

Interim PET in lymphoma: a step towards standardization

Michel Meignan

Published online: 22 July 2010
© Springer-Verlag 2010

In this issue, Barrington et al. propose well-defined criteria for interim PET/CT reporting in lymphoma [1]. These criteria, based on a five-point scale, introduce two main concepts which are highly relevant for interim PET analysis: (1) the intensity of the residual uptake frequently observed at interim is graded relative to various levels of reference background (RB)—mediastinal blood pool (MBP) and liver—and (2) this grading process is independent of the size of the residual tumour. Furthermore, the authors reported that this five-point scale assessed on a sample of 50 patients with advanced stage Hodgkin's lymphoma (HL) by 4 pairs of experts from 4 European centres: (1) seems a reproducible tool for guiding therapeutic strategy after 2 cycles of ABVD (Adriamycin, bleomycin, vinblastine, dacarbazine) when liver is chosen as an RB and (2) is robust enough to be used in multicentre trials for interim PET reporting.

The authors have decided to use this scale in the Response Adapted Treatment in Hodgkin Lymphoma (RATHL) trial, an interim PET-driven therapeutic strategy in patients with advanced stage HL. Although many trials are ongoing worldwide in HL and non-Hodgkin's lymphoma (NHL) to test whether the results of an interim PET performed after one to four cycles of chemotherapy can tailor a response-adapted therapy, standardized interim PET criteria are still missing. Interim PET requires indeed specific criteria which should be different from those defined for end treatment by the International Harmonization Project (IHP) in Lymphoma [2]. In the context of drug

development and survival improvement observed in lymphoma, the rationale of interim PET is to define an a posteriori prognostic index reflecting tumour chemosensitivity during first-line therapy, allowing early changes of therapy (escalation or de-escalation) adapted to specific situations. In many trials the strategy can be modified according to the results of two interim PET exams performed at different times during induction treatment. As a matter of fact, depending on the characteristics of tumour growth and histological type the evolution of PET under treatment is different. In diffuse large B-cell lymphoma (DLBCL) neoplastic cells constitute more than 90% of the total cell population. The metabolic activity during therapy decreases in a continuous mode, with cell killing. Interim PET after one or two cycles of chemotherapy evaluates the response of the cells with the highest mitotic index, thereby providing an early evaluation of chemosensitivity and identifying responders and non-responders [3, 4]. On the other hand, after three or four cycles of therapy, FDG uptake is more dependent on the tumour regrowth [5]. In HL, the Reed-Sternberg (RS) cells account for less than 1% of the tumour. Bystander, lymphomononuclear cells work, through cytokines, as an amplifier of the PET detection power. These cells are switched off very early (one to two cycles) by chemotherapy [6]. Therefore, the amount of residual activity at interim varies with the kinetics of tumoral destruction, the histology of lymphoma, the tumoral microenvironment, the treatment and the time point (two cycles or more) at which PET is performed. These variations are better described by a quantitative analysis by reference to the baseline PET than by using a visual dichotomous analysis (positive or negative) which results in a loss of information [3, 5]. In this regard the five-point scale proposed by Barrington et al. is going towards the right direction by giving the possibility to grade

M. Meignan (✉)
Department of Nuclear Medicine, Hôpital Henri Mondor,
Assistance Publique - Hôpitaux de Paris, University Paris Est,
Créteil 94010, France
e-mail: michel.meignan@hmn.aphp.fr

the residual uptake at different times using the same set of criteria and to describe the changes observed under treatment.

However, the practical question is the amount of residual FDG uptake that we can tolerate to define the threshold of response/non-response to the treatment and a prognostic index after two cycles of chemotherapy? This is challenging since the residual uptake can be due to residual tumour cells, to inflammatory reaction to chemotherapy but also to cells of the lymph node microenvironment with either good (stromal reaction I) or bad (stromal reaction II) prognostic implications in DLBCL or deleterious prognostic significance (CD68⁺ macrophages) in HL [7, 8]. The treatment efficacy is usually judged visually by comparing the residual activity to that of an RB which can be nearby background, MBP or liver. This visual comparison is somewhat difficult. Contrast depends on the activity of background surrounding the object and on the sharpness of its boundaries. Therefore, two structures with the same activity can be judged visually different if their background activity is different, which could have deleterious therapeutic consequences. For these reasons the RB should be well defined and the comparison with the residual activity reproducible among the observers. This RB should not be too low particularly when interim PET is performed after two cycles of chemotherapy either in NHL or HL. After two ABVD a minimal residual uptake with an SUV_{max} of between 2 and 3.5 was considered PET negative in the series of Gallamini et al. on advanced HL [9]. These values are very high but with these criteria the 2-year progression-free survival in patients with positive PET-2 was 12.8 vs 95% in patients with negative PET-2. In our series of 92 patients with aggressive NHL after two cycles of chemotherapy, an optimal cut-off value of 5.0 (SUV_{max}) was observed [3].

Unfortunately, in most recent clinical trials using interim PET, criteria derived from the IHP criteria have been used. These rely on two RBs with moderate to low activity: the MBP when the residual tumour is larger than 2 cm, and the nearby background for residual tumour smaller than 2 cm. These criteria designed for end-of-treatment evaluation are not recommended for early interim PET reporting [2]. MBP activity is relatively moderate with an SUV_{max} of around 1.6 and is not so simple to assess visually due to the possible uptake in vascular walls and its undefined limits. It is not surprising that using these criteria some studies have reported very high rates of positivity of interim PET in NHL with a positive predictive value (PPV) as low as 26% [10] in discrepancy with the first published series [11]. Moreover, in limited stage HL the rate of positive PET after two ABVD in early stage patients with HL was found with these criteria to be equal to the rate reported for advanced stage patients [12]. Having different RBs depending on

residual tumour size adds another challenge. The interobserver reproducibility of size measurement for tumours in the range of 15 to 20 mm is low. In a Groupe d'Etude des Lymphomes de l'Adulte (GELA) trial on NHL 17% of the discordances between the experts were explained by differences in the assessment of the size of residual tumour.

Barrington et al. have chosen in the RATHL trial the liver as the RB. Even if they cannot give at this time any information on the follow-up to validate this choice, they clearly show in their study that using the liver background (PET was judged positive if patients had residual uptake grade 4–5, that is moderately or markedly increased compared to liver) gives a much higher intercentre reproducibility than using the MBP with a $\kappa = 0.89$. They conclude that these criteria are sufficiently robust to be used in a multicentre setting. The liver as an RB for interim PET reporting has indeed many advantages over the MBP. Liver has well-defined limits with a relatively high SUV_{max} of around 2.5 which stays stable after two cycles of chemotherapy. Using the five-point scale and the liver as cut-off, we have observed a reduction of 45% of the PET-positive rate reported in the H10 trial using the IHP criteria based on MBP, nearby background and size measurement (Bardet et al., 2010, Second International Workshop on Interim PET in Lymphoma, Menton, personal communication).

In 38 NHL patients Horning et al. had reported a much lower interobserver agreement ($\kappa = 0.50$) using the five-point scale for interim PET reporting after three cycles of rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone (R-CHOP) [13]. These results are difficult to compare to Barrington et al.'s study because lymphoma types and PET procedures are different (3 vs 2 cycles, 60 vs 90 min post-injection). The reading conditions were slightly different. In Horning et al.'s study, the interim scans were read independently by three experts on the same software on different computers. In Barrington et al.'s study, the opinion of each one of the four centres resulted from the readings of two experts within the centre, using their custom workstation and software. An intercentre kappa evaluation is expected to be a bit higher than an interobserver kappa. With this liver cut-off 30% of patients have a positive PET after consensus, which is much lower than when the MBP cut-off was used (39%). However, it is much higher than the 19% reported by Gallamini et al. in similar populations of patients with advanced stage HL, but the cut-off was an SUV_{max} of 3.5, which is one unit SUV more than the liver SUV_{max} [9]. If this figure is maintained in a larger population, this will probably result in a lower PPV than the 93% reported [9]. One of the main advantages of the five-point scale is to allow the outcome of patients with different levels of residual FDG uptake to be analysed to determine if the separation of patients according to the degree of FDG uptake is meaningful and to grade the

evolution of PET during therapy. However, to achieve results comparable to the previous studies [9], it would probably be necessary to increase the upper threshold of the scale above the activity of the liver and set the positivity cut-off at grade 5, that is markedly increased uptake above the liver. Itti et al. have recently shown in NHL that after two cycles it was necessary to set the cut-off at 125% of the liver SUV_{max} for achieving a significant separation in patient outcome [14]. The criteria presented by Barrington et al. have many advantages. They are simple to apply without taking into account the size of the tumour. The grading allows one to adapt the cut-off according to lymphoma subtype, the objective of the trial (escalation or de-escalation) and the time of the interim PET. Therefore, this five-point scale was approved during the First International Workshop on Interim PET in Deauville in 2009 where it was suggested that the added value of an SUV analysis be tested [15]. Two international validation studies have been launched to validate the five-point scale in retrospective series of patients with HL and NHL. By establishing these criteria Barrington et al. have performed an inestimable work which has opened the way to the standardization of interim PET reporting.

References

- Barrington SF, Qian W, Somer EJ, Franceschetto A, Bagni B, Brun E, et al. Concordance between four European centres of PET reporting criteria designed for use in multicentre trials in Hodgkin lymphoma. *Eur J Nucl Med Mol Imaging* 2010. doi:10.1007/s00259-010-1490-5.
- Juweid ME, Stroobants S, Hoekstra OS, Mottaghy FM, Dietlein M, Guermazi A, 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):571–8.
- Lin C, Itti E, Haioun C, Petegnief Y, Luciani A, Dupuis J, 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(10):1626–32.
- Meignan M, Itti E, Gallamini A, Haioun C. Interim 18F-fluorodeoxyglucose positron emission tomography in diffuse large B-cell lymphoma: qualitative or quantitative interpretation—where do we stand? *Leuk Lymphoma* 2009;50(11):1753–6.
- Itti E, Lin C, Dupuis J, Paone G, Capacchione D, Rahmouni A, et al. Prognostic value of interim 18F-FDG PET in patients with diffuse large B-cell lymphoma: SUV-based assessment at 4 cycles of chemotherapy. *J Nucl Med* 2009;50(4):527–33.
- Gallamini A, Fiore F, Sorasio R, Meignan M. Interim positron emission tomography scan in Hodgkin lymphoma: definitions, interpretation rules, and clinical validation. *Leuk Lymphoma* 2009;50(11):1761–4.
- Lenz G, Wright G, Dave SS, Xiao W, Powell J, Zhao H, et al. Stromal gene signatures in large-B-cell lymphomas. *N Engl J Med* 2008;359(22):2313–23.
- Diehl V. Hematology. Are macrophages the bad guys in Hodgkin lymphoma? *Nat Rev Clin Oncol* 2010;7(6):301–2.
- Gallamini A, Hutchings M, Rigacci L, Specht L, Merli F, Hansen M, 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(24):3746–52.
- Moskowitz CH, Schöder H, Teruya-Feldstein J, Sima C, Iasonos A, Portlock CS, et al. Risk-adapted dose-dense immunochemotherapy determined by interim FDG-PET in advanced-stage diffuse large B-cell lymphoma. *J Clin Oncol* 2010;28(11):1896–903.
- Haioun C, Itti E, Rahmouni A, Brice P, Rain JD, Belhadj K, et al. [18F]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):1376–81.
- Andre MPE, Reman O, Fédérico M, Brice P, Brusamolino E, Girinski T, et al. First report on the H10 EORTC/GELA/IIL Randomized Intergroup Trial On Early FDG-PET Scan Guided Treatment Adaptation Versus Standard Combined Modality Treatment in Patients with Supra-Diaphragmatic Stage I/II Hodgkin's Lymphoma, for the Groupe d'Etude Des Lymphomes De l'Adulte (GELA), European Organisation for the Research and Treatment of Cancer (EORTC) Lymphoma Group and the Intergruppo Italiano Linfomi (IIL). *ASH Annu Meet Abstr* 2009;114(22):97.
- Horning SJ, Juweid ME, Schöder H, Wiseman G, McMillan A, Swinnen LJ, 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): 775–7.
- Itti E, Juweid M, Haioun C, Yeddes I, Hamza F, El Bez I, et al. Improvement of early FDG-PET interpretation in diffuse large B-cell lymphoma (DLBCL): importance of the reference background. *J Nucl Med* 2010;51:97P–8P.
- Meignan M, Gallamini A, Haioun C. Report on the First International Workshop on Interim-PET-Scan in Lymphoma. *Leuk Lymphoma* 2009;50(8):1257–60.