



**HAL**  
open science

## Association between textural and morphological tumor indices on baseline PET-CT and early metabolic response on interim PET-CT in bulky malignant lymphomas

Fayçal Ben Bouallègue, Yassine Al Tabaa, Marilyne Kafrouni, Guillaume Cartron, Fabien Vauchot, Denis Mariano-Goulart

### ► To cite this version:

Fayçal Ben Bouallègue, Yassine Al Tabaa, Marilyne Kafrouni, Guillaume Cartron, Fabien Vauchot, et al.. Association between textural and morphological tumor indices on baseline PET-CT and early metabolic response on interim PET-CT in bulky malignant lymphomas. *Medical Physics*, 2017, 44 (9), pp.4608-4619. 10.1002/mp.12349 . hal-02402829

**HAL Id: hal-02402829**

**<https://hal.science/hal-02402829>**

Submitted on 12 Dec 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Association between textural and morphological tumor indices on baseline PET-CT and early metabolic response on interim PET-CT in bulky malignant lymphomas

Fayçal Ben Bouallègue,<sup>a)</sup> Yassine Al Tabaa, and Marilyne Kafrouni

*Nuclear Medicine Department, Gui de Chauliac University Hospital, 80, Avenue Augustin Fliche, 34295 Montpellier Cedex 5, France*

Guillaume Cartron

*Haematology Department, Saint Eloi University Hospital, 80, Avenue Augustin Fliche, 34295 Montpellier Cedex 5, France*

Fabien Vauchot

*Nuclear Medicine Department, Gui de Chauliac University Hospital, 80, Avenue Augustin Fliche, 34295 Montpellier Cedex 5, France*

Denis Mariano-Goulart

*Nuclear Medicine Department, Gui de Chauliac University Hospital, 80, Avenue Augustin Fliche, 34295 Montpellier Cedex 5, France*

*U1046 INSERM - UMR9214 CNRS, CHU Arnaud de Villeneuve, 371 Avenue du Doyen Giraud, 34295 Montpellier Cedex 5, France*

**Purpose:** We investigated whether metabolic, textural, and morphological tumoral indices evaluated on baseline PET-CT were predictive of early metabolic response on interim PET-CT in a cohort of patients with bulky Hodgkin and non-Hodgkin malignant lymphomas.

**Methods:** This retrospective study included 57 patients referred for initial PET-CT examination. In-house dedicated software was used to delineate tumor contours using a fixed 30% threshold of SUV max and then to compute tumoral metabolic parameters (SUV max, mean, peak, standard deviation, skewness and kurtosis, metabolic tumoral volume (MTV), total lesion glycolysis, and area under the curve of the cumulative histogram), textural parameters (Moran's and Geary's indices, energy, entropy, contrast, correlation derived from the gray-level co-occurrence matrix, area under the curve of the power spectral density, auto-correlation distance, and granularity), and shape parameters (surface, asphericity, convexity, surfacic extension, and 2D and 3D fractal dimensions). Early metabolic response was assessed on interim PET-CT using the Deauville 5-point scale and patients were ranked according to the Lugano classification as complete or not complete metabolic responders. The impact of the segmentation method (alternate threshold at 41%) and image resolution (Gaussian postsmoothing of 3, 5, and 7 mm) was investigated. The association of the proposed parameters with early response was assessed in univariate and multivariate analyses. Their added predictive value was explored using supervised classification by support vector machines (SVM). We evaluated in leave-one-out cross-validation three SVMs admitting as input features (a) MTV, (b) MTV + histological type, and (c) MTV + histology + relevant texture/shape indices.

**Results:** Features associated with complete metabolic response were low MTV ( $P = 0.01$ ), low TLG ( $P = 0.003$ ), high power spectral density AUC ( $P = 0.007$ ), high surfacic extension ( $P = 0.006$ ), low 2D fractal dimension ( $P = 0.007$ ), and low 3D fractal dimension ( $P = 0.003$ ). The prognostic value of these metrics was optimal with the 30% segmentation threshold and overall was progressively altered with decreasing image resolution. In cross-validation, the SVM accounting for texture and shape achieved the highest predictive value with ROC AUC of 0.82 and 80% accuracy (compared with 0.68 and 61% for MTV, and 0.65 and 68% for MTV + histology).

**Conclusions:** The combination of usual prognostic factors with appropriately chosen textural and shape parameters evaluated on baseline PET-CT improves the prediction of early metabolic response in bulky lymphoma.

Key words:  $^{18}\text{F}$ FDG-PET, early metabolic response, lymphoma, morphology, texture

## 1. INTRODUCTION

Positron emission tomography coupled with computed tomography (PET-CT) using  $^{18}\text{F}$ -fluorodeoxyglucose

( $^{18}\text{F}$ FDG) has long been a reference imaging tool for the diagnosis and staging of lymphomatous disease,<sup>1,2</sup> particularly in  $^{18}\text{F}$ FDG avid histological types like Hodgkin disease and diffuse large B cell lymphomas. Baseline PET-CT is used to

characterize the metabolic tumor burden and detect nodal and extra-nodal involvement. PET-CT is also pertinent for therapeutic evaluation, enabling the assessment of tumor metabolic response to chemotherapy.<sup>3</sup> Interim PET-CT after two to three cycles of chemotherapy has proven to be an independent predictor of progression-free and overall survival in Hodgkin and non-Hodgkin lymphomas.<sup>4,5</sup> Currently, the prognostic information deduced from initial and interim PET examinations is based on qualitative review (number and topography of nodal sites, visceral involvement), uptake quantification and follow-up using standardized uptake value (SUV), and volumetric measurements (metabolic tumor size or volume).<sup>6–8</sup> This information enables therapy to be tailored to the individual patient, via minimization or intensification for low- and high-risk patient populations respectively.<sup>9,10</sup>

Tumor heterogeneity at both microscopic and macroscopic scales is related to cancer growth and dissemination as it reflects such histopathological features as tumor cellularity, necrosis, hypoxia, and angiogenesis, as well as the co-existence of multiple cell variants within primary tumors.<sup>11</sup> In the past few decades, effort in oncological imaging has focused on quantitative image analysis and image-based feature extraction<sup>12</sup> as noninvasive methods to provide additional information on intratumoral heterogeneity. Owing in part to ongoing improvements in PET spatial resolution (time-of-flight acquisition, 3D iterative reconstruction, point spread function (PSF) correction), textural and morphological characterization of tumoral metabolic patterns in <sup>18</sup>F-DG-PET has become an emerging topic in nuclear medicine.<sup>13,14</sup> Several studies have investigated its potential interest in nonsmall-cell lung cancer for initial diagnosis and staging<sup>15–17</sup> and have shown that textural abnormalities are predictive of response to chemotherapy and survival.<sup>18</sup> Recent attempts have also been made to correlate textural features with pathological prognostic factors in breast cancer<sup>19</sup> and to exploit textural analysis for therapeutic response prediction and monitoring in esophageal,<sup>20</sup> cervical,<sup>21</sup> and breast<sup>22</sup> cancers and for planning external radiation beam therapy in head and neck cancers.<sup>23</sup> Beyond clinical considerations, the heterogeneity computed on PET-CT correlates well with that computed *ex vivo* from autoradiography and the spatial cell patterns observed on pathological slides.<sup>24</sup>

With regard to lymphomas, there is very little evidence that textural analysis provides additional diagnostic or prognostic information. Lartizien et al. showed that textural information extracted from <sup>18</sup>F-DG PET-CT might help to better discriminate between neoplastic and inflammatory hypermetabolic foci.<sup>25</sup> Recent research has tended to demonstrate that textural analysis of computed tomography (CT)<sup>26,27</sup> or magnetic resonance (MR) images<sup>28</sup> can predict therapeutic response and progression-free survival in Hodgkin and non-Hodgkin lymphomas. In the present study, we sought to investigate the extent to which textural and morphological tumoral indices evaluated on baseline <sup>18</sup>F-DG PET-CT were predictive of early metabolic response assessed on interim PET-CT in a cohort of Hodgkin and non-Hodgkin patients presenting with a bulky mass.

## 2. MATERIALS AND METHODS

### 2.A. Study population

A retrospective study was conducted with patients referred to the nuclear medicine department of Montpellier University Hospital, France, from June 2014 to January 2016 for the initial PET-CT evaluation of a malignant lymphoma. The study was approved by the local ethics committee and the requirement for individual informed consent was waived. The inclusion criteria were the presence of at least one bulky tumor (nodal or extra-nodal) with metabolic tumor volume over 50 cm<sup>3</sup> on the baseline PET-CT and histological proof of malignant lymphoma (either Hodgkin or non-Hodgkin). Fifty-seven patients between 12 and 93 yr old were included (median 52 yr). Table I summarizes the clinical characteristics of the study population as well as the histological classification, Ann Arbor staging, and tumor localization. Twenty-five percent of the patients presented with Hodgkin lymphomas (HL) and 75% with non-Hodgkin lymphomas (NHL), predominantly diffuse large B cell (DLBCL) and transformed follicular lymphomas. Almost two-thirds of the patients presented extra-nodal involvement (Ann Arbor stage IV). The tumor localization was nodal in 60% of the cases and extra-nodal (mostly splenic, hepatic, or osteo-medullary) in 40%.

All patients benefited from an interim PET-CT evaluation after 2–6 (median 4) cycles of chemotherapy (mostly ABVD or BEACOPP in HL, and R-CHOP in NHL). The

TABLE I. Clinical characteristics of the study population. Histological classification. Ann Arbor staging and bulky mass localization on baseline PET-CT. Metabolic response on interim PET-CT.

Gender	
Male	34 (60%)
Female	23 (40%)
Age (years)	52 ± 20
Histological type	
Hodgkin	14 (25%)
Non-Hodgkin	43 (75%)
Diffuse large B cell	31 (54%)
Follicular	9 (16%)
Mantle cell	1 (2%)
Marginal zone	1 (2%)
Burkitt	1 (2%)
Ann arbor stage	
I	2 (4%)
II	9 (16%)
III	10 (18%)
IV	36 (63%)
Bulky mass localization	
Nodal	34 (60%)
Extra-nodal	23 (40%)
Lugano classification	
Complete metabolic response	32 (56%)
No complete metabolic response	25 (44%)

interval between baseline and interim PET-CT evaluations ranged from 5 to 32 weeks (median 11 weeks). Interim PET-CT scans were interpreted and scored using the Deauville 5-point scale, and patients could thus be ranked according to the Lugano classification<sup>3</sup> as complete metabolic responders (CMR, for Deauville 1–3) or not complete metabolic responders (no-CMR, for Deauville 4–5). Thirty-two (56%) patients were rated CMR, while 25 (44%) were rated no-CMR.

All PET-CT examinations were performed with a Siemens Biograph mCT 20 Flow (four rings) scanner following IV injection of 3.5 MBq/Kg of <sup>18</sup>F-DG. The PET axial field of view systematically encompassed head, thorax, abdomen, and pelvis. PET data were acquired in time-of-flight mode and images were reconstructed using the manufacturer's dedicated software and specifications (3D OSEM using 21 subsets and two iterations including PSF correction followed by postfiltering with a 3 mm wide Gaussian kernel). Image matrices were sampled on a 400 × 400 grid with a voxel size of 2.04 × 2.04 × 2.03 mm<sup>3</sup>.

## 2.B. Metabolic, textural, and morphological analysis

Metabolic, textural, and shape analyses were performed on the baseline PET-CT examinations retrieved from the local archiving system using in-house dedicated software. The choice of the studied tumor site and the subsequent computations were conducted blind to the results of the interim PET-CT. For each patient, only the largest tumor site with a metabolic volume over 50 cm<sup>3</sup> was taken into account and postprocessed. Reconstructed PET data were smoothed using Gaussian postfiltering of increasing FWHM (3, 5, and 7 mm) in order to test the impact of image resolution and PSF correction. Then tumor boundaries were defined using a fixed threshold with respect to the maximal SUV and a connectivity constraint. Extratumoral uptake (myocardial, cerebral, urinary) was manually masked whenever necessary. Two threshold values (30% and 41%) were sequentially implemented in order to evaluate the impact of the segmentation technique.<sup>29</sup> Once the tumor volume of interest (VOI) was segmented, the following parameters were estimated.

### 2.B.1. Metabolic indices

Computed indices included maximal SUV (SUV max), peak SUV in a 1 cm<sup>3</sup> region around the maximal SUV voxel so that the mean SUV in that region was maximal (SUV peak), mean SUV in the tumor (SUV mean), second, third, and fourth moments of the SUV distribution in the tumor (respectively, SUV variance, skewness, and kurtosis), the area under the curve of the voxel cumulative histogram sampled between minimal and maximal SUV (histogram AUC), metabolic tumor volume (MTV) as the tumor volume in cm<sup>3</sup>, and total lesion glycolysis (TLG) as the product of MTV multiplied by the SUV mean.

### 2.B.2. Textural indices

Let  $D$  be the radioactive distribution indexed by voxel number  $i$  ( $i = 1 \dots N$ ). Moran's index ( $I$ ) and Geary's index ( $C$ ) were defined as:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (D_i - \bar{D})(D_j - \bar{D})}{\sum_i (D_i - \bar{D})^2}$$

$$C = \frac{N-1}{2 \sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (D_i - D_j)^2}{\sum_i (D_i - \bar{D})^2}$$

with  $\bar{D}$  the average value of  $D$  and  $w$  a local weighting kernel. Moran's and Geary's indices are measures of spatial autocorrelation and have the ability to detect repeating patterns in the studied image.

The gray-level co-occurrence matrix (GLCM)  $G$  was defined as:<sup>30</sup>

$$G_{ab} = \sum_{i=1 \dots N} \sum_{j \in \omega(i)} \begin{cases} 1 & \text{if } \hat{D}_i = a \text{ and } \hat{D}_j = b \\ 0 & \text{otherwise} \end{cases}$$

where  $\omega(i)$  stood for the three-dimensional 26-neighbourhood of voxel  $i$ ,  $\hat{D}$  was the discrete resampled distribution, and with  $a, b = 1 \dots L$  where  $L$  is the number of gray levels. The matrix was built based on the assumption that the textural information in an image is contained in the average spatial relationship that the gray tones in the image have to one another. This relationship was described using GLCM-based textural features defined as follows:

$$\text{Energy: } E = \sum_a \sum_b G_{ab}^2$$

$$\text{Entropy: } S = - \sum_a \sum_b G_{ab} \log(G_{ab})$$

$$\text{Contrast: } T = \sum_a \sum_b (a - b)^2 G_{ab}$$

Correlation:  $O = \sum_a \sum_b \frac{(a - \mu_a)(b - \mu_b)}{\sigma_a \sigma_b} G_{ab}$  with  $\mu_a, \mu_b, \sigma_a, \sigma_b$ , respectively, the marginal means and standard deviations along the  $a$  and  $b$  dimensions.

The GLCM was computed after resampling of the voxel SUV values using two distinct resampling techniques.<sup>31</sup> The first technique was based on relative resampling using  $L$  gray levels equally distributed between minimal and maximal SUV in the tumor:  $\hat{D} = \text{round}\left\{L \frac{D - \text{SUV}_{\min}}{\text{SUV}_{\max} - \text{SUV}_{\min}}\right\}$ . The second technique relied on absolute resampling in which the lower and upper SUV bounds were set to 0 and 100:  $\hat{D} = \text{round}\left\{L \frac{D}{100}\right\}$ . Three discretization levels ( $L = 25, 50$  and 100) were implemented in order to evaluate the influence of the number of gray levels on the computed parameters.

Let  $\hat{D}$  be the 3D Fourier transform of  $D$ . The normalized power spectral density (PSD) was defined as  $PSD(v) = \kappa \frac{1}{2\pi v^2} \iint_{S_v} |\hat{D}|^2$  where  $S_v$  was the 2-sphere of radius  $v$  and  $\kappa$  a normalization coefficient such that  $PSD(0) = 1$ . The area under the curve of the PSD (PSD AUC) was normalized by the maximal (Nyquist) frequency and stood as a

interval between baseline and interim PET-CT evaluations ranged from 5 to 32 weeks (median 11 weeks). Interim PET-CT scans were interpreted and scored using the Deauville 5-point scale, and patients could thus be ranked according to the Lugano classification<sup>3</sup> as complete metabolic responders (CMR, for Deauville 1–3) or not complete metabolic responders (no-CMR, for Deauville 4–5). Thirty-two (56%) patients were rated CMR, while 25 (44%) were rated no-CMR.

All PET-CT examinations were performed with a Siemens Biograph mCT 20 Flow (four rings) scanner following IV injection of 3.5 MBq/Kg of <sup>18</sup>F-DG. The PET axial field of view systematically encompassed head, thorax, abdomen, and pelvis. PET data were acquired in time-of-flight mode and images were reconstructed using the manufacturer's dedicated software and specifications (3D OSEM using 21 subsets and two iterations including PSF correction followed by postfiltering with a 3 mm wide Gaussian kernel). Image matrices were sampled on a 400 × 400 grid with a voxel size of 2.04 × 2.04 × 2.03 mm<sup>3</sup>.

## 2.B. Metabolic, textural, and morphological analysis

Metabolic, textural, and shape analyses were performed on the baseline PET-CT examinations retrieved from the local archiving system using in-house dedicated software. The choice of the studied tumor site and the subsequent computations were conducted blind to the results of the interim PET-CT. For each patient, only the largest tumor site with a metabolic volume over 50 cm<sup>3</sup> was taken into account and postprocessed. Reconstructed PET data were smoothed using Gaussian postfiltering of increasing FWHM (3, 5, and 7 mm) in order to test the impact of image resolution and PSF correction. Then tumor boundaries were defined using a fixed threshold with respect to the maximal SUV and a connectivity constraint. Extratumoral uptake (myocardial, cerebral, urinary) was manually masked whenever necessary. Two threshold values (30% and 41%) were sequentially implemented in order to evaluate the impact of the segmentation technique.<sup>29</sup> Once the tumor volume of interest (VOI) was segmented, the following parameters were estimated.

### 2.B.1. Metabolic indices

Computed indices included maximal SUV (SUV max), peak SUV in a 1 cm<sup>3</sup> region around the maximal SUV voxel so that the mean SUV in that region was maximal (SUV peak), mean SUV in the tumor (SUV mean), second, third, and fourth moments of the SUV distribution in the tumor (respectively, SUV variance, skewness, and kurtosis), the area under the curve of the voxel cumulative histogram sampled between minimal and maximal SUV (histogram AUC), metabolic tumor volume (MTV) as the tumor volume in cm<sup>3</sup>, and total lesion glycolysis (TLG) as the product of MTV multiplied by the SUV mean.

### 2.B.2. Textural indices

Let  $D$  be the radioactive distribution indexed by voxel number  $i$  ( $i = 1 \dots N$ ). Moran's index ( $I$ ) and Geary's index ( $C$ ) were defined as:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (D_i - \bar{D})(D_j - \bar{D})}{\sum_i (D_i - \bar{D})^2}$$

$$C = \frac{N-1}{2 \sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (D_i - D_j)^2}{\sum_i (D_i - \bar{D})^2}$$

with  $\bar{D}$  the average value of  $D$  and  $w$  a local weighting kernel. Moran's and Geary's indices are measures of spatial autocorrelation and have the ability to detect repeating patterns in the studied image.

The gray-level co-occurrence matrix (GLCM)  $G$  was defined as:<sup>30</sup>

$$G_{ab} = \sum_{i=1 \dots N} \sum_{j \in \omega(i)} \begin{cases} 1 & \text{if } \hat{D}_i = a \text{ and } \hat{D}_j = b \\ 0 & \text{otherwise} \end{cases}$$

where  $\omega(i)$  stood for the three-dimensional 26-neighbourhood of voxel  $i$ ,  $\hat{D}$  was the discrete resampled distribution, and with  $a, b = 1 \dots L$  where  $L$  is the number of gray levels. The matrix was built based on the assumption that the textural information in an image is contained in the average spatial relationship that the gray tones in the image have to one another. This relationship was described using GLCM-based textural features defined as follows:

$$\text{Energy: } E = \sum_a \sum_b G_{ab}^2$$

$$\text{Entropy: } S = - \sum_a \sum_b G_{ab} \log(G_{ab})$$

$$\text{Contrast: } T = \sum_a \sum_b (a-b)^2 G_{ab}$$

Correlation:  $O = \sum_a \sum_b \frac{(a-\mu_a)(b-\mu_b)}{\sigma_a \sigma_b} G_{ab}$  with  $\mu_a, \mu_b, \sigma_a, \sigma_b$ , respectively, the marginal means and standard deviations along the  $a$  and  $b$  dimensions.

The GLCM was computed after resampling of the voxel SUV values using two distinct resampling techniques.<sup>31</sup> The first technique was based on relative resampling using  $L$  gray levels equally distributed between minimal and maximal SUV in the tumor:  $\hat{D} = \text{round}\left\{L \frac{D - \text{SUV}_{\min}}{\text{SUV}_{\max} - \text{SUV}_{\min}}\right\}$ . The second technique relied on absolute resampling in which the lower and upper SUV bounds were set to 0 and 100:  $\hat{D} = \text{round}\left\{L \frac{D}{100}\right\}$ . Three discretization levels ( $L = 25, 50$  and 100) were implemented in order to evaluate the influence of the number of gray levels on the computed parameters.

Let  $\hat{D}$  be the 3D Fourier transform of  $D$ . The normalized power spectral density (PSD) was defined as  $PSD(v) = \kappa \frac{1}{2\pi v^2} \iint_{S_v} |\hat{D}|^2$  where  $S_v$  was the 2-sphere of radius  $v$  and  $\kappa$  a normalization coefficient such that  $PSD(0) = 1$ . The area under the curve of the PSD (PSD AUC) was normalized by the maximal (Nyquist) frequency and stood as a

measure of spatial frequency distribution inside the tumor VOI. Using the Wiener-Khintchine theorem, the autocorrelation (AC) function of  $D$  can be written as  $R_D = \widetilde{PSD}$  where  $\widetilde{\phantom{x}}$  stands for the 1D inverse Fourier transform. The AC distance  $\bar{d}$  was defined as the distance such that  $R_D(d) < 0.1 R_D(0)$  for  $d \geq \bar{d}$ . Granularity  $Y$  was computed as  $Y = -\log\left(\frac{\bar{d}^3}{MTV}\right)$  and was a measure of how fine the image grain was compared to the tumor dimensions.

### 2.B.3. Morphological (shape) indices

Surface in  $\text{cm}^2$  was the surface of the 3D contour of the tumor VOI computed on a voxel basis. Asphericity was defined as  $A = \frac{1}{36\pi} \frac{Surf^3}{MTV^2} - 1$  with  $Surf$  the surface of the tumor VOI and  $36\pi$  the value of the ratio  $\frac{Surf^3}{MTV^2}$  for a perfectly spherical VOI. It depicted how the tumor VOI departed from a sphere. The convex hull of a tumor VOI was defined as the smallest convex set of voxels containing the VOI. Convexity was defined as the ratio of the volume of the tumor VOI to the volume of its convex hull. Surfacic extension was defined as the ratio of the surface of the convex hull to the surface of the tumor VOI. The Minkowski-Bouligand fractal dimension of an  $m$ -dimensional hypersurface  $\Phi$  was defined as  $\dim(\Phi) = \lim_{\varepsilon \rightarrow 0} \frac{\log(n_\varepsilon)}{\log(1/\varepsilon)}$  with  $n_\varepsilon$  the total number of  $m$ -dimensional boxes of side length  $\varepsilon$  required to cover  $\Phi$ . It was evaluated by computing  $n_\varepsilon$  for decreasing values of  $\varepsilon$  using a box-counting method, then estimating the slope of  $\{\log(\frac{1}{\varepsilon}); \log(n_\varepsilon)\}$  through a linear regression. In two dimensions (2D), it stood for the fractal dimension of the surface of the tumor VOI. In three dimensions (3D), it stood for the fractal dimension of the 3D radioactive distribution inside the tumor VOI processed as a 3D hypersurface.

### 2.C. Statistical analysis

The metabolic, textural and shape indices are given as interquartile range (IQR) because the number of patients in each group was relatively small and most of the computed parameters followed a non-normal distribution. The correlation of the computed indices with MTV and SUV max was assessed using Pearson's coefficient  $R$ . Their association with histological type (HL or NHL) was evaluated using a Wilcoxon rank sum test.

The distributions of metabolic, textural and shape indices in CMR and no-CMR patients were compared using a Wilcoxon rank sum test. For each parameter that significantly differed between groups, a multivariate analysis was conducted to compensate the potential confusion related to MTV and histological type using stepwise linear regression with an entry criterion of  $P < 0.05$  and a removal criterion of  $P > 0.10$ . The first regression included the tested parameter and MTV as explicative factors. The second regression included, in addition, the histological type and Ann Arbor stage. A two-sided  $P$ -value  $\leq 0.05$  was considered statistically significant. In order to account for multiple testing, the

Benjamini–Hochberg procedure was employed to correct for type I error inflation.

To test for predictive capabilities, a receiver operating characteristic (ROC) analysis was conducted for each parameter. The area under the ROC curve (AUC) reflected the global performance of the tested parameter in discriminating complete responders from noncomplete responders. The optimal accuracy was obtained by maximizing Youden's index (sensitivity + specificity - 1). The added prognostic value of textural and shape indices was further investigated using support vector machines (SVM). An SVM is a nonprobabilistic binary linear classifier based on a set of hyperplanes that are optimized using a training dataset.<sup>32</sup> Here, we studied three SVMs which admitted as input features: (a) MTV, (b) MTV and histological type, and (c) MTV, histological type, textural, and shape features. The performances of these SVMs were evaluated in self-validation and leave-one-out cross-validation. In self-validation, both the training and testing sets were constituted of the whole patient population. In leave-one-out cross-validation, each patient was classified using SVMs trained based on all patients except that one. The superiority of the SVM accounting for textural and shape features was assessed by comparing ROC AUCs using a one-sided test. Accuracy, sensitivity, specificity, and negative and predictive positive values were estimated, along with their 95% confidence interval computed using the Wilson score interval with continuity correction.

All statistical computations were performed using Matlab R2013 (The Math Works, Natick, MA, USA).

## 3. RESULTS

Table II summarizes the characteristics and performances of metabolic, textural and shape parameters. It details the interquartile ranges for the CMR and no-CMR patients, the statistical significance of the difference between the two groups in the univariate and multivariate analyses, the ROC AUC and accuracy, the correlation with MTV and SUV max, and the association with histological type. The GLCM-derived indices mentioned in the table are those computed using relative resampling and 50 gray levels. The number of gray levels used to compute the GLCM affected the amplitude of the GLCM-derived indices but did not have a significant influence on their correlation with metabolic response. The GLCM indices resulting from absolute resampling, which are not reported in the table, were highly correlated with SUV max ( $|R| \sim 0.80$  for entropy, energy, and contrast), and none of them was associated with early metabolic response (ROC AUC of 0.57, 0.57, 0.54, and 0.62 for entropy, energy, contrast, and correlation respectively). All the results shown in Table II are those obtained on the basis of 3 mm postsmoothed PET images segmented using a 30% SUVmax threshold, which will be considered as the reference results in the sequel. Figure 1 displays the distribution of MTV, and TLG in the CMR and no-CMR patients, as well as the six indices that were significantly associated with metabolic response in both the univariate and multivariate analyses

TABLE II. Association between metabolic, textural, and morphological indices on baseline PET-CT (3 mm postfiltering, segmentation threshold 30%) and tumor metabolic response on interim PET-CT. Correlation with tumor volume, SUV max, and histological type.

	Association with early metabolic response										
	CMR IQR	No-CMR IQR	p	(adjusted p)	ROC analysis		Correlation with MTV		Correlation with SUV max		Assoc. histol. p
					AUC	Acc.	R	p	R	p	
<b>Metabolic indices</b>											
SUV max	[14.8 29.1]	[19.2 37.2]	ns		0.63	64%	0.07	ns	–	–	0.008*
SUV peak	[10.3 20.4]	[14.6 26.5]	0.03	(ns)	0.67	68%	0.17	ns	0.97	<0.001*	0.001*
SUV mean	[6.2 14.5]	[8.9 16.0]	ns		0.64	68%	0.09	ns	0.95	<0.001*	0.001*
SUV std. dev.	[2.7 15.3]	[4.0 16.8]	ns		0.58	61%	0.00	ns	0.93	<0.001*	0.007*
SUV skewness	[0.25 1.24]	[0.42 1.16]	ns		0.50	55%	–0.22	ns	–0.17	ns	0.01*
SUV kurtosis	[2.6 4.9]	[2.6 4.9]	ns		0.53	58%	–0.05	ns	–0.12	ns	Ns
Cum. hist. AUC (%)	[69.6 82.7]	[70.3 82.8]	ns		0.54	58%	–0.14	ns	–0.16	ns	Ns
MTV (mL)	[89 408]	[207 1066]	0.01*	(ns)	0.70	69%	–	–	0.07	ns	Ns
TLG (mL)	[960 3870]	[2560 11200]	0.003*	(ns)	0.74	72%	0.81	<0.001*	0.47	<0.001*	0.001*
<b>Textural indices</b>											
Moran's I	[0.87 0.93]	[0.87 0.93]	ns		0.51	60%	0.21	ns	0.26	0.05	Ns
Geary's C	[0.02 0.07]	[0.02 0.05]	ns		0.56	59%	–0.13	ns	–0.58	<0.001*	0.01*
Entropy	[6.1 6.6]	[5.8 6.5]	ns		0.63	62%	–0.25	ns	0.07	ns	Ns
Energy	[5.0 9.8]	[5.7 11.8]	ns		0.62	66%	0.19	ns	–0.04	ns	Ns
Contrast	[22.3 39.0]	[18.6 30.3]	0.02	(0.01–0.008)	0.69	70%	–0.47	<0.001*	0.12	ns	Ns
Correlation	[36.3 39.7]	[36.6 42.6]	ns		0.63	67%	0.50	<0.001*	–0.11	ns	Ns
PSD AUC (%)	[12.3 17.3]	[9.3 14.1]	0.007*	(0.005*–0.05)	0.71	69%	–0.78	<0.001*	–0.04	ns	0.01*
AC distance (mm)	[7 9]	[7 10]	ns		0.51	58%	0.42	0.001*	0.11	ns	ns
Granularity	[5.20 6.72]	[6.00 7.55]	0.02	(0.02–0.05)	0.68	68%	0.62	<0.001*	–0.08	ns	Ns
<b>Shape indices</b>											
Surface (cm <sup>2</sup> )	[269 623]	[507 1339]	0.002*	(ns)	0.73	73%	0.77	<0.001*	–0.04	ns	0.02
Asphericity	[5.2 25.2]	[8.6 60.3]	ns		0.63	69%	0.08	ns	–0.06	ns	Ns
Convexity	[0.31 0.53]	[0.31 0.58]	ns		0.55	59%	0.15	ns	0.00	ns	Ns
Surf. extension	[0.95 1.11]	[0.71 1.03]	0.006*	(0.01–0.02)	0.71	70%	–0.43	0.001*	0.07	ns	Ns
2D fractal dim.	[1.99 2.12]	[2.07 2.20]	0.007*	(0.003*–0.003)	0.71	66%	0.41	0.002*	–0.09	ns	Ns
3D fractal dim.	[2.96 3.08]	[3.03 3.19]	0.003*	(0.001*–0.01)	0.73	72%	0.69	<0.001*	–0.05	ns	0.04

GLCM-derived features were obtained using relative resampling. CMR, complete metabolic response; IQR, interquartile range; AUC, area under the ROC curve; Acc., balanced accuracy. *P*-values are adjusted on MTV, then on MTV, histology and Ann Arbor stage.

\*remained significant after correction for multiple testing.

(contrast, PSD AUC, granularity, surfacic extension, 2D and 3D fractal dimensions).

MTV and TLG were significantly lower in the early metabolic responders in the univariate analysis, but not after adjustment for histological type and disease stage since histological type appeared as the main determinant of metabolic response ( $P < 0.001$ ). Optimal cut-off values derived from the ROC analysis were 604 mL and 3040 mL for MTV and TLG respectively. For the GLCM-derived textural indices, the contrast was significantly higher in the CMR patients ( $P = 0.02$ , adjusted 0.008). For the PSD-derived indices, PSD AUC was significantly increased ( $P = 0.007$ , adjusted 0.05) and granularity significantly decreased ( $P = 0.02$ , adjusted 0.05) in the CMR patients. As for the shape parameters, the metabolic response was significantly associated with surfacic extension (higher for CMR patients,  $P = 0.006$ , adjusted 0.02), 2D fractal dimension (lower in CMR patients,

$P = 0.007$ , adjusted 0.003) and 3D fractal dimension (lower in CMR patients,  $P = 0.003$ , adjusted 0.01). In the ROC analysis, MTV and TLG were the best metabolic predictors of early response (AUCs 0.70 and 0.74, accuracies 69% and 72% respectively). Contrast, PSD AUC, surfacic extension and 3D fractal dimensions achieved similar predictive performances with ROC AUCs ranging from 0.69 to 0.73 and accuracies ranging from 69% to 72%.

Table III details the impact of segmentation threshold (41%, using 3 mm postsmoothed PET data) and postsmoothing level (5 or 7 mm, using a 30% segmentation threshold) on MTV, TLG, and relevant texture and shape indices. For each parameter, the table shows the bias and Pearson's correlation with respect to the reference parameter (3 mm postsmoothing and 30% thresholded VOIs) and the strength of the univariate association with early metabolic response. For comparison, the corresponding *p*-values for the reference parameter are

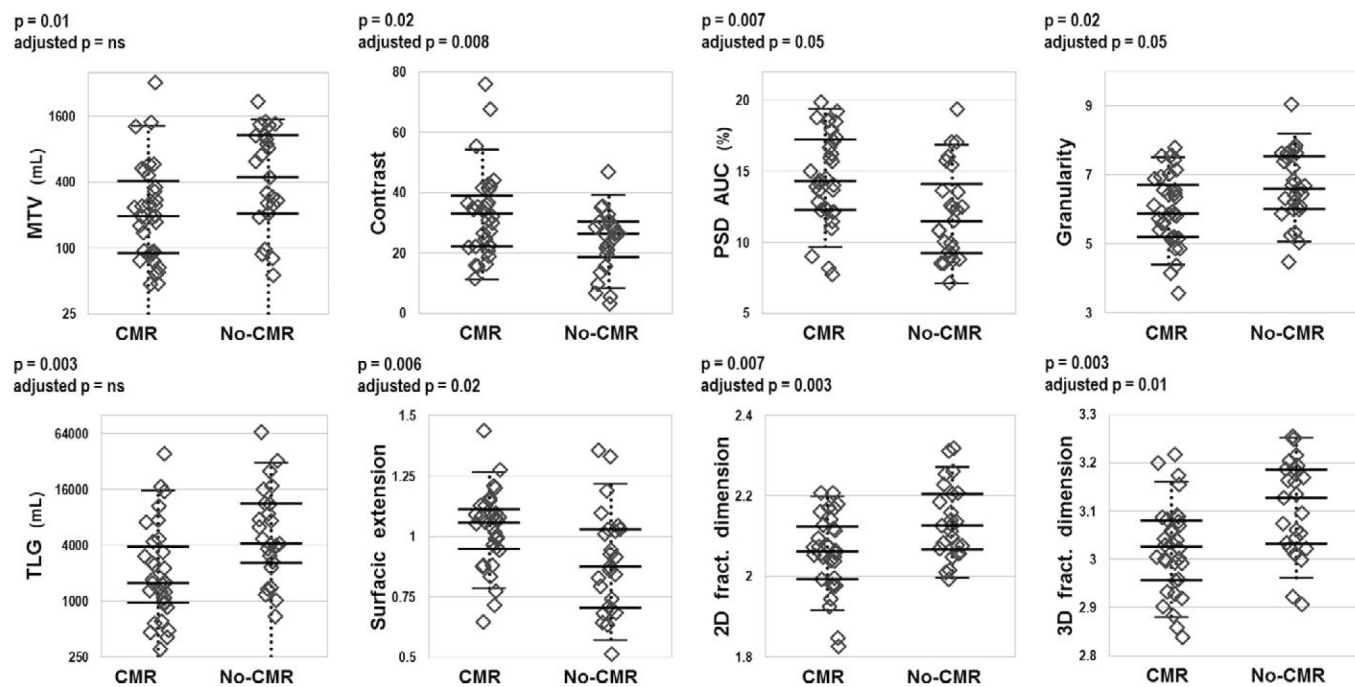


FIG. 1. Association between metabolic, textural, and morphological indices on baseline PET-CT (3 mm postfiltering, segmentation threshold 30%) and metabolic response on interim PET-CT. CMR: complete metabolic response. Thick lines: median and interquartile range; thin lines: mean  $\pm$  1.5 st. dev. Note the logarithmic scale for MTV and TLG.

TABLE III. Impact of segmentation threshold (41%) and image resolution (postsMOOTHING 5 and 7 mm). For each combination of threshold and smoothing level, the first two columns show the relative bias and Pearson's correlation with respect to the reference parameters computed using 3 mm postsmoothed images and 30% thresholded VOIs (\*:  $P < 0.001$ , \*\*:  $P < 0.01$ ). The third column gives the strength of the association with early metabolic response.

	Reference (3 mm, 30%) Assoc. w/response	Threshold 41% (postsMOOTHING 3 mm)			PostsMOOTHING 5 mm (threshold 30%)			PostsMOOTHING 7 mm (threshold 30%)		
		Bias	Pearson	Assoc. w/response	Bias	Pearson	Assoc. w/response	Bias	Pearson	Assoc. w/response
MTV	$P = 0.01$	-34%	0.74*	$P = 0.06$	0%	0.99*	$P = 0.009$	0%	0.99*	$P = 0.01$
TLG	$P = 0.003$	-23%	0.91*	$P = 0.02$	-2%	0.99*	$P = 0.001$	-4%	0.99*	$P = 0.002$
Contrast	$P = 0.02$	+26%	0.72*	$P = 0.02$	-5%	0.97*	$P = 0.02$	-13%	0.92*	$P = 0.03$
PSD AUC	$P = 0.007$	+19%	0.72*	$P = 0.03$	-6%	0.91*	$P = 0.003$	-9%	0.91*	$P = 0.003$
Granularity	$P = 0.02$	0%	0.85*	$P = 0.02$	-8%	0.96*	$P = 0.03$	-15%	0.93*	$P = 0.04$
Surf. extension	$P = 0.006$	-1%	0.82*	$P = 0.03$	+6%	0.98*	$P = 0.01$	+11%	0.93*	$P = 0.02$
2D fractal dim.	$P = 0.007$	+2%	0.38**	$P = 0.02$	+1%	0.71*	$P = 0.04$	1%	0.44*	$P = 0.35$
3D fractal dim.	$P = 0.003$	-1%	0.66*	$P = 0.01$	+1%	0.82*	$P = 0.01$	1%	0.74*	$P = 0.02$

recalled in the first column of the table. Textural and morphological parameters showed globally higher statistical association with early metabolic response when computed with the 30% thresholded VOIs, compared with those based on a 41% threshold. Image postsMOOTHING markedly impacted the prognostic value of the shape parameters, since their association with metabolic response was progressively altered with increasing postsMOOTHING level. The relevant textural features (contrast, PSD AUC, and granularity) remained significantly associated with early response within the tested range of smoothing levels (3–7 mm).

Figure 2 presents the ROC curves resulting from the evaluation of the implemented SVMs in self-validation (left) and cross-validation (right). The combination of texture and

shape parameters that gave the best added prognostic value to the third SVM was contrast, PSD AUC, surfacic extension, and the 2D and 3D fractal dimensions. Replacing MTV with TLG or taking into account the Ann Arbor stage, SUV max or SUV peak did not bring substantial improvement to any of the three tested SVMs. Table IV details the performances of the three SVMs in terms of ROC AUC, balanced accuracy, sensitivity, specificity, negative and positive predictive values, and risk ratio. For comparison, classifying patients based on histological type alone (HL versus NHL) yielded 67% accuracy, 96% sensitivity, 39% specificity, 92% negative predictive value, 56% positive predictive value, and 7.3 risk ratio. The SVM including textural and shape parameters yielded greater predictive power than those based on MTV



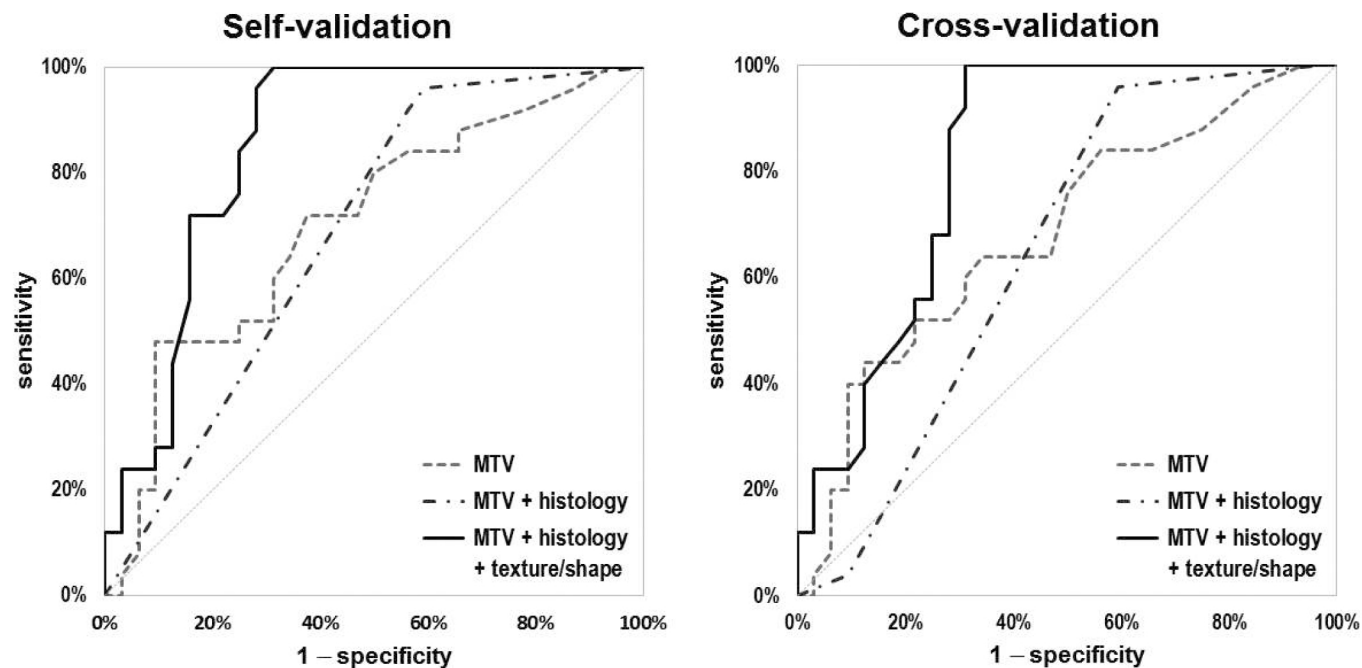


FIG. 2. ROC curves in self-validation (left) and leave-one-out cross-validation (right) for the three tested SVM classifiers. The textural and shape parameters included in the model were contrast, PSD AUC, surfacic extension, and 2D and 3D fractal dimensions (3 mm postfiltering, segmentation threshold 30%). The gray dotted line stands for the random classifier.

TABLE IV. Performances of the SVM classifiers in self- and cross-validation.

	Self-validation			Cross-validation		
	MTV	MTV + histology	MTV + histology + texture/shape	MTV	MTV + histology	MTV + histology + texture/shape
ROC AUC	0.70	0.69	0.86 <sup>(*)</sup>	0.68	0.65	0.82 <sup>(*)</sup>
Accuracy	65% [51% 77%]	68% [55% 80%]	84% [72% 93%]	61% [47% 74%]	68% [55% 80%]	80% [67% 90%]
Sensitivity	40% [22% 61%]	96% [78% 100%]	100% [83% 100%]	32% [16% 53%]	96% [78% 100%]	92% [73% 100%]
Specificity	91% [74% 99%]	41% [24% 59%]	69% [50% 84%]	91% [74% 99%]	41% [24% 59%]	69% [50% 84%]
NPV	66% [50% 79%]	93% [64% 100%]	100% [82% 100%]	63% [48% 77%]	93% [64% 100%]	92% [72% 100%]
PPV	77% [46% 96%]	56% [40% 71%]	71% [53% 85%]	73% [39% 94%]	56% [40% 71%]	70% [51% 84%]
Risk ratio	2.3	7.8	n/a	2.0	7.8	8.4

ROC AUC, area under the ROC curve; PPV, positive predictive value; NPV, negative predictive value; The textural and shape parameters included in the model were contrast, PSD AUC, surfacic extension, and 2D and 3D fractal dimensions (3 mm postfiltering, segmentation threshold 30%).

<sup>\*</sup>one-tailed  $P \leq 0.05$  vs MTV and MTV + histology. n/a: not applicable (no false negative).

and histology, with a high accuracy of 80% in cross-validation (compared with 61% and 68% for MTV and MTV + histology, respectively).

For illustrative purposes, Fig. 3 shows three examples of baseline PET-CT for one Hodgkin and two non-Hodgkin lymphomas. On each subplot the PET-CT fused slices and the PET images (gray-level) with superimposed tumor contours (green line) are displayed. Relevant metabolic, textural, and morphological indices are summarized on the right side of each subplot.

#### 4. DISCUSSION

To our knowledge, this is the first study to investigate the potential prognostic value of textural and shape parameters

evaluated on baseline PET-CT examination in malignant lymphomas. We deliberately chose to focus on patients presenting with a bulky mass since it is well-established that the complementary information provided by textural analysis strongly depends on tumor size and substantially increases with large volumes.<sup>33</sup> Although numerous textural metrics are known to be dependent on metabolic volume,<sup>33,34</sup> we considered a wide range of tumor volumes (from 50 cm<sup>3</sup> to over 3000 cm<sup>3</sup>) in order to match the wide variability of tumor presentation in routine clinical settings. An automatic segmentation method using a relative threshold at 30% of the SUV max was chosen for the sake of robustness and reproducibility. The 30% value was empirically selected as the one providing globally the most satisfying contours with regard to coregistered morphological slices. Furthermore, a

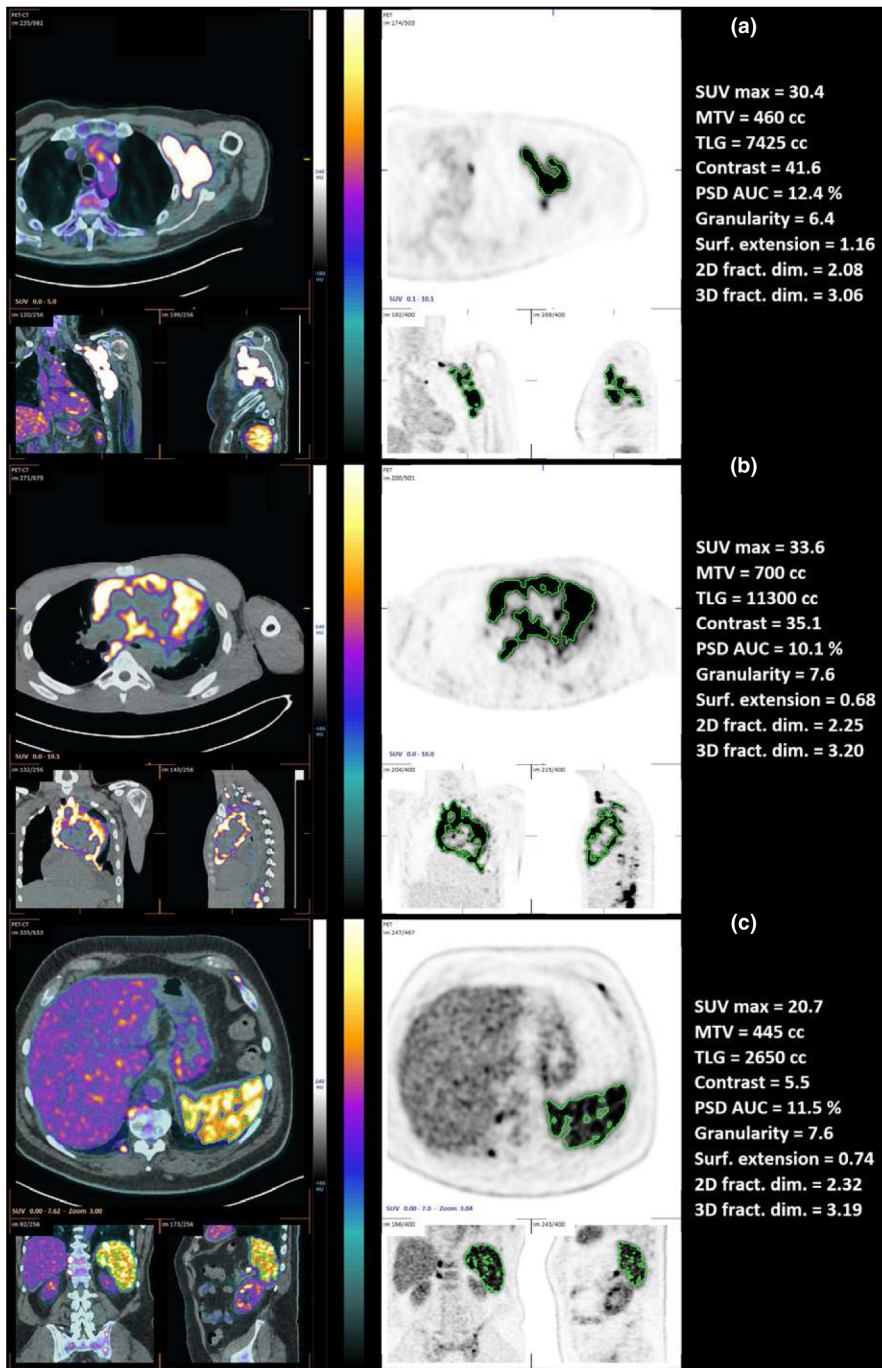


FIG. 3. Examples of baseline PET-CT for three bulky lymphomas. a) 84 year-old patient with Ann-Arbor stage III non-Hodgkin lymphoma showing complete metabolic response on interim PET-CT (Deauville 1). b) 23 year-old patient with Ann Arbor stage IV non-Hodgkin lymphoma showing partial metabolic response on interim PET-CT (Deauville 5). c) 69 year-old patient with Ann Arbor stage IV Hodgkin lymphoma showing partial metabolic response on interim PET-CT (Deauville 4). All three patients were correctly classified based on MTV, histology and textural/shape features. Based on MTV alone (cut-off 604 mL), patients A and C were classified as responders and patient B as nonresponder. Based on MTV and histology, patients A and B were classified as nonresponders and patient C as responder.

recent study concluded that a 30% threshold produces pertinent volumetric segmentation and heterogeneity quantification for highly contrasted bulky tumors.<sup>35</sup> An alternative threshold at 41% of SUVmax was also tested, in line with recent research regarding the prognostic value of metabolic tumor volume.<sup>29,36</sup> The main limitations of the present study are related to its retrospective design and the limited number

of patients. The Hodgkin and non-Hodgkin patients were pooled in order to maximize the statistical power of the study, as has been done in recent CT-based textural analysis.<sup>27</sup> Long-term follow-up was not available for survival assessment since most of the PET-CT examinations dated back only a few months when the statistical analysis was conducted.

#### 4.A. Metabolic indices

Unsurprisingly, none of the first-order metrics describing the SUV distribution was associated with tumor volume. As has already been reported,<sup>37</sup> the SUV values were significantly lower in Hodgkin lymphomas compared with non-Hodgkin, likely due to dissimilar pathological processes leading to <sup>18</sup>F<sub>2</sub> accumulation (higher inflammatory activity in HL). SUV peak was moderately associated with metabolic response in univariate analysis. It has been established that SUV max measured on baseline PET-CT is predictive of long-term progression-free and overall survival, particularly in DLBC lymphomas,<sup>38</sup> since it reflects tumor aggressiveness and propensity to relapse. There is, however, no evidence that baseline SUV max or SUV peak correlates with tumor sensitivity to first-line chemotherapy.

Several studies have clearly demonstrated the influence of total MTV and TLG on progression-free survival.<sup>36,38,39</sup> In our particular case tumor MTV and TLG were significantly associated with early metabolic response, although they were only representative of the metabolic burden of the bulky mass under consideration, and not of the total metabolic burden of the disease. Optimal cut-off values for MTV are known to be dependent on the study population and the method of MTV measurement.<sup>39</sup> The optimal MTV cut-off around 600 mL observed in our population was similar to that of 550 mL proposed by Sasanelli et al. for progression-free and overall survival prediction in DLBCL.<sup>36</sup> The lower optimal MTV cutoff of 400 mL recently reported by Mikhaeel et al. on a large cohort of DLBCL patients was likely consecutive to the tumor segmentation method that relied on absolute SUV thresholding.<sup>39</sup>

#### 4.B. Textural indices

Textural parameters based on the gray-level co-occurrence matrix were variably correlated with tumor volume, as discussed elsewhere.<sup>34</sup> Our results may appear somewhat different from previously reported correlation degrees, likely due to the fact that correlation with MTV tends to recede for increasing volumes.<sup>31</sup> Among the GLCM-derived indices, contrast was the only independent predictor of metabolic response, with higher contrast values associated with better responses. Dissimilarity (which is a GLCM-derived index akin to contrast) has already been highlighted as a good predictor of therapeutic response and long-term survival in esophageal and non-small-cell lung cancers.<sup>33,40</sup>

To our knowledge, this is the first attempt to evaluate the prognostic value of PSD-related indices in oncological PET. Their performances should therefore be interpreted with caution and deserve further assessment on a broader range of tumor types. PSD AUC was naturally negatively correlated with tumor volume since an increase in object size translates into a narrower spectral distribution. AC distance was positively correlated with tumor volume due to its duality with PSD. Granularity was by definition negatively correlated with tumor volume. After univariate and multivariate analyses,

both PSD AUC and granularity appeared as independent predictors of early metabolic response. Tumors with higher granularity, i.e., fine-grained tumors with higher global heterogeneity, corresponded to a lower probability of complete response. One explanation for the association between higher PSD AUCs and better metabolic responses is that the high-frequency content of the spectrum essentially accounts for tumor edges, with sharp and well-defined edges resulting in heavy-tailed PSDs with higher AUC. It might be hypothesized that fine-grained overall texture and crumbled edges reveal histopathological features such as cellular heterogeneity, areas of hypoxia or necrosis, and patchy vascularity, which would negatively affect tumoral response to chemotherapy.

These findings are in line with recent results showing the potential interest of textural features extracted from morphological images to assess disease status and risk-stratify patients with Hodgkin and non-Hodgkin lymphomas. Texture analysis of MR images has been helpful for chemotherapy response monitoring in NHL.<sup>28</sup> Contrast-enhanced CT textural features proved useful in identifying nodal residual disease in HL.<sup>26</sup> More importantly, Ganeshan et al. recently demonstrated that heterogeneity metrics extracted from pre-treatment CT provide prognostic information that is complementary to interim PET-CT and that these metrics are independent predictors of progression-free survival in both HL and NHL.<sup>27</sup>

#### 4.C. Morphological indices

Quantitative morphological characterization of tumoral lesions, essentially through fractal analysis, has been shown to bring additional diagnostic and prognostic value in the areas of computed tomography<sup>41,42</sup> and magnetic resonance imaging.<sup>43</sup> Morphological analysis of PET images postulates that <sup>18</sup>F<sub>2</sub> distribution, after appropriate segmentation, reflects the actual boundaries of the metabolically active tumoral tissue, the produced indices rendering the morphological complexity and heterogeneity of the tumor. Among the shape parameters evaluated in the present study, surfacic extension, and 2D and 3D fractal dimensions were independent predictors of early metabolic response. The surfacic extension depicted how, for a given surface, the tumoral volume spreads over space. Higher surfacic extension was associated with better metabolic response, possibly due to wider peripheral vascularization and hence greater exposure of the tumor shell to chemotherapeutic agents. The 2D fractal dimension reflects the global complexity of the tumor surface by estimating how details on the surface change with the scale at which it is measured, with its value tending toward 2 for smooth surfaces. The 3D fractal dimension accounts for both surface complexity and intratumoral uptake heterogeneity, with its value tending toward 3 for smooth surfaces with smooth tracer distribution. Lower values of both the 2D and 3D fractal dimensions were significantly associated with higher rates of complete metabolic response. One can reasonably infer that fractal dimensions reflect histological patterns

and tumoral perfusion distribution at the macroscopic scale, with cellular and perfusion homogeneity being favorable factors for therapeutic response. These findings are consistent with previously reported results using perfusion CT showing a higher fractal dimension in malignant colic lesions compared with normal colon, with a higher fractal dimension in low perfusion areas,<sup>44</sup> and better response of hepatocellular carcinoma to anti-angiogenic therapy for lesions with lower baseline fractal dimension.<sup>41</sup>

#### 4.D. Impact of segmentation threshold and image resolution

Phantom and patient studies showed that a threshold at 41% of maximal SUV was optimal for tumor volume estimation.<sup>29</sup> This value was derived using older generation PET-CT systems without TOF acquisition or PSF correction and have not been optimized for textural analysis. Due to spatial resolution considerations, a non-negligible part of the metabolic information related to the tumor is expected to lie outside the limits of the 41% thresholded VOI. The information loss is even more critical when a connectivity constraint is imposed on the tumor boundaries, since the higher the threshold value the higher the probability of yielding fragmented VOIs. In this study, the aim was not to achieve accurate volume measurements but rather to assess the prognostic relevance of texture and shape parameters; therefore, a low threshold value of 30% was deliberately chosen to systematically encompass the entire metabolic area pertaining to the tumor. The derived metrics were not those of the morphological tumor volume but those of its metabolic rendering using <sup>18</sup>FDG PET. The results presented in Tables II and III clearly show that the texture and shape indices extracted from the 30% thresholded VOIs, although highly correlated with those obtained using the 41% threshold (except for the 2D fractal dimension), were better predictors of early metabolic response. It has already been pointed out that the variability in the textural feature range depends on the segmentation method employed.<sup>34</sup> The accuracy of shape metrics such as fractal dimension extracted from metabolic images still needs to be assessed and deserves further investigation using phantom experiments.

The quantification of heterogeneity is dependent on PET data resolution, and PSF-corrected images are known to provide parameters with wider ranges of variation than noncorrected images. This is due to better resolution recovery and a slightly lower signal to noise ratio, which directly affect the textural information contained in the image.<sup>45,46</sup> Our results tend to demonstrate that the textural features extracted from smoothed images remain highly correlated with those extracted from nonsmoothed images and that their prognostic pertinence is robust to image resolution in a range of low to moderate postsmoothing levels. Shape parameters were significantly impacted by postsmoothing due to the close relationship between image resolution and the regularity of tumor boundaries induced by the segmentation method. In particular, the 2D fractal

dimension, as a marker of the complexity of tumor boundaries, lost its prognostic value when computed on strongly postsmoothed images.

#### 4.E. SVM predictive models

Predictive models including MTV and MTV + histology achieved globally similar performances in terms of ROC AUC and balanced accuracy. High MTV was very specific of an incomplete metabolic response, in agreement with recent results indicating that baseline MTV is a strong predictor of overall and progression-free survival in diffuse large B cell lymphoma.<sup>36,39</sup> Accounting for histological type resulted in slightly higher accuracy related to the high sensitivity of histology to poor metabolic response. This can mainly be explained by the observation that most of the HL patients were good responders, whereas response was more variable among the NHL patients. When the prognostic performances of textural and shape indices was examined, the combination of parameters that yielded the highest added predictive value was contrast, PSD AUC, surfacic extension, and 2D and 3D fractal dimensions. In cross-validation, the predictive accuracy reached 80%, with sensitivity above 90% and specificity around 70%. In particular, specificity was substantially higher than that provided by MTV + histology.

Prognosis assessment is of utmost importance for therapeutic adjustment and planning for monitoring in Hodgkin and non-Hodgkin lymphomas. Patient follow-up is ongoing to evaluate whether pertinent textural or shape indices, in association with usual prognostic factors like baseline tumor volume and early metabolic response on interim PET-CT, might help to improve the prediction of progression-free and overall survival. Further investigations are required to assess the reproducibility of the proposed metrics as a function of the PET scanner model and acquisition protocol.

### 5. CONCLUSION

These preliminary results of a retrospective study of 57 patients with bulky lymphomatous disease demonstrate that textural and shape analyses conducted on baseline <sup>18</sup>FDG PET-CT scans might be valuable tools for further assessment of tumor aggressiveness and forecasting sensitivity to chemotherapy. In particular, morphological features like surfacic extension, and 2D and 3D fractal dimensions, as potential macroscopic surrogates for histologic heterogeneity and tumoral perfusion, appear to be innovative and promising quantitative features in oncological PET. Most of the computed metrics are sensitive to the segmentation method and image resolution, their association with early response being optimal when high resolution images with PSF correction are considered, provided that a relatively low threshold is used for VOI segmentation. When combined with usual prognostic factors like tumor volume and histology, relevant texture and shape parameters allow for better prediction of early metabolic response. This suggests that texture and shape analyses of bulky lymphomas might eventually provide relevant

information for patient-tailored initial therapy planning and thus deserve further exploration in order to evaluate their ability to predict long-term outcome.

## CONFLICTS OF INTEREST

The authors have no conflicts of interest to disclose.

<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: faybenb@hotmail.com; Telephone: (+33)0467338598.

## REFERENCES

1. Bangert M, Moog F, Buchmann I, et al. Whole-body 2-[<sup>18</sup>F]-fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) for accurate staging of Hodgkin's disease. *Ann Oncol*. 1998;9:1117–1122.
2. Tatsumi M, Kitayama H, Sugahara H, et al. Whole-body hybrid PET with <sup>18</sup>F-FDG in the staging of non-Hodgkin's lymphoma. *J Nucl Med*. 2001;42:601–608.
3. Cheson BD, Fisher RI, Barrington SF, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol*. 2014;32:3059–3068.
4. Mikhaeel NG, Hutchings M, Fields PA, O'Doherty MJ, Timothy AR. FDG-PET after two to three cycles of chemotherapy predicts progression-free and overall survival in high-grade non-Hodgkin lymphoma. *Ann Oncol*. 2005;16:1514–1523.
5. Hutchings M, Loft A, Hansen M, et al. FDG-PET after two cycles of chemotherapy predicts treatment failure and progression-free survival in Hodgkin lymphoma. *Blood*. 2006;107:52–59.
6. Ceriani L, Martelli M, Zinzani PL, et al. Utility of baseline 18FDG-PET/CT functional parameters in defining prognosis of primary mediastinal (thymic) large B-cell lymphoma. *Blood*. 2015;126:950–956.
7. Evens AM, Kostakoglu L. The role of FDG-PET in defining prognosis of Hodgkin lymphoma for early-stage disease. *Blood*. 2014;124:3356–3364.
8. Song MK, Yang DH, Lee GW, et al. High total metabolic tumor volume in PET/CT predicts worse prognosis in diffuse large B cell lymphoma patients with bone marrow involvement in rituximab era. *Leuk Res*. 2016;42:1–6.
9. Danielewicz I, Małkowski B, Zaucha R, Zalewska M, Leśniewski-Kmak K, Zaucha JM. Early treatment intensification with escalated BEACOPP in patients with Hodgkins lymphoma not responding to ABVD therapy. *Acta Oncol*. 2014;53:286–288.
10. Pardal E, Coronado M, Martín A, et al. Intensification treatment based on early FDG-PET in patients with high-risk diffuse large B-cell lymphoma: a phase II GELTAMO trial. *Br J Haematol*. 2014;167:327–336.
11. Fidler IJ. Tumor heterogeneity and the biology of cancer invasion and metastasis. *Cancer Res*. 1978;38:2651–2660.
12. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Van Stiphout RG, Granton P. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012;48:441–446.
13. Chicklore S, Goh V, Siddique M, Roy A, Mardsen PK, Cook GJR. Quantifying tumour heterogeneity in <sup>18</sup>F-FDG PET/CT imaging “by texture analysis. *Eur J Nucl Med Mol Imaging*. 2013;40:133–140.
14. Buvat I, Orlhac F, Soussan M. Tumor texture analysis in PET: where do we stand? *J Nucl Med*. 2015;56:1642–1644.
15. Van Velden FHP, Cheesbumon P, Yaqub M, et al. Evaluation of a cumulative SUV-volume histogram method for parametrizing heterogeneous intratumoral FDG uptake in non-small cell lung cancer PET studies. *Eur J Nucl Med Mol Imaging*. 2011;38:1636–1647.
16. Miwa K, Inubushi M, Wagatsuma K, et al. FDG uptake heterogeneity evaluated by fractal analysis improves the differential diagnosis of pulmonary nodules. *Eur J Radiol*. 2014;83:715–719.
17. Gao X, Chu C, Li Y, et al. The method and efficacy of support vector machine classifiers based on texture features and multi-resolution histogram from <sup>18</sup>F-FDG PET-CT images for the evaluation of mediastinal lymph nodes in patients with lung cancer. *Eur J Radiol*. 2015;84:312–317.
18. Cook GJR, Yip C, Siddique M, Goh V, Chicklore S, Roy A. Are pre-treatment <sup>18</sup>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.
19. Soussan M, Orlhac F, Boubaya M, et al. Relationship between tumor heterogeneity measured on FDG-PET/CT and pathological prognostic factors in invasive breast cancer. *PLoS ONE*. 2014;9:e94017.
20. Tixier F, Le Rest CC, Hatt M, et al. Intratumor heterogeneity characterized by textural features on baseline 18f-fdg pet images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med*. 2011;52:369–378.
21. Yang F, Thomas MA, Dehdashti F, Grigsby PW. Temporal analysis of intratumoral metabolic heterogeneity characterized by textural features in cervical cancer. *Eur J Nucl Med Mol Imaging*. 2013;40:716–727.
22. Son SH, Kim DH, Hong CM, et al. Prognostic implication of intratumoral metabolic heterogeneity in invasive ductal carcinoma of the breast. *BMC Cancer*. 2014;14:585–594.
23. Yu H, Caldwell C, Mah K, Mozeg D. Coregistered FDG PET/CT-based textural characterization of head and neck cancer for radiation treatment planning. *IEEE Trans Med Imaging*. 2009;28:374–383.
24. Orlhac F, Thézé B, Soussan M, Boisgard R, Buvat I. Multi-scale texture analysis: from 18F-FDG PET images to histologic images. *J Nucl Med*. 2016;57:1823–1828.
25. Lartizien C, Rogez M, Niaf E, Ricard F. Computer-aided staging of lymphoma patients with FDG PET/CT imaging based on textural information. *IEEE J Biomed Health Inform*. 2014;18:946–955.
26. Knogler T, El-Rabadi K, Weber M, Karanikas G, Mayerhoefer ME. Three-dimensional texture analysis of contrast enhanced CT images for treatment response assessment in Hodgkin lymphoma: comparison with F-18-FDG PET. *Med Phys*. 2014;41:121904.
27. Ganeshan B, Miles KA, Babikir S, et al. CT-based texture analysis potentially provides prognostic information complementary to interim fdg-pet for patients with hodgkin's and aggressive non-hodgkin's lymphomas. *Eur Radiol*. 2017;27:1012–1020.
28. Harrison LC, Luukkaala T, Pertovaara H, et al. Non-Hodgkin lymphoma response evaluation with MRI texture classification. *J Exp Clin Cancer Res*. 2009;28:87.
29. Meignan M, Sasanelli M, Casasnovas RO, et al. Metabolic tumour volumes measured at staging in lymphoma: methodological evaluation on phantom experiments and patients. *Eur J Nucl Med Mol Imaging*. 2014;41:1113–1122.
30. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern*. 1973;3:610–621.
31. Orlhac F, Soussan M, Chouahnia K, Martinod E, Buvat I. 18F-FDG PET-derived textural indices reflect tissue-specific uptake pattern in non-small cell lung cancer. *PLoS ONE*. 2015;10:e0145063.
32. Vapnik VN. An overview of statistical learning theory. *IEEE Trans Neural Netw*. 1999;10:988–999.
33. Hatt M, Majdoub M, Vallières M, et al. <sup>18</sup>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.
34. Orlhac F, Soussan M, Maisonneuve JA, Garcia CA, Vanderlinden B, Buvat I. Tumor texture analysis in <sup>15</sup>F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis. *J Nucl Med*. 2014;55:414–422.
35. Lee H, Paeng JC, Hong SH, et al. Appropriate margin thresholds for isocontour metabolic volumetry of fluorine-18 fluorodeoxyglucose PET in sarcoma: a hybrid PET/MRI study. *Nucl Med Commun*. 2016;37:1088–1094.
36. Sasanelli M, Meignan M, Haioun C, et al. Pretherapy metabolic tumour volume is an independent predictor of outcome in patients with diffuse large B-cell lymphoma. *Eur J Nucl Med Mol Imaging*. 2014;41:2017–2022.
37. Ngeow JY, Quek RH, Ng DC, et al. High SUV uptake on FDG-PET/CT predicts for an aggressive B-cell lymphoma in a prospective study of

- primary FDG-PET/CT staging in lymphoma. *Ann Oncol.* 2009;20:1543–1547.
38. Xie M, Wu K, Liu Y, Jiang Q, Xie Y. Predictive value of F-18 FDG PET/CT quantization parameters in diffuse large B cell lymphoma: a meta-analysis with 702 participants. *Med Oncol.* 2015;32:446.
  39. Mikhaeel NG, Smith D, Dunn JT, et al. Combination of baseline metabolic tumour volume and early response on PET/CT improves progression-free survival prediction in DLBCL. *Eur J Nucl Med Mol Imaging.* 2016;43:1209–1219.
  40. Lovinfosse P, Janvier 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;43:1453–1460.
  41. Hayano K, Lee SH, Yoshida H, Zhu AX, Sahani DV. Fractal analysis of CT perfusion images for evaluation of antiangiogenic treatment and survival in hepatocellular carcinoma. *Acad Radiol.* 2014;21:654–660.
  42. Jauhari RK, Trivedi R, Munshi P, Sahni K. Fractal characterization of brain lesions in CT images. *Med Phys.* 2005;32:3661–3665.
  43. Smitha KA, Gupta AK, Jayasree RS. Fractal analysis: fractal dimension and lacunarity from MR images for differentiating the grades of glioma. *Phys Med Biol.* 2015;60:6937–6947.
  44. Goh V, Sanghera B, Wellsted DM, Sundin J, Halligan S. Assessment of the spatial pattern of colorectal tumour perfusion estimated at perfusion CT using two-dimensional fractal analysis. *Eur Radiol.* 2009;19:1358–1365.
  45. Yan J, Chu-Shern JL, Loi HY, et al. Impact of image reconstruction settings on texture features in 18FFDG PET. *J Nucl Med.* 2015;56:1667–1673.
  46. Lasnon C, Majdoub M, Lavigne B, et al. 18F-FDG PET/CT heterogeneity quantification through textural features in the era of harmonisation programs: a focus on lung cancer. *Eur J Nucl Med Mol Imaging.* 2016;43:2324–2335.