Journal of Nuclear Medicine, published on December 18, 2020 as doi:10.2967/jnumed.120.247924

Comparison of interim PET response to second-line vs. to first-line treatment in classical Hodgkin lymphoma – contribution to the development of response criteria for relapsed or progressive disease

Regine Kluge[1], Tim Wittig[2], Thomas W Georgi[1], Lars Kurch[1], Osama Sabri[1], W Hamish Wallace[3], Tomasz Klekawka[4], Ana Fernández-Teijeiro[5], Francesco Ceppi[6], Jonas Karlén[7], Jane Pears[8], Michaela Cepelová[9], Alexander Fosså[10], Auke Beishuizen[11], Lisa Lyngsie Hjalgrim[12], Dieter Körholz[13], Christine Mauz-Körholz[13]*, Dirk Hasenclever[14]*

*Christine Mauz-Körholz and Dirk Hasenclever contributed equally

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

Corresponding author
Regine Kluge
Liebigstr. 18
04103 Leipzig (Germany)
Telephone: +49-341-9718000
Fax: +49-341-9718009
Mail: regine.kluge@medizin.uni-leipzig.de

Word count: 2299

Short running title: PET criteria in Hodgkin lymphoma relapse

ABSTRACT

Rationale: In first-line treatment of Hodgkin lymphoma (HL) Deauville scores 1-3 define complete metabolic remission. Interim 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 treatment response to first-line and to second-line treatment in 127 patients with classical Hodgkin lymphoma who experienced relapse. 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. Empirical cumulative distribution functions (ECDF) of the qPET values were used to systematically analyze the response to first-line and second-line treatment.

Results and conclusion: Individual patients responded variably to first-line and second line treatment. However, the ECDF of the qPET values from all patients were nearly superimposable. This supports that first-line and second-line treatment in HL do not require different response criteria.

Key words: Hodgkin lymphoma – relapse – PET response criteria

INTRODUCTION

A large body of published studies consistently documented the value of 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 with either escalation in inadequate response or de-escalation in very good responders in both adult (3) and pediatric treatment optimization trials (4).

The majority of 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 FDG-PET prior to autologous stem cell transplantation has been shown to be highly outcome predictive (7). Even in primary refractory HL achieving a CMR based on PET may overcome the generally poor prognosis with 10-year-event-free survival of 68% in PET-negative vs. 33% in PET-positive patients (8). PET is also used for relapse treatment guidance: Patients with residual PET-positivity after first-salvage chemotherapy were switched to second-line salvage chemotherapy and if they achieved CMR they had a similar post-transplant outcome to patients achieving CMR with first salvage treatment (9). In the EuroNet-PHL-R1 trial medium-risk children and adolescents achieving CMR after two cycles of second-line chemotherapy were de-escalated to standard chemotherapy plus radiotherapy only while inadequate responders received high-dose chemotherapy and autologous stem cell transplantation (10).

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 (11,12). The proportion of patients receiving de-escalated (or escalated) treatment is strongly influenced by the criteria defining CMR. Thus, identification of the best cut-off representing the border between inadequate or adequate response is highly relevant. Several earlier studies observed an imbalance between relatively low positive and very high negative predictive values of interim PET in first-line treatment of HL. This led to a stepwise shift of the CMR criteria during the last ten years accepting higher intensities of residual FDG uptake (13-15). Today, the interpretation of PET-response is generally based on the 5-point Deauville scale with scores of 1 to 3 defined as CMR (16).

However, no special response criteria for second-line treatment have been established. To date, published studies used the transition between Deauville score (DS)2 and DS3 as cut-off value between inadequate and adequate response, corresponding to the criteria in first-line treatment at that time (7,10,17). Would it be justifiable to adopt the new response criteria for first-line treatment to second-line treatment? This could result in relevant reduction of treatment intensity of patients with DS3.

To the best of our knowledge, our study for the first time directly compares the iPET response behaviour in first-line and second-line treatment of HL patients. The used quantitative method allows to compare the distribution of the residual FDG uptake at both time-points by mathematical methods. A similar distribution of residual uptake values at both time points would

support the use of the same response criteria while a systematic shift would be an argument to develop separate criteria for relapse treatment.

MATERIALS AND METHODS

The EuroNet-PHL-C1 trial included newly-diagnosed children and adolescents aged under 18 years at time of diagnosis with classical HL between 2007 and 2013 (EudraCT number: 2006–000995–33). After two cycles of OEPA chemotherapy consisting of vincristine, etoposide, prednisone, and doxorubicin (18), 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, Sweden)(19). 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 ethic committees and regulatory authorities. Before patient accrual began, the trial was registered at clinicaltrials.gov (NCT00433459).

In case of progressive disease or relapse, the majority of patients received a second early response assessment with PET (iPET-RERA) after two cycles of a standard salvage chemotherapy with IEP-ABVD (20). Both ERA- and RERA iPETscans acquired on the same PET scanner are available on the central server from 131 patients developing progressive disease or relapse. In

four patients, one of these two scans was not evaluable due to 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 vs RERA (group relapse).
As a control group (group NO relapse), we use N= 725 EuroNet-PHL-C1 patients without relapse within 72 months, already included in Hasenclever et al. 2014 (21).

).

We used the qPET method for quantitative evaluation of the iPETs. The method was published in 2014 (21) and is now used as 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 SUVpeak in the tumour residual with the highest remaining glucose uptake (calculated in the hottest connected voxels forming a volume of about 0.2 ml) divided by the SUVmean of the liver (calculated in a squared 30 ml volume of interest in the right liver lobe). qPET values can be translated in DS: qPET of 0.95 and 1.3 assign the border between DS 2/3 and DS 3/4, respectively (20). According to the Lugano criteria (16), a CMR is defined by DS 1 – 3, corresponding to qPET values < 1.3. qPET values ≥ 1.3 correspond to DS 4 or 5 and represent inadequate response.
Density curves and empirical cumulative distribution functions were produced using R 3.6.3 (22).

RESULTS

Main clinical characteristics of the patients in group “relapse” and group “NO relapse” are given in Table 1. No significant differences occurred with regard to sex, age, stage of HL and B symptoms between both groups.

At early response assessment, a completely negative PET scan without any remaining FDG-accumulating tumour residuals (DS1) was detected in 144/725 (20%) patients of the “group NO relapse”, in 22/127 (17%) patients of “group relapse” at ERA and in 35/127 (28%) of “group relapse” at RERA. In these clearly PET-negative patients, no SUVpeak of a tumour residual can be measured and no qPET value was derived. Thus, measurable qPET values were available in the “group NO relapse” from 581 patients and in the “group relapse” from 105 patients at ERA and 92 patients at RERA.

The distribution of the measurable qPET values in the three groups is shown in Fig. 1 A and B and Table 2. There is a clear shift to higher qPET values in the group of patients destined to relapse compared to the patient population without relapse ($p<0.0001$, Kruskal test) predominantly with a reduced proportion of qPET values below 1.3 and increased proportion of clearly abnormal qPET values above 2 (~DS 5). In contrast, there is no difference in the distribution of qPET values within the relapse group between first-line (ERA) and second-line (RERA) treatment ($p=0.972$, paired Wilcoxon test).

Including the completely PET-negative results in the distribution function (Figure 1C), a slight shift towards lower residual uptake at second-line treatment compared to first-line treatment in

the relapsing patients is detectable. This is due to the higher proportion of completely PET-negative iPET at second-line treatment. This difference is, however, not significant (p=0.164, Wilcoxon test).

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

DISCUSSION

The Deauville five-point scale is recommended in international guidelines and is widely accepted for response assessment of iPET in HL (3,16). A quantitative scale instead of an only ordinal one, however, allows additional analysis which is not feasible using an ordinal scale, e.g. comparing density and empirical 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 (21). The qPET-method is a semiautomatic, easy and fast approach. Only one mouse click is required to determine the FDG-uptake in an exactly standardized hottest volume of the most FDG-avid tumour residuum, a second mouse click is used to determine the mean 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 cut-off 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 identically first-line treated HL patients (15, 21).

As expected, qPET values in first-line treatment of patients destined to relapse (red curve in Figure 1) tend to be higher than qPET values in a cohort of patients without later relapse (green) (Table 2). This reflects the well-known prognostic value of interim PET in HL.

Our data – for the first time – allows comparing the first line interim PET response to induction chemotherapy to second-line response to salvage therapy in HL. Individual patients showed variable response patterns (Figure 2 and Table 3): while 36% of the patients achieved CMR at both treatments, 25% did not respond well to both first- and second-line chemotherapy. CMR at interim assessment only in first- or only in second-line treatment was seen in 20 % of the patients, each. The proportion of completely negative iPET scans (without residual uptake in any former tumour lesion, DS 1) is higher after two courses of relapse treatment than after two courses of first-line treatment (28% vs. 17%). This may be due to the different treatment or to frequent follow-up allowing earlier diagnosis of the relapsing disease.

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

We obtained the results with data from a study including patients under the age of 18 years at first diagnosis. However, we think that our results should at least apply to young adults. The age–incidence curves in Hodgkin lymphoma are bimodal with a first peak at 20-25 years and a second peak at 75-80 years. There is little evidence of a different etiology between adolescents and younger adults (23). We therefore expect that our results are transferable also to adult HL patients. A verification in older adults above 45-55 years of age would be interesting.

Our results influenced the recently published risk- and response-adapted treatment guidelines of the EuroNet-PHL-group for managing first relapsed classical Hodgkin lymphoma in children and young people (10). These guidelines recommend defining CMR in first relapse as DS 1-3 or qPET < 1.3 – similarly to first-line treatment. In our data, using DS 1-3 instead of DS 1-2 for CMR apply to about 1/6 of the progressing or relapsed cases. These patients will now receive less intensive treatment with reduced risk of treatment-related side and late effects. However, clinical trials are needed to verify the utility of this cut-off for guiding relapse treatment and whether subsets of patients will fare well with less intensive treatment.

DISCLOSURE STATEMENT

No potential conflicts of interest relevant to this article exist.

ACKNOWLEDGEMENTS

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

KEY POINTS

QUESTION: Is the early PET response to second-line chemotherapy in Hodgkin lymphoma systematically different from the response to first-line therapy?
PERTINENT FINDINGS: A pairwise comparative analysis of the residual FDG uptake after two cycles of first-line and of second-line chemotherapy showed no differences in a group of 127 patients with classical Hodgkin lymphoma.
IMPLICATIONS FOR PATIENT CARE: Our result encourages to adopt the revised response criteria for first-line treatment also to the second-line treatment. This would cause less intensive relapse treatment in about 1/6 of patients.

REFERENCES

1 Gallamini A, Zwarthoed C. Interim FDG-PET imaging in lymphoma. *Semin Nucl Med*. 2017;48:17–27.

2 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.

3 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.

4 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.

5 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.

6 Mauz-Koerholz C, Metzger ML, Kelly KM, et al. Pediatric Hodgkin lymphoma. *J Clin Oncol*. 2015;33:2975–2985.

7 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.

8 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;175:1–8.

9 Moskowitz CH, Matasar MJ, Zelenetz AD, et al. Normalization of pre-ASCT, 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.

10 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.

11 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.

12 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.

13 Isik EG, Kuyumcu S, Kebudi R, et al. Prediction of outcome in pediatric Hodgkin lymphoma based on interpretation of 18FDG-PET/CT according to ΔSUVmax, Deauville 5-point scale and IHP criteria. *Ann Nucl Med*. 2017;31:660–668.

14 Fallanca 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 (IHPC).
Eur J Nucl Med Mol Imaging. 2016;43:1837–1848.

15 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.

16 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.

17 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.

18 Schellong G, Pötter R, Brämswig 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.

19 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.

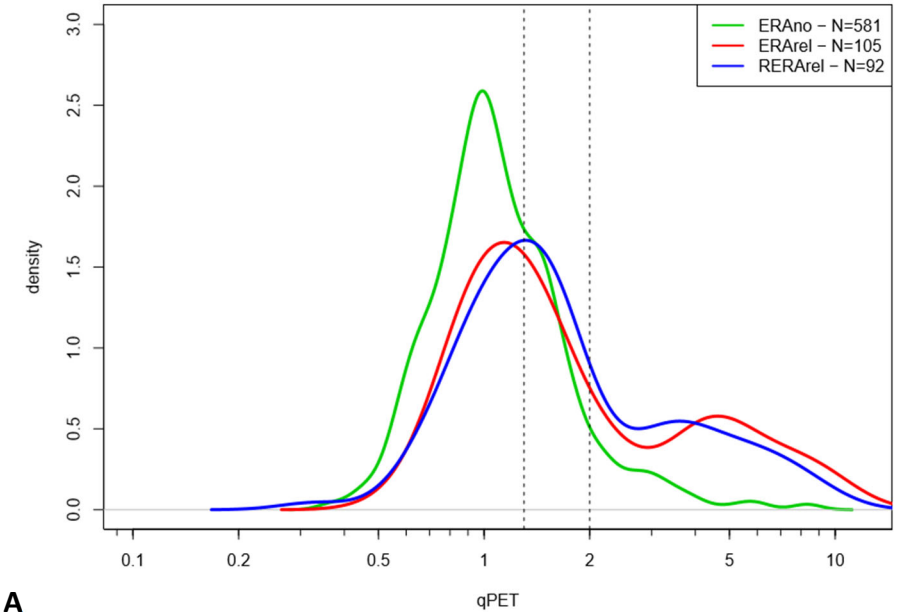
20 Schellong G, Dörffel W, Claviez A, et al. Salvage therapy of progressive and recurrent Hodgkin's disease: results from a multicenter study of the pediatric DAL/GPOH-HD study group. *J Clin Oncol*. 2005;23:6181–6189.

21 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.

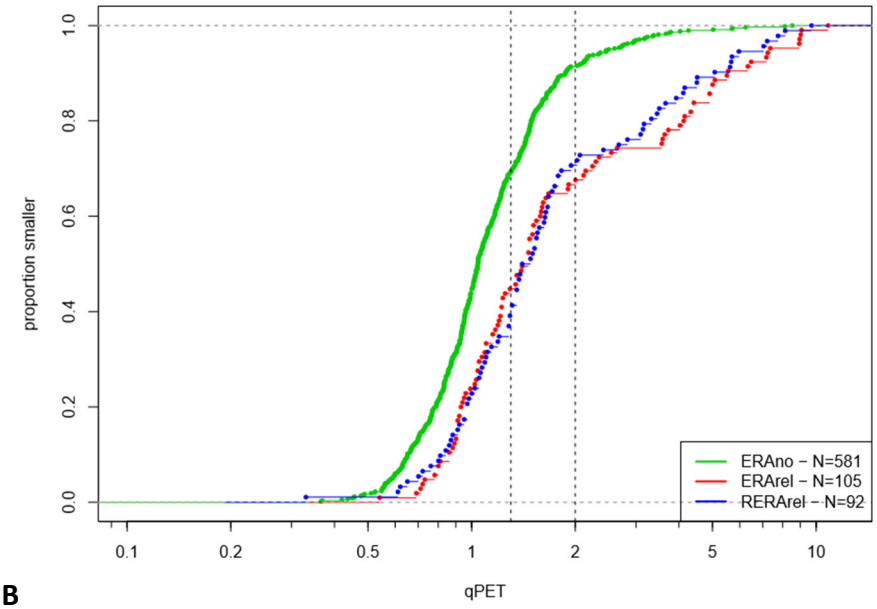
22 R core Team. R. A language and environment for statistical computing. 2020; https://www.R-project.org/.

23 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.

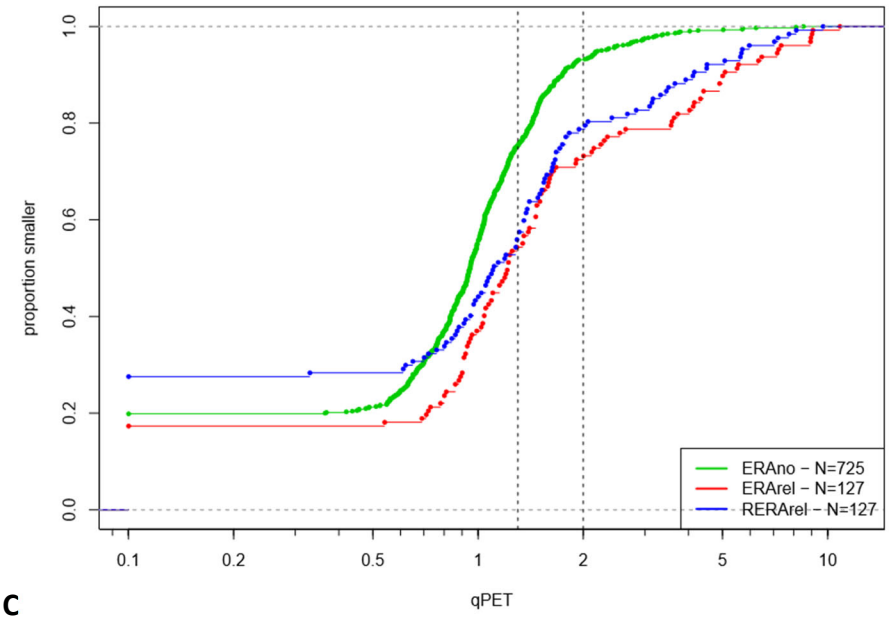
Figures and tables



A



B



C

Figure 1: Distribution of qPET values in relapsing patients with HL at first-line treatment (ERArel, red curve) and second-line treatment (RERArel, blue curve). The green curve shows distribution of qPET values at first-line treatment in a patient group without later relapse under identical first-line treatment (ERAno). A) Density curves of all measurable qPET values B) and C) Empirical cumulative distribution functions. B) Patients without qPET values due to completely normalized FDG uptake are not included. C) Patients without qPET values due to completely normalized FDG uptake are included.

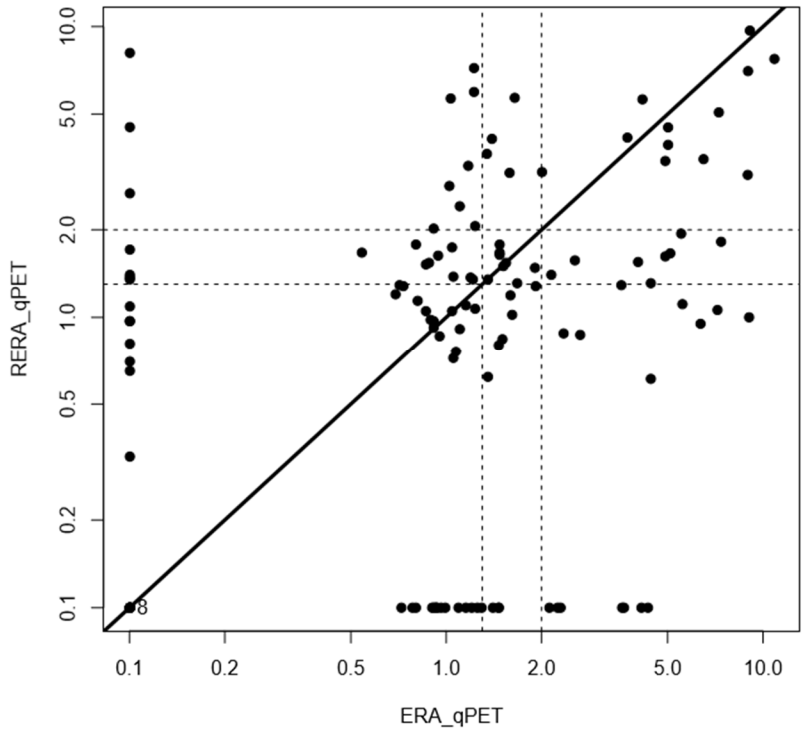


Figure 2: Pairwise comparison of qPET values at first-line and second-line treatment of 127 relapsing patients. In patients with completely negative PETs (DS 1) the qPET-value was set at 0.1.

Table 1: Clinical characteristics of patient groups with and without relapse.

	Group „relapse“		Group „NO relapse“	
	n	%	n	%
SEX male	67	52.8	374	51.6
female	60	47.2	351	48.4
AGE at first diagnosis				
< 13	35	27.6	203	28.0
13 -17	92	72.4	522	72.0
STAGE Stage I	1	0.8	19	2.6
Stage II	71	55.9	375	51.7
Stage III	28	22.0	145	20.0
Stage IV	27	21.3	186	25.7
B-symptoms	51	40.2	282	38.9

Table 2: Quartiles of the measurable qPET values in interim PET response assessment of the "group NO relapse" (ERAno) and of the "group relapse" at first-line treatment (ERArel) and at second-line treatment (RERArel).

	ERAno	ERArel	RERArel
median	1.04	1.46	1.44
Perc25	0.83	1.03	1.05
Perc75	1.42	3.56	2.71

Table 3: Patient-based comparison of PET response to first-line (ERA) and second-line (RERA) treatment.

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