Interim PET in Diffuse Large B-Cell Lymphoma

Lars Kurch[1], Andreas Hüttmann[2], Thomas W. Georgi[1], Jan Rekowski[3], Osama Sabri[1], Christine Schmitz[2], Regine Kluge[1], Ulrich Dührsen[*2], and Dirk Hasenclever[*4]

[1]Klinik und Poliklinik für Nuklearmedizin, Universitätsklinikum Leipzig, Leipzig, Germany; [2]Klinik für Hämatologie, Universitätsklinikum Essen, Essen, Germany; [3]Institut für Medizinische Informatik, Biometrie und Epidemiologie, Universität Duisburg-Essen, Duisburg, Germany; and [4]Institut für Medizinische Informatik, Statistik und Epidemiologie, Universität Leipzig, Leipzig, Germany

In diffuse large B-cell lymphoma, early assessment of treatment response by ^{18}F-FDG PET may trigger treatment modification. Reliable identification of good and poor responders is important. We compared 3 competing methods of interim PET evaluation. **Methods:** Images from 449 patients participating in the "PET-Guided Therapy of Aggressive Non-Hodgkin Lymphomas" trial were reanalyzed by applying the visual Deauville score and the SUV-based qPET (q = quantitative) and ΔSUV_{max} scales to interim PET scans performed after 2 cycles of chemotherapy. qPET relates residual lymphoma ^{18}F-FDG uptake to physiologic liver uptake, converting the ordinal Deauville scale into a continuous scale and permitting a direct comparison with the continuous ΔSUV_{max} scale, which is based on SUV_{max} changes between baseline and interim scans. Positive and negative predictive values were calculated for progression-free survival. **Results:** When established thresholds were used to distinguish between good and poor responders (visual Deauville score 1–3 vs. 4–5; $\Delta SUV_{max} > 66\%$ vs. $\leqq 66\%$), the positive predictive value was significantly lower with Deauville than ΔSUV_{max} (38.4% vs. 56.6%; $P = 0.03$). qPET and ΔSUV_{max} were strongly correlated on the log scale (Pearson $r = 0.75$). When plotted along corresponding percentiles, the positive predictive value curves for qPET and ΔSUV_{max} were superimposable, with low values up to the 85th percentile and a steep rise thereafter. The recommended threshold of 66% SUV_{max} reduction for the identification of poor responders was equivalent to qPET = 2.26, corresponding to score 5 on the visual Deauville scale. The negative predictive value curves were also superimposable but remained flat between 80% and 70%. **Conclusion:** Continuous scales are better suited for interim PET–based outcome prediction than the ordinal Deauville scale. qPET and ΔSUV_{max} essentially carry the same information. The proportion of poor-risk patients identified is less than 15%.

Key Words: diffuse large B-cell lymphoma; PET; interim evaluation; Deauville scale; $\Delta SUVmax$; qPET; PETAL trial

Diffuse large B-cell lymphoma (DLBCL) is the most frequent cancer of the immune system (1). It is cured in about two thirds of patients (2). Treatment response is among the most important factors determining outcome. Remissions are more durable in rapid than slow responders, as was demonstrated first by CT (3) and later by PET using the tracer ^{18}F-FDG (4,5). To adapt treatment to treatment response, reliable identification of good and poor responders is of the utmost importance.

Current guidelines recommend the Deauville scale for PET-based evaluation of early treatment response (6). This scale is based on a visual comparison of residual lymphoma-related uptake with areas of physiologically increased activity, such as in the mediastinal blood pool or liver (7). At the present time, any residual uptake exceeding that of the liver is considered a poor metabolic response.

A drawback of the Deauville scale is its ordinal nature, with no more than 5 response categories. An alternative way of evaluating interim scans is a quantitative comparison of the SUV_{max} before and during treatment. The ratio between the 2 values (ΔSUV_{max}) results in a continuous scale, which can be dichotomized to distinguish between good and poor responders (8). In DLBCL, thresholds of 66% SUV_{max} reduction after 2 cycles of rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) and 73% after 4 cycles of R-CHOP lead to maximum separation of the survival curves of good and poor responders (8–11).

The SUV_{max} method requires both baseline and interim PET scans. To restrict interim analysis to a single scan while maintaining the advantage of a continuous scale, we developed the qPET method (q = quantitative) where the SUV_{mean} of the 4 most intense connected voxels of residual lymphoma-related uptake are put into relation with the SUV_{mean} of a large volume in the liver. Pioneered in Hodgkin lymphoma, this approach converted the ordinal Deauville scale into a well-defined quantitative scale (12–14).

The goal of the present study was to apply the qPET approach to DLBCL and compare it with the visual Deauville scale and the ΔSUV_{max} method. To this end, we reanalyzed the data of the "PET-Guided Therapy of Aggressive Non-Hodgkin Lymphomas" (PETAL) trial, which set out—and failed—to improve treatment by adapting it to the response to the first 2 cycles of R-CHOP (15). Because none of the PET-driven treatment changes had an impact on outcome compared with standard R-CHOP, all treatment arms were combined for this analysis.

MATERIALS AND METHODS

Study Design

The PETAL trial (ClinicalTrials.gov NCT00554164; EudraCT 2006-001641-33) was a multicenter study for newly diagnosed aggressive non-Hodgkin lymphomas (15). The study was approved by the

TABLE 1

TABLE 1
Baseline Characteristics of Patients Included in Present Analysis in Comparison to Excluded Patients and All DLBCL Patients Participating in PETAL Trial

Characteristic	Patients included	Patients excluded	All patients
No. of patients	449	160	609
Median age (y)	62 (range, 18–80)	59.5 (range, 18–79)	62 (range, 18–80)
Age > 60 y	236 (52.6%)	78 (48.8%)	314 (51.6%)
Male sex	249 (55.5%)	93 (58.1%)	342 (56.2%)
ECOG performance status ≥ 2	48 (10.7%)	11 (6.9%)	59 (9.7%)
Ann Arbor stage III or IV	258 (57.5%)	100 (62.5)	358 (58.8%)
Extranodal sites > 1	148 (33.0%)	50 (31.2%)	198 (32.6%)
Lactate dehydrogenase > ULN	257 (57.4%)	78 (48.8%)	335 (55.1%)
International Prognostic Index			
Low risk	160 (35.7%)	64 (40.0%)	224 (36.8%)
Low-intermediate risk	111 (24.8%)	47 (29.4%)	158 (26.0%)
High-intermediate risk	102 (22.8%)	25 (15.6%)	127 (20.9%)
High risk	75 (16.7%)	24 (15.0%)	99 (16.3%)

ECOG = Eastern Cooperative Oncology Group; ULN = upper limit of normal.

Data are given as number of patients affected, followed by percentage of total number of patients with documented data, unless otherwise noted.

Federal Institute for Drugs and Medical Devices and the ethics committees of all participating sites. All patients gave written informed consent, including permission to use their data for post hoc scientific analyses.

Patients were treated with biweekly R-CHOP, with a 3-wk interval between cycles 2 and 3 to prevent false-positive results at interim staging uniformly performed after cycle 2. Patients with a favorable interim PET response received 4 more cycles of R-CHOP or the same treatment plus 2 extra doses of rituximab. Patients with an unfavorable response were randomly assigned to receive 6 additional cycles of R-CHOP or 6 blocks of a more intensive Burkitt lymphoma protocol (15).

PET/CT Imaging and Evaluation

Imaging conditions have been described previously (15). Baseline and interim scans (median chemotherapy-free interval, 20 d) were evaluated by local investigators using the ΔSUV_{max} method. Scans were then pseudonymized and transferred to a central server for reanalysis, including verification of the ΔSUV_{max} findings and evaluation according to the Deauville criteria by experienced nuclear medicine physicians (15).

In the present analysis, archived images were reanalyzed by a single physician (>10 y of working experience, >5,000 evaluated PET scans from lymphoma patients) using 3 different methods of interim scan evaluation. To ensure that visual Deauville scoring was not affected by quantitative measurements, assessment by purely visual criteria was performed before qPET and ΔSUV_{max}. The results were compared with the reports of the initially involved physicians (15), and any inconsistencies were resolved by further image evaluation taking account of clinical data available at the time of the first analysis. Discrepant results will be the subject of a future report.

The Deauville scale comprises 5 categories, which are defined as score 1, no residual uptake; 2, residual uptake not exceeding mediastinal uptake; 3, residual uptake above mediastinal but not exceeding liver uptake; 4, residual uptake above liver uptake; and 5, residual uptake markedly above liver uptake or new lesions (7). qPET was calculated by dividing the SUV_{mean} of the 4 hottest connected voxels (SUV_{peak}) of the hottest residual lesion by the SUV_{mean} of a 30-cm^3 volume of interest in the right lobe of the liver (12). ΔSUV_{max} was determined by dividing the SUV_{max} of the hottest residual lesion on the interim scan by the SUV_{max} of the hottest lesion on the baseline scan (8).

Statistical Analysis

The endpoint of the current analysis was progression-free survival, defined as the time from interim PET scanning to disease progression, relapse, or death from any cause, subsequently referred to as treatment failure. Progression and relapse were defined by clinical and imaging criteria and confirmed by biopsy in most cases. For simplicity, we treated progression-free survival as a binary variable (events within 60 mo). This decision appeared justified because, in DLBCL, most events occur within the first 2 y (16) and, with a median follow-up of 52 mo, the data were mature (15).

We plotted empiric cumulative distribution functions of qPET by visual Deauville scores and used receiver-operating-characteristic (ROC) analysis and the Youden index to derive plausible thresholds between individual scores of the visual scale. These cutoffs were compared with the thresholds found in a study with nearly 900 pediatric Hodgkin lymphoma patients (12).

With regard to ΔSUV_{max}, we used $1 - \Delta SUV_{max}$, that is, the remaining proportion of maximum ^{18}F-FDG uptake. This method allowed us to use the log scale (no negative values) and ensured the correlation with qPET to be positive. The area under the ROC curve was used to quantify the prognostic value of interim scanning. All analyses were performed using R, version 3.5.1 (R Core Team).

RESULTS

Patient Characteristics

Of 862 patients treated in the PETAL trial, 609 had DLBCL (15). Baseline and interim PET scans for post hoc analyses were available from 449 patients. In 65 cases, the scans were not
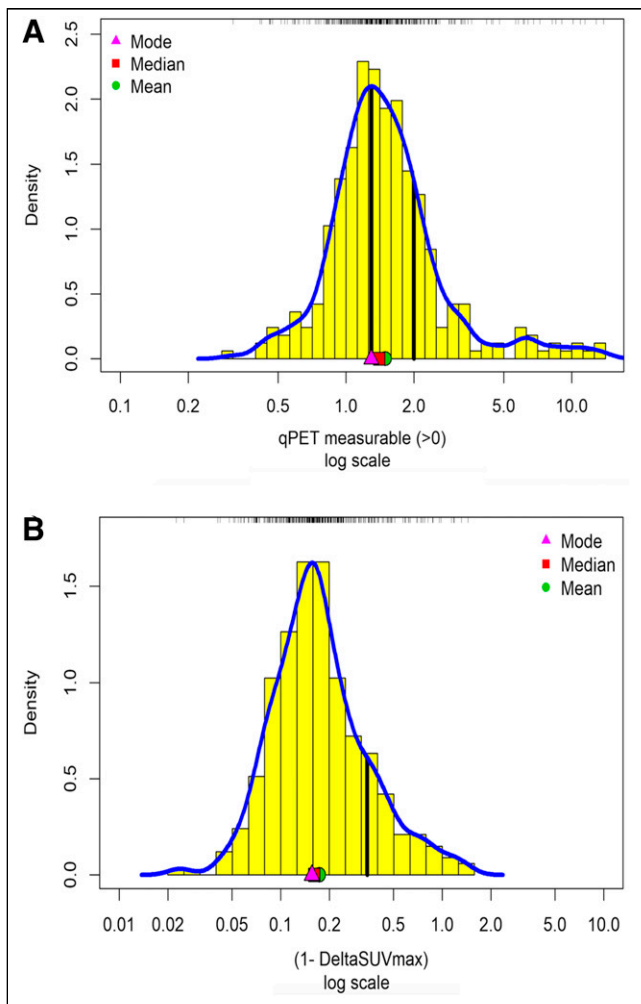
FIGURE 1. Density histograms for patients with measurable residual uptake at interim PET scanning ($n = 332$) evaluated by qPET (A) or ΔSUV_{max} (B) on log scale. First and second vertical lines in A indicate published thresholds between visual Deauville scores 3 and 4 (1.3) and 4 and 5 (2.0), respectively. ΔSUV_{max} in B is expressed as $1 - \Delta SUV_{max}$; vertical line indicates published threshold of 0.66, here $1 - 0.66 = 0.34$.

transferred to the central server; in 75, the transferred data were incomplete; and in 20, quantitative evaluation was not possible for technical reasons.

The baseline features and treatment results of the subgroup studied here were similar to the subgroup excluded from the analysis and the entire DLBCL population of the PETAL trial (Table 1, Supplemental Table 1, and Supplemental Fig. 1 [supplemental materials are available at http://jnm.snmjournals.org]; no statistically significant differences).

Response Assessment

Among the 449 patients included in this study, 117 (26.0%) were assigned to visual Deauville score 1, 42 (9.4%) to score 2, 113 (25.2%) to score 3, 120 (26.7%) to score 4, and 57 (12.7%) to score 5.

Patients with a visual score of 1 ($n = 117$) had a qPET value of zero, because the interim scan showed no measurable lymphoma-related activity. The qPET distribution of the 332 scans with measurable activity was symmetric on the log scale, with some outliers (representing poorly responding patients) on the right, and a mode (most frequent value) near 1.3 (Fig. 1A).
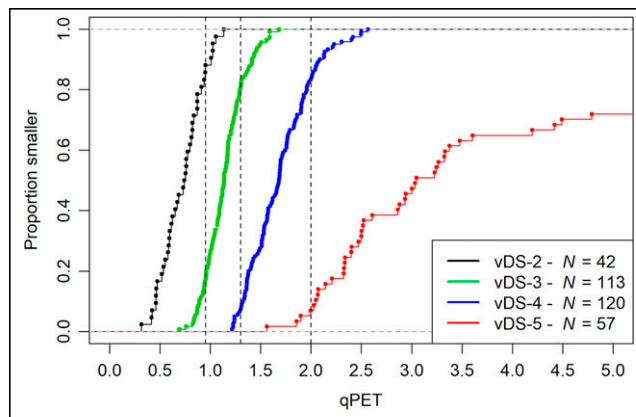

FIGURE 2. Empirical cumulative distribution functions of qPET measurements by visual Deauville categories. Vertical lines indicate published thresholds to map qPET values to individual categories. vDS = visual Deauville score.

qPET measurements within a single visual score category were well separated (Fig. 2). Some overlap was observed between neighboring categories, particularly between scores 3 and 4, which were located near the mode of the distribution. Optimal qPET thresholds between individual visual categories were 0.87 for the distinction between scores 2 and 3, 1.31 for scores 3 and 4, and 2.02 for scores 4 and 5, confirming the thresholds previously determined in pediatric Hodgkin lymphoma (0.95, 1.3, and 2.0, respectively) (*12*). These thresholds allow translation of qPET measurements into quantitative Deauville scores (1, qPET not measurable; 2, qPET measurable, but <0.95; 3, qPET 0.95 to <1.3; 4, qPET 1.3 to <2.0; 5, qPET ≥ 2). The concordance between visual and quantitative scores was 82.4% (Table 2).

As for ΔSUV_{max}, 100% SUV_{max} reduction was found in 117 patients. In the remainder, the distribution of measurable $1 - \Delta SUV_{max}$ resembled the qPET distribution, with symmetry on the log scale, outliers on the right, and a mode at 0.156 corresponding to 84.4% SUV_{max} reduction (Fig. 1B).

Outcome Prediction

Using the recommended thresholds for the visual Deauville scale (scores 1–3 vs. 4–5) and the quantitative ΔSUV_{max} scale (>66% vs. ≦66% SUV_{max} reduction) to distinguish between good and poor responders, the positive predictive values were 38.4% and 56.6% ($P = 0.03$) and the negative predictive values were 75.4% and 73.5%, respectively. Raising the threshold to score 5 (vs. 1–4) of the Deauville scale improved the positive predictive

TABLE 2
Comparison of Visual and Quantitative Deauville Scores

| Visual | Quantitative | | | | | |
	1	2	3	4	5	Sum
1	117	0	0	0	0	117
2	0	35	7	0	0	42
3	0	17	74	22	0	113
4	0	0	9	91	20	120
5	0	0	0	4	53	57
Sum	117	52	90	117	73	449

TABLE 3
Positive and Negative Predictive Values and Proportion of High-Risk Patients Identified by Interim PET: Comparison of
Methods and Thresholds

Definition of high-risk patients	PPV	NPV	Proportion of high-risk patients
Visual Deauville score 4 or 5	38.4%	75.4%	39.4%
Visual Deauville score 5	50.9%	73.0%	12.7%
Quantitative Deauville score 4 or 5	38.4%	76.1%	42.3%
Quantitative Deauville score 5	49.3%	73.7%	16.3%
qPET $\geq$ 2.26	54.7%	73.2%	11.8%
ΔSUV_{max}, $\leqq$66% SUV_{max} reduction	56.6%	73.5%	11.8%

PPV = positive predictive value; NPV = negative predictive value.

value, with a concomitant decrease in the proportion of high-risk patients (Table 3).

Treatment responses measured by qPET and ΔSUV_{max} were highly correlated on the log scale (Pearson $r = 0.75$; 95% CI, 0.70–0.80), suggesting that the methods provided similar information (Fig. 3). Neither of them reliably distinguished between patients in continued remission and patients who progressed, relapsed, or died, except at very high qPET and $1 - \Delta SUV_{max}$, where patients experiencing treatment failure were enriched. The similarities between qPET and ΔSUV_{max} were confirmed by ROC analysis yielding superimposable curves (Fig. 4). The area under the ROC curve was low for both methods (0.623 and 0.612, respectively), consistent with limited overall prognostic value.

To compare the positive and negative predictive values of qPET and ΔSUV_{max} at comparable thresholds, the values were plotted along their respective percentiles. Again, the curves were superimposable (Fig. 5). The positive predictive value was low up to the 85th percentile, followed by a steep increase (Fig. 5A). The

negative predictive value curves remained flat between 80% and 70% (Fig. 5B).

Corresponding percentiles were used to translate between the ordinal Deauville scale and the qPET and ΔSUV_{max} scales. Table 4 displays the clinically relevant thresholds on the visual 5-point scale and the 66% SUV_{max} reduction threshold on the ΔSUV_{max} scale, which identifies patients with a high risk of treatment failure.

Outcome According to Prognostic Group

Figure 6 shows the progression-free survival of patients categorized according to the Deauville scales (5 categories) or the ΔSUV_{max} scale, which was divided into 3 categories (SUV_{max} reduction by 100%, <100 to >66%, and $\leqq$66%). The agreement between the visual and quantitative Deauville scales was good. Progression-free survival did not differ between scores 2, 3, and 4, whereas score 5 was associated with significantly worse outcome (visual scale: hazard ratio [HR] of 2.56 and 95% CI of 1.68–3.90, $P < 0.0001$; quantitative scale: HR of 2.52 and 95% CI of 1.71–3.73, $P < 0.0001$), similar to
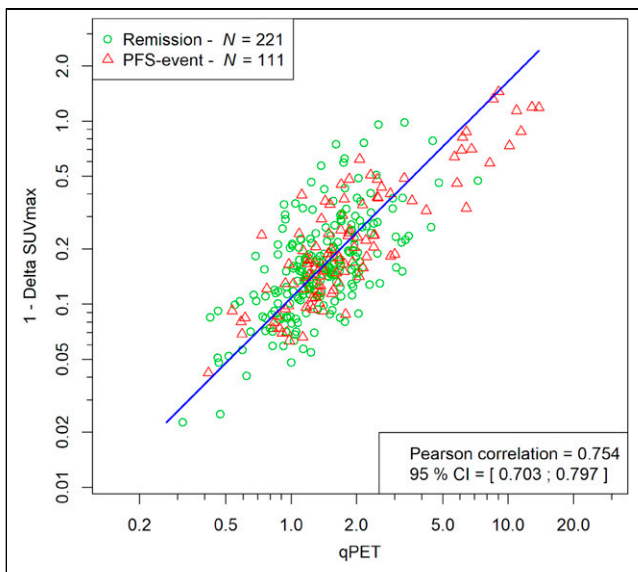

FIGURE 3. Scatterplot of qPET and ΔSUV_{max}. Triangles refer to patients experiencing treatment failure, whereas circles refer to patients who remained in remission. Blue line is the principal axis illustrating correlation. PFS = progression-free survival.
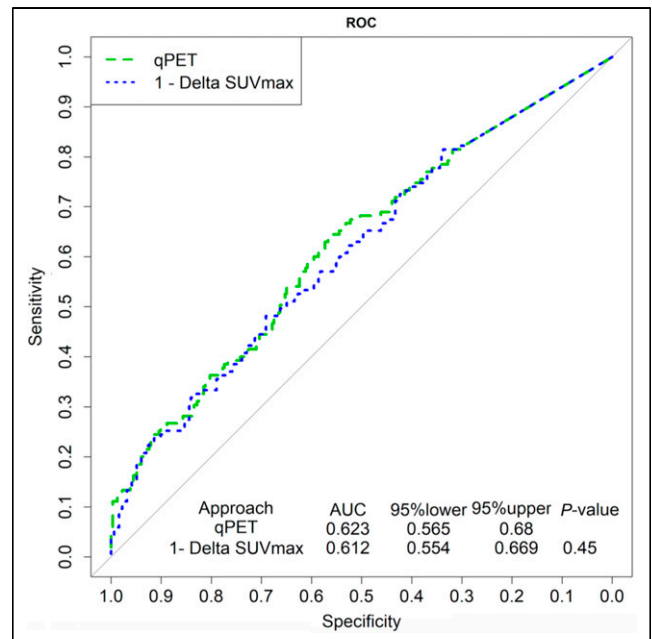

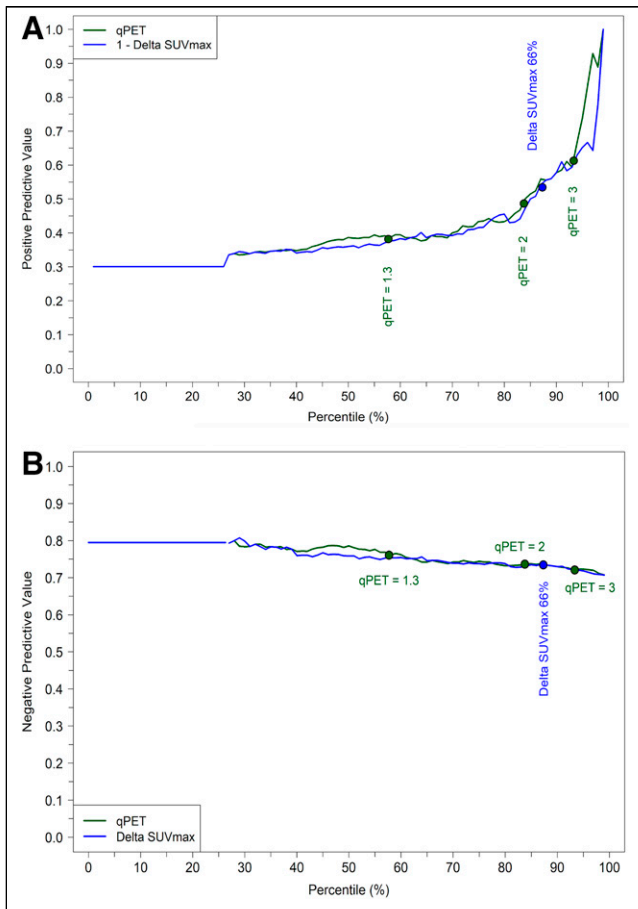**FIGURE 4.** ROC curves of qPET and $1 - \Delta SUV_{max}$ for progression-free survival.

FIGURE 5. Positive predictive value (A) and negative predictive value (B) of corresponding percentiles of qPET and ΔSUV_{max} measurements. Constant part of curves at low percentiles is due to inclusion of nonmeasurable values set at zero ($n = 117$).

an SUV_{max} reduction of 66% or less on the ΔSUV_{max} scale (HR of 3.27 and 95% CI of 2.16–4.96, $P < 0.0001$).

Patients with complete normalization of the interim scan (Deauville score 1, 100% SUV_{max} reduction) tended to have a better outcome than patients with a good response but residual activity. This observation was of borderline statistical significance (visual Deauville scale: HR of 1.63 and 95% CI of 1.03–2.60, $P = 0.036$; quantitative Deauville scale: HR of 1.57 and 95% CI of 0.98–2.51, $P = 0.054$; ΔSUV_{max} scale: HR of 1.57 and 95% CI of 0.99–2.51, $P = 0.051$).

TABLE 4
Selected Corresponding Thresholds in Categoric and Continuous Interim PET Response Scales

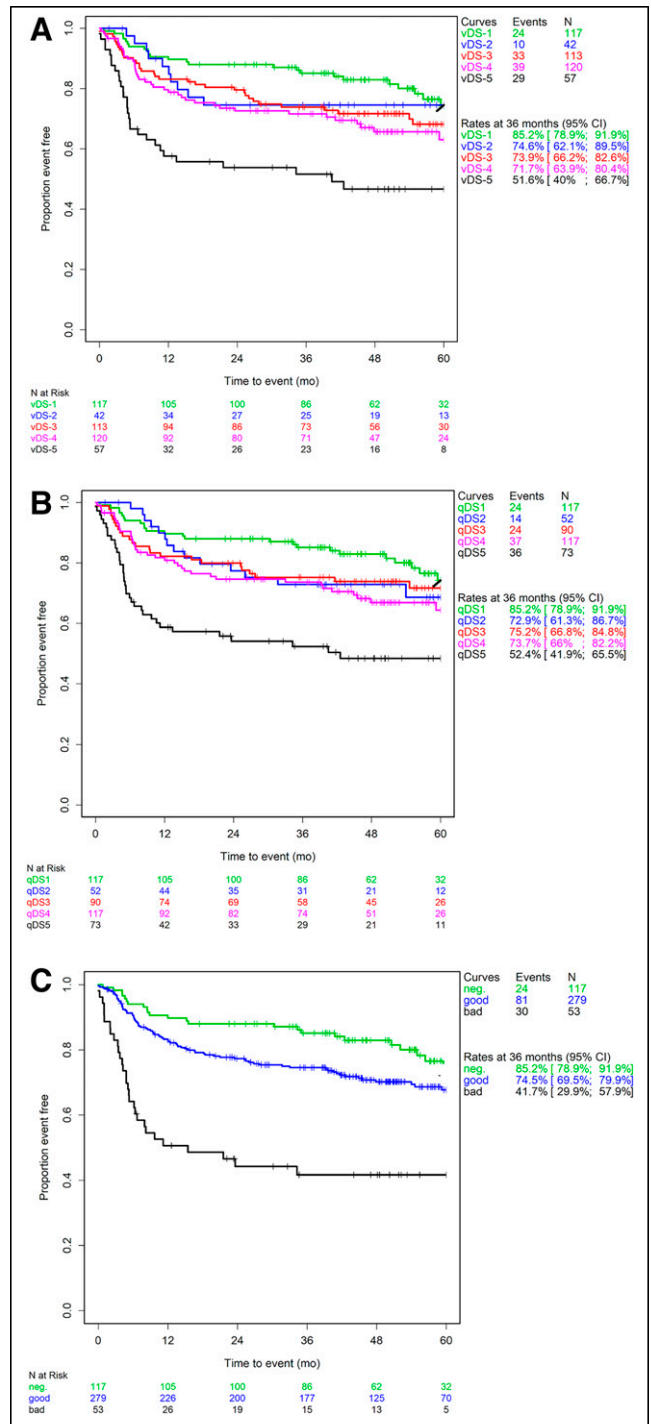| vDS	qPET	ΔSUV_{max}
2/3	0.95	91%
3/4	1.30	85%
4/5	2	73%
5	2.26	66%

vDS = visual Deauville score.

FIGURE 6. Progression-free survival in prognostic subgroups derived from visual Deauville scale (A), quantitative Deauville scale (B), or ΔSUV_{max} scale (C) (Kaplan–Meier analysis). vDS = visual Deauville score; qDS = quantitative Deauville score.

The 66% ΔSUV_{max} threshold identified a high-risk group comprising 13% of the total PETAL trial population (Fig. 5A) (15). The equivalent qPET value of 2.26 detected the same percentage but not the same individuals. Of 75 high-risk patients, 31 were tested positive by both methods, whereas 22 each were tested positive only by qPET or ΔSUV_{max}. Single-positive patients had a better outcome than double-positive patients ($P = 0.0011$). In our
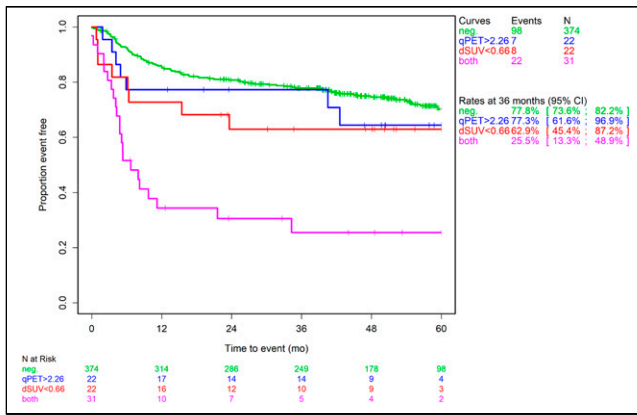
FIGURE 7. Progression-free survival in patients with good interim PET response according to both qPET and ΔSUV_{max}, only qPET, or only ΔSUV_{max}, or with poor interim PET response according to both methods (Kaplan–Meier analysis). dSUV = ΔSUV_{max}.

interpretation, qPET and ΔSUV_{max} are different methods measuring the same response concept, that is, replicates with independent measurement errors. Taking the average of the percentiles to reduce measurement errors, the mean percentiles were 0.47 for double-negative patients, 0.83 for single-positive patients, and 0.95 for double-positive patients ($P < 0.0001$). Thus, the observed outcomes (Fig. 7 corresponded to what was expected from the superimposable curves shown in Figure 5A.

DISCUSSION

The main result of our study is that, in DLBCL, qPET carries the same prognostic information as ΔSUV_{max}. The ROC and predictive value curves as a function of percentiles were superimposable. Thus, the methods can be used interchangeably. Combining them may help in individual cases, but only to reduce measurement errors by averaging.

The results obtained with both quantitative methods imply that the currently recommended threshold to identify high-risk patients by virtue of the visual Deauville scale (score 1–3 vs. 4–5) is of limited value. This conclusion complements our previous finding that ΔSUV_{max} is superior to Deauville for interim PET–based outcome prediction when recommended thresholds are used (*17*). A cut-off between scores 4 and 5 of the visual scale may be more appropriate, in particular when interim PET is used to select patients for more aggressive therapies. The current definition of Deauville score 5, however, is imprecise (*7*). Therefore, a better alternative is using a quantitative scale.

The visual Deauville scale is easy to use, but standardization remains difficult because of physiologic limitations of the human eye. The perception of light intensity depends on the surrounding background. In addition, any visual comparison is compounded by the distance between the areas of interest. Therefore, the reproducibility of visual assessments remains limited (*18–20*). The qPET method circumvents these problems because it relies on objective measurements rather than subjective impressions. Perhaps more importantly, it converts the Deauville categories into a continuous scale, allowing the definition of risk groups independent of the somewhat arbitrary thresholds of the visual scale. qPET is similar to rPET (r = ratio), which compares the SUV_{max}, that is, the single most intense voxels, in residual lymphoma and liver. In 2 small studies evaluating rPET after 2 cycles of R-CHOP, the best

threshold for prognostic dichotomization was determined to be 1.4 or 1.6 times the SUV_{max} of the liver (*21,22*). After 1 or 4 treatment cycles, the most appropriate thresholds were 3.1 or 1.4 times the liver SUV_{max}, respectively (*20,23*). Advanced image reconstructions, however, may overestimate SUV_{max} compared with SUV_{peak} and SUV_{mean} (*24*). We therefore chose the SUV_{mean} of a large volume within the liver as the reference standard, and we chose a very small SUV_{peak} volume, comprising a low number of connected voxels, to represent the residual lymphoma lesion (*12*). Because of the systemic nature of DLBCL, its often rapid response to therapy, and the small size of posttreatment remnants, larger SUV_{peak} volumes commonly used in solid tumors appeared less suitable.

ΔSUV_{max} is more firmly established for quantitative interim scan evaluation than is qPET. In contrast to ΔSUV_{max}, interim PET interpretation by qPET is based on a single scan, which minimizes the influence of factors known to impair SUV_{max} measurements. Prominent examples are blood glucose levels, adipose tissue, plasma clearance, paravenous injection, calibration and correction errors, and reconstruction algorithms (*25*).

While the Deauville scale distinguishes 5 response categories, dichotomization of the ΔSUV_{max} scale results in only 2 groups. In our study, patients with Deauville scores 2–4 did not differ in outcome. Interestingly, with all 3 methods, patients with complete interim PET normalization, comprising more than 25% of the total population, tended to fare better than patients with a good response but remaining uptake. Our conclusion that interim PET may identify 3 rather than 2 prognostic groups needs to be confirmed in an independent dataset.

Both qPET and ΔSUV_{max} convey the same information, but the proportion of patients identified to be at high risk of treatment failure is less than 15%. Most patients eventually failing therapy remain undetected. Treatment response is only one of several factors determining outcome. Others include lymphoma burden and distribution, which can readily be assessed at baseline PET scanning (*26,27*); gene expression (*28*); and genetic abnormalities (*29*). Combining one or several of these factors with early response assessment is likely to improve outcome prediction (*10,30*). Future studies will show whether radiation exposure can be eliminated by substituting serial measurement of circulating tumor DNA for PET/CT (*31*).

Strengths of our study include rigorously defined conditions for PET performance and treatment delivery, a large sample size encompassing the entire spectrum of DLBCL, and reevaluation of all scans by a single specialist whose interpretation was reconciled with previous assessments of the same scans. In a comparative study, overall interobserver agreement was found to be almost perfect for the ΔSUV_{max} approach but no more than substantial for the Deauville scale (*19*). As for qPET, data on interobserver concordance are not yet available. In rPET, relying on principles similar to those of qPET, interobserver agreement was found to be almost perfect (*20*).

CONCLUSION

The currently recommended method for the identification of high-risk patients at interim PET scanning appears of limited value in DLBCL. The visual Deauville scale should be replaced by one of the quantitative methods, such as qPET or ΔSUV_{max}, that minimize the confounding factors of visual assessment and permit outcome prediction on a continuous scale.

KEY POINTS

QUESTION: Should the visual Deauville scale be replaced by a quantitative method of interim ^{18}F-FDG PET evaluation in DLBCL?

PERTINENT FINDINGS: In a post hoc analysis of the PETAL trial, the positive predictive value of the categoric Deauville scale was lower than that of the continuous ΔSUV_{max} and qPET scales. The continuous scales conveyed similar information.

IMPLICATIONS FOR PATIENT CARE: ΔSUV_{max} and qPET are better suited for the identification of high-risk DLBCL patients than is the visual Deauville scale.

REFERENCES

1. Morton LM, Wang SS, Devesa SS, Hartge P, Weisenburger DD, Linet MS. Lymphoma incidence patterns by WHO subtype in the United States, 1992-2001. *Blood.* 2006;107:265–276.
2. Cunningham D, Hawkes EA, Jack A, et al. Rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisolone in patients with newly diagnosed diffuse large B-cell non-Hodgkin lymphoma: a phase 3 comparison of dose intensification with 14-day versus 21-day cycles. *Lancet.* 2013;381:1817–1826.
3. Armitage JO, Weisenburger DD, Hutchins M, et al. Chemotherapy for diffuse large-cell lymphoma: rapidly responding patients have more durable remissions. *J Clin Oncol.* 1986;4:160–164.
4. Mikhaeel NG, Hutchings M, Fields PA, O'Doherty MJ, Timothy AR. FDG-PET after two to three cycles of chemotherapy predicts progression-free and overall survival in high-grade non-Hodgkin lymphoma. *Ann Oncol.* 2005;16:1514–1523.
5. Haioun C, Itti E, Rahmouni A, et al. [^{18}F]fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) in aggressive lymphoma: an early prognostic tool for predicting patient outcome. *Blood.* 2005;106:1376–1381.
6. 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.
7. Barrington SF, Mikhaeel NG, Kostakoglu L, et al. 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.* 2014;32:3048–3058.
8. Lin C, Itti E, Haioun C, et al. Early ^{18}F-FDG PET for prediction of prognosis in patients with diffuse large B-cell lymphoma: SUV-based assessment versus visual analysis. *J Nucl Med.* 2007;48:1626–1632.
9. Casasnovas RO, Meignan M, Berriolo-Riedinger A, et al. Groupe d'étude des lymphomes de l'adulte (GELA). SUVmax reduction improves early prognosis value of interim positron emission tomography scans in diffuse large B-cell lymphoma. *Blood.* 2011;118:37–43.
10. Schmitz C, Hüttmann A, Müller SP, et al. Dynamic risk assessment based on positron emission tomography scanning in diffuse large B-cell lymphoma: post-hoc analysis from the PETAL trial. *Eur J Cancer.* 2020;124:25–36.
11. Itti E, Lin C, Dupuis J, 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;50:527–533.
12. 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.
13. Georgi TW, Kurch L, Hasenclever D, et al. Quantitative assessment of interim PET in Hodgkin lymphoma: an evaluation of the qPET method in adult patients in the RAPID trial. *PLoS One.* 2020;15:e0231027.
14. 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.
15. Dührsen U, Müller S, Hertenstein B, et al. Positron Emission Tomography-Guided Therapy of Aggressive Non-Hodgkin Lymphomas (PETAL): a multicenter, randomized phase III trial. *J Clin Oncol.* 2018;36:2024–2034.
16. Maurer MJ, Ghesquières H, Jais JP, et al. Event-free survival at 24 months is a robust end point for disease-related outcome in diffuse large B-cell lymphoma treated with immunochemotherapy. *J Clin Oncol.* 2014;32:1066–1073.
17. Rekowski J, Hüttmann A, Schmitz C, et al. Interim PET evaluation in diffuse large B-cell lymphoma using published recommendations: comparison of the Deauville 5-point scale and the ΔSUV_{max} method. *J Nucl Med.* 2021;62:37—42.
18. Horning SJ, Juweid ME, Schöder H, et al. Interim positron emission tomography scans in diffuse large B-cell lymphoma: an independent expert nuclear medicine evaluation of the Eastern Cooperative Oncology Group E3404 study. *Blood.* 2010;115:775–777.
19. 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 $\Delta SUVmax$. *Eur J Nucl Med Mol Imaging.* 2013;40:1312–1320.
20. Toledano MN, Vera P, Tilly H, Jardin F, Becker S. Comparison of therapeutic evaluation criteria in FDG-PET/CT in patients with diffuse large-cell B-cell lymphoma: prognostic impact of tumor/liver ratio. *PLoS One.* 2019;14:e0211649.
21. Itti E, Juweid ME, Haioun C, et al. Improvement of early ^{18}F-FDG PET interpretation in diffuse large B-cell lymphoma: importance of the reference background. *J Nucl Med.* 2010;51:1857–1862.
22. Fan Y, Zhang Y, Yang Z, et al. 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.* 2017;58:1–9.
23. Mylam KJ, Kostakoglu L, Hutchings M, et al. ^{18}F-fluorodeoxyglucose-positron emission tomography/computed tomography after one cycle of chemotherapy in patients with diffuse large B-cell lymphoma: results of a Nordic/US intergroup study. *Leuk Lymphoma.* 2015;56:2005–2012.
24. Kaalep A, Sera T, Rijnsdorp S, et al. Feasibility of state of the art PET/CT systems performance harmonisation. *Eur J Nucl Med Mol Imaging.* 2018;45:1344–1361.
25. Boellaard R, O'Doherty MJ, Weber WA, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imaging.* 2010;37:181–200.
26. Sasanelli M, Meignan M, Haioun C, et al. Pretherapy metabolic tumour volume is an independent predictor of outcome in patients with diffuse large B-cell lymphoma. *Eur J Nucl Med Mol Imaging.* 2014;41:2017–2022.
27. Cottereau AS, Nioche C, Dirand AS, et al. ^{18}F-FDG PET dissemination features in diffuse large B-cell lymphoma are predictive of outcome. *J Nucl Med.* 2020;61:40–45.
28. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature.* 2000;403:503–511.
29. Chapuy B, Stewart C, Dunford AJ, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat Med.* 2018;24:679–690.
30. Kim J, Lee JO, Paik JH, Lee WW, Kim SE, Song YS. Different predictive values of interim ^{18}F-FDG PET/CT in germinal center like and non-germinal center like diffuse large B-cell lymphoma. *Ann Nucl Med.* 2017;31:1–11.
31. Kurtz DM, Scherer F, Jin MC, et al. Circulating tumor DNA measurements as early outcome predictors in diffuse large B-cell lymphoma. *J Clin Oncol.* 2018;36:2845–2853.