

Intratumor Heterogeneity Characterized by Textural Features on Baseline ^{18}F -FDG PET Images Predicts Response to Concomitant Radiochemotherapy in Esophageal Cancer

Florent Tixier¹, Catherine Cheze Le Rest^{1,2}, Mathieu Hatt¹, Nidal Albarghach^{1,3}, Olivier Pradier^{1,3}, Jean-Philippe Metges^{3,4}, Laurent Corcos⁴, and Dimitris Visvikis¹

¹INSERM, U650, LaTIM, CHU Morvan, Brest, France; ²Department of Nuclear Medicine, CHU Morvan, Brest, France; ³Institute of Oncology, CHU Morvan, Brest, France; and ⁴INSERM, U613, Faculty of Medicine, Brest, France

^{18}F -FDG PET is often used in clinical routine for diagnosis, staging, and response to therapy assessment or prediction. The standardized uptake value (SUV) in the primary or regional area is the most common quantitative measurement derived from PET images used for those purposes. The aim of this study was to propose and evaluate new parameters obtained by textural analysis of baseline PET scans for the prediction of therapy response in esophageal cancer. **Methods:** Forty-one patients with newly diagnosed esophageal cancer treated with combined radiochemotherapy were included in this study. All patients underwent pretreatment whole-body ^{18}F -FDG PET. Patients were treated with radiotherapy and alkylating agents (5-fluorouracil-cisplatin or 5-fluorouracil-carboplatin). Patients were classified as nonresponders (progressive or stable disease), partial responders, or complete responders according to the Response Evaluation Criteria in Solid Tumors. Different image-derived indices obtained from the pretreatment PET tumor images were considered. These included usual indices such as maximum SUV, peak SUV, and mean SUV and a total of 38 features (such as entropy, size, and magnitude of local and global heterogeneous and homogeneous tumor regions) extracted from the 5 different textures considered. The capacity of each parameter to classify patients with respect to response to therapy was assessed using the Kruskal–Wallis test ($P < 0.05$). Specificity and sensitivity (including 95% confidence intervals) for each of the studied parameters were derived using receiver-operating-characteristic curves. **Results:** Relationships between pairs of voxels, characterizing local tumor metabolic nonuniformities, were able to significantly differentiate all 3 patient groups ($P < 0.0006$). Regional measures of tumor characteristics, such as size of nonuniform metabolic regions and corresponding intensity nonuniformities within these regions, were also significant factors for prediction of response to therapy ($P = 0.0002$). Receiver-operating-characteristic curve analysis showed that tumor textural analysis can provide nonresponder, partial-responder, and complete-responder patient identification with higher sensitivity (76%–92%) than any SUV measurement. **Conclusion:** Textural features of tumor metabolic distribution extracted from baseline ^{18}F -FDG PET images allow for the best

stratification of esophageal carcinoma patients in the context of therapy-response prediction.

Key Words: ^{18}F -FDG PET; esophageal cancer; textural analysis; predictive value; response to therapy

J Nucl Med 2011; 52:369–378

DOI: 10.2967/jnumed.110.082404

Esophageal cancer is associated with high mortality. In patients with early-stage disease at presentation, esophagectomy is the treatment of choice and is potentially curative. Unfortunately most patients at presentation have already locally advanced esophageal cancer or distant metastases. In locally advanced esophageal cancer, preoperative chemotherapy or radiochemotherapy will improve survival in patients who respond to induction therapy (1,2). On the other hand, patients who do not respond to neoadjuvant therapy may be affected unnecessarily by the toxicity of an inefficient therapy. Therefore, the development of a diagnostic test capable of noninvasively predicting response to therapy early in the course of treatment is of great interest, potentially allowing personalization of patient management. In patients treated by exclusive conventional combined radiochemotherapy, assessment of response is equally of great interest, because it could allow an early change in the management of nonresponding patients. Such assessment becomes more critical when one considers the availability of new targeted therapies that could be tested with higher efficiency if applied early in diagnosis (3,4).

^{18}F -FDG PET is already well established for the initial staging of esophageal cancer, because it is associated with a better sensitivity and specificity than combined use of CT and echoendoscopy, especially regarding detection of distant metastasis (5).

^{18}F -FDG PET has been also used to assess response to therapy and patient outcome prognosis (4,6). Within this context, few studies have explored the potential prognostic value of pretreatment ^{18}F -FDG PET, demonstrating that the level of activity concentration on preoperative PET,

Received Aug. 17, 2010; revision accepted Dec. 13, 2010.
For correspondence or reprints contact: Florent Tixier, LaTIM, INSERM U650, CHU Morvan, 5 Avenue Foch, 29609 Brest, France.
E-mail: florent.tixier@etudiant.univ-brest.fr
COPYRIGHT © 2011 by the Society of Nuclear Medicine, Inc.

although not statistically significant, tends to predict overall survival (7–9).

On the other hand, several studies have evaluated the role of PET in assessing treatment response based on ^{18}F -FDG uptake changes between a pre- and a posttreatment PET scan obtained during or after the treatment completion. Studies considering a second PET scan after treatment completion have shown that a complete metabolic response is associated with better outcome (10–12). However, because that information is of limited interest in patient management if acquired late, different attempts have been made to determine whether ^{18}F -FDG PET could be used for assessing response to therapy earlier (usually within a few weeks) in the course of treatment (13–15), showing some promising results that need to be confirmed in multicenter trials (4). One of the highlighted issues is that early response prediction during combined chemoradiotherapy, in contrast to chemotherapy alone, may be compromised by increased ^{18}F -FDG tumor uptake resulting from radiotherapy-induced inflammatory processes (4).

An alternative to monitoring changes during treatment is the potential of predicting response to therapy from the baseline ^{18}F -FDG PET scan alone, which may allow the best available therapy regime to be chosen for a given patient. However, to date there is only limited evidence that a measure of tumor activity concentration on a baseline PET scan in esophageal cancer can differentiate groups of patient response (8,9). Within the same context, parameters derived from pretreatment ^{18}F -FDG PET have shown the potential to differentiate between responders and nonresponders (NRs) in non-Hodgkin lymphoma patients (16).

The PET image index predominantly used in such studies for assessment of metabolic response is the normalized mean tumor activity concentration known as the mean standardized uptake value (SUV_{mean}), within a region of interest around the tumor, or the maximum standardized uptake value corresponding to the highest-activity pixel value (SUV_{max}). However, ^{18}F -FDG tumor uptake has been associated not only with increased metabolism but also with several other physiologic parameters such as perfusion, cell proliferation (17), tumor viability, aggressiveness, or hypoxia (18,19), all of which may in turn be responsible for tumor uptake heterogeneity. Therefore, the hypothesis can be made that characterizing tumor ^{18}F -FDG distribution, through its relationship to underlying tumor biologic characteristics, may be useful in predicting therapy response. ^{18}F -FDG tumor activity distribution may be assessed in a global, regional, or local fashion, allowing in turn the assessment of corresponding global, regional, or local patterns of biologic heterogeneity. Although the measurement of such features have been previously explored in anatomic imaging (20–22), they have not to date been widely used in PET. Until now, only 1 study has considered the use of some textural features to predict treatment outcome from

baseline ^{18}F -FDG PET images, with encouraging results in cervical and head and neck cancer (23), and the assessment of spatial heterogeneity was also shown to be significantly associated with survival in sarcoma patients (24). However, the potential predictive value of tumor heterogeneity characterization on a baseline ^{18}F -FDG PET scan has never been assessed.

The objective of this current study was, therefore, to assess the predictive value of ^{18}F -FDG uptake heterogeneity characterized by textural features extracted from pretherapy ^{18}F -FDG PET images of patients with esophageal carcinoma by assessing the ability of each parameter to identify different categories of responders. The predictive value of these parameters was compared with the use of standard image activity concentration indices (SUV_{max} , SUV_{mean}). The potential prognostic value of such image-derived parameters for assessing overall patient survival was not assessed in this study.

MATERIALS AND METHODS

Patients

Forty-one patients with newly diagnosed esophageal cancer treated with exclusive radiochemotherapy between 2003 and 2008 were included in this study. The characteristics of the patients are summarized in Table 1. The mean age at the time of diagnosis was 66 ± 10 y (median, 69 y; range, 45–84 y), and 85% of patients were male. Most of the tumors were squamous cell carcinoma (76%), and most of the patients had a well or moderately differentiated tumor (56%). Most of the tumors originated from the middle and lower esophagus (76%). Twenty-six patients had a T3 or T4 primary lesion, 25 had N1 (61%) lymph node metastases, and 17 had distant metastases (Table 1). All patients were treated with external-beam radiotherapy and chemotherapy with alkylating agents (5-fluorouracil-cisplatin or 5-fluorouracil-carboplatin). A median radiation dose of 60 Gy was delivered in 180-cGy daily fractions (5 d/wk and 6–7 wk in total). One month after the completion of the treatment, patients were reassessed to determine response to therapy using thoracoabdominal CT and endoscopy. Patients were subsequently classified as complete responders (CR), partial responders (PR), stable disease, or progressive disease. Response was assessed using pretreatment and posttreatment CT scans by evaluating the increase (or decrease) in the sum of the longest diameters for all target lesions and the appearance, persistence, or disappearance of nontarget lesions, according to the Response Evaluation Criteria in Solid Tumors (RECIST) (25). Considering the small number of patients in the stable disease (7) and progressive disease (4) groups, these patients were eventually combined into an NR group.

All patients underwent pretreatment whole-body ^{18}F -FDG PET for staging purposes. Patients were instructed to fast for a minimum of 6 h before the injection of ^{18}F -FDG. The dose of administered ^{18}F -FDG was 5 MBq/kg, and static emission images were acquired from thigh to head, on average 54 min after injection, on a Gemini PET/CT scanner (Philips). In addition to the emission PET scan, a low-dose CT scan was acquired for attenuation-correction purposes. Images were reconstructed with the 3-dimensional (3D) row-action maximum-likelihood algorithm using standard clinical protocol parameters (2 iterations, relaxation parameter of 0.05, and 3D gaussian postfiltering of 5 mm in

[Table 1]

TABLE 1
Characteristics of Patients ($n = 41$)

Characteristic	No. of patients
Sex	
Male	35 (85)
Female	6 (15)
Primary site	
Upper esophagus	10 (24)
Middle esophagus	15 (37)
Lower esophagus	16 (39)
Tumor cell type	
Squamous cell carcinoma	31 (76)
Adenocarcinoma	10 (24)
Histologic grade	
Well differentiated	12 (29)
Moderately differentiated	11 (27)
Poorly differentiated	3 (7)
Unknown	15 (37)
TNM stage	
T1	6 (15)
T2	7 (17)
T3	21 (51)
T4	7 (17)
N0	16 (39)
N1	25 (61)
M0	24 (59)
M1	17 (41)
AJCC stage	
I	4 (10)
IIa	6 (15)
IIb	5 (12)
III	12 (29)
IVa	4 (10)
IVb	10 (24)
RECIST	
CR	9 (22)
PR	21 (51)
Stable disease (NR)	7 (17)
Progressive disease (NR)	4 (10)

Data in parentheses are percentages.

full width at half maximum). The current data analysis was performed after approval by the institutional review board.

Tumor Analysis

For each patient, primary tumors were identified on ^{18}F -FDG PET images by an experienced nuclear physician. Tumors were then delineated automatically using the previously validated fuzzy locally adaptive Bayesian algorithm (26). All parameters were subsequently extracted from this delineated volume. Only the primary tumors were considered because textural analysis cannot be reliably performed on small lesions (nodal or distant metastases) because of the small number of voxels involved.

Standardized Uptake Value (SUV) Analysis

The following SUV parameters were extracted from each patient's baseline PET images: SUV_{\max} ; peak SUV (SUV_{peak}), defined as the mean of the voxel of maximum value and its 26 neighbors (in 3 dimensions); and mean SUV within the delineated tumor (SUV_{mean}). The SUV_{peak} was considered in addition to SUV_{\max} to investigate the impact of reducing the potential bias in the SUV_{\max} measurements as a result of its sensitivity to noise.

Textural Analysis

We define texture as a spatial arrangement of a predefined number of voxels allowing the extraction of complex image properties, and we define a textural feature as a measurement computed using a texture matrix. The method used was realized in 2 steps. First, matrices describing textures on images were extracted from tumors, and textural features were subsequently computed using these matrices. All these parameters characterize in some way tumor heterogeneity at local and regional (using texture matrices) or global scales (using image-voxel-intensity histograms).

Several different textures (Table 2, left column) were computed. Voxel values within the segmented tumors (Fig. 1A and 1B) were resampled to yield a finite range of values (Fig. 1C), allowing textural analysis using: [Table 2] [Fig. 1]

$$V(x) = \left[2^s \frac{I(x) - \min_{i \in \Omega} i}{\max_{i \in \Omega} i - \min_{i \in \Omega} i + 1} \right] \quad \text{Eq. 1}$$

where 2^s represents the number of discrete values (16–128), I is the intensity of the original image, and Ω is the set of voxels in the delineated volume. This resampling step on the delineated tumor volume, necessary for the computation of the textural analysis, has 2 effects: it reduces the noise in the image by clustering voxels with similar intensities and it normalizes the tumor voxel intensities across patients, which in turn facilitates the comparison of the textural features. Local and regional features were computed with different resampling considering 16, 32, 64, and 128 discrete values to investigate the potential impact of this resampling parameter.

All considered textures were originally described for 2 dimensions (27–30) and were therefore adapted in this work for 3 dimensions. The cooccurrence matrix (M1, Fig. 1D(a)) describing pairwise arrangement of voxels, and the matrix describing the alignment of voxels with the same intensity (M2, Fig. 1D(b)), were computed considering 13 different angular directions. Finally, 3D matrices describing differences between each voxel and its neighbors (M3, Fig. 1D(c)) and characteristics of homogeneous zones (M4, Fig. 1D(d)) were computed considering for each voxel the neighbors in the 2 adjacent planes, adapting the normalizing factors to 3 dimensions.

From each of the extracted texture matrices, different features summarized in Table 2 (middle column) were computed. Depending on the way the matrix is analyzed, it is possible to extract features of a local or regional nature. Six features highlighting local variations of voxel intensities within the image were extracted from the cooccurrence matrices M1 (Fig. 2C). For example, using the matrix M1, the local entropy and homogeneity are calculated using Equations 2 and 3, respectively: [Fig. 2]

$$\text{Local entropy} = -\sum_{i,j} M1(i,j) \log(M(i,j)) \quad \text{Eq. 2}$$

$$\text{Local homogeneity} = \sum_{i,j} \frac{M1(i,j)}{1+|i-j|} \quad \text{Eq. 3}$$

where M1 is a cooccurrence matrix, i, j are the rows and columns index, and $M1(i,j)$ is an element of the matrix.

In addition, M3 matrices were used to extract busyness (quantifying sharp-intensity variations) and contrast and coarse-

TABLE 2
Texture Type and Associated Features

Type	Feature	Scale
Features based on intensity histogram	Minimum intensity Maximum intensity Mean intensity Variance SD Skewness Kurtosis	Global
Features based on voxel-alignment matrix (M2)	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	Regional
Features based on intensity-size-zone matrix (M4)	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	Regional
Features based on cooccurrence matrices (M1)	Second angular moment Contrast (inertia) Entropy Correlation Homogeneity Dissimilarity	Local
Features based on neighborhood intensity-difference matrix (M3)	Coarseness Contrast Busyness	Local

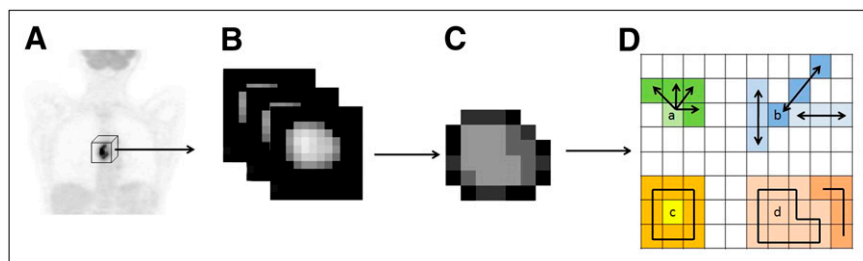
ness (quantifying tumor granularity). These features allow extracting measurements describing tumor local heterogeneity proportional to variations of ^{18}F -FDG uptake between individual voxels.

On the other hand, the M2 and M4 matrices were used to extract regional tumor uptake characteristics, representing regional heterogeneity, such as variation of intensity between regions and in the size and alignment of homogeneous areas. For example, the M4 matrix links the homogeneous tumor regions to their intensity (Fig.

2B). It was hence used to calculate the variability in the size and the intensity of identified homogeneous tumor zones according to Equations 4 and 5, respectively:

$$\text{Size-zone variability} = \frac{1}{\Theta} \sum_{m=1}^M \left[\sum_{n=1}^N M4(m, n) \right]^2 \quad \text{Eq. 4}$$

FIGURE 1. Whole-body ^{18}F -FDG PET scan (A), tumor segmentation (B), and voxel-intensity resampling (C) allowing extraction of different features (D) by analysis of consecutive voxels in a direction (for cooccurrence matrices) (a), alignment of voxels with same intensity (b), difference between voxels and their neighbors (c), and zones of voxels with same intensity (d).



RGB

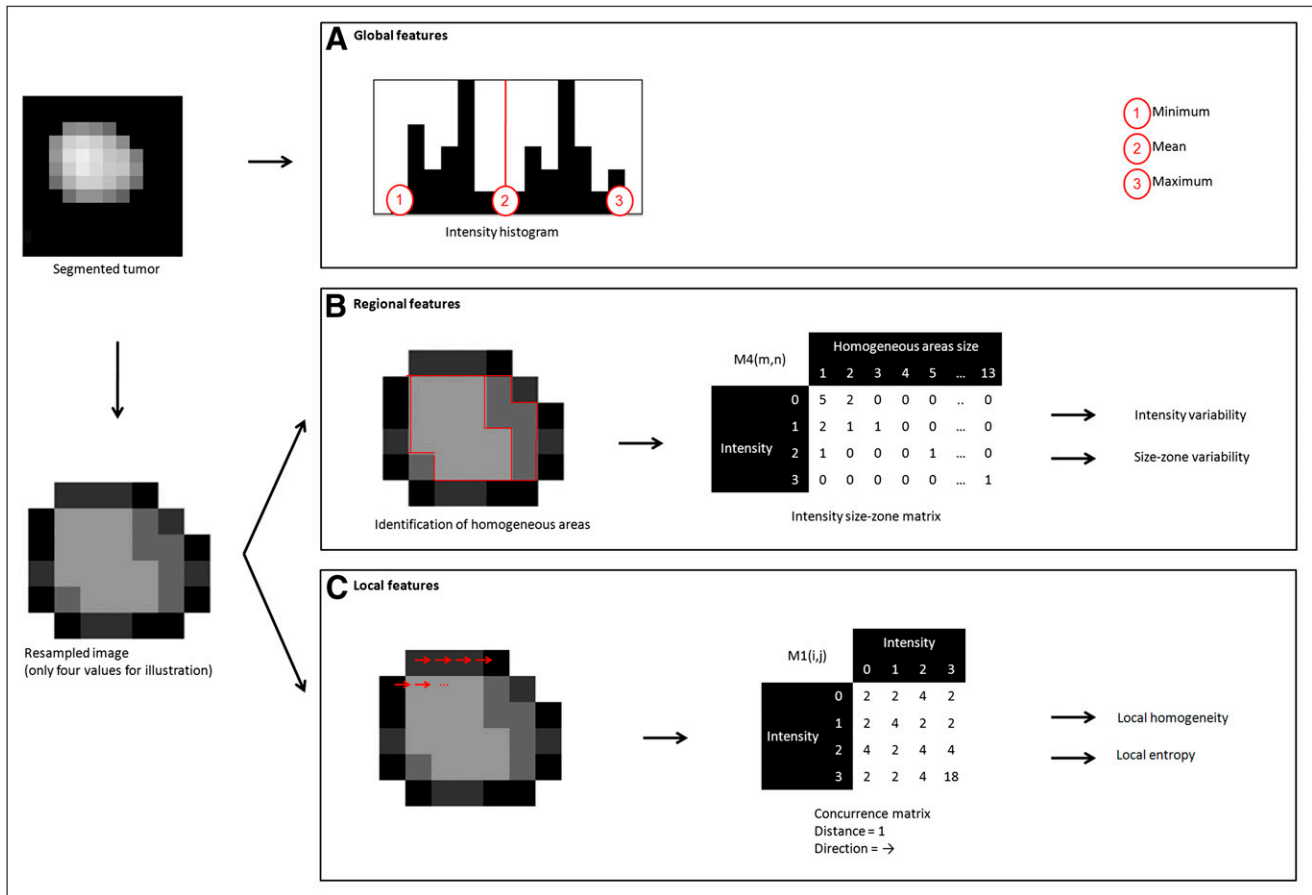


FIGURE 2. Examples of features extracted from tumor resampled on 4 values: 3 global features computed using intensity histogram, 2 regional features computed using M4 matrix, and 2 local features computed using M1 texture matrices.

$$\text{Intensity variability} = \frac{1}{\Theta} \sum_{n=1}^N \left[\sum_{m=1}^M M4(m,n) \right]^2, \quad \text{Eq. 5}$$

where Θ represents the number of homogeneous areas in the resampled tumor, M the number of distinct intensity values within the tumor, and N the size of the largest homogeneous area in the matrix $M4$.

Finally, global features are computed on the original image voxels' intensity distribution by analyzing the characteristics of the intensity value histogram within the segmented tumor (Fig. 2A).

Thirty-eight features were extracted from the 4 different texture matrices and intensity histograms. Seven of the 38 features characterize the uptake distribution within the entire tumor (using the intensity histogram), 9 describe local voxel arrangements (using matrices $M1$ and $M3$), and 22 are related to the organization of voxels at a regional scale (using matrices $M2$ and $M4$).

Statistical Analysis

The capacity of each feature to classify patients with respect to therapy response was investigated on the primary tumor using the Kruskal–Wallis test (8). P values of less than 0.05 were considered statistically significant. Specificity and sensitivity (including 95% confidence intervals [CIs]) for each of the studied parameters were derived using receiver operating characteristic (ROC) curves measuring associated areas under the ROC curves (AUC). Texture results

were compared with those of SUV_{\max} , SUV_{mean} , and SUV_{peak} for their ability to distinguish among responders (PR and CR) and NRs, CRs and non-CRs (PR, NR), and all 3 groups separately.

RESULTS

Patients were evaluated 1 mo after the completion of combined radiochemotherapy. Nine patients (22%) had no evidence of disease after treatment and were considered CRs. Radiochemotherapy led to partial response in 21 (51%) patients, whereas 11 (27%) were stable or progressed under treatment according to RECIST (25).

Results of the Kruskal–Wallis test show that SUV_{\max} (Fig. 3) and SUV_{mean} were capable of differentiating only [Fig. 3] CRs from NRs and PRs. Within this context, all SUV measurements were significant predictive factors of response ($P = 0.034$, 0.044 , and 0.012 for the SUV_{\max} , SUV_{mean} , and SUV_{peak} , respectively). However, only SUV_{peak} was a significant predictive factor ($P = 0.045$) when considering the differentiation of 3 patient response groups (i.e., NR, PR, and CR), whereas SUV_{\max} and SUV_{mean} were not ($P > 0.05$).

Figure 4 shows examples of different extracted features [Fig. 4] and associated values for tumors of CRs, PRs, and NRs. The Kruskal–Wallis tests revealed no statistically significant differences in the textural parameters derived using

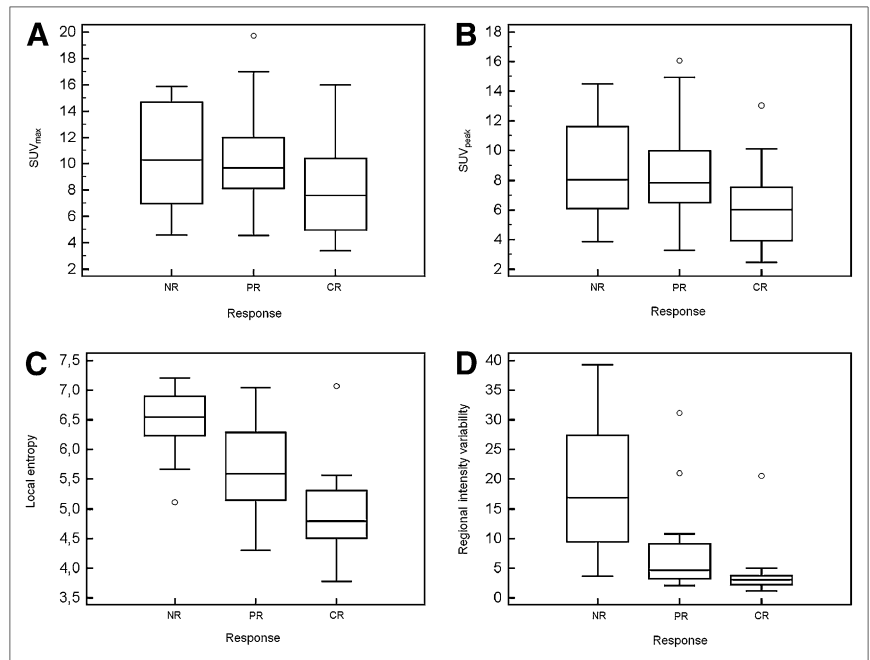


FIGURE 3. Box-plot representation of parameters' values in function of patient response (0, NR; 1, PR; and 2, CR) for SUV_{max} ($P = 0.106$) (A), SUV_{peak} ($P = 0.045$) (B), local entropy ($P = 0.0006$) (C), and regional intensity variability ($P = 0.0002$) (D).

different resampling values (16, 32, 64, or 128 discrete values). All subsequent reported results were obtained using 64 discrete values in the resampling normalization process. This value was chosen because it allows for 0.25 SUV increments, which were considered sufficient given the range of SUVs encountered (from ~ 4 –20).

None of the global features extracted from the intensity histogram within the tumor was a significant predictive factor of response to therapy. However, considering local variation of ^{18}F -FDG uptake, a high predictive value ($P < 0.0007$) was found using the cooccurrence features, particularly considering the use of the average feature values

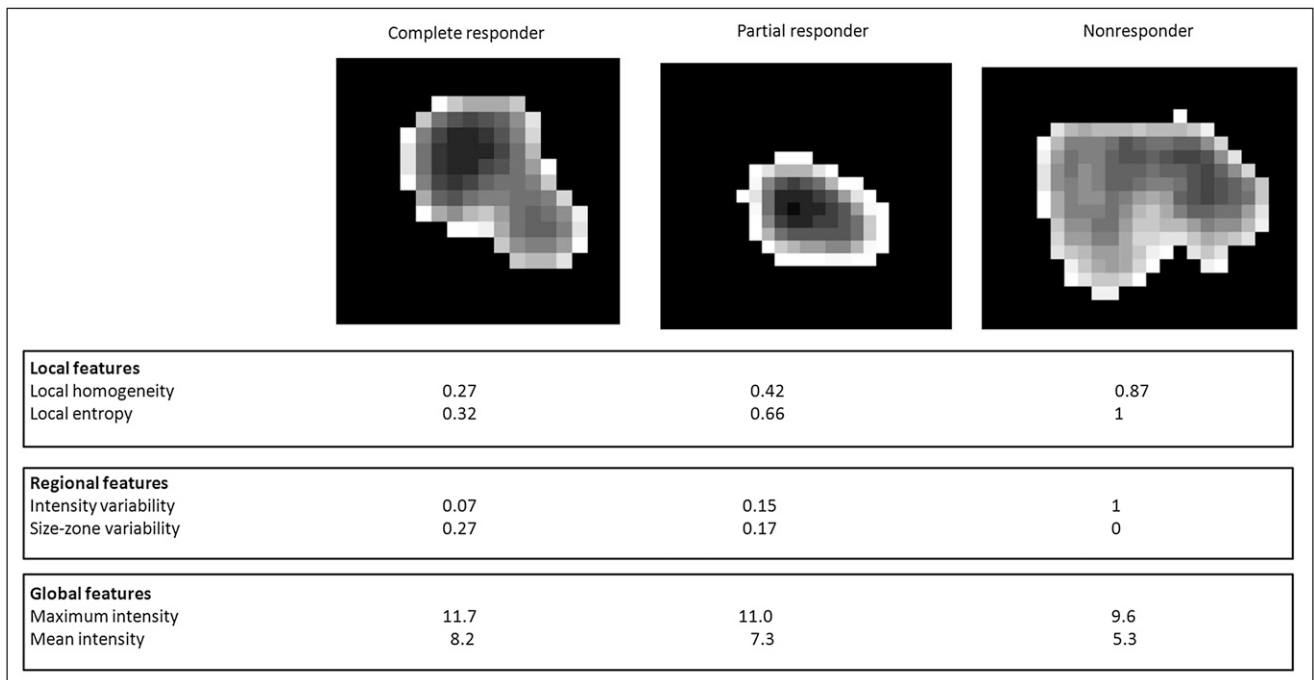


FIGURE 4. Example of different extracted features and associated values for tumors of CRs, PRs, and NRs (results are normalized in [0–1] interval using range of observed values for local and regional parameters).

[Table 3] computed using M1 matrices (Table 3). All these features offered statistically significant differentiation of NRs and responders (considering both CR and PR).

Regarding local features, the busyness and contrast computed on M3 matrices were not statistically significant predictive factors of response, but the coarseness, reflecting the local granularity of the tumor functional image, was found to be significant ($P = 0.0002$). Among the local measures of functional tumor characteristics computed using M1 matrices, the measure of local entropy was the only measure allowing statistically significant differentiation of all 3 patient groups ($P = 0.0006$, Fig. 3).

Because the features computed on M2 and M4 matrices, used to highlight regional variability in the ^{18}F -FDG distribution, were strongly correlated ($r > 0.9$), only features based on M4 were used in the subsequent analysis. Regional measures of tumor characteristics extracted from these M4 matrices, such as the variability in the size and the intensity of identified homogeneous tumor zones, were statistically significant in predicting therapy response ($P = 0.0002$), allowing the differentiation of all 3 patient response groups (Fig. 3).

The ROC curve analysis for SUV_{max} , SUV_{peak} , local homogeneity, local entropy, and regional tumor characteristics such as the variability in size and intensity of identified homogeneous tumor areas is presented in Figure 5. Table 3 summarizes the ROC curve analysis results, comparing the performance of the different studied parameters in terms of sensitivity and specificity in, on the one hand, identifying complete-response patients and, on the other hand, differentiating responders (PR and CR).

First, based on the ROC curve analysis, textural parameters can identify CRs better than can the SUV-based measurements, as demonstrated by the respective AUCs

(Fig. 5). For example, SUV_{max} , with an AUC of 0.7, allowed the identification of CRs, with a maximum sensitivity of 46% and specificity of 91%, using a threshold of 6. On the other hand, the variability in the size of the uniform tumor zones (AUC, 0.85) allowed for the extraction of CR patients with the best accuracy (sensitivity, 92%; specificity, 69%).

Second, as Figure 5 shows, textural features were most efficient in identifying responders (CRs and PRs), whereas for the same task the performance of SUV measurements was limited. For the differentiation of the patient responders, the AUC was less than 0.6 for the different SUV parameters, compared with an AUC of more than 0.82 for the use of the texture parameters. For example, the AUC of the SUV_{max} was 0.59, allowing a sensitivity of 53% and specificity of 73% in the differentiation of responders using an optimal threshold of 9.1. On the other hand, for the same task the local homogeneity had a specificity and sensitivity of 88% and 73%, respectively (AUC, 0.89).

DISCUSSION

Assessment of tumor response to therapy plays a central role in drug development and patient clinical management. Currently, response is mainly assessed by measuring anatomic tumor size and classifying tumor shrinkage according to standard criteria. Because metabolic changes often occur before morphologic changes, metabolic imaging appears to be a valuable tool for monitoring various treatments in different cancer types. Within this context, ^{18}F -FDG PET has shown promising results in assessing response to therapy and prognosis. In esophageal cancer, quantitative changes in ^{18}F -FDG uptake at 2 wk after the start of therapy have been shown to correlate well with

[Fig. 5]

TABLE 3
Sensitivity and Specificity (Along with Corresponding 95% Confidence Intervals) of 3 SUV-Based Measurements, 2 Cooccurrence Features, and 2 Size-Zone Features

Comparison	Parameters	Sensitivity (%)	95% confidence interval (%)	Specificity (%)	95% confidence interval (%)
NR vs. PR + CR	SUV_{max}	53	35.1–70.2	73	39.0–94.0
	SUV_{mean}	71	52.5–84.9	45	16.7–76.6
	SUV_{peak}	56	37.9–72.8	73	39.0–94.0
	Local homogeneity	88	71.8–96.6	73	39.0–94.0
	Local entropy	79	61.1–91.0	82	48.2–97.7
	Size-zone	76	58.8–89.8	91	58.7–99.8
	Intensity variability	76	58.7–89.3	91	58.7–99.8
NR + PR vs. CR	SUV_{max}	46	19.2–74.9	91	75.0–98.0
	SUV_{mean}	62	31.6–86.1	81	63.6–92.8
	SUV_{peak}	62	31.6–86.1	81	63.6–92.8
	Local homogeneity	92	61.5–99.8	56	37.7–73.6
	Local entropy	92	61.5–99.8	69	50.0–83.9
	Size-zone	92	64.0–99.8	69	50.0–83.9
	Intensity variability	85	54.6–98.1	75	56.6–88.5

Data in top part of table are evaluation of parameters to distinguish PR or CR; data on bottom part of table are evaluation of parameters to differentiate CRs.

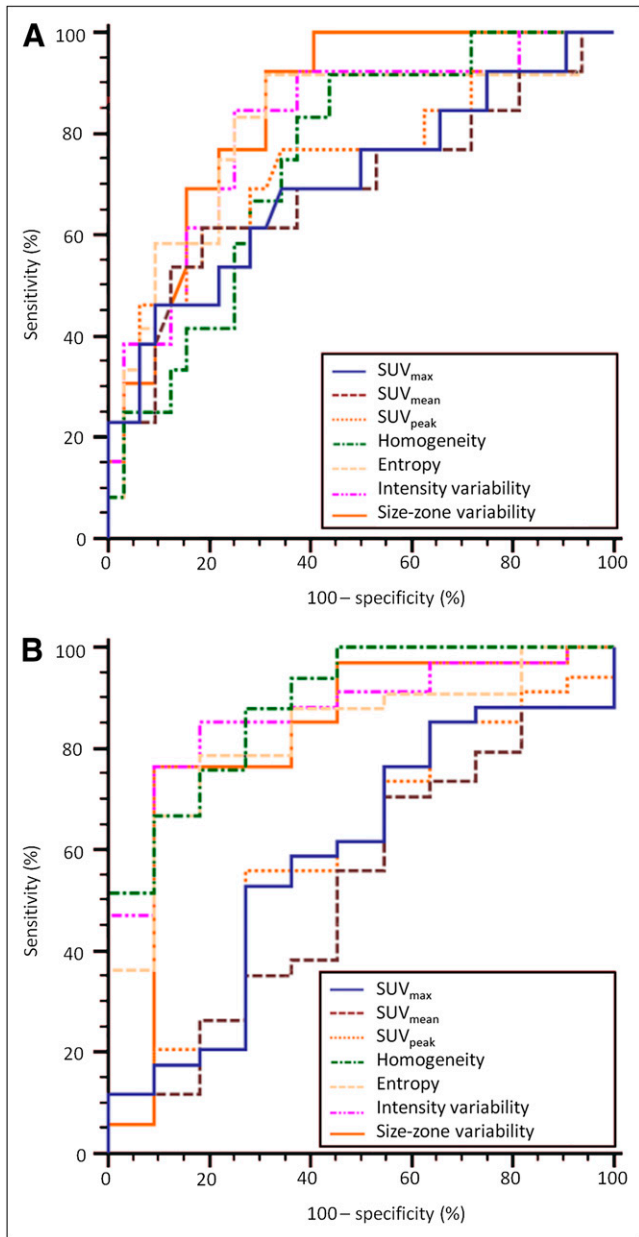


FIGURE 5. ROC curves for SUV_{max}, SUV_{mean}, SUV_{peak}, local homogeneity, uniform tumor areas, intensity variability, and size-zone variability for identification of CRs (A) and PRs or CRs (B).

subsequent tumor shrinkage and patient survival (4). This approach still has limitations, especially if patients undergo radiotherapy treatment. Hautzel et al. have shown that even low irradiation may enhance tumor uptake, and inflammatory changes may contribute early to this increase, yielding inaccurate information about treatment response (31). Within the same context, induced ulceration may also impair response assessment using PET (32).

On the other hand, the prediction of response before treatment initiation may be of great interest to the optimization of patient management. With such an endpoint, few authors have studied the predictive value of initial ¹⁸F-FDG uptake for therapy response. Rizk et al. reported an SUV_{max}

of more than 4.5 to be a reliable predictor of pathologic response (9), whereas Javeri et al. (8) demonstrated in a larger group of patients a trend of greater rate of response obtained after combined chemoradiotherapy in patients who had an initial SUV_{max} higher than 10. Similarly in our study, initial SUV_{mean}, SUV_{max}, and SUV_{peak} were also predictors of complete response. However, in general these indices did not allow differentiating NRs from PRs, a distinction that could be useful for patient management. For instance, within the patient population of our study the identification of PRs before any treatment could allow the definition of a subpopulation for which the use of conventional radiochemotherapy should be directly replaced by another option, such as a new targeted therapy.

A few studies have already focused on the link between image analysis and tumor biologic parameters. Gillies et al. (33) suggested that imaging can longitudinally characterize spatial variations in the tumor phenotype and its microenvironment so that the system dynamics over time can be quantitatively captured. Segal et al. (22) showed that contrast-enhanced CT image characteristics (such as texture heterogeneity score or estimated percentage of necrosis) correlate with most of the liver global gene expression profiles, revealing cell proliferation, liver synthetic function, and patient prognosis. Within the same context, Diehn et al. (34) mapped neuroimaging parameters with gene-expression patterns in glioblastoma, whereas Strauss et al. (35) combined dynamic PET kinetic parameters with gene-array techniques. Finally, Eary et al. (24) previously demonstrated that a globally assessed ¹⁸F-FDG distribution heterogeneity in sarcoma is a potential prognostic factor.

In our study, the value of textural feature analysis was explored on the pretreatment ¹⁸F-FDG PET scans for predicting response to combined chemoradiotherapy. Global tumor metabolic features based on the intensity histogram were computed directly on the original image. As such, they were therefore highly correlated with ¹⁸F-FDG uptake, which could explain why these textures could only predict CRs but could not distinguish NRs from PRs, similar to the SUV measurements. The other features evaluated in this study highlight tumor heterogeneity at a local and regional level, characterized in several ways, depending on the type of matrix used and the kind of feature computed on this matrix. Consequently, whereas a single feature cannot be directly linked to a specific biologic process, one could assume that a combination of textural parameters may be closely related to underlying physiologic processes, such as vascularization, perfusion, tumor aggressiveness, or hypoxia (18,19). Therefore textural features could be correlated to physiologic processes related to response to combined radiochemotherapy. For example, one could reasonably expect that a tumor exhibiting a heterogeneous, compared with a homogeneous, ¹⁸F-FDG distribution may respond less favorably to a uniformly distributed radiotherapy dose. We could also hypothesize that underlying neoangiogenesis contributes to tumor ¹⁸F-FDG uptake heterogeneity,

although it is now widely accepted that neoangiogenesis is associated with reduced effectiveness of conventional chemotherapy. However, the exact relationship between the proposed image-derived indices and underlying tumor biology can be established only on carefully designed prospective studies.

In this work, the cooccurrence features analyzing interrelationships between pairs of voxels, corresponding to the characterization of local nonuniformities, were able to significantly differentiate NRs from other patient groups. The measurement of local homogeneity and entropy gave the best results for this class of textures. Although in most cases responders (PR and CR) were associated with greater local heterogeneity than NRs, these features were less efficient in discriminating CRs from PRs.

The 2 features facilitating the best patient stratification were both associated with regional tumor characterization. Both the intensity and the size variability of uniform zones identified within the tumor, representing a measure of regional tumor heterogeneity, were significant predictors of response to therapy. ROC curve analysis showed that the performance of these features is similar to that of cooccurrence features in identifying NRs, but they can in addition distinguish between PRs and CRs with higher sensitivity and specificity than SUV measurements. These results suggest that regional (in terms of intensity and size of homogeneous areas) rather than local heterogeneity offers a superior differentiation of esophageal carcinoma patient groups in terms of response to combined chemoradiotherapy treatment than does any other global tumor metabolic activity measurement currently used in routine clinical practice, such as SUVs.

A limitation of the present study is that it is retrospective, considering a relatively small patient cohort. Therefore, the potential of new image-derived indices characterizing tumor ^{18}F -FDG distribution for prediction of response to therapy studies demonstrated in this work needs to be validated by a prospective study on a larger patient cohort.

CONCLUSION

We have demonstrated that textural analysis of the intratumor tracer uptake heterogeneity on baseline ^{18}F -FDG PET scans can predict response to combined chemoradiation treatment in esophageal cancer. Textural features derived from cooccurrence matrices strongly differentiated NRs from PRs, providing useful information for personalizing patient management. These results suggest that regional and local characterization of ^{18}F -FDG PET tracer heterogeneity in tumors, exploring processes underlying the ^{18}F -FDG uptake and distribution within tumors, are more powerful than global measurements currently used in clinical practice, holding the potential to revolutionize the predictive role of PET in cancer treatment. Finally, although only ^{18}F -FDG images in esophageal cancer have been considered here, clearly the same indices applied in other PET radiotracer studies in the same or different tumor types may

help create even stronger links between imaging and underlying tumor biology.

ACKNOWLEDGMENTS

This study was supported by a grant from the Ligue Contre le Cancer (Finistère and Côtes d'Armor Committees), IFR148-ScInBioS, and a fellowship from the French Ministry of Education and Research.

REFERENCES

1. GebSKI V, Burmeister B, Smithers BM, et al. Survival benefits from neoadjuvant chemoradiotherapy or chemotherapy in oesophageal carcinoma: a meta-analysis. *Lancet Oncol*. 2007;8:226–234.
2. Cunningham D, Allum WH, Stenning SP, et al. MAGIC trial participants: perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *N Engl J Med*. 2006;355:11–20.
3. Di Fabio F, Pinto C, Rojas Llimpe FL, et al. The predictive value of ^{18}F -FDG PET early evaluation in patients with metastatic gastric adenocarcinoma treated with chemotherapy plus cetuximab. *Gastric Cancer*. 2007;10:221–227.
4. Krause BJ, Herrmann K, Wieder H, zum Büschenfelde CM. ^{18}F -FDG PET and ^{18}F -FDG PET/CT for assessing response to therapy in esophageal cancer. *J Nucl Med*. 2009;50:89S–96S.
5. van Westreenen H, Westerterp M, Bossuyt P, et al. Systematic review of the staging performance of ^{18}F -fluorodeoxyglucose positron emission tomography in esophageal cancer. *J Clin Oncol*. 2004;22:3805–3812.
6. Ben-Haim S, Eil P. ^{18}F -FDG PET and PET/CT in the evaluation of cancer treatment response. *J Nucl Med*. 2009;50:88–99.
7. Rizk N, Downey RJ, Akhurst T, et al. Preoperative ^{18}F FDG positron emission tomography standardized uptake values predict survival after esophageal adenocarcinoma resection. *Ann Thorac Surg*. 2006;81:1076–1081.
8. Javeri H, Xiao L, Rohren E, et al. Influence of the baseline ^{18}F FDG positron emission tomography results on survival and pathologic response in patients with gastroesophageal cancer undergoing chemoradiation. *Cancer*. 2009;115:624–630.
9. Rizk NP, Tang L, Adusumilli PS, et al. Predictive value of initial PET SUVmax in patients with locally advanced esophageal and gastroesophageal junction adenocarcinoma. *J Thorac Oncol*. 2009;4:875–879.
10. Flamen P, Van Cutsem E, Lerut A, et al. Positron emission tomography for assessment of the response to induction radiochemotherapy in locally advanced esophageal cancer. *Ann Oncol*. 2002;13:361–368.
11. Downey RJ, Akhurst T, Ilson D, et al. Whole body ^{18}F FDG-PET and the response of esophageal cancer to induction therapy: results of a prospective trial. *J Clin Oncol*. 2003;21:428–432.
12. Kim MK, Ryu JS, Kim SB, et al. Value of complete metabolic response by ^{18}F -fluorodeoxyglucose-positron emission tomography in oesophageal cancer for prediction of pathologic response and survival after preoperative chemoradiotherapy. *Eur J Cancer*. 2007;43:1385–1391.
13. Weber WA, Ott K, Becker K, et al. Prediction of response to preoperative chemotherapy in adenocarcinomas of the esophagogastric junction by metabolic imaging. *J Clin Oncol*. 2001;19:3058–3065.
14. Lordick F, Ott K, Krause BJ, et al. PET to assess early metabolic response and to guide treatment of adenocarcinoma of the oesophagogastric junction: the MUNICON phase II trial. *Lancet Oncol*. 2007;8:797–805.
15. Ott K, Weber WA, Lordick F, et al. Metabolic imaging predicts response, survival, and recurrence in adenocarcinomas of the esophagogastric junction. *J Clin Oncol*. 2006;24:4692–4698.
16. Cazaentre T, Morschhauser F, Vermandel M, et al. Pre-therapy ^{18}F -FDG PET quantitative parameters help in predicting the response to radioimmunotherapy in non-Hodgkin lymphoma. *Eur J Nucl Med Mol Imaging*. 2010;37:494–504.
17. Vesselle H, Schmidt RA, Pugsley JM, et al. Lung cancer proliferation correlates with ^{18}F FDG uptake by positron emission tomography. *Clin Cancer Res*. 2000;6:3837–3844.
18. Rajendran JG, Schwartz DL, O'Sullivan J, et al. Tumour hypoxia imaging with ^{18}F fluoromisonidazole positron emission tomography in head and neck cancer. *Clin Cancer Res*. 2006;12:5435–5441.
19. Kunkel M, Reichert TE, Benz P, et al. Overexpression of Glut-1 and increased glucose metabolism in tumours are associated with a poor prognosis in patients with oral squamous cell carcinoma. *Cancer*. 2003;97:1015–1024.

20. Xu Y, Sonka M, McLennan G, Guo J, Hoffman EA. MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies. *IEEE Trans Med Imaging*. 2006;25:464–475.
21. Tesar L, Shimizu A, Smutek D, Kobatake H, Shigeru N. Medical image analysis of 3D CT images based on extension of Haralick texture features. *Comput Med Imaging Graph*. 2008;32:513–520.
22. Segal E, Sirlin CB, Ooi C, et al. Decoding global gene expression programs in liver cancer by noninvasive imaging. *Nat Biotechnol*. 2007;25:675–680.
23. El Naqa I, Grigsby PW, Aptea A, et al. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. *Pattern Recognit*. 2009;42:1162–1171.
24. Eary JF, O'Sullivan F, O'Sullivan J, Conrad EU. Spatial heterogeneity in sarcoma ¹⁸F-FDG uptake as a predictor of patient outcome. *J Nucl Med*. 2008;49:1973–1979.
25. Therasse P, Arbuick SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst*. 2000;92:205–216.
26. Hatt M, Cheze le Rest C, 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.
27. Amadasun M, King R. Textural features corresponding to textural properties. *IEEE Trans Syst Man Cybern*. 1989;19:1264–1274.
28. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern*. 1973;3:610–621.
29. Loh H, Leu J, Luo R. The analysis of natural textures using run length features. *IEEE Trans Ind Electron*. 1988;35:323–328.
30. Thibault G, Fertil B, Navarro C, et al. Texture indexes and gray level size zone matrix: application to cell nuclei classification. *Pattern Recognition Inf Process*. 2009;140–145.
31. Hautzel H, Müller-Gärtner HW. Early changes in fluorine-18-FDG uptake during radiotherapy. *J Nucl Med*. 1997;38:1384–1386.
32. Erasmus JJ, Munden RF, Truong MT, et al. Preoperative chemoradiation-induced ulceration in patients with esophageal cancer: a confounding factor in tumor response assessment in integrated computed tomography-positron emission tomography imaging. *J Thorac Oncol*. 2006;1:478–486.
33. Gillies RJ, Anderson AR, Gatenby RA, Morse DL. The biology underlying molecular imaging in oncology: from genome to anatome and back again. *Clin Radiol*. 2010;65:517–521.
34. Diehn M, Nardini C, Wang DS, et al. Identification of noninvasive imaging surrogates for brain tumor gene-expression modules. *Proc Natl Acad Sci USA*. 2008;105:5213–5218.
35. Strauss LG, Pan L, Koczan D, et al. Fusion of positron emission tomography (PET) and gene array data: a new approach for the correlative analysis of molecular biological and clinical data. *IEEE Trans Med Imaging*. 2007;26:804–812.