

SUVmax reduction improves early prognosis value of interim positron emission tomography scans in diffuse large B-cell lymphoma

René-Olivier Casasnovas,¹ Michel Meignan,² Alina Berriolo-Riedinger,³ Stéphane Bardet,⁴ Anne Julian,⁵ Catherine Thieblemont,⁶ Pierre Vera,⁷ Serge Bologna,⁸ Josette Brière,⁶ Jean-Philippe Jais,⁹ Corinne Haioun,² Bertrand Coiffier,¹⁰ and Franck Morschhauser,¹¹ on behalf of the Groupe d'étude des lymphomes de l'adulte (GELA)

¹Hopital Le Bocage, Dijon, France; ²Hopital Henri Mondor, Créteil, France; ³Centre Georges-Francois Leclerc, Dijon, France; ⁴Centre François Baclesse, Caen, France; ⁵Hopital Purpan, Toulouse, France; ⁶Hopital Saint-Louis, Paris, France; ⁷Centre Henri Becquerel, Rouen, France; ⁸CHU Nancy, Vandoeuvre les Nancy, France; ⁹Hopital Necker, Paris, France; ¹⁰Hopital Lyon Sud, Pierre Bénite, France; and ¹¹Hopital Claude Huriez and EA 4481 GRIIOT, Lille 2 University, Lille, France

The prognostic value of interim positron emission tomography (PET) interpreted according to visual criteria is a matter of debate in diffuse large B-cell lymphoma (DLBCL). Maximal standardized uptake value reduction (Δ SUVmax) may better predict outcome. To compare the prognostic value of both methods, we analyzed PET done at baseline (PET0) and after 2 (PET2) and 4 (PET4) cycles in 85 patients with high-risk DLBCL enrolled on a prospective multicenter trial. All images were centrally reviewed and inter-

preted visually according to the International Harmonization Project criteria and by computing Δ SUVmax between PET0 and PET2 (Δ SUVmaxPET0-2) or PET4 (Δ SUVmaxPET0-4). Optimal cutoff to predict progression or death was 66% for Δ SUVmaxPET0-2 and 70% for Δ SUVmaxPET0-4. Outcomes did not differ significantly whether PET2 and PET4 were visually positive or negative. Inversely, Δ SUVmaxPET0-2 analysis ($> 66\%$ vs $\leq 66\%$) identified patients with significantly different 2-year progression-free

survival (77% vs 57%; $P = .0282$) and overall survival (93% vs 60%; $P < .0001$). Δ SUVmaxPET0-4 analysis ($> 70\%$ vs $\leq 70\%$) seemed even more predictive for 2-year progression-free survival (83 vs 40%; $P < .0001$) and overall survival (94% vs 50%; $P < .0001$). Δ SUVmax analysis of sequential interim PET is feasible for high-risk DLBCL and better predicts outcome than visual analysis. The trial was registered at <http://clinicaltrials.gov> as NCT00498043. (*Blood*. 2011;118(1):37-43)

Introduction

Fluorine-18 fluorodeoxyglucose (¹⁸F-FDG) positron emission tomography (PET) was shown to improve both primary staging¹ and response assessment at completion of first-line therapy of diffuse large B-cell lymphoma (DLBCL),^{2,3} and it was implemented in the standardized response criteria for lymphoma.⁴ However, there is an increasing interest in using interim PET performed after 1-4 cycles of chemotherapy⁵ to predict response to induction treatment and to drive consolidation therapy.

The prognostic value of interim PET on the basis of visual analysis remains controversial in DLBCL. In a prospective study of 90 patients with DLBCL, a positive PET after 2 cycles (PET2) identified poor responders regardless of their treatment or age-adjusted international prognostic index (aaIPI) score.⁶ Similarly, PET after 4 cycles of induction treatment (PET4) was also shown to predict outcome.⁷ Conversely, in a recent report on 97 patients with DLBCL with aaIPI 1-3, progression-free survival (PFS) was similar for patients with either a positive or a negative PET4. Moreover, only 5 of the 38 patients with a positive PET4 had biopsy-proven active disease.⁸ These discrepancies on the predicting value of interim PET may either be because of the heterogeneity of the visual criteria used so far⁵ or reflect the lack of interobserver reproducibility in interpreting PET images on the basis of entirely visual criteria.⁹ Semiquantification of standardized uptake values (SUVs) may reduce false-positive interim PET interpretations, but whether SUV analysis better predicts outcome than visual analysis

has not been clearly established yet.^{10,11} Therefore, a comparison of interim PET results based on SUV analysis to results based on visual criteria could be helpful to establish their respective prognostic value.

In 2007, the Groupe d'étude des lymphomes de l'adulte started a prospective multicenter trial in previously untreated young patients with high-risk DLBCL. The intensity of consolidation was driven by a centralized assessment of both PET2 and PET4 with the use of the most recently published International Harmonization Project (IHP) visual criteria at that time. An exploratory analysis of SUVmax reduction (Δ SUVmax) between baseline PET and either PET2 or PET4 was performed simultaneously during the central review process. Here, we report the comparison between the results of the visual and Δ SUVmax analysis and demonstrate that the latter semiquantitative approach better predicts outcome than visual analysis.

Methods

LNH2007-3B study design

The LNH2007-3B trial was a prospective multicenter, randomized phase 2 trial of 2 induction regimens, R-CHOP14 (rituximab plus cyclophosphamide doxorubicin, vincristine, and prednisone every 14 days) versus R-ACVBP (rituximab plus doxorubicin, cyclophosphamide, vindesine, bleomycin, and prednisone), followed by a PET-driven consolidation

Submitted December 25, 2010; accepted April 10, 2011. Prepublished online as *Blood* First Edition paper, April 25, 2011; DOI 10.1182/blood-2010-12-327767.

Presented part of results in oral session at the 52nd annual meeting of the American Society of Hematology, Orlando, FL, December 4-7, 2010.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2011 by The American Society of Hematology

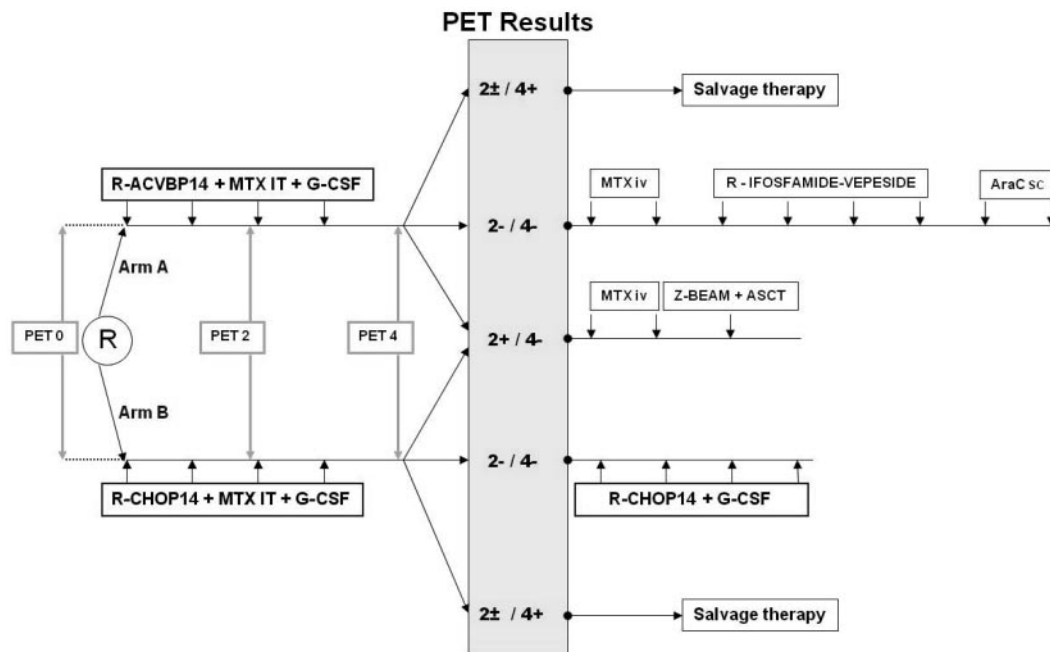


Figure 1. Scheme of the treatment according to the LNH2007-3B trial. MTX indicates methotrexate; IT, intrathecally; IV, intravenously; AraC, cytosine arabinoside; sc, subcutaneously; and Z-BEAM, carmustine, etoposide, cytarabine, and melphalan with Zevalin (ibrutinomab tiuxetan).

treatment in previously untreated young patients with high-risk DLBCL. The primary endpoint was the complete response rate according to the revised International Working Group criteria⁴ after 4 cycles of induction. To detect a complete response rate > 50% after 4 cycles of R-ACVBP or R-CHOP14, we calculated that a sample size of 101 assessable patients in each randomization arm would provide 85% power at an overall 2.5% (1-sided) significance level. The overall sample size was brought up to 222 patients, including 111 patients in each arm, to allow for a 10% drop-out rate. An interim analysis was planned after the inclusion of 52 assessable patients in each induction arm. The secondary endpoints included toxicity, overall survival (OS), and PFS. This study was approved by the ethics committee of Lyon and the national regulatory agency according to French regulatory laws. All patients provided written informed consent in accordance with the Declaration of Helsinki. The study was registered as NCT 00498043 at www.clinicaltrials.gov.

Eligibility criteria

Patients eligible for the present study were 18-59 years old with a previously untreated histologically proven CD20⁺ DLBCL and an aaIPI score of 2 or 3. A baseline PET scan (PET0) was mandatory with ≥ 1 evaluable hypermetabolic lesion. All patients had to be eligible for high-dose therapy followed by autologous stem cell transplantation (ASCT). Patients with known positive HIV status, active viral hepatitis B and C, or CNS involvement by lymphoma were excluded.

Treatment

Patients were randomly assigned to receive as induction treatment 4 cycles of either R-ACVBP14 (rituximab [375 mg/m²], cyclophosphamide [1200 mg/m²], and doxorubicin [75 mg/m²], given intravenously on day 1; vindesine [2 mg/m²] and bleomycin [10 mg] given intravenously on days 1 and 5; prednisone [60 mg/m²] given orally on days 1 through 5; and intrathecal methotrexate [15 mg] on day 2, recycling at day 14) or R-CHOP14 (rituximab [375 mg/m²], cyclophosphamide [750 mg/m²], doxorubicin [50 mg/m²], and vincristine [1.4 mg/m²] given intravenously on day 1; prednisone [60 mg/m²] given orally on days 1 through 5; and intrathecal methotrexate [15 mg] on day 2; recycling at day 14). The consolidation treatment was driven by centrally reviewed PET assessment after 2 and 4 cycles of induction immunochemotherapy interpreted according to visual criteria (Figure 1). Patients who were classified as PET2

and PET4 negative received consolidation with sequential conventional dose immunochemotherapy consisting of the following: in the R-ACVBP arm, 2 cycles of high-dose methotrexate (3 g/m²), then 4 cycles of rituximab (375 mg/m²), ifosfamide (1.5 g/m²), etoposide (300 mg/m²) given intravenously on day 1 and 2 cycles of cytarabine (100 mg/m² subcutaneous on days 1 through 4); in the R-CHOP14 arm, 4 additional cycles of R-CHOP14. Patients classified as PET2 positive and PET4 negative received 2 cycles of high-dose methotrexate (3 g/m²) and then a consolidative high-dose therapy (carmustine, etoposide, cytarabine, and melphalan with or without Zevalin [ibrutinomab tiuxetan]) followed by ASCT. For these PET2-positive patients, peripheral stem cell harvest was performed with G-CSF after the third cycle of induction treatment. Patients classified as PET4 positive were removed from the study no matter what the results of PET2 were, and they were treated at the discretion of the investigator. A biopsy of the residual hypermetabolic mass was recommended whenever possible.

PET restaging

Two PET examinations at mid (PET2) and end (PET4) of induction were required for full assessment and scheduled 2 weeks after the second and the fourth cycle of immunochemotherapy, respectively. G-CSF was stopped 48 hours before PET. Each patient was scanned on the same camera for baseline and subsequent PET scans. A whole-body acquisition was started 60 \pm 10 minutes after a 5-MBq/Kg injection of ¹⁸F-FDG, working from groin up to the head. The administered activity of FDG, the time of injection, and the time of the scan beginning were recorded.

Visual analysis of PET

A blinded central review in real time of the PET images was organized with the positroscope network.¹² For each patient, the data and images of the PET0, PET2, and PET4 were sent within 24 hours of the examination to ≥ 2 of 3 PET experts (M.M., A.B.-R., or S. Bardet) composing the central panel. PET2 and PET4 were binary interpreted as positive or negative. Interpretation criteria used the rules proposed by the IHP in Lymphoma³ with the following precisions: the "clearly increased activity relative to the reference background" which defines positive residual uptake in the IHP criteria should be $\geq 25\%$ higher than this background. A first central review was performed within 72 hours of receiving PET2 images, and the final result was sent back to the investigator to allow planning of stem cell harvest after cycle 3 in case PET2 was positive. A second central review

was done within 72 hours of receiving PET4 images, and the final result was sent back to the investigator together with the per-protocol recommended consolidation treatment allocation. In addition, a central exploratory analysis of PET2 and PET4 with the use of the Deauville criteria¹³ was done post hoc on all study patients to see whether they would better predict outcome than IHP criteria.

SUV-based assessment of ¹⁸F-FDG uptake

An analysis of the Δ SUV_{max} between baseline PET and PET2 (Δ SUV_{max}PET0-2) or PET4 (Δ SUV_{max}PET0-4) was performed during the central review process, with no influence on the consolidation treatment allocation. For each PET, the tumor with the most intense ¹⁸F-FDG uptake was identified among all foci with the use of a graded color scale. The hottest volumetric region was determined, and the SUV_{max} was calculated as previously described.¹⁰ To assess the Δ SUV_{max}, the hottest tumor in any region or organ on PET2 or PET4 was used for comparison, even if its location differed from the initial hottest tumor in PET0.

Statistics

The level of agreement on PET visual interpretation between the on-site and the review panel was analyzed with nonweighted κ statistics.¹⁴

Receiver-operating-characteristics (ROC) analysis¹⁵ was used to determine an optimal cutoff for Δ SUV_{max}PET0-2 and Δ SUV_{max}PET0-4 in predicting disease progression or death. For Δ SUV_{max}PET0-2, ROC analysis identified that the 2 cutoffs of 62% and 66%, respectively, had the best sensitivity and specificity to predict an event occurrence. Because a 66% cutoff had been previously identified on prior independent series,^{10,16} this threshold was chosen to analyze our series. For Δ SUV_{max}PET0-4 the cutoff identified by ROC curve was 70%.

PFS was defined as the time from randomization to first progression, relapse, and either death, whatever the cause, or last follow-up. OS was defined as the time from randomization to death from any cause, or last follow-up. Estimates of survival were calculated according to the Kaplan-Meier method and compared with the log-rank test.

Differences between the results of comparative tests were considered significant if the 2-sided *P* value was < .05. All statistical analyses were performed with the Statistical Application System software (SAS Version 9.1.3; SAS Institute).

Results

Patients of the planned interim analysis evaluable for PET analysis

One hundred thirteen patients were enrolled in 45 centers and were randomly assigned between October 2007 and April 2009. Their characteristics are detailed in Table 1. Two patients were removed from the study before treatment because of the investigator's decision (*n* = 1) and the patient's withdrawal of consent, 2 patients were prematurely withdrawn because of treatment toxicity before any PET restaging, and 7 patients completed induction treatment without performing PET2 (*n* = 2), PET4 (*n* = 1), or both (*n* = 4), leaving 102 patients (90%) evaluable for PET analysis. With a median follow-up of 19 months (range, 3-28 months), 20 patients progressed of whom 11 died of disease progression and 2 additional patients died without progression.

PET2 analysis

Among the 102 patients who underwent PET2 examination, PET2 was negative in 40 (39%) and positive in 62 (61%) patients according to the on-site interpretation, whereas central review concluded to 35 (34%) negative and 67 (67%) positive cases. So central review was in agreement with the on-site interpretation in

89% of cases (Table 2), leading to a κ coefficient of 0.769 (95% CI, 0.64-0.898). PET was assessed by both visual and Δ SUV_{max} analysis for 85 of these 102 patients (Table 1). For the 17 remaining patients, Δ SUV_{max} calculation was impossible because of the lack of images with attenuation correction in 2 cases or some PET technical data in 15 cases, including errors in the recorded ¹⁸F-FDG-injected activity. After cycle 2, the median SUV_{max} was 3.1 (range, 1.3-16.2) corresponding to a median Δ SUV_{max}PET0-2 of 81.4% (range, 21.3%-96.5%). Seventy patients (82%) of whom 25 were PET2 negative displayed a Δ SUV_{max}PET0-2 > 66% (Table 3). Forty-five (78%) of the 58 PET2-positive patients achieved a Δ SUV_{max}PET0-2 > 66%. Inversely, only 2 of the 15 patients (18%) with Δ SUV_{max}PET0-2 \leq 66% had negative PET2 according to visual criteria. These 2 patients had low baseline SUV_{max} of, respectively, 3.8 and 4.9.

PET4 analysis

At the end of induction treatment, 98 patients (96%) underwent PET4 examination. PET4 was negative according to on-site and review board interpretation for, respectively, 56 (57%) and 50 (51%) patients (Table 2). Review was in agreement with on-site conclusions in 92% of cases, leading to a κ coefficient of 0.836 (95% CI, 0.728-0.945).

After cycle 4, the median SUV_{max} was 2.5 (range, 1-19) for the 84 assessable patients corresponding to a median Δ SUV_{max}PET0-4 of 85.7% (range, 20%-97.1%). Seventy-four patients (88%), including 41 patients (55%) who had a negative PET4, achieved a Δ SUV_{max}PET0-4 > 70% (Table 3). Thirty-three (80%) of the 41 PET4 positive patients showed a Δ SUV_{max}PET0-4 > 70%. Again, among the 10 patients (12%) with a Δ SUV_{max}PET0-4 \leq 70%, PET4 was considered negative in the 2 cases with low baseline SUV_{max} (3.8 and 4.9).

Ten patients who had a positive PET4 underwent a biopsy of the residual hypermetabolic mass. In 2 cases the biopsy showed an active lymphoma disease and was associated with a Δ SUV_{max}PET0-4 \leq 70%. The remaining 8 patients, all with a Δ SUV_{max}PET0-4 > 70%, had no evidence of lymphoma.

PFS and OS according to PET2 results

PET2 results assessed by visual analysis according to IHP criteria had no influence on PFS (Figure 2A) and OS (*P* = .5861 and *P* = .336, respectively): the 2-year estimates for PFS and OS were 73% and 93%, respectively, for patients with a negative PET2 compared with 77% and 84% for patients achieving a positive PET2. Similar results were observed with Deauville criteria: the 2-year PFS estimate for patients with a PET2 residual mass showing a FDG uptake higher than the liver was 79% compared with 88% (*P* = .825) for patients with a lower uptake. Conversely, Δ SUV_{max}PET0-2 identified 2 groups of patients with significantly different PFS (*P* = .0282; Figure 2B) and OS (*P* < .0001): the 2-year estimates for PFS and OS were 57% and 60%, respectively, for patients with a Δ SUV_{max}PET0-2 \leq 66%, compared with 77% and 93% for patients achieving a Δ SUV_{max}PET0-2 > 66%. Patients who remained PET2 positive and had a Δ SUV_{max}PET0-2 \leq 66% had a significantly poorer PFS (*P* = .014; Figure 2C) and OS (*P* < .0001; data not shown) than patients having a negative PET2 or achieving a Δ SUV_{max}PET0-2 > 66%.

PFS and OS according to PET4 results

Patients with visual positive PET4 according to IHP criteria had a trend to poorer outcome than patients achieving a negative

Table 1. Patient characteristics

| | All (n = 113) | Patients evaluable for PET analysis with visual criteria (n = 102) | Patients evaluable for PET analysis with visual and Δ SUVmax criteria (n = 85) |
|---|---------------|--|---|
| Age, median (range), y | 46 (18-59) | 45 (18-59) | 46 (19-59) |
| Sex, ratio of male to female | 1.4 | 1.3 | 1.1 |
| Ann Arbor stage | | | |
| I-II, n (%) | 4 (4) | 4 (4) | 3 (4) |
| III-IV, n (%) | 109 (96) | 98 (96) | 82 (96) |
| LDH | | | |
| \leq Normal, n (%) | 8 (7) | 7 (7) | 7 (8) |
| > Normal, n (%) | 105 (93) | 95 (93) | 78 (92) |
| No. of extranodal sites | | | |
| \leq 1, n (%) | 37 (33) | 28 (34) | 28 (33) |
| > 1, n (%) | 76 (67) | 57 (68) | 57 (67) |
| ECOG | | | |
| < 2, n (%) | 85 (76) | 78 (77) | 64 (76) |
| \geq 2, n (%) | 27 (24) | 23 (23) | 20 (24) |
| Bulky disease (> 10 cm) | | | |
| Yes, n (%) | 22 (19) | 21 (21) | 15 (18) |
| No, n (%) | 91 (81) | 81 (79) | 70 (82) |
| Bone marrow involvement | | | |
| Involved, n (%) | 25 (22) | 22 (22) | 21 (25) |
| Not involved, n (%) | 83 (73) | 76 (75) | 60 (71) |
| Unspecified, n (%) | 5 (5) | 4 (4) | 4 (5) |
| Age-adjusted IPI | | | |
| 1, n (%) | 5 (4) | 5 (5) | 4 (5) |
| 2, n (%) | 87 (78) | 79 (78) | 66 (79) |
| 3, n (%) | 20 (18) | 17 (17) | 14 (17) |
| Subtotal 0-1, n (%) | 5 (4) | 5 (5) | 4 (5) |
| Subtotal 2-3, n (%) | 107 (96) | 96 (95) | 80 (95) |
| Lymphoma subtype (pathology review), n (%) | | | |
| Diffuse large B-cell lymphoma (%) | 75 | 74 | 73 |
| T cell-rich B-cell lymphoma (%) | 4 | 5 | 5 |
| Primary mediastinal B-cell lymphoma (%) | 16 | 16 | 16 |
| Unclassifiable (%) | 2 | 3 | 3 |
| Follicular lymphoma (%) | 1 | 1 | 2 |
| Chemotherapy regimen | | | |
| R-ACVBP, n (%) | 58 (51) | 50 (49) | 46 (54) |
| R-CHOP14, n (%) | 55 (49) | 52 (51) | 39 (46) |

LDH indicates lactate dehydrogenase; and ECOG, Eastern Cooperative Oncology Group.

PET4 both in term of PFS ($P = .0615$) (Figure 3A) or OS ($P = .054$): the 2-year estimates for PFS and OS were 73% and 83% for PET4-positive patients and 81% and 94% for PET4-negative patients. With the use of the Deauville criteria, the trend was similar with a 69% 2-year PFS estimate for patients with a PET4 residual mass showing a FDG uptake higher than the liver compared with 82% ($P = .065$) for patients with a lower uptake. Δ SUVmaxPET0-4 (> 70% vs \leq 70%) was more accurate to identify patients with significantly different 2-year

PFS (83% vs 40%) or OS (94% vs 50%) ($P < .0001$ for both): the median PFS and OS were 5 and 13 months, respectively, for patients with Δ SUVmaxPET0-4 \leq 70% and were not reached for patients achieving a Δ SUVmaxPET0-4 > 70% (Figure 3B). Six of the 8 patients who remained PET4 positive with a Δ SUVmaxPET0-4 \leq 70% relapsed within 8 months of diagnosis of whom 5 died of progression, whereas patients with a Δ SUVmaxPET0-4 > 70% or a negative PET had a 2-year PFS of > 90% (Figure 3C).

Table 2. Agreement between on-site and review board conclusions for PET2 and PET4 visual assessment

| On-site PET | PET review | | All, n (%) |
|-------------|-----------------|-----------------|------------|
| | Negative, n (%) | Positive, n (%) | |
| PET2 | | | |
| Negative | 32 (31) | 8 (8) | 40 (39) |
| Positive | 3 (3) | 59 (58) | 62 (61) |
| All | 35 (34) | 67 (66) | 102 |
| PET4 | | | |
| Negative | 49 (50) | 7 (7) | 56 (57) |
| Positive | 1 (1) | 41 (42) | 42 (43) |
| All | 50 (51) | 48 (58) | 98 |

Table 3. PET results according to visual and SUVmax reduction assessment

| PET visual analysis (central review result) | Δ SUVmaxPET0-2 | | | Δ SUVmaxPET0-4 | | |
|---|-----------------------|------------|---------|-----------------------|------------|---------|
| | > 66% | \leq 66% | All | > 70% | \leq 70% | All |
| PET2 | | | | | | |
| Negative, n (%) | 25 (29) | 2 (2) | 27 (32) | — | — | — |
| Positive, n (%) | 45 (53) | 13 (15) | 58 (68) | — | — | — |
| All, n (%) | 70 (82) | 15 (17) | 85 | — | — | — |
| PET4 | | | | | | |
| Negative, n (%) | — | — | — | 41 (49) | 2 (2) | 43 (51) |
| Positive, n (%) | — | — | — | 33 (39) | 8 (10) | 41 (49) |
| All, n (%) | — | — | — | 74 (88) | 10 (12) | 84 |

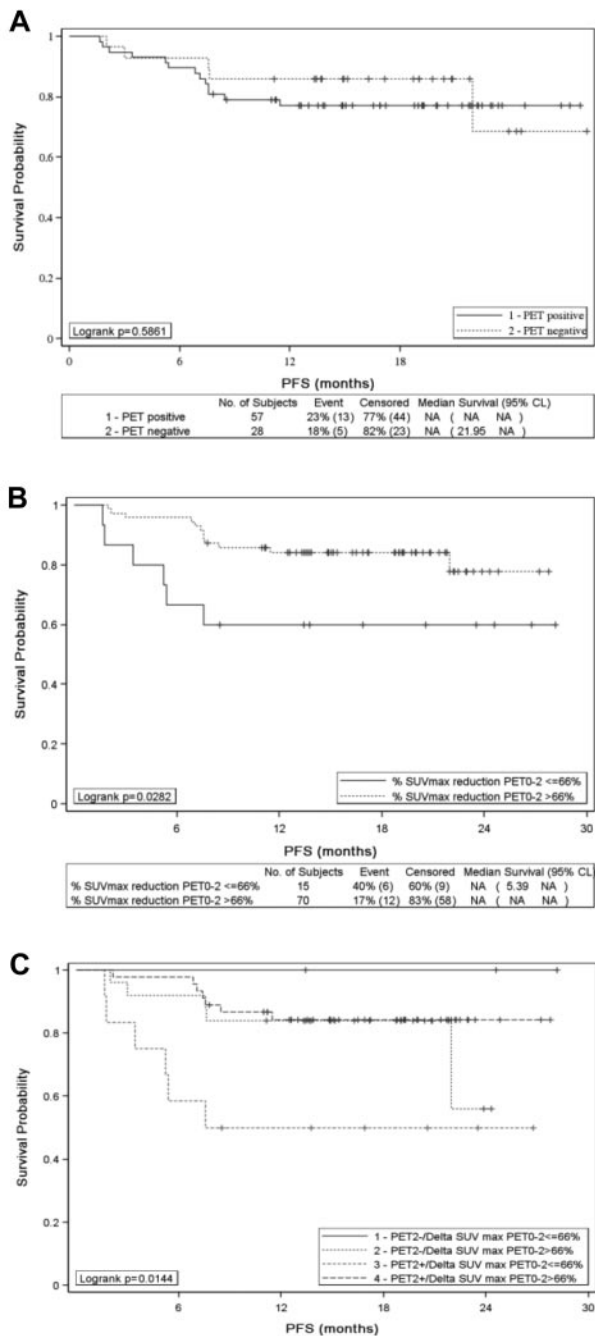


Figure 2. PFS of 85 patients. PFS according to (A) PET2 visual assessment (positive vs negative), (B) SUVmax reduction between PET0 and PET2 (> 66% vs ≤ 66%), and (C) PET2 analysis combining visual and quantitative assessment.

Effect of postinduction therapy on PFS according to PET results

With the use of IHP criteria, 2-year PFS was similar in the groups of patients who received ASCT or sequential consolidation but significantly worse for PET4-positive patients given salvage therapy ($P = .0065$; Figure 4A). With the use of Δ SUVmax, patients with Δ SUVmaxPET0-2 ≤ 66% given consolidative ASCT and patients with Δ SUVmaxPET0-4 ≤ 70% given salvage therapy had a significantly worse PFS than patients with Δ SUVmaxPET0-2 > 66% (53% vs 100%; $P = .0164$) and patients with Δ SUVmaxPET0-4 > 70% (0% vs 83%; $P < .0001$) given the same post induction therapy, respectively (Figure 4B).

Discussion

The present analysis on 85 patients enrolled in a prospective multicenter trial with central PET assessment shows that semiquantitative analysis with Δ SUVmax after 2 and 4 courses of induction treatment better predicts PFS and OS than visual analysis that is based on IHP criteria. Visual analysis produced an excess of positive results for PET2 and PET4, leading to a poor predictive value for PFS and OS. With Δ SUVmax analysis, 78% of PET2-positive and 80% of PET4-positive patients had a Δ SUVmax over

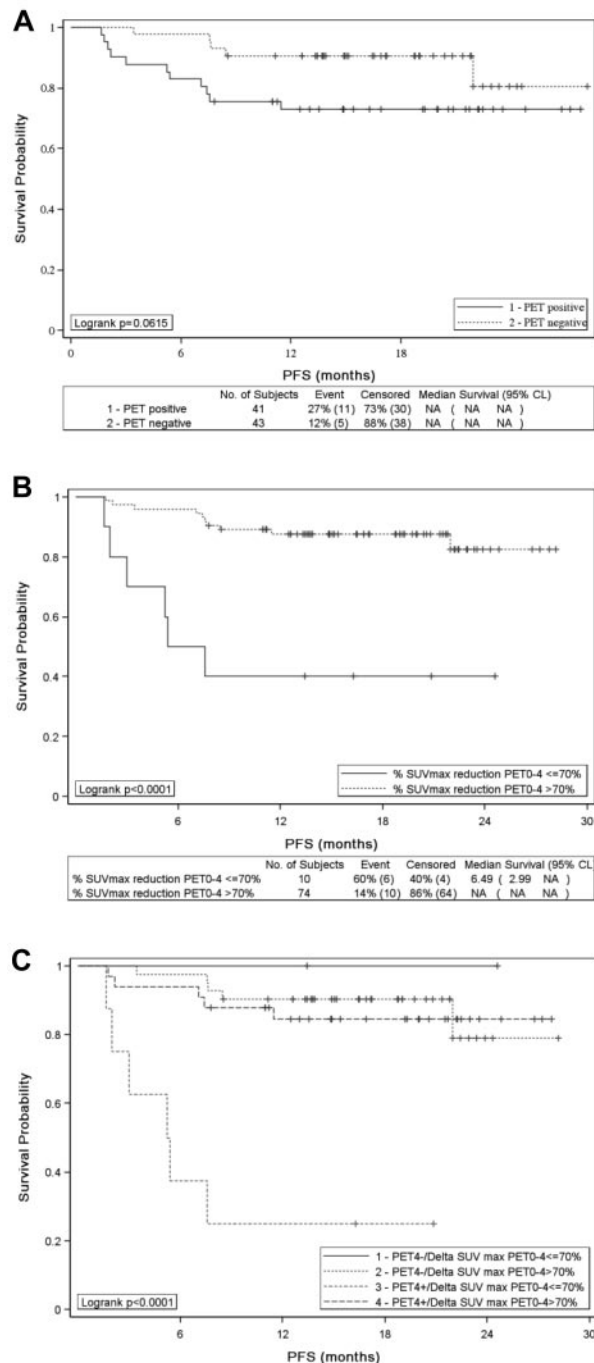


Figure 3. PFS of 84 patients. PFS according to (A) PET4 visual assessment (positive vs negative), (B) SUVmax reduction between PET0 and PET4 (> 70% vs ≤ 70%), and (C) PET4 analysis combining visual and quantitative assessment.

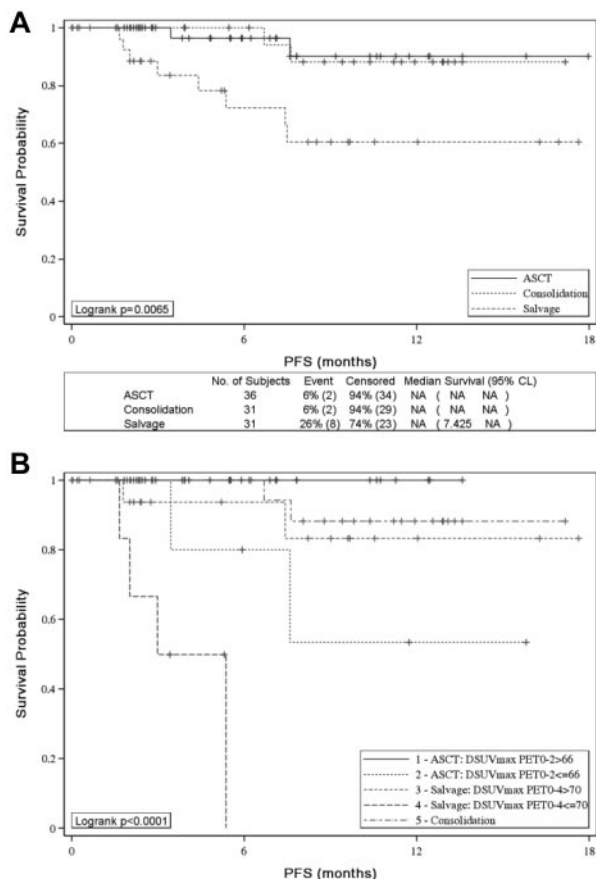


Figure 4. Effect of postinduction treatment on patients' PFS. Effect of postinduction treatment on patients' PFS according to (A) PET-driven postinduction treatment and (B) PET-driven postinduction treatment and SUVmax reduction analysis.

the cutoff value and a favorable 2-year PFS estimate of 77% and 83%, respectively. Thus, these patients classified as poor responders to immunochemotherapy according to visual analysis were indeed good responders and identified as such by Δ SUVmax analysis. Interestingly, the 80% of PET4-positive patients reclassified as good responders with the use of Δ SUVmaxPET0-4 in our series is consistent with the 87% patients with false-positive PET4 that was based on visual analysis in the study by Moskowitz et al.⁸

The Δ SUVmax cutoff values estimated by ROC analysis and used to distinguish good and bad responders were similar in our series to those previously reported in independent cohorts after either 2 or 4 cycles of induction treatment.^{10,11,16} Thus, these thresholds appear to be robust and reproducible regardless of age and IPI in patients with DLBCL treated with either CHOP or CHOP-like regimen combined with or without rituximab.

The disappointing positive predictive value of early PET with the use of modified IHP criteria was not related to discrepancies between readers because the reproducibility between on-site and centralized PET interpretation was quite satisfactory, with, respectively, a good and a very good agreement for PET2 and PET4, according to κ statistic. The agreement between readers appears to be much better in our study than the one observed by Horning et al.⁹ Moreover, in our series the few discrepancies between on-site and expert readers were overcome with the use of a real-time PET review process. Another hypothesis would be that postinduction may have affected outcome, especially high-dose therapy that may have improved outcome of PET2-positive patients, thereby erasing the predictive value of visual PET. However, patients who received

ASCT or sequential consolidation had similar PFS. Moreover, PET2-positive patients who received high-dose consolidative therapy and PET4-positive patients who received salvage therapy still could be split into good and poor prognostic subsets with the use of Δ SUVmax (Figure 4B).

False-positive results that were based on visual PET assessment could proceed from numerous other reasons. First, FDG uptake is not specific for lymphoma cells and can also be observed as well in inflammatory as well as infectious processes or after bone marrow stimulation. However, with the same tracer, semiquantitative analysis reduces dramatically the risk of attributing a positive result to residual lymphoma. In our series, the 8 PET4-positive patients who achieved a Δ SUVmaxPET0-4 > 70% and underwent a biopsy of the residual hypermetabolic mass had no evidence of lymphoma. In addition, the mediastinal blood pool area or the nearby background might not be the optimal reference background to visually compare the residual uptake in early PET. The liver could be a better reference background and was shown to generate less false-positive PET2 results.^{16,17} However, the Deauville criteria applied to our series did not significantly improve the accuracy of PET to identify subgroups of patients with different outcomes. In fact, visual assessment may lead to inaccurate interpretations regardless of the background tissue used, specifically when the unique minimal residual FDG uptake on restaging PET is close to that of the reference tissue and also in case of residual tumor with a size \sim 2 cm. In all these different situations, Δ SUVmax calculation is less subjective and helps distinguish which positive results may be related to significant residual lymphoma and affects outcome.

To a lesser extent Δ SUVmax analysis can also generate false-positive results. This occurred in 2 patients, when baseline SUVmax was low, leading to a Δ SUVmax lower than the defined cutoff value. Both cases were easily identified because PET2 and PET4 were negative according to visual analysis. The main drawback of the Δ SUVmax analysis is related to the absolute requirement of a baseline PET to allow a Δ SUVmax calculation. This could be a concern in high-risk patients with DLBCL who need a pressing treatment, specifically in a multicenter trial setting. In this prospective trial, cooperation between the hematologists and the nuclear medicine physicians was good, and the PET0 requirement did not bias recruitment because even patients with clinical features requiring urgent treatment such as bulk (18%) or poor performance status (24%) were enrolled (Table 1). In a multicenter trial setting, a last restriction to perform a quantitative PET assessment remained the quality of technical data transmitted to allow the SUVmax calculation; specifically, the weight of the patient at time of PET examination and the injected activity of ¹⁸F-FDG are critical. In this study some of these data were lost during either the data anonymization process or the data loading from on-site to review panel computers because of software bugs.

With a median follow-up of 19 months, Δ SUVmaxPET0-4 analysis allowed us to pick out the worst group of patients who experienced induction failure or early relapse. Most progressive diseases were identified before the sixth month after randomization, suggesting a weak effect of consolidative high-dose therapy and conventional salvage strategies in these poor-risk patients, as reported in the CORAL study.¹⁸ Thus, Δ SUVmaxPET0-4 seems to be a good way to identify patients who could be candidates for alternative experimental strategy after induction treatment. Conversely, longer follow-up is needed to conclude on the value of Δ SUVmaxPET0-2 in the context of the risk-adapted consolidative therapy. In addition, the value of analyzing sequential interim PET

and the effect on outcome of the kinetic of response remains to be examined and will be presented at the final analysis of the LNH2007-3B trial, in the whole population and in each induction treatment arm.

In conclusion, these encouraging results suggest the use of Δ SUVmax in addition to visual analysis to interpret interim PET for patients with DLBCL, specifically when a therapeutic decision is to be guided by interim PET results. Longer follow-up and analysis of the whole trial population are warranted to confirm the role of sequential interim PET in the context of the risk-adapted consolidation treatment.

Acknowledgments

We thank GELARC (Groupe d'Etude des Lymphomes de l'Adulte – Recherche Clinique) for the clinical management of this work.

This work was supported by the French government (grant PHRC 2008).

References

- Stumpe KD, Urbinelli M, Steinert HC, et al. Whole-body positron emission tomography using fluorodeoxyglucose for staging of lymphoma: effectiveness and comparison with computed tomography. *Eur J Nucl Med*. 1998;25(7):721-728.
- Spaepen K, Stroobants S, Dupont P, et al. Prognostic value of positron emission tomography (PET) with fluorine-18 fluorodeoxyglucose (18 F)FDG) after first-line chemotherapy in non-Hodgkin's lymphoma: is 18 F)FDG-PET a valid alternative to conventional diagnostic methods? *J Clin Oncol*. 2001;19(2):414-419.
- 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(5):571-578.
- Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol*. 2007;25(5):579-586.
- 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(11):1906-1914.
- 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(4):1376-1381.
- Dupuis J, Itti E, Rahmouni A, Hemery F, et al. Response assessment after an inductive CHOP or CHOP-like regimen with or without rituximab in 103 patients with diffuse large B-cell lymphoma: integrating 18 F)fluorodeoxyglucose positron emission tomography to the International Workshop Criteria. *Ann Oncol*. 2009;20(3):503-507.
- Moskowitz CH, Schöder H, Teruya-Feldstein J, et al. Risk-adapted dose-dense immunotherapy determined by interim FDG-PET in advanced-stage diffuse large B-Cell lymphoma. *J Clin Oncol*. 2010(11);28:1896-1903.
- 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(4):775-777.
- 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(10):1626-1632.
- 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(4):527-533.
- Meignan M, Itti E, Bardet S, et al. Development and application of a real-time on-line blinded independent central review of interim PET scans to determine treatment allocation in lymphoma trials. *J Clin Oncol*. 2009;27(16):2739-2741.
- Meignan M, Gallamini A, Haioun C. Report on the First International Workshop on Interim-PET-Scan in Lymphoma. *Leuk Lymphoma*. 2009;50(12):1257-1260.
- Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37-46.
- Zweig MH, Campbell G. Receiver-operating characteristics (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem*. 1993;39(4):561-577.
- Casasnovas RO, Saverot AL, Berriolo Riedinger A, et al. The 18 F-FDG SUVmax reduction after two cycles of R-CHOP regimen predicts progression free survival of patients with diffuse large B-cell lymphoma [abstract]. *Blood*. 2009;114(22):Abstract 2931.
- 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(12):1857-1862.
- Gisselbrecht C, Glass B, Mounier N, et al. Salvage regimens with autologous transplantation for relapsed large B-cell lymphoma in the rituximab era. *J Clin Oncol*. 2010;28(27):4184-4190.

Authorship

Contribution: R.-O.C. and F.M. were responsible for the conduct of the study and drafted the report that all co-authors critically revised for significant scientific content; R.-O.C., M.M., A.B.-R., S. Bardet, A.J., C.T., P.V., S. Bologna, J.B., J.-P.J., C.H., B.C., and F.M. contributed research data to the study; and R.-O.C., M.M., A.B.-R., J.-P.J., C.H., and F.M. contributed to data analysis and interpretation.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

A complete list of the members of the Groupe d'étude des lymphomes de l'adulte can be found in the supplemental Appendix (available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article).

Correspondence: René-Olivier Casasnovas, Hématologie Clinique, Hôpital Le Bocage – CHU Dijon, 10 Bd de Latre de Tassigny, 21079 Dijon Cedex, France; e-mail: olivier.casasnovas@chu-dijon.fr.