Interim PET in Hodgkin Lymphoma: Is It So Useless?

TO THE EDITOR: We read with interest the above letter from Adams and Kwee when it was published online ahead of print, and we think it important to comment on the strong and recognized impact of interim PET in the management of Hodgkin lymphoma. More than 2,000 patients with advanced disease were included in 3 well-designed prospective trials, in which patients who—after 2 cycles of ABVD (doxorubicin, bleomycin, vinblastine, dacarbazine)—were PET-positive (Deauville score $\geq$ 4) were escalated to BEACOPP (bleomycin, etoposide, adriamycin, cyclophosphamide, vincristine, procarbazine, and prednisone) whereas those who were PET-negative continued with the standard regimen (4 cycles of ABVD) (*1–3*). The PET positivity rate using this Deauville score as the cutoff was similar across these 3 studies, as was the outcome of both PET-positive and PET-negative patients. The study of the Southwest Oncology Group (*1*) enrolled 336 patients, 18% of whom were PET-positive. The "Response Adapted Therapy in Advanced Hodgkin Lymphoma" study (*2*) enrolled 1,119 patients, 25% of whom were PET-positive. The Italian GITIL/FIL HD0607 trial (*3*) enrolled 773 patients, 19% of whom were PET-positive. Overall, the 2-y progression-free survival of the PET-positive patients ranged between 64% and 67%. PET-negative patients who followed the standard regimen had a 2-y progression-free survival ranging between 82% and 89%, higher than the progression-free survival of the whole population.

A major advantage of this strategy is that PET-negative patients (>80% of the total population) can be spared from the adverse effects of BEACOPP. As pointed out by Adams and Kwee, none of the interim ^{18}F-FDG PET/CT–adapted trials that have been performed so far had a control arm; that is, none continued standard ABVD in PET-positive patients. Apart from the fact that this arm would have been ethically difficult to defend, series have found 2-y progression-free survivals of 12%–27% in advanced-stage Hodgkin lymphoma patients PET-positive after 2 ABVD cycles—much lower than the 64%–67% found when these patients were intensified with BEACOPP (*4,5*), with this strategy therefore clearly having an advantage in improving outcome and decreasing the number of events. This finding was further confirmed by the results of the EORTC/LYSA/FIL H10 trial (*6*), which included 1,950 patients with early-stage Hodgkin lymphoma randomized between an experimental arm and a standard arm. Patients PET-positive after 2 ABVD cycles had a 5-y progression-free survival of 77.4% in the experimental arm that received standard ABVD plus involved-node radiotherapy, and survival improved to 90.6% in the experimental arm that received BEACOPP escalation plus involved-node radiotherapy (hazard ratio, 0.42; 95% confidence interval, 0.23–0.74; $P = 0.002$).

A deescalation trial (*7*) in 823 patients with advanced Hodgkin lymphoma (AHL2011 LYSA trial) has also recently confirmed the major role of an interim PET–guided strategy. Patients PET-negative after 2 BEACOPP escalations were moved to ABVD. Their progression-free survival was not inferior to that of the standard arm, in which BEACOPP was continued. Thus, deescalation avoids the toxic effect of BEACOPP while producing the same outcome.

In view of the cited evidence, we strongly disagree with Adams and Kwee and support the recommended role of interim PET in Hodgkin lymphoma management, which allows better tailoring of treatment to individual patients (*8*).

REFERENCES

1. Press OW, Li H, Schoder H, et al. US intergroup trial of response-adapted therapy for stage III to IV Hodgkin lymphoma using early interim fluorodeoxyglucose-positron emission tomography imaging: Southwest Oncology Group S0816. *J Clin Oncol.* 2016;34:2020–2027.
2. Johnson P, Federico M, Kirkwood A, et al. Adapted treatment guided by interim PET-CT scan in advanced Hodgkin's lymphoma. *N Engl J Med.* 2016;374:2419–2429.
3. Gallamini A, Rossi A, Patti C, et al. Interim PET-adapted chemotherapy in advanced Hodgkin lymphoma (HL): results of the second interim analysis of the Italian GITIL/FIL HD0607 trial [abstract]. *Hematol Oncol.* 2015;33 (suppl):163.
4. Gallamini A, Hutchings M, Ritacco L, et al. Early interim 2-[^{18}F]fluoro-2-deoxy-D-glucose positron emission tomography is prognostically superior to international prognostic score in advanced-stage Hodgkin's lymphoma: a report from a joint Italian-Danish study. *J Clin Oncol.* 2007;25:3746–3752.
5. Gallamini A, Barrington SF, Biggi A, et al. The predictive role of interim positron emission tomography for Hodgkin lymphoma treatment outcome is confirmed using the interpretation criteria of the Deauville five-point scale. *Haematologica.* 2014;99:1107–1113.
6. Andre MP, Girinsky T, Federico M, et al. Early positron emission tomography response–adapted treatment in stage I and II Hodgkin lymphoma: final results of the randomized EORTC/LYSA/FIL H10 trial. *J Clin Oncol.* March 14, 2017 [Epub ahead of print].
7. Casasnovas O, Brice P, Bouabdallah R, et al. Randomized phase III study comparing an early PET driven treatment de-escalation to a not PET-monitored strategy in patients with advanced stages Hodgkin lymphoma: interim analysis of the AHL2011 Lysa study [abstract]. *Blood.* 2015;126:577.
8. Johnson PW. Response-adapted frontline therapy for Hodgkin lymphoma: are we there yet? *Hematology Am Soc Hematol Educ Program.* 2016;2016:316–322.

Michel Meignan*
Anne Ségolène Cottereau
Bénédicte Deau
Salim Kanoun
Alina Berriolo-Riedinger
Olivier Casasnovas6
**Hospital Henri Mondor*
51, avenue du Marechal de Lattre de Tassigny
Creteil, 94010, France
E-mail: Michel.meignan@aphp.fr

Published online Mar. 2, 2017.
DOI: 10.2967/jnumed.117.190462

REPLY: We thank Meignan et al. for their interest in our letter, in which we document the limitations and low necessity of interim ^{18}F-FDG PET/CT imaging in lymphoma. Although Meignan et al. seem to agree with us that interim ^{18}F-FDG PET/CT has low clinical value in non-Hodgkin lymphoma (*1,2*), they disagree on the value of interim ^{18}F-FDG PET/CT in Hodgkin lymphoma, which we will therefore discuss in this reply.

Hodgkin lymphoma is usually divided into early- and advanced-stage disease, which are treated differently and have a different prognosis. Studies have shown that the value of interim ^{18}F-FDG PET/CT for predicting outcome is not homogeneous in these different disease entities. In early-stage Hodgkin lymphoma, the value of interim ^{18}F-FDG PET/CT can be considered low: patients with positive interim ^{18}F-FDG PET/CT findings have been reported to have a generally good progression-free survival (range,

30%–100%) and an excellent overall survival (range, 85.2%–100%) after standard, nonintensified, therapies, with most studies estimating long-term progression-free survival of higher than 80% (*3*). Consequently, it has to be concluded that most patients with positive interim ^{18}F-FDG PET/CT findings remain disease-free after finishing nonintensified treatment and that second- and third-line therapies can cure most patients in whom first-line therapy fails. This seriously questions whether early treatment intensification based on interim ^{18}F-FDG PET/CT results is justified. Results from the randomized EORTC/LYSA/FIL H10 trial (*4*) showed that interim ^{18}F-FDG PET/CT–positive patients treated with intensified regimens (2 cycles of ABVD [doxorubicin, bleomycin, vinblastine, dacarbazine] plus 2 cycles of BEACOPP-escalated therapy [bleomycin, etoposide, adriamycin, cyclophosphamide, vincristine, procarbazine, and prednisone] plus radiation therapy) had a better 5-y progression-free survival than those treated with standard therapy (3 cycles of ABVD plus radiation therapy) (91% vs. 77%), but overall survival did not significantly differ between these two groups (*4*), supporting our aforementioned statement. On the other hand, although the relapse rate of early-stage Hodgkin lymphoma patients with negative interim ^{18}F-FDG PET/CT results treated with standard therapies is low from an absolute point of view (*5–7*), it is actually high considering the generally good prognosis of these patients (long-term progression-free survival, ~93% (*8*)), which underlines that a negative interim ^{18}F-FDG PET/CT result cannot reliably exclude residual disease (*9*). Although randomized studies applying interim ^{18}F-FDG PET/CT–based treatment deescalation (*5,6*) have shown that interim ^{18}F-FDG PET/CT–negative patients have a generally good outcome after being treated with less intensive therapies, this is more likely a reflection of the generally good prognosis of the disease rather than of the negative predictive value of interim ^{18}F-FDG PET/CT (*10*). From a relative point of view, disease relapse occurs much more frequently in patients treated with deescalated therapy than in those who continue standard therapy despite negative interim ^{18}F-FDG PET/CT results (hazard ratios of up to 9.36 have been reported (*5,6,10*)). Considering the low positive and negative predictive values, it remains questionable whether an interim ^{18}F-FDG PET/CT–based therapeutic approach is justified in early-stage Hodgkin lymphoma. Not the least of the reasons for this question is the fact that in this disease several other biomarkers, inexpensive and easily available (e.g., the risk models of the European Organisation for Research and Treatment of Cancer, German Hodgkin Study Group, and National Comprehensive Cancer Network (*11*)), have shown prognostic value equaling that of interim ^{18}F-FDG PET/CT and may be a better surrogate for risk-adapted trials. ^{18}F-FDG PET/CT scans are expensive, expose patients to potentially harmful ionizing radiation, are uncomfortable for patients, and are not available in all institutions (particularly in non-Western countries). Therefore, it is not unlikely that interim ^{18}F-FDG PET is useless in early-stage Hodgkin lymphoma.

In advanced-stage Hodgkin lymphoma, results on the predictive value of interim ^{18}F-FDG PET/CT are less consistent. Two studies by Gallamini et al. (*12,13*) reported interim ^{18}F-FDG PET/CT to have excellent positive and negative predictive values. Patients with positive interim ^{18}F-FDG PET/CT results had a dismal progression-free survival of 12.8% in one study and 28% in the other, whereas patients with negative interim ^{18}F-FDG PET/CT results had an excellent progression-free survival of 95% in both studies after finishing standard ABVD therapy (*12,13*). However, both studies had a major methodologic flaw: histologic confirmation was available for only a small minority of cases of relapse, with relapse being documented by follow-up imaging in most cases (*12–15*). Posttreatment and follow-up ^{18}F-FDG PET/CT studies have a strikingly high number of false-positive results, as has been reported for several lymphoma subtypes (*16–19*), including Hodgkin lymphoma (*20*). Consequently, the studies by Gallamini et al. (*12,13*) are methodologically seriously biased. The predictive value of interim ^{18}F-FDG PET/CT was generally lower in other comparable studies (*21*), and 2 recent studies (*22,23*) including advanced-stage lymphoma as part of their patient population showed interim ^{18}F-FDG PET/CT to have minor or no value in predicting prognosis. Three recent studies (*24–26*) on treatment intensification in interim ^{18}F-FDG PET/CT–positive patients were published, all lacking a randomized control arm with nonintensified treatments. Consequently, the true benefit of treatment intensification in these patients could not be assessed. In addition, comparisons with historical studies that suffered from inadequate methodology and heterogeneous results are futile (*21*). We individually criticized all 3 of these studies for these issues (*27–29*). On the other hand, multiple, recently published, large-scale studies (*14,24–26*) unambiguously showed that (in contrast to the studies by Gallamini et al. (*12,13*)) a high proportion of the large group of patients with negative interim ^{18}F-FDG PET/CT results develops disease relapse during follow-up and that, therefore, a negative interim ^{18}F-FDG PET/CT result cannot exclude residual disease. In other words, most relapses occur after a negative interim ^{18}F-FDG PET/CT result (*14,24–26*). One should consider this concern when interpreting the interim results of the study by Casasnovas et al. (currently published only in abstract form (*30*)) on treatment deescalation in ^{18}F-FDG PET/CT–negative patients, which is the only randomized study yet available claiming that ^{18}F-FDG PET/CT–based treatment deescalation is feasible (excepting a study by Johnson et al. (*24*), who made a minor change in treatment—omitting bleomycin in interim ^{18}F-FDG PET/CT–negative patients—without finding a significant increase in relapse rate).

In conclusion, interim ^{18}F-FDG PET/CT is not justified in early-stage Hodgkin lymphoma. The value of positive interim ^{18}F-FDG PET/CT results in advanced-stage Hodgkin lymphoma is not well established because of methodologic issues in historical studies and the lack of a control or randomization arm in recent ^{18}F-FDG PET/CT–adapted trials. On the other hand, most disease relapses in advanced-stage Hodgkin lymphoma occur in the large group of interim ^{18}F-FDG PET/CT–negative patients, underlining that residual disease cannot be excluded and that treatment deescalation in these patients is highly questionable. Except for the interim results of the AHL2011 LYSA trial by Casasnovas et al. (*30*), there are no data confirming that treatment deescalation in interim ^{18}F-FDG PET/CT–negative advanced-stage Hodgkin lymphoma is feasible. Therefore, there is currently no convincing evidence to support use of interim ^{18}F-FDG PET/CT, either for prognostication or for treatment adaptation, in the routine clinical care of patients with early- or advanced-stage Hodgkin lymphoma.

REFERENCES

1. Adams HJ, Kwee TC. Prognostic value of interim FDG-PET in R-CHOP-treated diffuse large B-cell lymphoma: systematic review and meta-analysis. *Crit Rev Oncol Hematol.* 2016;106:55–63.
2. Adams HJ, Nievelstein RA, Kwee TC. Prognostic value of interim and end-of-treatment FDG-PET in follicular lymphoma: a systematic review. *Ann Hematol.* 2016;95:11–18.

3. Adams HJ, Kwee TC. Will treatment intensification in early-stage Hodgkin lymphoma patients with a positive interim FDG-PET improve outcome? *Pediatr Hematol Oncol.* 2016;33:1–4.
4. Andre MP, Girinsky T, Federico M, et al. Early Positron Emission Tomography Response-Adapted Treatment in Stage I and II Hodgkin Lymphoma: Final Results of the Randomized EORTC/LYSA/FIL H10 Trial. *J Clin Oncol.* 2017: JCO2016686394.
5. Raemaekers JM, André MP, Federico M, et al. Omitting radiotherapy in early positron emission tomography-negative stage I/II Hodgkin lymphoma is associated with an increased risk of early relapse: clinical results of the preplanned interim analysis of the randomized EORTC/LYSA/FIL H10 trial. *J Clin Oncol.* 2014;32:1188–1194.
6. Radford J, Illidge T, Counsell N, et al. Results of a trial of PET-directed therapy for early-stage Hodgkin's lymphoma. *N Engl J Med.* 2015;372:1598–1607.
7. Adams HJ, Nievelstein RA, Kwee TC. Prognostic value of interim FDG-PET in Hodgkin lymphoma: systematic review and meta-analysis. *Br J Haematol.* 2015;170:356–366.
8. Engert A, Plutschow A, Eich HT, et al. Reduced treatment intensity in patients with early-stage Hodgkin's lymphoma. *N Engl J Med.* 2010;363:640–652.
9. Adams HJ, Kwee TC. A negative ^{18}F-FDG-PET scan can never exclude residual disease. *Nucl Med Commun.* 2016;37:102–103.
10. Adams HJ, Kwee TC. Negative PET: no guarantee of good prognosis in Hodgkin's lymphoma. *Ann Hematol.* 2015;94:1609–1610.
11. Klimm B, Goergen H, Fuchs M, et al. Impact of risk factors on outcomes in early-stage Hodgkin's lymphoma: an analysis of international staging definitions. *Ann Oncol.* 2013;24:3070–3076.
12. Gallamini A, Hutchings M, Ritacco L, et al. Early interim 2-[^{18}F]fluoro-2-deoxy-D-glucose positron emission tomography is prognostically superior to international prognostic score in advanced-stage Hodgkin's lymphoma: a report from a joint Italian-Danish study. *J Clin Oncol.* 2007;25:3746–3752.
13. Gallamini A, Barrington SF, Biggi A, et al. The predictive role of interim positron emission tomography for Hodgkin lymphoma treatment outcome is confirmed using the interpretation criteria of the Deauville five-point scale. *Haematologica.* 2014;99:1107–1113.
14. Agostinelli C, Gallamini A, Stracqualursi L, et al. The combined role of biomarkers and interim PET scan in prediction of treatment outcome in classical Hodgkin's lymphoma: a retrospective, European, multicentre cohort study. *Lancet Haematol.* 2016;3:e467–e479.
15. Adams HJ, Kwee TC. Prevention of large-scale implementation of unnecessary and expensive predictive tests in Hodgkin's lymphoma. *Lancet Haematol.* 2017;4:e63–e64.
16. Adams HJ, Kwee TC. Proportion of false-positive lesions at interim and end-of-treatment FDG-PET in lymphoma as determined by histology: systematic review and meta-analysis. *Eur J Radiol.* 2016;85:1963–1970.
17. Avivi I, Zilberlicht A, Dann EJ, et al. Strikingly high false positivity of surveillance FDG-PET/CT scanning among patients with diffuse large cell lymphoma in the rituximab era. *Am J Hematol.* 2013;88:400–405.
18. Nakayama H, Aisa Y, Ito C, et al. Importance of histologic verification of positive positron emission tomography/computed tomography findings in the follow-up of patients with malignant lymphoma. *Clin Lymphoma Myeloma Leuk.* 2015;15:753–760.
19. Zinzani PL, Tani M, Trisolini R, et al. Histological verification of positive positron emission tomography findings in the follow-up of patients with mediastinal lymphoma. *Haematologica.* 2007;92:771–777.
20. El-Galaly TC, Mylam KJ, Brown P, et al. Positron emission tomography/computed tomography surveillance in patients with Hodgkin lymphoma in first remission has a low positive predictive value and high costs. *Haematologica.* 2012;97:931–936.
21. Adams HJ, Kwee TC. Controversies on the prognostic value of interim FDG-PET in advanced-stage Hodgkin lymphoma. *Eur J Haematol.* 2016;97:491–498.
22. Bakhshi S, Bhethanabhotla S, Kumar R, et al. Posttreatment PET/CT rather than interim PET/CT using Deauville criteria predicts outcome in pediatric Hodgkin lymphoma: a prospective study comparing PET/CT with conventional imaging. *J Nucl Med.* 2017;58:577–583.
23. Mesguich C, Cazeau AL, Bouabdallah K, et al. Hodgkin lymphoma: a negative interim-PET cannot circumvent the need for end-of-treatment-PET evaluation. *Br J Haematol.* 2016;175:652–660.
24. Johnson P, Federico M, Kirkwood A, et al. Adapted treatment guided by interim PET-CT scan in advanced Hodgkin's lymphoma. *N Engl J Med.* 2016;374:2419–2429.
25. Press OW, Li H, Schoder H, et al. US intergroup trial of response-adapted therapy for stage III to IV Hodgkin lymphoma using early interim fluorodeoxyglucose-positron emission tomography imaging: Southwest Oncology Group S0816. *J Clin Oncol.* 2016;34:2020–2027.
26. Zinzani PL, Broccoli A, Gioia DM, et al. Interim positron emission tomography response-adapted therapy in advanced-stage Hodgkin lymphoma: final results of the phase II part of the HD0801 study. *J Clin Oncol.* 2016;34:1376–1385.
27. Adams HJ, Kwee TC. Interim PET-CT scan in advanced Hodgkin's lymphoma [comment]. *N Engl J Med.* 2016;375:999.
28. Adams HJ, Kwee TC. Predictive value of interim [^{18}F]fluorodeoxyglucose-positron emission tomography in advanced-stage Hodgkin lymphoma is not well established. *J Clin Oncol.* 2017;35:370–371.
29. Adams HJ, Kwee TC. Does interim ^{18}F-FDG-PET response-adapted therapy really benefit advanced-stage Hodgkin lymphoma patients? *Nucl Med Commun.* 2016;37:1333–1334.
30. Casasnovas O, Brice P, Bouabdallah R, et al. Randomized phase III study comparing an early PET driven treatment de-escalation to a not PET-monitored strategy in patients with advanced stages Hodgkin lymphoma: interim analysis of the AHL2011 Lysa study [abstract]. *Blood.* 2015;126:577.

Hugo J.A. Adams*
Thomas C. Kwee
**Deventer Ziekenhuis*
Nico Bolkesteinlaan 75
7416 SE Deventer, The Netherlands
E-mail: h.j.a.adams@gmail.com

Published online Mar. 9, 2017.
DOI: 10.2967/jnumed.117.192294

A Common Mistake in Assessing the Diagnostic Value of a Test: Failure to Account for Statistical and Methodologic Issues

TO THE EDITOR: I was interested to read the paper by Anand et al. in the December 2016 edition of *The Journal of Nuclear Medicine* (*1*). The purpose of the authors was to assess the impact of variability in scanning speed and in vendor-specific γ-camera settings on the reproducibility and accuracy of the automated bone scan index (BSI) (*1*). They measured reproducibility as the absolute difference between repeated BSI values, and they measured accuracy as the absolute difference between observed BSI values and phantom BSI values. Descriptive statistics were used to compare the generated data.

Reproducibility (reliability) and accuracy (validity), as two completely different methodologic issues, should be assessed using appropriate tests. It is crucial to be aware that, regarding reliability, one should use the intraclass correlation coefficient for quantitative variables and the weighted κ-test for qualitative variables. However, regarding validity, one should use the interclass correlation coefficient (Pearson r) for quantitative variables whereas the most appropriate tests for qualitative variables may include sensitivity, specificity, positive and negative predictive value, positive and negative likelihood ratio, diagnostic accuracy, and odds ratio. Moreover, in analyzing reliability, one should apply an individual-based approach using single-measure intraclass correlation coefficient agreement because applying a global-average approach (absolute difference) can be misleading. A test may indicate high validity, yet there may be no reliability at all (*2–8*).

Anand et al. enrolled 25 patients in each of 3 groups and observed a significantly lower reproducibility for group 2 (mean ± SD, 0.35 ± 0.59) than for group 1 (0.10 ± 0.13; $P < 0.0001$) or group 3 (0.09 ± 0.10; $P < 0.0001$). However, no significant difference in reproducibility was observed between group 3 and group 1 ($P = 0.388$) (*1*). Statistical significance and clinical