



Pleomorphic carcinoma of the lung: Prognostic models of semantic, radiomics and combined features from CT and PET/CT in 85 patients

Chohee Kim^{a,1}, Hwan-ho Cho^{b,1}, Joon Young Choi^c, Teri J. Franks^d, Joung-ho Han^e,
Yeonu Choi^a, Se-Hoon Lee^f, Hyunjin Park^{g,h,*}, Kyung Soo Lee^{a,**}

^a Department of Radiology, Samsung Medical Center, Sungkyunkwan University School of Medicine (SKKU-SOM), Seoul, South Korea

^b Department of Electronic and Computer Engineering, Sungkyunkwan University, Suwon, South Korea

^c Department of Nuclear Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine (SKKU-SOM), Seoul, South Korea

^d Department of Pulmonary and Mediastinal Pathology, Department of Defense, The Joint Pathology Center, Silver Spring, MD, USA

^e Department of Pathology, Samsung Medical Center, Sungkyunkwan University School of Medicine (SKKU-SOM), Seoul, South Korea

^f Division of Hematology-Oncology, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine (SKKU-SOM), Seoul, South Korea

^g School of Electronic and Electrical Engineering, Sungkyunkwan University, Suwon, South Korea

^h Center for Neuroscience Imaging Research, Institute for Basic Science, Suwon, South Korea

ARTICLE INFO

Keywords:

Lung
Non-small cell carcinoma
Pleomorphic carcinoma
Prognosis
Radiomics

ABSTRACT

Introduction: To demonstrate semantic, radiomics, and the combined risk models related to the prognoses of pulmonary pleomorphic carcinomas (PCs).

Methods: We included 85 patients (M:F = 71:14; age, 35–88 [mean, 63 years]) whose imaging features were divided into training (n = 60) and test (n = 25) sets. Nineteen semantic and 142 radiomics features related to tumors were computed. Semantic risk score (SRS) model was built using the Cox-least absolute shrinkage and selection operator (LASSO) approach. Radiomics risk score (RRS) from CT and PET features and combined risk score (CRS) adopting both semantic and radiomics features were also constructed. Risk groups were stratified by the median of the risk scores of the training set. Survival analysis was conducted with the Kaplan-Meier plots.

Results: Of 85 PCs, adenocarcinoma was the most common epithelial component found in 63 (73 %) tumors. In SRS model, four features were stratified into high- and low-risk groups (HR, 4.119; concordance index ([C-index], 0.664) in the test set. In RRS model, five features helped improve the stratification (HR, 3.716; C-index, 0.591) and in CRS model, three features helped perform the best stratification (HR, 4.795; C-index, 0.617). The two significant features of CRS models were the SUVmax and the histogram feature of energy ([CT Firstorder Energy]).

Conclusion: In PCs of the lungs, the combined model leveraging semantic and radiomics features provides a better prognosis compared to using semantic and radiomics features separately. The high SUVmax of solid portion (CT Firstorder Energy) of tumors is associated with poor prognosis in lung PCs.

Abbreviations: C-index, Concordance index; CRS, Combined risk score; DL, Deep learning; GCLM, Gray-level co-occurrence matrix; HR, Hazard ration; ICC, Intra-class correlation; ISZM, Intensity size zone matrix; KRAS, Kirsten rat sarcoma viral oncogene homolog; LASSO, Least absolute shrinkage and selection operator; LDA, Low density area; MRI, Magnetic resonance imaging; MTV, Metabolic tumor volume; PC, Pleomorphic carcinoma; PET/CT, Positron emission tomography/Computed tomography; ROI, Region of interest; RRS, Radiomics risk score; SRS, Semantic risk score; SUVavg, Average standardized uptake value; SUVmax, Maximum standardized uptake value; TLG, Total lesion glycolysis; VOI, Volume of interest.

* Corresponding author at: School of Electronic and Electrical Engineering, Sungkyunkwan University, Suwon, 16419, South Korea.

** Corresponding author at: Department of Radiology, Samsung ChangWon Hospital, Sungkyunkwan University School of Medicine (SKKU-SOM), ChangWon, Gyeongsangnam-Do, 51353 South Korea.

E-mail addresses: hyunjinp@skku.edu (H. Park), kyungs.lee@samsung.com (K.S. Lee).

¹ These two authors contributed equally to this study.

<https://doi.org/10.1016/j.ejro.2021.100351>

Received 4 April 2021; Received in revised form 3 May 2021; Accepted 8 May 2021

2352-0477/© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Pleomorphic carcinoma (PC) of the lung is one of a group of neoplasms sometimes referred to as ‘Sarcomatoid Carcinomas of the Lung [1].’ PC is defined as a poorly differentiated non-small cell carcinoma, namely squamous cell carcinoma, adenocarcinoma or undifferentiated non-small cell carcinoma that contains at least 10 % spindle and/or giant cells, or a carcinoma consisting only of spindle and giant cells [2]. The PC has been known as a rare lung cancer. Even though the tumor has been known to account for 0.1–0.4% of all lung malignant tumors [1], its real incidence is not known. The tumors mainly occur in heavy-smoking men of 60 years in age and are reported to show aggressive clinical behavior as compared with non-small cell lung cancer [2–5].

Computed tomography (CT) findings representing PCs of the lungs include a large peripheral lung tumor, the presence of a central low attenuation area within the tumor or peritumoral ground-glass opacity, and frequent pleural and chest wall invasion [6,7]. The presence of a large area of low-attenuation within the tumor on enhanced CT scans was associated with a poorer prognostic factor [8]. However, these clinicopathologic and CT data regarding PCs are based on a small number of cases; almost all reports were based on tumors of 70 cases or less in number.

At 18 Fluorine-fluorodeoxyglucose (FDG) positron emission tomography (PET), PCs show a tendency to present with intense FDG accumulation, and the measured maximum standardized uptake value (SUVmax) is high in proportion to the extent of programmed cell death ligand 1 (PD-L1) expression and is high with the presence of Kirsten rat sarcoma viral oncogene homolog (KRAS) expression. The total lesion glycolysis (TLG) of the primary tumor and KRAS mutation were independent prognostic factors of PCs [9]. In an additional PET/CT study [10], tumor SUVmax was not different in terms of tumor dimensions, histology, pathological stage, and pattern of recurrence. Pathologic stage, surgical completeness, and the presence of vascular or lymphatic invasion were related factors for patient survival.

Radiomics is a high-dimensional analysis approach where large amounts of mineable quantitative imaging features are analyzed from medical images in a high throughput fashion [11]. Because tumor heterogeneity is embedded in imaging data, high-dimensional features extracted from CT and PET can be effectively handled with quantification and analyzed in radiomics to identify important diagnostic and prognostic features. For example, according to a report [12], a tumor could be divided into necrotic and viable portions by incorporating diffusion-weighted magnetic resonance imaging (MRI) data into PET data and the division showed good correlation with histopathology.

In this study, we hypothesized that the radiomics analyses obtained from a relatively large number of lung PCs could lead to efficient prognostic prediction. Thus, the purpose of this report was to demonstrate effective radiomics models from CT and PET/CT that are keenly related to the prognoses of lung PCs in a relatively large number of patients.

2. Materials and methods

The institutional review board at Samsung Medical Center (Seoul, Korea) approved this retrospective study (IRB number; 2020-10-098-001). The requirement to obtain informed consent regarding the acquisition of CT and PET/CT data was waived.

2.1. Inclusion criteria

Between 1994 and 2018, 278 patients with surgically resected PC of the lung were identified from the files of the department of pathology at our institution. Among them, 85 cases fulfilled the following inclusion criteria: (1) availability of contrast-enhanced chest CT and PET/CT scan performed prior to surgery, (2) availability of clinical follow-up data

including patients’ survival, and (3) the absence of neoadjuvant therapy history. Because our study was performed in the span of several decades, there were many scanners involved. We limited CT and PET/CT machines to those from General Electric to reduce the effects of different scanner types. Of the 85 patients, images from 60 patients belonged to the training set and those from the remaining 25 patients were attributed to the test set (Fig. 1).

2.2. CT imaging

All presurgical contrast-enhanced chest CT examinations were performed using one of multidetector row CT scanners; Genesis Hispeed RP, Light Speed QX/i, Light Speed VCT, and Discovery CT750 HD (GE Healthcare, Chalfont St Giles, England). The following parameters were used as CT scanning: detector collimation, 3–8 mm; beam pitch, 0.75–1.5; reconstruction thickness, 3–8 mm; reconstruction interval, 2–8 mm; tube voltage, 120 kVp; tube current, 114–275 mAs; and the reconstruction kernel, standard soft-tissue reconstruction algorithm. A total of 100–120 mL of non-ionic low osmotic iodine contrast medium was administered intravenously at a rate of 2–3 mL/sec using a power injector (MCT Plus; Medrad, Pittsburgh, PA) in all patients.

One thoracic radiologist (C.K., with six-year experience in thoracic CT interpretation) evaluated both the primary tumor lesion and associated findings on enhanced CT scans. These features were referred to as semantic features from CT to distinguish them from radiomics features. Each primary tumor was analyzed after reviewing both transverse- and coronal-reformatted CT images. For size measurement, the equator images on transverse- and coronal-reformatted images were selected and the maximum diameter of the tumor was chosen as the tumor diameter. Tumor attenuation values were measured on the maximum diameter images by using a region of interest (ROI). The attenuation value of the ROI lower than 25 HU was defined as low density area (LDA) [7]. Because the LDA of the tumor was frequently encountered, the attenuation values were measured both in necrotic and nonnecrotic regions. The largest diameter of the necrosis was measured. Presence or absence of the cavity of the tumor, which was defined as a gas-filled space within the tumor was recorded [13]. Other image analyses included the location, contour (round, lobulated, and spiculated), and tumor margin (well versus poorly defined). Location was classified into being central and peripheral. The central tumors were defined on CT as those that originated from or were located in the segmental bronchus or proximal to the segmental bronchus, while peripheral tumors were defined as those distal to the segmental bronchus with or without peribronchial extension.

2.3. PET/CT acquisition and analyses

Before PET examination, all patients fasted for at least 6 h. Blood glucose levels were measured before the injection of FDG and were required to be < 200 mg/dL. Whole-body PET and unenhanced CT images were acquired using two types of PET/CT scanners (Discovery LS, GE Healthcare, Milwaukee, WI, USA; Discovery STe, GE Healthcare, Milwaukee, WI, USA), 60 min after the injection of FDG (5.0 MBq/Kg). After the CT scan, an emission scan was obtained from head to middle thigh for 4 min (for LS) or 2.5 min (for STe) per frame in 2-dimensional (for LS) or 3-dimensional (for STe) mode. Attenuation-corrected PET images were reconstructed from the CT data using an ordered-subset expectation maximization (OSEM) algorithm (28 subsets, 2 iterations for LS; 20 subsets, 2 iterations for STe).

One nuclear medicine physician (J.Y.C. with 17-years of experience in PET/CT interpretation), who was blinded to the clinical results, visually and quantitatively analyzed the PET images and recorded their findings. Semiquantitative and volumetric analyses were conducted using volume viewer software on a GE Advantage Workstation 4.4, which provided a convenient and automatic method to delineate the volume of interest (VOI), using an isocontour threshold method based on

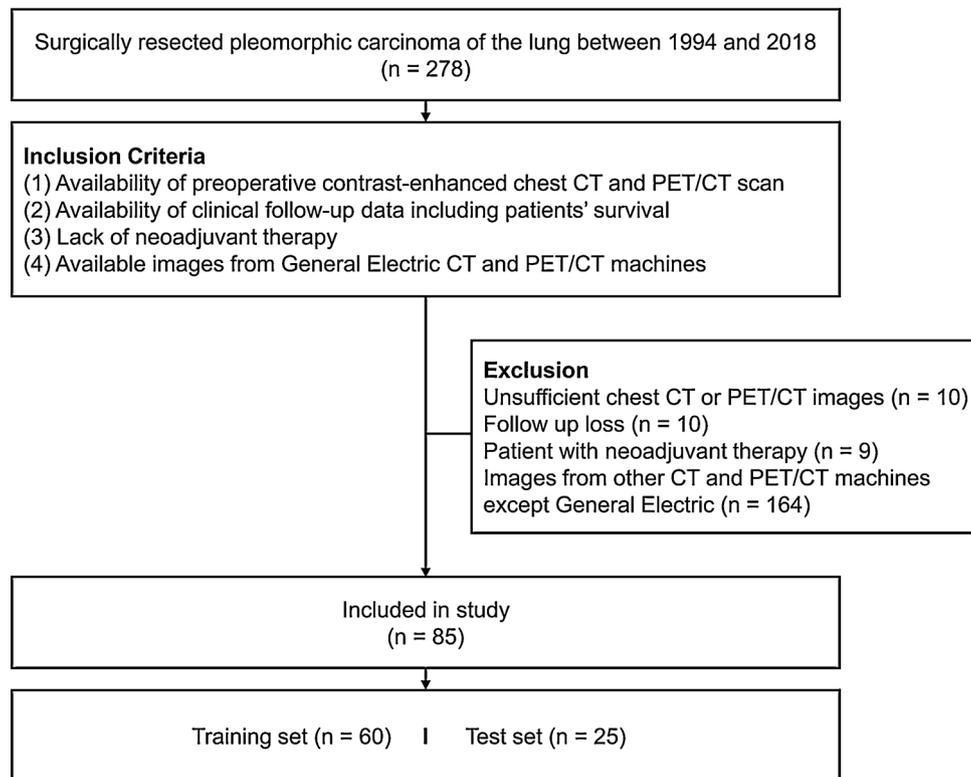


Fig. 1. Flow diagram of patient inclusion.

the SUV. The nuclear medicine physician placed an automatically delineated VOI over the primary PC lesion, after which the software allowed for the calculation of metabolic tumor volume (MTV), maximum SUV (SUVmax), and average SUV (SUVavg) within the entire primary cancer. MTV was defined as the total tumor volume segmented via threshold SUV 3.6 [14,15]. In addition, we calculated a composite parameter, the total lesion glycolysis (TLG), which was determined by multiplying MTV by SUVavg [16]. These features were included as semantic features from PET.

2.4. Semantic features analysis

The following 15 features from CT and histology were analyzed: 1) epithelial cell types, 2) epithelial cell component (%; volume percentage), 3) spindle/giant cell component (%), 4) tumor necrosis component (%), 5) epithelial cell component without necrosis (%), 6) spindle/giant cell component without necrosis (%), 7) location (central vs. peripheral), 8) contour (round, lobulated, and spiculated), 9) margin (well vs. poorly defined), 10) the presence of cavity, 11) calcification, 12) ground-glass opacity (GGO), 13) tumor invasion to the pleura or the chest wall, 14) tumor stage, and 15) presence of adjuvant therapy. The following four features from PET were analyzed: 1) metabolic tumor volume, 2) maximum SUV, 3) average SUV, and 4) total lesion glycolysis (TLG).

2.5. Radiomics analysis

Computer-based in-house software was used for lesion segmentation. For each tumor on CT, ROIs were delineated on axial images to generate a VOI that eventually contained the entire tumor with a semiautomatic approach. A manual adjustment was performed to exclude bronchovascular bundles and inflammatory lesions surrounding the primary tumor by two thoracic radiologists (C.K. and Y.C., both with five-year experience in thoracic CT interpretation). Each of two radiologists drew VOIs on primary tumors, and the drawn first set (by C.K.) was used

for radiomics analysis. Both sets (drawn by C.K. and Y.C.) were used to assess the reproducibility (intra-class correlation, ICC) of VOIs and radiomics features. A total of 72 radiomics features were extracted using open-source radiomics software, PyRadiomics, for each imaging modality [17]. CT radiomics features were extracted using the manually specified VOIs. PET radiomics features were extracted from VOIs specified by applying the threshold value of SUV, 3.6. Several PET radiomics features (i.e., MTV, SUVmax, and SUVavg) were already included in the semantic feature analysis and thus those were not considered. The features were grouped into shape (14 features), histogram-based (18 features), intensity size zone matrix (ISZM, 16 features), and gray-level co-occurrence matrix (GLCM, 24 features) features. The latter two, namely ISZM and GLCM, were texture features. The histogram-based features were computed from 256-bin histograms for CT and 32-bin histograms for PET calculated over the intra-tumoral intensity range. The GLCM features assess textural information and reflect intra-tumoral heterogeneity using a 2D histogram with 256 bins (CT) and 32 bins (PET). A total of 26 matrices corresponding to 26 3D directions with an offset of one were computed and then averaged to yield a single matrix. The averaged matrix was used to compute the GLCM features. The ISZM features were also related to texture using blobs of similar intensity and differing sizes. We constructed 32×256 matrix (CT) and 8×256 matrix (PET) in which the first dimension was binned intensity and the second dimension was the size of the blobs. We also considered one CT semantic feature, the ratio of the LDA defined as a ratio of LDA over the total VOI volume, known to be related to PC [8]. In total, we computed 73 features from CT, and 69 features from PET.

2.6. Pathological analysis

On gross pathologic specimens, the presence and the extent (nearest to the 10 %) of tumor necrosis within the primary tumors were determined by a lung pathologist (J.H. with 26-year experience in lung pathology). By histologic examination, the same pathologist decided the components of epithelial (adenocarcinoma, squamous cell carcinoma,

adenosquamous carcinoma, etc) and spindle cell/giant cell carcinomas, regardless of the presence and extent of tumor necrosis. The cancer components were recorded to the nearest 10 %. The final cancer compositions and their extent were determined in consideration of tumor necrosis extent; thus the epithelial and spindle/giant cell components (percentages), respectively, were calculated by $(100 - \text{necrosis percentage}) \times \text{percentage of epithelial and spindle/giant cell carcinomas} / 100$. The percentages of epithelial and spindle/giant cell carcinomas were rounded off and were calculated to the nearest to the 5%.

2.7. Model building and survival analysis

Having reproducible features is an important requirement in radiomics models and thus we filtered radiomics features with low reproducibility. Features whose ICC was lower than 0.75 were filtered out [18]. All radiomics features were z-score normalized using the mean and standard deviation of the training set. We adopted the Cox-Least Absolute Shrinkage and Selection Operator (LASSO) to select important features and build the radiomics risk score (RRS) from the remaining features (ICCs, 0.75 or greater) of both CT and PET. The optimal regularization penalty term of Cox-LASSO was determined using cross-validation within the training set. The RRS of each patient was defined as the relative risk of the multi-variate Cox regression model and we applied Kaplan-Meier analysis using the RRS. Patients were divided into low- and high-risk groups based on the median of the RRS of the training set; the same stratification threshold was also applied to the test set. Hazard ratio (HR) and p-value of the log-rank test were used to measure the survival difference in low- and high-risk groups. We applied the same procedure of Cox-LASSO to build the semantic risk score (SRS) using the semantic features from CT and PET and clinical variables from pathological analysis. The 19 clinical variables considered were described previously under the subheading of *Semantic Features Analysis* in this Methods section. A third model was built from the combined set of semantic and radiomics features using the same Cox-LASSO procedure and referred to as the combined risk score (CRS). The same analysis methods adopted in RRS calculation including hazard ratio and log-rank test were applied to SRS and CRS calculation.

2.8. Statistics

ICC was computed to assess the reproducibility of features using two sets of VOIs for CT radiomics features. ICC was not computed for PET radiomics as the VOIs were defined automatically with no variability. Cohen's kappa was used to assess the reliability of the VOIs in CT. For RRS, we used stable radiomics features only with ICC over 0.75 [19].

To compare demographic information, categorical variables were analyzed using a chi-square or Fisher's exact test, while continuous variables were analyzed using Student's *t*-test. To explore the association between radiomics features and survival, Kaplan-Meier analysis was conducted. Kaplan-Meier curves between subgroups were generated using the 'statistics and machine learning' toolbox in Matlab (Natic, MA: The MathWorks Inc.). HR, concordance index (C-index), and p-value of log-rank test were used to measure the difference in low- and high-risk groups.

3. Results

3.1. Demographics and clinicopathologic features

Of the 85 patients, 71 were males and 14 were females (male to female ratio: 5.1:1); patients ages ranged from 35 to 88 years (mean, 63 years). The pathologic stage according to the tumor-node-metastasis classification at the time of surgery was stage I in 38 cases (45 %) (IA: 10, IB: 28), stage II in 21 cases (25 %) (IIA: 5, IIB: 16), stage III in 20 cases (24 %) (IIIA: 18, IIIB: 2), and stage IV in five cases (6%).

Information regarding the post-surgical treatments was obtained for

41 patients. Two patients received radiation therapy only, 28 patients underwent chemotherapy only, and 11 patients underwent concurrent chemoradiation therapy. The median follow-up period was 1136 days (range, 2–3769 days). The deaths of 37 patients were confirmed. The three-year overall survival and five-year overall survival in the patients were 47 % and 27 %, respectively.

The maximum diameters of the tumors ranged from 0.4–17 cm (median, 4.3 cm; mean, 4.8 cm). Microscopic examination of the 85 tumors revealed the following epithelial components; adenocarcinoma was found in 63 (73 %) of the cases, squamous cell carcinoma in 15 (18 %), large cell carcinoma in 3 (4%), adenosquamous carcinoma in one (1%), and sclerosing mucoepidermoid carcinoma in one (1%). Tumor necrosis was observed in 58 cases (68 %), with a range of 10–95% (median, 10 %; mean, 26 %).

The marginal characteristics of the tumors on CT were well defined in 84 patients and poorly defined in one. The contour of tumors was rounded in four (5%), lobulated in 52 (61 %), and spiculated in 29 (34 %). Twenty tumors (24 %) showed frequent cavity representing definite central necrosis. Surrounding areas of ground-glass attenuation were noted in 29 tumors (34 %) (Table 1).

3.2. Semantic model for prognosis

As shown in Table 2, our SRS model showed that tumors with the following properties were associated with lower risk: 1) centrally located tumors, and 2) tumors being managed with adjuvant therapy. Tumors with the following properties were associated with higher risk: 1) presence of cavity, and 2) high SUVmax. Fig. 2 shows the Kaplan-Meier plot using the SRS model for both training and test sets. The risk groups of both sets were stratified by applying the median SRS of the training set. The risk group stratified by the SRS showed a significant difference in the training set but not in the test set. We observed a HR of 2.651 (95 % Confidence Interval [CI]: 1.238–5.678), a C-index of 0.709 and a p-value of 0.021 for the training set, while a HR of 4.119 (95 % CI: 1.089–15.577), a C-index of 0.664 and a p-value of 0.081 were observed for the test set. For the training set, the 3-year survival was 0.682 and 0.475 in the low and high risk groups respectively, and 5-year survival was 0.640 and 0.185 in the low and high risk groups respectively. For the test set, the 3-year survival was 0.831 and 0.563 in the low- and high-risk groups, respectively and the 5-year survival was 0.831 and 0.450 in the low- and high-risk groups, respectively.

3.3. Radiomics model for prognosis

The CT VOIs used in radiomics were reproducible with Cohen Kappa of 0.867. The full details regarding ICC values were reported in the Supplemental Data 1. Table 2 shows the features involved in the RRS for prognosis. The RRS model was built using three histogram-based features and two texture features; energy from CT, root mean squared of CT, total energy from PET, cluster shade of GLCM from PET, and small area low gray level emphasis of GLSZM from PET.

Fig. 2 shows the Kaplan-Meier plot using the RRS model. The risk groups of both sets were stratified by applying the median RRS of the training set. The risk group stratified by the RRS showed a significant difference in the training set but not in the test set. We observed a HR of 2.402 (95 % CI: 1.114–5.086), a C-index of 0.704 and a p-value less than 0.036 for the training set, while a HR of 3.716 (95 % CI: 0.851–16.222), a C-index of 0.591 and a p-value of 0.171 were observed for the test set. For the training set, the 3-year survival was 0.743 and 0.420 in the low- and high-risk groups, respectively, and the 5-year survival was 0.547 and 0.378 in the low- and high-risk groups, respectively. For the test set, the 3-year survival was 0.733 and 0.617 in the low- and high-risk groups, respectively and the 5-year survival were 0.733 and 0.309 in the low- and high-risk groups, respectively.

Table 1
Demographic Information.

Variables	Training set	Test set	p Values	Test Applied
n	60	25		
Age, Mean (STD)	63.4167 (10.4837)	65.2000 (8.8882)	0.458	T-test
Sex			0.105	Chi-squared test
Male	53	18		
Female	7	7		
Stage			0.7686	Fisher's exact test
NA	1	0		
1	25	13		
2	15	6		
3	16	4		
4	3	2		
Epithelial cell			0.532	Fisher's exact test
Squamous cell carcinoma	12	3		
Adenocarcinoma	44	19		
Large cell carcinoma	2	1		
Adenosquamous carcinoma	2	1		
Sclerosing mucoepidermoid carcinoma	0	1		
Epithelial cell component (%), Mean (STD)	41.3 (27.2)	43.2 (28.2)	0.767	T-test
Spindle/giant cell component (%), Mean (STD)	36.1 (25.3)	32.4 (23.5)	0.535	T-test
Necrosis component (%), Mean (STD)	22.7 (26.0)	24.4 (30.9)	0.792	T-test
Location			0.552	Fisher's exact test
Central	3	0		
Peripheral	57	25		
Contour			0.192	Fisher's exact test
Round	2	2		
Lobulated	40	12		
Spiculated	18	11		
Margin			0.294	
Ill-defined	0	1		
Well-defined	60	24		
Cavity			1.000	Chi-squared test
No	46	19		
Yes	14	6		
Calcification			1.000	Fisher's exact test
No	57	24		
Yes	3	1		
Peripheral GGO			0.790	Chi-squared test
No	39	17		
Yes	21	8		
Invasion			0.328	Fisher's exact test
No	49	23		
Yes	11	2		
Adjuvant Therapy			0.959	Fisher's exact test
None	30	14		
CCRT	8	3		
CTX	20	8		
RTx	2	0		

Note. - STD = standard deviation, GGO = ground-glass opacity, CCRT = concurrent chemoradiation therapy, CTX = chemotherapy, RTx = radiation therapy.

Table 2
Features Appeared to Be Significant in Various Risk Score Models.

Models	Significant features	Cox-Lasso coefficients
Semantic risk score model	Location center	-0.334
	Cavity	0.140
	Adjuvant Therapy	-0.492
	SUVmax	0.226
Radiomics risk score model	CT Firstorder Energy	0.090
	CT Firstorder RootMeanSquared	0.039
	SUV Firstorder Total Energy	0.014
	SUV GLCM Cluster Shade	0.046
	SUV GLSZM Small Area Low Gray Level Emphasis	-0.002
Combined risk score model	CT Firstorder Energy	0.033
	Adjuvant Therapy	-0.084
	SUVmax	0.013

Note - The name of the radiomics features follows the format of modality, category, and detailed name. For example, CT Firstorder Energy refers to the energy feature belonging to first-order histogram category computed from CT.

3.4. Combined model for prognosis

Table 2 shows the features involved in the CRS for prognosis. The selected features were one radiomics feature and two semantic features. The radiomics feature was energy from CT. The semantic features were SUVmax and getting adjuvant therapy (Figs. 3 and 4). All features were already identified in the previous models. Fig. 2 shows the Kaplan-Meier plot using the CRS model. The risk groups of both sets were stratified by applying the median CRS of the training set. The risk group stratified by the CRS showed a significant difference in both training and test sets. We observed a HR of 2.438 (95 % CI: 1.146–5.189), a C-index of 0.677 and a p-value of 0.034 for the training set, while a hazard ratio of 4.795 (95 % CI: 1.282–17.937), a C-index of 0.617 and a p-value under 0.046 were observed for the test set. For the training set, the 3-year survival was 0.677 and 0.482 in the low- and high-risk groups respectively, and the 5-year survival was 0.625 and 0.193 in the low- and high-risk groups, respectively. For the test set, the 3-year survival was 0.909 and 0.556 in the low- and high-risk groups, respectively, and the 5-year survival was 0.9097 and 0.463 in the low- and high-risk groups, respectively. Performance comparisons for various risk score models were summarized in Table 3.

4. Discussion

Our results showed that radiomics analysis combined with the semantic features annotated by experts was effective for PCs in the lung. The identified features of the CRS could provide new insights to assess risk of the PCs.

Interestingly, in our semantic model, PCs containing epithelial cell components of large cell carcinoma or adenosquamous carcinoma were associated with the low-risk group. In previous studies [8], the presence of PD-L1 (high SUVmax with high PD-L1 expression) and KRAS mutation (high SUVmax with the presence of KRAS mutation) expression.

The RRS model contained five important radiomics features. Two features were from CT and three were from PET, thus demonstrating that multimodal approaches were necessary to assess prognosis. Energy from CT measures intra-tumoral intensity which may reflect solid components [20]; root mean squared of CT may be related to indolence [21]; total energy from PET is a multiplication of MTV with the sum of the square of intra-tumoral SUV values, thus it can be related to MTV and extent of metabolism; cluster shade of GLCM measures the skewness of

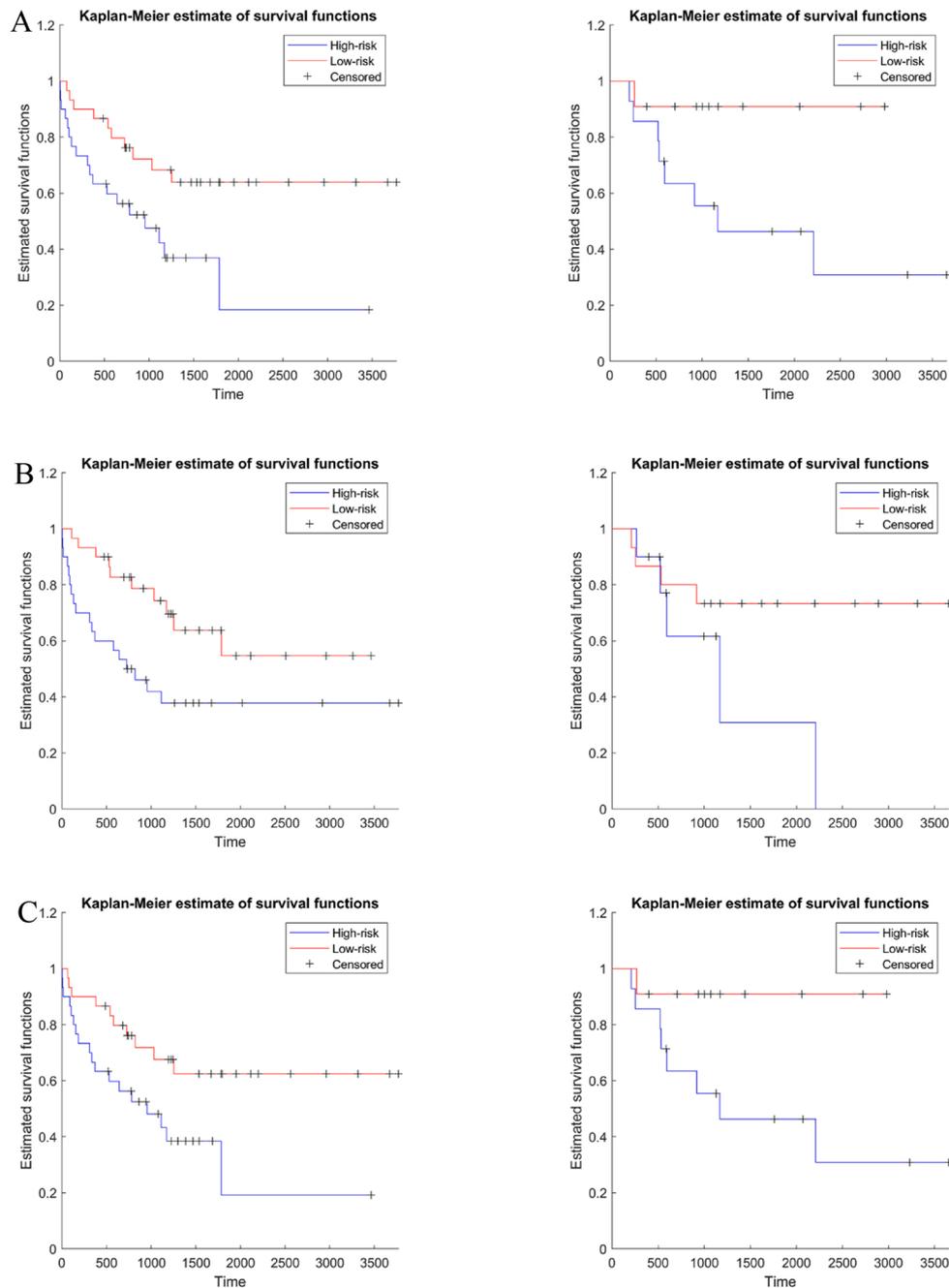


Fig. 2. Kaplan-Meier plots for (a) semantic risk score, (b) radiomics risk score, and (c) combined risk score. The left plots are the survival plots for the training set and the right plots are the survival for the test set.

the GLCM [17]; and small area low gray-level emphasis of GLSZM has ambiguous physical meaning but GLSZM textures have often been associated with prognosis of lung cancer in previous studies. The most significant feature was the histogram-based feature of energy from CT, reflecting the magnitude of pixels from the ROI, and the feature suggested that solid portion of the tumor could be associated with a poor prognosis. The cluster shade of GLCM derived from PET is related to intensity heterogeneity. Malignant transformation of tumors exhibits intratumoral biological heterogeneity associated with cellular and molecular characteristics such as cellular proliferation, necrosis, fibrosis, differences in blood flow and angiogenesis, cellular metabolism, hypoxia and expression of specific receptors [22]. Similarly, previous PET studies also reported that intratumoral metabolic heterogeneity within tumors is associated with more aggressive behavior, poorer response to treatment and worse prognosis [23–25].

The CRS model contained one radiomics feature and two semantics features. The four significant features in the SRS model were reduced to two important features which were the SUVmax and history of receiving adjuvant therapy. The five significant features in the RRS were reduced to only one important feature. The texture features from PET were removed, and the histogram feature of CT energy reflecting solid portion of the tumor remained to be important in patients’ prognostication. These changes in the significance of CRS model could be summarized as follows: the SUVmax of the solid component of the PCs is crucial in determining patients’ prognosis. The importance of CT energy in histogram feature (solid tumor portion) for cancer prognostication could be seen in thoracic MRI study for lung cancers. The low apparent diffusion coefficient (ADC) value of a lung cancer is associated with high pathological tumor grade and advanced metastatic nodal stage [26,27]. The low ADC value of the tumor in MRI is correlated histologically with

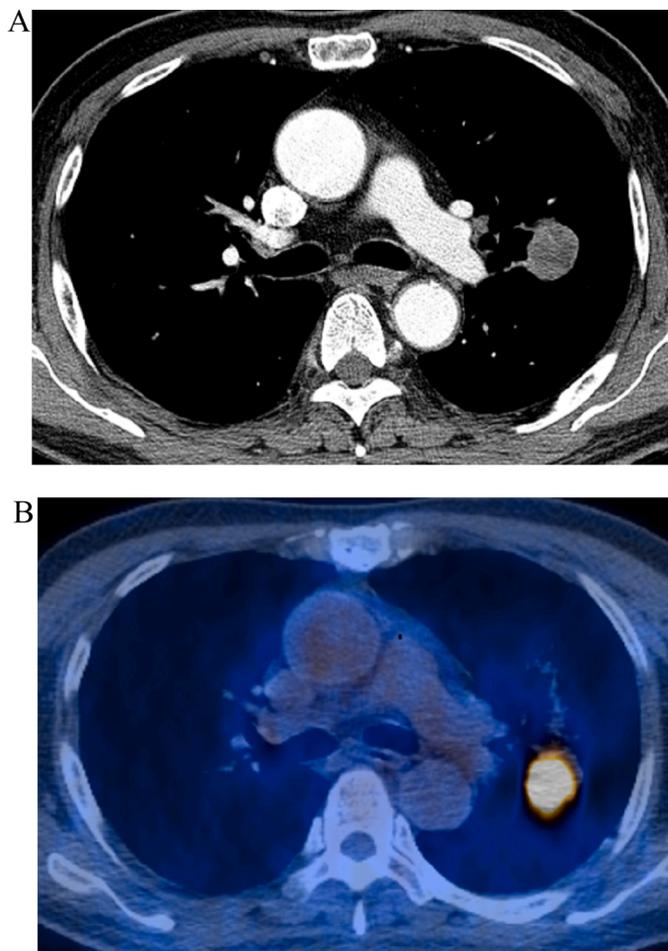


Fig. 3. An example of combined risk score features predicting good survival. (a) CT and (b) PET/CT fusion images show an approximately 30-mm-sized tumor in left upper lobe in a 69-year-old man. Patient’s TNM stage was T1cN0M0 (stage IA). This tumor was pathologically confirmed as pleomorphic carcinoma composed of spindle (80 %) cells and adenocarcinoma (10 %) cells. Prognosis was expected to be good with low risk (CT energy; 61158963, SUVmax; 9.9). The patient remained alive five years after surgical management.

dense cellularity and large cellular size in MRI [26].

It was reported that the presence of a large necrotic area or a cavity within the tumor on enhanced CT scans is associated with poor prognosis in lung PCs [8]. In our study, the presence of a cavity remained as an important SRS model; however, the semantic feature was removed from the list of CRS model. We cannot explain exquisitely the difference in results between our study and the previous one [10]. But the multivariate nature of our radiomics or combined model, not univariate nature, might help explain the difference in the results.

One may argue that the radiomics features only slightly contributed to the semantic score’s performance, given the fact that semantic scores, or in our study case, the SUVmax and adjuvant therapy, appeared to perform very well in estimating survival analysis. Thus, we repeated the same methodology to two subsets, namely, those who received adjuvant therapy (n = 41) and those who did not (n = 45). For each subset, we trained the radiomics model only using 30 samples. In the test set, the HR for the combined risk model was 2.08 for the non-adjuvant group and the HR was 3.95 for the adjuvant group (see SUPPLEMENTAL DATA 2). These HR values were smaller than the HR of 4.80 obtained using the full data. However, the results from the subsets were statistically underpowered and hence unstable. Because the radiomics models are high dimensional, the results of the subsets should be interpreted with care.

There are limitations to our study. Our study is a single-center study

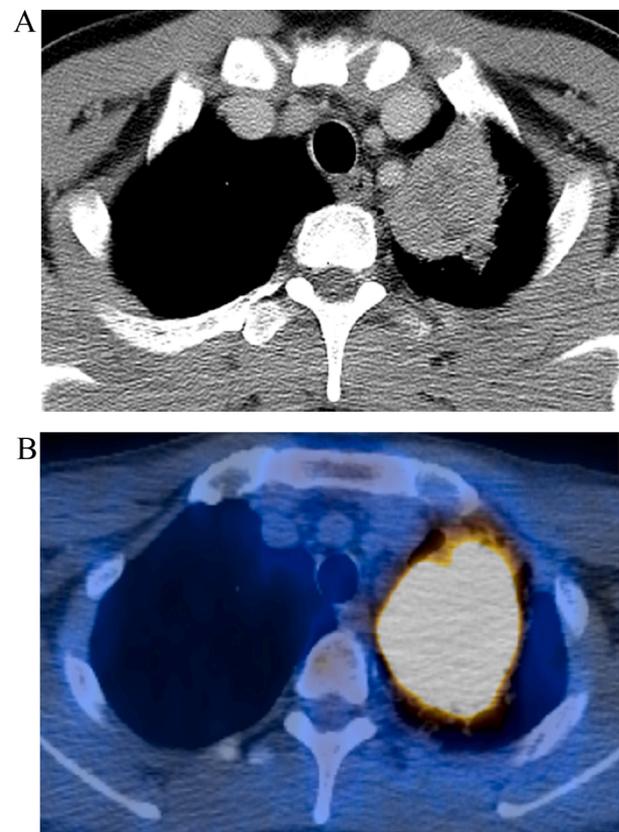


Fig. 4. Example of combined risk score features predicting poor survival. (a) CT and (b) PET/CT fusion images demonstrate an approximately 50-mm-sized mass in left lower lobe in a 48-year-old man. Patient’s TNM stage was T2bN0M0 (stage IIA). This mass was pathologically confirmed as pleomorphic carcinoma composed of spindle (10 %) cells and large cell carcinoma (10 %) cells. Prognosis was expected to be poor with high risk (CT energy; 976767333, SUVmax; 33.5). This patient died 101 days after surgery.

Table 3
Performance Comparisons for Various Risk Score Models.

		Hazard ratio	95 % confidence interval	Concordance index	p Value
Training	Semantic risk score	2.651	1.238 – 5.678	0.709	0.021
	Radiomics risk score	2.402	1.114 – 5.086	0.704	0.036
	Combined risk score	2.438	1.146 – 5.189	0.677	0.034
Test	Semantic risk score	4.119	1.089 – 15.577	0.664	0.081
	Radiomics risk score	3.716	0.851 – 16.222	0.591	0.171
	Combined risk score	4.795	1.282 – 17.937	0.617	0.046

lacking independent validation. Because PCs are a rare subgroup of lung cancer, the inclusion period of our study spanned for several decades in order to collect enough samples for the analysis. This led to having variability in scanners involved. It has been well known that radiomics features are affected by scanner types and acquisition parameters including reconstruction kernel. Therefore, we limited our analysis only to preoperative CT and PET/CT studies obtained with GE scanners and this may have led to inevitably including rather small number of cases and selection bias. With this keeping in mind, we split our data into training and tests set to build an analytic model. Future studies using

independent validation are needed to confirm our results. Also, there are harmonization methods that can reduce such deleterious effects arising from different scanners and reconstruction methods that could be explored in future studies [28]. Deep learning (DL) is a disruptive technology capable of extracting fine details from the imaging data and thus could be effective in analyzing PCs. However, DL approaches require more samples than the radiomics approach and thus could not be used in our study. Exploring DL for analyzing PCs of the lung could be a promising avenue for future research.

In conclusion, the combined model leveraging both semantic and radiomics features in PCs of the lungs provides a better prognosis compared to using semantic and radiomics features separately. Those features that are important in prognostication are CT Firstorder Energy and SUVmax. In other words, the high SUVmax of the solid component of the PCs is crucial in determining cancer prognosis. In addition, managing patient with adjuvant therapy does work and help lengthen patient survival.

Disclaimer

The views expressed in this article are those of the authors and do not reflect the official policy of the Department of Army, Navy, or Air Force; Department of Defense; or U.S. government (TJF).

This research was supported by the National Research Foundation (NRF-2020M3E5D2A01084892), Institute for Basic Science (IBS-R015-D1), Ministry of Science and ICT (IITP-2021-2018-0-01798), IITP grant funded by the AI Graduate School Support Program (2019-0-00421), and ICT Creative Consilience program (IITP-2020-0-01821).

Ethical statement

Nothing to declare.

Funding statement

Nothing to declare.

Transparency document

The [Transparency document](#) associated with this article can be found in the online version.

CRedit authorship contribution statement

Chohee Kim: Data curation, Formal analysis, Investigation, Methodology, Project administration, Visualization, Writing - original draft, Writing - review & editing. **Hwan-ho Cho:** Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing - original draft. **Joon Young Choi:** Data curation, Investigation, Visualization, Writing - original draft. **Teri J. Franks:** Data curation, Investigation, Visualization, Writing - original draft. **Joung-ho Han:** Data curation, Investigation, Visualization, Writing - original draft. **Yeonu Choi:** Data curation, Investigation, Visualization, Writing - original draft. **Se-Hoon Lee:** Data curation, Investigation, Visualization, Writing - original draft. **Hyunjin Park:** Conceptualization, Data curation, Investigation, Methodology, Software, Supervision, Visualization, Writing - original draft, Writing - review & editing. **Kyung Soo Lee:** Conceptualization, Data curation, Investigation, Methodology, Project administration, Supervision, Visualization, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgments

We are grateful for the librarian Jaero Park for his dedicated support of manuscript formatting. The librarian is working at the Samsung Medical Information & Media Services of Samsung Medical Center located in Seoul, South Korea.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.ejro.2021.100351>.

References

- [1] Y.L. Chang, et al., Pulmonary pleomorphic (spindle) cell carcinoma: peculiar clinicopathologic manifestations different from ordinary non-small cell carcinoma, *Lung Cancer* 34 (1) (2001) 91–97.
- [2] T. Mochizuki, et al., Pleomorphic carcinoma of the lung: clinicopathologic characteristics of 70 cases, *Am. J. Surg. Pathol.* 32 (11) (2008) 1727–1735.
- [3] T. Yuki, et al., Pleomorphic carcinoma of the lung: a surgical outcome, *J. Thorac. Cardiovasc. Surg.* 134 (2) (2007) 399–404.
- [4] F. Chen, et al., Clinicopathological characteristics of surgically resected pulmonary pleomorphic carcinoma, *Eur. J. Cardiothorac. Surg.* 41 (5) (2012) 1037–1042.
- [5] K. Ito, et al., Clinical characteristics of pleomorphic carcinoma of the lung, *Lung Cancer* 68 (2) (2010) 204–210.
- [6] T.H. Kim, et al., Pleomorphic carcinoma of lung: comparison of CT features and pathologic findings, *Radiology* 232 (2) (2004) 554–559.
- [7] T.S. Kim, et al., CT findings of surgically resected pleomorphic carcinoma of the lung in 30 patients, *AJR Am. J. Roentgenol.* 185 (1) (2005) 120–125.
- [8] A. Nishida, et al., Clinicoradiological outcomes of 33 cases of surgically resected pulmonary pleomorphic carcinoma: correlation with prognostic indicators, *Eur. Radiol.* 26 (1) (2016) 25–31.
- [9] X. Wu, et al., (18)F-FDG PET/CT imaging in pulmonary sarcomatoid carcinoma and correlation with clinical and genetic findings, *Ann. Nucl. Med.* 33 (9) (2019) 647–656.
- [10] C. Rapietta, et al., Primary sarcomatoid carcinoma of the lung: radiometabolic ((18)F-FDG PET/CT) findings and correlation with clinico-pathological and survival results, *Lung* 194 (4) (2016) 653–657.
- [11] G. Lee, et al., Radiomics and imaging genomics in precision medicine, *Precis. Future Med.* 1 (2017) 10–31.
- [12] M.R. Divine, et al., A population-based gaussian mixture model incorporating 18F-FDG PET and diffusion-weighted MRI quantifies tumor tissue classes, *J. Nucl. Med.* 57 (3) (2016) 473–479.
- [13] D.M. Hansell, et al., Fleischner Society: glossary of terms for thoracic imaging, *Radiology* 246 (3) (2008) 697–722.
- [14] N.C. Nguyen, et al., Is there a common SUV threshold in oncological FDG PET/CT, at least for some common indications? A retrospective study, *Acta Oncol.* 50 (5) (2011) 670–677.
- [15] K. Kim, et al., Prognostic value of volumetric parameters measured by F-18 FDG PET/CT in surgically resected non-small-cell lung cancer, *Nucl. Med. Commun.* 33 (6) (2012) 613–620.
- [16] H.Y. Lee, et al., Volume-based parameter of (18)F-FDG PET/CT in malignant pleural mesothelioma: prediction of therapeutic response and prognostic implications, *Ann. Surg. Oncol.* 17 (10) (2010) 2787–2794.
- [17] J.J.M. van Griethuysen, et al., Computational radiomics system to decode the radiographic phenotype, *Cancer Res.* 77 (21) (2017) e104–e107.
- [18] P.E. Shrout, J.L. Fleiss, Intraclass correlations: uses in assessing rater reliability, *Psychol. Bull.* 86 (2) (1979) 420–428.
- [19] C. Costa-Santos, et al., The limits of agreement and the intraclass correlation coefficient may be inconsistent in the interpretation of agreement, *J. Clin. Epidemiol.* 64 (3) (2011) 264–269.
- [20] A. Depeursinge, et al., Predicting adenocarcinoma recurrence using computational texture models of nodule components in lung CT, *Med. Phys.* 42 (4) (2015) 2054–2063.
- [21] Y. She, et al., The predictive value of CT-based radiomics in differentiating indolent from invasive lung adenocarcinoma in patients with pulmonary nodules, *Eur. Radiol.* 28 (12) (2018) 5121–5128.
- [22] Z. Yang, et al., Pretreatment (18)F-FDG uptake heterogeneity can predict survival in patients with locally advanced nasopharyngeal carcinoma—a retrospective study, *Radiat. Oncol.* 10 (2015) 4.
- [23] J.F. Eary, et al., Spatial heterogeneity in sarcoma 18F-FDG uptake as a predictor of patient outcome, *J. Nucl. Med.* 49 (12) (2008) 1973–1979.
- [24] S.R. Kang, et al., Intratumoral metabolic heterogeneity for prediction of disease progression after concurrent chemoradiotherapy in patients with inoperable stage III non-small-cell lung cancer, *Nucl. Med. Mol. Imaging* 48 (1) (2014) 16–25.
- [25] S. Chicklore, et al., Quantifying tumour heterogeneity in 18F-FDG PET/CT imaging by texture analysis, *Eur. J. Nucl. Med. Mol. Imaging* 40 (1) (2013) 133–140.

- [26] A.A. Razek, A. Fathy, T.A. Gawad, Correlation of apparent diffusion coefficient value with prognostic parameters of lung cancer, *J. Comput. Assist. Tomogr.* 35 (2) (2011) 248–252.
- [27] A.A. Razek, Diffusion magnetic resonance imaging of chest tumors, *Cancer Imaging* 12 (3) (2012) 452–463.
- [28] J.C. Beer, et al., Longitudinal ComBat: a method for harmonizing longitudinal multi-scanner imaging data, *Neuroimage* 220 (2020) 117129.