Ann Nucl Med (2017) 31:660–668
DOI 10.1007/s12149-017-1196-x

ORIGINAL ARTICLE

Prediction of outcome in pediatric Hodgkin lymphoma based on interpretation of ^{18}FDG-PET/CT according to ΔSUV_{max}, Deauville 5-point scale and IHP criteria

Emine Göknur Isik[1] · Serkan Kuyumcu[1] · Rejin Kebudi[2] · Yasemin Sanli[1] · Zeynep Karakas[3] · Fatma Betul Cakir[3] · Seher Nilgün Unal[1]

Received: 31 August 2016 / Accepted: 9 July 2017 / Published online: 24 July 2017

Abstract
Objective Minimizing side effects by using response-adopted therapy strategies plays an important role in the management of pediatric Hodgkin lymphoma (HL); however, the criteria for the definition of adequate or inadequate response are controversial. The aim of this study is to compare different methods of interpretation of ^{18}F-FDG-PET/CT (PET) in the prediction of disease outcome in order to determine the optimum method in this regard.
Methods Baseline, interim and post-treatment PET scans of 72 children were interpreted according to revised International Harmonization Project criteria (IHP) and Deauville criteria. Cut-off values for changes in interim and post-treatment FDG uptake (ΔSUV_{max}) in the prediction of progression-free survival (PFS) were measured using ROC analysis. Quantitative and visual data were compared with each other in the prediction of PFS.
Results Mean interim and post-treatment ΔSUV_{max} of the primary lesions were 77.4 ± 19.5 and 68.8 ± 30.4% and respective cut-off values were 82 and 73%. However, only post-treatment ΔSUV_{max} yielded statistically significant results in the prediction of 3-year PFS ($p = 0.043$). Interim ΔSUV_{max} was further analyzed according to the values reported in the literature (66 and 77%) yet statistically significant results were not reached ($p = 0.604$ and 0.431). For interim evaluation, IHP criteria was correlated to Deauville criteria ($p = 0.002$ and $p = 0.001$) and ΔSUV_{max} ($p = 0.03$), whereas for post-treatment evaluation, significant correlation with ΔSUV_{max} ($p = 0.04$) but marginally significant ($p = 0.055$ and $p = 0.058$) correlation with Deauville criteria were achieved. Overall, 1, 3 and 5-year PFS were 95.7 ± 0.2, 89.6 ± 0.4 and 80.8 ± 0.7%, respectively. All methods demonstrated comparable performance in the prediction of 3-year PFS; however, interim PET using Deauville criteria and post-treatment PET using IHP criteria were statistically significant. All methods demonstrated high negative-predictive value but substantially low positive-predictive value.
Conclusions Deauville criteria are superior to other methods in the prediction of pediatric HL outcome using interim PET data. On the other hand, quantitative evaluation and visual evaluation by IHP can be used reliably at the end of the treatment. In this regard, we report the optimal cut-off value of SUV_{max} reduction as 73%.

Keywords Deauville · Pediatric Hodgkin lymphoma · IHP criteria · FDG-PET/CT

Introduction

Hodgkin lymphoma is one of the most frequent but curable hematological malignancies in children over 10 years [1]. However, an excellent outcome is compromised by long-term side effects [2]. Therefore, optimizing treatment to minimize subsequent risks while maintaining the chance of cure is of paramount importance. Supporting evidence has led to the integration of ^{18}F-FDG-PET/CT (PET) imaging into the routine staging and restaging algorithm of lymphomas [3, 4]. While PET has greater sensitivity for sites

✉ Serkan Kuyumcu
serkan.kuyumcu@istanbul.edu.tr

1 Department of Nuclear Medicine, Istanbul Faculty of Medicine, Millet cad. ÇAPA, 34093 Istanbul, Turkey

2 Division of Pediatric Oncology, Oncology Institute, Istanbul University, Istanbul, Turkey

3 Department of Pediatric Hematology and Oncology, Istanbul Medical Faculty, Istanbul, Turkey

of extra-nodal involvement and correspondingly has been found to improve baseline staging, an interim PET scan after one or two cycles of chemotherapy can implement response-adapted therapy in malignant lymphoma [5–7]. It can be used to identify patients with an inadequate response who could benefit from treatment intensification or reduce treatment burden [8, 9].

In 2007, response criteria were first published by the International Harmonization Project [10] and revised (IHP) by Cheson et al. [3]. However, IHP criteria, do not relate specifically to the interim situation after induction CT. In 2009, at the First International Workshop on interim PET in lymphoma, a 5-point scoring scale, the so-called Deauville criteria (DC), for visual evaluation of interim PET scans was proposed, which currently is considered to be the international standard [11, 12]. The Deauville criteria rely on comparison of residual uptake with the uptake in uninvolved organs within the same image in order to reduce FDG uptake variations rather than quantitative evaluation.

Although PET [13, 14] has demonstrated a similar diagnostic performance in children, most of the studies have been conducted in adolescent patients; yet, studies in the pediatric cohort are limited. Therefore, the criteria for the definition of adequate or inadequate response in children with HL are still being discussed [15]. This study reviews visual and quantitative baseline, interim and post-treatment PET scan data in children with HL and in this regard compares ΔSUV_{max}, Deauville criteria and IHP in the prediction of disease outcome.

Materials and methods

The institutional review board approved this retrospective study and the need for written informed consent was waived. PET scans of 105 patients with pediatric HL were reviewed, and 72 who had baseline, interim and post-treatment PET studies were included in the study. The diagnosis was established by lymph node biopsy in all patients and classified according to histopathological subtypes as determined by the WHO classification revised in 2008 [16]. Staging was performed according to the Ann Arbor staging system. Baseline PET was performed following histopathological diagnosis prior to treatment. All patients were administered chemotherapy as determined by the department of pediatric hematology and oncology in our institution. Interim PET was performed followed by 2–4 cycles of chemotherapy just before the next cycle, which was not earlier than 14 or later than 21 days. Response to treatment was evaluated with post-treatment PET, which was performed at least 15 days after the last cycle of chemotherapy or after 8 weeks if radiotherapy was performed.

^{18}F-FDG-PET/CT imaging protocol and analysis

PET imaging was performed on a dedicated PET/CT scanner (Siemens Biograph^{TM} TruePoint^{TM} PET/CT) as determined by the Pediatric Committee of the European Association of Nuclear Medicine [17]. All pathological lymph nodes and sites of extra-nodal involvement were evaluated. Maximum semiquantitative uptake values (SUV_{max}) of the five target lesions with highest uptake selected at baseline PET were followed up at interim and post-treatment PET. The reviewers outlined regions of interest for each lesion using semi-automated tools provided in Siemens Syngo image analysis software. Subsequently, percent reduction in interim and post-treatment PET SUV_{max} from baseline PET values were calculated and expressed as interim and post-treatment ΔSUV_{max}, respectively.

PET data were qualitatively interpreted using IHP [3] as the primary and Deauville 5-PS [12] as the secondary reading criteria and compared to ΔSUV_{max} for both interim and post-treatment PET data. The response definition criteria are summarized in Table 1. In addition, Deauville 5-PS was carried out in two subcategories as conservative and sensitive. In the conservative category, a score of 4 or 5 was positive, whereas in the sensitive category a score of 3 was also determined as positive.

Statistical analysis

IBM SPSS 21.0 software was used to perform the statistical analysis. A p value less than 0.05 was considered statistically significant. In the descriptive statistics of age and SUV_{max}, data were expressed as mean ± SD (standard deviation). Non-parametric data were given as median ± SD. SUV_{max}, Interim ΔSUV_{max}, post-treatment ΔSUV_{max} of all pathological lesions were determined, and the performance of ΔSUV_{max} in the prediction of PFS was assessed by receiver operating characteristic (ROC) curves generated by plotting sensitivity vs. 1-specificity. Sensitivity, specificity, positive-predictive value (PPV), negative-predictive value (NPV) and accuracy of interim and post-treatment PET scans in the prediction of PFS were calculated according to visual interpretation criteria proposed by Deauville criteria and IHP and as well as according to cut-off values obtained by ROC analysis. Kaplan–Meier test was used for the survival analysis and Spearman's correlation tests for correlation, while Mann–Whitney U, independent samples t test, Kruskal–Wallis

Table 1 IHP and Deauville 5-point scale

IHP criteria		Deauville criteria	
Response	Definition	Score	Definition
CR	Disappearance of all evidence of disease	1	No uptake above the background
PR	Regression of measurable disease and no new sites	2	Uptake $\leq$ mediastinum
SD	Failure to attain CR/PR or PD	3	Uptake $>$ mediastinum but $\leq$ liver
PD	Any new lesion or increase by 50% of previously involved sites from nadir	4	Uptake moderately increased compared to the liver at any site
		5	Uptake markedly increased compared to the liver at any site
		X	New areas of uptake unlikely to be related to lymphoma

CR complete remission, *PR* partial remission, *SD* stable disease, *PD* progressive disease

and Chi-square tests were used to analyze the correlation between methods.

Results

Patient characteristics are described in Table 2.

SUV_{max} and ΔSUV_{max}

Mean SUV_{max} of the primary lesions in baseline, interim and post-treatment PET studies were 11.3 ± 5.5 (2.46–26), 2.1 ± 1.7 (0.4–8.7) and 2.8 ± 2.2 (0.3–10.6), respectively. On the other hand, Interim ΔSUV_{max} and post-treatment ΔSUV_{max} were 77.4 ± 19.5% (29.8–97.1%) and 68.8 ± 30.4% (37.6–97.73%). There was no significant difference between Interim ΔSUV_{max} of early and advanced stage patients ($p = 0.073$). However, mean SUV_{max} of primary lesions on baseline PET was significantly higher in the advanced stage (8.3 ± 10.5 vs. 12.7 ± 4.7, $p = 0.003$).

Progression-free survival

Progression-free survival was defined as the time from the initial diagnosis until the first recurrence. Mean follow-up duration was 33.6 ± 17.8 months (4.4–68.7). Overall, 1, 3 and 5-year PFS were 95.7 ± 0.2, 89.6 ± 0.3 and 80.8 ± 0.6%, respectively. The follow-up time ranged from 4.4 to 68.7 months, and therefore some patients were followed for only a short period. Accordingly, a sufficient number of patients to calculate 5-year PFS was not reached for all methods when patients were allocated into groups according to the interpretation results. However, 3-year

Table 2 Baseline clinical characteristics

Characteristic	*n*	
Total number	72	
Gender		
Male	45	
Female	27	
Mean age (range)	12.9 (3–18)	
Histopathological subtype		
Nodular sclerosing	32	
Mixed cellular	26	
Lymphocyte rich	2	
Lymphocyte poor	2	
Unspecified	10	
Ann Arbor staging		
1A	4	6
1B	2	
2A	27	36
2B	9	
3A	5	15
3B	10	
4A	10	15
4B	5	
Early	30	
Advanced	42	
Mean follow-up duration (months)	33.6 (4.4–68.7)	
Treatment protocol		
ABVD + RT	41	
OEPA + COPP + RT	31	
Total	72	

ABVD doxorubicin, bleomycin, vinblastine, dacarbazine, *OEPA* vincristine, procarbazine, prednisone, doxorubicin, *COPP* cyclophosphamide, vincristine, procarbazine, prednisone, *RT* radiotherapy

PFS could be calculated for all methods. Performance of interpretation methods that predicted 3-year PFS with statistical significance are given in Table 4. Overall survival (OS) was 100% at the time of the study.

PFS analysis by Interim ΔSUV_{max}

Cut-off value for Interim ΔSUV_{max} was 82% (Fig. 1) with a sensitivity and specificity of Interim ΔSUV_{max} of 71.4 and 55.3%, respectively, in the prediction of PFS; however, it was not statistically significant (AUC = 0.587, 95% CI = 0.444–0.719, $p = 0.434$). Therefore, the significance of Interim ΔSUV_{max} was also analyzed based on the cut-off values (66 and 77%) reported in the literature [18–20]. Consequently, Interim ΔSUV_{max} predicted 1-year PFS as 81.8 ± 1.2 and 95 ± 0.3%, respectively, for cut-off values 66% ($p = 0.604$) and 77% ($p = 0.431$), yet no statistical significance was found (Fig. 2).

PFS analysis by post-treatment ΔSUV_{max}

Cut-off value for post-treatment ΔSUV_{max} was 73% (Fig. 1), with a sensitivity and specificity of 66.7 and 76.3% in the prediction of PFS (AUC = 0.706; 95% CI = 0.550–0.834, $p = 0.0437$), respectively. In patients with post-treatment ΔSUV_{max} below 73%, 1-year PFS was 93% and 4-year PFS was 55%.

Application of IHP criteria

In complete responders as determined by IHP mean SUV_{max} was 1.5 for interim PET and 1.8 for post-treatment PET. Accordingly, mean interim and post-treatment ΔSUV_{max} were 80 and 78%, respectively. In complete responders, 3- and 5-year PFS were 100 and 84.4% by interim PET ($p = 0.057$) (Fig. 2). Post-treatment PET predicted 1-, 2- and 3-year PFS as 95.8, 95 and 86.3%, respectively, in complete responders ($p = 0.014$) (Fig. 2). In partial responders, mean SUV_{max} was 2.6 for interim PET and 1.9 for post-treatment PET while respective mean interim and post-treatment ΔSUV_{max} were 71 and 72%. In partial responders, 1-year PFS was 77.8%. In non-responders as concluded by post-treatment PET, mean SUV_{max} was 6.1 and mean ΔSUV_{max} was 9% which were statistically significant values when compared to those of responders ($p = 0.007$). Consequently, 1-year PFS for non-responders was 60%.

Application of Deauville criteria

When Deauville criteria were applied using sensitive reading, mean SUV_{max} was 1.3 for interim PET and 2.1 for post-treatment PET in PET-negative patients. In conservative reading, mean SUV_{max} was 1.5 and 2.1 for interim and post-treatment PET, respectively. For sensitive reading in the PET-positive group, mean SUV_{max} were 3.2 and 4.1 for interim and post-treatment PET, respectively, in contrast to respective conservative results of 3.9 and 4.3. Conservative and sensitive evaluation results were similar with no statistically significant differences.

Statistically significant prediction of 3-year PFS with a sensitivity, specificity and NPV of 80, 63.4 and 97.1% was performed by interim PET (sensitive). However, respective measured values by post-treatment PET (both conservative and sensitive) were not statistically significant.

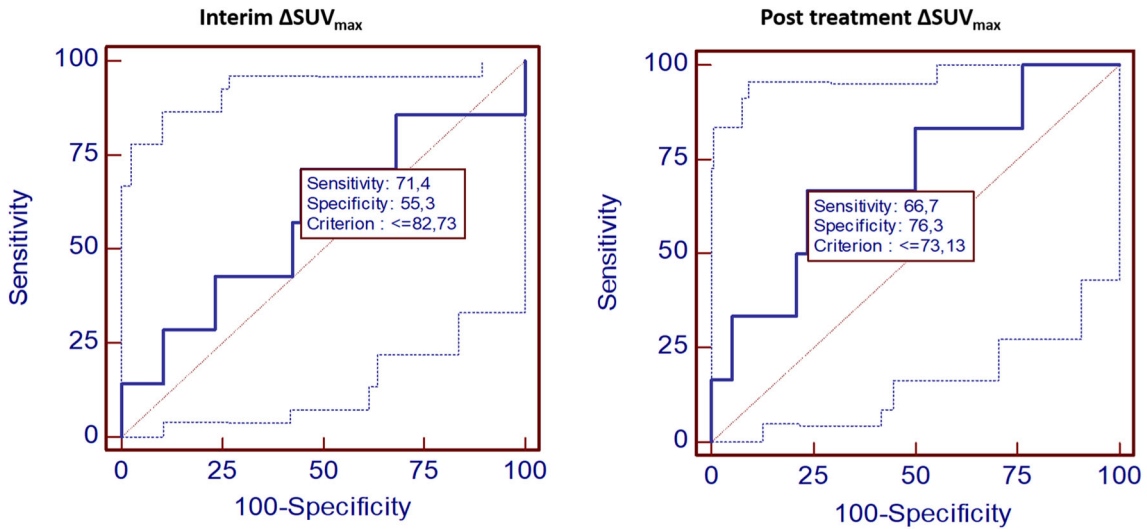
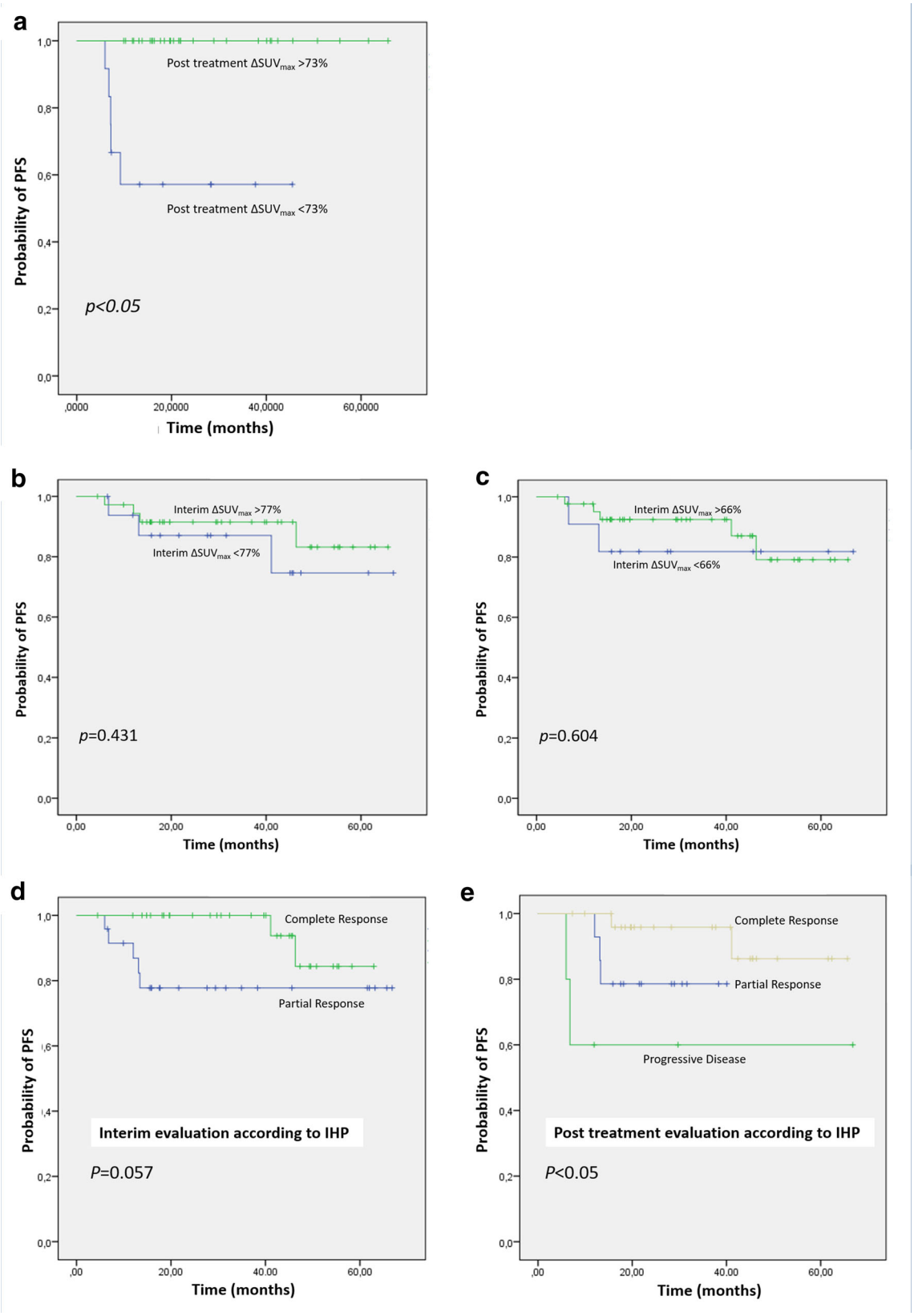


Fig. 1 The cut-off values for ΔSUV_{max} obtained by ROC analysis for interim and post-treatment which were 82.7% and 73.1, respectively


a
Post treatment ΔSUVmax >73%
Post treatment ΔSUVmax <73%
p<0.05
Probability of PFS
Time (months)
b
Interim ΔSUVmax >77%
Interim ΔSUVmax <77%
p=0.431
Probability of PFS
Time (months)
c
Interim ΔSUVmax >66%
Interim ΔSUVmax <66%
p=0.604
Probability of PFS
Time (months)
d
Complete Response
Partial Response
Interim evaluation according to IHP
P=0.057
Probability of PFS
Time (months)
e
Complete Response
Partial Response
Progressive Disease
Post treatment evaluation according to IHP
P<0.05
Probability of PFS
Time (months)

◀**Fig. 2** Kaplan–Meier curves of PFS: a statistically significant cut-off value for ΔSUV_{max} was only calculated for post-treatment PET (**a**). On the other hand, using interim PET this study failed to report a significant cut-off value and further analysis (**b**, **c**) using reported values in the literature (66 and 77%) also failed to predict PFS. Application of IHP criteria, yielded statistically significant results for post-treatment but not interim PET data (**d**, **e**)

Comparison of methods

Mean SUV_{max} of patients according to response criteria is demonstrated in Table 3. In the interim interpretation, PET-negative and positive groups by Deauville criteria were ($p = 0.002$ and $p = 0.001$ for sensitive and conservative readings, respectively) related to complete and partial responders by IHP. On the other hand, the relation was marginally statistically significant regarding post-treatment PET ($p = 0.055$ and $p = 0.058$ for sensitive and conservative readings, respectively). There was also a strong correlation ($r = -0.471$, $p = 0.000$) between Interim ΔSUV_{max} and Deauville criteria in Spearman's correlation test. Interim and post-treatment ΔSUV_{max} were also correlated to partial and complete responders by IHP which was also statistically significant ($p = 0.03$). In addition, post-treatment ΔSUV_{max} was also correlated with progressive disease ($p = 0.004$). Relation of methods with one another is demonstrated in Table 4. In the prediction of 3-year PFS, Deauville criteria had higher sensitivity while IHP criteria showed better specificity. On the other hand, all methods demonstrated high NPV; however, PPV was substantially low for all methods, reaching 40% at its best by post-treatment evaluation using IHP (Table 5).

Discussion

Hodgkin lymphoma accounts for 10% of pediatric cancers and is usually curable, with long-term survival rates exceeding 90% after treatment. Superiority of PET over conventional imaging has been shown by several studies for staging of pediatric HL. Response to treatment with interim PET is used as a criterion to identify non-responders, and thus the need for treatment intensification, particularly RT. Moreover, responders can be spared the long-term side effects of unnecessary treatment.

In this regard, interim PET provides prognostic information in HL and can predict early response to treatment; however, the number of studies in the pediatric group is limited. Given that the clinical prognostic factors used for adults are not used in children, interim and post-treatment evaluation come into question. In a study by Cerci et al. that investigated the predictive value of interim PET in 104 patients, 3-year PFS was 53 and 90.5% in PET-positive and negative patients, respectively [21]. Likewise, Gallamini et al. [22] found 2-year PFS 13 and 95%, respectively, in PET-positive and negative groups in 260 adult advanced HL patients. In addition, 2-year follow-up by PET was superior to the International Prognostic Score (IPS) in the prediction of prognosis and interim PPV and NPV were 86 and 95%, respectively. In a study by Zinzani et al. 24.5% of interim PET-positive and 92% of interim PET-negative patients were complete responders in a cohort of 304 HL patients [23]. In this study, complete response was also high in patients with negative interim PET; however, it was also high in patients with positive interim PET when compared to reported values in the literature. We assume this has accounts for the high PFS rate in patients with positive interim PET. This discrepancy may be due to the limited number of patients in our cohort as well as patient selection bias because of exclusion of patients without entire PET data.

In a review of seven studies investigating the correlations between early response and survival, interim PET predicted OS with 79% sensitivity, 92% specificity, 90% PPV, 81% NPV and 85% accuracy [24]. Hutchings et al. [25] investigated the prognostic value of interim PET in 85 patients and found it to predict PFS independently with an

Table 4 Correlation of IHP criteria to Deauville criteria and ΔSUV_{max} by Chi-square test

	Interim PET	Post-treatment PET
Deauville sensitive reading	$p = 0.002$	$p = 0.055$
Deauville conservative reading	$p = 0.001$	$p = 0.058$
ΔSUV_{max}	$p = 0.03$	$p = 0.004$

Table 3 Mean SUV_{max} of patients in respective categories

	IHP complete responders	PET-negative by Deauville		IHP partial responders	IHP non-responders	PET-positive by Deauville	
Interim PET	1.5	1.3 (sensitive)	1.6 (conservative)	2.6	n/a	3.2 (sensitive)	3.9 (conservative)
Post-treatment PET	1.8	2.1 (sensitive)	2.1 (conservative)	1.9	6.1	4.1 (sensitive)	4.3 (conservative)

Table 5 Performance of interpretation methods in the prediction of 3-year PFS

	Interim PET		Post-treatment PET			
	Deauville sensitive	Deauville conservative	Deauville sensitive	Deauville conservative	IHP criteria	ΔSUV_{end} 73%
Sensitivity	80.0	28.6	80.0	80.0	28.6	50.0
Specificity	63.5	78.0	65.9	70.7	95.4	71.4
PPV	17.4	15.4	22.2	25.0	40.0	21.5
NPV	97.1	88.6	96.4	96.7	92.5	90.0
Accuracy	51.4	56.9	43.1	45.8	88.9	41.7

NPV of 95% which is consistent with the result of our study. Similarly, the prognostic value of ΔSUV_{max} has also been documented in the literature [18]. However, in our study, the calculated cut-off value for Interim ΔSUV_{max} (83%) in the prediction of PFS was not statistically significant ($p = 0.434$). Calculation of PFS according to reported values in the literature [19] did not yield significant results either ($p = 0.604$ and 0.431, for Interim ΔSUV_{max} 66 and 77%, respectively). Consequently, we are unable to report a significant cut-off value for Interim ΔSUV_{max} in the prediction of PFS. In agreement with our findings; in a study by Furth et al. [26], which was the first prospectively conducted pediatric study, a cut-off value of 58% was found in a cohort of 40 children; however, it was reported that it could not reliably predict the response to treatment. Of note, NPV of Interim ΔSUV_{max} in the determination of 3-year PFS was 88.6% and PPV 16.6% which is in accordance with a number of studies reported in the literature. On the contrary, some studies have reported significant results in the prediction of PFS by Interim ΔSUV_{max}. In a study by Torizuka et al., a cut-off value of 60% could reliably predict the response to treatment [27]. They also found Interim ΔSUV_{max} to be significantly higher in patients with complete response than those without (81 and 35%, respectively). Itti et al. calculated 2-year PFS for a cut-off value of ΔSUV_{max} below and above 72.9%, as 32 and 79%, respectively [28]. In this regard, this study failed to identify a statistically significant ΔSUV_{max} value for interim PET to predict PFS. This particular contrast with previous studies discussed above could be attributed to a variety of causes, such as the retrospective design of the study or the relatively limited number of patients available. We also assume that unanticipated high PFS rate below the calculated Interim ΔSUV_{max} and the lack of predetermined PET timing due to the retrospective design of the study may be other potential factors. Nevertheless, further prospective studies, which consider these variables, should be undertaken.

On the other hand, contrary to Interim ΔSUV_{max}, post-treatment ΔSUV_{max} predicted PFS with strong statistical significance ($p = 0.000$) for a calculated value of 73%. No recurrence was observed in patients with post-treatment ΔSUV_{max} above 73%, while under 73% 1-year PFS was 93% and 4-year PFS 55%. To the best of our knowledge, the correlation between PFS and post-treatment ΔSUV_{max} in a pediatric cohort has not been investigated. However, in accordance with our results Barnes et al. [29] also found that interim PET failed to predict 4-year PFS ($p = 0.57$) and OS ($p = 0.09$), in contrast to post-treatment PET which successfully predicted PFS and OS ($p < 0.0001$) in adolescent patients.

IHP vs. Deauville criteria

Using IHP criteria on interim PET data could not statistically predict 3-year PFS ($p > 0.05$) in contrast to Deauville criteria interpretation. However, there was a statistically significant correlation when PET-positive and negative groups divided by Deauville criteria were, respectively, compared to complete and partial responders as determined by IHP. This comparison was made using both sensitive ($p = 0.002$) and conservative reading ($p = 0.001$) which both yielded high NPV (97.06 and 88.84%, respectively) in the prediction of 3-year PFS. Similarly, in a study by Furth et al. in which 39 pediatric HL patients were evaluated by visual analysis, NPV of Deauville criteria using sensitive reading was found in 96% [26].

In contrast to interim evaluation, using IHP criteria for post-treatment evaluation predicted 3-year PFS ($p = 0.014$) with high NPV and specificity. Prediction based on Deauville criteria was also statistically significant. Interestingly, when IHP and Deauville criteria were compared, the p value was at the limit of significance ($p = 0.057$ and $p = 0.058$ for conservative and sensitive reading, respectively). There was also a strong correlation ($r = -0.471$, $p = 0.000$) between Deauville criteria and ΔSUV_{max} and there was not a significant difference in the prediction of PFS between the two methods. Likewise, in a study by Pilkington et al., no statistically significant difference was found between Deauville criteria and

ΔSUV_{max} [30]. High PPV is needed to better guide management; however, PPV has been found to be substantially low by all interpretation methods in this study. PPV reached 40% at its best by post-treatment evaluation using IHP criteria. Low PPV may be a result of a typical false-positive paradox, that is, false-positive results are more probable than true positive when the overall population has a low incidence of a condition. The limited number of recurrences due to the excellent cure rate of pediatric HL is another factor which relates to a small number of true-positives, and thus low PPV in terms of statistics.

Another reason may be the lack of controlled PET timing due to the retrospective design of the study. Consequently, interim PET has not been performed following the same cycle of chemotherapy in all patients while post-treatment PET timing differed especially in patients undergoing RT. This variability in the timing of interim PET scans may have led to heterogeneity in the interim data. The Deauville criteria were adopted to minimize false-positive results and to allow for a relatively continuous reading scheme. However, higher PPV was not achieved with them either. Patient selection bias such as the exclusion of patients with entire PET data and relatively small sample size due to insufficient high numbers of patients in respective interpretation categories may be potential causes in addition to the ones discussed above. Despite all the limitations, a negative PET scan is a strong prognostic indicator of disease, whereas a positive scan should be validated with other imaging modalities or biopsy.

Conclusions

Our findings indicate the potential of improving the prognostic value of post-treatment ^{18}F-FDG-PET by using SUV-based analysis in pediatric Hodgkin lymphoma. The optimal cut-off value for SUV_{max} reduction from baseline to end of therapy is 73% for predicting PFS. Although quantitative evaluation is comparable to visual interpretation at the end of therapy with high NPV, our results failed to demonstrate a significant cut-off value for interim PET where Deauville criteria are superior to other methods. On the other hand, IHP criteria perform better than Deauville at the end of therapy. Potential implications for patient care will be to provide a more reproducible assessment of interim PET studies particularly to improve PPV and eventually, to guide risk-adapted therapies.

Compliance with ethical standards

Conflict of interest The authors have no conflicts of interest to declare.

References

1. Swerdlow AJ. Epidemiology of Hodgkin's disease and non-Hodgkin's lymphoma. Eur J Nucl Med Mol Imaging. 2003;30(Suppl 1):3–12.
2. Jorgov L, Montravers F, Balogova S, Ragu C, Pacquement H, Leblanc T, et al. Paediatric and adolescent Hodgkin lymphoma: information derived from diffuse organ uptake of ^{18}F-fluorodeoxyglucose on pre-treatment and on interim PET/CT. Eur J Nucl Med Mol Imaging. 2016;43(7):1220–30.
3. Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, et al. Revised response criteria for malignant lymphoma. J Clin Oncol. 2007;25:579–86.
4. Kostakoglu L, Cheson BD. Current role of FDG PET/CT in lymphoma. Eur J Nucl Med Mol Imaging. 2014;41(5):1004–27.
5. Gallamini A, Patti C, Viviani S, Rossi A, Fiore F, Di Raimondo F, et al. Early chemotherapy intensification with BEACOPP in advanced stage Hodgkin lymphoma patients with an interim-PET positive after two ABVD courses. Br J Haematol. 2011;152(5):551–60.
6. Körholz D, Kluge R, Wickmann L, Hirsch W, Lüders H, Lotz I, et al. Importance of F18-fluorodeoxy-D-2-glucose positron emission tomography (FDGPET) for staging und therapy control of Hodgkin's lymphoma in childhood and adolescence—consequences for the GPOH-HD 2003 protocol. Onkologie. 2003;26(5):489–93.
7. Kluge R, Körholz D. Role of FDG-PET in staging and therapy of children with Hodgkin lymphoma. Klin Padiatr. 2011;223(6):315–9.
8. Kostakoglu L, Gallamini A. Interim ^{18}F-FDG PET in Hodgkin lymphoma: would PET-adapted clinical trials lead to a paradigm shift? J Nucl Med. 2013;54(7):1082–93.
9. Gallamini A, Rigacci L, Merli F, Nassi L, Bosi A, Capodanno I, et al. The predictive value of positron emission tomography scanning performed after two courses of standard therapy on treatment outcome in advanced stage Hodgkin's disease. Haematologica. 2006;91(4):475–81.
10. Juweid ME, Stroobants S, Hoekstra OS, Mottaghy FM, Dietlein M, Guermazi A, et al. Use of positron emission tomography for response assessment of lymphoma: consensus of the Imaging Subcommittee of International Harmonization Project in Lymphoma. J Clin Oncol. 2007;25(5):571–8.
11. Meignan M, Gallamini A, Haioun C. Report on the first international workshop on interim-PET-scan in lymphoma. Leuk Lymphoma. 2009;50(8):1257–60.
12. Gallamini A, Fiore F, Sorasio R, Meignan M. Interim positron emission tomography scan in Hodgkin lymphoma: definitions, interpretation rules, and clinical validation. Leuk Lymphoma. 2009;50(11):1761–4.
13. Rhodes MM, Delbeke D, Whitlock JA, et al. Utility of FDG-PET/CT in follow-up of children treated for Hodgkin and non-Hodgkin lymphoma. J Pediatr Hematol Oncol. 2006;28(5):300–6.
14. Cheng G, Servaes S, Zhuang H. Value of (18)F-fluoro-2-deoxy-D-glucose positron emission tomography/computed tomography scan versus diagnostic contrast computed tomography in initial staging of pediatric patients with lymphoma. Leuk Lymphoma. 2013;54(4):737–42.
15. Hasenclever D, Kurch L, Mauz-Körholz C, et al. qPET—a quantitative extension of the Deauville scale to assess response in interim FDG-PET scans in lymphoma. Eur J Nucl Med Mol Imaging. 2014;41(7):1301–8.
16. Swerdlow SH, Campo E, Harris NL, et al. WHO classification of tumours of haematopoietic and lymphoid tissues. Lyon: IARC Press; 2008.

17. Stauss J, Franzius C, Pfluger T, et al. Guidelines for ^{18}F-FDG PET and PET-CT imaging in paediatric oncology. Eur J Nucl Med Mol Imaging. 2008;35(8):1581–8.
18. Casasnovas RO, Meignan M, Berriolo-Riedinger A, et al. SUV_{max} reduction improves early prognosis value of interim positron emission tomography scans in diffuse large B-cell lymphoma. Blood. 2011;118(1):37–43.
19. Lin C, Itti E, Haioun C, Petegnief Y, et al. Early 18F-FDG PET for prediction of prognosis in patients with diffuse large B-cell lymphoma: SUV-based assessment versus visual analysis. J Nucl Med. 2007;48(10):1626–32.
20. Itti E, Meignan M, Berriolo-Riedinger A, et al. An international confirmatory study of the prognostic value of early PET/CT in diffuse large B-cell lymphoma: comparison between Deauville criteria and ΔSUV_{max}. Eur J Nucl Med Mol Imaging. 2013;40(9):1312–20.
21. Cerci JJ, Pracchia LF, Linardi CC, et al. 18F-FDG PET after 2 cycles of ABVD predicts event-free survival in early and advanced Hodgkin lymphoma. J Nucl Med. 2010;51(9):1337–43.
22. Gallamini A, Hutchings M, Rigacci L, Specht L, Merli F, Hansen M, et al. Early interim 18F-FDG positron emission tomography is prognostically superior to international prognostic score in advanced stage Hodgkin's lymphoma: a report from a joint Italian-Danish study. J Clin Oncol. 2007;24:3746–52.
23. Zinzani PL, Rigacci L, Stefoni V, Broccoli A, Puccini B, Castagnoli A, et al. Early interim18 F-FDG PET in Hodgkin's lymphoma: evaluation on 304 patients. Eur J Nucl Med Mol Imaging. 2012;39(1):4–12.
24. Jerusalem G, Hustinx R, Beguin Y, Fillet G. Evaluation of therapy for lymphoma. Semin Nucl Med. 2005;35(3):186–96.
25. Hutchings M, Mikhaeel NG, Fields PA, Nunan T, Timothy AR. Prognostic value of interim FDG- PET after two or three cycles of chemotherapy in Hodgkin lymphoma. Ann Oncol. 2005;16(7):1160–8.
26. Furth C, Steffen IG, Amthauer H, Ruf J, Misch D, Schönberger S, et al. Early and late therapy response assessment with [^{18}F]fluorodeoxyglucose positron tomography in pediatric Hodgkin's lymphoma: analysis of a prospective multicenter trial. J Clin Oncol. 2009;27(26):4385–91.
27. Torizuka T, Nakamura F, Kanno T, Futatsubashi M, Yoshikawa E, Okada H, et al. Early therapy monitoring with FDG-PET in aggressive non-Hodgkin's lymphoma and Hodgkin's lymphoma. Eur J Nucl Med Mol Imaging. 2004;31(1):22–8.
28. Itti E, Lin C, Dupuis J, Paone G, Capacchione D, Rahmouni A, et al. Prognostic value of interim ^{18}F-FDG PET in patients with diffuse large B-cell lymphoma: SUV-based assessment at 4 cycles of chemotherapy. J Nucl Med. 2009;50(4):527–33.
29. Barnes JA, LaCasce AS, Zukotynski K, Israel D, Feng Y, Neuberg D, et al. End-of-treatment but not interim PET scan predicts outcome in nonbulky limited stage Hodgkin's lymphoma. Ann Oncol. 2011;22(4):910–5.
30. Pilkington Woll JP, García Vicente AM, Talavera Rubio MP, Palomar Muñoz AM, Jiménez Londoño G, León Martín A, et al. Quantitative and qualitative evaluation of the interim PET/CT in lymphoma treatment in the prediction of complete metabolic response. Rev Esp Med Nucl Imagen Mol. 2013;32(2):70–6.