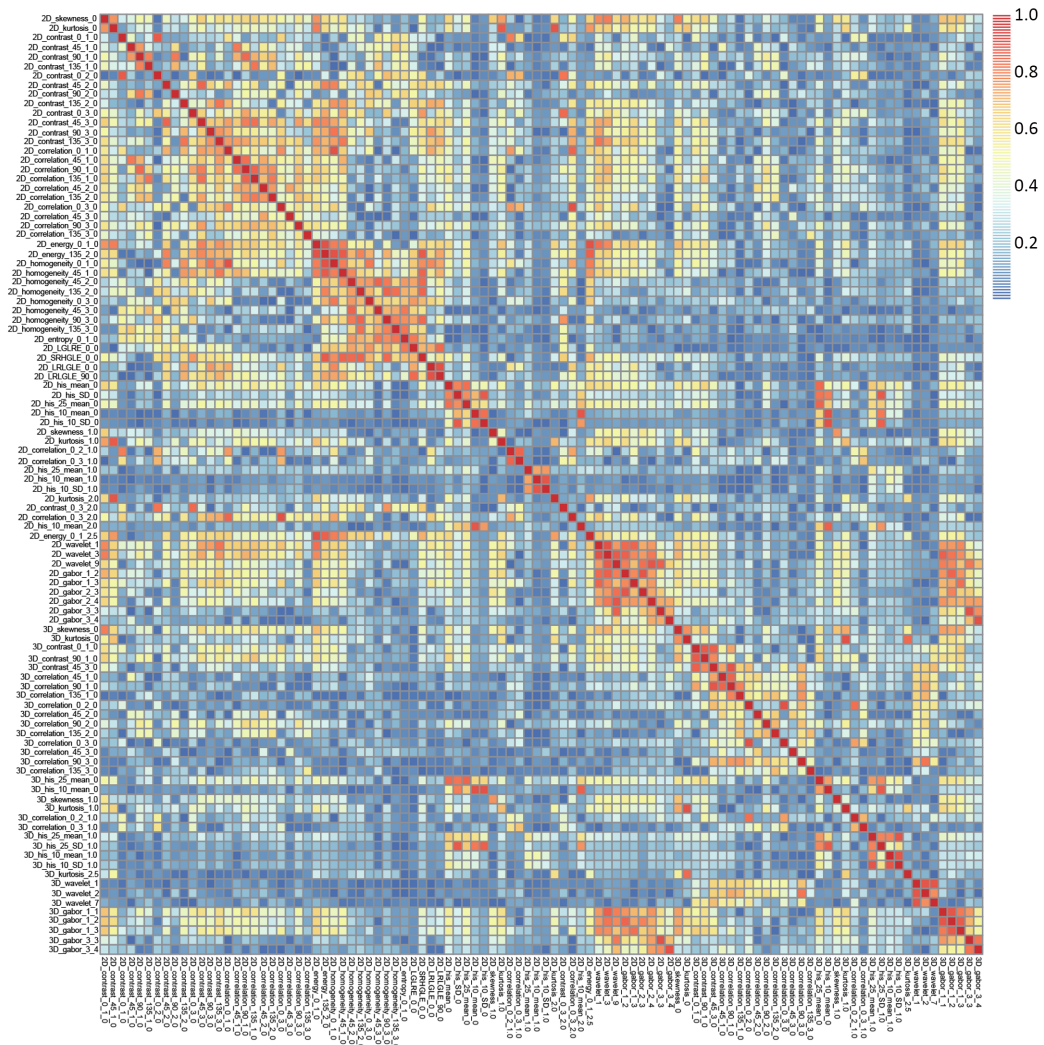
Figure S1 Correlation mapping for those radiomics features which were independent from other features, as determined by the Pearson's correlation coefficient among the features (cut-off value of 0.9).
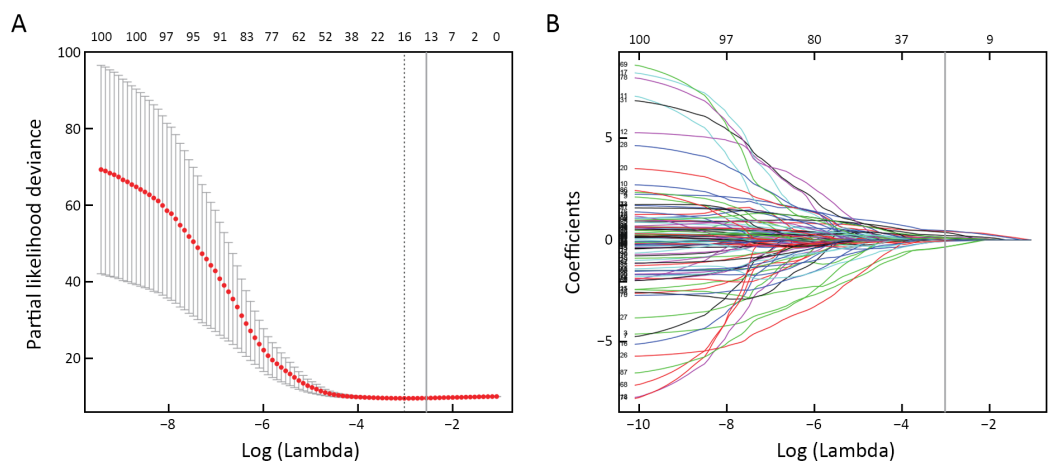
Figure S2 Feature selection in discovery cohort using the LASSO penalized Cox proportional regression analysis. LASSO, the least absolute shrinkage and selection operator method.
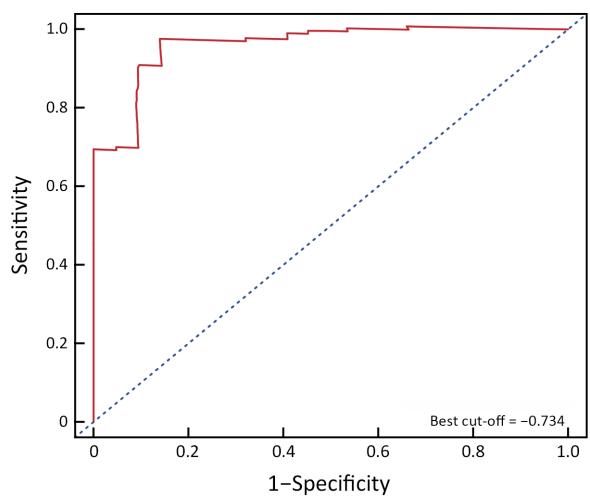

Figure S3 Time-dependent ROC curve at 1 year in discovery cohort (Best cut-off=−0.734). ROC, receiver-operating characteristic.