ORIGINAL ARTICLE

Early risk stratification for diffuse large B-cell lymphoma integrating interim Deauville score and International Prognostic Index

Sung Kyun Yim [1] · Ho-Young Yhim [1,2] · Yeon-Hee Han [3] · So-Yeon Jeon [1,2] · Na-Ri Lee [1,2] · Eun-Kee Song [1,2] · Hwan-Jeong Jeong [2,3] · Hee Sun Kim [4] · Jae-Yong Kwak [1,2]

Abstract

The aim of this study was to evaluate the prognostic relevance of early risk stratification in diffuse large B-cell lymphoma (DLBCL) using interim Deauville score on positron emission tomography-computed tomography (PET-CT) scan and baseline International Prognostic Index (IPI). This retrospective study included 220 patients (median age, 64 years; men, 60%) diagnosed with DLBCL between 2007 and 2016 at our institution, treated with rituximab-based chemotherapy. Interim PET-CT was performed after three cycles of immuno-chemotherapy. Interim Deauville score was assessed as 4 or 5 in 49 patients (22.3%), and 94 patients (42.7%) had high-intermediate or high-risk IPI scores. In multivariate analysis, interim Deauville score (1–3 and 4–5) and baseline IPI (low/low-intermediate and high-intermediate/high) were independently associated with progression-free survival (for Deauville score, hazard ratio [HR], 1.00 *vs.* 2.96 [95% confidence interval (CI), 1.83–4.78], $P < 0.001$; for IPI, HR, 1.00 *vs.* 4.84 [95% CI, 2.84–8.24], $P < 0.001$). We stratified patients into three groups: low-risk (interim Deauville scores 1–3 and low/low-intermediate IPI), intermediate-risk (Deauville scores 1–3 with high-intermediate/high IPI or Deauville scores 4–5 with low/low-intermediate IPI), and high-risk (Deauville scores 4–5 and high-intermediate/high IPI). This early risk stratification showed a strong association with progression-free survival (HR, 1.00 *vs.* 3.98 [95% CI 2.10–7.54] *vs.* 13.97 [95% CI 7.02–27.83], $P < 0.001$). Early risk stratification using interim Deauville score and baseline IPI predicts the risk of disease progression or death in patients with DLBCL. Our results provide guidance with interim PET-driven treatment intensification strategies.

Keywords Diffuse large B-cell lymphoma · Positron emission tomography-computed tomography · International Prognostic Index · Risk stratification · Deauville score

✉ Jae-Yong Kwak
jykwak@jbnu.ac.kr

1   Department of Internal Medicine, Chonbuk National University Medical School, 20 Geonji-ro, Deokjin-gu, Jeonju 54907, Republic of Korea

2   Research Institute of Clinical Medicine of Chonbuk National University-Biomedical Research Institute of Chonbuk National University Hospital, Jeonju, Republic of Korea

3   Department of Nuclear Medicine, Chonbuk National University Medical School, Jeonju, Republic of Korea

4   College of Nursing, Chonbuk National University, Jeonju, Republic of Korea

Introduction

Despite significant improvements in the treatment of diffuse large B-cell lymphoma (DLBCL) with the introduction of rituximab, approximately 30 to 40% of patients still experience treatment failure [1]. High-dose chemotherapy with autologous stem cell transplantation has been the backbone of current salvage treatments for younger patients with relapsed or refractory DLBCL; however, only a third of patients achieve long-term survival [2]. Therefore, numerous attempts have been undertaken to allow early identification of patients at risk of treatment failure and to provide risk-stratified therapy prior to this failure.

Positron emission tomography-computed tomography (PET-CT) with ^{18}F-fluorodeoxyglucose (^{18}F-FDG) has shown high sensitivity in detection of viable DLBCL lesions at baseline staging [3] and in response assessments after the end of primary treatment [4]. The Deauville 5-point scale has become the standard tool for assessment of treatment response in lymphoma [5], which compare ^{18}F-FDG uptake of the lesions with the uptake in the mediastinal pool and liver [6]. Interim PET-CT scans that were performed during treatment have been evaluated to identify patients at high risk of treatment failure, and the results of meta-analysis [7] along with multiple retrospective [8–12] and prospective [13, 14] studies have demonstrated an association of interim PET-CT scan with prognosis in patients with DLBCL. However, these studies have also shown that a substantial number of patients might experience prolonged survival in remission despite positive interim PET-CT scans, indicating a low positive predictive value of interim scans in DLBCL [7–14]. Accordingly, the overall accuracy of interim PET-CT adopting the Deauville score in predicting treatment outcomes is considerably low in DLBCL. Moreover, in a recent randomized phase 3 trial evaluating the influence on outcomes of interim PET-driven treatment intensification in patients with aggressive lymphoma, treatment intensification in interim PET-positive patients was not associated with improvement in outcomes [15]. These observations call into question the application of first-line risk-stratified strategies using interim PET-CT results in DLBCL and suggest that additional tools might be required to reduce the rate of false-positive interim PET-CT scans.

The International Prognostic Index (IPI) is meaningfully associated with treatment outcomes in DLBCL and considered to reflect the biologic aggressiveness of lymphoma [16]. The strong association between IPI scores and treatment outcomes of DLBCL might imply that FDG-avid lesions on interim PET-CT scans have a different influence on outcomes according to baseline IPI score. Given the low disease progression rates in patients with lower risk IPI scores, a significant portion of FDG-avid lesions on interim scans in these patients might be false positives, potentially associated with chemotherapy-induced inflammatory changes around tumor tissues [17]. In contrast, considering the higher rates of disease progression in patients with higher risk IPI scores, the interim PET-CT positive lesions may indicate an actual chemoresistant status of DLBCL. This idea is supported by the observations of a recent phase 3 trial [15], in which survival of interim PET-positive patients with high IPI scores was substantially lower than the survival of those with other IPI scores. Thus, we hypothesized that combined assessments using interim Deauville score as an estimate of early metabolic response and baseline IPI as an indicator of biologic aggressiveness might improve early prediction of outcomes in patients with DLBCL. The aim of the present study, therefore, was to determine the predictive value of risk stratification with integration of the Deauville score on interim PET-CT scan and IPI at diagnosis in patients with DLBCL treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) immuno-chemotherapy.

Material and methods

Study cohort

We conducted a retrospective study of data from patients who were diagnosed with DLBCL in the Chonbuk National University Hospital, Jeonju, South Korea. Between January 2007 and June 2016, patients with newly diagnosed, histologically proven DLBCL, who were treated with R-CHOP immuno-chemotherapy and for whom imaging data for both baseline and interim (after 3 cycles of R-CHOP) ^{18}F-FDG PET-CT scans were available, were consecutively enrolled in this analysis. Patients diagnosed with recurrent or secondary transformed DLBCL were excluded, as were patients with primary or secondary central nervous system involvement at baseline.

Baseline clinical assessments included medical history; determination of Eastern Cooperative Oncology Group (ECOG) performance status; full laboratory work-up with lactate dehydrogenase; CT scans of the chest, abdomen, and pelvis; bilateral bone marrow trephine biopsies; and PET-CT scan. Patients were staged according to the Ann Arbor staging system, and the IPI at baseline was determined for prognosis. Bone marrow involvement was defined only when lymphoma infiltration was identified histopathologically by trephine biopsy, and bulky disease was defined as any mass with a maximum diameter greater than 10 cm or any mediastinal mass exceeding 1/3 of the maximum transthoracic diameter. Standard R-CHOP treatment consisted of six or eight cycles of rituximab 375 mg/m^2, cyclophosphamide 750 mg/m^2, doxorubicin 50 mg/m^2, and vincristine 1.4 mg/m^2, intravenously at first day, and prednisone 100 mg orally for 5 days every 21 days. Before 2013, granulocyte colony-stimulating factor (G-CSF) was not administered if patients did not

experience the neutropenic side effects of chemotherapy. After January 2014, pegfilgrastim (6 mg, subcutaneously) was recommended for all patients with standard R-CHOP on day 2 of each treatment. Treatment was performed as planned, and no therapy changes were made on the basis of the interim PET-CT scan results unless disease progression was clearly documented. The study protocol was approved by the institutional review board of Chonbuk National University Hospital, which granted a waiver of informed consent because this study was a retrospective analysis with minimal risk for patients.

PET-CT procedures

All patients included underwent PET-CT scans prior to the administration of the first R-CHOP (baseline) and after three cycles of R-CHOP (interim). Interim PET-CT scan was generally scheduled at third week after three cycles of R-CHOP, but postponed if G-CSF was administered within 48 h of scheduled scan. End-of-treatment PET-CT scan was scheduled approximately 4 to 6 weeks after the end of R-CHOP treatment.

The PET-CT protocol has been previously described in detail [18]. Briefly, patients fasted for at least 6 h, and blood glucose levels of < 140 mg/dL were required in all patients prior to the intravenous injection of ^{18}F-FDG. Scanning was performed approximately 45–60 min after the injection of ^{18}F-FDG (5.5 MBq/Kg) using one of two dedicated PET-CT scanners (Biograph TruePoint 40 or Biograph 16; Siemens Medical Solutions, Knoxville, TN, USA). PET data were reconstructed iteratively using ordered-subset expectation maximization algorithm, and initial CT data were used for attenuation correction. Interim and end-of-treatment PET-CT scans were performed with the same camera and reconstruction algorithm as used in the baseline scan.

Interpretation criteria for PET-CT

Interim and end-of-treatment PET-CT scans were compared with baseline scans and assessed according to the Deauville criteria on the 5-point scale [6]. An experienced nuclear medicine physician (Han YH) reviewed all PET-CT images and interpreted interim and end-of-treatment scans using the Deauville criteria. Another experienced nuclear medicine physician (Jeong HJ) determined the scores of PET-CT scans that were assigned a score of 2, 3, or 4. Cases of differences in the Deauville score between two nuclear medicine physicians were resolved by direct contact and consensus through mutual discussion. They were blinded to treatment and survival outcomes.

Statistical analysis

The purpose of this analysis was to evaluate the association between progression-free survival (PFS) and interim PET-CT results in combination with baseline IPI. PFS and overall survival (OS) were defined as the time from the date of diagnosis to the date of first documented progression of disease, death from any cause, or last follow-up, as appropriate. PFS and OS were estimated using the Kaplan-Meier method. A refractory disease was defined as progressive disease during first-line treatment, stable disease as best response to ≥ 4 cycles of immunochemotherapy, or relapse ≤ 12 months after completion of first-line treatment [19]. Univariate analysis was performed to determine the association between PFS and clinical variables using the log-rank test. The clinical variables with $P < 0.05$ in univariate analysis were included in the multivariable analysis using the Cox proportional hazard model, and the results were reported as a hazard ratio (HR) and 95% confidence intervals (CIs). Descriptive analysis was expressed as a percentage for categorical variables and as median and interquartile range (IQR) for continuous variables. A two-sided $P < 0.05$ was considered to be significant, and all data analyses were performed using SPSS software, version 19.0 (SPSS Inc., Chicago, IL).
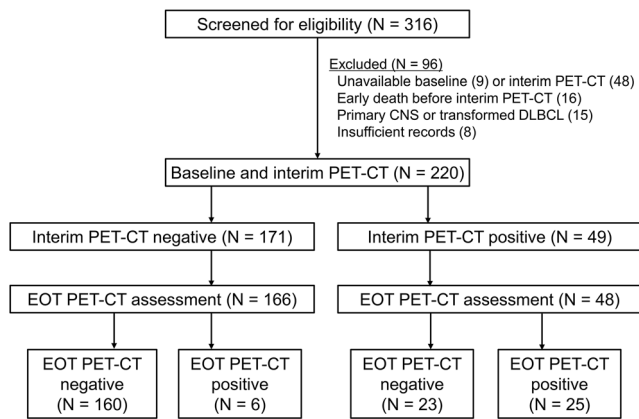
Results

Baseline characteristics of the study cohort

A total of 316 patients with newly diagnosed DLBCL were screened for eligibility. Of these, 96 patients were ineligible for this study; finally, the study cohort included 220 patients (Fig. 1).

The pretreatment characteristics of the 220 patients are summarized in Table 1. The median patient age was 64 years (range, 19–87 years) at diagnosis, and men were predominant ($N = 132$, 60%). Most patients had good performance status (ECOG grade 0 or 1, 79.1%) and involvement of fewer than two extranodal sites (75.5%). Thus, more than half of patients were classified as low or low-intermediate risk based on the IPI (57.3%). A pegfilgrastim was used prior to interim PET scanning in 77 patients (35.0%).

Interim Deauville scores, survival outcomes, and prognostic factors

In 220 evaluable interim scans, median time to perform interim PET-CT from preceding R-CHOP was 17 days (range, 12-33 days). Reviewers determined the Deauville score on all interim scans as follows: 1 ($N = 67$, 30.5%), 2 ($N = 65$, 29.5%), 3 ($N = 39$, 17.7%), 4 ($N = 36$, 16.4%), and 5 ($N = 13$, 5.9%). Thus, 49 patients (22.3%) with score of 4 or 5 were regarded as having positive interim PET-CT scans based on the Deauville criteria (Fig. 1). The strategy to determine the Deauville score (positive

```
                    Screened for eligibility (N = 316)
                                    |
                                    |  Excluded (N = 96)
                                    |     Unavailable baseline (9) or interim PET-CT (48)
                                    |     Early death before interim PET-CT (16)
                                    |     Primary CNS or transformed DLBCL (15)
                                    |     Insufficient records (8)
                                    |
                    Baseline and interim PET-CT (N = 220)
                          |                              |
                          |                              |
         Interim PET-CT negative (N = 171)    Interim PET-CT positive (N = 49)
                          |                              |
                          |                              |
         EOT PET-CT assessment (N = 166)      EOT PET-CT assessment (N = 48)
                    |         |                     |         |
                    |         |                     |         |
        EOT PET-CT      EOT PET-CT       EOT PET-CT      EOT PET-CT
        negative        positive         negative        positive
        (N = 160)       (N = 6)          (N = 23)        (N = 25)
```

Fig. 1 Study flow. All PET-CT scans were assessed according to Deauville criteria on 5-point scale, and a Deauville score of 4 or 5 was regarded as positive. *CNS*, central nervous system; *DLBCL*, diffuse large B cell lymphoma; *EOT*, end-of-treatment; *PET-CT*, positron emission tomography-computed tomography

or negative) in our study showed an overall agreement in 87.5% between two reviewers (Table S1).

At the time of this analysis, 70 patients had experienced relapse or progression and 61 had died, including three patients with non-disease-related deaths. With a median follow-up of 56.6 months (IQR, 36.0–71.8), the estimated 5-year PFS rate was 65.2% (95% CI, 58.1–72.3) and the OS rate was 69.9% (95% CI, 63.2–76.6). The 5-year PFS and OS for patients with negative interim PET-CT were 72.6% (95% CI, 65.0–80.2) and 78.1% (95% CI, 71.0–85.2), respectively, compared with 39.3% (95% CI, 25.0–53.6) and 42.4% (95% CI, 27.3–57.5), respectively, for patients with positive interim PET-CT. Comparisons based on interim Deauville score revealed significantly better PFS and OS in patients with negative interim PET-CT scans than in patients with positive scan results (PFS, HR 3.37, 95% CI 2.09–5.43, $P < 0.001$; OS, HR 3.53, 95% CI 2.13–5.56, $P < 0.001$; Fig. 2a, b). Comparison of survival outcomes according to IPI score revealed that PFS and OS were significantly longer in patients with low or low-intermediate IPI scores than in those with high-intermediate or high IPI scores (PFS, HR 5.21, 95% CI 3.07–8.85, $P < 0.001$; OS, HR 6.31, 95% CI 3.46–11.50, $P < 0.001$; Fig. 2c, d).

Table 1 Pre-treatment characteristics of the 220 enrolled patients

		Study cohort (N = 220)
Age	Median (range)	64 (19–87)
	≤ 60 years	86 (39.1)
	> 60 years	134 (60.9)
Sex	Male	132 (60.0)
	Female	88 (40.0)
Ann Arbor stage	I–II	96 (43.6)
	III–IV	124 (56.4)
Performance status	ECOG 0 or 1	174 (79.1)
	ECOG ≥ 2	46 (20.9)
B symptoms	Absence	162 (73.6)
	Presence	58 (26.4)
Serum LDH	Normal	88 (40.0)
	Increased	132 (60.0)
No. of extranodal organ involvements	0–1	166 (75.5)
	≥ 2	54 (24.5)
Bulky disease	No	200 (90.9)
	Yes	20 (9.1)
Bone marrow involvement	No	174 (79.1)
	Yes	46 (20.9)
Treatment	R-CHOP alone	202 (91.8)
	R-CHOP plus radiotherapy	18 (8.2)
IPI	Low or low-intermediate	126 (57.3)
	0–1	68 (30.9)
	2	58 (26.4)
	High-intermediate or high	94 (42.7)
	3	43 (19.5)
	4–5	51 (23.2)

ECOG Eastern Cooperative Oncology Group; *LDH* lactate dehydrogenase; *R-CHOP* rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone; *IPI* International Prognostic Index
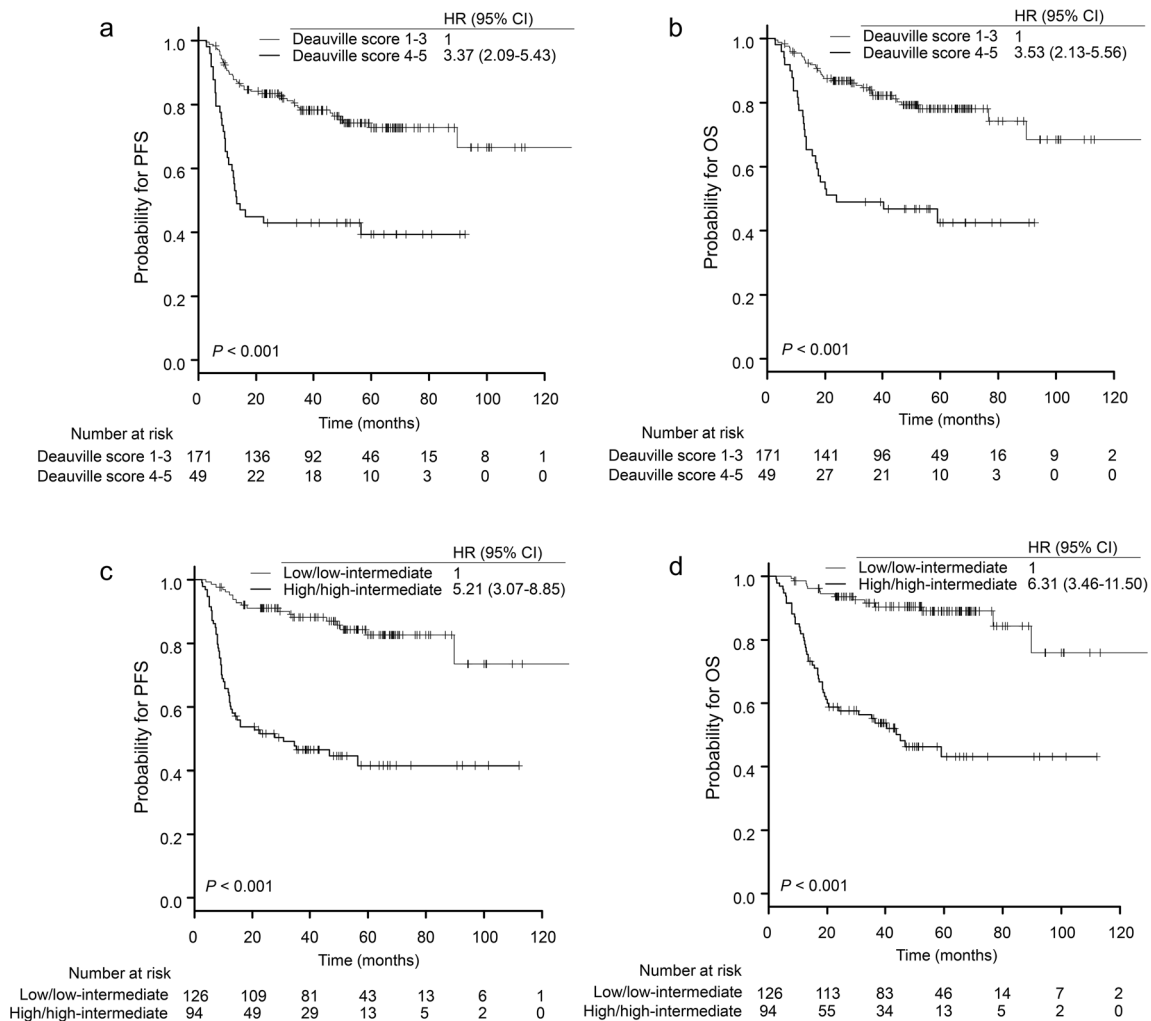
Fig. 2 Progression-free survival and overall survival according to interim Deauville score on PET-CT scan (**a**, **b**) and the baseline International Prognostic Index (**c**, **d**). PET-CT, positron emission tomography- computed tomography; PFS, progression-free survival; OS, overall survival; HR, hazard ratio; CI, confidence interval

Univariate analysis for PFS and OS was performed with clinical variables including individual IPI components, IPI score, sex, presence of B symptoms, bulky disease, and bone marrow involvement, in addition to interim Deauville score, and demonstrated that all variables were significantly associated with PFS and OS, with the exception of sex and bulky disease (Table S2). However, based on multivariate analysis, high-intermediate or high risk on the IPI score and interim Deauville scores of 4–5 were identified as independent prognostic factors for worse PFS and OS (Table 2).

Association between survival outcomes and interim Deauville score in combination with IPI

To determine the impact of interim Deauville score combined with IPI on survival outcomes, we stratified patients into four categories according to interim Deauville score and baseline IPI (Figure S1) and compared PFS and OS among these categories (Fig. 3a, b). We found that 66 of 171 patients with

interim Deauville scores of 1–3 and 28 of 49 patients with Deauville scores of 4–5 were high-intermediate or high risk based on the IPI score. Among patients with interim Deauville scores of 4–5, the 5-year PFS rate was significantly better in patients with low or low-intermediate IPI scores than in those with high-intermediate or high IPI scores (71.4% [95% CI, 52.0–90.8] *vs.* 14.3% [95% CI, 0–29.6], $P < 0.001$; Fig. 3a). The 5-year PFS rate in IPI-defined low or low-intermediate risk patients with interim Deauville scores of 4–5 was significantly worse than that in low or low-intermediate risk patients with Deauville scores of 1–3 (71.4% [95% CI, 52.0–90.8] *vs.* 84.5% [95% CI, 76.1–92.9], $P = 0.037$), but was similar to that of high-intermediate- or high-risk patients with Deauville scores of 1–3 (71.4% [95% CI, 52.0–90.8] *vs.* 54.0% [95% CI, 40.9–67.1], $P = 0.238$; Fig. 3a).

On the basis of these findings, we created three risk groups according to interim Deauville score and baseline IPI scores. The low-risk group ($N = 105$, 47.7%) included IPI-defined low or low-intermediate risk patients with interim Deauville

Table 2 Multivariate analysis for progression-free survival and overall survival

Variables		HR (95% CI)	P
Progression-free survival			
IPI	Low or low-intermediate	1	< 0.001
	High-intermediate or high	4.84 (2.84–8.24)	
Interim Deauville score	1–3	1	< 0.001
	4–5	2.96 (1.83–4.78)	
Overall survival			
IPI	Low or low-intermediate	1	< 0.001
	High-intermediate or high	5.75 (3.14–10.51)	
Interim Deauville score	1–3	1	< 0.001
	4–5	2.98 (1.79–4.97)	

HR hazard ratio, *CI* confidence interval, *IPI* International Prognostic Index

scores of 1–3, and the intermediate-risk group (*N* = 87, 39.5%) included low/low-intermediate risk patients with Deauville scores of 4–5 or high-intermediate/high-risk patients with Deauville scores of 1–3. Finally, the high-risk group (*N* =
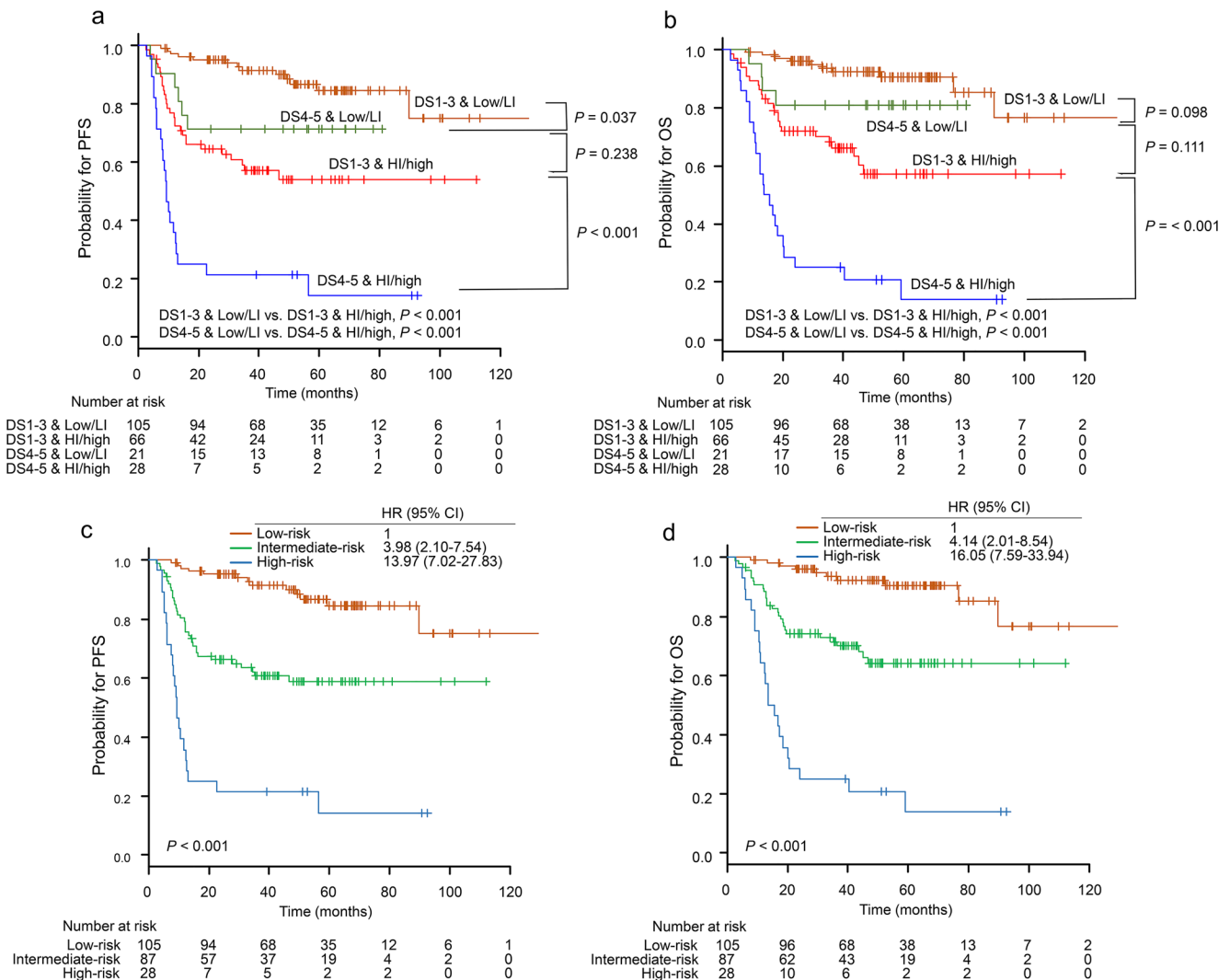
Fig. 3 Risk stratification and survival analysis. Progression-free survival (**a**) and overall survival (**b**) according to risk stratification based on interim Deauville score and baseline IPI scores. Progression-free survival (**c**) and overall survival (**d**) according to three-risk-group model. IPI, International Prognostic Index; PFS, progression-free survival; DS, Deauville score; OS, overall survival; HR, hazard ratio; CI, confidence interval; LI, low-intermediate; HI, high-intermediate

28, 12.7%) included high-intermediate- or high-risk patients with Deauville scores of 4–5. The respective 5-year PFS rates were 84.5% (95% CI, 76.1–92.9) for the low-risk group, 58.7% (95% CI, 47.7–69.7) for the intermediate-risk group, and 14.3% (95% CI, 0–29.6) for the high-risk group (P < 0.001, Fig. 3c). Consistent with these findings for PFS, there was also a significant association of the combined three-risk-group stratification with OS (P < 0.001, Fig. 3d).

Correlation between interim and end-of-treatment PET-CT results according to combined three-risk-group model

Among the 220 patients included in this analysis, 214 (97.3%) underwent end-of-treatment PET-CT scans (Fig. 1). Of the 166 patients with negative interim PET-CT findings, 160 patients (96.4%) retained their metabolic response on the end-of-treatment PET-CT scan, whereas 23 (47.9%) of 48 patients with positive interim PET-CT scans converted to negative on the end-of-treatment PET-CT scan (Table 3).

Among 27 high-risk patients according to our risk stratification, 10 patients (37.0%) converted to negative on the end-of-treatment PET-CT scan; in contrast, the number of patients converting from positive interim PET-CT to negative end-of-treatment scans was substantially higher in the intermediate-risk group (13 [61.9%]/21 patients; Table 3). Conversely, although most patients with negative interim PET-CT scans remained negative on the end-of-treatment scans, more interim negative patients among those in the intermediate-risk group (5 [8.2%]/61 patients) progressed to positive on the end-of-treatment PET-CT scans than did those in the low-risk group (1 [1.0%]/105 patients; Table 3).

At the time of this analysis, 41 of 171 patients (low-risk group, N = 13; intermediate-risk group, N = 28) had progressed despite negative interim PET-CT, and 29 of 49 patients (intermediate-risk group, N = 6; high-risk group, N = 23) with interim positive scans had failed. The positive predictive value (PPV) and negative predictive value (NPV) of interim PET-CT scan by progression were 59.1% and 76.0%, respectively, with a sensitivity of 41.4% and a specificity of 86.7% in the entire study cohort. Compared with these predictive values in the whole study cohort, the PPV of interim PET-CT by progression was much higher in the high-risk group (82.1%); in contrast, the NPV was highest in the low-risk group (87.6%). Among 70 progressions, 37 progressions were regarded as refractory disease in our cohort. A positive interim PET-CT was significantly associated with refractory disease (positive vs. negative; 22/49 [44.9%] vs. 15/171 [8.8%], P < 0.001; Table S3).

Discussion

The prognostic impact of interim PET-CT on treatment outcomes remains uncertain in patients with DLBCL, largely owing to the low PPV of this imaging tool [7–14]. In our study cohort, all of whom were treated with R-CHOP immuno-chemotherapy and underwent interim PET-CT scan after three cycles of R-CHOP, we observed a significant association of the Deauville score for interim PET-CT scan with survival outcomes. However, consistent with previous reports [7–14], a substantial proportion of patients with positive interim PET-CT results eventually achieved complete remission with long-term survival, indicating a considerable portion of false-positive interim PET-CT scans. Even though it is not a new finding, it is noteworthy that baseline IPI was significantly associated with outcomes of the present cohort, regardless of interim PET-CT results. IPI has been a widely accepted clinical prognostication tool for patients with DLBCL and is suggested to reflect biologic aggressiveness of the disease before treatment [16]. Based on these observations, we stratified our patients with DLBCL with interim Deauville scores of 1–3 or 4–5 into three risk groups according to baseline IPI scores and revealed a strong association between this risk stratification

Table 3 Correlation between interim and end-of-treatment PET-CT scans according to new risk stratification

Interim Deauville score	No. of patients	End-of-treatment Deauville score	
		Deauville scores of 1–3 (N, %)	Deauville scores of 4–5 (N, %)
Whole study cohort (N = 214)			
Deauville scores of 1–3	166	160 (96.4)	6 (3.6)
Deauville scores of 4–5	48	23 (47.9)	25 (52.1)
Low-risk group (N = 105)			
Deauville scores of 1–3	105	104 (99.0)	1 (1.0)
Intermediate-risk group (N = 82)			
Deauville scores of 1–3	61	56 (91.8)	5 (8.2)
Deauville scores of 4–5	21	13 (61.9)	8 (38.1)
High-risk group (N = 27)			
Deauville scores of 4–5	27	10 (37.0)	17 (63.0)

and PFS and OS. In addition, our data showed that predictive values of interim PET-CT scans were substantially improved in low- and high-risk groups based on our risk stratification, indicating the potential for reduction of false-positive and false-negative interim PET-CT scans through combination of baseline IPI scores. These data indicate that interim PET-CT response combined with baseline IPI scores is an important predictor of long-term treatment outcomes, and interim PET-CT-driven treatment strategies should therefore be investigated in the context of integration of baseline IPI scores.

Combined assessment using interim PET-CT and baseline IPI in the present study has important clinical implications. First, as observed in previous studies [20, 21], our data confirmed that negative interim PET-CT scan in patients with low or low-intermediate IPI scores was associated with excellent outcomes with low rates of treatment failure. Second, it is interesting that interim PET-CT-positive patients had significantly different risks of disease progression according to baseline IPI scores. In the present study, we observed remarkably high rates of disease progression or death among interim PET-CT-positive patients with high-intermediate or high IPI scores (5-year PFS rate, 14.3%), compared to the 5-year PFS rate of 71.4% among those with low or low-intermediate IPI scores. In addition, among interim PET-CT-positive patients with low or low-intermediate IPI scores, approximately two thirds of patients eventually converted to negative on end-of-treatment PET-CT scan. This finding indicated that a significant proportion of positive interim PET-CT scans were false positive, if they had low or low-intermediate IPI scores at baseline. In contrast, interim PET-CT-positive patients with high-intermediate or high IPI scores had extremely poor outcomes, indicating impending treatment failure in this group of patients. Thus, our data suggested that IPI was a strong indicator that identified patients with poor long-term prognosis among patients with positive interim PET-CT scans. Third, although treatment failure was relatively infrequent among patients with negative interim PET-CT results, patients with high-intermediate or high IPI scores experienced more disease progression than those with low or low-intermediate IPI score, including a rate of early disease progression as high as 8.2% on end-of-treatment PET-CT scan, consistent with the findings of a previous study [22]. These results indicated that some DLBCLs with aggressive biology could not be detected if small residual deposits were present under the detection limit of PET-CT [23], but recent data suggested that such a property might be partially expected with integration of baseline characteristics [11, 14, 22]. In fact, there has recently been great interest in predicting treatment outcomes of patients with DLBCL by combining interim PET results with pre-treatment characteristics to stratify the risk of disease progression or death. In a 147-patient cohort study by Mikhaeel et al. [11], they showed that combining baseline metabolic tumor volume with interim Deauville score improved the prediction

of interim PET-CT responses. In addition, de Oliveira Costa et al. [14] showed the prognostic role of cell-of-origin determination by immunohistochemistry in interim PET-CT-negative DLBCL patients. However, to date, no specific tool has been widely accepted yet for better discrimination of interim PET-CT-negative patients who would be likely to experience relapse or disease progression. Therefore, interim PET-CT-negative patients with high-intermediate or high IPI scores need to undergo careful follow-up until other tools become available to stratify this group of patients.

Another issue in the treatment of DLBCL is the role of treatment intensification based on interim PET-CT results. The benefit of treatment intensification in interim PET-CT-positive patients is assumed on the basis of promising initial observations [24, 25]. However, the results from recent prospective trials of interim PET-CT-guided treatment strategies have raised some questions [15, 26]. In a randomized phase 3 trial conducted to evaluate the impact of treatment intensification on survival outcomes using interim PET-CT following two cycles of R-CHOP immuno-chemotherapy, treatment intensification using the Burkitt protocol failed to improve outcomes in interim PET-positive patients [15]. In this study, however, approximately half of the patients with positive interim PET findings had a low or low-intermediate IPI scores and, as observed in our report, approximately 40% of these patients ultimately converted to negative on the end-of-treatment PET scan. Not surprisingly, these patients had excellent outcomes with 2-year PFS and OS rates of 84.2% and 100%, respectively. Given these data, the results of this trial did not indicate the ineffectiveness of interim PET-based approach, but merely highlighted the relevance of adequate selection of target patients who would most benefit from novel treatment intensification strategies. In the present study, the 5-year PFS rate of 14.3% among patients in the high-risk group based on our new risk stratifications suggests that room for further improvement still exists in the interim-PET-based treatment intensification strategies. Our new risk stratification may serve as a useful indicator that identifies patients at high risk for disease progression who may be candidates for studies evaluating up-front novel treatment intensification strategies. An additional important explanation for the negative results in the phase 3 trial is the possibility that conventional cytotoxic chemotherapy using the Burkitt protocol might not be sufficient to capture the improvement in outcomes in patients at risk of treatment failure. Thus, other novel approaches not based on conventional cytotoxic chemotherapy urgently require investigation.

Several limitations influencing the interpretation of our results should be considered in the present study. First, our study is a retrospective study, and the number of patients was not based on statistical considerations but on all feasible patients during the study period. Therefore, our data may contain unexpected bias and should be validated in prospective

multicenter studies. Second, interim PET-CT imaging was acquired from after one cycle to after four cycles of chemotherapy, but best timing of interim scan has not been determined so far. Because we uniformly performed interim PET-CT scan after three cycles of R-CHOP treatment, an evaluation of the impact of timing of interim scan on clinical outcomes is beyond the scope of this study. However, the best timing for interim PET-CT scan should be reevaluated together with the issues related to effectiveness of interim scan in future studies. Third, the use of pegfilgrastim may affect the interpretation of interim PET-CT. The short interval between the administration of pegfilgrastim and PET-CT scanning may cause false-positive uptake on interim PET-CT, particularly in cases with bone marrow, spleen, and Waldeyer's ring involvement [27]. Although our study showed high inter-observer agreement in Deauville score between two experienced reviewers, which was consistent in a previous report [28], the impact of G-CSF on the interpretation of interim PET-CT needs to be further clarified in the future study. Fourth, we considered a Deauville score of 4 or 5 to indicate a positive interim PET-CT, but data from other studies have suggested that a Deauville score of 5 was appropriate for cutoff on an interim scan [9]. Moreover, recent data have shown that quantitative analysis using the change in maximum standard uptake value might be superior to visual analysis using the Deauville criteria for interim PET-based outcome prediction [15, 26]. Thus, our new risk stratifications based on a Deauville cutoff score of 4 should be further investigated in prospective studies with larger patient populations.

In conclusion, our study showed that combined assessments using interim Deauville score and baseline IPI could improve prediction of the risk for disease progression or death in patients with previously untreated DLBCL. Our data will have an immediate clinical impact because we have identified a distinct subgroup with extremely poor outcomes (i.e., positive interim PET-CT and high-intermediate/high IPI scores), suggesting potential candidates for studies evaluating novel treatment intensification strategies. Alternatively, we have also identified a large subgroup of patients with excellent outcomes (i.e., negative interim PET-CT and low/low-intermediate IPI scores), in whom standard R-CHOP immuno-chemotherapy is sufficient for cure. These data suggest that interim PET-CT-based treatment intensification strategies should therefore be evaluated with incorporation of IPI scores at diagnosis.

Compliance with ethical standards

References

1. Kwak JY (2012) Treatment of diffuse large B cell lymphoma. Korean J Intern Med 27(4):369–377. https://doi.org/10.3904/kjim.2012.27.4.369

2. Gisselbrecht C, Glass B, Mounier N, Singh Gill D, Linch DC, Trneny M, Bosly A, Ketterer N, Shpilberg O, Hagberg H, Ma D, Briere J, Moskowitz CH, Schmitz N (2010) Salvage regimens with autologous transplantation for relapsed large B-cell lymphoma in the rituximab era. J Clin Oncol 28(27):4184–4190. https://doi.org/10.1200/JCO.2010.28.1618

3. Stumpe KD, Urbinelli M, Steinert HC, Glanzmann C, Buck A, von Schulthess GK (1998) Whole-body positron emission tomography using fluorodeoxyglucose for staging of lymphoma: effectiveness and comparison with computed tomography. Eur J Nucl Med 25(7): 721–728

4. Spaepen K, Stroobants S, Dupont P, Van Steenweghen S, Thomas J, Vandenberghe P, Vanuytsel L, Bormans G, Balzarini J, De Wolf-Peeters C, Mortelmans L, Verhoef G (2001) Prognostic value of positron emission tomography (PET) with fluorine-18 fluorodeoxyglucose ([18F]FDG) after first-line chemotherapy in non-Hodgkin's lymphoma: is [18F]FDG-PET a valid alternative to conventional diagnostic methods? J Clin Oncol 19(2):414–419. https://doi.org/10.1200/JCO.2001.19.2.414

5. Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, Lister TA, Alliance AL, Lymphoma G, Eastern Cooperative Oncology G, European Mantle Cell Lymphoma C, Italian Lymphoma F, European Organisation for R, Treatment of Cancer/Dutch Hemato-Oncology G, Grupo Espanol de Medula O, German High-Grade Lymphoma Study G, German Hodgkin's Study G, Japanese Lymphorra Study G, Lymphoma Study A, Group NCT, Nordic Lymphoma Study G, Southwest Oncology G, United Kingdom National Cancer Research I (2014) Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. J Clin Oncol 32(27):3059–3068. https://doi.org/10.1200/JCO.2013.54.8800

6. Barrington SF, Mikhaeel NG, Kostakoglu L, Meignan M, Hutchings M, Mueller SP, Schwartz LH, Zucca E, Fisher RI, Trotman J, Hoekstra OS, Hicks RJ, O'Doherty MJ, Hustinx R, Biggi A, Cheson BD (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(27):3048–3058. https://doi.org/10.1200/JCO.2013.53.5229

7. Burggraaff CN, de Jong A, Hoekstra OS, Hoetjes NJ, Nievelstein RAJ, Jansma EP, Heymans MW, de Vet HCW, Zijlstra JM (2019) Predictive value of interim positron emission tomography in diffuse

large B-cell lymphoma: a systematic review and meta-analysis. Eur J Nucl Med Mol Imaging 46(1):65–79. https://doi.org/10.1007/s00259-018-4103-3

8. Fan Y, Zhang Y, Yang Z, Ying Z, Zhou N, Liu C, Song Y, Zhu J, Wang X (2017) Evaluating early interim fluorine-18 fluorodeoxyglucose positron emission tomography/computed tomography with the SUVmax-liver-based interpretation for predicting the outcome in diffuse large B-cell lymphoma. Leuk Lymphoma 58(9):1–9. https://doi.org/10.1080/10428194.2016.1277384

9. Kim J, Song YS, Lee JS, Lee WW, Kim SE (2018) Risk stratification of diffuse large B-cell lymphoma with interim PET-CT based on different cutoff Deauville scores. Leuk Lymphoma 59(2):340–347. https://doi.org/10.1080/10428194.2017.1339877

10. Kong Y, Qu L, Li Y, Liu D, Lv X, Han J (2016) Predictive significance of a new prognostic score for patients with diffuse large B-cell lymphoma in the interim-positron emission tomography findings. Medicine (Baltimore) 95(6):e2808. https://doi.org/10.1097/MD.0000000000002808

11. Mikhaeel NG, Smith D, Dunn JT, Phillips M, Moller H, Fields PA, Wrench D, Barrington SF (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(7):1209–1219. https://doi.org/10.1007/s00259-016-3315-7

12. Nols N, Mounier N, Bouazza S, Lhommel R, Costantini S, Vander Borght T, Vekemans MC, Sonet A, Bosly A, Michaux L, Andre M, Van Den Neste E (2014) Quantitative and qualitative analysis of metabolic response at interim positron emission tomography scan combined with International Prognostic Index is highly predictive of outcome in diffuse large B-cell lymphoma. Leuk Lymphoma 55(4):773–780. https://doi.org/10.3109/10428194.2013.831848

13. Carr R, Fanti S, Paez D, Cerci J, Gyorke T, Redondo F, Morris TP, Meneghetti C, Auewarakul C, Nair R, Gorospe C, Chung JK, Kuzu I, Celli M, Gujral S, Padua RA, Dondi M, Group ILS (2014) Prospective international cohort study demonstrates inability of interim PET to predict treatment failure in diffuse large B-cell lymphoma. J Nucl Med 55(12):1936–1944. https://doi.org/10.2967/jnumed.114.145326

14. de Oliveira CR, Hallack Neto A, Siqueira S, Lage LA, de Paula HM, Coutinho AM, Pereira J (2016) Interim fluorine-18 fluorodeoxyglucose PET-computed tomography and cell of origin by immunohistochemistry predicts progression-free and overall survival in diffuse large B-cell lymphoma patients in the rituximab era. Nucl Med Commun 37(10):1095–1101. https://doi.org/10.1097/MNM.0000000000000553

15. Duhrsen U, Muller S, Hertenstein B, Thomssen H, Kotzerke J, Mesters R, Berdel WE, Franzius C, Kroschinsky F, Weckesser M, Kofahl-Krause D, Bengel FM, Durig J, Matschke J, Schmitz C, Poppel T, Ose C, Brinkmann M, La Rosee P, Freesmeyer M, Hertel A, Hoffkes HG, Behringer D, Prange-Krex G, Wilop S, Krohn T, Holzinger J, Griesshammer M, Giagounidis A, Raghavachar A, Maschmeyer G, Brink I, Bernhard H, Haberkorn U, Gaska T, Kurch L, van Assema DME, Klapper W, Hoelzer D, Geworski L, Jockel KH, Scherag A, Bockisch A, Rekowski J, Huttmann A, Investigators PT (2018) Positron emission tomography-guided therapy of aggressive non-Hodgkin lymphomas (PETAL): A Multicenter, Randomized Phase III Trial. J Clin Oncol 36(20):2024–2034. https://doi.org/10.1200/JCO.2017.76.8093

16. Ziepert M, Hasenclever D, Kuhnt E, Glass B, Schmitz N, Pfreundschuh M, Loeffler M (2010) Standard International prognostic index remains a valid predictor of outcome for patients with aggressive CD20+ B-cell lymphoma in the rituximab era. J Clin Oncol 28(14):2373–2380. https://doi.org/10.1200/JCO.2009.26.2493

17. Adams HJA, Kwee TC (2016) Proportion of false-positive lesions at interim and end-of-treatment FDG-PET in lymphoma as determined by histology: systematic review and meta-analysis. Eur J Radiol 85(11):1963–1970. https://doi.org/10.1016/j.ejrad.2016.08.011

18. Han YH, Lim ST, Jeong HJ, Sohn MH (2016) Clinical value of a one-stop-shop low-dose lung screening combined with (18)F-FDG PET/CT for the detection of metastatic lung nodules from colorectal cancer. Nucl Med Mol Imaging 50(2):144–149. https://doi.org/10.1007/s13139-015-0387-4

19. Crump M, Neelapu SS, Farooq U, Van Den Neste E, Kuruvilla J, Westin J, Link BK, Hay A, Cerhan JR, Zhu L, Boussetta S, Feng L, Maurer MJ, Navale L, Wiezorek J, Go WY, Gisselbrecht C (2017) Outcomes in refractory diffuse large B-cell lymphoma: results from the international SCHOLAR-1 study. Blood 130(16):1800–1808. https://doi.org/10.1182/blood-2017-03-769620

20. Mamot C, Klingbiel D, Hitz F, Renner C, Pabst T, Driessen C, Mey U, Pless M, Bargetzi M, Krasniqi F, Gigli F, Hany T, Samarin A, Biaggi C, Rusterholz C, Dirnhofer S, Zucca E, Martinelli G (2015) Final results of a prospective evaluation of the predictive value of interim positron emission tomography in patients with diffuse large B-cell lymphoma treated with R-CHOP-14 (SAKK 38/07). J Clin Oncol 33(23):2523–2529. https://doi.org/10.1200/JCO.2014.58.9846

21. Chow A, Phillips M, Siew T, Cull G, Augustson B, Ward M, Joske D (2013) Prognostic nomogram for diffuse large B-cell lymphoma incorporating the International Prognostic Index with interim-positron emission tomography findings. Intern Med J 43(8):932–939. https://doi.org/10.1111/imj.12194

22. Kwon SH, Kang DR, Kim J, Yoon JK, Lee SJ, Jeong SH, Lee HW, An YS (2016) Prognostic value of negative interim 2-[(1)(8)F]-fluoro-2-deoxy-d-glucose PET/CT in diffuse large B-cell lymphoma. Clin Radiol 71(3):280–286. https://doi.org/10.1016/j.crad.2015.11.019

23. Adams HJA, Kwee TC (2017) An evidence-based review on the value of interim FDG-PET in assessing response to therapy in lymphoma. Semin Oncol 44(6):404–419. https://doi.org/10.1053/j.seminoncol.2018.01.005

24. Mikhaeel NG, Hutchings M, Fields PA, O'Doherty MJ, Timothy AR (2005) FDG-PET after two to three cycles of chemotherapy predicts progression-free and overall survival in high-grade non-Hodgkin lymphoma. Ann Oncol 16(9):1514–1523. https://doi.org/10.1093/annonc/mdi272

25. Haioun C, Itti E, Rahmouni A, Brice P, Rain JD, Belhadj K, Gaulard P, Garderet L, Lepage E, Reyes F, Meignan M (2005) [18F]fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) in aggressive lymphoma: an early prognostic tool for predicting patient outcome. Blood 106(4):1376–1381. https://doi.org/10.1182/blood-2005-01-0272

26. Casasnovas RO, Ysebaert L, Thieblemont C, Bachy E, Feugier P, Delmer A, Tricot S, Gabarre J, Andre M, Fruchart C, Mounier N, Delarue R, Meignan M, Berriolo-Riedinger A, Bardet S, Emile JF, Jais JP, Haioun C, Tilly H, Morschhauser F (2017) FDG-PET-driven consolidation strategy in diffuse large B-cell lymphoma: final results of a randomized phase 2 study. Blood 130(11):1315–1326. https://doi.org/10.1182/blood-2017-02-766691

27. Jacene HA, Ishimori T, Engles JM, Leboulleux S, Stearns V, Wahl RL (2006) Effects of pegfilgrastim on normal biodistribution of 18F-FDG: preclinical and clinical studies. J Nucl Med 47(6):950–956

28. Burggraaff CN, Cornelisse AC, Hoekstra OS, Lugtenburg PJ, De Keizer B, Arens AIJ, Celik F, Huijbregts JE, De Vet HCW, Zijlstra JM, Group HIW (2018) Interobserver agreement of interim and end-of-treatment (18)F-FDG PET/CT in diffuse large B-cell lymphoma: impact on clinical practice and trials. J Nucl Med 59(12):1831–1836. https://doi.org/10.2967/jnumed.118.210807