European Radiology
https://doi.org/10.1007/s00330-020-06815-8

NUCLEAR MEDICINE

Baseline ^{18}F-FDG PET radiomic features as predictors of 2-year event-free survival in diffuse large B cell lymphomas treated with immunochemotherapy

Nicolas Aide[1,2,3] · Christophe Fruchart[4] · Catherine Nganoa[1] · Anne-Claire Gac[5] · Charline Lasnon[3,6]

Received: 9 December 2019 / Revised: 27 February 2020 / Accepted: 16 March 2020

Abstract

Objectives To explore the prognostic value of positron emission tomography (PET) radiomic features in the field of diffuse large B cell lymphoma (DLBCL) treated with a first-line immunochemotherapy.

Methods One-hundred thirty-two patients newly diagnosed with DLBCL were retrospectively included. PET studies were reconstructed using an ordered subset expectation maximisation algorithm with point spread function modelling. The total metabolic tumour volume (MTV) was recorded for each patient, and the volume of interest structure of the largest target lesion was used to compute ^{18}F-FDG textural parameters. Data was randomly split into training and validation datasets. Optimal cutoff values were determined by means of 2-year event-free survival (EFS) ROC analyses. Two-year EFS analyses were performed using Kaplan-Meier survival analyses and univariable and multivariable Cox regression models.

Results The median follow-up was 27 months, and the 2-year event-free survival (2y-EFS) was 77.3% in the entire population. ROC analyses for the 2y-EFS reached statistical significance for total MTV as well as four second-order metrics (homogeneity, contrast, correlation, dissimilarity) and five third-order metrics (LZE (Long-Zone Emphasis), LZLGE (Long-Zone Low-Grey Level Emphasis), LZHGE (Long-Zone High-Grey Level Emphasis), GLNU (Grey-Level Non-Uniformity) and ZP (Zone Percentage)). LZHGE displayed the highest ROC analysis accuracy (acc. = 0.76) and the best discriminant value on univariable Kaplan-Meier analysis ($p < 0.0001$, HR = 4.54). On multivariable analysis, including IPIaa, total MTV and LZHGE, LZHGE was the only independent predictor of 2y-EFS. These results were confirmed on the validation dataset.

Conclusions Baseline ^{18}F-FDG PET heterogeneity of the largest lymphoma lesion is a promising predictor of 2y-EFS in newly diagnosed DLBCL treated with immunochemotherapy.

Key Points

- *^{18}F-FDG metabolic heterogeneity emerges as a new tool for survival prognostication of patients and has been explored in many solid tumours with promising results.*
- *Baseline ^{18}F-FDG PET heterogeneity of the largest lymphoma lesion is an independent predictor of 2y-EFS in newly diagnosed DLBCL treated with immunochemotherapy.*
- *DLBCL patients presenting with a heterogeneous tumour displayed a worse prognosis.*

Keywords Lymphoma, large B cell, diffuse · Positron emission tomography computed tomography · Fluorodeoxyglucose F18 · Image processing, computer-assisted

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-020-06815-8) contains supplementary material, which is available to authorized users.

✉ Charline Lasnon
c.lasnon@baclesse.unicancer.fr

[1] Nuclear Medicine Department, Caen University Hospital, Caen, France

[2] Normandie University, Caen, France

[3] INSERM 1086 ANTICIPE, Normandie University, Caen, France

[4] Haematology Institute, François Baclesse Cancer Centre, Caen, France

[5] Haematology Institute, Caen University Hospital, Caen, France

[6] Nuclear Medicine Department, François Baclesse Cancer Centre, 3 Avenue du Général Harris, BP 45026, 14076 Caen Cedex 5, France

Published online: 04 April 2020

Abbreviations

^{18}F-FDG	^{18}Fluoro-deoxy-glucose
AUC	Area under the curve
CNIL	National Commission on Informatics and Liberty
DLBCL	Diffuse large B cell lymphoma
EFS	Event-free survival
GLNU	Grey-level non-uniformity
GLSZM	Grey-level size zone matrix
HGZE	High grey-level zone emphasis
IPIaa	Age-adjusted international prognostic index
LGZE	Low grey-Level Zone Emphasis
LZE	Long-Zone Emphasis
LZHGE	Long-Zone High Grey-level Emphasis
LZLGE	Long-Zone Low Grey-level Emphasis
MTV	Metabolic tumour volume
NHL	Non-Hodgkin's lymphoma
OS	Overall survival
OSEM	Ordered subset expectation maximisation
PET	Positron emission tomography
PSF	Point spread function
ROC	Receiver operating characteristic
SD	Standard deviation
SUV	Standardised uptake value
SZE	Short-Zone Emphasis
SZHGE	Short-Zone High Grey-level Emphasis
SZLGE	Short-Zone Low Grey-level Emphasis
VOI	Volume of interest
ZLNU	Zone length non-uniformity
ZP	Zone percentage

Introduction

Most lymphomas are ^{18}fluoro-deoxy-glucose (^{18}F-FDG) avid, and this avidity for ^{18}F-FDG concerns greater than 90% of Hodgkin lymphomas, diffuse large cell lymphomas and follicular lymphomas, which collectively account for 70% of all lymphomas [1]. DLBCL is part of non-Hodgkin's lymphoma (NHL) and is the most common NHL (40% of all cases of this type of disease) [2]. Currently, ^{18}F-FDG is commonly used for initial staging of DLBCL, as it has been demonstrated that performing a PET scan at diagnosis changed the disease stage in 8 to 32% of cases and the therapeutic decision in 5 to 25% of cases [3]. Moreover, baseline SUV_{max} and [4, 5] total metabolic tumour volume (MTV) are prognosticators in lymphoma [3, 6]. However, these parameters are not yet used in clinical practice and not integrated in haematological guidelines because prospective studies with larger cohort of patients and methodological harmonisation [7] are needed to define robust and worldwide applicable cutoff values. More recently, ^{18}F-FDG metabolic heterogeneity, a new PET parameter, has emerged as a new tool for survival prognostication of patients. Also referred to as textural features, this type of quantitative ^{18}F-FDG PET data has been explored in many solid tumours with promising results [8–12]. Few data are available to date in the field of lymphoma even though it is known to be a heterogeneous disease with multiple histopathologic, phenotypic and genetic patterns leading to different outcomes [13]. In a retrospective study of 57 patients with bulky lymphomatous disease, Ben Bouallègue [14] et al demonstrated that textural and shape analyses conducted on baseline ^{18}F-FDG PET-CT scans might be valuable tools for further assessment of tumour aggressiveness and forecasting sensitivity to chemotherapy. More recently, in a retrospective study of 82 patients, Parvez et al [15, 16] found that tumour texture features could not predict therapy response, whereas kurtosis correlated with overall survival (OS) with a one-unit increase associated with a 67% increase in expected risk of death. However, these studies were conducted on non-homogeneous populations in which patients presented different histological subtypes of aggressive NHL. In addition, in the study by Parvez et al with OS as end-point, only 6 events occurred in a population of 82 patients, which calls into question the robustness of the results. Therefore, we aimed to further support these results on a histologically homogeneous population of 132 patients newly diagnosed with DLBCL using 2-year event-free survival (2y-EFS) as a gold standard, which is a robust end-point for disease-related outcomes in DLBCL treated with immunochemotherapy.

Materials and methods

Population

All patients diagnosed with diffuse large B cell lymphoma referred to our Haematology Institute from December 2008 to December 2017 were retrospectively included. All procedures performed in this study were conducted according to the principles expressed in the Declaration of Helsinki. We sought approval to collect data for our study from the national committee for data privacy and the National Commission on Informatics and Liberty (CNIL) with registration no. 2081250 v 0. Informed consent was obtained from all patients. Age, gender and IPIaa were recorded for each patient. Cell of origin according to the immunohistochemistry-based Hans algorithm [17] was sought for all patients as well as CD5, BCL2 and MYC expressions because of their prognostic value [18] according to the WHO classification of DLBCL.

PET acquisition and reconstruction parameters

After a 15-min rest in a warm room, patients who had been fasting for 6 h were injected intravenously with ^{18}F-FDG. The injected dose, capillary glycaemia at the injection time and the

exact delay between injection and the start of the acquisition were recorded for each patient.

All PET imaging studies were performed on a Biograph TrueV (Siemens Healthineers) with a 6-slice spiral CT (computed tomography) component. CT acquisition was performed with the following parameters: 60 mAs, 130 kVp, pitch 1 and 6 × 2 mm collimation. Subsequently, PET emission acquisition was performed with 2 min 40 s and 3 min 40 s per bed position for low- to normal-weight patients and overweight to obese patients, respectively. Patients were scanned from the skull to the mid-thighs. All examinations were reconstructed using an OSEM algorithm with point spread function (PSF) modelling (HD; TrueX, Siemens Healthineers) with three iterations and 21 subsets without filtering. The matrix size was 168 × 168 voxels, resulting in isotropic voxels of 4.07 × 4.07 × 4.07 mm^3. Scatter and attenuation corrections were applied.

PET tumour delineation

The metabolic tumour volume (MTV) protocol of LifeX software (www.lifexsoft.org) [19] was used to standardise the procedure amongst three board-certified nuclear physicians who performed the images analysis. Initial VOI ranges were maintained as default: 2.2 SUV units and 0.5 mL with no upper limits. A VOI of 100 cc was drawn on the right liver lobe to determine the liver SUV_{max} value. Then, a SUV/liver activity ratio method was applied with a ratio set to 1.0. All non-lymphoma ^{18}F-FDG uptakes were manually excluded. The total MTV was recorded for each patient, and the VOI structure of the largest target lesion was used to further compute ^{18}F-FDG textural parameters using the texture protocol of LifeX software (www.lifexsoft.org). Lesions < 10 cc were not considered [20]. Three board-certified PET readers conducted the analysis.

The following PET radiomic features were extracted:

- Conventional PET parameters: SUV_{max}, total MTV
- Grey level co-occurrence matrix (GLCM) parameters: homogeneity, energy, contrast, correlation, entropy_log10 and dissimilarity
- Grey-level size zone matrix (GLSZM) parameters: SZE (Short-Zone Emphasis), LZE (Long-Zone Emphasis), LGZE (Low Grey-level Zone Emphasis), HGZE (High Grey-Level Zone Emphasis), SZLGE (Short-Zone Low Grey-level Emphasis), SZHGE (Short-Zone High Grey-level Emphasis), LZLGE (Long-Zone Low Grey-level Emphasis), LZHGE (Long-Zone High Grey-level Emphasis), GLNU (Grey-Level Non-Uniformity for zone) ZLNU (Zone Length Non-Uniformity) and ZP (Zone Percentage).

All textural features were compliant with the benchmark of the image biomarkers standardisation initiative [21]. Absolute resampling using 64 bins between 0 and the maximum SUV value recorded in the entire population was used [22, 23].

Statistical analysis

Quantitative data are presented as the mean ± standard deviation (SD), otherwise specified. Clinical and PET characteristics of all patients with no events at 2 years (2y-EFS 0) and patients who experienced an event at 2 years (2y-EFS 1) were compared using chi-square tests or Mann-Whitney tests as appropriate. For EFS, the end-point was defined as the time from diagnosis until relapse or progression, unplanned retreatment of lymphoma or death as a result of lymphoma. Data was randomly split into 2 groups: 80% of patients were attributed to the training dataset and 20% remaining patients to the validation dataset. Receiver operating characteristic (ROC) curves for 2-year event-free survival (EFS) were generated on the training dataset to define area under the curve (AUC) and optimal cutoff values of each metric by selecting the best Youden index. Two-year event-free survival analyses were then performed using univariable Kaplan-Meier survival analyses with log-rank tests to compare survival curves, as well as univariable and multivariable Cox regression models. Finally, 2-year event-free survival analyses were also performed using univariable Kaplan-Meier survival analyses on the validation dataset, using the predefined cutoff values.

Graphs and statistical analysis were performed using XLSTAT Software (XLSTAT 2019: Data Analysis and Statistical Solution for Microsoft Excel. Addinsoft). For all statistical tests, a two-tailed p value of less than 0.05 was considered statistically significant.

Results

Patient and PET characteristics

Amongst 152 included patients, 132 were analysed: four patients had missing data, one patient had diffuse liver involvement, one patient exhibited diffuse inflammatory post-surgery hypermetabolisms, three patients had only small lesions < 10 cc and eleven patients had no initial target. One hundred and eighteen patients (89.4%) were treated with R-CHOP, and 14 patients (10.6%) were treated with R-ACVBP. The median follow-up was 27 months, and the 2y-EFS was 77.3% in the entire population. There were 30 recorded events in the entire population over the first 2 years of follow-up. Patient's clinical and PET characteristics are displayed in Table 1. A significant difference in age-adjusted international prognostic index (IPIaa) scores was identified between patients with no events at 2 years (2y-EFS 0) and patients who experienced an event

Table 1 Patient clinical and PET characteristics for the entire population and according to their 2-year EFS status

Parameters	Whole population ($n = 132$)	2y-EFS 0* ($n = 102$)	2y-EFS 1** ($n = 30$)	p value
Age (year), mean (range)	62 (20–81)	61 (20–81)	63 (23–78)	0.44[ǂ]
Gender, n (%)				0.68[†]
Female	71 (53.8)	56 (54.9)	15 (50.0)	
Male	61 (46.2)	46 (45.1)	15 (50.0)	
IPIaa, n (%)				0.01[†]
0–1	59 (44.7)	52 (51.0)	7 (23.3)	
2–3	73 (55.3)	50 (49.0)	23 (76.7)	
Injected dose (MBq), mean (SD)	290 (69)	295 (70)	275 (65)	0.14[ǂ]
Glycaemia (g/L), mean (SD)	1.04 (0.26)	1.06 (0.28)	0.97 (0.14)	0.28[ǂ]
^{18}F-FDG Uptake time (min), mean (SD)	61 (5)	60 (5)	62 (4)	0.05[ǂ]

*2y-EFS 0: patients with no event at 2 years

**2y-EFS 1: patients who experienced an event at 2 years

Statistical tests used: [ǂ] Mann-Whitney test, [†] Chi-square tests

at 2 years (2y-EFS 1) ($p = 0.035$). No significant difference in other clinical characteristics was observed (Table 1). On average, the VOI drawn on the largest lesion covered 77% of the total MTV. The correlation between total MTV and the VOI drawn on the largest lesion was very good: $\rho = 0.88$ and $p < 0.0001$.

Baseline PET metrics of interest determined on the training dataset (105 patients)

ROC analyses for 2y-EFS reached statistical significance for total MTV as well as four second-order metrics and five third-order metrics: homogeneity, contrast, correlation, dissimilarity, LZE, LZLGE, LZHGE, GLNU and ZP (Table 2). Interestingly, SUV_{max} ROC analysis did not reach significance and therefore, SUV_{max} was not further explored. Optimal cutoff values for total MTV, homogeneity, contrast, correlation, dissimilarity, LZE, LZLGE, LZHGE, GLNU and ZP were 111.0; 0.55; 3.76; 0.60; 1.39; 24,829.55; 1873.01; 1,264,925.92; 21.19 and 0.08 with area under the curve (AUC) equal to 0.66, 0.62, 0.62, 0.64, 0.62, 0.67, 0.64, 0.69, 0.64 and 0.64, respectively. The parameter displaying the highest accuracy (acc. = 0.76) and area under the curve (AUC = 0.69) was LZHGE. All ROC analysis results with 2y-EFS as reference standard are displayed in Table 2.

There were 13 patients (12.4%) presenting tumour bulk defined as a single nodal mass of 6 cm or greater than the third of the transthoracic diameter at any level of thoracic vertebrae on CT [3, 24, 25]. There were no significant differences in homogeneity, contrast, correlation, dissimilarity, LZE, LZLGE, LZHGE, GLNU and ZP between bulky and non-bulky patients (Fig. 1).

Cell of origin according to the immunohistochemistry-based Hans algorithm [17] could be determined in 86 patients (81.9%): 52 patients were assigned as germinal centre derived B cells (GCB) and 34 as activated B cells (ABC). There were no significant differences observed between these two groups of patients for total MTV, homogeneity, contrast, correlation, dissimilarity, LZE, LZLGE, LZHGE, GLNU and ZP (Fig. 1). Neither was there any significant difference for these parameters between $CD5^+$ ($n = 12$) and $CD5^-$ ($n = 59$) patients. BCL2 and MYC could not be tested in the same way because of too much missing data.

Training dataset (105 patients) survival analyses

Univariable Kaplan-Meier survival analyses found that total MTV, homogeneity, contrast, correlation, dissimilarity, LZE, LZLGE, LZHGE, GLNU and ZP were all predictors of 2y-EFS. Kaplan-Meier curves of these ten parameters and IPIaa are displayed in Fig. 2. Patients with high total MTV (> 111.0 cc) had a 2y-EFS of 68.5%, whereas patients with low total MTV (≤ 111.0 cc) had a 2y-EFS of 96.4% ($p = 0.005$, HR = 10.21 (95% CI = 1.38–75.48)). Moreover, patients with high LZHGE (> 1,264,925.92) had a 2y-EFS of 49.5%, whereas patients with low LZHGE (≤ 1,264,925.92) had a 2y-EFS of 86.6% ($p < 0.0001$, HR = 4.54 (95% CI = 2.03–10.14)). Table 3 displays Cox univariable analyses of all ten parameters of interest and IPIaa. Of note, all patients ($n = 28$) with low total MTV (≤ 111.0 cc) had low LZHGE values (≤ 1,264,925.92). On the contrary, amongst the 77 patients with large total MTV (> 111.0 cc), 36 (46.8%) displayed low LZHGE values (≤ 1,264,925.92) and 41 (53.2%) high LZHGE values (> 1,264,925.92).

Table 2 ROC analyses for 2-year event-free survival

PET variable		Value	Se	Lower bound (95%)	Upper bound (95%)	Sp	Lower bound (95%)	Upper bound (95%)	Accuracy	AUC	Lower bound (95%)	Upper bound (95%)	*p* value
Basic	SUV_{max}	< 32.21	0.72	0.52	0.86	0.49	0.38	0.59	0.54	0.56	0.44	0.68	0.30
	MTV_{total}	*> 111.00*	*0.96*	*0.79*	*1.00*	*0.34*	*0.24*	*0.45*	*0.49*	*0.66*	*0.55*	*0.78*	*0.01*
GLCM	Homogeneity	*> 0.55*	*0.76*	*0.56*	*0.89*	*0.55*	*0.44*	*0.65*	*0.60*	*0.62*	*0.51*	*0.74*	*0.04*
	Energy	> 0.03	0.68	0.48	0.83	0.59	0.48	0.69	0.61	0.58	0.46	0.70	0.17
	Contrast	*< 3.76*	*0.72*	*0.52*	*0.86*	*0.56*	*0.45*	*0.67*	*0.60*	*0.62*	*0.50*	*0.73*	*0.04*
	Correlation	*> 0.60*	*0.60*	*0.4*	*0.77*	*0.70*	*0.59*	*0.79*	*0.68*	*0.64*	*0.51*	*0.77*	*0.03*
	Entropy	< 1.81	0.76	0.56	0.89	0.46	0.36	0.57	0.53	0.58	0.46	0.70	0.21
	Dissimilarity	*< 1.39*	*0.76*	*0.56*	*0.89*	*0.53*	*0.42*	*0.63*	*0.58*	*0.62*	*0.50*	*0.73*	*0.04*
GLDZM	SZE	> 0.49	0.60	0.41	0.77	0.58	0.47	0.68	0.58	0.59	0.46	0.72	0.18
	LZE	*> 24,829.55*	*0.72*	*0.52*	*0.86*	*0.69*	*0.58*	*0.78*	*0.70*	*0.67*	*0.55*	*0.79*	*0.01*
	LGZE	> 0.07	0.64	0.44	0.80	0.59	0.48	0.69	0.60	0.59	0.46	0.71	0.19
	HGZE	< 80.83	1.00	0.84	1.00	0.16	0.10	0.26	0.36	0.55	0.43	0.68	0.39
	SZLGE	> 0.04	0.64	0.44	0.80	0.66	0.55	0.76	0.66	0.62	0.49	0.75	0.06
	SZHGE	< 31.48	0.80	0.60	0.91	0.38	0.28	0.48	0.48	0.56	0.44	0.68	0.34
	LZLGE	*> 1873.01*	*0.68*	*0.48*	*0.83*	*0.69*	*0.58*	*0.78*	*0.69*	*0.64*	*0.51*	*0.77*	*0.04*
	LZHGE	*> 1,264,925.92*	*0.60*	*0.41*	*0.77*	*0.81*	*0.71*	*0.88*	*0.76*	*0.69*	*0.57*	*0.81*	*0.002*
	GLNU	*> 21.19*	*0.96*	*0.79*	*1.00*	*0.31*	*0.22*	*0.42*	*0.47*	*0.64*	*0.52*	*0.76*	*0.02*
	ZLNU	27.61	0.84	0.65	0.94	0.36	0.27	0.47	0.48	0.59	0.47	0.71	0.16
	ZP	*< 0.08*	*0.84*	*0.65*	*0.94*	*0.44*	*0.33*	*0.55*	*0.53*	*0.64*	*0.52*	*0.76*	*0.02*

Italicized lines correspond to statistically significant results

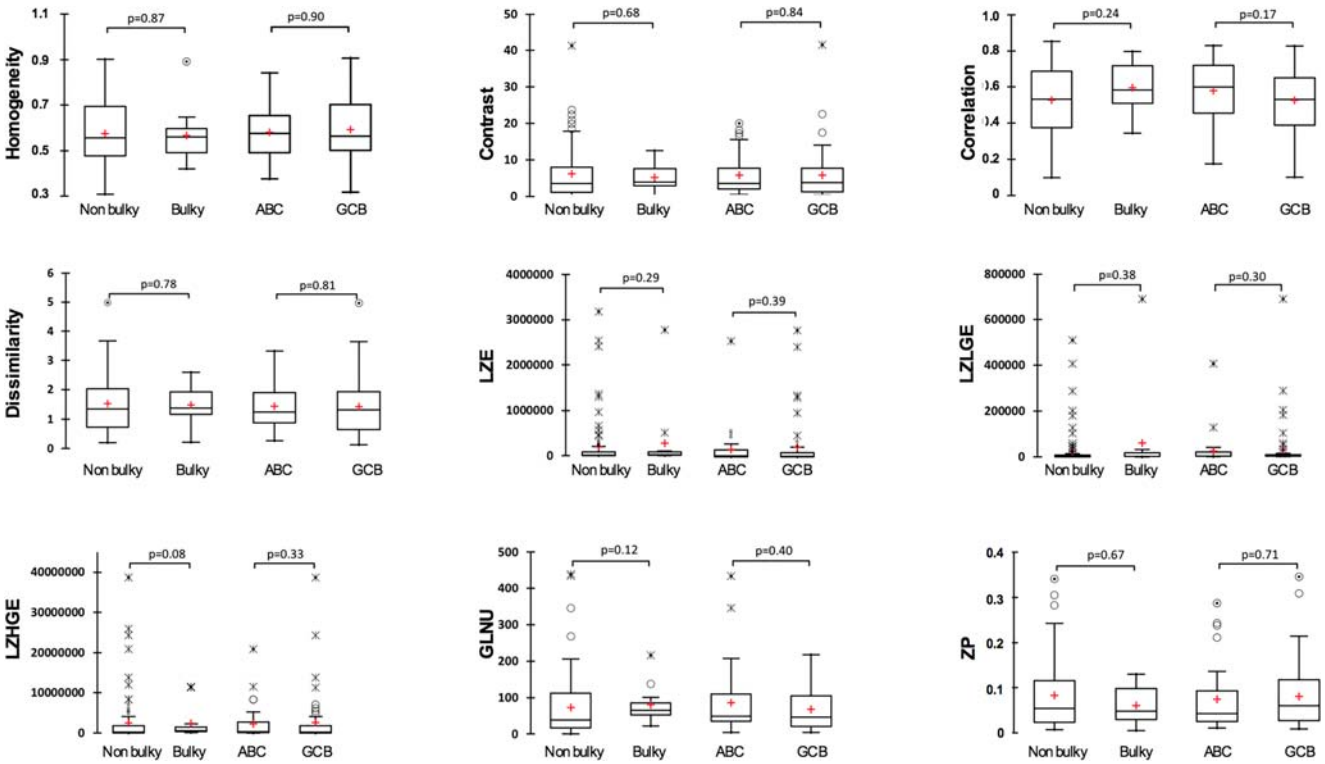

Fig. 1 Comparison of textural features between non-bulky versus bulky patients and between ABC versus GCB patients. The red crosses correspond to the means. The central horizontal bars are the medians. The lower and upper limits of the box are the first and third quartiles, respectively. Points above or below the whiskers' upper and lower bounds may be considered as outliers. Mann-Whitney tests were used

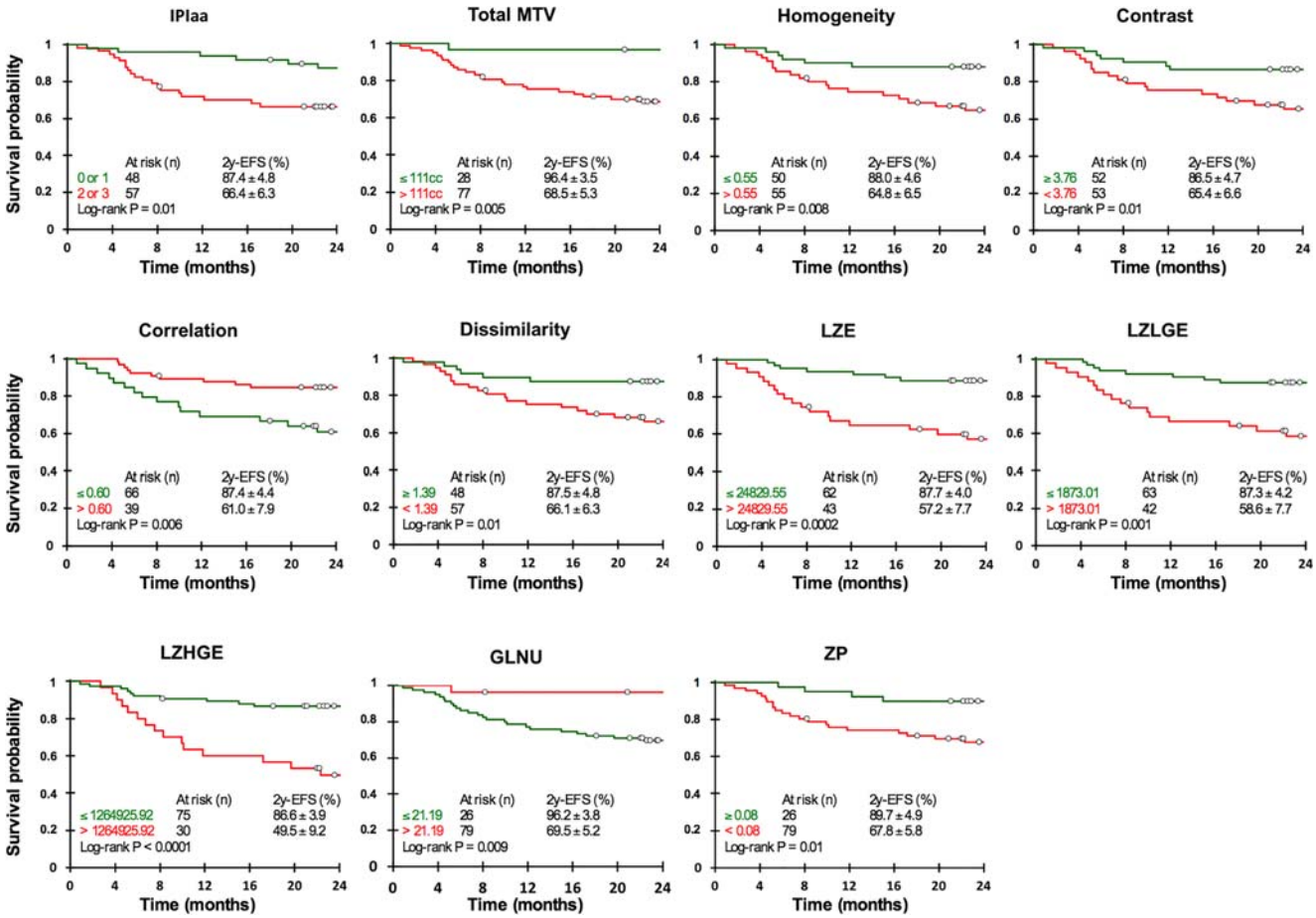

Fig. 2 Kaplan-Meier estimates of 2-year event-free survival (EFS) in the training dataset according to IPIaa, total metabolic tumour volume (MTV), homogeneity, contrast, correlation, dissimilarity, LZE, LZLGE, LZHGE, GLNU and ZP

For multivariable Cox analysis, known clinical and PET prognostic factors (IPIaa and total MTV) as well as LZHGE that displayed the highest ROC analysis accuracy and the best discriminant value on univariable Kaplan-Meier analysis were included. LZHGE appeared to be the only independent predictor of 2y-EFS (Table 4). Representative images are displayed in Fig. 3 for each group of patients.

Table 3 Cox univariable analyses for 2-year-event-free survival

Variable		Value	Standard error	Wald chi-square	*p* value	Hazard ratio	HR, lower bound (95%)	HR, upper bound (95%)
Clinical	IPIaa 2 or 3	1.15	0.47	6.01	0.01	3.16	1.26	7.91
PET Basic	MTV > 111.00	2.32	1.02	5.18	0.02	10.21	1.38	75.48
PET GLCM	Homogeneity > 0.55	1.18	0.47	6.32	0.01	3.25	1.30	8.14
	Contrast < 3.76	1.04	0.45	5.49	0.02	2.84	1.19	6.81
	Correlation > 0.60	1.08	0.41	6.95	0.01	2.94	1.32	6.54
	Dissimilarity < 1.39	1.09	0.47	5.41	0.02	2.97	1.19	7.45
PET GLDZM	LZE > 24,829.55	1.53	0.44	11.82	0.001	4.64	1.93	11.12
	LZLGE > 1873.01	1.35	0.42	9.89	0.002	3.86	1.66	8.96
	LZHGE > 1,264,925.92	1.51	0.40	13.67	0.0002	4.54	2.03	10.14
	GLNU > 21.19	2.19	1.02	4.61	0.03	8.94	1.21	66.10
	ZP < 0.08	1.29	0.55	5.60	0.02	3.64	1.25	10.60

Table 4 Cox regression analysis for 2-year event-free survival

Variable	Value	Standard error	Wald chi-square	*p* value	Hazard ratio	HR, lower bound (95%)	HR, upper bound (95%)
IPIaa 2 or 3	0.49	0.49	1.01	0.31	1.64	0.63	4.27
MTV > 111.00	1.56	1.08	2.09	0.15	4.76	0.57	39.41
LZHGE > 1,264,925.92	1.04	0.43	5.97	0.01	2.84	1.23	6.56

Validation dataset (27 patients) survival analyses

Using same cutoff values as determined above, Kaplan-Meier survival analysis found that LZHGE was a predictor of 2y-EFS. Patients with high LZHGE (> 1,264,925.92) had a 2y-EFS of 60.0%, whereas patients with low LZHGE (≤ 1,264,925.92) had a 2y-EFS of 94.1% ($p = 0.03$, HR = 7.47 (95% CI = 0.83–66.99)). Kaplan-Meier curve of LZHGE as well as those of IPIaa, total MTV, homogeneity, contrast, correlation, dissimilarity, LZE, LZLGE, GLNU and ZP can be seen in Supplemental Fig. 2.

Discussion

Despite a tremendous interest in PET heterogeneity features, these new metrics will draw the attention of the haematology community and eventually will find their way into clinical

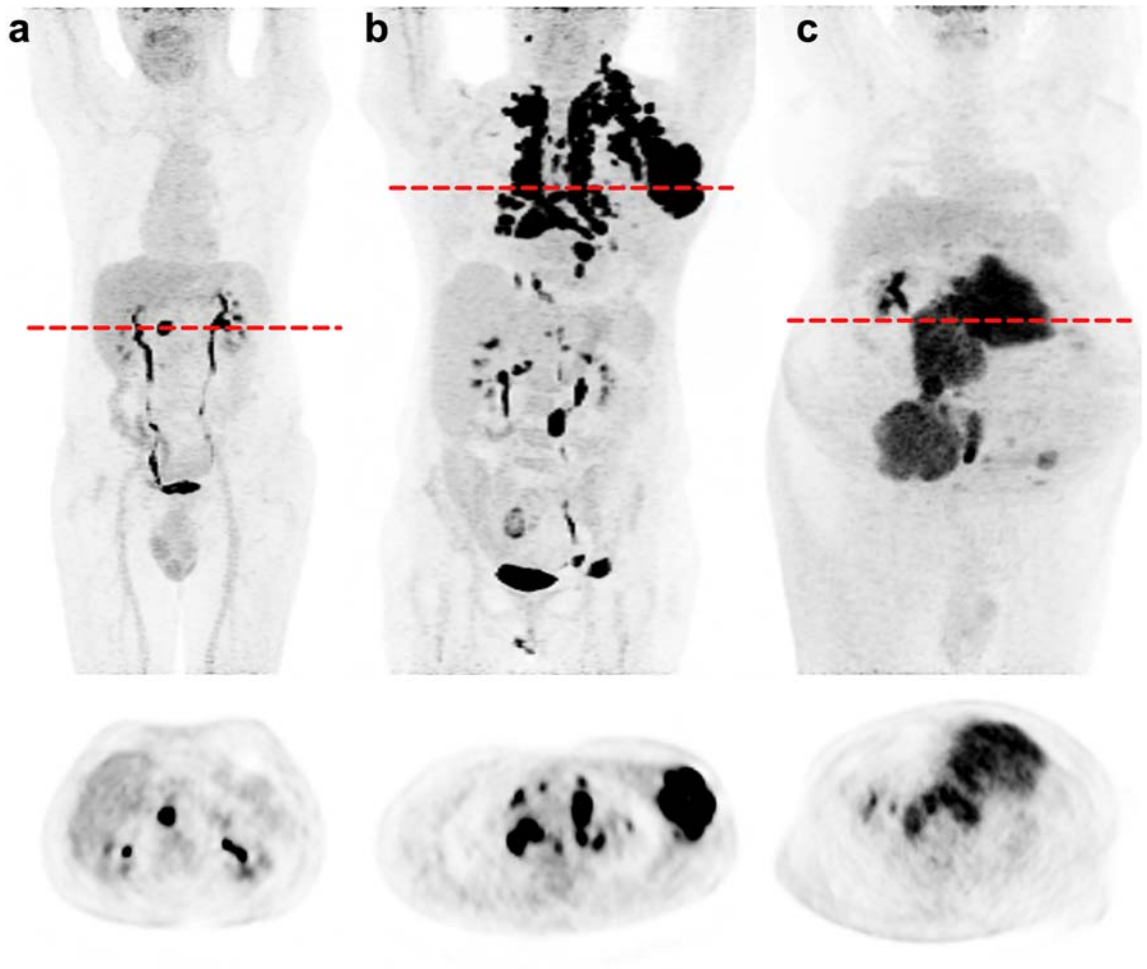

Fig. 3 Representative images of one low total MTV (≤ 111.0 cc) and low LZHGE (≤ 1,264,925.92) patient (**a**), one high total MTV (> 111.0 cc) and low LZHGE (≤ 1,264,925.92) patient (**b**) and one high total MTV (> 111.0 cc) and high LZHGE (> 1,264,925.92) patient (**c**). For each patient, the maximum intensity projection (MIP) image and a representative axial slice are displayed. Patient **c** experienced an event at 3 months (2y-EFS 1), whereas patients **a** and **b** were free of events at 2 years (2y-EFS 0). All images are scaled based on SUV_{max} 10

practice if they are shown to be independent prognosticators compared with well-validated clinical scores. In addition, the availability and user-friendliness of these metrics in busy PET centres is important. Characteristics of our population are similar to those typically observed in DLBCL patients regarding age, sex and IPIaa score [26]. In particular, our 2y-EFS of 77.3% was similar to those values previously reported [27]. Few studies have investigated the role of baseline SUV_{max} in predicting treatment outcome in DLBCL. In our study, baseline SUV_{max} was not associated with 2y-EFS. In contrast, Gallicchio R et al demonstrated that a baseline $SUV_{max} > 13$ predicts a poor outcome [5], and Miyazaki et al concluded that the SUV_{max} on ^{18}F-FDG-PET for newly diagnosed patients with DLBCL is an important predictor of disease progression [28]. As expected, we found that total MTV was a prognosticator of 2y-EFS in the training dataset. This finding is consistent with many studies [6, 29–32]. However, we noticed that statistical significance was not reach in the validation dataset presumably because of its small size. For the delineation of total MTV, a SUV/liver SUV_{max} ratio method with a ratio set to 1.0 was used, based on the Deauville score which states that, to be considered as viable tumour, lymphoma uptake must be superior to liver uptake [33]. Various total MTV delineation methods have been used. Recently, Ilyas et al demonstrated that SUV ≥ 2.5, SUV ≥ 41% and SUV ≥ mean liver uptake (PERCIST) methods all predicted patient's prognosis but that the SUV ≥ 2.5 method had the best interobserver agreement [34]. In our study, the median liver SUV_{max} was equal to 3.7 (range 1.6–5.4), which is nearly consistent with the 2.5 threshold recommended. We assume that using the same delineation method for both MTV and TFs quantification will facilitate and spread its use in clinical practice if it proves its efficiency. To go further into the quantification process of textural features, a relative resampling was not used, because it results in a high correlation between many textural features and metabolic volumes [22, 23]. This is why only the absolute resampling adapted to the reconstruction used was considered in this study. Nine textural features extracted from GLCM and GLSZM were predictors of 2y-EFS on univariable analyses: homogeneity, contrast, correlation, dissimilarity, LZE, LZLGE, LZHGE, GLNU and ZP. For multivariable Cox analysis, known clinical and PET prognostic factors (IPIaa and total MTV) as well as LZHGE were included. IPIaa is a well-established prognostic index based on age, serum lactate dehydrogenase level, performance status and disease stage. To avoid the redundancy and/or correlation of tested data and to lower the number of parameters tested in view of the limited size of our cohort, age, serum lactate dehydrogenase level, performance status and disease stage were therefore not independently included in the Cox multivariate analysis. Total MTV was considered because of its great interest in nuclear medicine and haematological communities, even though not yet validated for multicentre clinical trial and LZHGE because it displayed the highest ROC analysis accuracy and the best discriminant value on univariable Kaplan-Meier and Cox analyses. Finally, on multivariable analysis including well-known prognosticators (i.e. total MTV and IPIaa), LZHGE appeared to be the only independent predictor of 2y-EFS. This parameter extracted from grey-level zone length matrix (GLSZM) counts in 3 dimensions the number of zones of voxels presenting with the same grey level discretisation. It emphasises zone counts in the lower right quadrant of the GLSZM, where large zone sizes and high grey levels are located [35]. GLSZM was privileged to GLRLM because it does not include a step of matrices merging for 3D computation that could influence feature values and lowered the consistency of data [36]. Interestingly in our validation dataset, all patients with low total MTV had low LZHGE values demonstrating that textural feature characterisation could have some additional prognostic value, especially in patients presenting large total MTV. Of note according to the Zwannemburg et al meta-analysis, LZHGE showed average performance in terms of sensitivity in regard to diverse factors variation (i.e. acquisition, reconstruction, segmentation, processing, computation), worse than that of observed for SUV_{max} [21]. However, harmonisation efforts have to be made as per SUV and to this end several methods have been proposed [37] with promising results. Together with the compliance to standard recommendations [21, 38], they could certainly pave the way for multicentre data analysis using various reconstruction protocols and eventually to the implementation of radiomics in clinical routine. There is a growing interest of the haematological community in DLBCL heterogeneity. Gene expression profiling has identified 2 main entities: ABC and GCB subtypes that display different prognosis, ABC being the more aggressive subtype. In the present study, no association between ABC/GCB status and LZHGE was found. Nevertheless, these subtypes are also themselves genetically heterogeneous [39]. More recently, some studies have even addressed the question of intra-tumoural and inter-tumoural diversity within a same patient [40, 41] with a potential impact on survival [42]. In this field, it seems that it might be more relevant to explore the DLBCL heterogeneity at the scale of whole-body rather than with a single tissue biopsy. To this end, ^{18}F-FDG PET metabolic imaging could be a precious tool. For example, Bohers et al demonstrated an association between total MTV and mutational tumour heterogeneity determined by means of cell-free DNA [43]. In our study, we can suppose that the observed metabolic heterogeneity can be a non-invasive representation of the lymphoma tumoural burden and therefore a powerful prognostic indicator. This study has some limitations. First, we conducted our study in a series of newly diagnosed DLBCL patients treated with rituximab and anthracycline-based chemotherapy. Of note, 14 patients (10.6%) did not receive R-CHOP but R-ACVBP, which is a limitation of the study. Secondly, in the

present study, for ^{18}F-FDG heterogeneity computation, only the largest lesion was considered to keep the method straightforward, in view of potential future clinical applications. Taking into account multiple lesions per patient would have led to more sophisticated analysis that would certainly not be applicable in clinical routine and would have required taking into account data clustering, i.e. the fact that several lesions in the same patients are not statistically independent. Also, it seems important to highlight that the VOI of the largest lesions covered on average 77% of the total MTV and that volumes were very well correlated. Besides the VOI of the largest lesion still reached significance on ROC and Kaplan-Meier analyses (Supplemental Fig. 1). We could therefore hypothesise that it might be representative of the whole disease. Nevertheless, we have to keep in mind that some information has surely been omitted. Thirdly, we chose not to use a complex multivariable model, again to keep the method straightforward, in view of potential future clinical applications. Lastly, we chose to focus only on 2y-EFS and not on overall survival or progression-free survival based on the results of the large prospective and multicentric study of Maurer MJ et al stating that it should be considered as an end-point for future studies of newly diagnosed DLBCL [27].

Conclusions

Baseline ^{18}F-FDG PET heterogeneity of the largest lymphoma lesion (expressed as Large-Zone of High Grey-level Emphasis in our study) appeared to be a promising independent predictor of 2y-EFS in newly-diagnosed DLBCL treated with immunochemotherapy. Further studies and validation in an independent cohort of patients are nevertheless needed to confirm these results and more deeply explore the biological background of these results.

Funding information The authors state that this work has not received any funding.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Charline Lasnon.

Conflict of interest The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry One of the authors has significant statistical expertise.

No complex statistical methods were necessary for this paper.

Informed consent Written informed consent was obtained from all subjects (patients) in this study.

Ethical approval Institutional Review Board approval was not required because in accordance with European regulation, French observational studies without any additional therapy or monitoring procedure do not need the approval of an ethical committee. Nonetheless, we sought approval to collect data for our study from the national committee for data privacy, the National Commission on Informatics and Liberty (CNIL), with the registration no. 2081250 v 0.

Methodology
- Retrospective
- Diagnostic or prognostic study
- Performed at one institution

References

1. Weiler-Sagie M, Bushelev O, Epelbaum R et al (2010) (18)F-FDG avidity in lymphoma readdressed: a study of 766 patients. J Nucl Med 51:25–30
2. Tilly H, Gomes da Silva M, Vitolo U et al (2015) Diffuse large B-cell lymphoma (DLBCL): ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 26(Suppl 5): v116–v125
3. Barrington SF, Mikhaeel NG, Kostakoglu L et al (2014) Role of imaging in the staging and response assessment of lymphoma: consensus of the International Conference on Malignant Lymphomas Imaging Working Group. J Clin Oncol 32:3048–3058
4. Chihara D, Oki Y, Onoda H et al (2011) High maximum standard uptake value (SUVmax) on PET scan is associated with shorter survival in patients with diffuse large B cell lymphoma. Int J Hematol 93:502–508
5. Gallicchio R, Mansueto G, Simeon V et al (2014) F-18 FDG PET/CT quantization parameters as predictors of outcome in patients with diffuse large B-cell lymphoma. Eur J Haematol 92:382–389
6. Meignan M, Cottereau AS, Versari A et al (2016) Baseline metabolic tumor volume predicts outcome in high-tumor-burden follicular lymphoma: a pooled analysis of three multicenter studies. J Clin Oncol 34:3618–3626
7. Aide N, Lasnon C, Veit-Haibach P, Sera T, Sattler B, Boellaard R (2017) EANM/EARL harmonization strategies in PET quantification: from daily practice to multicentre oncological studies. Eur J Nucl Med Mol Imaging 44:17–31
8. Blanc-Durand P, Van Der Gucht A, Jreige M et al (2018) Signature of survival: a (18)F-FDG PET based whole-liver radiomic analysis predicts survival after (90)Y-TARE for hepatocellular carcinoma. Oncotarget 9:4549–4558
9. Garcia-Vicente AM, Molina D, Perez-Beteta J et al (2017) Textural features and SUV-based variables assessed by dual time point 18F-FDG PET/CT in locally advanced breast cancer. Ann Nucl Med 31: 726–735
10. Lovinfosse P, Janvary ZL, Coucke P et al (2016) FDG PET/CT texture analysis for predicting the outcome of lung cancer treated by stereotactic body radiation therapy. Eur J Nucl Med Mol Imaging 43:1453–1460
11. Lovinfosse P, Polus M, Van Daele D et al (2018) FDG PET/CT radiomics for predicting the outcome of locally advanced rectal cancer. Eur J Nucl Med Mol Imaging 45:365–375
12. Park S, Ha S, Lee SH et al (2018) Intratumoral heterogeneity characterized by pretreatment PET in non-small cell lung cancer patients predicts progression-free survival on EGFR tyrosine kinase inhibitor. PLoS One 13:e0189766
13. Li S, Young KH, Medeiros LJ (2018) Diffuse large B-cell lymphoma. Pathology 50:74–87

14. Ben Bouallegue F, Tabaa YA, Kafrouni M, Cartron G, Vauchot F, Mariano-Goulart D (2017) Association between textural and morphological tumor indices on baseline PET-CT and early metabolic response on interim PET-CT in bulky malignant lymphomas. Med Phys 44:4608–4619
15. Parvez A, Tau N, Hussey D, Maganti M, Metser U (2018) (18)F-FDG PET/CT metabolic tumor parameters and radiomics features in aggressive non-Hodgkin's lymphoma as predictors of treatment outcome and survival. Ann Nucl Med 32:410–416
16. Parvez A, Tau N, Hussey D, Maganti M, Metser U (2018) Publisher correction to: (18)F-FDG PET/CT metabolic tumor parameters and radiomics features in aggressive non-Hodgkin's lymphoma as predictors of treatment outcome and survival. Ann Nucl Med 32:417
17. Hans CP, Weisenburger DD, Greiner TC et al (2004) Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. Blood 103:275–282
18. Swerdlow SH, Campo E, Pileri SA et al (2016) The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood 127:2375–2390
19. Nioche C, Orlhac F, Boughdad S et al (2018) LIFEx: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. Cancer Res 78:4786–4789
20. Hatt M, Majdoub M, Vallieres M et al (2015) 18F-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 56:38–44
21. Zwanenburg A (2019) Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis. Eur J Nucl Med Mol Imaging 46:2638–2655
22. Orlhac F, Nioche C, Soussan M, Buvat I (2017) Understanding changes in tumor texture indices in PET: a comparison between visual assessment and index values in simulated and patient data. J Nucl Med 58:387–392
23. Orlhac F, Soussan M, Chouahnia K, Martinod E, Buvat I (2015) 18F-FDG PET-derived textural indices reflect tissue-specific uptake pattern in non-small cell lung cancer. PLoS One 10:e0145063
24. Cheson BD, Fisher RI, Barrington SF et al (2014) Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. J Clin Oncol 32:3059–3068
25. Pfreundschuh M, Ho AD, Cavallin-Stahl E et al (2008) Prognostic significance of maximum tumour (bulk) diameter in young patients with good-prognosis diffuse large-B-cell lymphoma treated with CHOP-like chemotherapy with or without rituximab: an exploratory analysis of the MabThera International Trial Group (MInT) study. Lancet Oncol 9:435–444
26. International Non-Hodgkin's Lymphoma Prognostic Factors Project (1993) A predictive model for aggressive non-Hodgkin's lymphoma. N Engl J Med 329:987–994
27. Maurer MJ, Ghesquieres H, Jais JP et al (2014) Event-free survival at 24 months is a robust end point for disease-related outcome in diffuse large B-cell lymphoma treated with immunochemotherapy. J Clin Oncol 32:1066–1073
28. Miyazaki Y, Nawa Y, Miyagawa M et al (2013) Maximum standard uptake value of 18F-fluorodeoxyglucose positron emission tomography is a prognostic factor for progression-free survival of newly diagnosed patients with diffuse large B cell lymphoma. Ann Hematol 92:239–244
29. Cottereau AS, Versari A, Loft A et al (2018) Prognostic value of baseline metabolic tumor volume in early-stage Hodgkin lymphoma in the standard arm of the H10 trial. Blood 131:1456–1463
30. Mikhaeel NG, Smith D, Dunn JT et al (2016) 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 43:1209–1219
31. Sasanelli M, Meignan M, Haioun C et al (2014) Pretherapy metabolic tumour volume is an independent predictor of outcome in patients with diffuse large B-cell lymphoma. Eur J Nucl Med Mol Imaging 41:2017–2022
32. Song MK, Chung JS, Shin HJ et al (2012) Clinical significance of metabolic tumor volume by PET/CT in stages II and III of diffuse large B cell lymphoma without extranodal site involvement. Ann Hematol 91:697–703
33. Barrington SF, Kluge R (2017) FDG PET for therapy monitoring in Hodgkin and non-Hodgkin lymphomas. Eur J Nucl Med Mol Imaging 44:97–110
34. Ilyas H, Mikhaeel NG, Dunn JT et al (2018) Defining the optimal method for measuring baseline metabolic tumour volume in diffuse large B cell lymphoma. Eur J Nucl Med Mol Imaging 45:1142–1154
35. IBSI (2019) Image biomarker standardisation initiative reference manual (V11). IBSI, International. Available via https://ibsi.readthedocs.io/en/latest/. Accessed 21 Jan 2020
36. LIFEx (2019) Texture User Guide, LIFEx, France. Available via https://www.lifexsoft.org/index.php/resources/documentation. Accessed 21 Jan 2020
37. Lasnon C, Majdoub M, Lavigne B et al (2016) (18)F-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 43:2324–2335
38. Boellaard R, Delgado-Bolton R, Oyen WJ et al (2015) FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging 42:328–354
39. Dubois S, Viailly PJ, Mareschal S et al (2016) Next-generation sequencing in diffuse large B-cell lymphoma highlights molecular divergence and therapeutic opportunities: a LYSA study. Clin Cancer Res 22:2919–2928
40. Nissen MD, Kusakabe M, Wang X et al (2019) Single cell phenotypic profiling of 27 DLBCL cases reveals marked intertumoral and intratumoral heterogeneity. Cytometry A. https://doi.org/10.1002/cyto.a.23919
41. Rizzo D, Viailly PJ, Mareschal S et al (2017) Oncogenic events rather than antigen selection pressure may be the main driving forces for relapse in diffuse large B-cell lymphomas. Am J Hematol 92:68–76
42. Wang Y, Feng W, Liu P (2019) Genomic pattern of intratumor heterogeneity predicts the risk of progression in early stage diffuse large B-cell lymphoma. Carcinogenesis 40:1427–1434
43. Bohers E, Viailly PJ, Becker S et al (2018) Non-invasive monitoring of diffuse large B-cell lymphoma by cell-free DNA high-throughput targeted sequencing: analysis of a prospective cohort. Blood Cancer J 8:74