The use of Deauville 5-point score could reduce the risk of false-positive fluorodeoxyglucose-positron emission tomography in the posttherapy evaluation of patients with primary bone lymphomas

ABSTRACT
Primary bone lymphoma (PBL) is a rare disease. Little is reported about response evaluation procedures in these patients. Our aim was to evaluate response to therapy according to fluorodeoxyglucose-positron emission tomography (FDG-PET) results, and in particular to test the Deauville 5-point scale as compared to the visual evaluation of FDG-PET scans in PBL. In this single-center study, we diagnosed 31 consecutive patients with PBL, of which 24 were evaluated with end-of-treatment FDG-PET. Patients’ ages ranged from 19 to 82 years. Six patients were treated with chemotherapy, 24 with chemotherapy and radiotherapy, and one patient with radiotherapy alone. Six patients were affected by a pathological fracture. Four patients died within the range of 3 to 36 months after diagnosis. The average follow-up of the remaining patients was 70 (24–173) months. Overall survival was 87% at 5 years. The only positive prognostic factor was complete remission after chemotherapy. According to visual criteria, end-of-treatment FDG-PET was evaluated in 24 patients and it was positive in 11 (46%) and negative in 13 patients. We organized a retrospective central-blinded revision of end-of-treatment FDG-PET scans using the 5-point Deauville Score (DS). We reviewed 17 out of 24 patients and obtained the following results: at the end of therapy, 12 patients with DS score 2, three patients with DS score 3, one patient with DS score 4, and none with DS score 5. Considering that all the 24 patients achieved complete remission after treatment, visual interpretation produced 11/24 false-positive results, and DS interpretation produced 1/17 false-positive results, thus significantly reducing the number of false positives. In PBL, the final evaluation at the end of therapy with FDG-PET should be evaluated using Deauville 5-point scale in order to significantly reduce the risk of false-positive scans.

Keywords: Bone lymphomas, Deauville criteria, false positive

INTRODUCTION
Primary bone lymphoma (PBL) is a rare disease, representing only 2% of all bone tumors and 5% of all extranodal lymphomas.1,2 The definition of the disease itself is still somewhat controversial, but now there is a prevalent agreement on considering lymphomas as primary lymphomas of bone when the disease affects one or more bones with or without involvement of local lymph nodes, but with no evidence of disease in distant nodes or other extraosseous sites, i.e., Ann Arbor Stage I and Stage II disease;3 a lymphoma with the above-mentioned pattern can be considered a PBL, also when involvement of bone marrow is present at diagnosis, i.e., Ann Arbor Stage IV.4,5
PBL can occur in adults and children. A different behavior of the disease can be expected in this two different groups with a more rapid systemic spread, but still a better prognosis, in children. Therefore, adult and pediatric cases should be separately examined in the evaluation of treatment and prognosis.

Due to rarity of the disease, few series were reported in literature, and only retrospective, multicentric reviews involved a high number of patients.

The majority of studies in PBL patients address issues regarding treatment regimens, chemotherapy alone, or chemotherapy plus radiotherapy being the cornerstones. Little is reported in literature about imaging evaluation of final response to therapy in PBL. Usually, the same procedures applied in all other lymphomas are used. In particular, in the past years, 18F-fluorodeoxyglucose-positron emission tomography (FDG-PET) has been increasingly used for staging and to define response to therapy, according to the last Cheson response criteria.

At staging, FDG-PET has been demonstrated to be very useful in PBL, especially in defining bone localizations unrecognized at computed tomography (CT). However, the use of FDG-PET in therapy response evaluation is still controversial. Some authors suggest that FDG-PET scan should be interpreted with caution, due to a persistent FDG uptake in bone lesions even after remission in some primary bone-diffuse large B-cell lymphoma (PB-DLCL) patients. Frequently at the end of therapy, bone lesions continue to be FDG-PET positive and this situation could be particularly stressful for both patient and clinician, inducing to perform local biopsy or more intensive instrumental follow-up.

Therefore, we decided to retrospectively review our series of PBLs, with the aim to evaluate response to therapy according to FDG-PET results, and in particular to test a new interpretation method of FDG-PET scans, according to Deauville criteria. Our hypothesis is that the use of a semi-quantitative method such as the Deauville 5-point scale could elude, or at least decrease, the risk of FDG-PET false-positive results.

MATERIALS AND METHODS

Patients

Among the patients treated for lymphoma at the Haematological Department of Florence from 1999 to 2014, 31 adult patients fitted the above-mentioned criteria for PBL. There were 17 males and 14 females. Age at diagnosis ranged from 19 to 82 years (average: 53 years). Clinical characteristics are summarized in Table 1. The most common histological type was DLBCL, accounting for 90% of the cases (28 patients); two patients were affected by follicular lymphoma and one patient was affected by small lymphocytic lymphoma. Ann Arbor stage was Stage I in 11 patients, Stage II in two patients, and Stage IV in 18 cases.

In all cases, diagnosis was obtained by a percutaneous or open biopsy. Staging of the patients at presentation was accomplished by a contrast-enhanced CT scan of neck, chest, abdomen, and pelvis, a total-body FDG-PET scan, a bone marrow biopsy, and blood parameters’ evaluation, including lactate dehydrogenase and beta-2 microglobulin.

One patient underwent only radiotherapy, six patients were treated with chemotherapy alone, and the remaining 24 patients underwent a combined treatment, including both chemotherapy and radiotherapy.

Chemotherapy regimens used were CHOP or CHOP-like (cyclophosphamide, adriablastina, vincristina, and

| Table 1: Clinical characteristics and response to therapy |
|----------------------------------------------------------|
| **Characteristics**                                      | **n (%)** |
| Median age                                              | 53 (range 19-82) |
| Female                                                  | 14 (45) |
| Male                                                    | 17 (55) |
| Histotype                                               |          |
| DLBCL                                                   | 28 (90) |
| FL                                                      | 2 (6)   |
| SLL                                                     | 1 (4)   |
| Stage                                                   |          |
| I                                                       | 11 (35) |
| II                                                      | 2 (6)   |
| IV                                                      | 18 (59) |
| LDH normal                                              | 19 (61) |
| LDH abnormal                                            | 12 (39) |
| International Prognostic Index                         |          |
| 0-1                                                     | 7 (22)  |
| 2                                                       | 10 (34) |
| 3                                                       | 7 (22)  |
| 4-5                                                     | 7 (22)  |
| Therapy                                                 |          |
| R-CHOP and R-CHOP-like                                  | 24 (78) |
| R-MiCEP                                                 | 3 (10)  |
| R-MACOP-B                                               | 2 (6)   |
| Chlorambucil                                            | 1 (3)   |
| Radiotherapy                                            | 1 (3)   |
| Response to therapy                                     |          |
| Complete Remission                                      | 27 (87) |
| Partial Remission                                       | 3 (10)  |
| Non Responder                                           | 1 (3)   |

DLBCL: Diffuse Large B Cell Lymphoma; FL: Follicular Lymphoma; SLL: Small Lymphocytic Lymphoma
prednisone\textsuperscript{[19]} in 24 patients, MICEP (cyclophosphamide, mitoxantrone, etoposide, and prednisone)\textsuperscript{[20]} in 3, MACOP-B (methotrexate, Adriablastina, cyclophosphamide, vincristina, bleomycin, and prednisone)\textsuperscript{[21]} in 2 patients, and clorambucil in 1 patient.

Evaluation at the end of therapy was performed with total-body CT, magnetic resonance imaging (MRI) of the involved bone lesions, and FDG-PET. Twenty-four (77\%) patients were evaluated with FDG-PET at the end of treatment (FDG-PET scan was available since 2003) to evaluate response to therapy. Patients with a single localization unregardless to CT, MRI, or PET response were treated with involved field radiotherapy. Patients with multiple localizations did not perform radiotherapy: one of these patients was treated with intensification high-dose chemotherapy and peripheral blood stem cell reinfusion and other were strictly followed with CT or nuclear magnetic resonance on the sites of doubt-persistent disease. None of the patients with multiple localizations were treated with radiotherapy, even in case of persistent PET positivity.

Follow-up of the patients was accomplished by clinical examination and instrumental evaluation, alternating total-body CT and MRI of the involved bone lesions combined with blood test every 3 months in the 1\textsuperscript{st} year after the end of therapy. In the 2\textsuperscript{nd} and 3\textsuperscript{rd} years after therapy, clinical and instrumental examinations were performed every 6 months. FDG-PET scan was used at the end of therapy to confirm the attainment of complete remission and in the follow-up only in the case of suspected recurrence.

**Complete response definition**

It is defined as complete disappearance of all detectable clinical evidence of disease and disease-related symptoms if present before therapy.

**Typically fluorodeoxyglucose-avid lymphoma**

in patients with no pretreatment PET scan or when the PET scan was positive before therapy, a posttreatment residual mass of any size is permitted as long as it is PET negative.

**Variably fluorodeoxyglucose-avid lymphomas/ fluorodeoxyglucose avidity unknown**

In patients without a pretreatment PET scan, or if a pretreatment PET scan was negative, all lymph nodes and nodal masses must have regressed on CT to normal size (1.5 cm in their greatest transverse diameter for nodes 1.5 cm before therapy). Previously involved nodes that were 1.1–1.5 cm in their long axis and more than 1.0 cm in their short axis before treatment must have decreased to 1.0 cm in their short axis after treatment.

The spleen and/or liver, if considered enlarged before therapy on the basis of a physical examination or CT scan, should not be palpable on physical examination and should be considered normal size by imaging studies, and nodules related to lymphoma should disappear. However, determination of splenic involvement is not always reliable because a spleen considered normal in size may still contain lymphoma, whereas an enlarged spleen may reflect variations in anatomy, blood volume, the use of hematopoietic growth factors, or causes other than lymphoma.

If the bone marrow was involved by lymphoma before treatment, the infiltrate must have cleared on repeat bone marrow biopsy. The biopsy sample on which this determination is made must be adequate (with a goal of 20 mm unilateral core). If the sample is indeterminate by morphology, it should be negative by immunohistochemistry. A sample that is negative by immunohistochemistry but that demonstrates a small population of clonal lymphocytes by flow cytometry will be considered a complete response until data become available demonstrating a clear difference in patient outcome.

**Fluorodeoxyglucose-positron emission tomography**

FDG-PET examinations were performed according to the European Association of Nuclear Medicine guidelines\textsuperscript{[22]} and acquired using a 16-slice PET/CT hybrid system (Philips Gemini TF 16 PET/CT). Briefly, patients were instructed to fast at least 4 h prior to the intravenous administration of 350–450 MBq of FDG. Blood glucose level was measured before tracer injection so as to ensure levels <160 mg/dl. Imaging started 60 ± 15 min after intravenous tracer administration. Unenhanced low-dose CT was performed at 120 kV and 50 mA for attenuation correction of emissive data and anatomical localization of PET data set. Emissive scan was performed in three-dimensional (3D) mode, shortly after CT acquisition, with a 2-min acquisition per bed position and field of view of 576 cm. PET images were reconstructed using an iterative reconstruction algorithm (3D LOR RMLA), matrix 512 × 512, voxel size of 4 mm × 4 mm × 4 mm.

Scans were performed starting from the orbital plane on to the mid-thigh, except for the cases where the clinical history demanded a whole-body, vertex-to-toes scan.

PET images were qualitatively interpreted according to Juweid criteria at the time of PET examination.

Moreover, FDG-PET examinations were retrospectively re-evaluated using Deauville 5-point scale by an experienced
reader blinded from qualitative interpretation and patients’ follow-up data.

Statistical analysis

Our data were analyzed using the Statistical Package for the Social Sciences (SPSS, Inc., Chicago, IL, USA). Curves for overall survival (OS) were estimated using the Kaplan–Meier method. The log-rank test was used to assess the significance of differences for each prognostic factor in univariate analysis. OS was calculated from the time of diagnosis until death for any cause or last contact. Cox proportional hazards regression models and logistic regression analyses were used to assess the ability of patients’ characteristics to predict OS. Specifically, we first performed a backward elimination with cutoff value of 0.1; then, variables with \( P = 0.05 \) were entered in the model. The limit of significance for all analyses was defined as \( P = 0.05 \). Two-sided tests were used in all calculations.

RESULTS

At the end of therapy, 27 patients (87%) reached a complete remission of the disease, 3 had a partial remission with an overall response rate of 97%, and one patient did not respond and progressed. Until 2003, the final evaluation of response was performed with total-body CT and MRI, both contrast enhanced. Persisting radiological abnormalities with osteolytic lesions after treatment is a common finding in PBL and it does not prevent from the assessment of a remission state, which is evaluated upon the disappearance of local (extraosseous mass, PET uptake, and pain) and systemic signs of lymphoma,\(^5,23\) with a strict instrumental follow-up.

According to qualitative visual criteria of PET evaluation at the end of therapy, PET examination was positive for persistence of disease in 11 (46%) patients and negative (indicating complete metabolic response) in 13 (54%) patients. Six patients with positive FDG-PET at visual assessment repeated the examination after 3 months: in 4 (67%) of them, PET persisted positive in the same sites and in two it was negative, confirming a false positivity of previous scan. In two patients with persistent positivity, a CT-guided biopsy was performed and the histological results were negative in both cases.

At the clinical follow-up, 22 patients (92%) achieved a complete remission of disease at 6 months after treatment, according to the criteria defined above, and after a median follow-up period of observation of 80 months (range: 24–140 months). One patient relapsed after 15 months from

| Table 2: Performance of visual assessment fluorodeoxyglucose-positron emission tomography |
|-----------------------------------------------|---------------|---------------|
| | Visual assessment FDG-PET |             |               |
| | Negative | Positive | Total |
| Clinical follow-up | | | |
| Negative | 13 | 11 | 24 |
| Positive | 0 | 0 | 0 |
| Total | 13 | 11 | 24 |

| Table 2a: Performance of Deauville 5-point scale evaluation at end-of-treatment fluorodeoxyglucose-positron emission tomography |
|-----------------------------------------------|---------------|---------------|
| | Deauville 5-point scale FDG-PET |             |               |
| | Negative | Positive | Total |
| Clinical follow-up | | | |
| Negative | 16 | 1 | 17 |
| Positive | 0 | 0 | 0 |
| Total | 16 | 1 | 17 |

| Table 2b: Direct comparison between visual assessment and Deauville 5-point scale evaluation at end-of-treatment fluorodeoxyglucose-positron emission tomography in 17 primary bone lymphoma patients |
|-----------------------------------------------|---------------|---------------|
| | Deauville 5-point scale FDG-PET |             |               |
| | Negative | Positive | Total |
| Visual assessment FDG-PET | | | |
| Negative | 9 | 0 | 9 |
| Positive | 7 | 1 | 8 |
| Total | 16 | 1 | 17 |
with DS 4 [Figure 2a and b], and none with DS 5. Setting the cutoff for negativity at DS 3, using the Deauville 5-point scale, PET examination was negative in 16 patients and positive in only one patient [Tables 2a and b]. Among the two patients who relapsed after more than 6 months, the first one presented a DS 2 at the end of therapy and the other patient had a DS 4.

The use of Deauville 5-point scale at end-of-treatment FDG-PET significantly reduced the number of false-positive results.

Considering all patients, four patients died of disease, at a time ranging from 3 to 36 months from diagnosis (average: 12.2 months). The average follow-up of the remaining patients was 70 months, ranging from 24 to 173 months. Kaplan–Meyer OS and progression-free survival were, respectively, 87% and 83% at 5 years [Figures 3 and 4].

**DISCUSSION**

In our series after a median follow-up period of 70 months, the OS of patients affected by PBLs was 87%. Different values were found by different authors in previous papers, ranging from 58% to as much as 90%.[9–13,22,24–27] An OS of 95% at 8 years in an homogeneous cohort of patients affected by DLBCLs treated with anthracycline-containing chemotherapeutic regimens with the addition of rituximab was reported.[28–31] The progression-free survival for our series was 83% with four patients who did not obtain a remission and only two patients relapsed.

Our series cannot add significant data about this issue, because of the retrospective nature and the limited number of patients and because it is characterized by a high heterogeneity of treatment regimens adopted as well. Nevertheless, our results are in line with those reported in
literature. Considering this limitation, the main aim of our work was to analyze the role of end-of-therapy evaluation with FDG-PET in a consecutive series of PBLs, as very few and often single case report or incomplete data are reported in literature about this issue.

FDG-PET has proven to be more specific and sensitive than conventional bone scintigraphy in identifying osseous involvement by malignant lymphoma. FDG-PET has been routinely used for staging and restaging in patients with Hodgkin lymphoma (HL) and high-grade non-HL (NHL). A lot of studies revealed that FDG-PET has higher sensitivity and specificity than CT in the staging evaluation of HL and DLBCL. The frequency of extranodal involvement is roughly 50%–60% in different casistics and the most affected sites were the lungs, bone, bone marrow, spleen, and liver in HL whereas the most affected sites in NHL were spleen, bone, bone marrow, Waldeyer’s ring, and soft tissues.

Coming to the main issue of this study, we suggested that PET/CT scan should be interpreted with caution due to a persistent FDG uptake of bone lesions even after remission in some PB-DLBCL patients and that the use of Deauville 5-point scale is more adequate to define response to therapy, significantly reducing the incidence of false-positive scans. Due to missing events in our casistic, it is not possible to define positive predictive value and negative predictive value. First, Israel et al. in 2002 reported that gallium scintigraphy has high sensitivity and specificity for the evaluation of response to therapy in PBL in comparison with CT and MRI; nevertheless, 58% of patients persisted positive to gallium scintigraphy at the end of therapy. In 2003, Zinzani et al. stated in a casistic of 52 patients that response to treatment is commonly hampered by the persistence of abnormalities on gallium scan and MRI after successful treatment. Ng et al. in 2007 evaluated the prognostic value of persistently positive end-of-therapy FDG-PET and concluded that a positive restaging FDG-PET at sites of bone involvement by DLBCL at diagnosis appears to be less predictive of disease progression than residual abnormality by PET in nodal or extra-nodal soft tissue. On the basis of our experience, the final evaluation of PB NHL using CT or MRI showed the persistence of diffuse bone structural alteration in the bone lesions with persistence of contrast enhancement in some cases. In our study, all patients were evaluated with a visual method according to Juweid criteria, resulting in a very high number of positive scans (11 patients – 46% - FDG-PET positive at the end of therapy). According to the newly proposed response criteria, they should be considered as partial remission. FDG-PET was positive after chemotherapy in five out of seven patients with localized disease (Stage I), and regardless to response, they were treated with radiotherapy. After chemotherapy, patients with multifocal disease and with negative final FDG-PET (eight patients) were followed up with CT or MRI. Patients with
positive PET (six patients) repeated this examination after 2 or 3 months: in two cases, the PET remained positive; all the other patients (four patients) were strictly followed up with MRI. Two of these patients were submitted to biopsy of the persistent PET-positive lesion, and in both cases, the result was histologically negative. To note, that all these patients showed an improvement of symptoms in particular, the most important symptom present at diagnosis, pain, has disappeared at the end of therapy. Moreover, none of these patients relapsed within 1 year from the end of therapy after a median follow-up period of observation of 70 months. The high incidence of false-positive PET scans induced us to reconsider the patients with final FDG-PET. It was possible to review 17 out of 24 FDG-PET scans, because of irrecoverable images. The end-of-treatment FDG-PET scans were evaluated according to Deauville 5-point score system and only one patient had a 4-point score and should be considered as persistent disease. Four patients had a score of 3 and 12 patients had a score of 2 and therefore should be considered as negative for disease. Retrospectively, the use of DS significantly reduced the number of false-positive scans, passing from 46% (11/24) to 6% (1/17) and increased the number of true-negative scans passing from 54% (13/24) to 94% (16/17).

Although our study was retrospective and due to the long period of accrual, only a portion of cases were evaluated for response with FDG-PET; therefore, we can suggest that an end-of-therapy FDG-PET evaluation is mandatory in PBL and that the use of the Deauville 5-point score system permits to reduce significantly the false-positive results. It is possible that the persistence of FDG uptake in bone lesions (a large part of patients with DS 3 and 4) could be associated with the activity of osteoblasts for damaged bone tissue regeneration. In our group of 17 patients reviewed with new criteria, only two patients relapsed, one after 15 months and with DS 2 at the end of induction therapy, so difficult to consider positive, and the other after 18 months with DS 4 at the end of therapy, it is very hard to believe that FDG-PET was able to predict a relapse after more than 1 year. Moreover, this patient was one of those biopsied in one of the positive sites and resulted histologically negative and the relapse was observed in a different bone localization. Relapse was confirmed histologically in a FDG-PET-positive site and the proliferative index (Ki67) was very high, i.e., >90%.

CONCLUSION

We would like to summarize our results pointing out that, in PBL, the end–of‑therapy FDG-PET should be evaluated using Deauville 5-point scale, to significantly reduce the risk of false-positive scans. Our results, although retrospective and in a limited group of patients, should be considered significant, given the rarity of the disease. Obviously, we think that only a multicentric prospective study could definitively respond to this issue.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

REFERENCES

1. Freeman C, Berg JW, Cutler SJ. Occurrence and prognosis of extranodal lymphomas. Cancer 1972;29:252-60.
2. Rudders RA, Ross ME, DeLellis RA. Primary extranodal lymphoma: Response to treatment and factors influencing prognosis. Cancer 1978;42:406-16.
3. Ostrowski ML, Unni KK, Banks PM, Shives TC, Evans RG, O’Connell MJ, et al. Malignant lymphoma of bone. Cancer
Lymphomas of the bone: A pathological and clinical study of Primary lymphoma of bone in adult patients. Cancer Extranodal Hodgkin disease: Spectrum of disease. Hydroxyldaunomycin (Adriamycin) combination chemotherapy Primary bone lymphoma: Evaluation of chemoimmunotherapy international workshop on positron emission Treatment of large cell lymphoma in elderly patients with a...
44. Israel O, Mekel M, Bar-Shalom R, Epelbaum R, Hermony N, Haim N, et al. Bone lymphoma: 67Ga scintigraphy and CT for prediction of outcome after treatment. J Nucl Med 2002;43:1295-303.

45. Zinzani PL, Carrillo G, Ascani S, Barbieri E, Tani M, Paulli M, et al. Primary bone lymphoma: Experience with 52 patients. Haematologica 2003;88:280-5.

46. Ng AP, Wirth A, Seymour JF, Lee M, Hogg A, Januszewicz H, et al. Early therapeutic response assessment by (18)FDG-positron emission tomography during chemotherapy in patients with diffuse large B-cell lymphoma: Isolated residual positivity involving bone is not usually a predictor of subsequent treatment failure. Leuk Lymphoma 2007;48:596-600.