Impact of [$^{18}$F]Fluorodeoxyglucose Positron Emission Tomography Response Evaluation in Patients With High–Tumor Burden Follicular Lymphoma Treated With Immunochemotherapy: A Prospective Study From the Groupe d’Études des Lymphomes de l’Adulte and GOELAMS


See accompanying editorial on page 4285

ABSTRACT

Purpose

[$^{18}$F]Fluorodeoxyglucose positron emission tomography (PET) is widely used for the staging and restaging of patients with aggressive lymphoma, but less is known about the utility of PET in patients with follicular lymphoma (FL). In a prospective study, we evaluated the prognostic value of PET performed during treatment and at the end of treatment in 121 patients with FL treated with first-line immunochemotherapy.

Patients and Methods

Patients with previously untreated high–tumor burden FL were treated with six cycles of R-CHOP (rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone) plus two cycles of rituximab, without rituximab maintenance. PET was performed before treatment, after four cycles of R-CHOP (interim PET), and at the end of treatment (final PET). PET scans were centrally reviewed.

Results

The total number of patients included was 121. Median age was 57 years. After central review, interim PET (n = 111) was negative in 76% of patients, and final PET (n = 106) was negative in 78%. With a median follow-up of 23 months, 2-year progression-free survival rates were 86% for interim PET–negative versus 61% for interim PET–positive patients (P = .0046) and 87% for final PET–negative versus 51% for final PET–positive patients (P < .001), respectively. Two-year overall survival also significantly differed according to final PET results: 100% versus 88% (P = .0128).

Conclusion

PET performed either after four cycles of R-CHOP or at the end of therapy was strongly predictive of outcome in this prospective study. Therapeutic intervention based on PET results during or after inductive treatment should be evaluated.

J Clin Oncol 30:4317-4322. © 2012 by American Society of Clinical Oncology

INTRODUCTION

Follicular lymphoma is the second most common non-Hodgkin lymphoma subtype. Although a slowly evolving form of cancer associated with long survival times, it remains incurable. No consensus exists regarding optimal initial therapy, but rituximab combined with chemotherapy is a widely recognized standard of care in high–tumor burden patients. Although initially associated with high response rates, this approach is followed by a continuous relapse pattern, with a median time to progression of approximately 4 years in patients receiving R-CHOP (rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone). Response duration is related to quality of response. Some patients either do not respond or relapse early after therapy and have a particularly dismal outcome. Early identification of this subgroup of patients could lead to early therapeutic changes and potentially to a better prognosis.

[$^{18}$F]Fluorodeoxyglucose (FDG) positron emission tomography (PET) is a widely employed technique for the staging and restaging of lymphoma patients.
method of metabolic imaging. In both Hodgkin and diffuse large B-cell lymphoma (DLBCL), residual abnormalities on PET after treatment are associated with an unfavorable outcome, and PET–response adapted therapy is being explored as a means of improving outcomes.\(^7\)

The International Harmonization Project (IHP) recommendations provide guidelines for PET imaging and response criteria for end-of-treatment evaluation in Hodgkin and DLBCL, but they do not recommend use of PET in follicular lymphoma outside of clinical trials.\(^8,9\)

The so-called Deauville criteria were specifically developed for the interpretation of PET performed early during the course of therapy.\(^16\)

Follicular lymphoma is almost always FDG avid, irrespective of tumor grade.\(^11-14\) Several studies have shown an adverse prognostic impact of positive post-therapeutic PET in patients with follicular lymphoma.\(^15-18\) However, these were small retrospective studies performed in heterogeneously treated patients, without homogeneous PET interpretation criteria. More recently, a study performed in 122 patients from the PRIMA (Primary Rituximab and Maintenance) trial indicated that at the end of initial treatment, positive PET (according to review of local investigators’ conclusions) was associated with a two-fold decrease in progression-free survival (PFS; 32.9% vs 70.7% at 42 months).\(^19\) However, this study was retrospective, and PET scans were not centrally reviewed.

The study presented herein is the first prospective study to our knowledge aiming to evaluate the prognostic value of PET in patients with high–tumor burden follicular lymphoma treated with first-line R-CHOP midtreatment (after four cycles) and at the end of therapy, with a centralized review of all PET scans.

### PATIENTS AND METHODS

This nonrandomized study was undertaken between September 2007 and November 2009 in one Italian and 13 French centers. Patients received four cycles of R-CHOP plus two cycles of R-CHOP and two cycles of rituximab in case of response based on conventional criteria (partial response [PR] or complete response [CR]/unconfirmed complete response [CRu]). They were evaluated before therapy, after four R-CHOP cycles, and at the end of therapy by conventional methods (clinical examination, computed tomography [CT], skull base. Images were reconstructed iteratively with and without attenuation correction by each camera. Patients fasted for at least 6 hours before each scan and had to have a blood glucose concentration ≤10 mmol/L. They were administered intravenous injections of 3.5 to 8 MBq/kg (minimal activity, 185 MBq) FDG or C virus infection were also excluded. Patients with known HIV infection or active hepatitis B or C virus infection were also excluded.

The study was approved by the ethics committee (Comité de Protection des Personnes) of Paris IX–Créteil and was undertaken in accordance with the Declaration of Helsinki. Patients were required to provide informed consent before registration.

### Treatment

Patients were to receive R-CHOP (rituximab 375 mg/m\(^2\) intravenously on day 1, cyclophosphamide 750 mg/m\(^2\) intravenously on day 1, vincristine 1 × 4 mg/m\(^2\) [capped at 2 mg] intravenously on day 1, doxorubicin 50 mg/m\(^2\) intravenously on day 1, and prednisone 40 mg/m\(^2\) orally on days 1 to 5, every 3 weeks) for six cycles, plus two additional cycles of rituximab 375 mg/m\(^2\)/week for 3 weeks. Four patients received an antiangiogenic agent in addition to R-CHOP in the setting of a separate phase IIb/I trial.\(^20\) No form of maintenance treatment was planned.

### Response Evaluation and PET Scan Modalities

Responses according to conventional diagnostic methods were qualified by each investigator in accordance with the 1999 International Workshop Criteria (IWC).\(^21\) No central reviews of bone marrow biopsies or CT scans could be planned in this study because of cost constraints; thus, conventional responses were based on investigator conclusion. After the end of treatment, patients were observed on a regular basis (clinical examination every 3 months and CT scan every 6 months during the first 2 years), and investigators were requested to deliver information on patient outcome on a 6-month basis.

PET was performed in each center on a dedicated PET scanner according to standardized modalities, taking into account the technical characteristics of each camera. Patients fasted for at least 6 hours before each scan and had to have a blood glucose concentration <10 mmol/L. They were administered intravenous injections of 3.5 to 8 MBq/kg (minimal activity, 185 MBq) FDG and were asked to lie in supine position for 1 hour to avoid muscular uptake. Imaging was performed to cover a volume starting from the upper thigh to the skull base. Images were reconstructed iteratively with and without attenuation correction. PET quality control (regular testing of image quality performed by a qualified physicist as recommended by the SFPF [French Society of Medical Physics]) was required from each center.

Central review was performed by three experienced nuclear medicine physicians (A.B.-R., A.J., M.M.) on a Positron workstation (Kéosys, Saint-Herblain, France).\(^24\) Differences between observers were resolved by majority view. PET results were reported using the Deauville 5-point scale.\(^10\) Two different thresholds were compared to define positivity and negativity: residual activity greater than the liver activity (scores 4 and 5 on 5-point scale), and residual activity greater than the mediastinal blood pool (scores 3, 4, and 5 on 5-point scale).

### Statistical Analysis

Quantitative variables were summarized in tables displaying sample size, mean, standard deviation, median, and range. Qualitative variables were described in terms of percentages of the number of patients examined.

Censored data were presented as Kaplan–Meier plots of time to first event and summary tables of Kaplan–Meier estimates for criterion rates at fixed time points, with 95% CIs.\(^25\) For the primary criterion, survival end points were analyzed using the log-rank test. Estimates of prognostic factors were expressed as hazard ratios based on the Cox proportional hazards model with 95% CIs.\(^26\) All statistical tests were two sided and performed using a 5% level of
RESULTS

A total of 121 patients were enrolled onto this study from September 18, 2007, to November 18, 2009. On review, seven patients did not fulfill the criteria for high tumor mass, although they had been deemed as requiring treatment by the investigator.

Demographic and Other Baseline Characteristics

The median age of the study population was 57 years, and 63% of patients were men. The main baseline characteristics are summarized in Table 1.

The study is summarized in Figure 1: Two patients were excluded from the full analysis set because of lack of data. Two patients did not receive any study treatment and were thus excluded from the intent-to-treat analysis: One patient had to undergo urgent splenectomy because of splenic rupture resulting from lymphoma, and the other died as a result of infectious complications after the first cycle. The other 116 patients completed the first four treatment cycles. Three died as a result of infectious complications after the first cycle. The analyzable patients received at least one treatment cycle. One patient was lost to follow-up before receiving any study treatment. Thus, 117 because of insufficient response (stable disease; n = 2) or investigator decision (n = 1). Six cycles of R-CHOP were administered to 113 patients, but only 107 received the first additional rituximab injection, and only 106 received the second one, mainly because of investigator decision.

PET Scan Conclusions

The initial PET scan was available for 117 treated patients. Only one initial PET was considered negative (no significant uptake above background). At cycle four, one patient had no PET review because of early death, and five had no review because of absence of adequate data transfer to reviewers. Of 111 centrally reviewed patients, 84 (76%) had a negative PET at the intermediate assessment. Eleven patients had no end-of-treatment PET review; four because PET was not performed on investigator decision, and seven because of inadequate data transfer to reviewers. Of 106 centrally reviewed end-of-treatment PETs, 83 (78%) were considered negative. Among those reviewed, 104 patients had complete data available at cycles four and eight. With a cutoff value of ≥ 4 to define PET positivity, among 78 patients with a negative PET at cycle four, 72 (92%) remained negative at end-of-treatment evaluation, but six (8%) reverted to positive at the end of treatment. On the other hand, nine of 26 positive patients at cycle four became negative; however, 17 (65%) of 26 remained positive.

Concordance Among Conclusions of Review Board Members

The mean κ coefficient between the three observers was calculated for each PET interpretation according to the chosen level of positivity: The mean κ coefficient was 0.707, corresponding to a good level of agreement (values > 0.7 are usually considered satisfactory) when using liver activity as a threshold level. When using mediastinal blood pool activity as a threshold to define positivity, the mean κ coefficient was compatible with a moderate level of agreement, with a value of 0.57.

Concordance in Response Designations Between 1999 IWC and PET at the End of Treatment

At end of treatment, 104 patients had both a centrally reviewed PET and a formal response designation according to IWC. On the basis of IWC, 54 had achieved CR; 20, CRu; 26, PR; and three, stable disease, whereas one had experienced disease progression. As expected, only one (2%) of 54 patients achieving CR had a positive PET. However, 50% of patients (10 of 20) achieving CRu and 34% of patients (nine of 26) achieving PR still had a positive

Table 1. Main Baseline Demographics and Clinical Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No.</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>28-76</td>
<td></td>
</tr>
<tr>
<td>≥ 60</td>
<td>54</td>
<td>45</td>
</tr>
<tr>
<td>Male sex</td>
<td>63</td>
<td>53</td>
</tr>
<tr>
<td>Ann Arbor stage III/IV</td>
<td>110</td>
<td>93</td>
</tr>
<tr>
<td>ECOG performance status ≤ 1</td>
<td>47</td>
<td>40</td>
</tr>
<tr>
<td>B symptoms</td>
<td>21</td>
<td>18</td>
</tr>
<tr>
<td>Bone marrow lymphoma involvement</td>
<td>64</td>
<td>58</td>
</tr>
<tr>
<td>Lactate dehydrogenase &gt; ULN</td>
<td>31</td>
<td>26</td>
</tr>
<tr>
<td>Hemoglobin &lt; 120 g/L</td>
<td>23</td>
<td>19</td>
</tr>
<tr>
<td>Four nodal areas</td>
<td>68</td>
<td>57</td>
</tr>
<tr>
<td>β2-microglobulin ≥ 3 mg/L</td>
<td>26</td>
<td>31</td>
</tr>
<tr>
<td>FLIPI score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low (none to one risk factors)</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>Intermediate (two risk factors)</td>
<td>50</td>
<td>43</td>
</tr>
<tr>
<td>High (three to five risk factors)</td>
<td>49</td>
<td>42</td>
</tr>
<tr>
<td>Initial local diagnosis of FL (other than grade 3B)</td>
<td>119</td>
<td>100</td>
</tr>
<tr>
<td>Central pathologic review performed</td>
<td>109</td>
<td>91</td>
</tr>
<tr>
<td>Confirmed FL (other than grade 3B)</td>
<td>97</td>
<td>88</td>
</tr>
<tr>
<td>Diagnosis of other lymphoma subtype</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Unclassifiable or not assessable for technical reasons</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

Abbreviations: ECOG, Eastern Cooperative Oncology Group; FL, follicular lymphoma; FLIPI, Follicular Lymphoma International Prognostic Index; ULN, upper limit of normal.

Because of insufficient response (stable disease; n = 2) patients did not continue with planned treatment after four cycles; the other 116 patients completed the first four treatment cycles. Three died as a result of infectious complications after the first cycle. The analyzable patients received at least one treatment cycle. One patient was lost to follow-up before receiving any study treatment. Thus, 117 received the first additional rituximab injection, and only 106 received the second one, mainly because of investigator decision.

PET Scan Conclusions

The initial PET scan was available for 117 treated patients. Only one initial PET was considered negative (no significant uptake above background). At cycle four, one patient had no PET review because of early death, and five had no review because of absence of adequate data transfer to reviewers. Of 111 centrally reviewed patients, 84 (76%) had a negative PET at the intermediate assessment. Eleven patients had no end-of-treatment PET review; four because PET was not performed on investigator decision, and seven because of inadequate data transfer to reviewers. Of 106 centrally reviewed end-of-treatment PETs, 83 (78%) were considered negative.

Among those reviewed, 104 patients had complete data available at cycles four and eight. With a cutoff value of ≥ 4 to define PET positivity, among 78 patients with a negative PET at cycle four, 72 (92%) remained negative at end-of-treatment evaluation, but six (8%) reverted to positive at the end of treatment. On the other hand, nine of 26 positive patients at cycle four became negative; however, 17 (65%) of 26 remained positive.

Concordance Among Conclusions of Review Board Members

The mean κ coefficient between the three observers was calculated for each PET interpretation according to the chosen level of positivity: The mean κ coefficient was 0.707, corresponding to a good level of agreement (values > 0.7 are usually considered satisfactory) when using liver activity as a threshold level. When using mediastinal blood pool activity as a threshold to define positivity, the mean κ coefficient was compatible with a moderate level of agreement, with a value of 0.57.

Concordance in Response Designations Between 1999 IWC and PET at the End of Treatment

At end of treatment, 104 patients had both a centrally reviewed PET and a formal response designation according to IWC. On the basis of IWC, 54 had achieved CR; 20, CRu; 26, PR; and three, stable disease, whereas one had experienced disease progression. As expected, only one (2%) of 54 patients achieving CR had a positive PET. However, 50% of patients (10 of 20) achieving CRu and 34% of patients (nine of 26) achieving PR still had a positive
PET at the end of treatment. Overall, 21 (20%) of 104 patients had discordant response designations, and 83 (80%) of 104 had concordant designations.

Survival Analysis

With date of last contact censored at stopping date (most recent date with 10% of patients presenting with an earlier date of last contact), the median duration of follow-up for the intent-to-treat population was 23 months. The estimated 2-year PFS for the whole population is 79.6%, and the estimated 2-year overall survival (OS) is 96.4%.

Using liver activity as a threshold to define positivity (scores 4 and 5 of 5-point scale) allowed the best separation in terms of prognosis (data not shown). Results are thus given using the following criteria. PFS differed significantly according to the results of both PET performed after four cycles and at the end of treatment. The estimated PFS at 2 years was 86% in patients with a negative PET at cycle four versus 61% in those with a positive PET (P = .0046) and 87% in patients with a negative PET at the end of treatment versus 51% in those with a positive PET (P < .001; Fig 2).

Two-year OS also significantly differed according to final PET results: 100% versus 88% (P = .0128), whereas the results of the PET performed at cycle four did not have a significant impact on OS (Fig 3).

Standard response evaluation using 1999 IWG criteria was less proficient in predicting PFS than response evaluation with PET. Two-year PFS was 67.7% for patients achieving PR according to 1999 IWG criteria, compared with 83.2% for patients achieving CR or CRu (P = .1063).

Patients who had a positive PET at four cycles and eventually converted to a negative PET (n = 9) did not have a significantly different PFS than those for whom both examinations were negative (n = 72; 2-year PFS, 72.9% vs 88.9%). Those patients who converted from a negative PET at four cycles to a positive end-of-treatment PET (n = 6) had, on the other hand, a PFS that did not significantly differ from that of patients for whom both examinations were positive (n = 16; 2-year PFS, 33.3% vs 51.9%). Three of these patients had histologic or cytologic disease transformation at relapse.

PFS and OS did not differ significantly according to FLIPI (Follicular Lymphoma International Prognostic Index), either when considering three categories or when considering only two groups (0 to 2 vs 3 to 5; data not shown). The influence of PET results on PFS was observed within each subgroup according to FLIPI (Fig 4).

DISCUSSION

This study confirms the prognostic value of PET in patients with follicular lymphoma, which has already been suggested by several retrospective studies.15-19 This prognostic value is maintained whatever the FLIPI category. Although there is no firm consensus on the optimal first-line regimen in high–tumor burden follicular lymphoma, the R-CHOP regimen used in this study is a widely used option chosen by 73% of investigators in the PRIMA trial5 and by 55% of physicians choosing rituximab plus chemotherapy as a treatment option in the US National LymphoCare Study.28 Our conclusions should be extended to patients treated with other approaches with caution, but the prognostic role of a positive PET has also been suggested for patients treated with radioimmunotherapy.29

None of our patients received maintenance treatment with rituximab, which has been shown to provide a PFS benefit in patients treated in the first line.5 In the retrospective analysis of PET scans
performed at the end of treatment in the PRIMA trial, a nonsignificant inferior PFS was observed in nine PET-positive patients among 47 receiving maintenance (55.6% vs 77.4%; P = .18). It was not expected that the PFS values observed in our study would be similar to those observed in the PRIMA study, given the fact that half of these patients received maintenance rituximab. This might in part be related to shorter follow-up. The prognostic value of PET in patients with follicular lymphoma should be studied prospectively in the future in patients receiving maintenance rituximab.

In our study, response evaluation with PET had a superior predictive power than investigator-assessed IWC response. However, almost every patient achieving CR according to IWC also achieved metabolic CR by PET (53 of 54 patients; Table 2); thus, PET could be reserved for patients who have not reached CR.

On the whole, these results strongly support moving away from response criteria that are based solely on CT, as recommended in the 2007 IHP for patients with Hodgkin lymphoma or DLBCL. In our opinion, end-of-treatment response evaluation by PET/CT should be included in future first-line follicular lymphoma trials. Using the Deauville 5-point scale for PET interpretation, defining positivity with a threshold of ≥ 4, provided the best separation in terms of PFS at the end of treatment. Interestingly, a significant separation in terms of PFS was also found using a lower cutoff (≥ 3; P < .001), which is roughly equivalent to using standard end-of-treatment IHP criteria (same reference background to define PET positivity). However, because using a threshold of ≥ 4 provided the best levels of concordance among observers, it seems logical to propose this cutoff value for future studies.

PET is able to provide a meaningful surrogate for PFS in this patient population, allowing for future exploration of response-adapted treatment strategies in follicular lymphoma, as they are now under investigation in patients with other lymphoma subtypes, 30,31

Evaluation by PET before four cycles, as has been evaluated in other lymphoma subtypes, has to the best of our knowledge never been evaluated in patients with follicular lymphoma. PET scans performed both after four cycles of therapy and at the end of treatment were able to predict PFS, but approximately 15% of patients (15 of 104) had discordant results between interim and final PET results, and interim PET was not predictive of OS in this study. Thus, on the basis of our data, we would suggest waiting until the end of induction treatment before performing PET and deciding on eventual treatment modifications.

High-dose therapy with stem-cell transplantation has been abandoned in the first line because its benefits in terms of PFS were counterbalanced by excessive late toxicity. However, it remains a frequently used option in relapsed patients, for whom it seems to offer better outcomes. 32 Such an approach obviously represents an appealing strategy in post-therapy PET-positive patients. Other approaches, such as using new drugs (eg, lenalidomide and so on) as maintenance therapy in PET-positive patients, also seem particularly attractive.

Table 2. Concordance in Response Designations Between International Workshop Criteria and PET at the End of Treatment

<table>
<thead>
<tr>
<th>PET Review Board Conclusion (No. of patients)</th>
<th>Negative</th>
<th>Positive</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete response</td>
<td>53</td>
<td>1</td>
<td>54</td>
</tr>
<tr>
<td>Unconfirmed complete response</td>
<td>10</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Partial response</td>
<td>17</td>
<td>9</td>
<td>26</td>
</tr>
<tr>
<td>Stable disease</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Progressive disease</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Not evaluated</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>83</td>
<td>23</td>
<td>106</td>
</tr>
</tbody>
</table>

Abbreviation: PET, positron emission tomography.
REFERENCES


