Early ^{18}F-FDG PET for Prediction of Prognosis in Patients with Diffuse Large B-Cell Lymphoma: SUV-Based Assessment Versus Visual Analysis

Chieh Lin[1], Emmanuel Itti[2], Corinne Haioun[3], Yolande Petegnief[4], Alain Luciani[1], Jehan Dupuis[3], Gaetano Paone[2], Jean-Noël Talbot[4], Alain Rahmouni[1], and Michel Meignan[2]

[1]Department of Radiology, H. Mondor Hospital, AP-HP/Paris 12 University, Créteil, France; [2]Department of Nuclear Medicine, H. Mondor Hospital, AP-HP/Paris 12 University, Créteil, France; [3]Department of Hematology, H. Mondor Hospital, AP-HP/Paris 12 University, Créteil, France; and [4]Department of Nuclear Medicine, Tenon Hospital, AP-HP/Paris 6 University, Paris, France

The purpose of this study was to assess the prognostic value of early ^{18}F-FDG PET using standardized uptake value (SUV) compared with visual analysis in patients with diffuse large B-cell lymphoma (DLBCL). **Methods:** Ninety-two patients with newly diagnosed DLBCL underwent ^{18}F-FDG PET prospectively before and after 2 cycles of chemotherapy (at midtherapy). Maximum SUV (SUVmax) and mean SUV (SUVmean) normalized to body weight and body surface area, as well as tumor-to-normal ratios, were computed on the most intense uptake areas. The SUVs, tumor-to-normal ratios, and their changes over time were compared with visual analysis for predicting event-free survival (EFS) and overall survival, using receiver-operating-characteristic (ROC) analysis. Survival curves were estimated with Kaplan–Meier analysis and compared using the log-rank test. **Results:** With visual analysis, the accuracy of early PET to predict EFS was 65.2%. The 2-y estimate for EFS was 51% (95% confidence interval [CI], 34%–68%) in the PET-positive group compared with 79% (95% CI, 68%–90%) in the PET-negative group ($P = 0.009$). An optimal cutoff value of 65.7% SUVmax reduction from baseline to midtherapy obtained from ROC analysis yielded an accuracy of 76.1% to predict EFS. The 2-y estimate for EFS was 21% (95% CI, 0%–42%) in patients with SUVmax reduction $\leq$ 65.7% compared with 79% (95% CI, 69%–88%) in those with reduction $>$ 65.7% ($P < 0.0001$). Fourteen patients considered as positive on visual analysis could have been reclassified as good responders. **Conclusion:** SUV-based assessment of therapeutic response during first-line chemotherapy improves the prognostic value of early ^{18}F-FDG PET compared with visual analysis in DLBCL.

Key Words: early PET; standardized uptake value; lymphoma; prognosis; response

J Nucl Med 2007; 48:1626–1632
DOI: 10.2967/jnumed.107.042093

PET with ^{18}F-FDG is a well-recognized diagnostic tool used for staging and monitoring response to therapy in most lymphomas (*1–10*). The superiority of PET over CT to identify active disease after therapy completion (*11–13*) has recently led to revision of response criteria, allowing elimination of the complete remission/unconfirmed category (*14*). However, specific criteria for midtherapy PET interpretation have not been defined yet. In non-Hodgkin's lymphoma, ^{18}F-FDG uptake was found to decrease as early as 1 d after the initiation of chemotherapy (*15*). Several recent studies have also demonstrated that early assessment of response during the first treatment cycles is important to appreciate chemosensitivity and may potentially guide further risk-adapted therapeutic strategies in aggressive lymphoma (*4,6,7*).

Assessment of early response relies most often on visual analysis, which is subjective to the dichotomous interpretation of an observer or a panel of observers. In contrast to staging studies, where qualitative assessment of ^{18}F-FDG uptake is usually sufficient, treatment monitoring may require objective quantification of ^{18}F-FDG uptake changes. Standardized uptake value (SUV) is currently a popular semiquantitative, easy-to-calculate and noninvasive index of ^{18}F-FDG metabolic rate (*16,17*). However, sufficient confidence in the technical aspects of SUV calculation is not established yet, and clinical evidence of whether SUV is superior to visual analysis for outcome prediction is still lacking (*18,19*). Furthermore, a clear cutoff for an adequate SUV reduction during treatment remains to be defined (*20*).

The purpose of our study was to assess the prognostic value of early ^{18}F-FDG PET during first-line chemotherapy, using SUV semiquantification and comparison with visual analysis, in a homogeneous series of patients with diffuse large B-cell lymphoma (DLBCL).

MATERIALS AND METHODS

Patients

Between January 2000 and December 2005, a multicenter study involving 4 Departments of Hematology of the Assistance

Received Mar. 3, 2007; revision accepted May 31, 2007.
For correspondence or reprints contact: Michel Meignan, MD, PhD, Service de Médecine Nucléaire, Hôpital Henri Mondor, 51 Ave. du Marechal de Lattre de Tassigny, 94010 Créteil, France.
E-mail: michel.meignan@hmn.aphp.fr

Publique-Hôpitaux de Paris (AP-HP) was conducted prospectively on 110 patients with newly diagnosed and histologically proven aggressive non-Hodgkin's lymphoma to assess the prognostic value of early ^{18}F-FDG PET after 2 cycles of chemotherapy. The study was approved by the AP-HP review board, and all patients gave informed written consent. Early PET results did not influence the scheduled first-line therapeutic strategy. Results based on visual analysis involving the initial 90 patients have recently been published (*7*). Among the 110 patients, 104 had DLBCL (6 T-cell) but complete attenuation-corrected PET raw data were not available in 12 (scans done at other institutions). Therefore, 92 homogeneous DLBCL patients were included in the current study. Patient characteristics and chemotherapy regimens are summarized in Table 1. The database was closed in May 2006, with a median follow-up of 42 mo among survivors.

^{18}F-FDG PET

Patients underwent PET before initiation (PET1) and after 2 cycles (PET2) of chemotherapy, with a median interval of 14 d after the second cycle (range, 8–37 d). Patients fasted for 6 h, and PET was performed after having controlled the blood glucose level, which was targeted ≤ 7 mM. Before February 2004, the first 81 patients were scanned on a dedicated C-PET camera (ADAC). They were injected with 2 MBq/kg ^{18}F-FDG and sat at rest for 83 ± 22 min before imaging. The acquisition consisted of 5–7 overlapping bed shifts, to cover a volume from the upper thigh to the skull base (25-cm field of view). For each bed position, a 6-min emission scan was acquired in a 3-dimensional (3D) coincidence mode, followed by a 1-min transmission scan (^{137}Cs source). Images (144 × 144 matrix; voxel size, 4 × 4 × 4 mm^3) were reconstructed using an iterative ordered-subsets expectation maximization (OSEM) algorithm with attenuation correction. The last 11 patients were scanned on a Gemini PET/CT system (Philips). They were injected with 5 MBq/kg ^{18}F-FDG and rested for 69 ± 9 min before imaging. The acquisition featured a low-dose transmission CT scan (100 kV; 40 mAs; slice thickness, 5 mm), followed by the emission scan in 9–11 overlapping bed shifts (18-cm field of view), each for 3-min duration. Images (144 × 144 matrix; voxel size, 4 × 4 × 4 mm^3) were reconstructed with a 3D row-action maximum-likelihood algorithm (RAMLA). All patients also underwent a concurrent diagnostic CT scan of the chest, abdomen, and pelvis before, during midtherapy, and after treatment completion and then every 6 mo for follow-up based on the criteria of Cheson et al. (*21*).

TABLE 1
Patient Characteristics and Chemotherapy Regimens

Characteristic	Total (n = 92)
Median age, y (range)	54 (19–78)
Sex: men/women	60/32
Ann Arbor stage, no. (%)	
I–II	11 (12)
III–IV	81 (88)
Standard IPI score, no. (%)	
Low risk (L)	20 (22)
Low-intermediate (LI)	18 (19)
High-intermediate (HI)	30 (33)
High (H)	24 (26)
Chemotherapy regimen, no. (%)	
CHOP	3 (3)
R-CHOP	27 (29)
ACVBP/ACE	42 (46)
R-ACVBP	20 (22)

IPI = International Prognostic Index; CHOP = cyclophosphamide, hydroxydaunomycin, oncovin (vincristine), prednisone; R = rituximab; ACVBP = adriamycin (doxorubicin), cyclophosphamide, vindesine, bleomycin, prednisone; ACE = adriamycin, cyclophosphamide, etoposide.

Visual Analysis of ^{18}F-FDG Uptake

PET images were analyzed by a consensus of 2 experienced observers who were unaware of clinical, radiologic, and follow-up data. All foci were scored for their extent and intensity using a 3-point scale (1 = low, 2 = moderate, 3 = high) (*3,7*). The extent was scored within each lymphatic area, organ, or skeletal region, depending on the number of nodes or volume involved; the intensity was scored compared with surrounding tissues after upper thresholding of the data to have the liver activity around 30% of the gray scale. Then, PET2 was scored as positive or negative in comparison with PET1. Negative was defined as having either no residual abnormal uptake or having a unique residual site (with an extent score of 1) associated with an intensity score of 1, whereas all other previously hypermetabolic sites had disappeared. Positive was defined as having at least 1 residual site (with an extent score of 1) associated with an intensity score of 2, or as having ≥2 residual sites with any extent and intensity scores.

SUV-Based Assessment of ^{18}F-FDG Uptake

For each PET dataset, the tumor (T) with the most intense ^{18}F-FDG uptake among all foci was carefully identified, relying on a graded color-scaled parametric analysis (Fig. 1). From the activity profile crossing the hottest point, an isocontour threshold was determined halfway between the background and the maximal pixel value (*22*) and was automatically propagated on adjacent slices to encompass the entire tumor volume. Maximal and mean counts-per-pixel were computed within the volumetric region of interest (ROI). If present, a central cold area was included. In addition, 2 large ROIs were manually drawn over gluteal muscle regions (Fig. 1), from which the counts-per-pixel were averaged to define the normal background (N).

To assess metabolic changes during chemotherapy, the hottest tumor in any region or organ on PET2 was used for comparison and as the indicator for disease status, even though its location differed from the initial hottest tumor on PET1. In cases in which all lesions had disappeared, ROIs were manually drawn in the same area on PET2 as that on PET1, with careful slice-to-slice comparison and by making sure that the ROI size was restricted to the baseline tumor. In addition, we also investigated ^{18}F-FDG uptake changes on PET2 within the initial hottest tumor site on PET1, even though PET2 demonstrated hottest foci on other locations.

SUVs were calculated from the counts-per-pixel and normalized to body weight (BW) and body surface area, defined as BSA (m^2) = 0.007184 weight (kg)$^{0.425}$ × height (cm)$^{0.725}$ (*23,24*), using the following formulas:

$$SUV_{BW} = \frac{\text{tissue activity (kBq/mL)}}{\text{injected activity}^*\text{(MBq)/weight (kg)}}, \quad \text{Eq. 1}$$

$$SUV_{BSA} = \frac{\text{tissue activity (kBq/mL)}}{\text{injected activity}^*\text{(MBq)/BSA (m}^2\text{)}}, \quad \text{Eq. 2}$$

where *activity was decay-corrected from the delay between injection and image acquisition. In addition, the tumor-to-normal

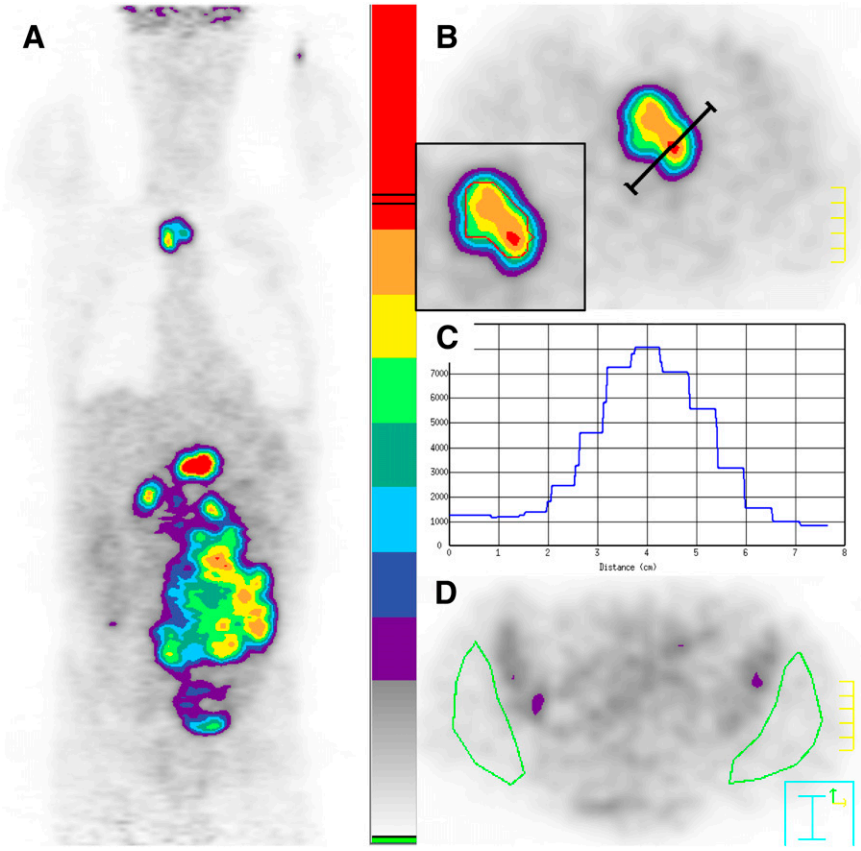

FIGURE 1. Selection of ROIs in 57-y-old patient before chemotherapy. (A) Graded color-scaled parametric analysis applied in reconstructed coronal PET image shows most active tumor in upper abdomen. (B) Transverse PET image with a higher scale reveals celiac tumor (T) with activity profile crossing the hottest point (red spot). (C) Corresponding activity profile in counts-per-pixel. Isocontours are drawn with lower autocontour threshold of 4,500 counts-per-pixel (red isocontour at inset in B). (D) Normal background tissue (N): 2 large ROIs are manually selected on gluteal muscles, avoiding iliac bone marrow activity.

uptake ratio (T/N ratio) for each PET image was computed as follows (Eq. 3):

$$\text{T/N ratio} = \frac{\text{maximal counts-per-pixel within T}}{\text{mean counts-per-pixel within N}}. \quad \text{Eq. 3}$$

Statistical Analysis

To evaluate the prognostic value of early PET, event-free survival (EFS) and overall survival (OS) were chosen as endpoints. EFS was defined as the date of enrollment to first evidence of progression, relapse, or death from any cause. Data were censored if the patients were alive and free of progression or relapse at last follow-up. OS was defined as the date of enrollment to death from any cause. Data were censored if the patients were alive at last follow-up. Receiver-operating-characteristic (ROC) analysis was performed to determine an optimal cutoff value of uptake on PET2 or an optimal cutoff value of uptake reduction from PET1 to PET2 in predicting EFS—event versus no event—and OS—dead versus alive. Differences in SUVs between groups were analyzed with an unpaired Student *t* test, and significance was obtained when the 2-sided *P* value was < 0.05. Survival according to visual analysis and SUV-based assessment of early PET was depicted using the Kaplan–Meier plots and compared using the log-rank test.

RESULTS

Patient Outcome

During a median follow-up period of 42 mo after inclusion, 60 patients had no event (EFS = 65.2%), whereas the remaining 32 patients progressed or died, with a median delay of 6.7 mo; in addition, 71 patients survived (OS = 77.2%), whereas the remaining 21 patients died, with a median delay of 9.1 mo.

Visual Analysis of Survival

All patients demonstrated intense foci of uptake on PET1, as expected from DLBCL. At midtherapy, PET2 was interpreted as positive in 34 patients and negative in 58 patients. The 2-y estimate for EFS was 51% (95% confidence interval [CI], 34%–68%) in PET2-positive patients compared with 79% (95% CI, 68%–90%) in PET2-negative patients ($P = 0.009$; Fig. 2A). Positive and negative predictive values (PPVs and NPVs, respectively), as well as accuracies in predicting EFS and OS (65.2% and 68.5%, respectively) are listed in Table 2. Of the 34 PET2-positive patients, 17 remained free of an event at last follow-up.

SUV-Based Assessment of Survival

There was no statistical difference between the SUVs computed from the C-PET system and those obtained from the Gemini PET/CT, on both PET1 and PET2 ($P = 0.6$ and 0.3, respectively). At baseline, $SUV_{BW}max$ averaged 13.2 ± 4.8, whereas at midtherapy $SUV_{BW}max$ decreased to 3.4 ± 2.7, corresponding to a mean reduction of 71.7%. Among the 34 PET2-positive patients, the hottest tumor remained in the same site as PET1 in 50%, whereas the location changed in 50%, either because the tumor responded

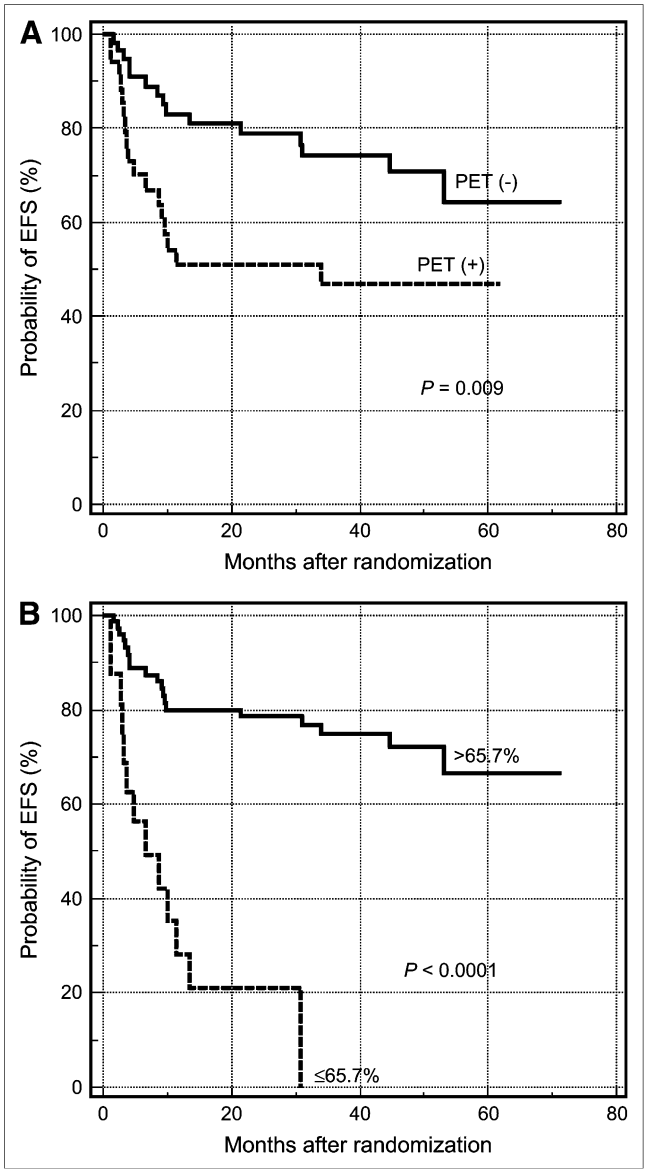

FIGURE 2. Kaplan–Meier plots for estimation of probability of EFS according to PET status at midtherapy. (A) Survival curves based on visual analysis. (B) Survival curves based on percentages of SUV_{BW}max reduction.

better to therapy than other locations ($n = 15$) or, in the case of progression, because another location demonstrated higher uptake ($n = 2$). SUV_{BW}max reduction averaged 55.4% in PET2-positive patients versus 81.2% in PET2-negative patients ($P < 0.0001$). All SUV_{BW}, SUV_{BSA} values and T/N ratios are displayed in Table 3.

With ROC analysis, an optimal cutoff value of SUV_{BW}max of 5.0 at PET2 could predict EFS with higher PPV, NPV, and accuracy (75.0%) than visual analysis (Table 2). SUV_{BSA}max performed even better for predicting outcome, with an accuracy reaching 88.0% for OS. At midtherapy, most patients showed SUV_{BW}max ≤ 5.0, and these patients ($n = 79$) tended to have lower baseline SUV_{BW}max (12.8 ± 4.7) than those ($n = 13$) with SUV_{BW}max > 5.0 (15.4 ± 5.0; $P = 0.073$). Furthermore, SUV_{BW}max at PET1 failed to demonstrate a significant predictive value for EFS (area under ROC curve = 0.525, $P = 0.7$).

The percentage of SUV_{BW}max reduction from PET1 to PET2 averaged $60.7\% \pm 32.6\%$ in the 32 patients whose disease progressed or who died versus $77.5\% \pm 12.4\%$ in the 60 patients who remained free of disease ($P < 0.0006$). ROC analysis yielded an optimal cutoff value of 65.7% SUV_{BW}max reduction at midtherapy for predicting EFS. The overall accuracy increased to 76.1% (Table 2). In patients with SUV_{BW}max reduction $\leq 65.7\%$ ($n = 16$), the 2-y estimate for EFS was only 21% (95% CI, 0%–42%) compared with 79% (95% CI, 69%–88%) in those with SUV_{BW}max reduction $> 65.7\%$ ($n = 76$) ($P < 0.0001$; Fig. 2B). Results obtained from both the reduction of SUV_{BW}mean and T/N ratio showed slightly poorer accuracies and PPVs—that is, more false-positive scans. SUVmax reduction, whether normalized to BW or BSA, also gave slightly better PPVs and accuracies for OS (Table 2).

When considering the ^{18}F-FDG uptake change with regard to the most active focus on the baseline scan only, ROC analysis led to an 83.3% PPV, 72.5% NPV, and 73.9% accuracy for EFS. The optimal cutoff value was 65.7%.

DISCUSSION

In the present study, we emphasize, in a homogeneous series of 92 patients with DLBCL, that SUV-based assessment of glucose metabolic changes after 2 treatment cycles improves the prognostic value of early ^{18}F-FDG PET, compared with visual analysis, with a median follow-up of 42 mo.

A negative interim scan based on visual analysis during first-line chemotherapy has proven to be an independent indicator of favorable outcome among patients with low-risk or high-risk disease based on the International Prognostic Index (*7*). Moreover, a more recent study showed that a significant survival difference exists between patient groups on the basis of early PET results but not on the gene-expression profiles (*25*). However, visual analysis of PET may be improved because some PET-positive patients still have a good outcome (*7,26,27*). If response on interim ^{18}F-FDG PET is to be used to guide second-line risk-adapted therapeutic strategies in the future, efforts should be made to decrease the false-positive scans so that patients are not overtreated—that is, to improve the PPV of PET (*27*). In the present series, 17 of 34 patients with residual ^{18}F-FDG uptake, considered visually positive on PET2, remained free of an event. The same issue was also raised in a recent study with advanced-stage Hodgkin's disease (*28*). Such patients could have already shown significant reduction of ^{18}F-FDG uptake after first-line chemotherapy but still presented with an increased activity compared with the surrounding normal tissue by visual analysis.

TABLE 2
Outcome Prediction by Means of Visual Analysis and SUVs in 92 Patients

Parameter	EFS			OS		
	PPV (%)	NPV (%)	Accuracy* (%)	PPV (%)	NPV (%)	Accuracy* (%)
Visual analysis	50.0	74.1	65.2	38.2	86.2	68.5
SUV-based assessment						
PET2 values						
SUV_{BW}max	84.6	73.4	75.0 (0.661)	84.6	87.3	87.0 (0.740)
SUV_{BSA}max	91.7	73.8	76.1 (0.655)	91.7	87.5	88.0 (0.735)
SUV_{BW}mean	58.6	76.2	70.7 (0.671)	63.2	87.7	82.6 (0.727)
SUV_{BSA}mean	64.0	76.1	72.8 (0.666)	70.6	88.0	84.8 (0.718)
T/N ratio	64.0	76.1	72.8 (0.633)	56.5	88.4	80.4 (0.696)
Uptake reduction from PET1 to PET2						
SUV_{BW}max	81.3	75.0	76.1 (0.664)	73.3	87.0	84.8 (0.689)
SUV_{BSA}max	81.3	75.0	76.1 (0.672)	73.3	87.0	84.8 (0.695)
SUV_{BW}mean	70.8	77.9	76.1 (0.684)	54.2	88.2	79.3 (0.683)
SUV_{BSA}mean	72.0	79.1	77.2 (0.692)	56.5	88.4	80.4 (0.694)
T/N ratio	68.2	75.7	73.9 (0.665)	68.8	86.8	83.7 (0.680)

*The areas under the ROC curves are presented in parentheses.

Results of SUV-based assessment are obtained from ROC analyses. EFS = event-free survival; OS = overall survival; PPV = positive predictive value; NPV = negative predictive value; BW = body weight; BSA = body surface area; T/N ratio = tumor-to-normal ratio.

To be able to quantify the ^{18}F-FDG metabolic rate instead of interpreting images only as positive or negative, Römer et al. demonstrated in 11 lymphoma patients that Patlak analysis of ^{18}F-FDG kinetics may provide superior information in therapy monitoring (*1*). However, quantification of glucose metabolic rate, which requires dynamic imaging on a restricted field of view and measurement of the arterial input function, is generally regarded as too complex in routine practice (*23,24*). Moreover, it may not be suitable in DLBCL when the most active lesion indicating tumor viability can be outside the field of view (18–25 cm). In Hodgkin's lymphoma, Hutchings et al. demonstrated that SUV analysis of interim PET may help in patient stratification (*29*). Our study also shows that an easy-to-calculate semiquantitative index, such as the SUVmax, is adapted to assess early response and predict long-term outcome. The PPV of early PET for EFS can be improved from 50% with visual analysis to 81.3% when using SUV_{BW}max reduction of 65.7% from PET1 to PET2 as a cutoff value. Fourteen patients could have been considered good responders in this case without altering NPV (at the expense of 4 more false-negative scans), among whom 11 were in complete remission at the end of first-line therapy and remain free of an event at the last follow-up (Fig. 3). The overall accuracy for EFS prediction based on SUV_{BW}max reduction compared with visual analysis is 76.1% versus 65.2%, with even higher performance in predicting OS, as shown in Table 2. More importantly, Kaplan–Meier analysis demonstrates much higher statistical significance between EFS curves using the SUV approach.

One of the reasons why the usefulness of SUV in clinical application remained controversial is that being a simplified semiquantitative method, SUV is prone to many sources of variability from one institution to another. In our series, SUV_{BW}max in the most intense lesion at baseline averaged 13.2 ± 4.8, which compares well with the value of

TABLE 3
Mean SUV_{BW}, SUV_{BSA}, and T/N Ratio

Parameter	PET1	PET2		
	Total (n = 92)	Total (n = 92)	PET2-positive (n = 34)	PET2-negative (n = 58)
SUV_{BW}max	13.2 ± 4.8 (4.8–24.8)	3.4 ± 2.7 (0.9–20.8)	5.4 ± 3.5 (1.9–20.8)	2.2 ± 0.7 (0.9–4.5)
SUV_{BW}mean	9.5 ± 3.6 (3.4–18.0)	2.4 ± 2.1 (0.6–16.2)	4.1 ± 2.8 (1.5–16.2)	1.5 ± 0.5 (0.6–2.9)
SUV_{BSA}max	0.35 ± 0.13 (0.11–0.67)	0.09 ± 0.07 (0.03–0.53)	0.15 ± 0.09 (0.05–0.53)	0.06 ± 0.02 (0.03–0.11)
SUV_{BSA}mean	0.25 ± 0.10 (0.08–0.49)	0.07 ± 0.06 (0.02–0.42)	0.11 ± 0.07 (0.04–0.42)	0.04 ± 0.01 (0.02–0.08)
T/N ratio	19.5 ± 8.0 (6.6–45.7)	4.7 ± 3.3 (1.2–21.5)	7.4 ± 4.1 (2.9–21.5)	3.1 ± 0.9 (1.2–4.9)

BW = body weight; BSA = body surface area; T/N ratio = tumor-to-normal ratio.

Values are mean ± SD (range).

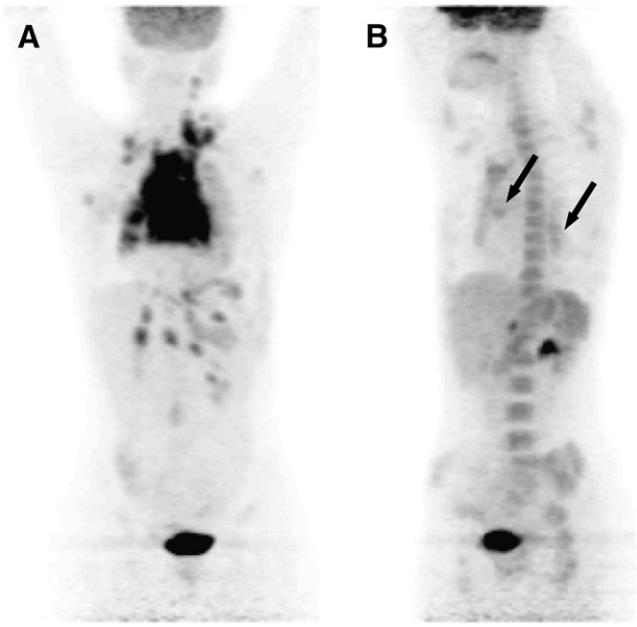

FIGURE 3. A 25-y-old woman with bulky mediastinal uptake on PET1 (A, anterior view), who was interpreted as positive visually on PET2 (B, left anterior oblique view). Two residual foci were diagnosed (arrows), scored as extent = 1, intensity = 1 for retrosternal site and extent = 2, intensity = 1 for left hilar site. However, SUV_{BW}max reduction was measured at 80.7%—that is, above the cutoff value of 65.7%, indicating good early response. Patient remained in complete remission after 36 mo of follow-up.

17.2 ± 9.7 reported by Schöder et al. in 63 patients with aggressive lymphoma (43 new cases of disease) (*30*). Our SUVs are relatively lower but appear more homogeneous with regard to the lower SD. Variability in SUV measurement can stem from many factors, including the imaging delay after ^{18}F-FDG injection, the partial-volume effect, and the applied normalization scheme (*17,22,31*).

With regard to the imaging delay, we generally performed image acquisition >60 min after injection, which is considered the time required for the ^{18}F-FDG uptake to reach a plateau (*16*). Although the time interval varied slightly between the C-PET and Gemini acquisitions for logistics reasons inherent to a multicenter study, no scan was performed earlier than 48 min.

The partial-volume effect is related primarily to the limited spatial resolution of PET (*32*) and results in an underestimation of ^{18}F-FDG uptake in small lesions. Previous studies have shown that SUV measurement is more reliable for an object with a diameter at least 2.5-fold the PET intrinsic spatial resolution (*31,32*). For the same reason, SUVmax is regarded as better than SUVmean as a metabolic index, especially in small lesions (*31*). In our study, SUVmax, normalized either to BW or to BSA, demonstrated PPVs and accuracies for outcome prediction that were superior to those of SUVmean and T/N ratio, and we recommend the use of SUVmax. The method of ROI definition may also have an impact on the partial-volume effect: ROIs obtained from the PET activity profiles, as it was done in our study, correspond most closely to the actual tumor size (*22*). Moreover, we have investigated the differences in SUVs between PET-based and CT-based ROI definitions on coregistered PET/CT images of a simple phantom (a series of syringes of increasing diameters ranging from 0.4 to 2.5 cm, filled with a homogenated ^{18}F-FDG solution) and showed that CT-based ROIs would not improve SUV measurements (*33*).

As to the correction scheme, SUV normalized to BW for many tissues was found to have a strong positive correlation with weight (*34,35*) because ^{18}F-FDG uptake is lower in fat than in other tissues and, consequently, SUVs tend to be overestimated in heavier patients (*24*). Because body weight often changes during treatment, BSA normalization was proposed. In our study, SUV_{BSA}max values on PET2 alone showed a better PPV than SUV_{BW}max for EFS and OS prediction; however, when considering the percentage of SUV reduction, SUV_{BSA}max and SUV_{BW}max gave identical values and accuracies for EFS and OS prediction.

For all of the reasons discussed, it seems difficult to rely on one single SUV at a given time point to appreciate the therapeutic response and to predict outcome. Indeed, because a cutoff value for an absolute SUV can vary greatly between different institutions (here, SUV_{BW}max of 5.0 on PET2), the measurement of an interscan SUV reduction within the same institution is probably a better and more reproducible approach (here, a 65.7% SUV_{BW}max reduction). This is confirmed by the higher accuracy obtained for EFS from ROC analysis with the second analysis, for ^{18}F-FDG uptake change, over the first one.

One could regret that most patients were scanned on the C-PET and few on the Gemini, which could introduce quantification bias. This limitation is due to the prospective nature of our study, with subsequent evolution of the technology. Even though the calibration factors could vary slightly between scanners related to different attenuation correction methods or reconstruction algorithms, univariate analysis has proven that SUVs obtained from these 2 systems showed no statistical difference. Most important, each patient had his or her 2 PET scans on the same machine; therefore, the potential systemic bias could have been eliminated by computing the SUV_{BW}max differences. Another limitation of our study is the use of a post hoc response criterion for SUV-based analysis (obtained from the same patient population), instead of a predefined response criterion, as we did for visual analysis.

It must be pointed out that in our study, when a lesion different from the baseline tumor showed the most intense activity on PET2, which happened in 17 patients (18% of 92), we used its SUV as the index of ^{18}F-FDG uptake at midtherapy. When only the change of SUV_{BW}max within the initial tumor was considered in the analysis, more false-negative scans were noted in predicting EFS. Indeed, in this case the overall accuracy was slightly inferior (73.9% instead of 76.1%).

CONCLUSION

Our findings indicate the potential of improving the prognostic value of early ^{18}F-FDG PET by using SUV-based rather than visual analysis in DLBCL. The optimal cutoff value for $SUV_{BW}max$ reduction from baseline to midtherapy is 65.7% for predicting EFS. This cutoff value, however, may require refinement under circumstances of different treatment regimens and in other histologic types of lymphomas (*27*), and we look forward for its application in other study groups. Potential implications for patient care will be to provide a more reproducible assessment of early PET studies and, eventually, to guide risk-adapted therapies.

ACKNOWLEDGMENTS

We are indebted to the entire team of the AP-HP PET Center at Tenon Hospital, Paris, for their help with C-PET imaging. Particular thanks go to Marie-Joséphine Waryn and Sébastien Mrozowicz for their collaborations in the phantom study, Julien-Aymeric Simonnet for his review of statistical analysis, and Marie-Claude Bassene and Antoine Allain for their help with database management. This study was supported by the Délégation à la Recherche Clinique de l'Assistance Publique-Hôpitaux de Paris (PHRC-AOM00152) and the Société Française de Radiologie.

REFERENCES

1. Römer W, Hanauske AR, Ziegler S, et al. Positron emission tomography in non-Hodgkin's lymphoma: assessment of chemotherapy with fluorodeoxyglucose. *Blood.* 1998;91:4464–4471.
2. Jerusalem G, Beguin Y, Fassotte MF, et al. Persistent tumor ^{18}F-FDG uptake after a few cycles of polychemotherapy is predictive of treatment failure in non-Hodgkin's lymphoma. *Haematologica.* 2000;85:613–618.
3. Mikhaeel NG, Timothy AR, O'Doherty MJ, Hain S, Maisey MN. 18-FDG-PET as a prognostic indicator in the treatment of aggressive non-Hodgkin's lymphoma: comparison with CT. *Leuk Lymphoma.* 2000;39:543–553.
4. Kostakoglu L, Coleman M, Leonard JP, Kuji I, Zoe H, Goldsmith SJ. PET predicts prognosis after 1 cycle of chemotherapy in aggressive lymphoma and Hodgkin's disease. *J Nucl Med.* 2002;43:1018–1027.
5. Spaepen K, Stroobants S, Dupont P, et al. Early restaging positron emission tomography with ^{18}F-fluorodeoxyglucose predicts outcome in patients with aggressive non-Hodgkin's lymphoma. *Ann Oncol.* 2002;13:1356–1363.
6. Torizuka T, Nakamura F, Kanno T, 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;31:22–28.
7. 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.
8. 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.
9. Tatsumi M, Cohade C, Nakamoto Y, Fishman EK, Wahl RL. Direct comparison of FDG PET and CT findings in patients with lymphoma: initial experience. *Radiology.* 2005;237:1038–1045.
10. Rahmouni A, Luciani A, Itti E. MRI and PET in monitoring response in lymphoma. *Cancer Imaging.* 2005;5(A):S106–S112.
11. Jerusalem G, Beguin Y, Fassotte MF, et al. Whole-body positron emission tomography using ^{18}F-fluorodeoxyglucose for posttreatment evaluation in Hodgkin's disease and non-Hodgkin's lymphoma has higher diagnostic and prognostic value than classical computed tomography scan imaging. *Blood.* 1999;94:429–433.
12. Guay C, Lepine M, Verreault J, Benard F. Prognostic value of PET using ^{18}F-FDG in Hodgkin's disease for posttreatment evaluation. *J Nucl Med.* 2003;44:1225–1231.
13. Kazama T, Faria SC, Varavithya V, Phongkitkarun S, Ito H, Macapinlac HA. FDG PET in the evaluation of treatment for lymphoma: clinical usefulness and pitfalls. *Radiographics.* 2005;25:191–207.
14. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol.* 2007;25:579–586.
15. Yamane T, Daimaru O, Ito S, Yoshiya K, Nagata T, Uchida H. Decreased ^{18}F-FDG uptake 1 day after initiation of chemotherapy for malignant lymphomas. *J Nucl Med.* 2004;45:1838–1842.
16. Huang SC. Anatomy of SUV: standardized uptake value. *Nucl Med Biol.* 2000;27:643–646.
17. Thie JA. Understanding the standardized uptake value, its methods, and implications for usage. *J Nucl Med.* 2004;45:1431–1434.
18. Kostakoglu L, Goldsmith SJ. ^{18}F-FDG PET evaluation of the response to therapy for lymphoma and for breast, lung, and colorectal carcinoma. *J Nucl Med.* 2003;44:224–239.
19. Juweid ME, Stroobants S, Hoekstra OS, 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;25:571–578.
20. Kasamon YL, Jones RJ, Wahl RL. Integrating PET and PET/CT into the risk-adapted therapy of lymphoma. *J Nucl Med.* 2007;48(suppl 1):19S–27S.
21. Cheson BD, Horning SJ, Coiffier B, et al. Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas: NCI Sponsored International Working Group. *J Clin Oncol.* 1999;17:1244–1253.
22. Boellaard R, Krak NC, Hoekstra OS, Lammertsma AA. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. *J Nucl Med.* 2004;45:1519–1527.
23. Graham MM, Peterson LM, Hayward RM. Comparison of simplified quantitative analyses of FDG uptake. *Nucl Med Biol.* 2000;27:647–655.
24. Hoekstra CJ, Paglianiti I, Hoekstra OS, et al. Monitoring response to therapy in cancer using [^{18}F]-2-fluoro-2-deoxy-D-glucose and positron emission tomography: an overview of different analytical methods. *Eur J Nucl Med.* 2000;27:731–743.
25. Dupuis J, Gaulard P, Itti E, et al. Early response evaluation with 18FDG-PET scanning, but not phenotypic profile, are predictive of outcome in diffuse large B-cell lymphoma [abstract]. *Blood.* 2005;106:1914.
26. Castellucci P, Zinzani P, Pourdehnad M, et al. ^{18}F-FDG PET in malignant lymphoma: significance of positive findings. *Eur J Nucl Med Mol Imaging.* 2005;32:749–756.
27. Jerusalem G, Beguin Y. The place of positron emission tomography imaging in the management of patients with malignant lymphoma. *Haematologica.* 2006;91:442–444.
28. Gallamini A, Rigacci L, Merli F, 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;91:475–481.
29. Hutchings M, Loft A, Hansen M, et al. FDG-PET after two cycles of chemotherapy predicts treatment failure and progression-free survival in Hodgkin lymphoma. *Blood.* 2006;107:52–59.
30. Schöder H, Noy A, Gonen M, et al. Intensity of 18fluorodeoxyglucose uptake in positron emission tomography distinguishes between indolent and aggressive non-Hodgkin's lymphoma. *J Clin Oncol.* 2005;23:4643–4651.
31. Keyes JW Jr. SUV: standard uptake or silly useless value? *J Nucl Med.* 1995;36:1836–1839.
32. Hoffman EJ, Huang SC, Phelps ME. Quantitation in positron emission computed tomography. 1. Effect of object size. *J Comput Assist Tomogr.* 1979;3:299–308.
33. Lin C. *Optimization of SUV (Standardized Uptake Value) Measurement for the Evaluation of the Therapeutic Response in Aggressive Lymphoma* [MSc thesis]. Creteil, France: Paris XII University; 2006.
34. Sugawara Y, Zasadny KR, Neuhoff AW, Wahl RL. Reevaluation of the standardized uptake value for FDG: variations with body weight and methods for correction. *Radiology.* 1999;213:521–525.
35. Yeung HW, Sanches A, Squire OD, Macapinlac HA, Larson SM, Erdi YE. Standardized uptake value in pediatric patients: an investigation to determine the optimum measurement parameter. *Eur J Nucl Med Mol Imaging.* 2002;29:61–66.