Radiomics Signature: A Potential Biomarker for the Prediction of Disease-Free Survival in Early-Stage (I or II) Non–Small Cell Lung Cancer[1]

Yanqi Huang, MD
Zaiyi Liu, MD
Lan He, MPhil
Xin Chen, MD
Dan Pan, MD
Zelan Ma, MD
Cuishan Liang, MD
Jie Tian, PhD
Changhong Liang, MD

Purpose: To develop a radiomics signature to estimate disease-free survival (DFS) in patients with early-stage (stage I–II) non–small cell lung cancer (NSCLC) and assess its incremental value to the traditional staging system and clinical-pathologic risk factors for individual DFS estimation.

Materials and Methods: Ethical approval by the institutional review board was obtained for this retrospective analysis, and the need to obtain informed consent was waived. This study consisted of 282 consecutive patients with stage IA–IIB NSCLC. A radiomics signature was generated by using the least absolute shrinkage and selection operator, or LASSO, Cox regression model. Association between the radiomics signature and DFS was explored. Further validation of the radiomics signature as an independent biomarker was performed by using multivariate Cox regression. A radiomics nomogram with the radiomics signature incorporated was constructed to demonstrate the incremental value of the radiomics signature to the traditional staging system and other clinical-pathologic risk factors for individualized DFS estimation, which was then assessed with respect to calibration, discrimination, reclassification, and clinical usefulness.

Results: The radiomics signature was significantly associated with DFS, independent of clinical-pathologic risk factors. Incorporating the radiomics signature into the radiomics-based nomogram resulted in better performance ($P < .0001$) for the estimation of DFS (C-index: 0.72; 95% confidence interval [CI]: 0.71, 0.73) than with the clinical-pathologic nomogram (C-index: 0.691; 95% CI: 0.68, 0.70), as well as a better calibration and improved accuracy of the classification of survival outcomes (net reclassification improvement: 0.182; 95% CI: 0.02, 0.31; $P = .02$). Decision curve analysis demonstrated that in terms of clinical usefulness, the radiomics nomogram outperformed the traditional staging system and the clinical-pathologic nomogram.

Conclusion: The radiomics signature is an independent biomarker for the estimation of DFS in patients with early-stage NSCLC. Combination of the radiomics signature, traditional staging system, and other clinical-pathologic risk factors performed better for individualized DFS estimation in patients with early-stage NSCLC, which might enable a step forward precise medicine.

Online supplemental material is available for this article.

[1] From the Department of Radiology, Guangdong General Hospital, Guangdong Academy of Medical Sciences, 106 Zhongshan Er Road, Guangzhou 510080, China (Y.H., Z.L., L.H., D.P., Z.M., Cuishan Liang, Changhong Liang); Graduate College, Southern Medical University, Guangzhou, China (Y.H., Z.M., Cuishan Liang); School of Medicine, South China University of Technology, Guangzhou, Guangdong, China (L.H.); Department of Radiology, the Affiliated Guangzhou First People's Hospital, Guangzhou Medical University, Guangzhou, China (X.C.); Key Laboratory of Molecular Imaging, Chinese Academy of Sciences, Beijing, China (J.T.). Received October 27, 2015; revision requested January 14, 2016, and received January 27; accepted March 28; final version accepted April 6. **Address correspondence to** Changhong Liang. (e-mail: *cjr.lchh@vip.163.com*).

Supported by the National Natural Scientific Foundation of China (nos. 81271569, 81271654, and U1301258).

Y.H. and Z.L. contributed equally to this work.

Lung cancer is the leading cause of cancer-related mortality worldwide, and non–small cell lung cancer (NSCLC) accounts for 85% of cases (1). Early-stage (IA-IIB) NSCLC, although it accounts for only 25%–30% of lung cancer, theoretically provides the highest possibility of modifying the outcome of NSCLC (2,3). Surgical resection with a curative intent is regarded as the cornerstone of treatment for early-stage NSCLC, and tumor node metastasis (TNM) stage is traditionally considered to be the most important postoperative prognostic factor (4,5). However, the wide spectrum of survival times that exists even after complete resection of the same-staged NSCLCs demonstrates the imperative need for personalized medicine (3,6).

Improvements in the survival estimation have largely been made as a result of advances in biologic and genomic technologies that have allowed incorporation of survival-associated biologic or genetic signatures (7,8). However, the inability to obtain comprehensive information about heterogeneous tumors remains a limitation of these invasive methods (9,10).

The field of radiogenomics has attracted recent interest in defining association maps between genomic phenotypes and imaging texture features (11). To facilitate a direct estimation of outcomes using imaging features, "radiomics" has been brought into the evolving topic (12). By extracting high throughput of quantitative descriptors from routinely acquired computed tomography (CT) studies, radiomics enables the noninvasive profiling of tumor heterogeneity (12–14). Recent advances in radiomics have provided insights in personalized medicine in oncologic practice related to tumor detection, subtype classification, and therapeutic response assessment (13,15,16).

Although several texture features, such as uniformity and entropy, have been suggested to allow for risk stratification, initial attempts at survival estimation in NSCLC were largely limited to the use of a texture feature in isolation as a biomarker (12,17,18). However, in the prognostic setting, estimates of risk probabilities are rarely based on a single risk factor because of the insufficiency regarding reliable estimates. Estimation based on a multivariable model is now considered a more reliable way, as in the way that doctors naturally integrate several patient characteristics and symptoms to make estimations. A multi-marker signature is of utmost importance in the "-omics" setting, where huge numbers of markers are usually studied in a high-throughput manner (19–22). Although it is now recognized that a signature composed of multiple biomarkers holds higher value than a single biomarker (19), a radiomics signature associated with the survival outcome of early-stage NSCLC has not yet been described, to our knowledge.

Therefore, the aim of this study was to develop a radiomics signature to estimate the disease-free survival (DFS) in patients with early-stage (I–II) NSCLC and to assess its incremental value to the traditional staging system and clinical-pathologic risk factors for individual DFS estimation.

Advances in Knowledge

- The identified multi-feature–based radiomics signature was a significant biomarker of disease-free survival (DFS) in early-stage non–small cell lung cancer (NSCLC) ($P = .0182$, G-rho rank test), independent of clinical stage, histologic grade, age, and sex (multivariate Cox regression model, $P = .0055$; hazard ratio: 1.77, 95% confidence interval [CI]: 1.18, 2.65); the identified signature consisted of the following features: CE_kurtosis_0, CE_uniformity_0_0, CE_homogeneity_45_0, UE_uniformity_45_1.0, and CE_uniformity_0_1.5.
- The radiomics signature demonstrated incremental value to the traditional staging system and other clinical-pathologic risk factors for the estimation of individual DFS in patients with early-stage NSCLC when the radiomics signature–based nomogram was compared with the clinical-pathologic nomogram, with the former presenting a higher discrimination ability for DFS (C-index: 0.72 vs 0.691, $P < .0001$), a better calibration, an improved accuracy of classification of the survival outcomes (net reclassification improvement: 0.182; 95% CI: 0.02, 0.31; $P = .02$), and a greater clinical utility (as demonstrated by the decision curve analysis).

Implication for Patient Care

- The identified multi-feature–based radiomics signature may serve as a noninvasive biomarker for early prognostication and patient stratification for patients with early-stage NSCLC in human clinical trials or clinical precise medicine.

Materials and Methods

Patients

Ethical approval was obtained for this retrospective analysis, and the need to obtain informed consent was waived.

Published online before print
10.1148/radiol.2016152234 **Content codes:** CH OI

Radiology 2016; 281:947–957

Abbreviations:
AIC = Akaike information criterion
AJCC = American Joint Committee on Cancer
CI = confidence interval
DFS = disease-free survival
HR = hazard ratio
LASSO = least absolute shrinkage and selection operator
NSCLC = non–small cell lung cancer

Author contributions:
Guarantors of integrity of entire study, Z.L., X.C., D.P., J.T., Changhong Liang; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; manuscript final version approval, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, Y.H., Z.L., X.C., Cuishan Liang, Changhong Liang; clinical studies, Y.H., Z.L., L.H., X.C., D.P.; experimental studies, Z.L., X.C., Z.M., J.T.; statistical analysis, Y.H., Z.L., L.H., Z.M., J.T.; and manuscript editing, Y.H., Z.L., J.T., Changhong Liang

Conflicts of interest are listed at the end of this article.

This study comprised an evaluation of the institutional database for medical records from January 2007 to July 2012 to identify patients with histologically confirmed early-stage (IA, IB, IIA, and IIB) NSCLC who underwent surgical resection with curative intent. In total, 282 cases were identified (186 men and 96 women; mean age, 60.87 years ± 11.27; range, 21–85 years). Tumor staging was performed on the basis of the *American Joint Committee on Cancer TNM Staging Manual, 7th Edition* (23). Figure E1 (online) shows the patient recruitment pathway, along with the inclusion and exclusion criteria.

Computer-generated random numbers were used to assign half of the patients to the training data set and half to the validation data set.

Table 1

Demographic and Clinical Characteristics of Patients with NSCLC in the Training Data Set and Validation Data Set

Characteristic	Training Data Set	Validation Data Set
Sex		
Male	94 (66.7)	92 (65.2)
Female	47 (33.3)	49 (34.8)
Age (y)		
Median*	61 (21–84)	60 (28–85)
≤60	62 (44)	71 (50.4)
61–69	48 (34)	38 (27.0)
≥70	31 (22)	32 (22.7)
Stage		
IA	33 (23.4)	36 (25.5)
IB	86 (61)	77 (54.6)
IIA	2 (1.4)	4 (2.8)
IIB	20 (14.2)	24 (17.0)
T stage		
T1	35 (24.8)	37 (26.2)
T2	97 (68.8)	93 (66.0)
T3	9 (6.4)	11 (7.8)
N stage		
N0	130 (92.2)	123 (87.2)
N1	10 (7.1)	18 (12.8)
N2	1 (0.7)	0
Smoking status		
Yes	53 (37.6)	51 (36.2)
No	88 (62.4)	90 (63.8)
Histologic grade		
Well	6 (4.3)	10 (7.1)
Mediate	102 (72.3)	95 (67.4)
Poor	33 (23.4)	36 (25.5)
Follow-up time (mo)		
Median†	45 (27.5–68)	43 (31–67.5)
Maximum	98	95

Note.—Unless otherwise specified, data are numbers of patients, with percentages in parentheses. No difference was found between the training data set and the validation data set in either the clinical characteristics or the follow-up data ($P = .548–.980$).

* Data in parentheses are ranges.

† Data in parentheses are interquartile ranges.

Follow-up

The end point of this study was DFS, which was defined as the time from the date of the CT examination until either the date of relapse (event), which refers to tumor recurrence within or immediately adjacent to the treated field, mediastinal relapse, distant relapse, or death, or until the date that the patient was last known to be free of relapse (censored). The minimum follow-up period to ascertain the DFS was 36 months after the first CT study; while the maximum follow-up time was 98 months (median, 44 months). Patients were postoperatively followed up with chest CT every 6–12 months for the first 2 years and then annually, according to the follow-up protocol of our institution.

Baseline clinical-pathologic data, including age, sex, smoking status, stage (T stage, N stage, and clinical stage), histologic grade, and date of baseline CT imaging, were obtained from the medical records (Table 1).

Image Acquisition

All patients underwent unenhanced and contrast material–enhanced chest CT. Details regarding the acquisition parameters and CT image retrieval procedure are presented in Appendix E1 (online).

Imaging Texture Analysis

The interobserver reproducibility was initially analyzed with 50 randomly chosen images for region-of-interest–based texture feature extraction by two experienced radiologists (Y.H. and Z.L.; readers 1 and 2, with 10 and 12 years of clinical experience in chest CT study interpretation, respectively) in a blinded fashion. The workflow for the remaining images was completed by the first radiologist.

The corresponding algorithms for the region-of-interest settings, along with the image filtration process and feature generation, are described in Appendix E2 (online). Imaging texture analysis was applied to the pretreatment CT study by using in-house texture analysis algorithms implemented in Matlab 2010a (Mathworks, Natick, Mass). In total, 132 texture features from the category of histogram and gray-level co-occurrence matrix were finally extracted from one CT image (Table E1 [online]).

Statistical Analysis

The statistical analysis was performed with R software, version 3.0.1 *(http://www.R-project.org)* and X-tile

software, version 3.6.1 (Yale University School of Medicine, New Haven, Conn). The packages in R used in this study are described in Appendix E3 (online). The reported statistical significance levels were all two sided, with the statistical significance level set at .05.

The differences in age, sex, stage, smoking status, histologic grade, mean follow-up time, and mean survival time between the training and validation data sets were assessed by using an independent samples *t* test, χ^2 test, or Mann-Whitney *U* test, where appropriate.

The statistical analysis for assessment of interobserver reproducibility is described in Appendix E4 (online).

Construction of the radiomics score–based radiomics signature.—According to the Harrell guideline, the number of events should exceed the number of included covariates by at least 10 times in a multivariate analysis (24). To address this issue, the least absolute shrinkage and selection operator (LASSO) Cox regression model, which is suitable for the regression of high-dimensional data (25,26), was used to select the most useful prognostic features in the training data set. The selected imaging features were then combined into a radiomics signature. A radiomics score (Rad-score) was computed for each patient through a linear combination of selected features weighted by their respective coefficients.

Validation of radiomics signature.—The potential association of the radiomics signature with DFS was first assessed in the training data set and then validated in the validation data set by using Kaplan-Meier survival analysis. The patients were classified into high-risk or low-risk groups according to the Rad-score, the threshold of which was identified by using X-tile (27). The difference in the survival curves of the high-risk and low-risk groups was evaluated by using a weighted log-rank test (the G-rho rank test, rho = 1) for a substantial increase in efficiency because the event rate was relatively low compared with the sample size (28).

Stratified analyses were performed to explore the potential association of the radiomics signature with the DFS using subgroups within clinical-pathologic risk factors from the whole data set. The performances of the constructed radiomics signature in patients with stage I disease and in those with stage II disease were meanwhile investigated. The possible association of each selected feature with the DFS was also assessed, and the relative hazard ratio (HR) was calculated on the basis of the G-rho rank test in a univariate analysis.

Evaluation of the multi-feature–based radiomics signature as an independent biomarker was performed by integrating the following clinical-pathologic risk factors into the multivariable Cox proportional hazards model (backward step-down selection; the Akaike information criterion): TNM stage (IA, IB, IIA, or IIB), histologic grade (well, moderately, or poorly differentiated), sex (female or male), age (≤60, 61–69, or ≥70 years), and smoking status (yes or no).

Assessment of incremental value of radiomics signature in individual DFS estimation.—To demonstrate the incremental value of the radiomics signature to the traditional staging system and other clinical-pathologic risk factors for individualized assessment of DFS in patients with early-stage NSCLC, both a radiomics nomogram and a clinical-pathologic nomogram were presented in the validation data set. The radiomics nomogram incorporated the radiomics signature and the independent clinical-pathologic risk factors based on the multivariate Cox analysis. The clinical-pathologic nomogram incorporated only the independent clinical-pathologic risk factors.

The incremental value of the radiomics signature to the traditional staging system and other clinical-pathologic risk factors was assessed with respect to calibration, discrimination, reclassification, and clinical usefulness. The potential of the radiomics signature was compared with that of the traditional staging system (23). The performance of the radiomics nomogram was compared with that of both the traditional staging system and clinical-pathologic nomogram.

To compare the agreement between the observed outcomes and the DFS association of the radiomics nomogram and the clinical-pathologic nomogram, calibration curves were generated (29). To quantify the discrimination performance, the Harrell concordance index (C-index) was measured, along with the concordance probability estimate considering the high degree of censoring in our data (1 indicates perfect concordance; 0.5 indicates no better concordance than chance) (30,31). To quantify the improvement of usefulness added by the radiomics signature, a net reclassification improvement calculation was also applied (32). The Akaike information criterion (AIC) was calculated to assess the risk of overfitting. Finally, a decision curve analysis determined the clinical usefulness of the radiomics nomogram by quantifying the net benefits at different threshold probabilities (33).

Results

Clinical Characteristics and DFS

As of the last follow-up, 97 patients (34.4%) had experienced a confirmed disease relapse. The mean DFS was 27.66 months, and the median DFS was 23 months. The shortest DFS was 3 months.

No difference was found between the training data set and the validation data set in either clinical characteristics or follow-up data (P = .548–.980; Table 1).

The interobserver reproducibility of the texture features extraction was high (Appendix E4 [online]). Therefore, all outcomes were based on the measurements of the first radiologist.

Construction of the Radiomics Score-based Radiomics Signature

The texture features with a nonzero coefficient in the LASSO Cox regression model were as follows: CE_kurtosis_0, CE_uniformity_0_0, CE_homogeneity_45_0, UE_uniformity_45_1.0, and CE_uniformity_0_1.5 (Fig E2a and E2b [online]).

Figure 1

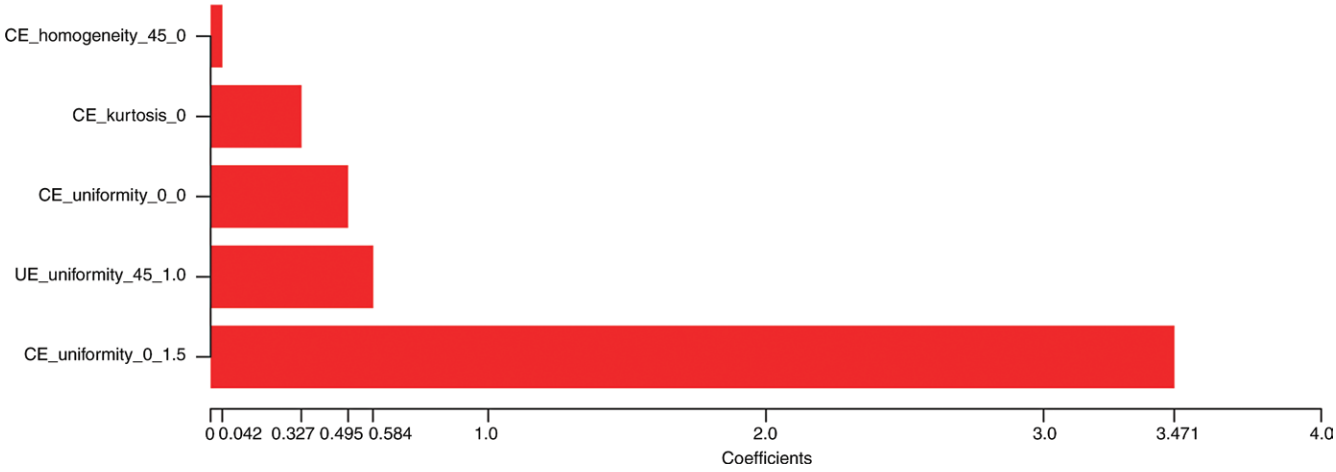

Figure 1: Histogram shows the role of individual features that contribute to the developed signature. The features that contribute to the radiomics signature are plotted on the y-axis, with their coefficients in the LASSO Cox analysis plotted on the x-axis.

Figure 2

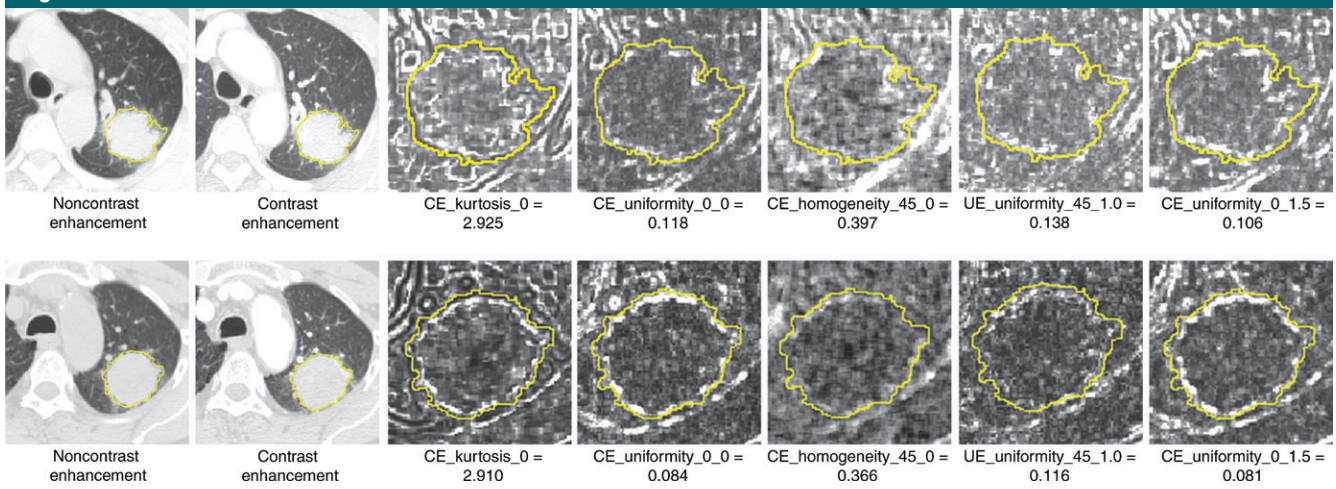

Figure 2: Radiomics feature maps of the five selected features. From left to right: The original unenhanced CT image; the contrast-enhanced CT image; and the feature maps of CE_kurtosis_0, CE_uniformity_0_0, CE_homogeneity_45_0, UE_uniformity_45_1.0, and CE_uniformity_0_1.5. Top row: Images in a patient (DFS = 19 months; Rad-score = 1.48; stage IB) who was classified into the high-risk group. Bottom row: Images in a patient (DFS = 78 months; Rad-score = 1.36; stage IB) who was classified into the low-risk group.

The radiomics signature was constructed, with a Rad-score calculated by using the following formula:

$$\begin{aligned}\text{Rad-score} = {} & \text{CE\_kurtosis\_0} \cdot 0.32742342 \\ & + \text{CE\_uniformity\_0\_0} \\ & \cdot 0.49479424 + \text{CE\_homogeneity\_45\_0} \\ & \cdot 0.04249791 + \text{UE\_uniformity\_45\_1.0} \\ & \cdot 0.58494321 + \text{CE\_uniformity\_0\_1.5} \\ & \cdot 3.47114758.\end{aligned}$$

The contribution of the selected parameters with their regression coefficients for the signature construction is presented in the form of a histogram in Figure 1. The optimum cutoff generated by the X-tile plot was 1.38 (Fig E2c [online]). Accordingly, patients were classified into a high-risk group (Rad-score $\geq$ 1.38) and a low-risk group (Rad-score $<$ 1.38). Radiomics maps of the above five features in two patients with stage IB disease who were classified into the high-risk group and low-risk group, respectively, are presented in Figure 2 . DFS and relapse rate in the high-risk group and the low-risk group in the training and validation data sets is listed in Table 2. The distributions of the Rad-scores and DFS in the validation data set are shown in Figure E2d (online).

Table 2

DFS and Recurrence Rate in High-Risk and Low-Risk Groups

	Training Data Set			Validation Data Set		
Parameter	High-Risk Group	Low-Risk Group	Total	High-Risk Group	Low-Risk Group	Total
No. of patients	91 (64.5)	50 (35.5)	141	83 (58.9)	58 (41.1)	141
3-Year DFS						
Median*	22 (10–46)	25.5 (7–57.25)	24 (9.5–46)	22.5 (7–38.25)	26 (13–38.5)	22.5 (7.75–37.75)
Shortest	4	3	3	3	3	3
No. of recurrences						
At 1 year	27 (29.7)	7 (14)	34 (24.1)	23 (27.7)	9 (15.5)	32 (22.7)
At 2 years	39 (42.9)	10 (20)	49 (34.8)	30 (36.1)	11 (19)	41 (29.1)
At 3 years	39 (42.9)	10 (20)	49 (34.8)	34 (41)	14 (24.1)	48 (34)

Note.—Unless otherwise specified, data in parentheses are percentages.

* Data in parentheses are interquartile ranges.

Validation of Radiomics Signature

The radiomics signature was associated with the DFS in the training data set (P = .0129; HR = 2.36, 95% confidence interval [CI]: 1.25, 4.47), and this finding was confirmed in the validation data set (P = .0182; HR = 2.09; 95% CI: 1.12, 3.91) (Fig 3). Patients with lower Rad-scores generally had better DFS: Although 14 (24.1%) patients were confirmed to have tumor relapses in the low-risk group (mean DFS, 30 months), 34 (41%) patients in the high-risk group were confirmed to have tumor relapse (mean DFS, 24.97 months). When the patients were stratified on the basis of clinical-pathologic risk factors, a significant association was found in one or more subgroups (Fig E3 [online]). Note the heterogeneity of the estimated probabilities for the time to recurrence within each American Joint Committee on Cancer (AJCC) stage. The performances of the constructed radiomics signature in patients with stage I disease and those with stage II disease, respectively, are presented in Appendix E5 (online). Univariate analysis revealed the possible association between each of the selected features and the DFS, as presented in Figure E4 (online). However, in the validation data set, none of these individual features could be used to successfully stratify patients into high-risk and low-risk groups in the validation data set, with P values of the G-rho rank test ranging from .0859 to .8254 (P value: CE_kurtosis_0, .0859; CE_uniformity_0_0, .8254; CE_homogeneity_45_0, .1077; UE_uniformity_45_1.0, .2115; and CE_uniformity_0_1.5, .1267).

We found that all of the female patients in the whole data set were nonsmokers (96 of 96). Considering that the inclusion of smoking status into the model along with sex may have introduced a confounding effect, we excluded smoking status from the candidate factors. A Cox regression analysis identified radiomics signature, clinical stage, histologic grade, sex, and age as independent risk factors (radiomics signature: HR: 1.77, 95% CI: 1.18, 2.65, P = .0055; clinical stage: HR: 1.72, 95% CI: 1.25, 2.35, P = .0008; histologic grade: HR: 4.20, 95% CI: 1.43, 12.37, P = .0092; sex: HR: 2.1285, 95% CI: 1.11, 4.09; P = .0236; and age: HR: 2.46, 95% CI: 1.09, 5.54, P = .0303).

Assessment of Incremental Value of Radiomics Signature in Individual DFS Performance

The radiomics nomogram and clinical-pathologic nomogram are presented in Figure 4a. The calibration curves of the nomograms for the probability of tumor recurrence at 1, 2, or 3 years after surgery are shown in Figure 4b and they showed better agreement between the estimation with the radiomics nomogram and actual observation. C-index, concordance probability, and AIC estimates for the different models, including the radiomics signature, traditional system, radiomics nomogram, and clinical-pathologic nomogram, are listed in Table 3. The radiomics signature and the AJCC staging system had similar C-indexes (0.629 for the tumor stage and 0.617 for the radiomics signature, P = .10). The discrimination performance of the radiomics signature improved when it was integrated into the radiomics nomogram along with the clinical-pathologic risk factors (C-index for the radiomics nomogram: 0.720; 95% CI: 0.71, 0.73). Compared with either the AJCC 7th staging system (C-index: 0.629; 95% CI: 0.62, 0.64) or the clinical-pathologic nomogram (C-index: 0.691; 95% CI: 0.68, 0.70), the radiomics nomogram showed a better discrimination capability ($P < .0001$ for each comparison). A histogram of the estimated freedom from tumor relapse obtained from the nomogram demonstrated that heterogeneity of the estimated probabilities of the time to recurrence existed within each traditional stage (Fig E5 [online]). Among all of the classifiers or models, the radiomics nomogram yielded the lowest AIC (417.6) and the highest concordance probability (0.738). Furthermore, the inclusion of the radiomics signature in the clinical-pathologic nomogram yielded a total net reclassification improvement of 0.182 (95% CI:

Figure 3

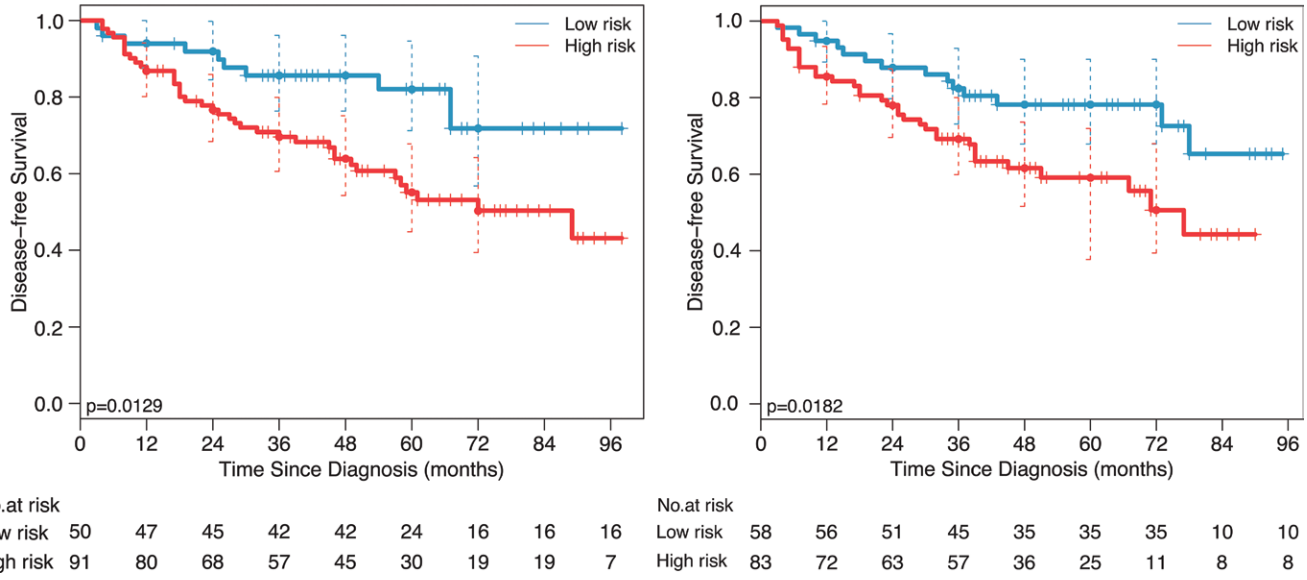

Figure 3: Graphs show results of Kaplan-Meier survival analyses according to the radiomics signature for patients in the training data set (left) and those in the validation data set (right). A significant association of the radiomics signature with the DFS was shown in the training data set, which was then confirmed in the validation data set. Dashed line = two-sided CI of the survival curves (solid line).

0.018, 0.308; P = .02), which showed improved classification accuracy for survival outcomes.

A decision curve analysis showed that the radiomics nomogram had a higher overall net benefit than the clinical-pathologic nomogram and the traditional staging system across the majority of the range of reasonable threshold probabilities (Fig 5).

Discussion

This study extends the analysis of individual imaging texture features to an "-omics"-based survival estimation approach. A multi-feature–based radiomics signature was identified to be an independent factor for DFS in patients with early-stage NSCLC in this study, with incremental value to the traditional staging system and other clinical-pathologic risk factors for individualized DFS estimation. The radiomics signature successfully stratified those patients into high-risk and low-risk groups, with significant differences in the 3-year DFS. The combination of the radiomics signature, traditional staging system, and other clinical-pathologic risk factors into a radiomics nomogram performed better than the traditional staging system and clinical-pathologic nomogram, which well demonstrated the incremental value of the radiomics signature for individualized DFS association in patients with early-stage NSCLC.

The identified signature consisted of the following features: CE_kurtosis_0, CE_uniformity_0_0, CE_homogeneity_45_0, UE_uniformity_45_1.0, and CE_uniformity_0_1.5, which are consistent with results of recent studies of risk stratification (12,17,18). According to the radiomics hypothesis, the intratumoral heterogeneity assessed through imaging could be the expression of genomic heterogeneity, which would indicate worse prognosis, as tumors with more genomic heterogeneity are more likely to develop a resistance to treatment and to metastasize (12,34). Findings of previous studies have supported the above hypothesis that proteogenomic and phenotypic information of the tumor can be inferred from radiologic images (35,36). However, interpreting the complex associations between the biologic processes and radiomics features still remains an intractable challenge (9). On one hand, the biologic processes involve multiple interacting components; on the other hand, the maximized information obtained from computer-based radiologic image analysis is far beyond the depiction by visual inspection, which is meanwhile hard to elaborate in the biologic fashion. Similar to the challenge encountered when selecting valuable biomarkers out of thousands of candidate genomic factors, it is difficult to correlate a single radiomics-based factor with a pathophysiologic basis in an intuitive way; instead, construction of multi-factor panels is a more common approach for outcome estimation in the "-omics" setting (35,36). As demonstrated in the current study, the identified signature could be suggestive of survival outcomes, which sustained the idea that the radiomics signature has the ability to capture intratumoural heterogeneity in a noninvasive way and thus is associated with patient prognosis. Further work in radiogeomics is needed to establish a biologic rationale for textural heterogeneity to determine potential radiomics-biologics correlates.

Figure 4

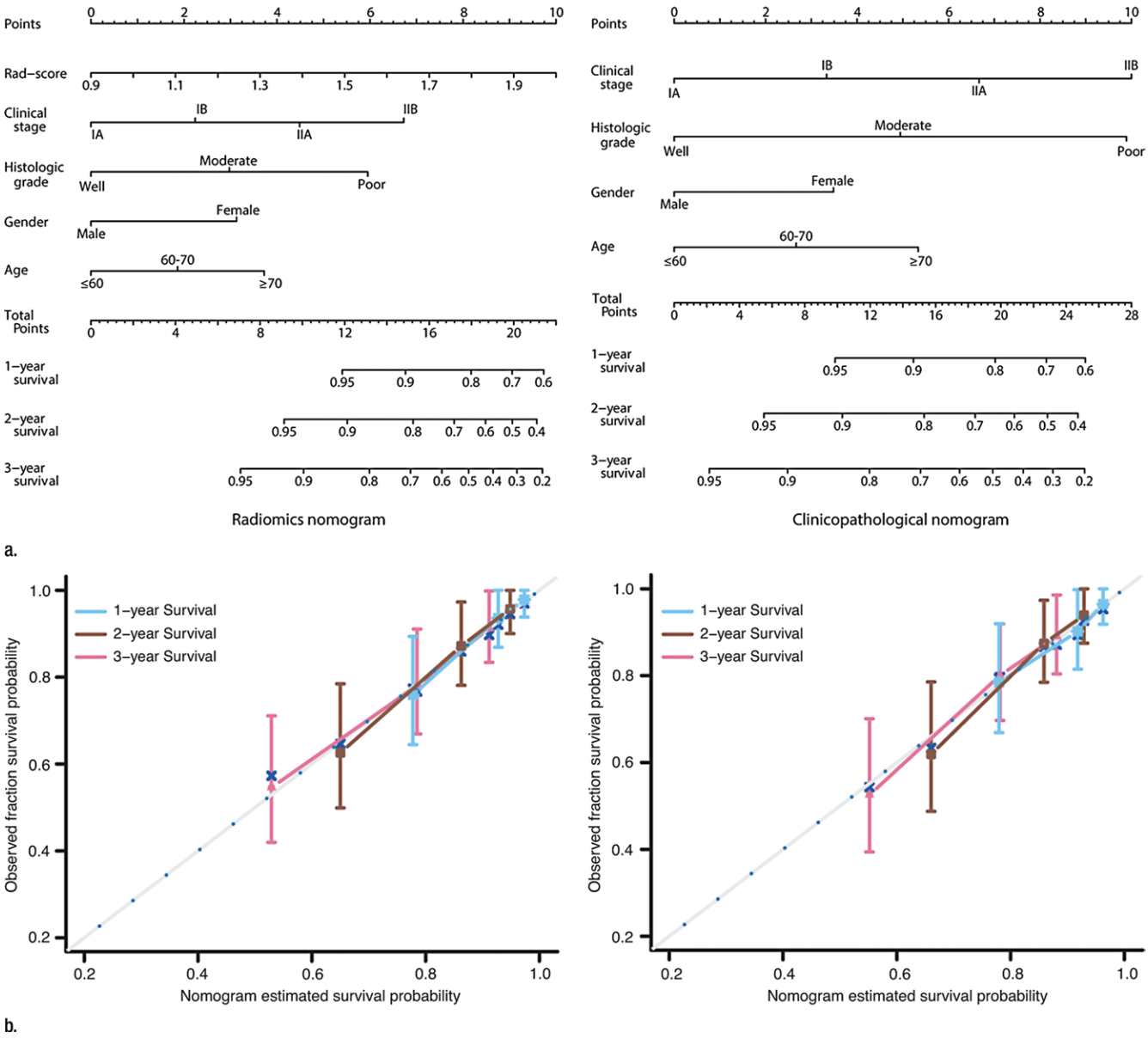

Figure 4: Use of the constructed radiomics nomogram and the clinical-pathologic nomogram to estimate the risk of cancer recurrence for early-stage NSCLC, along with the assessment of the model calibration. **(a)** Radiomics nomogram (left) and clinical-pathologic nomogram (right). Locate the patient's Rad-score on the Rad-score axis. Draw a line straight upward to the points' axis to determine how many points toward the probability of DFS the patient receives for his or her Rad-score. Repeat the process for each variable. Sum the points achieved for each of the risk factors. Locate the final sum on the Total Point axis. Draw a line straight down to find the patient's probability of DFS. **(b)** Calibration curves for the radiomics nomogram (left) and the clinical-pathologic nomogram (right) show the calibration of each model in terms of the agreement between the estimated and the observed 1-, 2-, and 3-year outcomes. Nomogram-estimated DFS is plotted on the x-axis; the observed tumor relapse rate is plotted on the y-axis. Diagonal dotted line = a perfect estimation by an ideal model, in which the estimated outcome perfectly corresponds to the actual outcome. Solid line = performance of the nomogram, a closer lining of which with the diagonal dotted line represents a better estimation.

Although none of these individual features successfully stratified patients into high-risk and low-risk groups in the validation data set, the multi-feature–based radiomics signature could be used to successfully estimate the 3-year DFS. This result is not surprising, because a nonsignificant statistical association with the outcome does not necessarily suggest its unimportance, given the potential nuances that may exist in the data set or the confounding effect by other risk factors

Table 3

Performance of Models

Model	C-Index	95% CI	*P* Value	Concordance Probability	AIC
Radiomics signature	0.617*	0.60, 0.63	.1000*	0.595	431.1
AJCC 7th Edition staging system	0.629*†	0.62, 0.64	<.0001†	0.652	429.8
Radiomics nomogram	0.720†‡	0.71, 0.73	<.0001‡	0.738	417.6
Clinicopathologic nomogram	0.691‡	0.68, 0.70	<.0001‡	0.716	422.9

* The C-index was compared between the radiomics signature and the AJCC 7th Edition staging system.

† The C-index was compared between the AJCC 7th Edition staging system and the radiomics nomogram.

‡ The C-index was compared between the radiomics nomogram and the clinicopathologic nomogram.

Figure 5

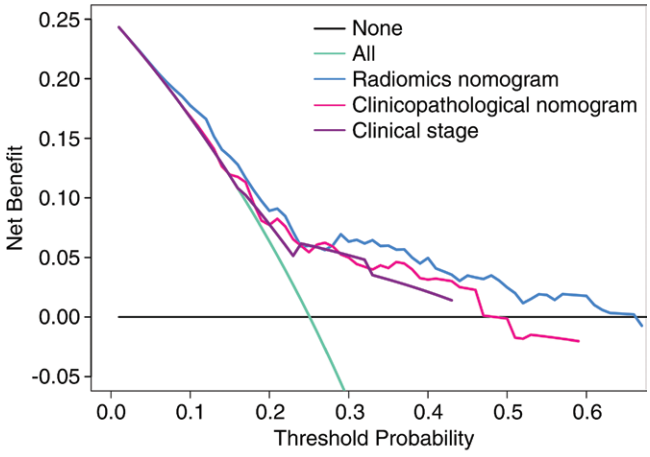

Figure 5: Decision curve analysis for each model. The y-axis measures the net benefit. The net benefit was calculated by summing the benefits (true-positive results) and subtracting the harms (false-positive results), weighting the latter by a factor related to the relative harm of an undetected cancer compared with the harm of unnecessary treatment. The radiomics model had the highest net benefit compared with both the other models and simple strategies such as follow-up of all patients (green line) or no patients (horizontal black line) across the full range of threshold probabilities at which a patient would choose to undergo imaging follow-up.

(37). In addition, the fact that doctors naturally integrate multiple manifestations of disease in the patients to make an estimation and determine consequent therapy rather than focus on a single symptom in clinical practice underscores the necessity of multivariable estimation (37,38). In the meanwhile, the complex nature and biologic processes of malignancy involve multiple interacting components, which may be better reflected when taking into account the interactions between different features. As the first attempt, to our knowledge, to address the issue of survival estimation using a multi-component radiomics signature, our study supported the suggestion that multiple variables could provide a more statistically robust approach (22,37,38). It is noteworthy that CE_kurtosis_0, CE_homogeneity_0_45, UE_uniformity_1.0_45, and UE_uniformity_1.5_0 were observed to be associated with DFS in the combined whole data set. Similarly, it was found that although only one of six selected micro-RNAs was associated with DFS in patients with stage II colon cancer in a testing data set, a validation data set with a larger sample size successfully exhibited the expected significant association (8). These findings underscore the limit in sample size of the validation data set that may have contributed to the failure of risk stratification.

Unlike prior prognostic investigations that mostly analyzed patients with all stages of disease, our current study focused exclusively on patients with early stage. Note that when patients were stratified by clinical disease stage, there was a difference in DFS for both clinical stage IA and IB, which suggests that heterogeneity existed in the survival outcomes, even within the same clinical stage (stage I). Within stages IA and IB, the radiomics signature successfully identified high-risk patients with poor survival outcomes, for whom more intensified treatment was needed. Unfortunately, the radiomics signature failed to discriminate subgroups within stage IIA (n = 6) and stage IIB (n = 44), which may largely be attributed to the limited size of the study population.

Furthermore, the performance of the radiomics signature and the radiomics nomogram was demonstrated for the estimation of DFS in our study. Compared with the long-term outcome overall survival, DFS is an end point that avoids extended follow-up and enables earlier adjustment of therapy (39). Thus, our study may present a more efficient tool that enables earlier personalized treatment.

It has been suggested that the use of a single considered strong risk factor, in the absence of any modeling, could hardly assess the comprehensive postoperative outcome of individual patients (29). A statistical model, such as the nomogram, which takes into account multiple risk factors by assigning a total number of points to each patient, is imperative. A recently established nomogram integrating multiple clinical-pathologic risk factors successfully estimated the overall survival for patients with resected NSCLC (C-index: 0.71) (40).

Recently, the combination of prognostic molecular signatures and clinical risk factors has shown enhanced prognostic accuracy (8,20). Incorporating the positron emission tomography standardized uptake value into a nomogram along with the clinical-pathologic risk factors to estimate the recurrence-free survival for lung adenocarcinoma has exhibited better estimation than conventional TNM staging (C-index: 0.632) (41). As the first study (to our knowledge) performed with a radiomics signature in survival estimation for patients with early-stage NSCLC, our study demonstrated that the combined radiomics-clinical-pathologic nomogram achieved superior prognostic performance than either the radiomics signature or the clinical-pathologic nomogram alone, with a higher C-index, better calibration, and positive net reclassification improvement. Although the usefulness of the proposed nomogram lacked external validation, the decision curve analysis demonstrated that the radiomics nomogram was superior to both the clinical-pathologic nomogram and the clinical staging system across the majority of the range of reasonable threshold probabilities, which in the meanwhile indicated that the radiomics signature added incremental value to the traditional staging system and other clinical-pathologic risk factors for individualized estimation.

The limitations of this study included the relatively small sample size, the retrospective nature of data collection, and the lack of external validation. Although the preferred design should be a prospective longitudinal cohort study, where there is full control for ensuring that no bias is introduced for all relevant risk factors and outcomes and minimizing the loss to follow-up (37), the protracted length of a prospective longitudinal cohort study in early-stage NSCLC (because of the long wait needed for survival outcomes) may make the research daunting (42). Although a large-scale independent prospective multicenter validation cohort is warranted to assess the generalizability of the reported findings, the decision curve analysis used in this study, which enables the evaluation of clinical relevance without the requirement for additional validation data in a traditional decision-analytic approach (33), justified that the identified radiomics signature and radiomics nomogram hold great potential for clinical application in postoperative outcome estimation.

In conclusion, the identified radiomics signature has the potential to be used as a biomarker for risk stratification for DFS in patients with early-stage NSCLC. The radiomics nomogram described here, which well demonstrated the incremental value of the radiomics signature to the traditional staging system and other clinical-pathologic risk factors for individualized DFS estimation, may serve as a potential tool to guide individual postoperative care for those patients, although this will require further external validation before widespread implementation in clinical practice.

Disclosures of Conflicts of Interest: Y.H. disclosed no relevant relationships. **Z.L.** disclosed no relevant relationships. **L.H.** disclosed no relevant relationships. **X.C.** disclosed no relevant relationships. **D.P.** disclosed no relevant relationships. **Z.M.** disclosed no relevant relationships. **Cuishan Liang** disclosed no relevant relationships. **J.T.** disclosed no relevant relationships. **Changhong Liang** disclosed no relevant relationships.

References

1. Brawley OW. Avoidable cancer deaths globally. CA Cancer J Clin 2011;61(2):67–68.
2. Chansky K, Sculier JP, Crowley JJ, et al. The International Association for the Study of Lung Cancer Staging Project: prognostic factors and pathologic TNM stage in surgically managed non-small cell lung cancer. J Thorac Oncol 2009;4(7):792–801.
3. Scott WJ, Howington J, Feigenberg S, Movsas B, Pisters K; American College of Chest Physicians. Treatment of non-small cell lung cancer stage I and stage II: ACCP evidence-based clinical practice guidelines (2nd edition). Chest 2007;132(3 Suppl):234S–242S.
4. Crinò L, Weder W, van Meerbeeck J, Felip E; ESMO Guidelines Working Group. Early stage and locally advanced (non-metastatic) non-small-cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2010;21(Suppl 5):v103–v115.
5. Vansteenkiste J, Crinò L, Dooms C, et al. 2nd ESMO Consensus Conference on Lung Cancer: early-stage non-small-cell lung cancer consensus on diagnosis, treatment and follow-up. Ann Oncol 2014;25(8):1462–1474.
6. Ost D, Goldberg J, Rolnitzky L, Rom WN. Survival after surgery in stage IA and IB non-small cell lung cancer. Am J Respir Crit Care Med 2008;177(5):516–523.
7. Halabi S, Lin CY, Kelly WK, et al. Updated prognostic model for predicting overall survival in first-line chemotherapy for patients with metastatic castration-resistant prostate cancer. J Clin Oncol 2014;32(7):671–677.
8. Zhang JX, Song W, Chen ZH, et al. Prognostic and predictive value of a microRNA signature in stage II colon cancer: a microRNA expression analysis. Lancet Oncol 2013;14(13):1295–1306.
9. Tran B, Dancey JE, Kamel-Reid S, et al. Cancer genomics: technology, discovery, and translation. J Clin Oncol 2012;30(6):647–660.
10. Hofman V, Ilie M, Long E, et al. Immunohistochemistry and personalised medicine in lung oncology: advantages and limitations [in French]. Bull Cancer 2014;101(10):958–965.
11. Kuo MD, Jamshidi N. Behind the numbers: decoding molecular phenotypes with radiogenomics--guiding principles and technical considerations. Radiology 2014;270(2):320–325.
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(4):441–446.
13. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
14. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. Magn Reson Imaging 2012;30(9):1234–1248.
15. Altazi B, Fernandez D, Zhang G, Biagioli M, Moros E, Moffitt HL. SU-E-J-258: prediction of cervical cancer treatment response using radiomics features based on F18-FDG uptake in PET images. Med Phys 2015;42(6):3326.
16. Li H, Lan L, Drukker K, Perou C, Giger M. TU-AB-BRA-08: radiomics in the analysis of breast cancer heterogeneity on DCE-MRI. Med Phys 2015;42(6):3588.
17. Ganeshan B, Panayiotou E, Burnand K, Dizdarevic S, Miles K. Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival. Eur Radiol 2012;22(4):796–802.

18. Win T, Miles KA, Janes SM, et al. Tumor heterogeneity and permeability as measured on the CT component of PET/CT predict survival in patients with non-small cell lung cancer. Clin Cancer Res 2013;19(13):3591–3599.

19. Birkhahn M, Mitra AP, Cote RJ. Molecular markers for bladder cancer: the road to a multimarker approach. Expert Rev Anticancer Ther 2007;7(12):1717–1727.

20. Mitra AP, Lam LL, Ghadessi M, et al. Discovery and validation of novel expression signature for postcystectomy recurrence in high-risk bladder cancer. J Natl Cancer Inst 2014;106(11):dju290.

21. Mitra AP, Pagliarulo V, Yang D, et al. Generation of a concise gene panel for outcome prediction in urinary bladder cancer. J Clin Oncol 2009;27(24):3929–3937.

22. Moons KG, Kengne AP, Woodward M, et al. Risk prediction models. I. Development, internal validation, and assessing the incremental value of a new (bio)marker. Heart 2012;98(9):683–690.

23. Rami-Porta R, Crowley JJ, Goldstraw P. The revised TNM staging system for lung cancer. Ann Thorac Cardiovasc Surg 2009;15(1):4–9.

24. Harrell FE Jr. Regression modeling strategies with applications to linear models, logistic and ordinal regression, and survival analysis. New York, NY: Springer-Verlag, 2015.

25. Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med 1997;16(4):385–395.

26. 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(13):3001–3008.

27. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clin Cancer Res 2004;10(21):7252–7259.

28. Buyske S, Fagerstrom R, Ying Z. A class of weighted log-rank tests for survival data when the event is rare. J Am Stat Assoc 2000;95(449):249–258.

29. Kattan MW. Judging new markers by their ability to improve predictive accuracy. J Natl Cancer Inst 2003;95(9):634–635.

30. Gonen M, Heller G. Concordance probability and discriminatory power in proportional hazards regression. Biometrika 2005;92(4):965–970.

31. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA 1982;247(18):2543–2546.

32. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med 2011;30(1):11–21.

33. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26(6):565–574.

34. Campbell PJ, Yachida S, Mudie LJ, et al. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature 2010;467(7319):1109–1113.

35. Rutman AM, Kuo MD. Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging. Eur J Radiol 2009;70(2):232–241.

36. Kuo MD, Gollub J, Sirlin CB, Ooi C, Chen X. Radiogenomic analysis to identify imaging phenotypes associated with drug response gene expression programs in hepatocellular carcinoma. J Vasc Interv Radiol 2007;18(7):821–831.

37. 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(1):W1–W73.

38. Hayward RA, Kent DM, Vijan S, Hofer TP. Multivariable risk prediction can greatly enhance the statistical power of clinical trial subgroup analysis. BMC Med Res Methodol 2006;6:18.

39. 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(7):619–626.

40. Liang W, Zhang L, Jiang G, et al. Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. J Clin Oncol 2015;33(8):861–869.

41. Yang HC, Kim HR, Jheon S, et al. Recurrence risk-scoring model for stage I adenocarcinoma of the lung. Ann Surg Oncol 2015;22(12):4089–4097.

42. Hellmann MD, Chaft JE, William WN Jr, et al. Pathological response after neoadjuvant chemotherapy in resectable non-small-cell lung cancers: proposal for the use of major pathological response as a surrogate endpoint. Lancet Oncol 2014;15(1):e42–e50.