Purpose: In the present study, we aimed to investigate the role of baseline, interim and end-of treatment positron emission tomography/computed tomography (PET/CT) in assessing the prognosis of follicular lymphoma (FL).

Methods: A total of 84 FL patients were retrospectively analyzed in this study. Baseline (n=59), interim (n=24, after 2–4 cycles) and end-of treatment (n=43) PET/CT images were re-evaluated, and baseline maximum standardized uptake value ($SUV_{\text{max}}$), total metabolic tumor volume (tMTV) and total lesion glycolysis (TLG) were recorded. Interim (I-PET) and end-of treatment (E-PET) PET/CT responses were interpreted by Deauville five-point scale (D-5PS) and International Harmonization Project criteria (IHP). Survival curves were calculated by Kaplan-Meier curves, and differences between groups were compared by log-rank test.

Results: The 2-year progression-free survival (PFS) of the high- and low-TLG groups was 57.14% and 95.56%, respectively ($p=0.0001$). The 2-year overall survival (OS) of the high- and low-TLG groups was 62.50% and 100%, respectively ($p<0.0001$). Multivariate analysis showed that TLG was an independent prognostic factor for PFS ($p=0.001$, HR=6.577, 95% CI=2.167–19.960) and OS ($p=0.030$, HR=19.291, 95% CI=2.689–137.947). Besides, Eastern Cooperative Oncology Group (ECOG) was the independent prognostic factor for OS (HR=8.924, 95% CI=1.273–62.559, $p=0.028$). Interim PET results based on D-5PS or IHP criteria were not significantly correlated with PFS (all $p>0.05$). However, E-PET results using D-5PS and IHP criteria were statistically significant ($p=0.0001$ and $p=0.006$). The D-5PS showed stronger prognostic value compared with IHP criteria. The optimal cutoff value of $\Delta SUV_{\text{max}}\%$ was 66.95% according to I-PET and 68.97% according to E-PET. However, only the $\Delta SUV_{\text{max}}\%$ from the baseline to the end-of therapy yielded statistically significant results in the prediction of PFS ($p=0.0002$).

Conclusion: Our findings indicated that the baseline TLG and E-PET results were significantly associated with prognosis in patients with FL.

Keywords: follicular lymphoma, PET/CT, MTV, TLG, D-5PS, IHP

Introduction

Follicular lymphoma (FL) is the second most common subtype of non-Hodgkin lymphoma (NHL) in Western countries, accounting for approximately 20% to 25% of NHL.¹ Although rituximab in combination with chemotherapy has improved the prognosis of FL patients, approximately 20% of patients have relapse within 2 years after first-line treatment, with a 5-year overall survival (OS) rate of only 50%.²,³ However, these patients are not easily identified by current clinical indices of risk, such as the Follicular Lymphoma International Prognostic Index (FLIPI) or FLIPI2...
scores. Therefore, early identification of high-risk factors that have a strong prognostic value for progression is particularly important.

The role of $^{18}$F-fluorodeoxyglucose positron emission tomography/computed tomography ($^{18}$F-FDG PET/CT) in the staging and response assessment of lymphoma has been widely established.\(^4\)–\(^6\) The maximum standardized uptake value ($SUV_{\text{max}}$) is the most widely studied parameter for assessing disease activity in lymphoma.\(^7\) However, reliability is affected by partial volume effect, blood glucose level and time after injection.\(^8\) Recently, several studies have shown that metabolic tumor volume (MTV) and total lesion glycolysis (TLG) are promising prognostic indices in lymphoma, such as diffuse large B-cell lymphoma (DLBCL)\(^4,9\) and Hodgkin lymphoma (HL)\(^10\) However, only very few studies have investigated the prognostic value of baseline total MTV and TLG in FL patients.

The International Harmonization Project criteria (IHP) and the Deauville five-point scale (D-5PS) are commonly used to assess treatment outcome during and after first-line therapy.\(^11,12\) However, the roles of interim PET/CT and end-of treatment PET/CT in determining the prognosis still remain controversial.\(^11,13,14\)

In the present study, we aimed to investigate the prognostic value of semi-quantitative parameters, tMTV, TLG and $SUV_{\text{max}}$, measured at baseline $^{18}$F-FDG PET/CT in FL. We also compared the two different criteria (D-5PS and IHP) for exploring the prognostic value of interim and end-of treatment $^{18}$F-FDG PET/CT.

**Materials and methods**

**Patients**

A retrospective analysis was performed in the present study, which consisted of 84 FL patients (age 25–80 years, mean age of 51 years) who were diagnosed between March 2013 and December 2018. Inclusion criteria were as follows: (1) age ≥18 years, (2) histologically confirmed as FL, (3) patients who underwent baseline PET/CT (B-PET), or/and interim PET/CT (I-PET) after 2–4 cycles of chemotherapy, or/and end-of treatment PET/CT (E-PET) after all planned first-line therapy. Clinical pathological features of patients were also determined, including epidemiological features (gender, age), clinical information [B symptoms, FLIPI score, LDH (lactate dehydrogenase) level, hemoglobin level], Ann Arbor stage, histologic grade, bone marrow biopsy, Eastern Cooperative Oncology Group (ECOG) performance status and imaging data. The FLIPI score was determined according to age ≥60 years, Ann Arbor stage III–IV, hemoglobin level <120 g/L, elevated LDH and number of extranodal sites ≥4.\(^15\)

Of the enrolled patients, 70 patients received (rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone) R-CHOP/R-CHOP-like regimens, eight patients didn’t receive treatment due to lack of indications, and six patients were treated with other treatment regimens, including R-FM (rituximab, fludarabine and mitoxantrone), R-CVP (rituximab, cyclophosphamide, vincristine and prednisone) and rituximab alone.

This study was approved by the institutional review board of the First Affiliated Hospital of Soochow University. Trial registration number: ChiCTR1900023183. Because the trial was a retrospective study, written informed consent for this study was waived by the ethics committee, and no personal information was disclosed. This study was in accordance with the Declaration of Helsinki.

**PET/CT acquisition**

All the FDG PET/CT images were obtained from the US GE Discovery STE 16 PET/CT scanner. The patients were fasted for at least 6 h prior to the intravenous injection of $^{18}$F-FDG (4.07–5.55 M Bq/kg), and the blood glucose level was lower than 11 mmol/L. After intravenous injection of $^{18}$F-FDG for an average of 60±10 min, imaging data were obtained using low-dose CT (140 kV, 120 mA, transaxial FOV 70 cm, pitch 1.75, rotation time 0.8 s, slice thickness 3.75 mm), followed by PET emission images, 2–3 min per bed position. Whole body CT and PET images were obtained on the Xeleris Functional Imaging workstation, and the coronal, axial and sagittal slices PET/CT fusion images were obtained by iterative reconstruction.

**Image analysis**

All PET/CT images were reviewed by two experienced nuclear medicine physicians using the Advantage Workstation 4.3_05 (AW4.3_05). In baseline PET, the highest FDG uptake was considered to be the $SUV_{\text{max}}$ of the patient. For the interim PET/CT and end-of PET/CT images, $SUV_{\text{max}}$ was measured in residual lesions. If the lesion was disappeared after treatment, a region of interest was drawn in the same area on the baseline PET. The percentage change of $SUV_{\text{max}}$ was calculated using the following equation: $\Delta SUV_{\text{max}}\% = [SUV_{\text{max}}(\text{baseline})-SUV_{\text{max}}(\text{post-therapy})/SUV_{\text{max}}(\text{baseline})] \times 100$. A
threshold of 41% $\text{SUV}_{\text{max}}$ was used to delineate the metabolic tumor volume (MTV), as recommended by the European Association of Nuclear Medicine. Total MTV (tMTV) referred to the sum of MTV of all lesions, and TLG was the sum of the product of MTV and its $\text{SUV}_{\text{mean}}$ in each lesion. Bone marrow involvement was considered in volume measurement only if there was focal uptake. Splenic involvement was considered if there was focal uptake or diffuse uptake higher than 150% of the liver background.

The interim and end-of treatment PET/CT results were assessed according to the IHP$^{18}$ and D-5PS$^{19}$ criteria. For IHP criteria, FDG uptake greater than the uptake of the mediastinum in lesions greater than or equal to 2 cm and more than the adjacent background tissue in lesions less than 2 cm represented residual disease. The D-5PS scoring system was used to quantitatively evaluate the treatment response as follows: (1) no uptake; (2) uptake $\leq$ mediastinal blood pool; (3) uptake $>$ mediastinal blood pool, but $\leq$ liver; (4) uptake moderately increased compared with the liver uptake at any site; (5) uptake markedly increased compared with the liver at any site. Scores of 4–5 were considered positive, while scores of 1–3 were considered negative.

Statistical analysis
All statistical analyses were performed using SPSS software (version 22, Chicago, IL). The Pearson chi-square test and Fisher’s exact test were used to analyze the relationships between the PET/CT results and clinical variables. Correlations between clinical characteristics and $\text{SUV}_{\text{max}}$, tMTV or TLG were assessed using the Spearman correlation test. The suitable cutoff points of $\text{SUV}_{\text{max}}$, tMTV, TLG and $\Delta\text{SUV}_{\text{max}}$% were obtained using the receiver operating characteristic (ROC) curve, and the area under the curve (AUC) was calculated. Progression-free survival (PFS) was defined from diagnosis to disease progression (increased uptake of FDG-PET/CT, increased tumor volume, changes in laboratory examinations and clinical symptoms), death or last follow-up. OS was defined from diagnosis to the date of death or last follow-up. Survival curve was plotted by Kaplan-Meier curves, and differences between groups were compared by log-rank test. The multivariate survival analysis was performed using the Cox proportional hazards model. OS was only analyzed with baseline PET/CT due to the small number of events. Differences in sensitivity, specificity, accuracy, positive predictive value (PPV) and negative predictive value (NPV) were compared using McNemar’s test. Differences in $\Delta\text{SUV}_{\text{max}}$% between different clinical outcomes were assessed using the Mann-Whitney U-test. A $p$-value of $<0.05$ was considered as statistically significant.

Results
Characteristics and outcomes of patients
A total of 84 patients were enrolled between March 2013 and December 2018. Table 1 lists the clinical characteristics of all patients. After a median follow-up of 34 months (range of 2–83 months), 22 patients showed relapse or disease progression at a median time of 18.5 months (range of 2–42 months), and six patients died at a median time of 19 months (2, 9, 14, 24, 34 and 36 months, respectively). The 1-year and 2-year PFS were 89.29% (75/84) and 83.33% (70/84), respectively, and the 1-year and 2-year OS were 97.62% (82/84) and 95.24% (80/84), respectively. The median PFS was 26.5 months (range of 2–75 months), and the median OS was 34 months (range of 2–84 months).

Baseline PET/CT
Clinical characteristics of patients in relation to tMTV, TLG and $\text{SUV}_{\text{max}}$
A total of 59 patients underwent B-PET. After a median follow-up of 34 months (range of 2–75 months), 13 patients showed relapse or disease progression (Figures 1 and 2), with
a median time of 24 months (range of 2–42 months), and five patients died with a median time of 24 months (range of 2–36 months).

**Table 2** lists the differences in clinical characteristics among the dichotomized tMTV, TLG and SUV$_{\text{max}}$ groups. Pearson chi-square test and Fisher’s exact test showed that the LDH level and FLIPI score were significantly associated with tMTV, TLG and SUV$_{\text{max}}$, whereas Ann Arbor stage, bone marrow biopsy, LodLIN (cm) and number of involved nodes were significantly associated with a higher tMTV and TLG, but not SUV$_{\text{max}}$. B symptoms and ECOG were significantly associated with a higher SUV$_{\text{max}}$ but not tMTV and TLG. Meanwhile, the number of extranodal sites was significantly associated with a higher tMTV and SUV$_{\text{max}}$, but not TLG.

**Correlation between clinical characteristics with semi-quantitative parameters**

Spearman correlation test showed that the Ann Arbor stage, number of extranodal sites, LDH level, FLIPI score, LodLIN (cm) and number of involved nodes were positively and significantly associated with SUV$_{\text{max}}$, tMTV or TLG, whereas bone marrow biopsy was positively and significantly associated with tMTV or TLG, but not SUV$_{\text{max}}$. On the other hand, there was a positive and significant association among SUV$_{\text{max}}$, tMTV and TLG. (Table 3)
Table 2 Comparison between low and high tMTV, TLG and SUV\textsubscript{max} groups

| Clinical Factors (n=59) | tMTV | TLG | SUV\textsubscript{max} |
|------------------------|------|-----|----------------------|
|                        | Low  | High| p        | Low  | High| p        | Low  | High| p        |
| Sex                    |      |     |          |      |     |          |      |     |          |
| Female (26)            | 17   | 9   | 0.151    | 22   | 4   | 0.152    | 19   | 7   | 0.417    |
| Male (33)              | 16   | 17  |          | 23   | 10  |          | 26   | 7   |          |
| Age                    |      |     |          |      |     |          |      |     |          |
| >60 (18)               | 9    | 9   | 0.372    | 13   | 5   | 0.431    | 12   | 6   | 0.205    |
| ≤60 (41)               | 24   | 17  |          | 32   | 9   |          | 33   | 8   |          |
| B symptoms             |      |     |          |      |     |          |      |     |          |
| Yes (18)               | 8    | 10  | 0.186    | 12   | 6   | 0.205    | 10   | 8   | 0.018*   |
| NO (41)                | 25   | 16  |          | 33   | 8   |          | 35   | 6   |          |
| Ann Arbor stage        |      |     |          |      |     |          |      |     |          |
| I/II (18)              | 18   | 0   | 0.000*   | 18   | 0   | 0.003*   | 15   | 3   | 0.311    |
| III/IV (41)            | 15   | 26  |          | 27   | 14  |          | 30   | 11  |          |
| ECOG                   |      |     |          |      |     |          |      |     |          |
| 0–1 (48)               | 29   | 19  | 0.133    | 38   | 10  | 0.236    | 40   | 8   | 0.015*   |
| ≥2 (11)                | 4    | 7   |          | 7    | 4   |          | 5    | 6   |          |
| No. of extranodal sites|      |     |          |      |     |          |      |     |          |
| <1 (36)                | 27   | 9   | 0.000*   | 30   | 6   | 0.101    | 32   | 4   | 0.006*   |
| ≥1 (23)                | 6    | 17  |          | 15   | 8   |          | 13   | 10  |          |
| BM                     |      |     |          |      |     |          |      |     |          |
| + (18)                 | 6    | 12  | 0.021*   | 10   | 8   | 0.018*   | 14   | 4   | 0.569    |
| - (41)                 | 27   | 14  |          | 35   | 6   |          | 31   | 10  |          |
| LDH                    |      |     |          |      |     |          |      |     |          |
| Increased (16)         | 5    | 11  | 0.021*   | 8    | 8   | 0.007*   | 8    | 8   | 0.007*   |
| - (43)                 | 28   | 15  |          | 37   | 6   |          | 37   | 6   |          |
| FLIPI                  |      |     |          |      |     |          |      |     |          |
| 1–2 (36)               | 28   | 8   | 0.000*   | 34   | 2   | 0.000*   | 32   | 4   | 0.006*   |
| 3–5 (23)               | 5    | 18  |          | 11   | 12  |          | 13   | 10  |          |
| LodLIN (cm)            |      |     |          |      |     |          |      |     |          |
| >6cm (14)              | 4    | 10  | 0.020*   | 6    | 8   | 0.002*   | 9    | 5   | 0.196    |
| <6cm (45)              | 29   | 16  |          | 39   | 6   |          | 36   | 9   |          |
| No. of involved nodes  |      |     |          |      |     |          |      |     |          |
| >4 (33)                | 10   | 23  | 0.000*   | 19   | 14  | 0.000*   | 23   | 10  | 0.152    |
| ≤4 (26)                | 23   | 3   |          | 36   | 0   |          | 22   | 4   |          |

Note: *Statistically significant.

Abbreviations: tMTV, total metabolic tumor volume; TLG, total lesion glycolysis; SUV\textsubscript{max}, maximum standardized uptake value; ECOG, Eastern Cooperative Oncology Group; BM, bone marrow; LDH, lactate dehydrogenase; FLIPI, Follicular Lymphoma International Prognostic Index.

Role of $^{18}$F-FDG PET/CT in predicting outcome

Table 4 summarizes the baseline PET/CT metabolic parameters for SUV$\textsubscript{max}$, tMTV and TLG. According to the Kaplan-Meier curve and the log-rank test, patients with high levels of tMTV and TLG had poorer clinical survival compared with those with low levels of tMTV and TLG.
Patients with a high level of tMTV had a 2-year PFS rate of 73.77%, while it became 96.97% in patients with a low level of tMTV (cutoff value 179.84 cm$^3$, $p=0.005$). The 2-year OS rates of the low- and high-tMTV groups were 100% and 70%, respectively (cutoff value 353.04 cm$^3$, $p=0.003$). Patients with a high level of TLG had a 2-year PFS rate of 57.14%, while it became 95.56% in patients with a low level of TLG (cutoff value 1,364.60, $p=0.0001$). The 2-year OS rates of the low- and high-TLG groups were 100% and 62.50%, respectively (cutoff value 2,398.00, $p<0.0001$).

By univariate analysis (Table 5), we found that Ann Arbor stage ($p=0.015$), SUV$\text{max}$ ($p=0.043$), tMTV ($p=0.005$), TLG ($p=0.000$), number of extranodal sites ($p=0.035$) and bone marrow biopsy ($p=0.019$) were significantly associated with PFS.

The tMTV ($p=0.003$), TLG ($p<0.0001$), ECOG ($p=0.009$) and bone marrow biopsy ($p=0.050$) were significantly associated with OS. Multivariate analysis showed (Table 6) that TLG was an independent prognostic factor for PFS ($p=0.001$, HR=6.577, 95% CI=2.167–19.960). TLG and ECOG were independent prognostic factors for OS ($p=0.030$, HR=19.291, 95% CI=2.689–137.947 and $p=0.028$, HR=8.924, 95% CI=1.273–62.559).

Interim PET/CT

Prognostic impact of D-5PS and IHP criteria

A total of 24 patients underwent I-PET after 3–4 cycles of chemotherapy (median of four cycles), among which 15 patients underwent I-PET after four cycles of chemotherapy, and nine patients were imaged after three cycles of chemotherapy. Kaplan–Meier survival analysis demonstrated that I-PET, using any of the evaluation criteria (D-5PS or IHP), was not a prognostic factor for PFS ($p>0.05$, Figure 4). By using the D-5PS criteria, 10

Table 4 Baseline $^{18}$F-FDG PET/CT metabolic parameters

| PFS          | OS           |
|--------------|--------------|
| AUC (95 %CI) | Sensitivity  | Specificity | Cutoff | p  |
| tMTV         | 0.727 (0.570–0.884) | 76.9% | 65.2% | 179.84 | 0.013 |
| TLG          | 0.712 (0.533–0.872) | 53.8% | 84.8% | 1364.60 | 0.020 |
| SUV$\text{max}$ | 0.602 (0.427–0.777) | 46.2% | 80.4% | 10.44 | 0.265 |
| AUC (95 %CI) | Sensitivity  | Specificity | Cutoff | p  |
| tMTV         | 0.759 (0.508–1.0) | 60% | 87% | 353.04 | 0.057 |
| TLG          | 0.696 (0.419–0.973) | 60% | 90.7% | 2398.00 | 0.149 |
| SUV$\text{max}$ | 0.552 (0.233–0.812) | 40% | 81.5% | 11.81 | 0.870 |

Abbreviations: tMTV, total metabolic tumor volume; TLG, total lesion glycolysis; SUV$\text{max}$, maximum standardized uptake value.
(41.67%) patients were categorized as score 4 or 5, and 14 (58.33%) patients were categorized as score 1 to 3. Moreover, the 2-year PFS rate was 92.86% for patients with a score of 1–3, and it became 70% for those with a score of 4–5 (p=0.059, Table 7, Figure 5). According to IHP criteria, 12 (50%) patients achieved CR. IHP-negative patients showed a higher PFS rate compared with IHP-positive patients, while the difference was not statistically significant (91.67 vs 75%, p=0.186, Table 7, Figure 5).

Table 8 illustrates the comparisons of the sensitivity, specificity, accuracy, PPV, NPV and AUC of D-5PS and IHP criteria. D-5PS criteria had a better specificity (68.4% vs 57.9%), accuracy (70.8% vs 62.5%), PPV (40% vs 33.3%) and NPV (92.9% vs 91.7%) compared with IHP.
|                                | 2-years PFS | \( \chi^2 \) | p   | 2-years OS | \( \chi^2 \) | p   |
|--------------------------------|-------------|-------------|-----|------------|-------------|-----|
| **B symptoms**                 |             |             |     |            |             |     |
| Yes                            | 83.33%      | 0.262       | 0.609| 94.44%     | 0.116       | 0.773|
| No                             | 87.80%      |             |     |            |             |     |
| **Ann Arbor stage**            |             |             |     |            |             |     |
| I/II                           | 100.00%     | 5.974       | 0.015*| 92.68%     | 3.347       | 0.067|
| III/IV                         | 80.49%      |             |     |            |             |     |
| **SUV\(_{\text{max}}\)**      |             |             |     |            |             |     |
| High                           | 64.29%      | 4.093       | 0.043*| 83.33%     | 0.812       | 0.368|
| Low                            | 93.33%      |             |     |            |             |     |
| **tMTV**                       |             |             |     |            |             |     |
| High                           | 73.77%      | 7.878       | 0.005*| 70.00%     | 9.139       | 0.003*|
| Low                            | 96.97%      |             |     |            |             |     |
| **TLG**                        |             |             |     |            |             |     |
| High                           | 57.14%      | 14.498      | 0.0001*| 62.50%     | 16.319      | <0.0001*|
| Low                            | 95.56%      |             |     |            |             |     |
| **FLIPI**                      |             |             |     |            |             |     |
| 1–2                            | 91.67%      | 1.452       | 0.228| 97.22%     | 0.103       | 0.748|
| 3–5                            | 78.26%      |             |     |            |             |     |
| **ECOG**                       |             |             |     |            |             |     |
| 0–1 (48)                       | 89.58%      | 1.897       | 0.168| 97.92%     | 6.766       | 0.009*|
| ≥2 (11)                        | 72.73%      |             |     |            |             |     |
| **No. of extranodal sites**    |             |             |     |            |             |     |
| <1                             | 91.67%      | 4.431       | 0.035*| 97.22%     | 1.306       | 0.253|
| ≥1                             | 78.26%      |             |     |            |             |     |
| **BM**                         |             |             |     |            |             |     |
| +                              | 72.22%      | 5.460       | 0.019*| 88.89%     | 3.827       | 0.050*|
| -                              | 92.68%      |             |     |            |             |     |
| **LDH**                        |             |             |     |            |             |     |
| Increased                      | 81.25%      | 0.275       | 0.600| 87.5%      | 0.489       | 0.484|
| Normal                         | 88.37%      |             |     |            |             |     |
| **LodLIN (cm)**                |             |             |     |            |             |     |
| >6 cm                          | 84.44%      | 1.153       | 0.283| 92.86%     | 0.019       | 0.892|
| <6 cm                          | 92.86%      |             |     |            |             |     |
| **No. of involved nodes**      |             |             |     |            |             |     |
| >4                             | 78.79%      | 2.656       | 0.103| 90.91%     | 0.308       | 0.579|
| ≤4                             | 96.15%      |             |     |            |             |     |

(Continued)
criteria, while the sensitivity (80% vs 80%) was similar between two criteria (all \( p > 0.05 \)).

**Prognostic impact of \( \Delta \text{SUV}_{\text{max}} \)%**

Of the 24 patients, 15 patients underwent baseline and interim PET/CT scans, and the \( \Delta \text{SUV}_{\text{max}} \)% between I-PET and B-PET was significantly different between the progression (\( n=4 \)) and progression-free groups (\( n=11 \)) (59.34% vs 76.22%, \( p=0.361 \)).

The ROC analysis showed that the optimal cutoff point was 66.95% (AUC: 0.659; sensitivity: 75%; specificity: 72.73%; accuracy: 73.33%; PPV: 50%; NPV: 88.89%) when using \( \Delta \text{SUV}_{\text{max}} \)% as a predictor of progression. Patients with lower \( \Delta \text{SUV}_{\text{max}} \)% (<66.95%) had low PFS compared with those with higher \( \Delta \text{SUV}_{\text{max}} \)% (>66.95%), while the difference was not statistically significant (66.67% vs 88.89%, \( \chi^2=2.096, p=0.148 \), Figure 6).

**End-of treatment PET/CT**

**Prognostic impact of D-5PS and IHP criteria**

A total of 43 patients underwent end-of treatment PET/CT after all planned first-line therapy. By using the D-5PS criteria, 14 (32.56%) patients were categorized as score 4–5, and 29 (67.44%) patients were categorized as score 1–3. According to IHP criteria, 22 (51.16%) patients showed CR. In patients with a negative PET/CT, the 2-year PFS rates for the D-5PS and IHP criteria were 93.10% and 95.45%, respectively, compared with the corresponding values of 35.71% and 57.14% in patients with a positive PET/CT, respectively (\( p=0.0001 \) for D-5PS, and \( p=0.006 \) for IHP, Figures 4 and 7).

D-5PS criteria had a better specificity (86.21% vs 65.52%), accuracy (81.40% vs 69.77%) and PPV (71.43% vs 52.38%) compared with IHP criteria (all \( p < 0.05 \)), while its sensitivity (71.43% vs 78.57%) and NPV (86.21% vs 86.36%) were slightly lower (all \( p > 0.05 \), Table 8).

**Prognostic impact of \( \Delta \text{SUV}_{\text{max}} \)%**

Of the 43 patients, 28 patients underwent both baseline and end-of treatment PET/CT scans, and the \( \Delta \text{SUV}_{\text{max}} \)% between E-PET and B-PET was significantly different between the progression (\( n=14 \)) and progression-free groups (\( n=14 \)) (41.70% vs 82.34%, \( p=0.003 \)).

The ROC analysis showed that the optimal cutoff point was 68.97% (AUC: 0.848; sensitivity: 77.78%; specificity: 89.47%; accuracy: 85.71%; PPV: 77.78%; NPV: 89.47%) when using \( \Delta \text{SUV}_{\text{max}} \)% as a predictor of progression. Patients with lower \( \Delta \text{SUV}_{\text{max}} \)% (<68.97%) had significantly lower PFS compared with those with higher \( \Delta \text{SUV}_{\text{max}} \)% (>68.97%) (22.2% vs 89.5%, \( \chi^2=13.774, p=0.0002 \), Figure 6).
**Discussion**

$^{18}$F-FDG PET/CT is currently a standard imaging technology for diagnosis, staging and prediction of prognosis in patients with HL or NHL. Over 95% of FL is FDG uptake. PET/CT is recommended for initial evaluation, staging and response assessment of FL in the Lugano 2014 International Conference on Malignant Lymphoma imaging consensus guidelines. However, the data of PET/CT in the baseline, interim and end-of-treatment of adult FL patients are very limited.
In the present study, we demonstrated that both baseline tMTV and TLG had the potential to predict PFS and OS for FL patients. Ann Arbor stage, number of extranodal sites, LDH level, FLIPI score, LodLIN (cm), number of involved nodes and bone marrow biopsy were positively and significantly associated with tMTV or TLG. Furthermore, TLG was the independent prognostic factor of PFS and OS. MTV and TLG included both anatomical and metabolic features, and

Table 8 Predictive values of interim and end-of treatment PET/CT

|                | Sensitivity | Specificity | Accuracy | PPV  | NPV  | AUC (95% CI) |
|----------------|-------------|-------------|----------|------|------|--------------|
| I-PET (n=24)   |             |             |          |      |      |              |
| D-5PS          | 80.00%      | 68.42%      | 70.83%   | 40.00% | 92.86% | 0.742 (0.499–0.985) |
| IHP            | 80.00%      | 57.89%      | 62.50%   | 33.33% | 91.67% | 0.689 (0.436–0.943) |
| E-PET (n=43)   |             |             |          |      |      |              |
| D-5PS          | 71.43%      | 86.21%*     | 81.40%*  | 71.43%* | 86.21% | 0.788 (0.629–0.947) |
| IHP            | 78.57%      | 65.52%      | 69.77%   | 52.38% | 86.36% | 0.720 (0.557–0.884) |

Note: *Statistically significant.
Abbreviations: PPV, positive predictive value; NPV, negative predictive value.

Figure 5 A patient with interim $^{18}$F-FDG PET/CT showed increased $^{18}$F-FDG uptake in the neck, axilla, mediastinum, abdominal and pelvis. D-5PS and IHP criteria were considered positive for patient, and the patient experienced relapse after 7 months of follow-up. A patient with an interim $^{18}$F-FDG PET/CT D-5PS (score 1) and negative IHP criteria did not show progression and survived at the end of the 28-month follow-up period (B).

Figure 6 Kaplan-Meier survival curves for the PFS of $\Delta \text{SUV}_{\text{max}}\%$. The I-PET results according to $\Delta \text{SUV}_{\text{max}}\%$ (A) were not associated with PFS; and the E-PET results according to $\Delta \text{SUV}_{\text{max}}\%$ (B) were associated with PFS.

Graphs showing progression-free survival (%) over time (mo) for different SUVmax% thresholds.
might be considered to be tumor invasive as well as expression of tumor volume. Recently, several studies have demonstrated the prognostic value of semi-quantitative parameters in lymphoma. Zhou et al have reported that higher TLG level is associated with a poorer survival in patients with DLBCL. Pak et al in a multicenter retrospective study have shown that high TLG is the only independent factor for predicting survival in extranodal nasal type NK/T cell lymphoma. Combination of early PET/CT response and baseline MTV or TLG improves the predictive power of interim PET in DLBCL. Other studies have also demonstrated the significant prognostic values of MTV and TLG in FL. Cottereau et al have found that MTV is an independent predictor of PFS and OS in patients with high-tumor burden FL. A pooled analysis consisting of 185 patients with high-tumor burden FL indicates that baseline tMTV is an independent predictor of PFS. Although we found the relationship between tMTV and prognosis in univariate analysis, the multivariate analysis indicated that TLG was an independent predictive factor of PFS and OS. This discrepancy might be attributed to many factors. First, TLG was not included in their study. Furthermore, PET/CT had low sensitivity when detecting bone marrow involvement, which might affect the calculation of tMTV. Last, the distributions of risk groups in patients were different.

Recent studies have shown that the optimal cutoff points for total MTV are different. This variation can be explained by different characteristics of the patients and the marginal threshold methods. Patients with advanced stages, bulky disease and high FLIPI score had higher tMTV compared with the patients with earlier stage, less bulky disease and low FLIPI score. For example, Meignan et al have included patients from three prospective multicenter trials, most with stage III-IV (92%), with LodLIN >6 cm in 47% and two or more extra-nodal sites in 38%, and found an optimal cutoff of 510 cm. Liang et al have found an optimal tMTV cutoff of 476.4 cm in patients, of which 75% have stage III-IV disease, 43.8% have FLIPI1 score of 3–5 and 20.8% have FLIPI2 score of 3–5. Song et al have found an optimal tMTV cutoff of 220 cm in patients with only stage II and III disease, 4.1% with bulky disease (>5 cm). Another factor affecting tMTV cutoff is the different ways used to measure the marginal threshold. In our study, the tMTV was measured using the 41% SUV threshold recommended by the European Association of Nuclear Medicine guidelines because of its high inter-observer reproducibility. However, tMTV is measured using absolute values (2.0, 2.5 and 3.0) as the threshold in Song’s studies. As previously reported, absolute threshold, such as SUV ≥2.5, may underestimate the volume of lesions with low SUV, which is less than the threshold value. Although our cutoff value was different, tMTV and TLG were robust prognostic indicators of patient survival.

Recently, the prognostic role of I-PET has been emphasized. Lu et al have conducted a retrospective analysis on 57 FL patients (grade 1, 2 and 3a), and demonstrated a poor prognostic value for the IHP criteria in
mid-treatment PET scans for the prediction of PFS and OS. Bishu et al\textsuperscript{32} have shown similar results that no significant difference in PFS between I-PET positive and negative patients. However, in a prospective study consisting of 112 patients, Dupuis et al\textsuperscript{30} have shown that the D-5PS criteria of the PET after four cycles of chemotherapy have a strong predictive value for PFS. In the present study, we confirmed that I-PET had no predictive value for PFS in FL patients, regardless of D-5PS or IHP criteria. Therefore, the utility of I-PET still remains controversial in evaluating the response to FL. These controversies might be attributed to many factors. First of all, FL is an incurable disease, the absence of FDG-uptake lesions in I-PET imaging does not mean the absence of tumor cells, while most of the tumor cells have responded and \textsuperscript{18}F-FDG-PET/CT cannot detect these cells.\textsuperscript{31} In addition, the follow-up time, the patient population, the treatment regimen, and the number of chemotherapy cycles can all cause these controversies.\textsuperscript{33}

The utility of PET/CT in assessing response after end-of treatment has been confirmed in several studies.\textsuperscript{6,11,19,32} In accordance with previous findings, our study confirmed that E-PET had predictive value for PFS in FL patients, and PFS in PET-positive patients was significantly lower compared with PET-negative patients. However, only very few studies have investigated the relationship between PET/CT results and OS at the end-of treatment.\textsuperscript{13} In our present study, a small number of patients died for progressive FL, and the OS was not analyzed. Therefore, the role of E-PET in FL remains controversial.

Visual interpretation of D-5PS seems to be a better prognostic value than IHP criteria, with a higher referenced background (liver).\textsuperscript{12,34–36} In our study, we found that the D-5PS criteria had a higher specificity, PPV and accuracy compared with the IHP criteria in the interim and the end-of treatment. Interestingly, we also found that the specificity, accuracy and PPV of PET/CT at the end-of chemotherapy, regardless of D-5PS or IHP criteria, were significantly higher compared with interim PET/CT. This finding was consistent with previous conclusions,\textsuperscript{37} indicating that E-PET might have a stronger diagnostic value than I-PET. However, additional studies are still necessary.

The percentage of $\Delta$SUV\textsubscript{max} is a semi-quantitative method with excellent inter-observer agreement and improved prognostic value of I-PET\textsuperscript{38,39} and E-PET.\textsuperscript{40,41} Rossi et al\textsuperscript{38} have performed an interim PET/CT after two cycles of chemotherapy in HL patients, and shown a promising prognostic value using the criteria of $\Delta$SUV\textsubscript{max} >71\%. Itti et al\textsuperscript{40} have found that $\Delta$SUV\textsubscript{max} >72.9\% is an important predictor of PFS at the end of treatment in DLBCL patients. In our study, we found that $\Delta$SUV\textsubscript{max} using a cutoff of 66.97\% had predictive value for FL patients after first-line chemotherapy. However, we failed to show a better predictive value for interim PET/CT when using a cutoff value of 66.95\%. To the best of our knowledge, studies on the evaluation criteria of $\Delta$SUV\textsubscript{max} in FL patients are limited, and a large number of prospective studies are required.

There are some limitations and shortcomings in this study. This was a single-center retrospective study, in which only 59 patients underwent PET/CT scans before treatment, only 24 patients had 2–4 cycles of chemotherapy, and only 43 patients had all planned first-line therapy. Since a significant proportion of patients had a good therapeutic response after chemotherapy and a small number of patients died for progressive FL, the OS was not analyzed with I-PET and E-PET.

**Conclusion**

Collectively, we demonstrated that baseline TLG was an independent predictor of PFS and OS in FL. E-PET results using D-5PS and IHP criteria were significantly associated with PFS, whereas I-PET results were not associated with PFS. D-5PS criteria showed a better sensitivity, accuracy and PPV compared with IHP criteria in I-PET or E-PET. In addition, the $\Delta$SUV\textsubscript{max} from the baseline to the end-of therapy could be used for precise prediction of patient prognosis.

**Data sharing statement**

The datasets generated during the current study are available from the corresponding authors for reasonable requests.

**Acknowledgments**

This work was financially supported by National Natural Science Foundation of China (No. 81601522), Natural Science Foundation of Jiangsu Province (No. BK20160348), and Medical Youth Talent Project of Jiangsu Province (No. QNRC2016749).

**Disclosure**

The authors report no conflicts of interest in this work.
References

1. Meignan M, Cottereau AS, Versari A, et al. Baseline metabolic tumor volume predicts outcome in high-tumor-burden follicular lymphoma: a pooled analysis of three multicenter studies. J Clin Oncol. 2016;34(30):3618–3626. doi:10.1200/JCO.2016.66.9440

2. Casulo C, Byrtek M, Dawson KL, et al. Early relapse of follicular lymphoma after rituximab plus cyclophosphamide, doxorubicin, vin-cristine, and prednisone defines patients at high risk for death: an analysis from the national lympho care study. J Clin Oncol. 2015;33:2516–2522. doi:10.1200/JCO.2014.59.7534

3. Luminari S, Biasoli I, Versari A, et al. The prognostic role of post-induction FDG-PET in patients with follicular lymphoma: a subset analysis from the FOLL05 trial of the FondazioneItalianaLinfomi (FIL). Ann Oncol. 2014;25:442–447. doi:10.1093/annonc/mdt562

4. Chang CC, Cho SF, Chuang YW, et al. Prognostic significance of total metabolic tumor volume on 18F-fluorodeoxyglucose positron emission tomography/computed tomography in patients with diffuse large B-cell lymphoma receiving rituximab-containing chemotherapy. Oncotarget. 2017;8(59):99587–99600. doi:10.18632/oncotarget.20447

5. 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(8):950–956. doi:10.1182/blood-2014-12-616474

6. Trotman J, Barrington SF, Belada D, et al. Prognostic value of end-of-induction PET response after first-line immunotherapy for follicular lymphoma (GALLI UM): secondary analysis of a randomised, phase 3 trial. Lancet Oncol. 2018;19:1530–1542. doi:10.1016/S1470-2045(18)30618-1

7. Kim HJ, Lee R, Choi H, et al. Application of quantitative indexes of FDG PET to treatment response evaluation in indolent lymphoma. Nucl Med Mol Imaging. 2018;52:342–349. doi:10.1007/s13139-018-0543-8

8. Albanò D, Bosio G, Pagani C, et al. Prognostic role of baseline 18F-FDG PET/CT metabolic parameters in Burkitt lymphoma. Eur J Nucl Med Mol Imaging. 2019;46:87–96. doi:10.1007/s00259-018-4173-2

9. Zhou M, Chen Y, Huang H, et al. Prognostic value of total lesion glycolysis of baseline 18F-fluorodeoxyglucose positron emission tomography/computed tomography in diffuse large B-cell lymphoma. Oncotarget. 2016;7:83544–83553. doi:10.18632/oncotarget.13180

10. Moskowitz AJ, Schöder H, Gavane S, et al. Prognostic significance of baseline metabolic tumor volume in relapsed and refractory Hodgkin lymphoma. Blood. 2017;130(20):2196–2203.

11. Annunziata S, Cuccaro A, Tisi MC, et al. FDG-PET/CT at the end of immuno-chemotherapy in follicular lymphoma: the prognostic role of the ratio between target lesion and liver SUV (rPET). Ann Nucl Med. 2018;32:372–377. doi:10.1007/s12419-018-1243-2

12. Minamimoto R, Fayad L, Martinelli G, et al. FDG-PET is prognostic and predictive for progression-free survival in relapsed follicular lymphoma: exploratory analysis of the GAUSS study. Leuk Lymphoma. 2017;58:372–381. doi:10.1080/10428194.2016.1196815

13. Adams HJA. No convincing evidence to support postinduction FDG-PET in follicular lymphoma. Blood. 2018;131:404–419. doi:10.1053/j.seminoncol.2018.01.005

14. Kostakoglu L, Goy A, Martinelli G, et al. FDG-PET is prognostic and predictive for progression-free survival in relapsed follicular lymphoma: exploratory analysis of the GAUSS study. Leuk Lymphoma. 2017;58:372–381. doi:10.1080/10428194.2016.1196815

15. Metser U, Hussey D. Impact of (18) F-FDG PET/CT on the staging and management of follicular lymphoma. Br J Radiol. 2014;87:20140360. doi:10.1259/bjr.20140360

16. Brady JL, Binkley JM, Haji C, et al. Definitive radiotherapy for localized follicular lymphoma staged by F-FDG PET-CT: a collaborative study by ILROG. Blood. 2019;133:237–245. doi:10.1182/blood-2018-04-843540

17. Adams HJA. Overestimated value of baseline total metabolic tumor volume and total lesion glycolysis measured on 18F-FDG PET/CT in patients with follicular lymphoma. Cancer Res Treat. 2019;12. doi:10.4143/crt.2018.649

18. Cottereau AS, Versari A, Luminari S, et al. Prognostic model for high-tumor-burden follicular lymphoma integrating baseline and end-induction PET. A LYSA/F IL study. Blood. 2018;131:2449–2453. doi:10.1182/blood-2017-11-816298

19. Adams HJ. Overestimated value of baseline total metabolic tumor volume at (18) F-labeled fluorodeoxyglucose positron emission tomography in follicular lymphoma. J Clin Oncol. 2017;35:918–919. doi:10.1200/JCO.2016.70.4890

20. Song MK, Chung JS, Shin HJ, et al. 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. 2012;91:697–703. doi:10.1007/s00277-011-1357-2

21. Barrington SF. Time to prepare for risk adaptation in lymphoma by standardising measurement of metabolic tumour burden. J Nucl Med. 2019;jnumed.119.227249. doi:10.2967/jnumed.119.227249

22. Dupuis J, Berriolo-Riedinger A, Julian A, et al. Impact of [18F] fluoro-deoxyglucose positron emission tomography response evaluation in patients with high-tumor burden follicular lymphoma treated with immuno-chemotherapy: a prospective study from the Grouped‘Etudes des Lymphomes de l’Adulte and GOELAMS. J Clin Oncol. 2012;30:2437–2442. doi:10.1200/JCO.2012.43.0934

23. Lu Z, Lin M, Downe P, et al. The prognostic value of mid- and post-treatment [(18) F]fluorodeoxyglucose (FDG) positron emission tomography (PET) in indolent follicular lymphoma. Ann Nucl Med. 2014;28:805–811. doi:10.1259/bjr.2014-0874-1

24. Bishu S, Quigley JM, Bishu SR, et al. Predictive value and diagnostic accuracy of F-18-fluoro-deoxy-glucose positron emission tomography treated grade 1 and 2 follicular lymphoma. Leuk Lymphoma. 2007;48:1548–1555. doi:10.1080/10428190701422059

25. Zhang Y, Fan Y, Ying Z, et al. Can the SUVmax-liver-based interpretation improve prognostic accuracy of interim and posttreatment 18F-FDG PET/CT in patients with diffuse large B-cell lymphoma? Leuk Lymphoma. 2018;59(3):669–666. doi:10.1080/10428194.2017.1357171
34. Lombion N, Robin P, Tempescul A, et al. Prognostic value of interim FDG PET-CT in patients older than 60 years with diffuse large B-cell lymphoma treated by PMitCEBO plus Rituximab: comparison between Deauville 5-point scale and International Harmonization Project criteria. *Q J Nucl Med Mol Imaging*. Epub 2016 Nov 30.

35. Jiang C, Su M, Kosik RO, et al. The Deauville 5-point scale improves the prognostic value of interim FDG PET/CT in extranodal natural killer/T-cell lymphoma. *Clin Nucl Med*. 2015;40:767–773. doi:10.1097/RLU.0000000000000892

36. Metser U, Mohan R, Beckley V, et al. FDG PET/CT response assessment criteria for patients with Hodgkin’s and Non-Hodgkin’s lymphoma at end of therapy: a multiparametric approach. *Nucl Med Mol Imaging*. 2016;50:46–53. doi:10.1007/s13139-015-0368-7

37. Del PuigCózar-Santiago M, García-Garzón JR, Moragas-Freixa M, et al. Optimisation of metabolic criteria in the prognostic assessment in patients with lymphoma. A multicentre study. *Rev Esp Med Nucl Imagen Mol*. 2017;36:304–311. doi:10.1016/j.remn.2017.03.003

38. Rossi C, Kanoun S, Berriolo-Riedinger A, et al. Interim 18F-FDG PET SUVmax reduction is superior to visual analysis in predicting outcome early in Hodgkin lymphoma patients. *J Nucl Med*. 2014;55:569–573. doi:10.2967/jnumed.113.130609

39. Fan Y, Zhang Y, Yang Z, et al. Evaluating early interim fluorine-18 fluorodeoxyglucose positron emission tomography/computed tomography with the SUV-based interpretation for predicting the outcome in diffuse large B-cell lymphoma. *Leuk Lymphoma*. 2017;58:1–9. doi:10.1080/10428194.2016.1277384

40. Itti E, Lin C, Dupuis J, et al. Prognostic value of interim 18F-FDG PET in patients with diffuse large B-Cell lymphoma: SUV-based assessment at 4 cycles of chemotherapy. *J Nucl Med*. 2009;50:527–533. doi:10.2967/jnumed.108.057703

41. Wei WX, Huang JJ, Li WY, et al. Prognostic values of interim and post-therapy 18F-FDG PET/CT scanning in adult patients with Burkitt’s lymphoma. *Chin. J Cancer*. 2015;34:608–613. doi:10.1186/s40880-015-0057-z