Interim ^{18}F-FDG PET SUVmax Reduction Is Superior to Visual Analysis in Predicting Outcome Early in Hodgkin Lymphoma Patients

Cédric Rossi[1], Salim Kanoun[2], Alina Berriolo-Riedinger[2], Inna Dygai-Cochet[2], Olivier Humbert[2], Caroline Legouge[1], Marie Lorraine Chrétien[1], Jean-Noel Bastie[1], François Brunotte[2], and René-Olivier Casasnovas[1]

[1]Hématologie Clinique, CHU Le Bocage, Dijon, France; and [2]Médecine Nucléaire, Centre G.F. Leclerc, Dijon, France

PET performed after 2 cycles of chemotherapy (PET2) allows prediction of outcome in most patients with Hodgkin lymphoma (HL). Visual analysis using a 5-point scale was proposed to assess PET response, but a semiquantitative approach using maximum standardized uptake value (SUVmax) reduction between baseline and interim PET was shown to be superior to the 5-point scale in patients with diffuse large B-cell lymphoma and may also improve the accuracy of interim PET interpretation in HL. To compare the clinical usefulness of both methods in HL patients, we analyzed PET2 according to visual and ΔSUVmax criteria in a retrospective single-center study. **Methods:** From 2007 to 2010, 59 consecutive patients with a first diagnosis of HL were treated with 4–8 cycles of anthracycline-based chemotherapy. Radiotherapy was performed in 19 responding patients with localized disease. PET was done at baseline (PET0) and after 2 cycles of chemotherapy, and treatment was not modified according to the PET2 result. PET2 was interpreted using the 5-point scale (positivity for score 4 or 5). The SUVmax reduction between PET0 and PET2 (ΔSUVmax) was computed for all patients, and patients with a ΔSUVmax greater than 71% were considered good responders. **Results:** When the 5-point scale was used, 46 patients (78%) achieved a negative PET2 result, 7 of whom failed treatment (negative predictive value, 85%). Forty-nine patients (83%) had a ΔSUVmax greater than 71%, 6 of whom failed treatment (negative predictive value, 88%). The PET2 positive predictive value was significantly better for ΔSUVmax (70%) than for the 5-point scale (46%). When ΔSUVmax was used, 6 (46%) of the 13 PET2-positive patients could be reclassified as good responders. Although visual PET2 positivity was related to a lower 4-y progression-free survival (45%) compared with PET2 negativity (81%, $P < 0.002$), ΔSUVmax (>71 vs ≤71%) was more accurate for identifying patients with different 4-y progression-free survivals (82% vs. 30%; $P < 0.0001$). In multivariate analysis using the international prognosis score and ΔSUVmax as covariates, ΔSUVmax remained the unique independent predictor for progression-free survival ($P = 0.0001$; relative risk, 8.1). **Conclusion:** Semiquantitative analysis was more accurate than visual analysis based on the 5-point scale to interpret PET2 and predict the outcome of HL patients. These encouraging results warrant further confirmation in larger and prospective series.

T he current treatment of patients with Hodgkin lymphoma (HL) allows the expectation of a high complete-response rate and an overall progression-free survival (PFS) of 80% at 5 y (*1–5*). In advanced disease, the risk of treatment failure is much higher and reaches about 30% in patients treated with the ABVD regimen (doxorubicin, bleomycin, vinblastine, and dacarbazine). Even if BEACOPPesc (bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone) can improve disease control, the early and long-term toxicity related to this regimen leads to similar overall survivals in patients treated with either BEACOPPesc or ABVD. In early stages of the disease, the landscape is quite different since the rate of treatment failure is low. The main goal of treatment is to reduce toxicity without impairing disease control. An early way of identifying patients who are at high risk of failing ABVD treatment is needed in both early and advanced stages of the disease. ^{18}F-FDG PET may help to identify these patients early (*1*).

Early evaluation of chemosensitivity through visual interpretation of interim PET has been shown to have a stronger predictive value than the currently available prognosis scores (*6*). However, most of the first reported studies used heterogeneous visual criteria that may affect the prognostic accuracy of interim PET. Indeed, the sensitivity of interim PET for identifying patients with different outcomes was found to range from 67% to 100% (*7*). The need for harmonized criteria to interpret interim PET led, in 2009, to an expert consensus that defined visual criteria based on a 5-point scale using liver uptake as a reference. Thus, a residual mass having ^{18}F-FDG uptake higher than that of the liver was considered a positive PET finding (*8*). However, the use of the 5-point scale did not preclude interobserver reproducibility issues (*9,10*). Alternative approaches to visual analysis were developed to improve the accuracy and reproducibility of interim PET, these being based mainly on analysis of maximum standardized uptake value (SUVmax). The SUVmax reduction between baseline and interim PET (*11*) was shown to be superior to visual analysis (*12*) in patients with high-risk diffuse large B-cell lymphoma.

The present study evaluated the influence of the interim-PET interpretation criteria, either ΔSUVmax or the 5-point scale, on predictions of the prognosis of HL patients.

MATERIALS AND METHODS

Patients

We retrospectively analyzed 59 consecutive patients with a first diagnosis of classic HL according to the 2008 World Health Organization classification of hematologic malignancies (13). The patients had been referred to the hospital of Dijon from January 2007 to January 2010. Patients with positive serology for HIV were excluded. The extent of disease was staged in accordance with the Ann Arbor classification using bone marrow biopsy and enhanced CT scans of the neck, thorax, abdomen, and pelvis. Patient characteristics are listed in Table 1.

The institutional review board of the hospital approved the study, and all patients provided written informed consent. The study procedures were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2008.

Treatment and Patient Outcomes

Patients were treated according to the recommendations of the Lymphoma Study Association (formerly Groupe d'étude des lymphomes de l'adulte): patients with stage I or II disease received 4–6 cycles of an anthracycline-based chemotherapy regimen, followed by 20–36 Gy of involved-field radiotherapy; patients with stage III or IV disease received 8 cycles of anthracycline-based chemotherapy.

Response was assessed using the revised criteria of Cheson et al. (14) at the end of the first line of treatment or, in patients with progressive disease, at the time of progression. PET/CT and CT scan assessment was used, along with bone marrow biopsy if bone marrow was involved at baseline. Five patients (8%) had progressive disease. Fifty-two (88%) and 2 (4%) patients achieved a complete and partial response, respectively, leading to a 92% overall response rate, and 5 patients (8%) relapsed. Five patients (8%) died: 3 from HL progression, 1 from hepatocellular cancer, and 1 from bleomycin-related pulmonary fibrosis. The median follow-up time was 50 mo (range, 22–71 mo).

PET Acquisition

PET was performed at baseline (PET0) and after 2 courses of chemotherapy (PET2) for all patients, according to the policy of systematic PET evaluation for ^{18}F-FDG–avid lymphoma in the Dijon Hematology Department since 2005. The therapeutic strategy was not changed on the basis of the PET2 results.

Whole-body PET was performed sequentially using a dedicated PET/CT system (Gemini GXL or Gemini TOF; Philips). CT scans were used for anatomic registration but also for attenuation correction. Emission data were corrected for dead time, random and scatter coincidences, and attenuation before reconstruction with the row-action maximum-likelihood algorithm iterative method. The image voxel counts were calibrated to activity concentration (Bq/mL) and were decay-corrected using the time of tracer injection as a reference.

All patients were instructed to fast for at least 6 h before the injection of ^{18}F-FDG. Serum glucose level was measured by the hexokinase method. Whole-body emission and transmission scans were acquired in 3-dimensional mode, 60 min after the intravenous administration of 5 (Gemini GXL) or 3 (Gemini TOF) MBq of ^{18}F-FDG per kilogram. The second PET scanner, used for 4 patients, had time-of-flight capability, and the improved signal-to-noise ratio was used to lower the administered activity from 5 to 3 MBq/kg. Because each patient was scanned on the same system at baseline and for further PET evaluations, the measured response was supposed to be almost independent of noise resolution and region-of-interest method. Diagnostic-quality unenhanced CT images were acquired before PET data acquisition. The CT, PET, and coregistered PET/CT images were reviewed in transaxial, coronal, and sagittal planes along with maximum-intensity-projection whole-body images.

TABLE 1
Characteristics of the 59 Patients

Characteristic	Data
Median age at diagnosis (y)	35.5 (16–76)
Sex	
Male	40 (68)
Female	19 (32)
Histologic type	
Lymphocyte-rich	5 (9)
Mixed cellularity	7 (12)
Nodular sclerosis	45 (76)
Unclassified*	2 (3)
Ann Arbor stage	
I	5 (9)
II	17 (29)
III	10 (17)
IV	27 (46)
Bulky tumor (mass > 10 cm)	9 (15)
International prognosis score	
0–2	23 (39)
≥3	36 (61)

*Patients with only extranodal biopsy available were defined as having unclassified HL according to World Health Organization 2008 classification.

Data are *n* followed by percentage in parentheses, except for age.

PET2 Analysis and ΔSUVmax Assessment

All PET0 and PET2 images were reviewed independently by 2 nuclear medicine physicians who were masked to the patients' outcomes. In cases of disagreement, the ^{18}F-FDG PET/CT findings were discussed until agreement was reached.

In PET0 and PET2, SUVmax was assessed by drawing a region of interest around the most intense area of pathologic ^{18}F-FDG uptake. The SUVmax reduction was computed by expressing the percentage reduction between the SUVmax in the tumor site with the most intense uptake at PET0 and the SUVmax in the tumor site with the most intense uptake at PET2 (ΔSUVmax PET0–PET2) (11). In cases of complete remission, a region of interest was drawn around the previous site of most intense uptake identified at PET0.

The ΔSUVmax PET0–PET2 was calculated for all patients and compared with a visual analysis using the Deauville 5-point scale (8). Briefly, a score of 1 indicated no residual uptake above the background level, a score of 2 indicated residual uptake less than or equal to the mediastinum, a score of 3 indicated residual uptake greater than the mediastinum but not greater than the liver, a score of 4 indicated residual uptake moderately increased compared with the liver, and a score of 5 indicated residual uptake markedly increased compared with the liver or new sites of disease. A PET2 study was considered visually positive when the residual ^{18}F-FDG uptake was superior to the liver uptake (score 4 or 5).

Statistical Analysis

To analyze the prognostic influence of ΔSUVmax PET0–PET2, its values were dichotomized by applying the receiver-operating-characteristics approach (15), based on its ability to predict treatment failure with the best sensitivity and specificity.

To compare the predictability of the semiquantitative and visual early PET analyses on treatment outcome, we computed their respective negative and positive predictive values and their sensitivity and specificity.

PFS and time to progression (TTP) were analyzed according to PET results based on the 5-point scale and ΔSUVmax criteria. The PFS was

defined as the time from the beginning of treatment until progression, relapse, or death from any cause or the date of last follow-up. The TTP was defined as the time from the date of the first course of chemotherapy to any treatment failure, including progression, relapse, or death related to lymphoma, or the date of last follow-up. Patients who died from a cause other than lymphoma were censored at the time of death.

Survival of patient subgroups as defined by the ΔSUVmax PET0–PET2 and 5-point scale were estimated using the Kaplan–Meier product limit method and compared using the log-rank test.

To construct a model for the prediction of PFS and TTP, 2 Cox proportional hazards regression models were set up including the international prognosis score and ΔSUVmax PET0–PET2 as explanatory variables.

RESULTS

Semiquantitative PET2 Analysis

The median SUVmax was 11.5 at baseline (range, 2.3–30.6) and decreased to 1.65 after 2 cycles of chemotherapy (range, 0.6–18.9), leading to a median ΔSUVmax PET0–PET2 of 85% (range, 47%–95%). The performance of the receiver-operating-characteristics curve in determining the optimal cutoff of ΔSUVmax PET0–PET2 for identifying good and poor responders is presented in Figure 1. The area under the receiver-operating-characteristics curve was 0.717 ($P < 0.025$; 95% confidence interval, 0.584–0.826), and the best cutoff according to the Youden index was 71%, with a sensitivity of 54% (95% confidence interval, 25–81) and a specificity of 94% (95% confidence interval, 82–99).

Comparison of Semiquantitative and Visual Analysis of PET2

Overall agreement between the 2 readers measured with the κ statistics was 0.85 (very good) for both the visual analysis (score 1–3 vs. 4–5) and the semiquantitative analysis (ΔSUVmax $\leq 71\%$ vs. $> 71\%$). In addition, the mean absolute difference of ΔSUVmax between the observers was 2.6%, with both readers finding the same proportion of patients with a ΔSUVmax greater than 71%.

Forty-six patients (78%) achieved a negative PET2 study on the basis of the visual analysis, and 49 patients (83%) reached a ΔSUVmax PET0–PET2 greater than 71%.

Among the 46 patients with a visually negative PET2 study, 43 had a ΔSUVmax PET0–PET2 greater than 71%. The remaining 3 patients had a low baseline SUVmax (2.3, 4.1, and 5.3). Among the 13 patients with a positive PET2, 7 patients (54%) had a ΔSUVmax PET0–PET2 of 71% or less. Thus, 6 patients could be reclassified as good responders according to the semiquantitative analysis.

Overall, the predictive performance was better for semiquantitative analysis than for visual analysis, with a significantly better positive predictive value for ΔSUVmax PET0–PET2 analysis (70%) than for visual analysis (46%) (Table 2). This result leads to better specificity and accuracy for the semiquantitative method. Inversely, the negative predictive value was similar with the 2 interpretation criteria: 7 of the 46 patients (15%) with a visually negative PET2 study experienced treatment failure (either progressive disease or a relapse), leading to a negative predictive value of 85% for visual analysis, whereas 6 of the 49 patients (12%) who achieved a ΔSUVmax PET0–PET2 greater than 71% failed treatment (negative predictive value, 88%) (Tables 2 and 3).

Influence of PET2 Results on Patient Outcomes According to Semiquantitative and Visual Analysis

Patients with a visually positive PET2 study had a lower PFS and TTP than PET2-negative patients (4-y PFS: 45% vs. 81%, respectively, $P < 0.002$, hazard ratio = 4.3; 4-y TTP: 51% vs. 83%, respectively, $P < 0.006$, hazard ratio = 4.1) (Fig. 2). The semiquantitative
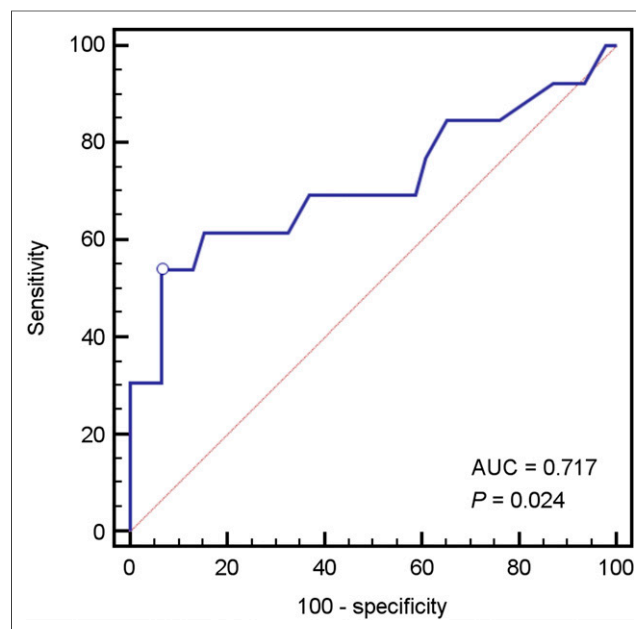

FIGURE 1. Receiver-operating-characteristics curve for determining ΔSUVmax PET0–PET2 cutoff value. Marked point corresponds to cutoff point with best Youden index. AUC = area under the curve.

approach allowed more accurate identification of patients with a high risk of treatment failure: patients who did not reach a ΔSUVmax PET0–PET2 greater than 71% had a significantly lower PFS and TTP than those who did (Fig. 3) (4-y PFS: 30% vs. 82%, respectively, $P < 0.0001$, hazard ratio = 6.55; 4-y TTP: 30% vs. 86%, respectively, $P < 0.0001$, hazard ratio = 8.51). Moreover, patients who reached a ΔSUVmax PET0–PET2 greater than 71% had a similar outcome whatever the visual PET2 result (4-y PFS: 82% and 83% for negative and positive PET2, respectively) (Table 3).

In multivariate analysis, ΔSUVmax PET0–PET2 remained the unique independent predictor for PFS ($P = 0.0001$; relative risk, 7.9; 95% confidence interval, 2.9–22.9) and TTP ($P = 0.0001$; relative risk, 9.1; 95% confidence interval, 3.4–31.5).

DISCUSSION

This single-center retrospective study showed that semiquantitative analysis of early PET response using SUVmax reduction is

TABLE 2
Outcome Prediction Using Semiquantitative or Visual Analysis for PET2 Interpretation*

Parameter	Semiquantitative analysis (ΔSUVmax $\leq 71\%$ vs. $> 71\%$)	Visual analysis (score 1, 2, 3 vs. 4, 5)
Sensitivity	54%	46%
Specificity	94%	84%
Negative predictive value	88%	85%
Positive predictive value	70%	46%
Accuracy	85%	76%

*59 patients; 10 progression or relapse events.

TABLE 3
Outcome of Subsets of Patients Defined by PET2 Results Combining Visual and Semiquantitative Analysis

PET2 Visual analysis	PET2 Semiquantitative analysis	n	Treatment failure (progression or relapse)	4-y PFS	4-y TTP
Score 1–3	ΔSUVmax > 71%	43	6 (14%)	82%	86%
Score 1–3	ΔSUVmax ≤ 71%	3	1 (33%)	67%	67%
Score 4–5	ΔSUVmax > 71%	6	0	83%	100%
Score 4–5	ΔSUVmax ≤ 71%	7	6 (86%)	14%	14%

more accurate than visual analysis based on the 5-point scale criteria to identify subsets of patients with significantly different outcomes.

The visual analysis produced an excess of positive results leading to a poor positive predictive value for treatment failure. With the ΔSUVmax analysis, 46% of PET2-positive patients had a ΔSUVmax over the cutoff value and favorable 4-y PFS and TTP estimates of 83% and 100%, respectively. Thus, these 6 patients with a positive PET2 study but a ΔSUVmax PET0–PET2 greater than 71% were indeed good responders since none relapsed during the current follow-up, and they were identified as good responders by ΔSUVmax analysis. All had a residual mass with relatively low ^{18}F-FDG uptake (median SUVmax of 2.89; range, 2.09–3.21), which remained superior to the liver uptake. The visual estimation of the PET2 liver uptake was checked by measuring the SUVmax of the liver and was in all these cases lower than the SUVmax of the residual mass, ruling out any visual misinterpretation. Therefore, these results suggest that the low positive predictive value of the visual interpretation in HL can be attributed to an excess of false-positive scans, as was previously shown in diffuse large B-cell lymphoma (11,12). These results also suggest that a complete metabolic response is not necessarily required after 2 cycles of chemotherapy. Conversely, a good reduction of ^{18}F-FDG uptake, which can be considered a marker of tumor chemosensitivity, could be a satisfactory endpoint for expecting a long-term failure-free outcome.

The positive predictive value of the visual analysis in the present study was lower (46%) than the 73% recently reported in a multicenter retrospective analysis by Biggi et al. of 260 patients using the same PET2 interpretation criteria (16). These discrepancies might be partly related to the selection of patients in this series, since only 260 of 440 patients were retained for analysis. Also, most patients had advanced-stage disease in that series whereas in our study 37% of enrolled patients had localized disease, for which early PET assessment was previously reported to have a lower positive predictive value (2). In addition, in the series of Biggi et al., there were discrepancies between the 6 reviewers in 18% of cases and the patients with discordant results were reclassified after a consensus meeting, indicating that the positive predictive value calculated for each individual reviewer would probably be significantly lower than 73%. Interestingly, the positive predictive value according to the local interpretations at most centers by the study investigators was about 45%.

False-positive results related to the visual PET evaluation could proceed from numerous causes. In HL, the neoplastic cells represent less than 1% of the total cell population of the tumor (17), and the ^{18}F-FDG tumor uptake observed is probably related more to the microenvironment surrounding the tumor cells than to the tumor cells themselves. In addition, the type of surrounding cells is heterogeneous and varies according to the pathologic subset of HL. Therefore, residual ^{18}F-FDG uptake may still be related mainly to

inflammatory cells. In this setting, we cannot exclude the possibility that the level of ^{18}F-FDG uptake used as the background reference, that is, the liver uptake, was too low to avoid an excess of positive scans. Thus, it was previously shown that a liver-based interpretation using 140% of the liver uptake as a cutoff was more suitable for interpreting interim PET in patients with diffuse large B-cell lymphoma (18). In the present study, the use of such a cutoff would have allowed elimination of most of the false-positive cases. Other processes can also stimulate ^{18}F-FDG uptake, such as infectious foci or bone marrow stimulation. However, in all these different situations, with the same tracer, semiquantitative analysis significantly reduces the risk of attributing a positive result to
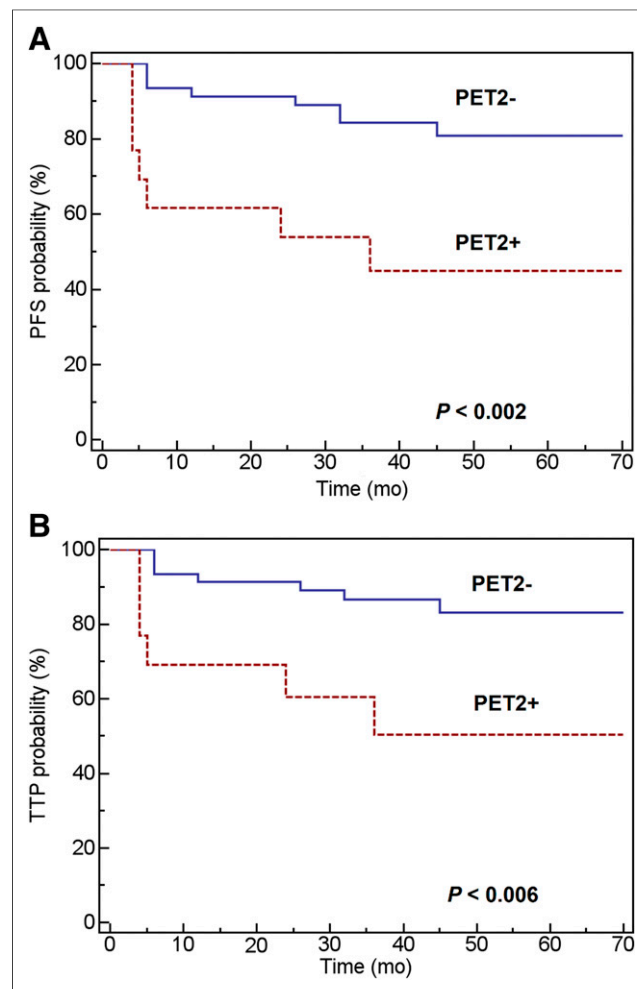

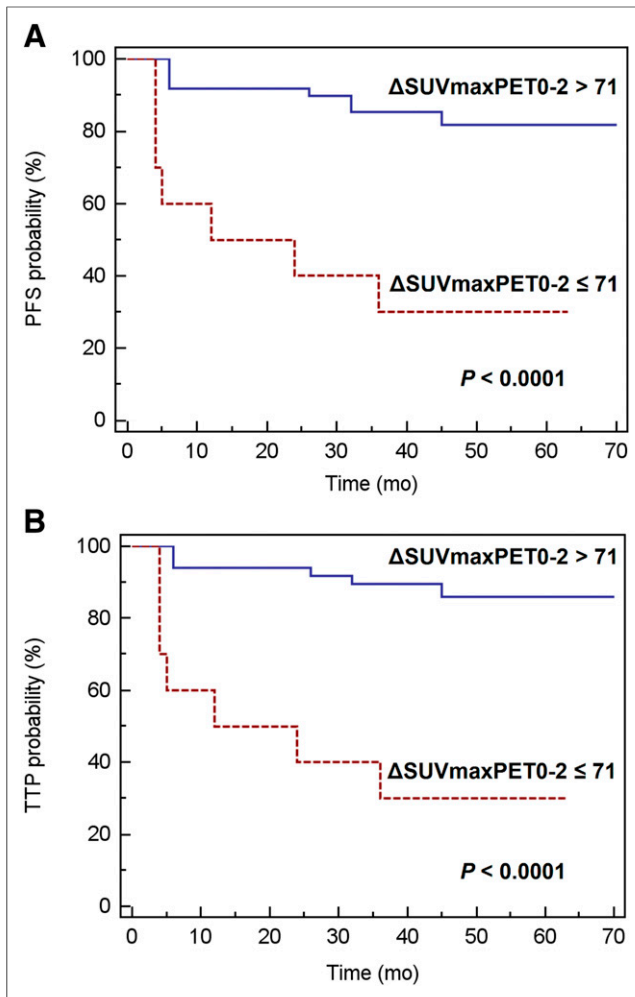**FIGURE 2.** PFS (A) and TTP (B) according to PET2 results on basis of 5-point scale analysis.

FIGURE 3. PFS (A) and TTP (B) according to PET2 results on basis of ΔSUVmax analysis.

residual lymphoma. Thus, ΔSUVmax calculation appears to be less subjective and helps distinguish which positive results may be related to significant residual lymphoma, leading to a better predictive value. To a lesser extent, ΔSUVmax analysis can also generate false-positive results. This occurred in 3 patients, when baseline SUVmax was low, leading to a ΔSUVmax lower than the defined cutoff value. These cases were easily identified since PET2 was negative according to visual analysis. Therefore, as previously suggested (*19*), in diffuse large B-cell lymphoma patients whose tumor exhibits a baseline SUVmax lower than 10 and for whom SUVmax reduction after 2 cycles of chemotherapy does not reach the defined cutoff, use of visual analysis can be recommended. In the present study, 7 patients had a baseline SUVmax less than 10 and a ΔSUVmax PET0–PET2 greater than 71%: 4 of these had a positive PET2 study according to the visual analysis.

With a median follow-up of 50 mo, semiquantitative analysis allowed identification of the poorest responders—patients who experienced induction failure or early relapse. Most progressive diseases were identified before the 10th month after the first-line treatment launch. Thus, the semiquantitative analysis seems to be a good method for early identification of patients who could be candidates for alternative therapeutic strategies.

CONCLUSION

These encouraging results suggest the use of semiquantitative analysis in addition to visual analysis to interpret early PET findings for HL patients, specifically for predicting with good confidence those patients who will have a poor outcome requiring alternative therapies. Larger and prospective series are warranted to confirm these preliminary results.

DISCLOSURE

The costs of publication of this article were defrayed in part by the payment of page charges. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734. No potential conflict of interest relevant to this article was reported.

REFERENCES

1. Hutchings M, Barrington SF. PET/CT for therapy response assessment in lymphoma. *J Nucl Med.* 2009;50(suppl):21S–30S.
2. 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.
3. 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.
4. Zijlstra JM, Boellaard R, Hoekstra OS. Interim positron emission tomography scan in multi-center studies: optimization of visual and quantitative assessments. *Leuk Lymphoma.* 2009;50:1748–1749.
5. Zinzani PL, Tani M, Fanti S, et al. Early positron emission tomography (PET) restaging: a predictive final response in Hodgkin's disease patients. *Ann Oncol.* 2006;17:1296–1300.
6. Gallamini A, Hutchings M, Rigacci L, et al. Early interim 2-[18F]fluoro-2-deoxy-D-glucose positron emission tomography is prognostically superior to international prognostic score in advanced-stage Hodgkin's lymphoma: a report from a joint Italian-Danish study. *J Clin Oncol.* 2007;25:3746–3752.
7. Terasawa T, Lau J, Bardet S, et al. Fluorine-18-fluorodeoxyglucose positron emission tomography for interim response assessment of advanced-stage Hodgkin's lymphoma and diffuse large B-cell lymphoma: a systematic review. *J Clin Oncol.* 2009;27:1906–1914.
8. Meignan M, Gallamini A, Meignan M, Gallamini A, Haioun C. Report on the first international workshop on interim-PET scan in lymphoma. *Leuk Lymphoma.* 2009;50:1257–1260.
9. 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.
10. 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 ΔSUVmax. *Eur J Nucl Med Mol Imaging.* 2013;40:1312–1320.
11. Lin C, Itti E, Haioun C, et al. Early 18F-FDG PET for prediction of prognosis in patients with diffuse large B-cell lymphoma: SUV-based assessment versus visual analysis. *J Nucl Med.* 2007;48:1626–1632.
12. Casasnovas R-O, Meignan M, Berriolo-Riedinger A, et al. SUVmax reduction improves early prognosis value of interim positron emission tomography scans in diffuse large B-cell lymphoma. *Blood.* 2011;118:37–43.
13. Swerdlow S, Campo E, Harris N. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues.* Lyon, France: International Agency for Research on Cancer; 2008.
14. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol.* 2007;25:579–586.
15. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem.* 1993;39:561–577.
16. Biggi A, Gallamini A, Chauvie S, et al. International validation study for interim PET in ABVD-treated, advanced-stage Hodgkin lymphoma: interpretation criteria and concordance rate among reviewers. *J Nucl Med.* 2013;54:683–690.
17. Canellos GP. Residual mass in lymphoma may not be residual disease. *J Clin Oncol.* 1988;6:931–933.
18. Itti E, Juweid ME, Haioun C, et al. Improvement of early 18F-FDG PET interpretation in diffuse large B-cell lymphoma: importance of the reference background. *J Nucl Med.* 2010;51:1857–1862.
19. Casasnovas R-O, Meignan M, Berriolo-Riedinger A, et al. Early interim PET scans in diffuse large B-cell lymphoma: can there be consensus about standardized reporting, and can PET scans guide therapy choices? *Curr Hematol Malig Rep.* 2012;7:193–199.