The Impact of Optimal Respiratory Gating and Image Noise on Evaluation of Intratumor Heterogeneity on ^{18}F-FDG PET Imaging of Lung Cancer

Willem Grootjans[1,2], Florent Tixier[2,3], Charlotte S. van der Vos[2], Dennis Vriens[1], Catherine C. Le Rest[3], Johan Bussink[4], Wim J.G. Oyen[5], Lioe-Fee de Geus-Oei[6], Dimitris Visvikis[7], and Eric P. Visser[2]

[1]Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands; [2]Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen, The Netherlands; [3]Department of Nuclear Medicine, DACTIM, University Hospital Poitiers, Poitiers, France; [4]Department of Radiation Oncology, Radboud University Medical Center, Nijmegen, The Netherlands; [5]Institute of Cancer Research and Royal Marsden NHS Foundation Trust, London, United Kingdom; [6]MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, Enschede, The Netherlands; and [7]INSERM, UMR1101, LaTIM, University of Brest, Brest, France

Accurate measurement of intratumor heterogeneity using parameters of texture on PET images is essential for precise characterization of cancer lesions. In this study, we investigated the influence of respiratory motion and varying noise levels on quantification of textural parameters in patients with lung cancer. **Methods:** We used an optimal-respiratory-gating algorithm on the list-mode data of 60 lung cancer patients who underwent ^{18}F-FDG PET. The images were reconstructed using a duty cycle of 35% (percentage of the total acquired PET data). In addition, nongated images of varying statistical quality (using 35% and 100% of the PET data) were reconstructed to investigate the effects of image noise. Several global image-derived indices and textural parameters (entropy, high-intensity emphasis, zone percentage, and dissimilarity) that have been associated with patient outcome were calculated. The clinical impact of optimal respiratory gating and image noise on assessment of intratumor heterogeneity was evaluated using Cox regression models, with overall survival as the outcome measure. The threshold for statistical significance was adjusted for multiple comparisons using Bonferroni correction. **Results:** In the lower lung lobes, respiratory motion significantly affected quantification of intratumor heterogeneity for all textural parameters ($P < 0.007$) except entropy ($P > 0.007$). The mean increase in entropy, dissimilarity, zone percentage, and high-intensity emphasis was 1.3% ± 1.5% ($P = 0.02$), 11.6% ± 11.8% ($P = 0.006$), 2.3% ± 2.2% ($P = 0.002$), and 16.8% ± 17.2% ($P = 0.006$), respectively. No significant differences were observed for lesions in the upper lung lobes ($P > 0.007$). Differences in the statistical quality of the PET images affected the textural parameters less than respiratory motion, with no significant difference observed. The median follow-up time was 35 mo (range, 7–39 mo). In multivariate analysis for overall survival, total lesion glycolysis and high-intensity emphasis were the two most relevant image-derived indices and were considered to be independent significant covariates for the model regardless of the image type considered. **Conclusion:** The tested textural parameters are robust in the presence of respiratory motion artifacts and varying levels of image noise.

The combined use of PET imaging with CT imaging has gradually evolved from a diagnostic tool toward a multirole imaging platform for the management of patients with lung cancer (1). The advantage of PET over other tomographic imaging modalities is the ability to characterize and quantify the biologic landscape of cancerous lesions with high sensitivity, making it possible to identify areas that are linked to therapy resistance or are more aggressive (1).

In this regard, it is becoming increasingly important to develop PET image–derived indices with the objective of extracting as much information from the images as possible (2). Traditional PET image–derived indices typically rely on quantification of lesion SUV and overall tumor volume, which have been shown to be independent prognostic factors for patient outcome and treatment response (3). Although useful, these parameters do not reveal the spatial distribution and specific pattern of radiotracer accumulation within the tumor, limiting the possibility of further characterizing its biologic behavior.

Interest in the quantification of intrinsic spatial and temporal heterogeneity within solid malignancies has been growing over the last few years. Particularly, there has been increasing recognition of the role of medical imaging in identifying specific tumor phenotypes (4), predicting treatment resistance (5–7), and projecting overall survival (OS) (8). This view fits our current knowledge of cancer, in which malignant lesions consist of heterogeneous cell populations with distinct molecular and microenvironmental differences (9). Hence arises the current interest in using medical imaging to repetitively assess intratumor spatial and temporal heterogeneity (4).

However, in order to characterize cancer lesions with high precision, it is essential to assess the accuracy of measurements made under different imaging conditions (10,11). The accuracy

TABLE 1
Patient Characteristics

Characteristic	Data
Sex (n)	
Male	43
Female	17
Age (y)	
Median	68
Range	49–85
Mean (±SD) body mass (kg)	77 ± 15
Mean (±SD) administered activity (MBq)	249 ± 48
Mean (±SD) postadministration interval (min)	79 ± 15
Histologic subtype (n)	
Small cell lung cancer	7
Squamous cell carcinoma	28
Adenocarcinoma	25
Disease stage (n)	
IA	5
IB	4
IIA	6
IIB	2
IIIA	17
IIIB	10
IV	16

Disease stage is according to seventh revision of International Association for Study of Lung Cancer.

and robustness of these measurements is related to the PET acquisition and reconstruction protocols, as well as variability in patient physiology. In particular, blurring of PET images due to respiratory motion can significantly influence quantification of lung lesions (12,13). In addition, PET textural parameters can be sensitive to variations in statistical quality and to normal stochastic variations in images (14). In this study, we investigated the clinical impact of respiratory gating and varying levels of image noise on quantification of several textural parameters in high-resolution time-of-flight PET imaging with the glucose analog ^{18}F-FDG.

MATERIALS AND METHODS

Patients

The institutional review board of Radboud University Medical Center approved this retrospective study, and the requirement to obtain informed consent was waived. From our fast-track outpatient diagnostic program, 60 patients with histologically proven lung cancer were chosen for the study. Only lesions at least 3 cm^3 in volume were included, since this minimum has been found necessary for calculation of meaningful and complementary information on intratumoral heterogeneity using textural analysis of ^{18}F-FDG PET images (15). The patient characteristics are summarized in Table 1.

Image Acquisition and Reconstruction

Whole-body ^{18}F-FDG PET imaging was performed using a 40-slice Biograph mCT PET/CT scanner (Siemens Medical Solutions). The amount of administered ^{18}F-FDG was adjusted to each patient's weight (3.2 ± 0.3 MBq/kg). Full details on image acquisition and reconstruction have been described previously (12,16). In short, imaging of the thorax and upper abdomen was performed in list mode at 6 min per bed position, and the respiratory signal was obtained using an AZ-733V respiratory gating system (Anzai Medical Co. Ltd.). Reconstruction was performed using 3 iterations, 21 subsets (the ultra HD·PET setting of the Biograph mCT), and a transaxial matrix of 400 × 400 voxels. Postreconstruction filtering was applied using a 3-dimensional gaussian filter kernel with a full width at half maximum of 3.0 mm.

Respiratory Gating

Respiratory gating was performed on the list-mode data using an amplitude-based optimal-respiratory-gating algorithm (HD·Chest) integrated in the *syngo* molecular imaging PET/CT software (version 2011A; Siemens Medical Solutions). The main user input for the algorithm is the percentage duty cycle—the percentage of the total acquired data used for image reconstruction (12). The optimal-respiratory-gating images were reconstructed with a duty cycle of 35% (ORG$_{35\%}$), which was previously found to provide the best balance between image quality and motion rejection (12). Nongated images equivalent in statistical quality to the ORG$_{35\%}$ images were reconstructed using the first 126 s (35%) of the acquired PET data (NG$_{35\%}$). Furthermore, nongated images using the full 360 s (100%) of the acquired PET data were reconstructed (NG$_{100\%}$). The different reconstructed images used for analysis are shown in Figure 1.

For each patient, additional gated images during maximum inspiration and maximum expiration (by restricting the amplitude range to the maximum and minimum amplitude of the respiratory signal, respectively) were reconstructed to calculate lesion displacement during the respiratory cycle. These images were reconstructed with a duty cycle of 20%
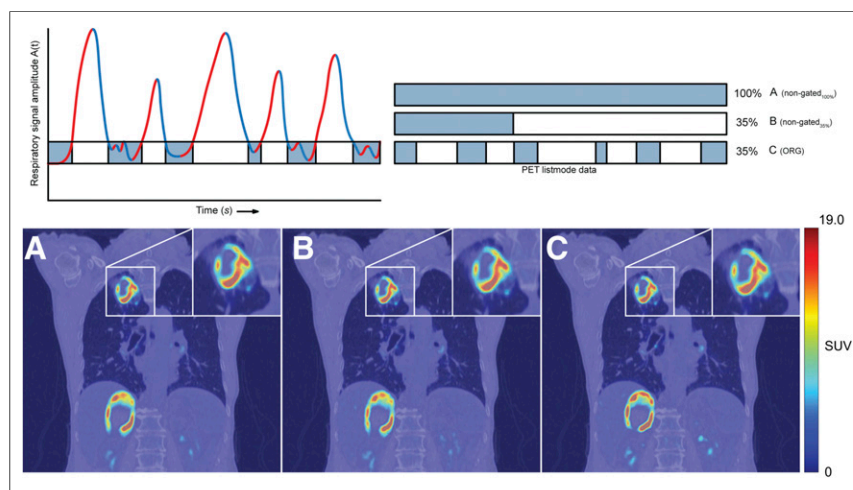
FIGURE 1. Reconstructed PET images used for analysis: NG$_{100\%}$ (A), NG$_{35\%}$ (B), and ORG$_{35\%}$ (C). Optimal respiratory gating is performed through rebinning of list-mode data according to simultaneously acquired respiratory signal.
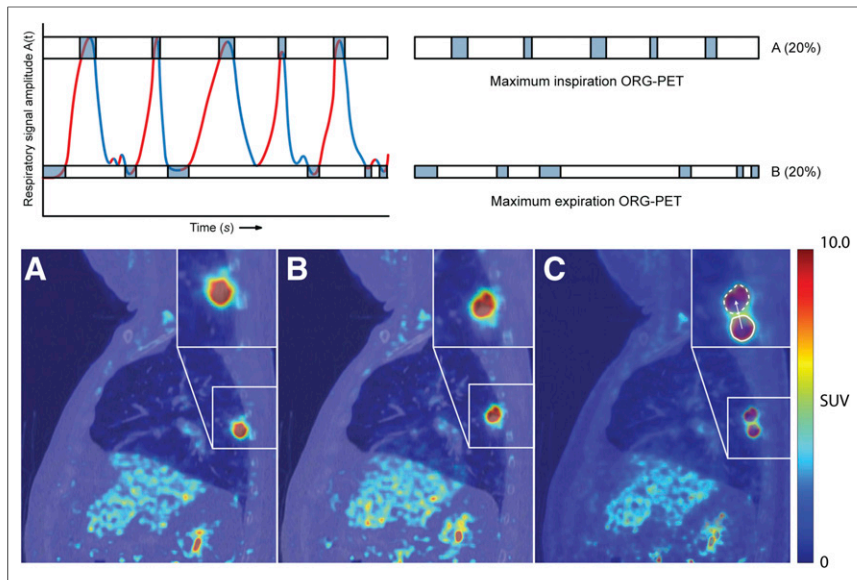
FIGURE 2. Analysis of tumor displacement due to respiratory cycle. Shown are ORG$_{20\%}$ images at maximum inspiration (A) and maximum expiration (B), along with a fusion of the two (C). Dashed contour indicates lesion at maximum expiration; solid contour indicates lesion at maximum inspiration.

(ORG$_{20\%}$) to minimize the impact of residual motion in the reconstructed PET images. The displacement vector was determined by delineating lesions on the maximum-expiration and maximum-inspiration images and calculating the distance between the lesion centers (Fig. 2).

Image Analysis

In order to determine the effect of respiratory motion, the textural parameters were calculated on the ORG$_{35\%}$ images and compared with the respective NG$_{35\%}$ images. The effect of image noise was investigated by comparing the NG$_{35\%}$ images with the NG$_{100\%}$ images. The lesions were delineated using fuzzy locally adaptive Bayesian segmentation with 2 segmentation classes (17) and were evaluated for 4 textural parameters: entropy, dissimilarity, high-intensity emphasis, and zone percentage. The first two describe local heterogeneity (variations in intensity between each voxel and its immediate neighbors averaged over the entire volume), and the third and fourth describe regional heterogeneity (at the level of groups of voxels and areas of various sizes and intensities). These parameters have been found useful for predicting prognosis in patients with non–small cell lung cancer (8). The voxel intensities were grouped using 64 levels of gray, and the local-heterogeneity parameters were computed over 13 directions (18). In addition to evaluating the 4 textural parameters, we extracted from the PET images 3 global image–derived indices: metabolic tumor volume, SUV$_{mean}$, and total lesion glycolysis (TLG).

The lesions were categorized as being located in the upper lung lobes, in the middle and lower lung lobes, or centrally in the lung hilum or mediastinum (12). Given that the absolute and particular spatial distribution of ^{18}F-FDG uptake can differ for different histologic subtypes (19), we compared 3 types of lung cancer: adenocarcinoma, squamous cell carcinoma, and small cell lung carcinoma.

Statistics

Statistics were analyzed using SPSS Statistics (version 21; IBM). Bonferroni adjustment was applied for multiple testing (i.e., $P_{critical} = P_\alpha/k$, where $P_{critical}$ is the threshold for statistical significance, P_α is the α-probability [0.05], and k is the number of performed tests). Statistical significance was then defined for $P < P_{critical}$. The Wilcoxon signed rank test was applied for paired measurements (7 comparisons, $P_{critical} = 0.007$). The Kruskal–Wallis H test was applied for group comparisons regarding tumor displacement (3 comparisons, $P_{critical} = 0.02$) and histologic subtype (8 comparisons, $P_{critical} = 0.006$). To analyze OS, we used univariate and multivariate Cox regression. These analyses included only patients with non–small cell lung cancer and used OS as the outcome measure (the interval between the PET acquisition and death). The closeout date was August 2015. Patients who were alive on that date were censored for OS on that date. The multivariate Cox regression models were obtained using semiautomated iterative forward and backward selection of image-derived parameters based on the likelihood-ratio criterion. The hazard ratios with their corresponding 95% confidence intervals are reported. A maximum of 4 covariates was chosen so that the number of events per covariate would be sufficient for reliable statistical assessment (20). In addition, Kaplan–Meier analysis was performed to determine the association of different PET-derived image indices with OS. All variables were split at their median to prevent data-driven dichotomization, yielding a low group and a high group of similar size. The Kaplan–Meier curves were compared using Mantel–Cox (log rank) statistics.

RESULTS

Lesion Displacement

Displacement during respiration was typically largest for lesions in the lower lobes—a statistically significantly difference from the other two anatomic groups ($P < 0.02$). Displacement of the centrally located lesions was more heterogeneous, with hilar lesions typically exhibiting considerable displacement but mediastinal lesions remaining almost stationary. Lesions in the upper lobe, particularly in the apical segments, showed almost no displacement. Lesion displacement as a function of anatomic location is summarized in Figure 3.
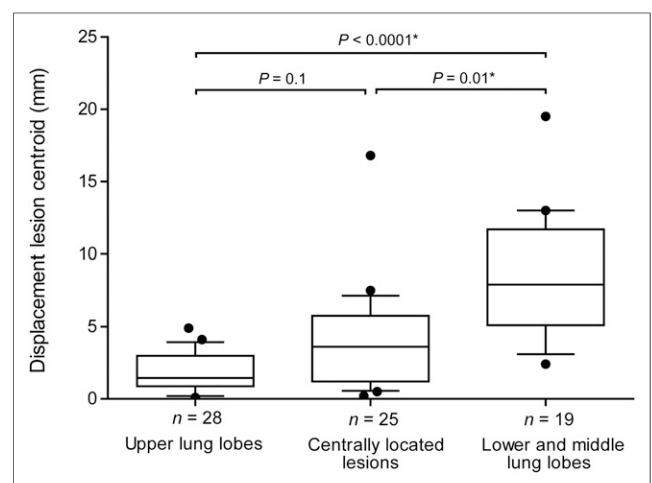
FIGURE 3. Quantification of lesion displacement as function of anatomic group. Whiskers indicate 10th and 90th percentiles; outliers are indicated by dot. *Statistically significant difference between groups.
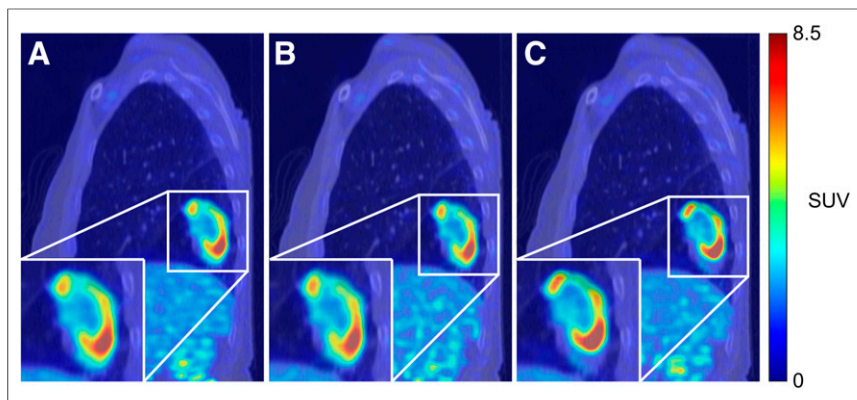
FIGURE 4. $NG_{100\%}$ (A), $NG_{35\%}$ (B), and $ORG_{35\%}$ (C) images of patient with non–small cell lung cancer lesion in right lower lung lobe.

Global Image–Derived Indices

In comparison with the $NG_{35\%}$ images, the $ORG_{35\%}$ images had a significant increase in SUV_{mean} (2.9% ± 13.0%, $P < 0.0001$) and decrease in metabolic tumor volume (3.6% ± 15.1%, $P < 0.0001$). However, there were no statistically significant differences in TLG between the $NG_{35\%}$ and $ORG_{35\%}$ images ($P > 0.007$).

Textural Parameters

In the cohort as a whole, respiratory gating did not result in statistically significant differences in any textural parameters between the $NG_{35\%}$ and $ORG_{35\%}$ images ($P > 0.007$). The mean increase in entropy, dissimilarity, zone percentage, and high-intensity emphasis was 0.3% ± 2.7% ($P = 0.5$), 3.6% ± 14.3% ($P = 0.2$), 0.5% ± 3.3% ($P = 0.3$), and 4.2% ± 21.4% ($P = 0.3$), respectively. For lesions in the middle and lower lobes, there was a statistically significant difference in all textural parameters except entropy ($P > 0.007$). The mean increase in entropy, dissimilarity, zone percentage, and high-intensity emphasis between the $NG_{35\%}$ and $ORG_{35\%}$ images was 1.3% ± 1.5% ($P = 0.02$), 11.6% ± 11.8% ($P = 0.006$), 2.3% ± 2.2% ($P = 0.002$), and 16.8% ± 17.2% ($P = 0.006$), respectively. Figure 4 presents the images of a patient with a lower-lobe lesion. For centrally located lesions, the mean increase in entropy, dissimilarity, zone percentage, and high-intensity emphasis was 0.58% ± 3.7% ($P = 0.6$), 5.0% ± 19.0% ($P = 0.4$), 0.59% ± 4.0% ($P = 0.9$), and 4.4% ± 27.8% ($P = 0.4$), respectively.

Lesions in the upper lobes showed a mean decrease of 0.35% ± 1.8% ($P = 0.3$), 1.0% ± 7.7% ($P = 0.3$), 0.4% ± 2.7% ($P = 0.5$), and 1.7% ± 13.2% ($P = 0.4$) in entropy, dissimilarity, zone percentage, and high-intensity emphasis, respectively. There was no significant correlation between lesion volume and the change in heterogeneity parameters between $NG_{35\%}$ and $ORG_{35\%}$ images.

Entropy and high-intensity emphasis were the parameters most affected by the change in the statistical quality of the images. Comparison of the $NG_{100\%}$ and $NG_{35\%}$ images showed a statistically significant difference in entropy (−0.6% ± 1.8%, $P = 0.002$) and high-intensity emphasis (9.1% ± 22.0%, $P < 0.0001$). Dissimilarity and zone percentage were not significantly affected ($P > 0.007$), with a difference of 0.7% ± 8.0% ($P = 0.4$) and −0.4% ± 3.4% ($P = 0.2$), respectively. Measurement variability due to respiratory motion typically exceeded that due to differences in noise level on the images.

Table 2 sorts the textural parameters by histologic subtype. There were no statistically significant differences in any textural parameters between histologic subtypes ($P > 0.006$). Furthermore, optimal respiratory gating did not influence the characterization of intratumor heterogeneity for any of the histologic subtypes, all of which had a similar data distribution on the $NG_{35\%}$ and $ORG_{35\%}$ images.

Univariate Analysis

Of the 60 patients, 53 were diagnosed with non–small cell lung cancer and could be included in the OS analysis. The median follow-up time was 35 mo (range, 7–39 mo). During this period, 38 of the 53 patients (72%) died, all due to cancer progression.

The Kaplan–Meier curves obtained for TLG were similar for all 3 image types, with no differences between the low and high TLG groups. Figure 5 depicts the Kaplan–Meier curves for the summed TLG of all intrapulmonary lesions larger than 3 cm³. The OS curves of summed TLG dichotomized at their median were significantly different, with a strong association of lower TLG values with longer OS ($P = 0.008$). Furthermore, there were only minor

TABLE 2
Intratumor Heterogeneity in Glucose Metabolism on $NG_{35\%}$ and $ORG_{35\%}$ Images, by Histologic Subtype

Parameter	NG_{35%}			Significance		ORG_{35%}			Significance	
	AC	SCC	SCLC	χ^2	P	AC	SCC	SCLC	χ^2	P
Entropy	7.29 ± 0.28	7.70 ± 0.31	7.21 ± 0.26	2.9	0.2	7.31 ± 0.29	7.35 ± 0.29	7.24 ± 0.28	1.1	0.6
Dissimilarity	8.62 ± 1.76	9.00 ± 1.92	9.11 ± 2.57	0.93	0.6	8.92 ± 1.85	9.12 ± 2.07	9.57 ± 2.42	0.64	0.7
Zone percentage	0.85 ± 0.04	0.86 ± 0.05	0.85 ± 0.05	2.3	0.3	0.86 ± 0.05	0.86 ± 0.05	0.87 ± 0.05	0.65	0.7
High-intensity emphasis	519 ± 182	554 ± 156	483 ± 128	1.6	0.5	539 ± 195	561 ± 157	477 ± 134	1.4	0.5

AC = adenocarcinoma ($n = 25$); SSC = squamous cell carcinoma ($n = 28$); SCLC = small cell lung cancer ($n = 7$).
Data are mean ± SD.
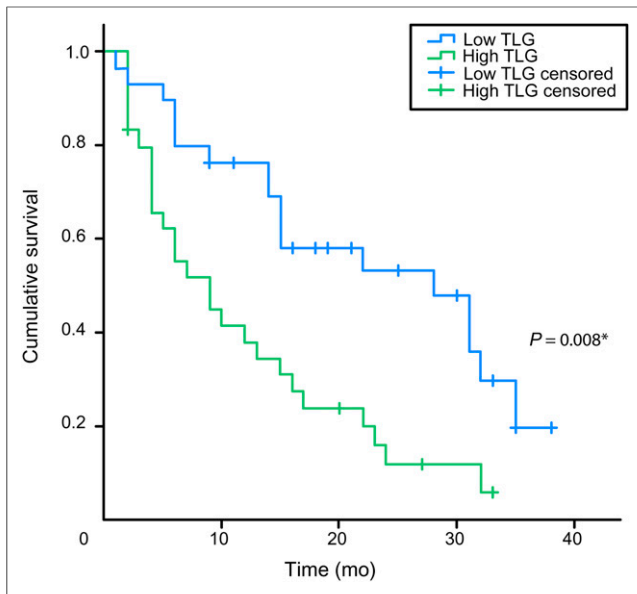
FIGURE 5. Kaplan–Meier survival curves for summed TLG.

differences in the Kaplan–Meier curves for entropy, dissimilarity, zone percentage, and high-intensity emphasis between the $ORG_{35\%}$ and $NG_{35\%}$ images. Table 3 summarizes the log rank comparison of the OS curves for several image-derived indices.

The covariates used in the univariate Cox regression analysis, along with their respective hazard ratios and significance levels, are summarized in Tables 4 and 5. Of the clinical covariates, treatment ($P < 0.0001$) and disease stage ($P = 0.02$) were the only significant predictors of OS. Of the image-derived indices, primary-tumor metabolic volume, primary-tumor TLG, and summed TLG were significant predictors of OS. The SUV_{mean} of the primary tumor was not significantly associated with OS for any of the 3 image types. Furthermore, none of the 4 textural parameters was significantly predictive of OS in this patient group.

Multivariate Analysis

Given that patient treatment and clinical stage had a statistically significant association with OS in this patient cohort and are generally known to have a strong association with OS, we forced

TABLE 3
P Values from Log Rank Comparison of Kaplan–Meier Curves for Different Image-Derived Indices

Index	$NG_{35\%}$	$ORG_{35\%}$	$NG_{100\%}$
Primary-tumor metabolic volume	0.2	0.06	0.2
Primary-tumor TLG	0.03*	0.03*	0.03*
Summed TLG	0.008*	0.008*	0.008*
Entropy	0.6	0.3	0.9
Dissimilarity	0.4	0.3	0.2
Zone percentage	0.6	0.2	0.4
High-intensity emphasis	0.6	0.1	0.1

*Statistically significant.

TABLE 4
Univariate Cox Regression Analysis of Clinical Covariates for OS

Covariate	Hazard ratio	95% confidence interval	*P*
Treatment			<0.0001*
Surgery	1		
Chemoradiotherapy	0.5	0.3–0.9	
Chemotherapy	1.7	1.0–2.9	
Expectative (none)	3.3	1.7–6.3	
Stage			0.02*
I	1		
II	0.4	0.07–1.9	
III	1.6	0.6–4.4	
IV	4.3	1.4–13	
Histologic subtype			
Adenocarcinoma	1		0.3
Squamous cell carcinoma	0.7	0.4–1.4	
Age	1.01	1.0–1.1	0.4

*Statistically significant.

these 2 covariates in the multivariate model. The multivariate models obtained through iterative backward selection of relevant covariates are summarized in Table 6. There were no differences in the multivariate models obtained by forward and backward selection of covariates for any of the 3 image types. In these models, high-intensity emphasis and summed TLG were included as independent prognostic image-derived parameters.

DISCUSSION

This study showed that respiration during PET imaging affects quantification of the heterogeneity of glucose metabolism within lesions in the lower lobes of the lung. Furthermore, the statistical quality of PET images was shown to affect textural parameters to a lesser extent than respiratory motion, with no significant differences found between images of different statistical quality. The blurring effect of respiratory motion on quantification of intratumor heterogeneity was shown in studies by Yip et al. (*13*), and Oliver et al. (*21*). Although our results are in line with theirs, the observed differences between the respiratory gated and nongated PET images were usually small. Furthermore, the clinical impact of the observed differences was limited, with no differences being found in the multivariate models for OS.

With optimal respiratory gating, we were able to evaluate solely the effect of respiratory motion on quantification of intratumor heterogeneity and to reduce the confounding effects of varying noise levels—a task that is more challenging with other respiratory gating methods (*12*). We showed that the blurring effect of respiratory motion had the largest impact on lesions in the middle and lower lobes but only a limited impact on lesions in the upper lobes. The effect of respiratory motion on blurring of centrally located lesions was more variable, with hilar lesions sometimes

TABLE 5
Univariate Analysis of Image-Derived Covariates for OS

Parameter	NG$_{35\%}$ Hazard ratio	P	ORG$_{35\%}$ Hazard ratio	P	NG$_{100\%}$ Hazard ratio	P
Primary-tumor metabolic volume	1.009 (1.003–1.016)	0.004*	1.010 (1.003–1.016)	0.003*	1.009 (1.003–1.015)	0.004*
Primary-tumor SUV$_{mean}$	1.00 (0.95–1.06)	0.9	1.00 (0.94–1.05)	0.8	1.00 (0.94–1.06)	0.9
Primary-tumor TLG	1.08 (1.02–1.14)	0.004*	1.08 (1.03–1.13)	0.003*	1.08 (1.03–1.14)	0.003*
Summed TLG	1.09 (1.03–1.14)	0.002*	1.08 (1.03–1.14)	0.002*	1.08 (1.03–1.14)	0.002*
Entropy	0.3 (0.2–1.8)	0.3	0.4 (0.1–1.2)	0.1	0.7 (0.3–2.0)	0.5
Dissimilarity	0.9 (0.70–1.1)	0.2	0.9 (0.7–1.0)	0.09	0.9 (0.8–1.1)	0.2
Zone percentage	0.01 (0–16)	0.2	0.002 (0–1.5)	0.07	0.3 (0–79)	0.7
High-intensity emphasis	0.998 (0.996–1.00)	0.1	0.998 (0.996–1.00)	0.06	0.998 (0.996–1.00)	0.07

*Statistically significant.
Data in parentheses are 95% confidence intervals. Hazard ratios are per 100-unit change for TLG and per 1-unit change for other parameters.

demonstrating more blurring than mediastinal lesions. These results were supported by the tumor-motion analysis, in which lower-lobe lesions demonstrated the largest displacements.

Although respiratory motion has been shown to have a considerable impact on quantification of SUV$_{mean}$ and metabolic tumor volume (12), the effect was limited in our patient population. Optimal respiratory gating did not significantly influence quantification of TLG, possibly because of preselection of only lesions larger than 3 cm^3 for the purpose of feature extraction. Therefore, there were only few lesions in the lower lobes available for analysis. Given that blurring has the greatest effect on lesions

in the lower lobes, this asymmetric distribution could have caused an underestimation of the effect of blurring. In addition, the increase in SUV$_{mean}$ in the ORG$_{35\%}$ images might be cancelled by the reduction in metabolic tumor volume, making TLG a parameter that is more robust in the presence of respiratory motion artifacts.

One limitation of the current study was the relatively small patient cohort that could be analyzed. This limitation, in combination with the retrospective character of the study, limits the possibility of identifying which parameters are truly associated with OS. Identification and validation of the image parameters associated with patient outcome and OS require a multicenter

TABLE 6
Multivariate Cox Regression Models for OS

Fixed input	NG$_{35\%}$ Hazard ratio	P	ORG$_{35\%}$ Hazard ratio	P	NG$_{100\%}$ Hazard ratio	P
Disease stage		0.1		0.1		0.1
I	1		1		1	
II	0.2 (0.03–1.1)		0.2 (0.03–1.1)		0.2 (0.04–1.2)	
III	1.7 (0.4–8.6)		1.4 (0.3–7.2)		1.8 (0.4–8.9)	
IV	1.8 (0.3–12)		1.7 (0.2–12)		1.9 (0.3–13)	
Treatment		<0.0001*		<0.0001*		<0.0001*
Surgery	1		1		1	
Chemoradiotherapy	0.3 (0.1–0.6)		0.3 (0.2–0.7)		0.4 (0.1–1.8)	
Chemotherapy	1.1 (0.5–2.4)		1.1 (0.5–2.5)		1.6 (0.3–8.3)	
Expectative (none)	5.2 (2.2–12)		4.5 (2.0–10)		7.7 (1.5–40)	
Iterative backward selection						
Summed TLG	1.07 (1.01–1.13)	0.02*	1.08 (1.02–1.14)	0.009*	1.07 (1.01–1.13)	0.04*
High-intensity emphasis	0.997 (0.995–0.999)	0.02*	0.997 (0.995–0.999)	0.02*	0.997 (0.995–0.999)	0.02*

*Statistically significant.
Data in parentheses are 95% confidence intervals. Hazard ratios are per 100-unit change for TLG and per 1-unit change for other parameters.

prospective study. Although these textural parameters were found to be associated with OS in a previous study (*8*), such was not our finding in the current patient cohort. This difference could be due to our relatively large number of patients with metastatic disease, which was an exclusion criterion in the previous study (*8*). In patients with metastatic disease, the characteristics of the primary tumor might have limited predictive value regarding OS, whereas indices containing more information about disease load (such as summed TLG) might more accurately predict OS.

The obtained multivariate models for OS were consistent over all 3 image types, suggesting that the accuracy of the calculated parameters is robust in the presence of respiratory motion artifacts and varying noise levels. In these multivariate models, TLG and high-intensity emphasis were the only independent image-derived covariates that were relevant for the clinical model, even after correction for treatment and disease stage.

Although we have studied the influence of important sources of measurement variability on quantification of intratumor heterogeneity, several other possible sources remain. Indeed, measurement variability in PET can stem from a myriad of factors ranging from acquisition and reconstruction settings to patient physiology (*22*). There have been reports of variability in ^{18}F-FDG PET radiomics parameters due to test–retest variability (*18,23*). In a study by Leijenaar et al., the investigated radiomics parameters showed high stability in an interobserver and test–retest setup (*23*). Furthermore, Tixier et al. reported that several textural parameters derived from PET images showed a high degree of reproducibility as determined on double-baseline PET acquisitions (*18*). It is essential to assess the clinical impact of such measurement variations on the development and validation of new PET image–derived indices to characterize cancer lesions.

CONCLUSION

The results of this study suggest that the tested textural parameters are robust in the presence of respiratory motion artifacts and varying levels of image noise.

DISCLOSURE

REFERENCES

1. Grootjans W, de Geus-Oei L-F, Troost EGC, Visser EP, Oyen WJG, Bussink J. PET in the management of locally advanced and metastatic NSCLC. *Nat Rev Clin Oncol.* 2015;12:395–407.
2. Visvikis D, Hatt M, Tixier F, Le Rest CC. The age of reason for FDG PET image-derived indices. *Eur J Nucl Med Mol Imaging.* 2012;39:1670–1672.
3. Im H-J, Pak K, Cheon G, et al. Prognostic value of volumetric parameters of ^{18}F-FDG PET in non-small-cell lung cancer: a meta-analysis. *Eur J Nucl Med Mol Imaging.* 2015;42:241–251.
4. Parmar C, Leijenaar RTH, Grossmann P, et al. Radiomic feature clusters and prognostic signatures specific for lung and head & neck cancer. *Sci Rep.* 2015;5:11044.
5. Cook GJR, Yip C, Siddique M, et al. Are pretreatment ^{18}F-FDG PET tumor textural features in non–small cell lung cancer associated with response and survival after chemoradiotherapy? *J Nucl Med.* 2013;54:19–26.
6. Cook GJR, 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;276:883–893.
7. Pyka T, Bundschuh RA, Andratschke N, et al. Textural features in pre-treatment [F18]-FDG-PET/CT are correlated with risk of local recurrence and disease-specific survival in early stage NSCLC patients receiving primary stereotactic radiation therapy. *Radiat Oncol.* 2015;10:100.
8. Tixier F, Hatt M, Valla C, et al. Visual versus quantitative assessment of intratumor ^{18}F-FDG PET uptake heterogeneity: prognostic value in non–small cell lung cancer. *J Nucl Med.* 2014;55:1235–1241.
9. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144:646–674.
10. Fave X, Cook M, Frederick A, et al. Preliminary investigation into sources of uncertainty in quantitative imaging features. *Comput Med Imaging Graph.* 2015;44:54–61.
11. Leijenaar RTH, Nalbantov G, Carvalho S, et al. The effect of SUV discretization in quantitative FDG-PET radiomics: the need for standardized methodology in tumor texture analysis. *Sci Rep.* 2015;5:11075.
12. Grootjans W, de Geus-Oei L-F, Meeuwis AW, et al. Amplitude-based optimal respiratory gating in positron emission tomography in patients with primary lung cancer. *Eur Radiol.* 2014;24:3242–3250.
13. Yip S, McCall K, Aristophanous M, Chen AB, Aerts HJWL, Berbeco R. Comparison of texture features derived from static and respiratory-gated PET images in non-small cell lung cancer. *PLoS One.* 2014;9:e115510.
14. 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.
15. Hatt M, Majdoub M, Vallières M, et al. ^{18}F-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi–cancer site patient cohort. *J Nucl Med.* 2015;56:38–44.
16. Grootjans W, Hermsen R, van der Heijden EHFM, et al. The impact of respiratory gated positron emission tomography on clinical staging and management of patients with lung cancer. *Lung Cancer.* 2015;90:217–223.
17. Hatt M, Le Rest CC, Turzo A, Roux C, Visvikis D. A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. *IEEE Trans Med Imaging.* 2009;28:881–893.
18. Tixier F, Hatt M, Le Rest CC, Le Pogam A, Corcos L, Visvikis D. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in ^{18}F-FDG PET. *J Nucl Med.* 2012;53:693–700.
19. Schuurbiers OCJ, Meijer TWH, Kaanders JHAM, et al. Glucose metabolism in NSCLC is histology-specific and diverges the prognostic potential of ^{18}FDG-PET for adenocarcinoma and squamous cell carcinoma. *J Thorac Oncol.* 2014;9:1485–1493.
20. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. *J Clin Epidemiol.* 1995;48:1503–1510.
21. Oliver JA, Budzevich M, Zhang GG, Dilling TJ, Latifi K, Moros EG. Variability of image features computed from conventional and respiratory-gated PET/CT images of lung cancer. *Transl Oncol.* 2015;8:524–534.
22. Boellaard R. Standards for PET image acquisition and quantitative data analysis. *J Nucl Med.* 2009;50(suppl):11S–20S.
23. Leijenaar RTH, Carvalho S, Velazquez ER, et al. Stability of FDG-PET radiomics features: an integrated analysis of test-retest and inter-observer variability. *Acta Oncol.* 2013;52:1391–1397.