Clinical Radiology 74 (2019) 467–473

Contents lists available at ScienceDirect

Clinical Radiology

journal homepage: www.clinicalradiologyonline.net

Pre-treatment ^{18}F-FDG PET-based radiomics predict survival in resected non-small cell lung cancer

H.K. Ahn [a], H. Lee [b], S.G. Kim [b], S.H. Hyun [c,*]

[a] *Division of Hematology and Oncology, Department of Internal Medicine, Gachon University Gil Medical Center, Incheon, Republic of Korea*
[b] *Department of Nuclear Medicine, Gachon University Gil Medical Center, Incheon, Republic of Korea*
[c] *Department of Nuclear Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea*

ARTICLE INFORMATION

Article history:
Received 24 October 2018
Accepted 12 February 2019

AIM: To assess the prognostic value of 2-[^{18}F]-fluoro-2-deoxy-D-glucose (FDG) positron-emission tomography (PET)-based radiomics using a machine learning approach in patients with non-small cell lung cancer (NSCLC).

MATERIALS AND METHODS: Ninety-three patients with stage I–III NSCLC who underwent combined PET/computed tomography (CT) followed by curative resection. A total of 35 unique quantitative radiomic features was extracted from the PET images, which included imaging phenotypes such as pixel intensity, shape, and texture. Radiomic features were ranked based on score according to their correlation with disease recurrence status within a 3-year follow-up. The recurrence risk classification performances of machine learning algorithms (random forest, neural network, naive Bayes, logistic regression, and support vector machine) using the 20 best-ranked features were compared using the areas under the receiver operating characteristic curve (AUC) and validated by the random sampling method.

RESULTS: Contrast and busyness texture features from neighbourhood grey-level difference matrix were found to be the two best predictors of disease recurrence. The random forest model obtained the best performance (AUC: 0.956, accuracy: 0.901, F1 score: 0.872, precision: 0.905, recall: 0.842), followed by the neural network model (AUC: 0.871, accuracy: 0.780, F1 score: 0.708, precision: 0.755, recall: 0.666).

CONCLUSION: A PET-based radiomic model was developed and validated for risk classification in NSCLC. The machine learning approach with random forest classifier exhibited good performance in predicting the recurrence risk. Radiomic features may help clinicians to improve the risk stratification for clinical practice.

Introduction

Non-small cell lung cancer (NSCLC) is the leading cause of cancer-related deaths with poor overall prognosis. More

* Guarantor and correspondent: S. H. Hyun, Department of Nuclear Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Irwonro 81, Gangnam-gu, Seoul, 06351, Republic of Korea. Tel.: +82 2 3410 2625; fax: +82 2 3410 2639.
E-mail address: hyunsh99@gmail.com (S.H. Hyun).

https://doi.org/10.1016/j.crad.2019.02.008
0009-9260/

than half of patients with NSCLC suffer from local or distant recurrence even after curative resection. The most important prognostic tool in NSCLC is the tumour–node–metastasis (TNM) stage at presentation, which also guides treatment decisions. Nevertheless, TNM stage cannot explain the wide variation of treatment responses and overall outcomes within each stage group. Therefore, prognostic biomarkers beyond the TNM staging system should be identified to stratify high-risk patients and develop risk-adapted treatment strategies.

High-throughput radiomic analysis has recently emerged as a powerful approach to identify imaging biomarkers in cancer, which refers to various mathematical methods extracting a large number of quantitative features describing imaging phenotypes such as pixel intensity, shape, and texture.[1] Radiomics can be applied to standard clinical imaging using a post-processing technique, thereby maximising the information derived from the routinely available imaging data. Several positron-emission tomography (PET)-based radiomic features, such as texture features, were previously shown as better imaging biomarkers than conventional imaging metrics in lung, oesophageal, and pancreatic cancers.[2–4]

In the present study, the prognostic value of PET-based radiomic features was investigated using a machine learning approach in patients with stage I–III NSCLC who underwent curative resection.

Materials and methods

Study population

A cohort of 93 patients newly diagnosed with stage I–III NSCLC who underwent pretreatment 2-[^{18}F]-fluoro-2-deoxy-D-glucose (FDG) PET with computed tomography (CT) followed by curative resection at Gachon University Gil Medical Center from January 2007 to December 2012 were reviewed. Patients who received neoadjuvant chemotherapy or radiotherapy were excluded. Adjuvant chemotherapy was administered in 50 patients (53.8%), and two of them received adjuvant radiotherapy.

Demographic and clinical characteristics and survival data were obtained from the medical records. The institutional review board approved this retrospective study, and the requirement of written informed consent was waived.

PET/CT imaging

All patients fasted for at least 6 hours before the PET/CT study. Blood glucose levels were required to be <150 mg/dl. Whole-body PET and unenhanced CT images were acquired using a PET/CT scanner (Biograph 6, Siemens Medical Solutions, Germany). After CT without contrast enhancement, an emission scan from the thigh to the basal skull was performed for 2 minutes per frame in the three-dimensional mode, 60 min after intravenous injection of FDG (370–555 MBq). PET images were reconstructed using CT for attenuation correction with the ordered subsets expectation maximisation algorithm (21 subsets, two iterations) with a voxel size of 4×4×2 mm. Standardised uptake value (SUV) was normalised based on the patient body weight.

Radiomic feature extraction

All combined PET/CT images were reviewed for tumour segmentation by a nuclear medicine physician with >10 years of experience who was blinded to the clinical information. The gradient-based segmentation method (PET edge) in MIM version 6.4 (MIM Software, Cleveland, OH, USA) was used for tumour delineation. A total of 35 unique quantitative radiomic feature was extracted from PET images using Chang-Gung Image Texture Analysis toolbox (CGITA, http://code.google.com/p/cigita), an open-source software package implemented in MATLAB (version 2012a; MathWorks, Natick, MA, USA).[5] Radiomic features included first-order textural features based on SUV histogram statistics (maximum, mean, standard deviation, skewness, kurtosis, and entropy), shape features (metabolic tumour volume [MTV], total lesion glycolysis [TLG]), higher-order textural features (grey-level run length matrix,

Table 1
List of 35 quantitative positron-emission tomography (PET)-based radiomic features.

Matrix	Index
Voxel statistics	Maximum SUV
	Mean SUV
	SUV SD
	SUV Skewness
	SUV Kurtosis
	Entropy
	MTV
	TLG
Grey-level run length matrix	Short run emphasis
	Long run emphasis
	Intensity variability
	Run-length variability
	Run percentage
	Low-intensity run emphasis
	High-intensity run emphasis
	Low-intensity short-run emphasis
	High-intensity short-run emphasis
	Low-intensity long-run emphasis
	High-intensity long-run emphasis
Neighbourhood grey-level difference matrix	Coarseness
	Contrast
	Busyness
	Complexity
	Strength
Grey-level size zone matrix	Short-zone emphasis
	Large-zone emphasis
	Intensity variability
	Size-zone variability
	Zone percentage
	Low-intensity zone emphasis
	High-intensity zone emphasis
	Low-intensity short-zone emphasis
	High-intensity short-zone emphasis
	Low-intensity large-zone emphasis
	High-intensity large-zone emphasis

TLG, total lesion glycolysis; MTV, metabolic tumour volume; SUV, standardised uptake value; SD, standard deviation.

Table 2
Clinical characteristics of 93 patients with non-small cell lung cancer.

Characteristic	Value
Age	63.4±10.8
Sex	
Women	38 (40.9%)
Men	55 (59.1%)
TNM stage	
I	45 (48.4%)
II	27 (29%)
III	21 (22.6%)
Histology	
Non-adenocarcinoma	35 (37.6%)
Adenocarcinoma	58 (62.4%)
Adjuvant chemotherapy	
No	43 (46.2%)
Yes	50 (53.8%)
Maximum SUV	8.4±4.7
Mean SUV	4±2.1
MTV	43.6±68.2
TLG	198.1±306.9

Data are number of patients (proportion) or mean values ± standard deviation.
SUV, standardised uptake value; SD, standard deviation; TLG, total lesion glycolysis; MTV, metabolic tumour volume.

neighbourhood grey-level difference matrix, and grey-level size zone matrix). Table 1 lists details of the radiomic features and extraction method described in a previous report.[4]

Statistical analysis and machine learning methods

The primary end-point was disease recurrence status. Disease-free survival (DFS) was measured from the date of surgery to either the date of radiologically documented recurrence or death from lung cancer.

A feature reduction procedure is necessary to select a subset of useful features that increase the prediction accuracy. Information gain ratio,[6] Gini index,[7] and chi-square statistics were used as ranking and filter-based feature selection methods to reduce the radiomic feature dimensions. Radiomic features were ranked based on the score according to their correlation with disease recurrence status within 3-year follow-up. A total of 20 best-ranked features were selected based on the overall ranks across the three scoring methods for the machine learning approach.

Five different machine learning algorithms for binary risk classification were compared, including random forest,

Table 3
Feature ranking of positron-emission tomography (PET)-based radiomics.

Overall rank	Radiomic feature	Gain ratio		Gini		χ^2	
		Rank	Value	Rank	Value	Rank	Value
1	Contrast	1	0.068	1	0.084	1	26.17
2	Busyness	2	0.065	2	0.082	2	25.72
3	Large-zone emphasis	9	0.052	5	0.068	3	23.34
4	High-intensity short-zone emphasis	3	0.056	3	0.072	12	17.13
5	TLG	5	0.053	4	0.069	10	20.01
6	Complexity	4	0.055	7	0.067	11	17.43
7	Size-zone variability	7	0.053	8	0.066	9	20.07
8	Intensity variability[a]	10	0.051	11	0.064	4	21.71
9	Intensity variability[b]	8	0.053	9	0.066	8	20.21
10	Run-length variability	11	0.051	12	0.063	6	21.14
11	High-intensity short-run emphasis	6	0.053	6	0.067	19	15.25
12	Coarseness	12	0.050	14	0.062	5	21.19
13	High-intensity large-zone emphasis	15	0.048	13	0.062	7	20.57
14	Short-zone emphasis	13	0.050	15	0.061	18	15.81
15	SUV Kurtosis	16	0.046	16	0.059	17	16.03
16	High-intensity run emphasis	14	0.050	10	0.065	28	6.30
17	Strength	19	0.042	18	0.054	15	16.46
18	Long-run emphasis	20	0.041	20	0.051	13	17.05
19	Short-run emphasis	21	0.040	21	0.050	16	16.46
20	MTV	22	0.038	22	0.050	14	16.54
21	Mean SUV	18	0.044	19	0.053	23	9.17
22	SUV SD	17	0.045	17	0.055	31	5.09
23	High-intensity zone emphasis	23	0.036	23	0.048	20	11.80
24	Zone percentage	25	0.029	26	0.038	21	11.52
25	Low-intensity zone emphasis	26	0.029	25	0.039	24	9.06
26	Low-intensity run emphasis	27	0.023	27	0.030	22	9.71
27	Maximum SUV	24	0.034	24	0.042	30	5.88
28	Low-intensity short-run emphasis	29	0.020	29	0.026	25	7.94
29	Low-intensity short-zone emphasis	28	0.021	28	0.027	27	7.70
30	High-intensity long-run emphasis	30	0.018	30	0.023	26	7.93
31	Low-intensity long-run emphasis	31	0.015	31	0.020	29	6.18
32	Run percentage	32	0.011	32	0.014	32	3.57
33	Entropy	33	0.009	33	0.012	34	0.30
34	SUV Skewness	34	0.006	34	0.008	33	0.44
35	Low-intensity large-zone emphasis	35	0.003	35	0.004	35	0.28

SUV, standardised uptake value; SD, standard deviation; TLG, total lesion glycolysis; MTV, metabolic tumour volume.
[a] Intensity variability from grey-level size zone matrix.
[b] Intensity variability from grey-level run length matrix.

Table 4
Classification performances of positron-emission tomography (PET)-based radiomics by ROC curve analyses for recurrence status within 3-year follow-up in patients with non-small cell lung cancer (n=93).

Method	AUC	Accuracy	F1	Precision	Recall
Random forest	0.956	0.901	0.872	0.905	0.842
Neural network	0.871	0.780	0.708	0.755	0.666
Naive Bayes	0.765	0.715	0.683	0.615	0.768
Logistic regression	0.663	0.655	0.507	0.591	0.443
Support vector machine	0.592	0.539	0.542	0.450	0.683

ROC, receiver operating characteristic; AUC, area under the ROC curve.

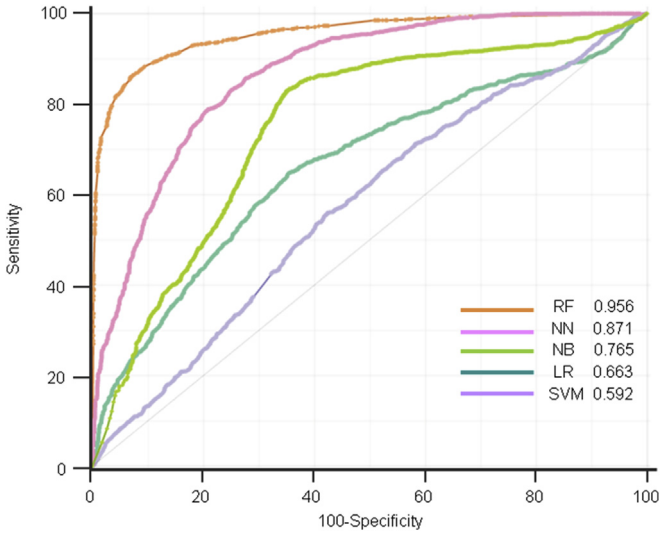

Figure 1 Classification performances of PET-based radiomics by ROC curve analyses for the recurrence status within 3-year follow-up in patients with NSCLC (n=93).

neural network, naive Bayes, logistic regression, and support vector machine. To compare their predictive performance for disease recurrence between models, the receiver operating characteristic (ROC) curve analysis was used and the areas under the ROC curve (AUC) were measured. The model performance was validated using the random sampling method, which randomly split data into the training and testing sets, and the whole procedure was repeated 100 times (training set size: 70%). The following performance measures were computed: AUC, accuracy, F1 score, precision (also called positive predictive value), and recall (also known as sensitivity). The F1 score (also known as F-score or F-measure) is the harmonic average of the precision and recall.

The prognostic association of the best-ranked radiomic features was assessed using the Cox proportional hazards regression model for additional validation. Categorical variables were dichotomised into two groups for statistical analyses. Tumours were histologically categorised as adenocarcinoma and non-adenocarcinoma. Squamous cell carcinomas were predominant among the non-adenocarcinomas group. Kaplan–Meier curves were generated with an optimal cut-off derived from maximally selected rank statistics and were compared using the log-rank test.[8]

The statistical analyses were performed using the R software version 3.5.1 (R Foundation, Vienna, Austria). The machine learning approach was performed using Orange version 3.16 (Bioinformatics Lab at University of Ljubljana, Slovenia), an open-source data mining and visualisation package.[9] All tests were two-sided. Confidence intervals (CIs) were reported at the 95% level, and $p<0.05$ was considered statistically significant.

Results

Patient characteristics

The patient demographic and clinical characteristics are summarised in Table 2. At the time of analysis, 36 (38.7%) recurrences and 18 (19.4%) deaths occurred during a median of 45 months of follow-up. The median DFS of the whole population was 61.3 months, and the 5-year DFS rate was 50.7%.

Radiomic feature selection and ROC curve analysis

Radiomic features were ranked according to the scoring methods (Table 3). Contrast and busyness texture features from neighbourhood grey-level difference matrix (NGDM) were found to be the two best predictors of disease

Table 5
Univariate and multivariate Cox regression analysis of disease-free survival in patients with non-small cell lung cancer (n=93).

Variable	Univariate analysis			Multivariate analysis		
	HR	95% CI	p-Value	HR	95% CI	p-Value
Age (1-year increase)	1.01	0.98–1.04	0.428	1.02	0.99–1.05	0.234
Sex, men versus women	0.80	0.43–1.49	0.492	0.92	0.44–1.95	0.834
TNM stage III versus I-II	2.61	1.38–4.93	0.003	2.13	1.09–4.16	0.028
Histology, adenocarcinoma versus non-adenocarcinoma	1.49	0.77–2.87	0.230	1.81	0.84–3.89	0.129
PET radiomic feature (contrast, $\log_2$ scale)	0.84	0.73–0.96	0.011	0.84	0.73–0.97	0.020

HR, hazard ratio; CI, confidence interval.

recurrence. The overall classification performances of five machine learning methods were compared by AUC using the 20 best-ranked radiomic features (Table 4, Fig 1). The random forest model obtained the best performance (AUC: 0.956, accuracy: 0.901, F1 score: 0.872, precision: 0.905, recall: 0.842), followed by neural network model (AUC: 0.871, accuracy: 0.780, F1 score: 0.708, precision: 0.755, recall: 0.666).

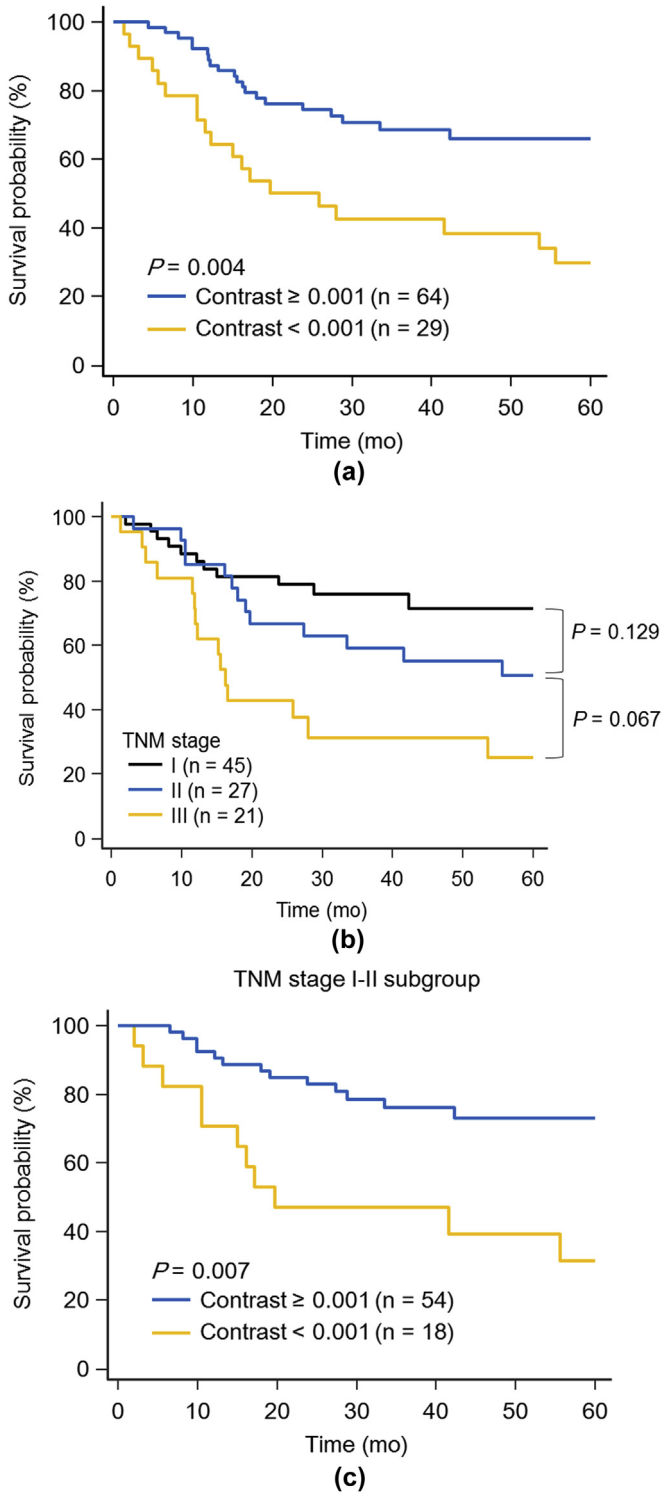

Figure 2 Kaplan–Meier survival curves of DFS according to contrast (a), TNM stages (b), and contrast in TNM stages I–II subgroup (c) in patients with NSCLC (n=93).

Survival analysis

As contrast texture feature from NGDM was considered the best-ranked radiomics, contrast was selected for further survival analysis. Contrast, age, sex, pathological TNM stage, and tumour histology were entered into Cox regression models. In univariate analyses, pathological TNM stage and contrast were significant predictors of DFS (Table 5). After adjusting for age, sex, pathological TNM stage, and tumour histology, multivariate Cox analysis demonstrated that contrast (HR, 0.84; 95% CI 0.73–0.97; p=0.020) was a significant independent prognostic imaging biomarker (Table 5).

Kaplan–Meier analysis of the entire cohort revealed significantly worse survival in low-contrast tumours (Fig 2a). Contrast as a PET imaging biomarker allowed a better risk stratification in early stage (TNM stages I–II) patients (Fig 2b and c). Representative cases are presented in Fig 3.

Discussion

A PET-based radiomic model was developed and validated for risk classification in NSCLC. The random forest classifier greatly predicted disease recurrence. Contrast and busyness texture features from NGDM had better prognostic value than conventional PET metrics such as MTV, TLG, or SUV.

The measurement of tumour texture features using PET-based radiomic analysis has been reported as a promising tool for prognostication in NSCLC; however, currently, the available data are limited.[10,11] Several PET-based radiomics measured by the first (entropy)-, second (dissimilarity)-, and higher-order (coarseness, contrast, busyness) textural features have been reported as independent predictors of outcome in patients with NSCLC.[3,12,13] The present results add to the accumulating evidence that contrast as a measurement of tumour texture is associated with survival. High contrast level from NGDM has been significantly associated with better progression-free survival in NSCLC.[3] Compared with non-responders to chemoradiotherapy, responders showed higher contrast at baseline PET. A recent study has shown the relationship between higher contrast and improved DFS in NSCLC treated with stereotactic body radiation therapy.[13]

Higher-order texture features from NGDM, such as contrast and busyness, describe the local texture features based on the differences between each voxel and the neighbouring voxels (Table 6). Contrast correlates with the difference between neighbouring regions of voxel intensities. High contrast within an image indicates that voxel intensity significantly differs between the neighbouring voxels. Busyness relates to the change rate between neighbourhood intensities weighted by the difference in intensities. A busy texture is characterised by rapid intensity changes in adjacent voxels.[14] In the present study,

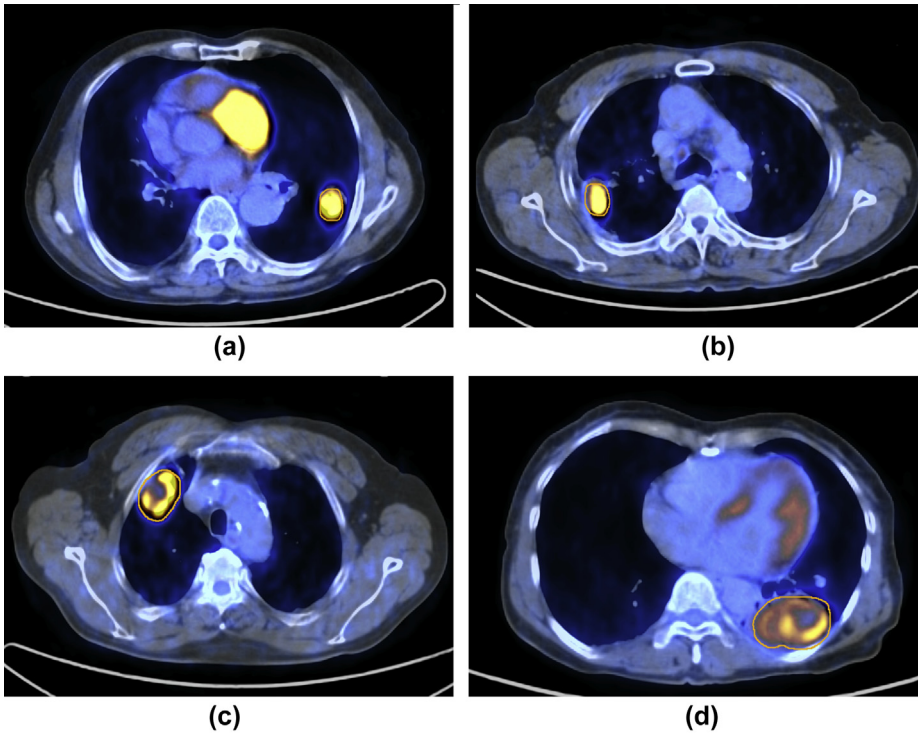

Figure 3 Representative images of NSCLC using the PET texture analysis. Axial fused ^{18}F-FDG PET/CT images show tumours with high contrast low busyness (a,b) and low contrast high busyness (c,d) texture features. (a) A 73-year-old man with stage IB squamous cell carcinoma; DFS=69 months, contrast=0.0026, busyness=0.0518, maximum SUV=11.0, mean SUV=5.6. (b) A 69-year-old man with stage IB squamous cell carcinoma; DFS=43 months, contrast=0.0019, busyness=0.0859, maximum SUV=10.3, mean SUV=4.8. (c) A 69-year-old man with stage IB large cell carcinoma; DFS=6 months, contrast=0.0007, busyness=0.2005, maximum SUV=10.9, mean SUV=4.4. (d) A 63-year-old woman with stage IIA adenocarcinoma; DFS=12 months, contrast=0.0001, busyness=0.4930, maximum SUV=6.8, mean SUV=3.3.

tumours with poor prognosis showed low contrast and high busyness texture features.

The present results have significant clinical implications. Combining information on PET-based radiomics and TNM stage allowed more powerful outcome prediction in patients with NSCLC. Currently, a standardised imaging method to quantify intratumoural heterogeneity is not available. PET-based radiomic analysis allowed assessment of the tumour texture features and intratumoural heterogeneity. This imaging biomarker can help improve risk stratification and optimise patient selection for clinical trials. Patients with high-risk radiomic features could benefit from more aggressive treatment such as (neo)adjuvant therapy with clinical-trial drugs. PET-based radiomic features also enable appropriate postoperative surveillance models for high-risk patients.

This study has several limitations. First, its retrospective nature and lack of external validation data inherently limited the general applications of results. Although data that support the clinical value of intratumoural heterogeneity measured using radiomic analysis were provided, the present results need to be validated in a larger external cohort. Second, several technical limitations exist in the methods of PET-based radiomic analysis. Texture feature calculations are affected by several factors, including signal-to-noise ratio, volume of interest definition, partial volume effect, and reconstruction settings. In the present study, commercially available PET edge software using a gradient-based segmentation method was used. Reproducibility and reliability of this method have already been validated in previous studies.[15,16] The PET-based radiomic analysis in small tumours with <1 cm as the largest diameter cannot be reliably performed because of the small number of voxels included, the partial volume effect, and the limited spatial resolution of PET scanner. Unfortunately, the partial volume effect correction was not performed because there was no

Table 6
Equations of contrast and busyness texture features derived from neighbourhood grey-level difference matrix.

Feature	Equation	Definition
Contrast	$\left[\frac{1}{N_g(N_g-1)}\sum_{i=0}^{G_h}\sum_{j=0}^{G_h}p_ip_j(i-j)^2\right]\left[\frac{1}{n^3}\sum_{i=0}^{G_h}s(i)\right]$, $p_i\neq 0, p_j\neq 0$	Measures the difference between neighbouring regions of voxel intensities.
Busyness	$\left[\sum_{i=0}^{G_h}p_is(i)\right]\Big/\left[\sum_{i=0}^{G_h}\sum_{j=0}^{G_h}(ip_i-jp_j)\right]$, $p_i\neq 0, p_j\neq 0$	Measures the change rate between neighbourhood intensities weighted by the difference in intensities.

widely accepted solution. It was assumed that the partial volume effect only slightly influenced the results because most patients (87%) had tumours with the largest diameter of >2 cm. Finally, intratumoural heterogeneity is a well-recognised feature of malignant tumours; however, the relationship between intratumoural heterogeneity in FDG uptake and its underlying biology remains to be elucidated. Malignancy has a very complex biology, and many components were involved in intratumoural heterogeneity, including tumour cell density distribution, cellular proliferation, metabolism, hypoxia, vascularity, necrosis, haemorrhage, and microenvironment.[17–19] Further studies integrating imaging phenotypes with pathological components and genetic features are required.

In conclusion, a PET-based radiomic model was developed and validated for risk classification in NSCLC. The machine learning approach with random forest classifier exhibited good performance in predicting the recurrence risk. Radiomic features may help clinicians to improve the risk stratification for clinical practice. Further prospective validation studies are required to confirm the practical applicability of this potential imaging biomarker.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education (2017R1 D1A1B03028735).

References

1. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2016;**278**:563–77.
2. Cheng NM, Fang YH, Lee LY, *et al*. Zone-size nonuniformity of ^{18}F-FDG PET regional textural features predicts survival in patients with oropharyngeal cancer. *Eur J Nucl Med Mol Imaging* 2015;**42**:419–28.
3. Cook GJ, Yip C, Siddique M, *et al*. Are pretreatment ^{18}F-FDG PET tumour textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? *J Nucl Med* 2013;**54**:19–26.
4. Hyun SH, Kim HS, Choi SH, *et al*. Intratumoural heterogeneity of F-FDG uptake predicts survival in patients with pancreatic ductal adenocarcinoma. *Eur J Nucl Med Mol Imaging* 2016;**43**:1461–8.
5. Fang YH, Lin CY, Shih MJ, *et al*. Development and evaluation of an open-source software package "CGITA" for quantifying tumour heterogeneity with molecular images. *Biomed Res Int* 2014;**2014**:248505.
6. Quinlan JR. Induction of decision trees. *Machine Learn* 1986;**1**:81–106.
7. Gini C. Concentration and dependency ratios. *Rivista di Politica Economica* 1909;**87**:769–89.
8. Lausen B, Schumacher M. Maximally selected rank statistics. *Biometrics* 1992:73–85.
9. Demsar J, Curk T, Erjavec A, *et al*. Orange: data mining toolbox in Python. *J Machine Learn Res* 2013;**14**:2349–53.
10. Kirienko M, Cozzi L, Antunovic L, *et al*. Prediction of disease-free survival by the PET/CT radiomic signature in non-small cell lung cancer patients undergoing surgery. *Eur J Nucl Med Mol Imaging* 2018;**45**:207–17.
11. Zhang Y, Oikonomou A, Wong A, *et al*. Radiomics-based prognosis analysis for non-small cell lung cancer. *Sci Rep* 2017;**7**:46349.
12. Cook GJ, O'Brien ME, Siddique M, *et al*. Non-small cell lung cancer treated with erlotinib: heterogeneity of (18)F-FDG uptake at PET-association with treatment response and prognosis. *Radiology* 2015;**276**:883–93.
13. Lovinfosse P, Janvary ZL, Coucke P, *et al*. FDG PET/CT texture analysis for predicting the outcome of lung cancer treated by stereotactic body radiation therapy. *Eur J Nucl Med Mol Imaging* 2016;**43**:1453–60.
14. Amadasun M, King R. Textural features corresponding to textural properties. *IEEE Trans Sys Man Cybern* 1989;**19**:1264–74.
15. Shah B, Srivastava N, Hirsch AE, *et al*. Intra-reader reliability of FDG PET volumetric tumour parameters: effects of primary tumour size and segmentation methods. *Ann Nucl Med* 2012;**26**:707–14.
16. Werner-Wasik M, Nelson AD, Choi W, *et al*. What is the best way to contour lung tumours on PET scans? Multiobserver validation of a gradient-based method using a NSCLC digital PET phantom. *Int J Radiat Oncol Biol Phys* 2012;**82**:1164–71.
17. Pugachev A, Ruan S, Carlin S, *et al*. Dependence of FDG uptake on tumour microenvironment. *Int J Radiat Oncol Biol Phys* 2005;**62**:545–53.
18. Gerlinger M, Rowan AJ, Horswell S, *et al*. Intratumour heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 2012;**366**:883–92.
19. Ganeshan B, Goh V, Mandeville HC, *et al*. Non-small cell lung cancer: histopathologic correlates for texture parameters at CT. *Radiology* 2013;**266**:326–36.