
Comparison of Interim PET Response to Second-Line Versus First-Line Treatment in Classic Hodgkin Lymphoma: Contribution to the Development of Response Criteria for Relapsed or Progressive Disease

Regine Kluge¹, Tim Wittig², Thomas W. Georgi¹, Lars Kurch¹, Osama Sabri¹, W. Hamish Wallace³, Tomasz Klekawka⁴, Ana Fernández-Teijeiro⁵, Francesco Ceppi⁶, Jonas Karlén⁷, Jane Pears⁸, Michaela Cepelová⁹, Alexander Fosså¹⁰, Auke Beishuizen¹¹, Lisa Lyngsie Hjalgrim¹², Dieter Köhrholz¹³, Christine Mauz-Körholz^{*13}, and Dirk Hasenclever^{*14}

¹Department of Nuclear Medicine, University Hospital of Leipzig, Leipzig, Germany; ²Department of Internal Medicine, University Hospital of Leipzig, Leipzig, Germany; ³University of Edinburgh and Royal Hospital for Sick Children, Edinburgh, Scotland, United Kingdom; ⁴Department of Pediatric Oncology and Hematology, University Children's Hospital of Krakow, and Institute of Pediatrics, Jagiellonian University Medical College, Krakow, Poland; ⁵Department of Pediatric Oncology and Hematology, University Hospital Virgen Macarena, Sevilla, Spain; ⁶Pediatric Hematology–Oncology Unit, Division of Pediatrics, Woman–Mother–Child Department, University Hospital of Lausanne, Lausanne, Switzerland; ⁷Pediatric Cancer Unit, Astrid Lindgren Children's Hospital, Karolinska University Hospital, Stockholm, Sweden; ⁸Our Lady's Children's Hospital, Dublin, Ireland; ⁹Department of Pediatric Hematology and Oncology, Faculty Hospital Motol, Prague, Czech Republic; ¹⁰Department of Medical Oncology and Radiotherapy, Rikshospitalet–Radiumhospitalet, Oslo, Norway; ¹¹Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands; ¹²Department of Childhood and Adolescence Medicine, Department of Pediatric Hematology and Oncology, Juliane Marie Center, University Hospital Rigshospitalet, Copenhagen, Denmark; ¹³Department of Pediatric Oncology and Hematology, University Hospital Giessen and Marburg, Giessen, Germany; and ¹⁴Institute for Medical Informatics, Statistics, and Epidemiology, University of Leipzig, Leipzig, Germany

In first-line treatment of Hodgkin lymphoma (HL), Deauville scores 1–3 define complete metabolic remission. Interim ¹⁸F-FDG PET is also used for relapse-treatment adaptation; however, PET response criteria are not validated for relapse treatment. **Methods:** We performed a pairwise comparative analysis of early response to first- and second-line treatments in 127 patients with classic HL who experienced relapse. The patients participated in the prospective, multicenter EuroNet-PHL-C1 study. Residual uptake was measured retrospectively using the qPET method, a validated semiautomatic quantitative extension of the Deauville score. Empiric cumulative distribution functions of the qPET values were used to systematically analyze the response to first- and second-line treatments. **Results:** Individual patients responded variably to first- and second-line treatments. However, the empiric cumulative distribution functions of the qPET values from all patients were nearly superimposable. **Conclusion:** The findings support that first- and second-line treatments in HL do not require different response criteria.

Key Words: Hodgkin lymphoma; relapse; PET response criteria

J Nucl Med 2021; 62:338–341

DOI: 10.2967/jnumed.120.247924

A large body of published studies has consistently documented the value of ¹⁸F-FDG PET/CT performed at an interim time point (iPET) during first-line chemotherapy to predict the long-term outcome of patients with Hodgkin lymphoma (HL) (1,2). Consequently, iPET-adapted treatment strategies are frequently applied, either with escalation when there is an inadequate response or with deescalation in very good responders in both adult (3) and pediatric treatment optimization trials (4).

Most HL patients are cured with first-line treatment; treatment failure rates with most effective therapy strategies are in the range of 10%–15% (5,6). In the relapse setting, achieving a complete metabolic remission (CMR) on ¹⁸F-FDG PET before autologous stem cell transplantation has been shown highly predictive of outcome (7). Even in primary refractory HL, achieving a CMR based on PET may overcome the generally poor prognosis, with 10-y event-free survival of 68% in PET-negative versus 33% in PET-positive patients (8). PET is also used to guide treatment after relapse: patients with residual PET positivity after first-line salvage chemotherapy were switched to second-line salvage chemotherapy, and if they achieved CMR, they had a posttransplant outcome similar to that of patients achieving CMR with first-line salvage treatment (7). In the EuroNet-PHL-R1 trial, medium-risk children and adolescents achieving CMR after 2 cycles of second-line chemotherapy were deescalated to standard chemotherapy plus radiotherapy only, whereas inadequate responders received high-dose chemotherapy and autologous stem cell transplantation (9).

Response-adapted strategies may allow tailoring of treatment intensity to individual needs. Avoidance of acute and late side effects from unnecessary treatment is especially important in young patients (10,11). The proportion of patients receiving deescalated (or

Received Apr. 23, 2020; revision accepted Jun. 30, 2020.

For correspondence or reprints contact: Regine Kluge, University Hospital of Leipzig, Liebigstrasse 18, 04103 Leipzig, Germany.

E-mail: regine.kluge@medizin.uni-leipzig.de

*Contributed equally to this work.

Published online Aug. 6, 2020.

COPYRIGHT © 2021 by the Society of Nuclear Medicine and Molecular Imaging.

TABLE 1
Clinical Characteristics of Patients With and Without Relapse

Characteristic	Relapse		No relapse	
	n	%	n	%
Sex				
Male	67	52.8	374	51.6
Female	60	47.2	351	48.4
Age at first diagnosis (y)				
<13	35	27.6	203	28.0
13–17	92	72.4	522	72.0
Stage				
I	1	0.8	19	2.6
II	71	55.9	375	51.7
III	28	22.0	145	20.0
IV	27	21.3	186	25.7
B symptoms	51	40.2	282	38.9

escalated) treatment is strongly influenced by the criteria defining CMR. Thus, identification of the best cutoff representing the border between inadequate and adequate response is highly relevant. Several earlier studies observed an imbalance between relatively low positive predictive values and very high negative predictive values for iPET in first-line treatment of HL. This finding led to a stepwise shift in the CMR criteria during the last 10 y, with higher intensities of residual ^{18}F -FDG uptake being accepted (12–14). Today, the interpretation of PET response is generally based on the 5-point Deauville scale, with scores of 1–3 defined as CMR (15).

However, no special response criteria for second-line treatment have been established. To date, published studies have used the transition between Deauville score (DS) 2 and DS 3 as the cutoff between inadequate and adequate response, corresponding to the criteria in first-line treatment at that time (7,9,16). Would it be justifiable for second-line treatment to adopt the new response criteria used in first-line treatment? This change could result in a relevant reduction of treatment intensity for patients with DS 3.

To the best of our knowledge, our study was the first to directly compare iPET response behavior between first-line and second-line treatments of HL patients. The quantitative method allows the

distribution of residual ^{18}F -FDG uptake to be mathematically compared between the 2 time points. A similar distribution of residual uptake at both time points would support the use of the same response criteria, whereas a systematic shift would be an argument for developing separate criteria for relapse treatment.

MATERIALS AND METHODS

The EuroNet-PHL-C1 trial included patients who were newly diagnosed with classic HL between 2007 and 2013 and who were younger than 18 y at the time of diagnosis (EudraCT number 2006-000995-33). After 2 cycles of chemotherapy consisting of vincristine, etoposide, prednisone, and doxorubicin (17), a mandatory PET scan for early response assessment (iPET ERA) was performed in order to decide on radiotherapy at the end of treatment. For patients from 14 of 16 participating European countries, images were assessed by an interdisciplinary central review board in Germany. The PET/CT scans were reference-evaluated at the University of Leipzig after transfer of the images to a central server (Hermes Medical Solutions) (18). All patients or their guardians gave written informed consent to trial participation, scientific analyses of the data, and image transfer. The trial was approved by the respective ethics committees and regulatory authorities. Before patient accrual began, the trial was registered at clinicaltrials.gov (NCT00433459).

In cases of progressive disease or relapse, most patients received a second ERA with PET (iPET RERA) after 2 cycles of standard salvage chemotherapy with ifosfamide, etoposide, prednisone, doxorubicin, bleomycin, vinblastine, and dacarbazine (19). Both ERA and RERA iPET scans acquired on the same PET scanner were available on the central server from 131 patients developing progressive disease or relapse. In 4 patients, 1 of these 2 scans was not evaluable because of intensive uptake in brown adipose tissue or poor image quality. The remaining 127 patients form the analysis set for a pairwise comparison of treatment response in ERA versus RERA (relapse group).

As a control group (no-relapse group), we used 725 EuroNet-PHL-C1 patients without relapse within 72 mo, already included in a publication by Hasenclever et al. (20).

We used the qPET method for quantitative evaluation of the iPET images. The method was published in 2014 (20) and is now used as a routine method for standardized determination of the DS and treatment stratification in the EuroNet-PHL-C2 trial (EudraCT number 2012-004053-88). qPET is the quotient of the SUV_{peak} in the tumor residual with the highest remaining glucose uptake (calculated in the hottest connected voxels forming a volume of about 0.2 mL) divided by the SUV_{mean} of the liver (calculated in a squared 30-mL volume of interest in the right liver lobe). qPET values can be translated into DS: qPET values of 0.95 and 1.3 assign the border between DS 2/3 and DS 3/4, respectively (19).

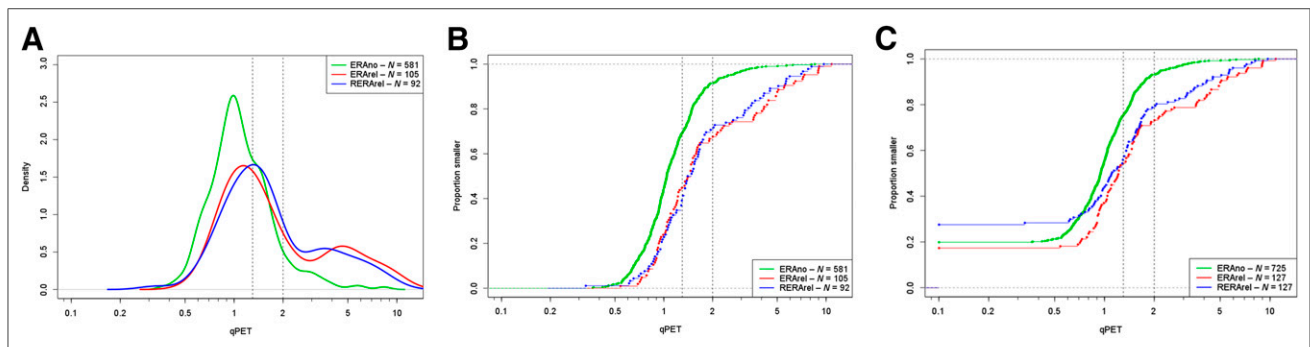


FIGURE 1. Distribution of qPET values in relapsing patients with HL at first-line treatment (ERAREl) and second-line treatment (RERAREl) and at first-line treatment in patient group without later relapse under identical first-line treatment (ERANo). (A) Density curves of all measurable qPET values. (B and C) Empiric cumulative distribution functions. Patients without qPET values due to completely normalized ^{18}F -FDG uptake are not included in B. Patients without qPET values due to completely normalized ^{18}F -FDG uptake are included in C.

TABLE 2

Quartiles of Measurable qPET Values in iPET Response Assessment of No Relapse vs. Relapse at ERA and RERA

Quartile	No relapse		Relapse	
	ERA	ERA	RERA	
Median	1.04	1.46	1.44	
25%	0.83	1.03	1.05	
75%	1.42	3.56	2.71	

According to the Lugano criteria (15), a CMR is defined by DS 1–3, corresponding to qPET values of less than 1.3. A qPET value of 1.3 or greater corresponds to DS 4 or 5 and represents inadequate response.

Density curves and empiric cumulative distribution functions were produced using R, version 3.6.3.

RESULTS

The main clinical characteristics of the patients in the relapse and no-relapse groups are given in Table 1. No significant differences occurred with regard to sex, age, stage of HL, or B symptoms between the 2 groups.

At ERA, a completely negative PET scan without any remaining ¹⁸F-FDG-accumulating tumor residuals (DS 1) was detected in 144 of 725 (20%) patients in the no-relapse group, in 22 of 127 (17%) patients in the relapse group at ERA, and in 35 of 127 (28%) patients in the relapse group at RERA. In these clearly PET-negative patients, no SUV_{peak} of a tumor residual could be measured and no qPET value was derived. Thus, measurable qPET values were available in the no-relapse group from 581 patients and in the relapse group from 105 patients at ERA and 92 patients at RERA.

The distribution of the measurable qPET values in the 3 groups is shown in Figures 1A and 1B and in Table 2. There was a clear shift to higher qPET values in the patients destined to relapse than in the patients without relapse ($P < 0.0001$, Kruskal test),

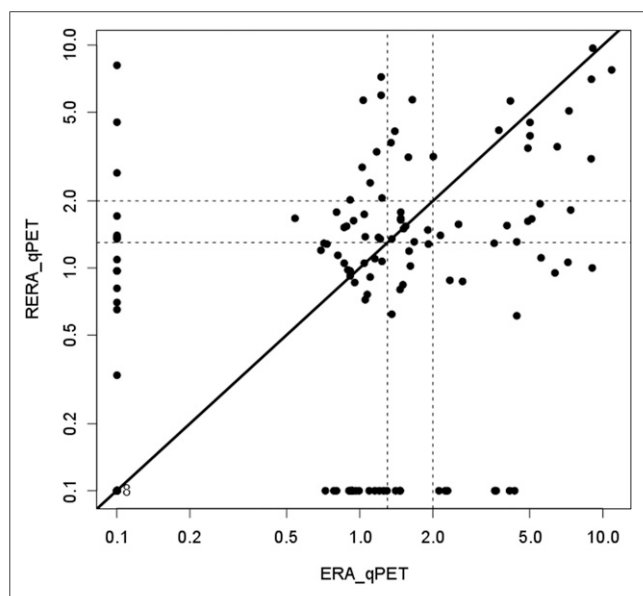


FIGURE 2. Pairwise comparison of qPET values at first- and second-line treatment of 127 relapsing patients. In patients with completely negative PET findings (DS 1), qPET value was set at 0.1.

predominantly with a reduced proportion of qPET values below 1.3 and an increased proportion of clearly abnormal qPET values above 2 (~DS 5). In contrast, there was no difference in the distribution of qPET values within the relapse group between ERA and RERA treatment ($P = 0.972$, paired Wilcoxon test).

Including the completely PET-negative results in the distribution function (Fig. 1C), we detected a slight shift toward lower residual uptake at second-line treatment than at first-line treatment in the relapsing patients. This shift was due to the higher proportion of completely PET-negative iPET exams at second-line treatment. This difference was not, however, significant ($P = 0.164$, Wilcoxon test).

On an individual basis, the response to first- and second-line therapy varied. Patients with CMR at first-line treatment (qPET < 1.3) had a nearly 2-fold chance of having CMR also at second-line treatment. However, 14 of the patients with CMR at first-line treatment showed high qPET values (>2.0) even at relapse treatment. In contrast, 11 patients with a poor response to first-line treatment (qPET > 2) achieved CMR with second-line treatment.

DISCUSSION

The Deauville 5-point scale is recommended in international guidelines and is widely accepted for response assessment of iPET in HL (3,15). A quantitative scale instead of only an ordinal one, however, allows additional analysis that is not feasible using an ordinal scale, such as comparing density and empiric cumulative distribution function curves and investigating choice of thresholds to positivity.

The qPET method provides a standardized and reproducible quantitative extension of the Deauville method (20). The qPET method is a semiautomatic, easy, and fast approach. Only 1 mouse click is required to determine the ¹⁸F-FDG uptake in an exactly standardized hottest volume of the most ¹⁸F-FDG-avid tumor residuum; a second mouse click is used to determine the mean ¹⁸F-FDG uptake in the liver. The qPET value is the quotient of both values. A qPET value of 1.3 was identified to represent the cutoff between visual DS 3 and 4. To the best of our knowledge, qPET has been the only method used for generating density and cumulative distribution curves of the residual uptake in a large group of HL patients whose first-line treatment was identical (14,20).

As expected, qPET values in first-line treatment tended to be higher in patients destined to relapse (Fig. 1) than in patients without later relapse (Table 2). This finding reflects the well-known prognostic value of iPET in HL.

Our data allow—for the first time, to our knowledge—a comparison between the first-line iPET response to induction chemotherapy and the second-line response to salvage therapy in HL. Individual patients showed variable response patterns (Fig. 2; Table 3): although 36% of the patients achieved CMR at both treatments, 25% did not respond well to either first- or second-line chemotherapy. CMR at interim assessment only in first-line treatment or only in second-line treatment was seen in 20% of the patients, each. The

TABLE 3

Patient-Based Comparison of PET Response to ERA and RERA Treatment

qPET values at RERA	qPET values at ERA	
	<1.3	≥1.3
<1.3	46	24
≥1.3	25	32

proportion of completely negative iPET scans (without residual uptake in any former tumor lesion, DS 1) was higher after 2 courses of relapse treatment than after 2 courses of first-line treatment (28% vs. 17%). This finding may be due to the difference in treatments or to the frequent follow-up, allowing earlier diagnosis of the relapsing disease.

Our main result, however, is that the distribution of measurable qPET values in patients destined to relapse at ERA are superimposable over the distribution of measurable qPET values at RERA for salvage treatment. This result supports that first- and second-line treatments do not require different response criteria.

We obtained the results with data from a study including patients under the age of 18 y at first diagnosis. However, we think that our results should apply at least to young adults. The age–incidence curves in HL are bimodal, with a first peak at 20–25 y and a second peak at 75–80 y. There is little evidence of a different etiology between adolescents and younger adults (21). We therefore expect that our results are transferable also to adult HL patients. A verification in older adults above 45–55 y old would be interesting.

Our results influenced the recently published risk- and response-adapted treatment guidelines of the EuroNet-PHL group for managing a first relapse of classic HL in children and young adults (9). These guidelines recommend defining CMR in a first relapse as DS 1–3 or a qPET value of less than 1.3—similarly to first-line treatment. In our data, using DS 1–3 instead of 1–2 for CMR applies to about one sixth of the progressing or relapsed cases. These patients will now receive less intensive treatment with a reduced risk of treatment-related side effects and late effects. However, clinical trials are needed to verify the utility of this cutoff for guiding treatment of relapse and whether subsets of patients will fare well with less intensive treatment.

CONCLUSION

The findings of this study support that first- and second-line treatments in HL do not require different response criteria.

DISCLOSURE

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

We thank our colleagues from the nuclear medicine and pediatric oncology departments of all participating centers of the EuroNet-PHL-C1 study for providing the PET imaging data of their patients. We also thank the Deutsche Krebshilfe and the Mitteldeutsche Kinderkrebsforschung.

KEY POINTS

QUESTION: Is the early PET response to second-line chemotherapy in HL systematically different from the response to first-line therapy?

PERTINENT FINDINGS: A pairwise comparative analysis of residual ^{18}F -FDG uptake after 2 cycles of first- and second-line chemotherapy showed no differences in a group of 127 patients with classic Hodgkin lymphoma.

IMPLICATIONS FOR PATIENT CARE: Our result indicates that it would be justifiable for second-line treatment to adopt the new response criteria used in first-line treatment. This change would result in less intensive relapse treatment in about one sixth of patients.

REFERENCES

- Gallamini A, Zwarthoed C. Interim FDG-PET imaging in lymphoma. *Semin Nucl Med.* 2018;48:17–27.
- Barrington SF, Kluge R. FDG PET for therapy monitoring in Hodgkin and non-Hodgkin lymphomas. *Eur J Nucl Med Mol Imaging.* 2017;44:97–110.
- Eichenauer DA, Aleman BMP, André M, et al. Hodgkin lymphoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2018;29(suppl 4):iv19–iv29.
- Kluge R, Kurch L, Georgi T, et al. Current role of FDG-PET in pediatric Hodgkin's lymphoma. *Semin Nucl Med.* 2017;47:242–257.
- Borchmann P, Goergen H, Kobe C, et al. PET-guided treatment in patients with advanced-stage Hodgkin's lymphoma (HD18): final results of an open-label, international, randomised phase 3 trial by the German Hodgkin study group. *Lancet.* 2018;390:2790–2802.
- Mauz-Körholz C, Metzger ML, Kelly KM, et al. Pediatric Hodgkin lymphoma. *J Clin Oncol.* 2015;33:2975–2985.
- Moskowitz CH, Matasar MJ, Zelenetz AD, et al. Normalization of preASCT, FDG-PET imaging with second-line, non-cross-resistant, chemotherapy programs improves event-free survival in patients with Hodgkin lymphoma. *Blood.* 2012;119:1665–1670.
- Shah GL, Yahalom J, Matasar MJ, et al. Risk factors predicting outcomes for primary refractory Hodgkin lymphoma patients treated with salvage chemotherapy and autologous stem cell transplantation. *Br J Haematol.* 2016;175:447.
- Daw S, Hasenclever D, Mascarin M, et al. Risk and response adapted treatment guidelines for managing first relapsed and refractory classical Hodgkin lymphoma in children and young people: recommendations from the EuroNet pediatric Hodgkin lymphoma group. *HemaSphere.* 2020;4:e329.
- Schaapveld M, Aleman BM, van Eggermond AM, et al. Second cancer risk up to 40 years after treatment for Hodgkin's lymphoma. *N Engl J Med.* 2015;373:2499–2511.
- van Nimwegen FA, Schaapveld M, Cutter DJ, et al. Radiation dose-response relationship for risk of coronary heart disease in survivors of Hodgkin lymphoma. *J Clin Oncol.* 2016;34:235–243.
- Isik EG, Kuyumcu S, Kebudi R, et al. Prediction of outcome in pediatric Hodgkin lymphoma based on interpretation of ^{18}F FDG-PET/CT according to $\Delta\text{SUV}_{\text{max}}$, Deauville 5-point scale and IHP criteria. *Ann Nucl Med.* 2017;31:660–668.
- Fallana F, Alongi P, Incerti E, et al. Diagnostic accuracy of FDG PET/CT for clinical evaluation at the end of treatment of HL and NHL: a comparison of the Deauville criteria (DC) and the International Harmonization Project Criteria (IHP). *Eur J Nucl Med Mol Imaging.* 2016;43:1837–1848.
- Kurch L, Hasenclever D, Kluge R, et al. Only strongly enhanced residual FDG uptake in early response PET (Deauville 5 or qPET ≥ 2) is prognostic in pediatric Hodgkin lymphoma: results of the GPOH-HD2002 trial. *Pediatr Blood Cancer.* 2019;66:e27539.
- Cheson BD, Fisher RI, Barrington SF, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol.* 2014;32:3059–3068.
- Moskowitz CH, Yahalom J, Zelenetz AD. High-dose chemo-radiotherapy for relapsed or refractory Hodgkin lymphoma and the significance of pretransplant functional imaging. *Br J Haematol.* 2010;148:890–897.
- Schellong G, Pötter R, Brämwig J, et al. High cure rates and reduced long-term toxicity in pediatric Hodgkin's disease: the German-Austrian multicenter trial DAL-HD-90. *J Clin Oncol.* 1999;17:3736–3744.
- Kurch L, Mauz-Körholz C, Bertling S, et al. The EuroNet pediatric Hodgkin network: modern imaging data management for real time central review in multicentre trials. *Klin Padiatr.* 2013;225:357–361.
- Schellong G, Dörrfel W, Claviez A, et al. Salvage therapy of progressive and recurrent Hodgkin's disease: results from a multicenter study of the pediatric DAL/GPOH-HD study group. *J Clin Oncol.* 2005;23:6181–6189.
- 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;41:1301–1308.
- Grotmol T, Bray F, Holte H, et al. Frailty modeling of the bimodal age-incidence of Hodgkin lymphoma in the Nordic countries. *Cancer Epidemiol Biomarkers Prev.* 2011;20:1350–1357.