Imaging features from pretreatment CT scans are associated with clinical outcomes in nonsmall-cell lung cancer patients treated with stereotactic body radiotherapy

Qian Li
Department of Radiology, National Clinical Research Center of Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin Medical University Cancer Institute and Hospital, Tianjin, China.
Department of Cancer Imaging and Metabolism, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA.

Jongphil Kim
Department of Biostatistics and Bioinformatics, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA.

Yoganand Balagurunathan
Department of Cancer Imaging and Metabolism, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA.

Ying Liu
Department of Radiology, National Clinical Research Center of Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin Medical University Cancer Institute and Hospital, Tianjin, China.

Kujtim Latifi
Department of Radiation Oncology, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA.

Olya Stringfield and Alberto Garcia
Department of Cancer Imaging and Metabolism, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA.

Eduardo G. Moros
Department of Cancer Imaging and Metabolism, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA.
Department of Radiation Oncology, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA.

Thomas J. Dilling
Department of Radiation Oncology, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA.

Matthew B. Schabath
Department of Cancer Epidemiology, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA

Zhaoxiang Ye[a)]
Department of Radiology, National Clinical Research Center of Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin Medical University Cancer Institute and Hospital, Tianjin, China.

Robert J. Gillies[a)]
Department of Cancer Imaging and Metabolism, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA.

(Received 18 September 2016; revised 29 March 2017; accepted for publication 12 April 2017; published 24 June 2017)

Purpose: To investigate whether imaging features from pretreatment planning CT scans are associated with overall survival (OS), recurrence-free survival (RFS), and loco-regional recurrence-free survival (LR-RFS) after stereotactic body radiotherapy (SBRT) among nonsmall-cell lung cancer (NSCLC) patients.
Patients and methods: A total of 92 patients (median age: 73 yr) with stage I or IIA NSCLC were qualified for this study. A total dose of 50 Gy in five fractions was the standard treatment. Besides clinical characteristics, 24 "semantic" image features were manually scored based on a point scale (up to 5) and 219 computer-derived "radiomic" features were extracted based on whole tumor segmentation. Statistical analysis was performed using Cox proportional hazards model and Harrell's C-index, and the robustness of final prognostic model was assessed using tenfold cross validation by dichotomizing patients according to the survival or recurrence status at 24 months.
Results: Two-year OS, RFS and LR-RFS were 69.95%, 41.3%, and 51.85%, respectively. There was an improvement of Harrell's C-index when adding imaging features to a clinical model. The model for OS contained the Eastern Cooperative Oncology Group (ECOG) performance status [Hazard Ratio (HR) = 2.78, 95% Confidence Interval (CI): 1.37–5.65], pleural retraction (HR = 0.27, 95% CI: 0.08–0.92), F2 (short axis × longest diameter, HR = 1.72, 95% CI: 1.21–2.44) and F186 (Hist-Energy-L1, HR = 1.27, 95% CI: 1.00–1.61); The prognostic model for RFS contained vessel attachment (HR = 2.13, 95% CI: 1.24–3.64) and F2 (HR = 1.69, 95% CI: 1.33–2.15); and the model for LR-RFS contained the ECOG performance status (HR = 2.01, 95% CI: 1.12–3.60) and F2 (HR = 1.67, 95% CI: 1.29–2.18).

Conclusions: Imaging features derived from planning CT demonstrate prognostic value for recurrence following SBRT treatment, and might be helpful in patient stratification. © *2017 American Association of Physicists in Medicine* [https://doi.org/10.1002/mp.12309]

Key words: computed tomography, image features, radiomics, semantics, stereotactic body radiotherapy (SBRT), survival

1. INTRODUCTION

Stereotactic body radiotherapy (SBRT) is a technique that allows delivery of very high doses of radiation per fraction to a focused target with relatively less irradiation of normal tissues. SBRT is a treatment option for early stage nonsmall cell lung cancer (NSCLC) patients who are medically inoperable or who refuse surgery, or have high surgical risk. The treatment outcomes of SBRT are comparable to surgery,[1–3] and the disease-free survival rates are reported to be 48.3% at 3 yr.[1] In patients that progressed after SBRT, 84% of cases occurred within the first 2 yr,[4] in which period radiation-induced lung injury is invariably observed. Distinguishing between recurrence and lung injury is problematic. Previous studies[5–7] have identified standardized uptake value of FDG-PET and CT-based "high-risk features" that were correlated with recurrence, but the most reliable features cannot predict recurrence until 12 months after treatment. A delay in detecting recurrence may lead to delays in implementation of salvage therapy. Thus, it is important to identify factors associated with recurrence so that patients may benefit early from salvage treatment or systemic therapy. Furthermore, tumor recurrence after SBRT may be due to dose insufficiency.[8] Hence, if tumor response could be predicted prior to treatment, dose prescription could be adjusted to improve local control. Prior studies on staging PET-CT images[9–12] have been inconclusive. A few studies of pretreatment CT[13–15] showed that tumor growth rate and pleural attachment were associated with survival or recurrence.

However, a thorough quantitative analysis of pretreatment CT images is lacking. There are a large number of features that can be used to characterize lung tumors; not only manually scored radiological features (semantics) but also computer-derived features (radiomics), which may improve predictive accuracy. Recently, the emerging quantitative imaging field, also known as radiomics, has shown great potential for prognosis in a variety of cancers.[16–19] Radiomics enables the high-throughput extraction of large numbers of quantitative features from medical images.[20,21] Most of these features cannot be detected by radiologists directly, thus provide complementary information to semantic features. A previous study[22] showed that the predictive power of clinical and size features could be significantly improved by incorporating radiomic features in stage III NSCLC treated with chemo-radiation therapy.

In this study, 24 radiological image traits were systematically scored by two radiologists on a point scale (up to 5) and 219 radiomic features were automatically extracted on the delineated regions of interest on the pretreatment CT images used for SBRT planning. Then clinical, semantic and radiomic features were combined to create multivariate prognostic models for overall survival (OS), recurrence free survival (RFS) and loco-regional recurrence free survival (LR-RFS) after SBRT. In this paper, we show that simulation CT scans and patient clinical data are prognostic in NSCLC patients treated with SBRT.

2. PATIENTS AND METHODS

2.A. Patients

This retrospective analysis was approved by the institutional review board (IRB #105996). The cohort had 92 evaluable patients (from 213 patients considered) treated with SBRT between January 2009 and July 2013. The inclusion criteria included patients with primary NSCLCs confirmed by biopsy without prior lung irradiation or prior lung tumor history, and TNM stage $\leq$ IIA (node negative). Excluded patients included those with more than one lung tumor or concurrent other tumors (N = 19), nonavailability of biopsy confirming malignancy (N = 9), or lack of data to confirm recurrence (N = 4). In 2 patients, the tumors were located at the bronchus and could not be scored. Patients who had been treated with lung cancer before were also excluded (N = 45). Patients were required to be followed up for at least 2 yr, or until they developed recurrence or passed away within 2 yr. Forty-two patients were excluded because of lack of follow-up. Clinical data included age at diagnosis, gender, clinical TNM stage, clinical T stage, smoker status, pack-years smoking history, O_2 dependence or not, and Eastern Cooperative Oncology Group (ECOG) performance status.

2.B. CT scanning and treatment protocol

The patients were placed in supine position, typically underwent abdominal compression, and immobilized with a BodyFix double-vacuum cradle (Elekta AB, Stockholm, Sweden). CT simulation scans were performed using a helical four dimensional (4D) CT scanner (Philips Brilliance CT, Philips Medical Systems, Cleveland, OH, USA). Scanning parameters were as follows: 120 kVp, 224 mA, and 3 mm reconstruction slice thickness. Average-CT or 50% phase-CT images were used for analysis. The CT scans were performed at a median time of 15 days before SBRT.

The heterogeneity corrected collapsed cone convolution (CCC) algorithm were used for planning. Either 3D conformal or volumetric arc therapy (VMAT) techniques were used, with photon beam energies ranging from 6 to 15 MV. The

patients were treated on a Trilogy or a TrueBeam medical linear accelerator (Varian Medical Systems, Palo Alto, CA, USA) equipped with a 120-leaf Millennium multi-leaf collimator (5-mm leaves in the central portion of the field). Daily image guidance was provided by cone beam computed tomography (CBCT), with alignment to the visible tumor on the planning scan. Dose voxel size was kept at 2 mm.

All patients were treated with SBRT using a risk-adapted fractionation scheme. Briefly, 50 Gy in five fractions was the standard treatment option used for most of the patients. A regimen of 48 Gy in four fractions was an alternative while 60 Gy in eight fractions was used for "central" lesions (near the hilum/proximal airways, peri-mediastinal/epicardial).

2.C. Follow-up and clinical outcomes

Follow-up evaluations were based on CT images and clinical examination performed every 3 months in the first 2 yr after SBRT, then every 4–6 months for the following 3 yr, and annually thereafter. An ^{18}F-FDG-PET/CT scan was recommended when recurrence or metastasis was suspected. Local recurrence was defined as progression of the original primary lesion or new tumors in the same lobe of the primary tumor. Regional recurrence was defined as hilar or mediastinal lymph node metastasis. Distant metastasis was defined as tumors in other lobes of the lung or outside the lung. Recurrence was confirmed by biopsy, PET/CT, or CT images at follow-up. The recurrence date was recorded as the date of first CT or PET/CT scan that showed signs of progression.

We evaluated three clinical endpoints in this study: overall survival (OS), recurrence-free survival (RFS), and loco-regional recurrence-free survival (LR-RFS). OS was calculated from the start date of SBRT to the last follow-up date (for censored cases) or date of death. RFS was calculated from the start date of SBRT to the date of local, regional or distant metastasis, or the date of death, or censored at the last follow-up date. LR-RFS was calculated from the start date of SBRT to the date of local or regional recurrence or death, or censored at the last follow-up date.

2.D. Image assessment

Pretreatment planning CT images were reviewed using both mediastinal (width, 350 HU; level, 40 HU) and lung (width, 1500 HU; level, −600 HU) window settings. A clinical radiologist with 4 yr of experience (Q. L.) in thoracic imaging interpreted all the CT images. In order to compare the semantic interpretations, a subset of 40 CT scans randomly chosen was reviewed by another radiologist (Y.L.). Both of them were blinded to clinical and histologic findings. The concordance between the two radiologists for the 24 semantic features was evaluated. Description of the features can be found in Table S1.

Automatic extraction of radiomic features was accomplished using Definiens Developer® (Munich, Germany) image analysis software. Firstly, the preprocessing performed automatic organ segmentation with the main goal of segmenting the aerated lung with correct identification of the pleural wall in order to facilitate the semi-automatic segmentation of juxtapleural lesions. In some cases, the pulmonary boundaries were further manually corrected. Then, whole tumor segmentation was done by adopting the ensemble click and grow segmentation method developed by our group.[23] The delineated region was later corrected by the radiologist (Q.L.) to encompass the tumor region. At last, the 219 3D image features from size, shape and texture categories were extracted in the delineated region of interest (Fig. 1, features and their definitions are listed in Table S2). Examples of tumor segmentations can be seen in Figs. 2(c) and 2(d).

2.E. Statistical analysis

Agreement between the two readers was measured by intra-class correlation coefficient (ICC) for continuous variables and (weighted) Kappa index for categorical variables (see Table S1 for the whole list of variables). The Kappa value was interpreted as follows: < 0: poor agreement; 0.01 to 0.2: slight agreement; 0.21 to 0.4: fair agreement; 0.41 to 0.6: moderate agreement; 0.61 to 0.8: substantial agreement; > 0.8: almost perfect agreement.[24]

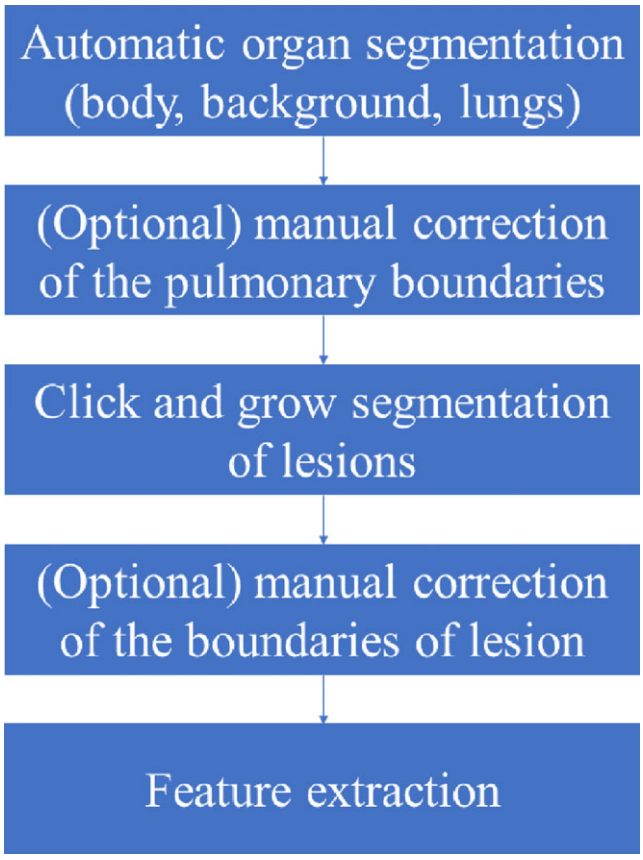

FIG. 1. The processes of radiomic feature extraction using definiens. [Color figure can be viewed at wileyonlinelibrary.com]

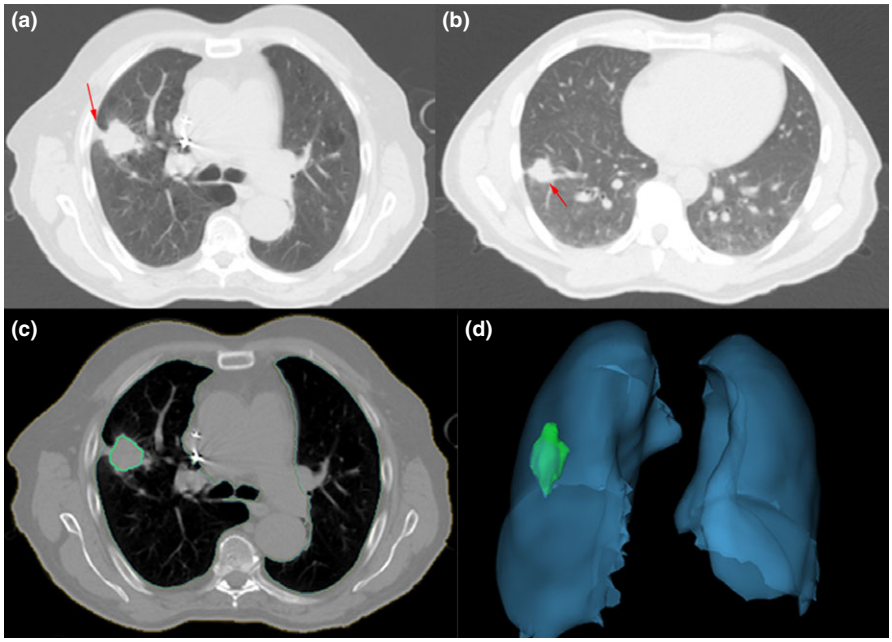

FIG. 2. Examples of CT images showing typical semantic features and tumor segmentation (a): tumor with pleural retraction; (b): tumor with vessel attachment: (c): one slice showing lung and tumor segmentation; (d): 3D view of lung and tumor segmentation). [Color figure can be viewed at wileyonlinelibrary.com]

2.E.1. Feature screening

Among all the semantic features, distribution, contour, attenuation, and calcification were excluded, because the majority of patients fell into the same category and thus these features are not prognostic of endpoints. For radiomic features, Pearson's correlation analysis was performed to identify redundant features and to address collinearity. As such, we eliminated 46 of the 219 features (Table S5) that were highly co-dependent (Pearson's correlation > 0.95). Hence, this study investigated associations of 20 semantic features and 173 radiomic features with patient outcomes.

2.E.2. Univariate and multivariate analysis

All statistical analyses were performed using SAS software (version 9.4, Cary, NC) and all *P*-values were two-sided. The Cox proportional hazards model[25] (see equation below) was used to explore the association with OS, RFS and LR-RFS. The false discovery rate (FDR) with *q*-value of < 0.1[26] was used to control the expected proportion of incorrect rejection for radiomic features. Features with *P*-value of < 0.1 in univariate analysis for clinical and semantic models and those with *q*-value of < 0.1 for radiomic model were incorporated into the multivariable analysis (either stepwise selection or backward elimination method was applied). The hazard ratio (HR) and 95% confidence interval (CI) were reported. The final multivariable model was built by combining clinical with imaging features (semantics and radiomics), and a stepwise selection method was utilized. Harrell's C (Concordance) index[27] was computed to assess the prognostic power of a model. The higher the C index is, the more accurate the Cox model predict. The model with highest Harrell's C-index was selected for the prognostic model. The risk scores of OS, RFS and LR-RFS were developed based on the Cox model, accordingly. Patients were dichotomized into low and high risk groups on the basis of their median risk score. Differences in the OS, RFS and LR-RFS between the two groups were estimated and compared by the Kaplan–Meier method.[28,29]

$$\lambda(\mathrm{t,z}) = \lambda_0(t)e^{\beta' z}$$

In this equation, $\lambda(\mathrm{t,z})$ is the hazard function for an individual with a *p*-vector z of covariates (which may include both clinical and imaging features), and $\lambda_0(t)$ is the baseline hazard function without any covariate effect. $\beta' = (\beta_1, \cdots, \beta_p)$ is the *p*-vector of regression coefficients. Exp (β') is the hazard ratio. The risk scores were just calculated based on β'.

2.E.3. Cross validation

The robustness of the final Cox model was accessed by dichotomizing the patients into short- and long-term survival according to the survival or recurrence status at 24 months (censored subjects within 24 months were excluded because those patients are neither short- nor long-term survivors, detailed information in Table S7). Multiple logistic regression model and tenfold cross validation were used. The area under the receiver operating curve (AUC) was calculated along with 95% CI.[30,31]

3. RESULTS

Among the 92 patients, 43 were females (46.7%) and 49 were males (53.3%) and the median age at diagnosis was 73.

The median follow-up time was 39.2 months (range: 24–74 months) in patients who were alive at the date of last follow-up and did not develop recurrence. At the end of this study, 25 patients (27.2%) developed loco-regional recurrence and 26 patients (28.3%) developed distant metastasis (Table S3). Two-year OS, RFS and LR-RFS were 69.95%, 41.3%, and 51.85%, respectively. Detailed information is provided in Table I.

3.A. Reader reproducibility

The agreements between two readers ranged from 0.529–1 (see Table S4). Among all the features, location, lobulation, nodules in primary tumor lobe, border definition, lymphadenopathy, attachment to vessel, peripheral emphysema, and pleural attachment had almost perfect agreement, while thickened adjacent bronchovascular bundles, bubble-like lucency, fissure attachment, nodules in nontumor lobes, spiculation, pleural retraction and contour were in substantial agreement. Peripheral fibrosis and vascular convergence showed moderate agreement. The ICCs for long and short axis diameter were 0.985 (0.972–0.992) and 0.974 (0.950–0.986), respectively.

TABLE I. Clinical and treatment characteristics of NSCLC patients treated with SBRT.

Variables	Level	Number	Percentage (%)
Gender	Female	43	46.7
	Male	49	53.3
TNM stage	IA	65	70.7
	IB	23	25.0
	IIA	4	4.3
T stage	1A	34	37.0
	1B	31	33.7
	2A	23	25.0
	2B	4	4.3
Smoker[a]	No	64	71.1
	Yes	26	28.9
Pack-years smoking[a]	< = 43	22	25.3
	44–79	41	47.1
	> = 80	24	27.6
O_2 dependence[a]	No	65	72.2
	Yes	25	27.8
ECOG performance status	0	21	22.8
	1	49	53.3
	2	21	22.8
	3	1	1.1
Dose and fractions	50 Gy in 5 fractions	82	89.1
	48 Gy in 4 fractions	4	4.3
	60 Gy in 5 or 8 fractions	6	6.5
Pathology	Adenocarcinoma	36	39.1
	Squamous cell carcinoma	33	35.9
	NSCLC, nonspecified	20	21.7
	Large cell carcinoma	3	3.3

ECOG, Eastern Cooperative Oncology Group; NSCLC, nonsmall cell lung cancer.

[a]The pack-years smoking history cannot be determined in 5 cases and 2 of them cannot be identified as smoker or nonsmoker. The O_2 dependence information cannot be confirmed in 2 subjects.

3.B. Clinical features and clinical outcomes

Univariable analysis of clinical features demonstrated that T stage and ECOG performance status were significantly associated ($P < 0.05$) with OS, RFS and LR-RFS, while O_2 dependence was marginally associated ($P < 0.1$) with OS and LR-RFS.

In multivariable analysis, only ECOG (HR = 2.66, 95% CI: 1.37–5.18; $P = 0.004$) was associated with OS and T-stage (HR = 2.49, 95% CI: 1.50–4.13; $P = 0.0004$) with RFS. Both T stage (HR = 1.81, 95% CI: 1.01–3.24; $P = 0.047$) and ECOG (HR = 1.99, 95% CI: 1.08–3.65; $P = 0.027$) were independently associated with LR-RFS.

3.C. Semantic features and clinical outcomes

Both long- and short-axis diameters were significantly associated with all three clinical outcome endpoints ($P < 0.05$) in univariable analysis. Additionally, vessel attachment was significantly associated ($P < 0.05$) with RFS and LR-RFS, border definition ($P < 0.0001$) was significantly related with OS and LR-RFS, and pleural retraction ($P = 0.06$) was marginally associated with OS.

In multivariate analysis, long axis diameter ($P = 0.001$, HR = 1.77, 95% CI: 1.26–2.50) and pleural retraction ($P = 0.048$, HR = 0.31, 95% CI: 0.09–0.99) remained independently associated with OS. Short axis diameter ($P = 0.001$, HR = 1.58, 95% CI: 1.21–2.05) and vessel attachment ($P = 0.003$, HR = 2.24, 95% CI: 1.31–3.81) correlated significantly with RFS. Long axis diameter ($P = 0.001$, HR = 1.65, 95% CI: 1.24–2.20) was the only independent indicator of LR-RFS.

3.D. Radiomic features and clinical outcomes

In the univariate analysis, 18, 68, and 8 radiomic features were found to be prognostic to OS, RFS, and LR-RFS, respectively. However, in multivariate analysis, only F2 [short axis × longest diameter (mm^2)] ($P < 0.001$) was independently related with OS (HR = 1.98, 95% CI: 1.44–2.72), RFS (HR = 1.85, 95% CI: 1.47–2.33) and LR-RFS (HR = 1.73, 95% CI: 1.35–2.23). A subsequent analysis with nonsize related features revealed that F15 (AV-Dist-COG-to-Border, $P = 0.001$, HR = 2, 95% CI: 1.31–3.05), F51 (avgLRE, $P = 0.033$, HR = 0.38, 95% CI: 0.15–0.92), F186 (Hist-Energy-L1, $P = 0.018$, HR = 2.60, 95% CI: 1.18–5.74), and F214 (3D-WaveP1-L2-25, $P = 0.027$, HR = 1.39, 95% CI: 1.04–1.85) were independently prognostic of OS. F17 (MIN_Dist_COG_To_Border, $P < 0.0001$, HR = 1.81, 95% CI: 1.39–2.36) and F93 (3D-Laws-35, $P = 0.036$, HR = 0.68, 95% CI: 0.47–0.97) were related with RFS. F17 (MIN_Dist_COG_To_Border, $P = 0.0004$, HR = 1.69, 95% CI: 1.26–2.25) also remained significant in prognosis of LR-

RFS. F15 (AV-Dist-COG-to-Border) and F17 (MIN_Dist_-COG_To_Border) stand for average or minimum distance of center of gravity of the ROI to border. F51 (avgLRE), that is, average long run emphasis, examines runs of similar gray values in an image, and long runs of the same gray value correspond to coarser textures. F186 (Hist-Energy-L1) is a feature based on histogram analysis, measures the energy of HU values within the lesion. F93 (3D Laws features L5 L5 W5 Layer 1) and F214 (3D Wavelet decomposition P1 L2 C3 Layer1) were based on law and wavelet analysis, respectively. These features described the tumor heterogenicity quantitatively and have been demonstrated highly reproducible (test-retest concordance correlation coefficient ≥ 0.90, dynamic range ≥ 0.90) by our previous work.[20] Detailed information about the significant features was listed in Table S6.

3.E. Prognostic model and Prognostic Index

The final prognostic models (Tables II and III) were built by combining the clinical and imaging features together. It can be seen that there was an improvement of the Harrell's C index by adding the imaging features [semantic, size related radiomics feature (F2) or nonsize related radiomics feature (F186)] to the clinical model. In an exploratory analysis, the OS risk score (OSRS), RFS risk score (RFRS) and LR-RFS risk score (LRRS) were developed accordingly (Kaplan–Meier curves in Fig. 3):

- OSRS == 0.541 × (F2–average (F2)) + 0.240 × (F186–average (F186)) + 1.021 (if ECOG was 2 or 3)–1.294 (if pleural retraction was positive)
- RFRS = 0.524 × (F2–average (F2)) + 0.755 (if vessel attachment was positive).
- LRRS = 0.515 × (F2–average (F2)) + 0.696 (if ECOG was 2 or 3).

When patients were dichotomized, these models still remained to be prognostic (Table II). According to the tenfold cross validation, the AUCs of OS, RFS and LR-RFS prognostic model were 0.728, 0.747, and 0.690, respectively.

4. DISCUSSION

In this study, we performed a comprehensive analysis of pretreatment planning CT images among lung cancer patients treated with SBRT and found that besides clinical features (ECOG performance status and T stage), tumor size (F2), pleural retraction, vessel attachment, and Hist-Energy-L1 were also independently associated with OS or recurrence related survival.

Various features related to tumor size, including clinical T stage, semantic features of long- or short-axis diameter, and the radiomic feature F2 (short axis × longest diameter), were consistently prognostic for OS, RFS and/or LR-RFS. These data strongly support the idea that tumor size is a key predictor of outcome in NSCLC patients treated with SBRT.[13] Bhatt et al[32] observed that T-stage was associated with tumor shrinkage during SBRT treatment, with T1 tumors showing greater decrease than T2 tumors. These data imply that patients with larger tumors might benefit from dose escalation (if normal tissue constraints allow). Compared with 2-D semantic features (long/short-axis diameter), 3-D radiomic feature (F2) showed higher HR and Harrell's index in predicting all three kinds of survival, which implies that accurate volumetric measurement of the tumor is key in predicting survival.

TABLE II. Features involved in prognostic models of OS, RFS and LR-RFS.

				Hazard ratio			Validation	
					95% CI			
	Features	Level	*P*-value	Point	Lower	Upper	Logistic regression analysis AUC (95% CI)	10-fold cross-validation AUC (95% CI)
OS	F2[a] (short axis × longest diameter)		0.002	1.72	1.21	2.44	0.753	0.728
	F186[a)] (Hist-Energy-L1)		0.048	1.27	1.00	1.61	(0.627–0.878)	(0.578–0.878)
	ECOG	0 or 1	Reference					
		2 or 3	0.005	2.78	1.37	5.65		
	Pleural retraction	0	Reference					
		1	0.035	0.27	0.08	0.92		
RFS	F2[a] (short axis × longest diameter)		< 0.0001	1.69	1.33	2.15	0.765	0.747
	Vessel attachment	0	Reference				(0.667–0.864)	(0.649–0.844)
		1	0.006	2.13	1.24	3.64		
LR-RFS	F2[a] (short axis × longest diameter)		0.0001	1.67	1.29	2.18	0.691	0.690
	ECOG	0 or 1	Reference				(0.577–0.804)	(0.573–0.807)
		2 or 3	0.020	2.01	1.12	3.60		

[a]Per 1 standard deviation increase. OS, overall survival; RFS, recurrence-free survival; LR-RFS, loco-regional recurrence-free survival; ECOG, Eastern Cooperative Oncology Group, AUC: area under the receiver operating curve, CI: confidence interval.

TABLE III. Harrell's C-index for prognostic models of OS, RFS and LR-RFS.

Models	Features	Harrell's C index		
			95% CI	
		Point	Lower	Upper
OS				
Clinical alone	ECOG	0.614	0.539	0.690
Clinical & imaging features (semantic and radiomics)	ECOG, Pleural retraction, F2[a], F186[a]	0.722	0.637	0.806
RFS				
Clinical alone	T-stage	0.614	0.559	0.669
Clinical & imaging features (semantic and radiomics)	Vessel attachment, F2[a]	0.702	0.642	0.762
LR-RFS				
Clinical alone	T-stage, ECOG	0.625	0.551	0.698
Clinical & imaging features (semantic and radiomics)	ECOG, F2[a]	0.659	0.585	0.733

[a]Per 1 standard deviation increase. OS, overall survival; RFS, recurrence-free survival; LR-RFS, loco-regional recurrence-free survival; F2: short axis × longest diameter [mm]; F186: Hist-Energy-L1; ECOG, Eastern Cooperative Oncology Group.

Shultz[14] and Yamamoto et al[15] found that contacting with the mediastinal pleura or broad attachment to the pleura was negative predictor of loco-regional control or distant metastasis. In our study, pleural attachment was also associated with poorer OS (HR = 1.66), RFS (HR = 1.45) and LR-RFS (HR = 1.3), but without reaching statistical significance ($P > 0.05$). Interestingly, pleural retraction showed prognostic value in OS. Pleural retraction is usually taken as a sign of malignancy.[33] However, Webb[34] suggested that it was not indicative of malignant lesions and did not aid in the differentiation of benign and malignant lesions. In our study, pleural retraction was associated with improved OS. Clearly, further study is needed.

Vessel attachment was an independent poor prognostic factor for RFS in our study. Blood vessel involvement is one of the steps of metastatic process.[35] A previous study of resected NSCLC[36] found that even microscopic vascular invasion was an indicator of poor survival. Tumor attachment to vessel in radiology may reflect pathologic vascular invasion. Tsuchiya et al[37] discovered that the prognosis of stage IA NSCLC with vessel invasion is similar to stage IB NSCLC, but can be improved significantly by postoperative chemotherapy. Therefore, in patients with the sign of vessel attachment, further treatment may be necessary.

ECOG performance status was negatively associated with RFS and also an independent prognostic marker for OS and LR-RFS in this analysis. Patients with worse ECOG performance status tended to have poorer survival, and maybe their dose ought to be adjusted accordingly. Compared with using clinical features only, when imaging

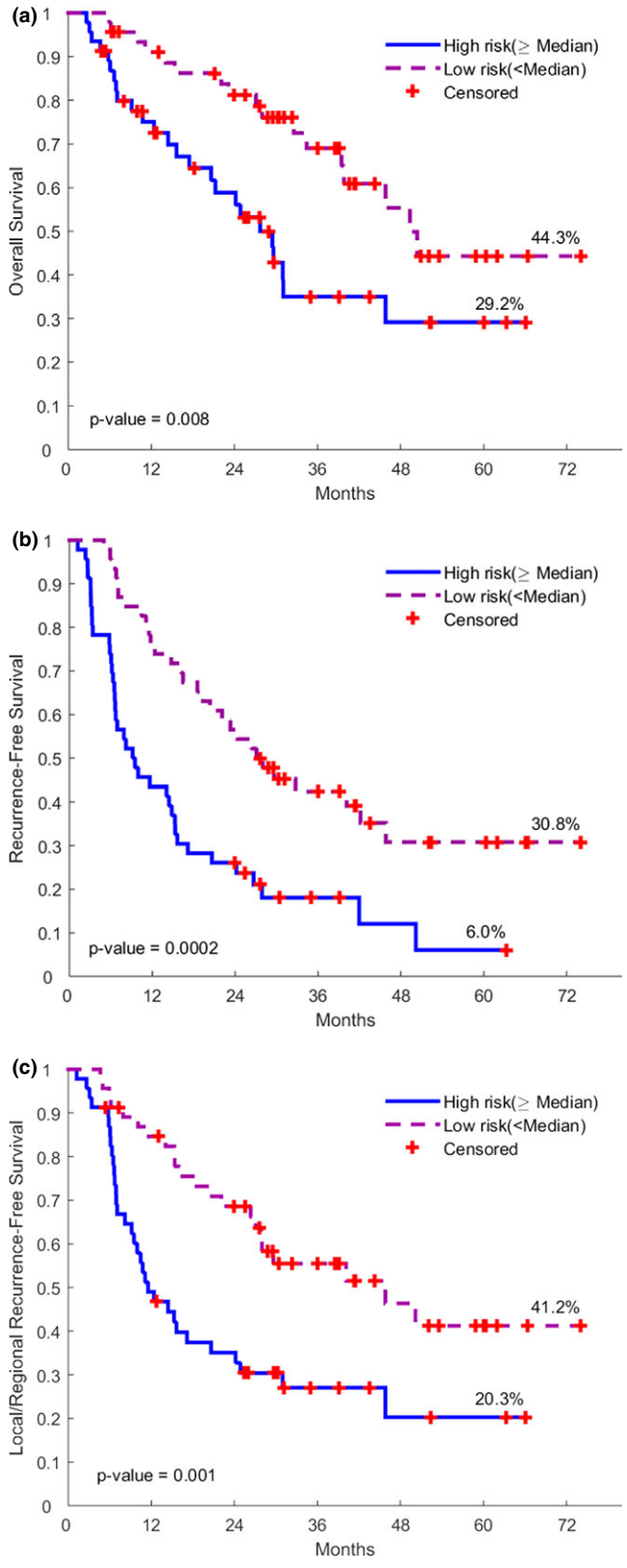

FIG. 3. Kaplan-Meier Plots of OS, RFS and LR-RFS according to the prognostic risk scores incorporating the clinical, semantic, and radiomic features. (OS: overall survival; RFS: recurrence free survival; LR-RFS: loco-regional recurrence free survival). [Color figure can be viewed at wileyonlinelibrary.com]

features were incorporated, there was an improvement of the outcome discriminating ability.

The clinical and semantic features were common features used in daily practice, while radiomic features were quite distinct. Semantically, we would categorize a tumor as homogeneous or heterogeneous. More specifically, we classified the tumor as solid or part-solid, and whether there was bubble-like lucency or calcification inside. In contrast, radiomics enables quantifying tumor (segmented region of interest) heterogenicity using imaging features extracted via high-throughput computing. For example, there are radiomic features about the distribution of the tumor density (histogram), and features that calculate the average of volumes of air spaces inside tumors (Table S2). Radiomic features are complementary to traditional features and has shown great potential in clinical decision support.[16] Previously, it has reported that texture measures of CT images following SBRT could predict recurrence,[6] and our study demonstrated that recurrence could be predicted using radiomic features even before treatment. These features, including AV-Dist-COG-to-Border, MIN_Dist_COG_To_Border, avgLRE, Hist-Energy-L1, along with another Laws and wavelet features, were all independent predictors either of OS, or RFS, or LR-RFS. Huynh et al[38] had similar findings, they also showed that pretreatment radiomic features were prognostic for some outcomes that conventional imaging metrics did not predict, which is different from the findings in our study where clinical, semantic and radiomic features were all prognostic and their combination was better than any category alone. The reason may be as follows. Firstly, more semantic features were used in our study. We performed a systemic semi-quantitative analysis of 24 semantic features (Table S1) by two radiologists, while Huynh et al[38] used only three conventional features related to tumor size in their study. Secondly, we used different radiomic features. The plurality and lack of standardization of radiomic features is one of the problems in the emerging field of the radiomics.[39]

There are several limitations in our study besides its retrospective nature. First, the sample size was comparatively small because of the strict inclusion and exclusion criteria; secondly, the resolution of simulation CT was limited and may have comparatively limited consistent texture related to prognosis; thirdly, an independent validation is needed to further confirm findings. However, we did perform extensive cross-validation analyses to conduct internal validation.

5. CONCLUSION

We showed that recurrence related survival of SBRT patients could be prognosticated prior to treatment from their imaging characteristics. We find tumor size, pleural retraction, vessel attachment along with some radiomic features to be useful in the prognostication. The image features derived from planning CT would be helpful in patient stratification and risk scores could be used to individualize radiotherapy planning for each patient. Finally, we also showed that prognostic models composed of clinical, conventional, and radiomic features performed better than models having only one of these feature categories.

ACKNOWLEDGMENTS

This research was supported by the National Cancer Institute (grants U01 CA143062 and P50 CA119997), State of Florida Department of Health, James & Esther King Biomedical Research Program (grant 2kT01), and in part by Biostatistics Core shared resources at the H. Lee Moffitt Cancer Center & Research Institute, an NCI designated Comprehensive Cancer Center (P30-CA076292).

CONFLICT OF INTEREST

Dr. Gillies reports grants from National Cancer Institute and nonfinancial support from HealthMyne. The other authors declare that they have no conflicts of interest.

[a)]Author to whom correspondence should be addressed. Electronic mails: robert.gillies@moffitt.org, yezhaoxiang@163.com.

REFERENCE

1. Timmerman R, Paulus R, Galvin J, et al. Stereotactic body radiation therapy for inoperable early stage lung cancer. *JAMA*. 2010;303:1070–1076.
2. Shirvani SM, Jiang J, Chang JY, et al. Comparative effectiveness of 5 treatment strategies for early-stage non-small cell lung cancer in the elderly. *Int J Radiat Oncol*. 2012;84:1060–1070.
3. Onishi H, Shirato H, Nagata Y, et al. Stereotactic body radiotherapy (SBRT) for operable stage I non–small-cell lung cancer: can SBRT be comparable to surgery? *Int J Radiat Oncol*. 2011;81:1352–1358.
4. Spratt DE, Wu AJ, Adeseye V, et al. Recurrence patterns and second primary lung cancers after stereotactic body radiation therapy for early-stage non–small-cell lung cancer: implications for surveillance. *Clin Lung Cancer*. 2016;17:177–183.
5. Zhang X, Liu H, Balter P, et al. Positron emission tomography for assessing local failure after stereotactic body radiotherapy for non-small-cell lung cancer. *Int J Radiat Oncol*. 2012;83:1558–1565.
6. Mattonen SA, Palma DA, Haasbeek CJ, Senan S, Ward AD. Early prediction of tumor recurrence based on CT texture changes after stereotactic ablative radiotherapy (SABR) for lung cancer. *Med Phys*. 2014;41:033502.
7. Huang K, Senthi S, Palma DA, et al. High-risk CT features for detection of local recurrence after stereotactic ablative radiotherapy for lung cancer. *Radiother Oncol*. 2013;109:51–57.
8. Chen F, Matsuo Y, Yoshizawa A, et al. Salvage lung resection for non-small cell lung cancer after stereotactic body radiotherapy in initially operable patients. *J Thorac Oncol*. 2010;5:1999–2002.
9. Na F, Wang J, Li C, Deng L, Xue J, Lu Y. Primary tumor standardized uptake value measured on F18-fluorodeoxyglucose positron emission tomography is of prediction value for survival and local control in non–small-cell lung cancer receiving radiotherapy: meta-analysis. *J Thorac Oncol*. 2014;9:834–842.
10. Clarke K, Taremi M, Dahele M, et al. Stereotactic body radiotherapy (SBRT) for non-small cell lung cancer (NSCLC): is FDG-PET a predictor of outcome? *Radiother Oncol*. 2012;104:62–66.
11. Burdick MJ, Stephans KL, Reddy CA, Djemil T, Srinivas SM, Videtic GM. Maximum standardized uptake value from staging FDG-PET/CT does not predict treatment outcome for early-stage non–small-cell lung cancer treated with stereotactic body radiotherapy. *Int J Radiat Oncol*. 2010;78:1033–1039.

12. Satoh Y, Nambu A, Onishi H, et al. Value of dual time point F-18 FDG-PET/CT imaging for the evaluation of prognosis and risk factors for recurrence in patients with stage I non-small cell lung cancer treated with stereotactic body radiation therapy. *Eur J Radiol*. 2012;81:3530–3534.
13. Atallah S, Cho BJ, Allibhai Z, et al. Impact of pretreatment tumor growth rate on outcome of early-stage lung cancer treated with stereotactic body radiation therapy. *Int J Radiat Oncol*. 2014;89:532–538.
14. Shultz DB, Trakul N, Abelson JA, et al. Imaging features associated with disease progression after stereotactic ablative radiotherapy for stage I non–small-cell lung cancer. *Clin Lung Cancer*. 2014;15:294–301. e293.
15. Yamamoto T, Kadoya N, Shirata Y, et al. Impact of tumor attachment to the pleura measured by a pretreatment CT image on outcome of stage I NSCLC treated with stereotactic body radiotherapy. *Radiat Oncol*. 2015;10:35.
16. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278:563–577.
17. Giger M, Deasy J. MO-FG-207B-00: state-of-the-art in radiomics in radiology and radiation oncology. *Med Phys*. 2016;43:3715–3715.
18. Lee G, Lee HY, Park H, et al. Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: state of the art. *Eur J Radiol*. 2017;86:297–307.
19. Sala E, Mema E, Himoto Y, et al. Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. *Clin Radiol*. 2017;72:3–10.
20. Balagurunathan Y, Gu Y, Wang H, et al. Reproducibility and prognosis of quantitative features extracted from CT images. *Transl Oncol*. 2014;7:72–87.
21. Aerts HJ. The potential of radiomic-based phenotyping in precision medicine: a review. *JAMA Oncol*. 2016;2:1636–1642.
22. Fried DV, Tucker SL, Zhou S, et al. Prognostic value and reproducibility of pretreatment ct texture features in stage III non-small cell lung cancer. *Int J Radiat Oncol*. 2014;90:834–842.
23. Gu Y, Kumar V, Hall LO, et al. Automated delineation of lung tumors from CT images using a single click ensemble segmentation approach. *Pattern Recognit*. 2013;46:692–702.
24. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005;37:360–363.
25. Harrell FE. *Regression Modeling Strategies*. New York: Springer-Verlag; 2001.
26. Storey JD. The positive false discovery rate: a Bayesian interpretation and the q-value. *Ann Stat*. 2003;31:2013–2035.
27. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361–387.
28. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53:457–481.
29. Rich JT, Neely JG, Paniello RC, Voelker CC, Nussenbaum B, Wang EW. A practical guide to understanding Kaplan-Meier curves. *Otolaryngol Head Neck Surg*. 2010;143:331–336.
30. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. *J Am Stat Assoc*. 1983;78:316–331.
31. Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinform*. 2006;7:1–8.
32. Bhatt AD, El-Ghamry MN, Dunlap NE, et al. Tumor volume change with stereotactic body radiotherapy (SBRT) for early-stage lung cancer: evaluating the potential for adaptive SBRT. *Am J Clin Oncol*. 2015;38:41–46.
33. Cohen JG, Reymond E, Lederlin M, et al. Differentiating pre-and minimally invasive from invasive adenocarcinoma using CT-features in persistent pulmonary part-solid nodules in Caucasian patients. *Eur J Radiol*. 2015;84:738–744.
34. Webb WR. The pleural tail sign. *Radiology*. 1978;127:309–313.
35. Kessler R, Gasser B, Massard G, et al. Blood vessel invasion is a major prognostic factor in resected non-small cell lung cancer. *Ann Thorac Surg*. 1996;62:1489–1493.
36. Ruffini E, Asioli S, Filosso PL, et al. Significance of the presence of microscopic vascular invasion after complete resection of stage I–II pT1-T2N0 non-small cell lung cancer and its relation with T-size categories: did the 2009 7th edition of the TNM staging system miss something? *J Thorac Oncol*. 2011;6:319–326.
37. Tsuchiya T, Akamine S, Muraoka M, et al. Stage IA non-small cell lung cancer: vessel invasion is a poor prognostic factor and a new target of adjuvant chemotherapy. *Lung Cancer*. 2007;56:341–348.
38. Huynh E, Coroller TP, Narayan V, et al. CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. *Radiother Oncol*. 2016;120:258–266.
39. Nyflot MJ, Yang F, Byrd D, Bowen SR, Sandison GA, Kinahan PE. Quantitative radiomics: impact of stochastic effects on textural feature analysis implies the need for standards. *J Med Imaging (Bellingham)*. 2015;2:041002.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

Table S1. Radiological scoring sheet of lung tumors.
Table S2. Radiomic features extracted from CT images using Definiens.
Table S3. The number of recurrences that confirmed by biopsy, PET-CT or follow-up assessment.
Table S4. Agreement of semantic features between two readers.
Table S5. Redundant radiomic features that were excluded based on Pearson's correlation analysis.
Table S6. Significant prognostic clinical, semantic and radiomics features in univariate and multivariate analysis.
Table S7. The number of patients dichotomized in either group.