



# Volumetric and texture analysis on FDG PET in evaluating and predicting treatment response and recurrence after chemotherapy in follicular lymphoma

Mitsuaki Tatsumi<sup>1,2</sup> · Kayako Isohashi<sup>2,3</sup> · Keiko Matsunaga<sup>2</sup> · Tadashi Watabe<sup>2</sup> · Hiroki Kato<sup>2</sup> · Yuzuru Kanakura<sup>4</sup> · Jun Hatazawa<sup>2</sup>

Received: 3 February 2019 / Accepted: 31 May 2019 / Published online: 5 June 2019  
© Japan Society of Clinical Oncology 2019

## Abstract

**Purpose** The purpose of this study was to determine if quantitative SUV-related, volumetric FDG PET parameters, and texture features (SPs, VPs, and TFs, respectively) were useful to evaluate and predict response and recurrence after chemotherapy in follicular lymphoma (FL).

**Methods** Pre- and posttreatment FDG PET examinations in 45 FL patients were analyzed retrospectively. In addition to SPs in the representative lesion, metabolic tumor volume (MTV) and total lesion glycolysis (TLG) were calculated as VPs for the representative and whole-body lesions. Six TFs were calculated in the pretreatment representative lesion. Response results with reduction of SPs or VPs after treatment ( $\Delta$ ) were compared to the Lugano classification based on visual assessment. SPs, VPs, and  $\Delta$  of them as well as TFs were also evaluated if they allow prediction of response and recurrence after chemotherapy.

**Results** Quantitative assessment with SPs and VPs provided 89% and 93–96% concordant results, respectively, with Lugano classification. Among pretreatment PET parameters, low gray-level zone emphasis (LGZE) in TFs solely showed statistical significance to predict complete response. All of posttreatment and  $\Delta$  of SPs and VPs were considered as the predictors of progression free survival in the univariate Cox regression analysis, but none of them was the predictor in the multivariate analysis.

**Conclusion** This study demonstrated that quantitative PET parameters were applicable to evaluate treatment response in FL. Texture analysis showed promise in predicting treatment response. Although posttreatment and  $\Delta$  of PET parameters were the candidates, all of them proved to have limited value in predicting recurrence after chemotherapy.

**Keywords** Volumetric parameters · Texture analysis · Quantitative evaluation · FDG PET · Follicular lymphoma

## Introduction

Follicular lymphoma (FL) is the second most common types of B-cell lymphoma. FL accounts for 20 - 30% of all non-Hodgkin's lymphomas, and its incidence is reported to be increasing [1, 2]. FL is classified as indolent lymphoma, and often grows slowly. As in other types of lymphoma, 2-deoxy-2-[F-18]fluoro-D-glucose (FDG) positron emission tomography (PET) or positron emission tomography/computed tomography (PET/CT) (hereinafter, both referred to as FDG PET) is now used to evaluate stage and treatment response in FL. However, FL often represents relatively low FDG uptake in PET examinations even before treatment, reflecting its indolent nature [3, 4]. This relatively low FDG

✉ Mitsuaki Tatsumi  
m-tatsumi@radiol.med.osaka-u.ac.jp

<sup>1</sup> Department of Radiology, Osaka University Hospital, 2-2-D1 Yamadaoka, Suita, Osaka 565-0871, Japan

<sup>2</sup> Department of Nuclear Medicine and Tracer Kinetics, Osaka University Graduate School of Medicine, Suita, Osaka, Japan

<sup>3</sup> Department of Radiology, Osaka Medical College, Takatsuki, Osaka, Japan

<sup>4</sup> Department of Hematology and Oncology, Osaka University Graduate School of Medicine, Suita, Osaka, Japan

uptake potentially results in underestimation of treatment response, which could cause confusion in clinical situations.

The revised response criteria proposed in 2007 followed by the Lugano classification in 2014 have been used to evaluate treatment response in lymphoma [5, 6]. They adopt FDG PET as a reliable imaging modality for evaluation, especially when residual soft tissues exist after treatment. FDG uptake is analyzed only visually to determine response after treatment even in the Lugano classification.

Standardized uptake value (SUV), the maximum value of SUV in the region-of-interest ( $SUV_{max}$ ) among others, has long been used as a useful quantitative parameter in oncologic FDG PET examinations.  $SUV_{max}$  is easily available and is considered to allow more objective evaluation of disease activity than visual analysis. However, there has been criticism in using  $SUV_{max}$ , because it is calculated only from one voxel of the region-of-interest. Metabolic tumor volume (MTV) is an area (voxels) above the threshold SUV and total lesion glycolysis is the parameter defined as MTV multiplied by the mean value of SUV within the target observed in PET images. Many researchers have reported that these volumetric parameters are useful to evaluate disease activity in several malignancies [7–9]. In addition to SUV-related or volumetric FDG PET parameters, texture analysis has gained attention to evaluate intratumoral heterogeneity associated with tumor characteristics in recent years [10–12]. Texture analysis is reported to be useful in predicting treatment response or prognosis in several malignancies as well. However, only limited number of reports has been available so far as to evaluation with volumetric parameters or texture features on FDG PET in lymphoma [13–16].

PET Response Criteria in Solid Tumors (PERCIST) was proposed in 2009 to evaluate changes of tumor metabolic activity observed in FDG uptake [17]. Treatment response is defined based on the quantitative changes of FDG uptake in tumors in PERCIST. Although included in the hematologic malignancies, lymphoma is evaluated with FDG PET in the same manner as solid malignant tumors in clinical situations. Thus, PERCIST may be used as a quantitative approach to evaluate treatment response in lymphoma like in solid malignant tumors.

In the present study, we aimed to determine if quantitative SUV-related or volumetric FDG PET parameters were applicable to evaluate treatment response in follicular lymphoma, comparing the results to ones from the Lugano classification based on visual evaluation of FDG uptake. We also evaluated if these quantitative FDG PET parameters as well as texture features were useful to predict treatment response and recurrence after chemotherapy in FL.

## Materials and methods

### Patients

Fifty-one patients with FL had FDG PET examinations before and after chemotherapy during October 2007 and December 2013. In addition to FDG PET examinations, routine physical examination, laboratory studies, contrast-enhanced computed tomography (CT) of the neck, chest, and abdomen/pelvis, and bone marrow biopsy were performed as staging procedures within 14 days before the start of chemotherapy. Patients received 6–8 cycles of chemotherapy. Chemotherapy regimen included cyclophosphamide, hydroxydaunorubicin, vincristine, prednisolone, and additional rituximab. Six patients were excluded from this study because of insufficient PET data. The 45 patients included 19 males and 26 females, and aged from 30 to 80 (mean 59 y.o.). Body weight ranged from 39 to 89 kg (mean 60 kg). Six, 25, and 14 patients represented stage 2, 3, and 4, respectively.

The institutional review board approved this retrospective study and written informed consent was waived.

### FDG PET imaging

FDG PET imaging was performed with an integrated PET-CT scanner (GEMINI GXL, Philips, Cleveland, USA). PET portion was acquired using the following parameters: 3D emission scan; 2 min scan/bed position  $\times$  11 positions; OSEM reconstruction; and 4.0 mm slice thickness/interval. Acquisition parameters for CT portion (16 slice CT) were as follows: breath-hold during normal expiration from the level of apex of lungs to the lower pole of kidneys; without intravenous nor oral contrast media; 120 kVp and 50 effective mAs; and 5.0 mm slice thickness/4.0 mm interval.

### Evaluation of treatment response and recurrence

Abnormal FDG uptake for lymphoma was defined as increased uptake relative to any parts of the liver according to the Lugano classification in malignant lymphoma [6], which was not due to physiological or inflammatory activity. FDG uptake obviously resulted from other malignant or benign lesions was excluded from the analysis in this study.

Treatment response was evaluated using posttreatment FDG PET images on the basis of the Lugano classification in malignant lymphoma. The classification adopts visual five-point scale as follows; 1, no uptake; 2, uptake equal to or less than that of mediastinum blood pool; 3, uptake equal to or less than that of liver; 4, moderately increased uptake more

than that of liver; and 5, markedly increased uptake more than that of liver and/or new lesions. A score of 1–3 was regarded as negative regardless of residual morphological abnormality, and a score of 4 or 5 was judged as positive. Cases with score 1, 2, or 3 were considered to achieve complete response (CR). Cases with score 4 or 5 were classified into partial response (PR), stable disease (SD), or progressive disease (PD) if the uptake represented decrease, no significant changes, or increase, respectively, as compared to pretreatment uptake. Further details are described in the Lugano classification in malignant lymphoma [6].

FDG PET and/or diagnostic CT were performed as follow-up examinations to evaluate disease status and to detect recurrence.

### Quantitative analysis

Volumes-of-interest (VOIs) were automatically set throughout the body with a threshold SUV slightly greater than  $SUV_{max}$  in the liver, and VOIs corresponding to FL lesions were then selected viewing PET, CT, and fused PET-CT images. When FDG uptake in FL lesions was difficult to separate from adjacent physiological activity, spherical or rectangular VOIs were reset manually in these particular lesions. These procedures were performed with a commercially available software (PETSTAT: AdIn Research, Tokyo, Japan).

In each examination, the representative lesion was defined as a single lesion with the highest FDG uptake, even though patients had multiple active lesions.  $SUV_{max}$  was calculated from a single voxel exhibiting the maximum SUV within a representative lesion.  $SUV_{peak}$  was the mean SUV of a 1 cm<sup>3</sup>-3 dimensional region-of-interest showing the highest value in the representative lesion. They were obtained in the same area as the pretreatment lesion in case of no abnormal uptake after treatment.  $SUV_{peak}$  normalized to the lean body mass ( $SUL_{peak}$ ) was used for analysis in this study.

MTV and total lesion glycolysis (TLG) were calculated as volumetric parameters in the representative lesion as well as in whole-body (wb) lesions. MTV was defined as a volume showing abnormal FDG uptake greater than any parts of the liver in this study. TLG is the product of MTV and mean SUV in a lesion. Reduction rate of post-treatment parameter values relative to pretreatment ones was defined as  $[(1 - \text{posttreatment parameter values}/\text{pretreatment values}) \times 100]$  (%) and was expressed using  $\Delta$  such as  $\Delta SUV_{max}$ .

Changes of SUV-related parameters after treatment were evaluated according to PERCIST [17] and those of volumetric parameters were with modified PERCIST. The quantitative evaluation results were classified into 4 categories as follows; complete metabolic response—complete resolution of abnormal FDG uptake (uptake equal to or less than that

of the liver), partial metabolic response—reduction of minimum of 30% in  $SUV_{max}$  and  $SUL_{peak}$  or minimum of 50% in volumetric parameters, stable metabolic disease—not complete metabolic response, partial metabolic response, or progressive metabolic disease, and progressive metabolic disease—30% increase in  $SUV_{max}$  and  $SUL_{peak}$  or 50% increase in volumetric parameters, or advent of new abnormal FDG uptake. Changes of volumetric parameter values are generally greater than those of  $SUV_{max}$  or  $SUL_{peak}$ , and thus a 50% threshold was chosen as a modification referring to the revised response criteria in lymphoma. Complete metabolic response, partial metabolic response, stable metabolic disease, and progressive metabolic disease were treated as CR, PR, SD, and PD in this study.

### Texture analysis

Intratumoral heterogeneity was evaluated quantitatively with texture analysis in the pretreatment lesions where  $SUV_{max}$  was obtained. Lesions to be evaluated should have MTV greater than 5 ml to avoid volume effects in small tumors [11, 18]. Texture features used in this study were selected according to the report by Orhac et al. [11].

Homogeneity and entropy were calculated from the co-occurrence matrix, short-run emphasis (SRE) and long-run emphasis (LRE) from the gray-level run length matrix, and low gray-level zone emphasis (LGZE) and high gray-level zone emphasis (HGZE) from the gray-level zone length matrix. These six texture features were reported to be the most robust with respect to the tumor segmentation method in each texture correlation group indices [11] and were obtained with the same software as the one to calculate volumetric parameters (PETSTAT).

### Statistical analysis

The diagnostic performance was compared using Fisher's exact test between cases with  $SUV_{max}$  less than 5 and cases with  $SUV_{max}$  equal to or greater than 5 to clarify if uptake intensity affects the results. Quantitative PET parameters were compared between CR and non-CR cases as well as non-recurrence and recurrence cases with the Mann–Whitney U test. The receiver operating characteristic (ROC) analysis was used for quantitative PET parameters to assess the association with treatment response or recurrence status and to obtain optimal cut-off values for comparison of high- and low-value subgroups. Association between quantitative PET parameters and treatment response was also assessed with the logistic regression analysis.

The progression free survival (PFS) was defined as the time from initiation of treatment to disease progression or recurrence. The follow-up period ended in five years in cases without disease progression or recurrence. PFS was

calculated by the Kaplan–Meier method and the log-rank test was used to compare PFS between the high- and low-value subgroups. Association between quantitative PET parameters and PFS was evaluated using univariate Cox proportional hazards regression analysis. Significant PET parameters identified by univariate analysis were included in a multivariate Cox proportional hazards regression analysis. A *p* value less than 0.05 was considered statistically significant.

## Results

Abnormal FDG uptake for lymphoma was observed in all cases before treatment.  $SUV_{max}$  in each case ranged from 3.8 to 17.8 with a mean of  $7.5 \pm 3.4$  (Table 1).  $SUV_{max}$  less than 5 was observed in 20 of 45 (44%) pretreatment examinations. MTV ranged from 5.9 to 686 ml (mean:  $91 \pm 135$  ml) and  $MTV_{wb}$  ranged from 5.9 to 3556 ml (mean:  $323 \pm 742$  ml)

(Table 1). The values of other pretreatment PET parameters and texture features are listed in Table 1.

## Evaluation of treatment response

The Lugano classification revealed CR in 23 cases, PR in 20, and PD in 2 after treatment (Table 2). Posttreatment  $SUV_{max}$ ,  $SUL_{peak}$ , and values of volumetric PET parameters as well as  $\Delta$  of these values are summarized in Table 1.

PERCIST with  $SUV_{max}$  provided concordant results with Lugano classification in 40 (89%) of the 45 cases, as CR in 23, PR in 15, and PD in 2 (Table 2). Discordant results between PERCIST with  $SUV_{max}$  and Lugano classification in 5 cases all accounted for the combination of SD in PERCIST and PR in Lugano classification (Fig. 1). Seventeen (85%) of 20 cases with  $SUV_{max}$  less than 5 and 23 (92%) of 25 cases with  $SUV_{max}$  equal to or greater than 5 showed concordant results. No statistical difference was observed between them.

**Table 1** SUV-related and volumetric PET parameters and texture features: pre- and posttreatment values and reduction rate

	Pretreatment	Posttreatment	$\Delta$ (%)
$SUV_{max}$	$7.5 \pm 3.4$ (3.8–17.8)	$2.8 \pm 2.0$ (1–10.3)	$59 \pm 25$ (– 15 to 90)
$SUL_{peak}$	$6.7 \pm 3.0$ (3.4–15.5)	$2.6 \pm 1.6$ (1–8.4)	$56 \pm 25$ (– 5 to 88)
MTV	$91 \pm 135$ (5.9–686)	$18 \pm 72$ (0–395)	$48 \pm 243$ (– 1358 to 100)
TLG	$399 \pm 609$ (9.4–3062)	$60 \pm 258$ (0–1496)	$69 \pm 129$ (– 662 to 100)
$MTV_{wb}$	$323 \pm 742$ (5.9–3556)	$24 \pm 81$ (0–420)	$85 \pm 50$ (– 173 to 100)
$TLG_{wb}$	$1168 \pm 2429$ (13–11,109)	$81 \pm 285$ (0–1573)	$88 \pm 41$ (– 133 to 100)
Homogeneity	$0.15 \pm 0.06$ (0.01–0.28)	na	na
Entropy	$4.8 \pm 2.1$ (0.11–7.2)	na	na
SRE	$0.98 \pm 0.01$ (0.94–1.0)	na	na
LRE	$1.1 \pm 0.06$ (1.0–1.3)	na	na
LGZE	$0.07 \pm 0.05$ (0.02–0.28)	na	na
HGZE	$689 \pm 319$ (251–1868)	na	na

$\Delta$ , reduction rate after treatment

$SUV_{max}$  maximum standardized uptake value,  $SUL_{peak}$  peak value of SUV corrected for the lean body mass, MTV metabolic tumor volume, TLG total lesion glycolysis,  $MTV_{wb}$  MTV for whole-body lesions,  $TLG_{wb}$  TLG for whole-body lesions, SRE short-run emphasis, LRE long-run emphasis, LGZE low gray-level zone emphasis, HGZE high gray-level zone emphasis, na not available

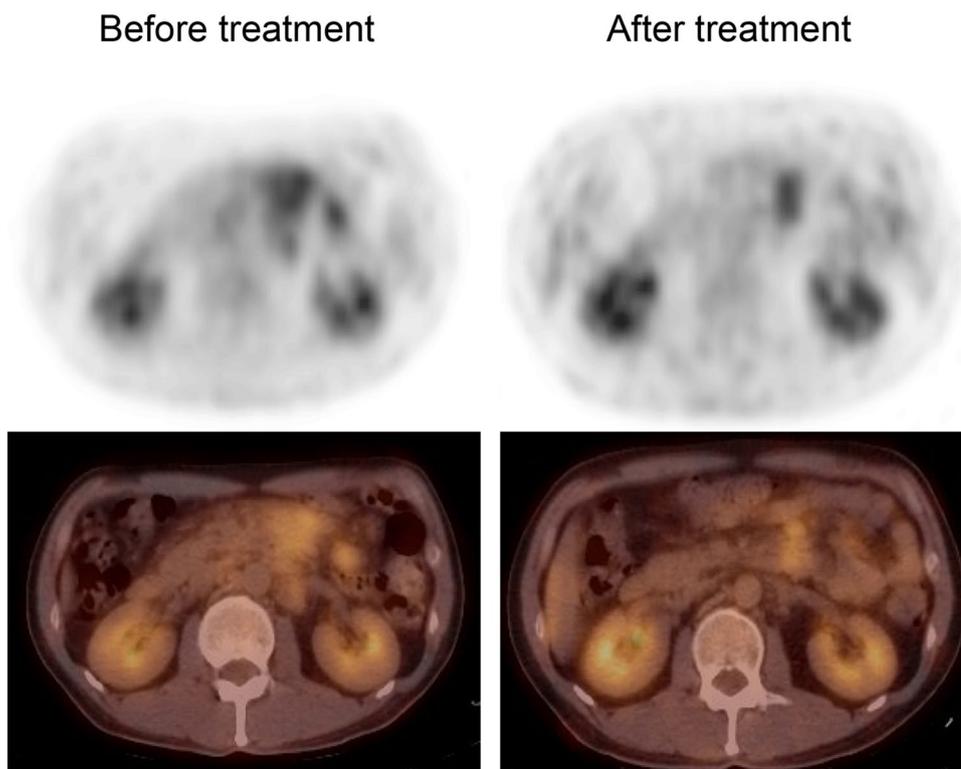
**Table 2** Comparison of treatment response between Lugano classification and PERCIST with various parameters

Treatment response	Lugano classification	PERCIST with parameters below					
		$SUV_{max}$	$SUL_{peak}$	MTV	TLG	$MTV_{wb}$	$TLG_{wb}$
CR	23	23	23	23	23	23	23
PR	20	15	15	17	18	18	18
SD	0	5 (5)	5 (5)	1 (1)	0	0	1 (1)
PD	2	2	2	4 (2)	4 (2)	4 (2)	3 (1)

Number of patients are shown. Parentheses indicate inaccurate evaluation

PERCIST PET response criteria in solid tumors,  $SUV_{max}$  maximum standardized uptake value,  $SUL_{peak}$  peak value of SUV corrected for the lean body mass, MTV metabolic tumor volume, TLG total lesion glycolysis,  $MTV_{wb}$  MTV for whole-body lesions,  $TLG_{wb}$  TLG for whole-body lesions, CR complete response, PR partial response, SD stable disease, PD progressive disease

**Fig. 1** A case with discordant results between PERCIST with  $SUV_{max}$  and Lugano classification. The abdominal lesion decreased in size after chemotherapy with residual FDG uptake (Lugano 4: partial response).  $SUV_{max}$  changed from 3.9 to 3.1 (21% reduction), which was considered as stable disease in PERCIST. PERCIST: PET Response Criteria in Solid Tumors



PERCIST with  $SUL_{peak}$  or volumetric PET parameters exhibited concordant results with Lugano classification as follows:  $SUL_{peak}$  40 cases (89%); MTV 42 cases (93%); TLG,  $MTV_{wb}$ , and  $TLG_{wb}$  43 (96%) of 45 cases (Table 2). PERCIST with volumetric parameters provided more concordant results with Lugano classification than PERCIST with  $SUV_{max}$  or  $SUL_{peak}$  did by reducing the number of discordant results due to SD in PERCIST and PR in Lugano classification.

### Prediction of treatment response

$SUV_{max}$ ,  $SUL_{peak}$ , values of volumetric PET parameters, or texture features before treatment were compared between CR and non-CR cases (Table 3). Although none of the SUV-related and volumetric PET parameters was different between CR and non-CR cases, low gray-level zone emphasis in texture features was significantly higher in CR cases. ROC analysis also indicated that the low gray-level zone emphasis solely showed statistical significance to predict CR after treatment ( $p < 0.005$ ; area under curve 0.72; sensitivity 70%; specificity 64%). The low gray-level zone emphasis was demonstrated to be the predictor of CR after treatment by the logistic regression analysis ( $p < 0.005$ ).

**Table 3** Comparison of SUV-related and volumetric PET parameters and texture features between CR and non-CR cases

	CR	Non-CR	<i>p</i> value
pre $SUV_{max}$	6.8 ± 3.5	6.8 ± 3.4	ns
pre $SUL_{peak}$	6.0 ± 3.0	6.1 ± 3.2	ns
preMTV	76.3 ± 106	73.8 ± 148	ns
preTLG	309 ± 417	346 ± 701	ns
pre $MTV_{wb}$	261 ± 734	272 ± 640	ns
pre $TLG_{wb}$	870 ± 2142	1058 ± 2389	ns
Homogeneity	0.06 ± 0.01	0.05 ± 0.01	ns
Entropy	4.4 ± 2.5	5.2 ± 1.5	ns
SRE	0.98 ± 0.01	0.98 ± 0.01	ns
LRE	1.1 ± 0.06	1.1 ± 0.05	ns
LGZE	0.09 ± 0.06	0.05 ± 0.02	< 0.01
HGZE	714 ± 407	663 ± 196	ns

pre- indicates pretreatment-

CR complete response,  $SUV_{max}$  maximum standardized uptake value,  $SUL_{peak}$  peak value of SUV corrected for the lean body mass, MTV metabolic tumor volume, TLG total lesion glycolysis,  $MTV_{wb}$  MTV for whole-body lesions,  $TLG_{wb}$  TLG for whole-body lesions, SRE short-run emphasis, LRE long-run emphasis, LGZE low gray-level zone emphasis, HGZE high gray-level zone emphasis, ns not significant

### Prediction of recurrence after treatment

At the time of analysis, disease progression was observed in 26 (58%) of 45 cases. Twenty of them showed progression within 2 years. The median PFS was 26.5 months (range 3.1–83.9 months).

Pre- and posttreatment  $SUV_{max}$ ,  $SUL_{peak}$ , values of volumetric PET parameters,  $\Delta$  of them, and pretreatment texture features were compared between recurrence and non-recurrence cases (Table 4). None of the pretreatment SUV-related

and volumetric PET parameters as well as texture features was different between recurrence and non-recurrence cases. However, all of the posttreatment and  $\Delta$  of SUV-related and volumetric PET parameters exhibited significant differences between recurrence and non-recurrence cases. They were demonstrated to show significance to predict recurrence after treatment in ROC analysis ( $p < 0.001$  for all). (Table 5).

The posttreatment and  $\Delta$  of the SUV-related and volumetric PET parameters were dichotomized into high and low groups with cutoff thresholds provided by ROC analysis

**Table 4** Comparison of PET-derived parameters between recurrence and non-recurrence cases

PET-derived parameter	Recurrence	Non-recurrence	<i>p</i> value	PET-derived parameter	Recurrence	Non-recurrence	<i>p</i> value
pre $SUV_{max}$	7.2 ± 3.3	6.3 ± 3.7	ns	post $SUV_{max}$	3.4 ± 2.3	1.7 ± 0.6	< 0.01
pre $SUL_{peak}$	6.4 ± 3.1	5.7 ± 3.1	ns	post $SUL_{peak}$	3.1 ± 1.8	1.7 ± 0.6	< 0.01
preMTV	74.6 ± 139	75.7 ± 113	ns	postMTV	25.4 ± 84.9	0.04 ± 0.2	< 0.01
preTLG	352 ± 654	292 ± 438	ns	postTLG	86 ± 305.4	0.11 ± 0.5	< 0.01
pre $MTV_{wb}$	248 ± 596	291 ± 802	ns	post $MTV_{wb}$	35.1 ± 94.8	0.05 ± 0.2	< 0.01
pre $TLG_{wb}$	996 ± 2229	915 ± 2320	ns	post $TLG_{wb}$	118 ± 335	0.1 ± 0.6	< 0.01
Homogeneity	0.16 ± 0.04	0.13 ± 0.07	ns	$\Delta SUV_{max}$	49.0 ± 30.3	65.6 ± 18.5	< 0.05
Entropy	5.1 ± 1.7	4.3 ± 2.5	ns	$\Delta SUL_{peak}$	45.4 ± 30.6	61.8 ± 22.0	< 0.05
SRE	0.98 ± 0.01	0.98 ± 0.01	ns	$\Delta MTV$	6.0 ± 302	99.3 ± 3.1	< 0.01
LRE	1.1 ± 0.05	1.1 ± 0.06	ns	$\Delta TLG$	33.0 ± 191	99.4 ± 2.5	< 0.01
LGZE	0.06 ± 0.02	0.09 ± 0.07	ns	$\Delta MTV_{wb}$	50.7 ± 149	99.7 ± 1.4	< 0.01
HGZE	626 ± 200	775 ± 423	ns	$\Delta TLG_{wb}$	51.9 ± 160	99.7 ± 1.3	< 0.01

pre- and post- indicate pretreatment- and posttreatment-, respectively.  $\Delta$ , reduction rate after treatment

$SUV_{max}$  maximum standardized uptake value,  $SUL_{peak}$  peak value of SUV corrected for the lean body mass,  $MTV$  metabolic tumor volume,  $TLG$  total lesion glycolysis,  $MTV_{wb}$  MTV for whole-body lesions,  $TLG_{wb}$  TLG for whole-body lesions,  $SRE$  short-run emphasis,  $LRE$  long-run emphasis,  $LGZE$  low gray-level zone emphasis,  $HGZE$  high gray-level zone emphasis, *ns* not significant

**Table 5** Results of ROC analysis with PET-derived parameters in progression-free survival

PET-derived parameter	Area under ROC curve	<i>p</i> value	PET-derived parameter	Area under ROC curve	<i>p</i> value
pre $SUV_{max}$	0.61 (0.43–0.78)	ns	post $SUV_{max}$	0.81 (0.68–0.94)	< 0.001
pre $SUL_{peak}$	0.59 (0.41–0.77)	ns	post $SUL_{peak}$	0.81 (0.68–0.94)	< 0.001
preMTV	0.53 (0.34–0.72)	ns	postMTV	0.75 (0.64–0.86)	< 0.001
preTLG	0.45 (0.27–0.64)	ns	postTLG	0.75 (0.64–0.86)	< 0.001
pre $MTV_{wb}$	0.56 (0.38–0.74)	ns	post $MTV_{wb}$	0.75 (0.64–0.86)	< 0.001
pre $TLG_{wb}$	0.57 (0.39–0.75)	ns	post $TLG_{wb}$	0.75 (0.64–0.86)	< 0.001
Homogeneity	0.60 (0.42–0.78)	ns	$\Delta SUV_{max}$	0.66 (0.50–0.82)	< 0.001
Entropy	0.56 (0.38–0.74)	ns	$\Delta SUL_{peak}$	0.68 (0.51–0.84)	< 0.001
SRE	0.52 (0.34–0.70)	ns	$\Delta MTV$	0.74 (0.63–0.85)	< 0.001
LRE	0.52 (0.34–0.70)	ns	$\Delta TLG$	0.74 (0.63–0.85)	< 0.001
LGZE	0.62 (0.45–0.80)	ns	$\Delta MTV_{wb}$	0.75 (0.64–0.86)	< 0.001
HGZE	0.58 (0.40–0.76)	ns	$\Delta TLG_{wb}$	0.75 (0.64–0.86)	< 0.001

pre- and post- indicate pretreatment- and posttreatment-, respectively.  $\Delta$ , reduction rate after treatment

Parentheses indicate 95% confidence intervals

ROC receiver operating characteristic,  $SUV_{max}$  maximum standardized uptake value,  $SUL_{peak}$  peak value of SUV corrected for the lean body mass,  $MTV$  metabolic tumor volume,  $TLG$  total lesion glycolysis,  $MTV_{wb}$  MTV for whole-body lesions,  $TLG_{wb}$  TLG for whole-body lesions,  $SRE$  short-run emphasis,  $LRE$  long-run emphasis,  $LGZE$  low gray-level zone emphasis,  $HGZE$  high gray-level zone emphasis, *ns* not significant

to generate Kaplan–Meier survival plots. Significant differences were observed in PFS between high and low groups in all of the posttreatment and  $\Delta$  of the SUV-related and volumetric PET parameters ( $p < 0.05$  for  $\Delta\text{SUV}_{\text{max}}$  and  $\Delta\text{SUL}_{\text{peak}}$ ;  $p < 0.001$  for others).

The univariate Cox regression analysis revealed that all of the posttreatment and  $\Delta$  of the SUV-related and volumetric PET parameters proved to be the predictors of PFS after treatment ( $p < 0.05$  for  $\Delta\text{SUV}_{\text{max}}$  and  $\Delta\text{SUL}_{\text{peak}}$ ;  $p < 0.001$  for others) (Table 6). However, none of them was considered to be the predictor of PFS after treatment in the multivariate analysis.

## Discussion

The present study demonstrated that FDG uptake observed in FL cases was generally low (mean of  $\text{SUV}_{\text{max}}$ : 7.5) and that approximately half of all cases showed low FDG uptake ( $\text{SUV}_{\text{max}}$  less than 5) before treatment. These results were as expected, and consistent with previous reports [3, 4]. This generally low FDG uptake is considered to reflect disease activity of FL.

Our concern was that underestimation of treatment response might occur in FL cases especially with low FDG uptake before treatment. This study revealed that quantitative approach with various PET parameters was applicable in the response assessment of FL. Quantitative assessment with  $\text{SUV}_{\text{max}}$  or  $\text{SUL}_{\text{peak}}$  provided almost 90% concordant results with the Lugano classification. Even in cases with  $\text{SUV}_{\text{max}}$

less than 5, 85% concordant results were observed between PERCIST with  $\text{SUV}_{\text{max}}$  and the Lugano classification, which was not statistically different from the results (92%) obtained in cases with  $\text{SUV}_{\text{max}}$  equal to or greater than 5. PERCIST is considered to be applicable in evaluating treatment response in FL patients.

Treatment response in PERCIST is determined based on the degree of change in FDG uptake except in cases showing CR. Thus, discordant results between PERCIST and the Lugano classification potentially occur in cases with residual FDG uptake after treatment. In the present study, this kind of discordance was observed in cases with lesions showing SD in PERCIST and PR in the Lugano classification. Volumetric PET parameters such as MTV or TLG provided more similar results to the Lugano classification than  $\text{SUV}_{\text{max}}$  or  $\text{SUL}_{\text{peak}}$  did in evaluating treatment response. Volumetric parameters reduced the number of discordant results due to insufficient reduction rate in SUV-related parameters. Theoretically changes of volumetric PET parameters are greater than those of SUV-related ones as tumor growth or shrinkage occurs three-dimensionally in general. Accordingly these prominent changes of volumetric parameters are expected to contribute significantly to accurate response evaluation, especially in cases with low FDG uptake before treatment. Sharma et al. [13] or Tateishi et al. [19] also reported the advantages of volumetric PET parameters over SUV-related ones in evaluating treatment response in lymphoma.

Texture features were demonstrated to be useful in predicting treatment response, but not in predicting recurrence after treatment in this study. There have been only a few published papers so far as to relationship between intratumoral heterogeneity of FDG uptake and treatment response or recurrence in lymphoma. Tateishi et al. [19] and Hanaoka et al. [20] both reported that the area under the curve of cumulative SUV-volume histogram (AUC-CSH) had no association with treatment response nor recurrence. Analyses in this study were performed with texture features of higher order than AUC-CSH, which might have provided the discrepant results between our study and theirs. Ko et al. [16] reported that texture analysis was useful in predicting PFS in nasal type extranodal natural killer/T-cell lymphoma (ENTNL). FL and ENTNL usually behave differently in terms of disease sites and activity. This possibly resulted in different results from our studies. Another reason may be the difference of texture features used in the studies. Orlich et al. recommend 6 texture features in clinical studies because of their robustness and reproducibility [11]. The texture features used in this study were same as their recommendation, while Ko et al. used other texture features.

Among texture features evaluated in this study, low gray-level zone emphasis was demonstrated to be useful in predicting treatment response. Cases showing CR after treatment had significantly higher low gray-level zone emphasis

**Table 6** Results of univariate analysis with PET-derived parameters in progression-free survival

PET-derived parameter	Hazard ratio	<i>p</i> value
post $\text{SUV}_{\text{max}}$	10.9 (4.1–29.1)	<0.001
post $\text{SUL}_{\text{peak}}$	10.9 (4.1–29.1)	<0.001
postMTV	9.1 (3.6–23.3)	<0.001
postTLG	9.1 (3.6–23.3)	<0.001
post $\text{MTV}_{\text{wb}}$	9.1 (3.6–23.3)	<0.001
post $\text{TLG}_{\text{wb}}$	9.1 (3.6–23.3)	<0.001
$\Delta\text{SUV}_{\text{max}}$	2.4 (1.0–5.3)	<0.05
$\Delta\text{SUL}_{\text{peak}}$	2.5 (1.1–5.5)	<0.05
$\Delta\text{MTV}$	7.1 (2.9–17.5)	<0.001
$\Delta\text{TLG}$	7.1 (2.9–17.5)	<0.001
$\Delta\text{MTV}_{\text{wb}}$	7.4 (3.0–18.2)	<0.001
$\Delta\text{TLG}_{\text{wb}}$	7.4 (3.0–18.2)	<0.001

post- indicates posttreatment-.  $\Delta$ , reduction rate after treatment

Parentheses indicate 95% confidence intervals

$\text{SUV}_{\text{max}}$  maximum standardized uptake value,  $\text{SUL}_{\text{peak}}$  peak value of SUV corrected for the lean body mass, MTV metabolic tumor volume, TLG total lesion glycolysis,  $\text{MTV}_{\text{wb}}$  MTV for whole-body lesions,  $\text{TLG}_{\text{wb}}$  TLG for whole-body lesions

values before treatment than cases showing non-CR did. Orlhac et al. reported that low gray-level zone emphasis exhibited higher values in visually homogeneous lesions than in heterogeneous lesions [21]. Thus, this study indirectly demonstrated that cases with homogeneous FL lesions were expected to have good treatment response.

Posttreatment SUV-related and volumetric PET parameters as well as  $\Delta$  of these PET parameters proved to be significant predictors of PFS after treatment in the univariate analysis. None of them was, however, decided as an independent predictor of PFS in the multivariate analysis. Regarding prediction of recurrence in FL, Meignan et al. [22] reported the prognostic value of pretreatment total-body MTV (whole-body MTV in this study). There have been several reports published on the prognostic value of pretreatment MTV in other types of lymphoma as well [13–15, 19]. In this study, no pretreatment SUV-related and volumetric PET parameters as well as texture features were associated with recurrence. This study dealt with cases with relatively low tumor burden (median: 37.7 ml) as compared to cases in the report of Meignan et al. (median: 297 ml) or others. This relatively low tumor burden might have provided the results that pretreatment PET parameters had no relationship with prognosis. In esophageal cancer, Nakajo et al. reported similar results to our study regarding prediction of treatment response and prognosis after treatment [23]. Texture features as well as volumetric parameters were useful in predicting treatment response but not in predicting prognosis after treatment in their study.

There were some limitations in this study. First, half of the patients gained CR after treatment. As CR was determined only visually in both PERCIST and Lugano classification, quantitative results in PERCIST had to be assessed in the rest of the patients. Second, threshold SUV for volumetric PET parameters was determined according to  $SUV_{max}$  in the liver in this study. SUV 2.5 has been frequently selected as a threshold according to previous reports dealing with various malignancies [24, 25], but Lugano classification defined active lymphoma lesions as ones with higher FDG uptake than liver. FL lesions generally represent relatively low FDG uptake even before treatment, and many lesions showed low FDG uptake equal to or less than that of the liver after treatment in this study. This resulted in MTV equal to 0 after treatment in many cases. Third, this study was conducted retrospectively and the number of patients analyzed was relatively small. They were also limitations in this study.

## Conclusion

This study demonstrated that quantitative SUV-related or volumetric FDG PET parameters were applicable to evaluate treatment response in FL. Texture analysis showed promise

in predicting treatment response. Although posttreatment and  $\Delta$  of PET parameters were the candidates, all of them proved to have limited value in predicting recurrence after chemotherapy.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Dreyling M, Ghielmini M, Rule S et al (2016) Newly diagnosed and relapsed follicular lymphoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 27(suppl 5):v83–v90
- Pyo J, Won Kim K, Jacene HA et al (2013) End-therapy positron emission tomography for treatment response assessment in follicular lymphoma: a systematic review and meta-analysis. *Clin Cancer Res* 19(23):6566–6577
- Schöder H, Noy A, Gönen ML et al (2005) Intensity of 18fluorodeoxyglucose uptake in positron emission tomography distinguishes between indolent and aggressive non-Hodgkin's lymphoma. *J Clin Oncol* 23(21):4643–4651
- Wu X, Pertovaara H, Korkola P et al (2014) Correlations between functional imaging markers derived from PET/CT and diffusion-weighted MRI in diffuse large B-cell lymphoma and follicular lymphoma. *PLoS One* 9(1):e84999
- Cheson BD, Pfistner B, Juweid ME et al (2007) Revised response criteria for malignant lymphoma. *J Clin Oncol* 25(5):579–586
- 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(27):3059–3067
- Pak K, Cheon GJ, Nam H-Y et al (2014) Prognostic value of metabolic tumor volume and total lesion glycolysis in head and neck cancer: a systematic review and meta-analysis. *J Nucl Med* 55(6):884–890
- Im H-J, Pak K, Cheon GJ et al (2015) Prognostic value of volumetric parameters of 18 F-FDG PET in non-small-cell lung cancer: a meta-analysis. *European journal of nuclear medicine and molecular imaging* 42(2):241–251
- Li Y-J, Dai Y-L, Cheng Y-S et al (2016) Positron emission tomography (18 F-fluorodeoxyglucose uptake and prognosis in patients with bone and soft tissue sarcoma: a meta-analysis. *Eur J Surg Oncol* 42
- Chicklore S, Goh V, Siddique M et al (2013) Quantifying tumour heterogeneity in 18 F-FDG PET/CT imaging by texture analysis. *Eur J Nucl Med Mol Imaging* 40(1):133–140
- Orlhac F, Soussan M, Maisonobe J-A et al (2014) Tumor texture analysis in 18F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis. *J Nucl Med* 55(3):414–422
- Hatt M, Tixier F, Pierce L et al (2017) Characterization of PET/CT images using texture analysis: the past, the present ... any future? *Eur J Nucl Med Mol Imaging* 44(1):151–165
- Sharma P, Gupta A, Patel C et al (2012) Pediatric lymphoma: metabolic tumor burden as a quantitative index for treatment response evaluation. *Ann Nucl Med* 26(1):58–66

14. Song M, Chung J, Lee J et al (2013) Metabolic tumor volume by positron emission tomography/computed tomography as a clinical parameter to determine therapeutic modality for early stage Hodgkin's lymphoma. *Cancer Sci* 104(12):1656–1661
15. Kim C-Y, Hong CM, Kim D-H et al (2013) Prognostic value of whole-body metabolic tumour volume and total lesion glycolysis measured on 18 F-FDG PET/CT in patients with extranodal NK/T-cell lymphoma. *Eur J Nucl Med Mol Imaging* 40(9):1321–1329
16. Ko K-Y, Liu C-J, Ko C-L et al (2016) Intratumoral heterogeneity of pretreatment 18F-FDG PET images predict disease progression in patients with nasal type extranodal natural killer/T-cell lymphoma. *Clin Nucl Med* 41(12):922–926
17. Wahl RL, Jacene H, Kasamon Y et al (2009) From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med* 50(Suppl 1):122S–150S
18. Tixier F, Le Rest CC, Hatt M et al (2011) Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med* 52(3):369–378
19. Tateishi U, Tatsumi M, Terauchi T et al (2015) Prognostic significance of metabolic tumor burden by positron emission tomography/computed tomography in patients with relapsed/refractory diffuse large B-cell lymphoma. *Cancer Sci* 106(2):186–193
20. Hanaoka K, Hosono M, Tatsumi Y et al (2015) Heterogeneity of intratumoral 111 In-ibritumomab tiuxetan and 18 F-FDG distribution in association with therapeutic response in radioimmunotherapy for B-cell non-Hodgkin's lymphoma. *EJNMMI Res* 5(1):10
21. Orhac F, Nioche C, Soussan M et al (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(3):387–392
22. Meignan M, Cottreau 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(30):3618–3626
23. Nakajo M, Jinguji M, Nakabeppu Y et al (2017) Texture analysis of 18 F-FDG PET/CT to predict tumour response and prognosis of patients with esophageal cancer treated by chemoradiotherapy. *Eur J Nucl Med Mol Imaging* 44(2):206–214
24. Larson SM, Erdi Y, Akhurst T et al (1999) Tumor treatment response based on visual and quantitative changes in global tumor glycolysis using PET-FDG imaging: the visual response score and the change in total lesion glycolysis. *Clin Positron Imaging* 2(3):159–171
25. Miyabe J, Hanamoto A, Tatsumi M et al (2017) Metabolic tumor volume of primary tumor predicts survival better than T classification in the larynx preservation approach. *Cancer Sci* 108(10):2030–2038

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.