Original Article

Acta Radiologica
0(0) 1–7

DOI: 10.1177/0284185117751279
journals.sagepub.com/home/acr
SAGE

Intra- and inter-observer agreement in the visual interpretation of interim 18F-FDG PET/CT in malignant lymphoma: influence of clinical information

Maya Kato Arimoto[1], Yuji Nakamoto[1], Tatsuya Higashi[2,3], Takayoshi Ishimori[1,4], Mana Ishibashi[5,6] and Kaori Togashi[1]

Abstract

Background: Interim PET/CT is widely performed in lymphoma patients in clinical practice and clinical trials. Visual assessment using a 5-point scale is proposed for PET/CT interpretation, but intra- and inter-observer variation is not fully investigated.

Purpose: To investigate intra- and inter-observer variations in the reporting of interim positron emission tomography/computed tomography (PET/CT) in lymphoma patients, and the influence of clinical information on the interpretation.

Material and Methods: Three expert readers from different institutions interpreted interim PET/CT images of 42 consecutive patients with malignant lymphoma twice, with and without clinical information. The intra- and inter-observer agreements were calculated using the kappa statistic on a patient and a region basis.

Results: On a patient basis, intra-observer agreement, inter-observer agreement without information, and inter-observer agreement with information were within the ranges 0.48–0.62, 0.51–0.62, and 0.42–0.76, respectively. In the evaluation of lymph nodes, intra-observer agreement, inter-observer agreement without information, and inter-observer agreement with information were within the ranges 0.78–0.92, 0.80–0.82, and 0.77–0.83, respectively. Observer agreements were in almost perfect to substantial agreement categories for most lymphatic organs, but were generally low for the other organs.

Conclusion: The intra- and inter-observer agreements in evaluating interim PET/CT were relatively low for extranodal lesions, but they were substantial to almost perfect when interpreting nodal regions in malignant lymphoma, irrespective of the provision of clinical information, although memory at the first interpretation might have affected the intra-observer results.

Keywords
Malignant lymphoma, interim, FDG-PET, PET/CT, observer variation

Date received: 30 June 2017; accepted: 8 December 2017

Introduction

Positron emission tomography (PET)/computed tomography (CT) with 18F-fluorodeoxyglucose (FDG) has been widely used for staging, re-staging, and the monitoring of responses to therapy in malignant lymphoma (1–6). Interim PET or PET/CT, performed early during chemotherapy, has been increasingly applied for malignant lymphoma such as Hodgkin lymphoma (HL) or diffuse large B-cell lymphoma (DLBCL), as a surrogate marker for chemosensitivity.

[1]Department of Diagnostic Imaging and Nuclear Medicine, Kyoto University, Graduate School of Medicine, Kyoto, Japan
[2]Shiga Medical Center Research Institute, Shiga, Japan
[3]National Institute of Radiological Sciences, Chiba, Japan
[4]Department of Radiology, Kurashiki Central Hospital, Okayama, Japan
[5]Department of Nuclear Medicine and Tracer Kinetics, Osaka University Graduate School of Medicine, Osaka, Japan
[6]Division of Radiology, Department of Pathophysiological and Therapeutic Science, Faculty of Medicine, Tottori University, Tottori, Japan

Corresponding author:
Yuji Nakamoto, Department of Diagnostic Imaging and Nuclear Medicine, Kyoto University Graduate School of Medicine, 54 Shogoin Kawahara-cho, Sakyo-ku, Kyoto 606-8507, Japan.
Email: ynakamo1@kuhp.kyoto-u.ac.jp

Better two-year progression-free survival of interim PET-negative patients than that of PET-positive patients is reported in HL and DLBCL (7–9). However, reports on the prognostic value of interim PET in other lymphoma subtypes, such as follicular lymphoma or T-cell lymphoma, are still limited (10–13).

With the increasing use of interim PET in clinical settings and multicenter trials, the reproducibility of visual assessments is very important. To standardize the visual interpretation procedures for interim PET, the International Workshop on Interim PET in Lymphoma proposed the simple and easily reproducible "Deauville" 5-point scale, which uses normal liver and mediastinal uptake as reference values (14). Good inter-observer agreement has been reported when the Deauville criteria are used for HL (15); however, data on observer agreements on interim PET for other lymphoma subtypes, including various levels of FDG uptake and sites of involvement, are currently not available (16). Additionally, although clinical information may affect the reproducibility of visual interpretations, it remains unknown how clinical information at interpretation affects intra- and inter-observer variation in the evaluation of interim PET/CT.

In this study, we conducted a reporting exercise on interim PET/CT in lymphoma patients, to reveal patterns of observer agreement and disagreement on patient-based and region-based analyses and the influence of clinical information on visual interpretation.

Material and Methods

Study design

Three board-certified nuclear medicine physicians (Readers A, B, and C with 18, 5, and 13 years of experience of PET or PET/CT interpretation, respectively) from three different institutions read the datasets of interim PET/CT obtained at another institute, so that none of the readers had previously seen the images. They interpreted the test dataset twice, without and with clinical information. Then, intra- and inter-observer agreements were assessed.

Patients

After excluding one patient where the involved site was resected before chemotherapy, the test dataset for the reporting exercise was obtained from 42 consecutive patients with biopsy proven malignant lymphoma (male:female = 12:30; age range = 8–82 years; mean age = 53 years) who underwent a PET/CT scan for early response assessment after two to four cycles of chemotherapy. For patients who had more than one interim PET scan during the period, only the first scan of each patient was retrieved. The pathological classification of patients is shown in Table 1. This retrospective study was approved by the ethics committee of our institute, and written informed consent was obtained from all patients for access to their data.

Table 1. Pathological classifications.

Classification	n
Hodgkin lymphoma	4
Non-Hodgkin lymphoma	38
Diffuse large B-cell	20
Follicular	4
Mantle cell	2
Primary mediastinal large B-cell lymphoma	1
Burkitt	1
Angioimmunoblastic T-cell lymphoma	2
T-cell lymphoma	1
Anaplastic large cell	1
Others	6

18F-FDG PET/CT scanning

PET/CT scanning was performed using a combined PET/CT scanner (Discovery ST Elite; GE Healthcare, Waukesha, WI, USA). Each patient fasted for at least 4 h before intravenous administration of 3.7 MBq/kg body weight of FDG. The administered dose of FDG for children (aged < 19 years) was determined in accordance with the revised European Association of Nuclear Medicine pediatric dosage card (17). Approximately 1 h after administration of 18F-FDG, whole-body PET was performed, with 2–3 min per bed position. The scanning range was from the upper thigh to the skull. CT data were used for attenuation correction, and images were reconstructed using the VUE Point Plus (GE Healthcare) three-dimensional iterative reconstruction algorithm. Scans were anonymized and saved in the Digital Imaging and Communications in Medicine (DICOM) format for loading onto image viewers at other institutions.

18F-FDG PET/CT interpretation

All physicians used the same image viewer (Vox-base, J-mac System, Sapporo, Japan) for interpretation. The viewer provides axial, coronal, and sagittal images of PET, CT, and PET-CT fusion. Standardized uptake value (SUV) scale and CT window can be changed manually. Maximum intensity projection (MIP) images are generated and SUV can be measured if needed, although quantitative analysis was not performed in this study.

Each observer interpreted the anonymized datasets twice. First, they visually interpreted the test datasets without clinical information (A1, B1, and C1), except for the patients' age. After the readers finished the first interpretation, they interpreted the test datasets again with clinical information (A2, B2, and C2). The clinical information included the PET/CT or CT image before initiation of chemotherapy, sites of lymphoma involvement before therapy, pathological classification of lymphoma, chemotherapy regimen, clinical symptoms at the time of the PET/CT examination, and any history of granulocyte stimulating factor (G-CSF) administration within one month of the PET/CT examination. Time duration between two series of interpretations by each reader was more than one week to avoid bias caused by memory.

The observers used a standardized form for rating individual nodal groups, lymphatic organs, and extranodal organs. Nodal groups included cervical, axillary, infraclavicular, mediastinal, hilar, para-aortic, pelvic, mesenteric, and inguinal. Nine nodal regions (five regions were bilateral: cervical, axillary, infraclavicular, pelvic, and inguinal) were scored with a 5-point scale that referred to the normal liver and mediastinal uptake: 1 = no uptake above background; 2 = uptake equal to or under mediastinum; 3 = uptake over mediastinum but equal to or under liver; 4 = uptake moderately increased compared with the liver; and 5 = uptake markedly increased compared with the liver (15). Lymphatic organs included tonsil and spleen, while extranodal sites included bone marrow/bone, lung, liver, stomach, and bowel. Lymphatic organs and extranodal sites were scored using a 3-point scale (0 = negative; 1 = equivocal; 2 = positive) regarding lymphoma involvement. Observers were asked to record any abnormal uptake suspicious of lymphoma involvement, in the other sites than listed above.

Statistical analysis

The intra-observer variations between the first and second interpretations, and the inter-observer variations between the readers for the first and second interpretations, were evaluated using the non-weighted Kappa statistic on a region and a patient basis. On region-based analyses, Kappa statistics were calculated according to negative (score of 1, 2, or 3) or positive (score of 4 or 5) results for each lymph node region, and to negative (score of 0 or 1) or positive (score of 2) results for lymphatic organs and extranodal sites. On a patient-based analyses, patients with at least one positive region were regarded as positive. Kappa values of > 0.81, 0.61–0.80, 0.41–0.60, 0.21–0.40, and 0–0.20 were defined as indicating almost perfect, substantial, moderate, fair, and slight agreement, respectively (18). Analyses were performed using JMP Pro 11.0 software (SAS Institute Inc., Cary, NC, USA).

Results

Per-patient analysis

The percentages of patients whose PET scans were diagnosed as positive by the different readers were as follows: Reader A1, 40.5%; Reader A2, 57.1%; Reader B1, 40.5%; Reader B2, 52.4%; Reader C1, 50.0%; Reader C2, 50.0%. Intra-observer agreement on the two intra-patient diagnoses with and without information was in the range of 0.49–0.62 (Table 2). Inter-observer agreements were in the range of 0.51–0.62 without information, and 0.42–0.76 with information (Suppl. Table 1), evaluated as moderate to substantial consistency.

Lymph node regions

Evaluation of the scores for each nodal site showed intra-observer agreements in the range of 0.78–0.92 (Table 2), and inter-observer agreements in the range of 0.80–0.82 without information and 0.77–0.83 with information (Suppl. Table 1). Both the intra- and inter-observer agreements for most nodal sites showed substantial to almost perfect results, except for the cervical nodes (0.59, B1 vs. C1), para-aortic nodes (0.55, B1 vs. B2; 0.55, C2 vs. A2), and mesenteric nodes (−0.02, C2 vs. A2).

Lymphatic organs

Most of the lymphatic organs, including tonsil and spleen, showed substantial to almost perfect agreement, except for the agreements between C2 and A2 (0.48) for tonsil.

Extranodal organs

Intra-observer agreements for evaluation of the bone marrow/bone were in the range of 0.43–0.83 (Table 2), while inter-observer agreements were in the range of 0.78–0.90 without information and 0.69–0.85 with information (Suppl. Table 1). Intra-observer agreements for evaluation of the lung were in the range of 0.36–0.88 (Table 2), while inter-observer agreements were in the range of 0.38–0.66 without information and 0.64–0.73 with information (Suppl. Table 1). Intra-observer agreements for evaluation of the liver were in the range of 0.66–1.00 (Table 2), while inter-observer agreements were in the range of -0.02–0.66 without information, to a perfect agreement of 1.00 with information (Suppl. Table 1). Intra-class

Table 2. Intra-observer variations with and without access to patient information.

	A1* vs. A2†	B1* vs. B2†	C1* vs. C2†
Patient	0.49 (0.24–0.74)	0.48 (0.22–0.74)	0.62 (0.38–0.86)
Nodal sites			
Cervical	0.85 (0.68–1.00)	0.77 (0.57–0.96)	0.93 (0.78–1.00)
IC	1.00 (1.00–1.00)	0.74 (0.39–1.00)	0.88 (0.66–1.00)
Axillary	0.66 (0.04–1.00)	1.00 (1.00–1.00)	N.A.
Mediastinal	0.66 (0.35–0.97)	0.73 (0.44–1.00)	1.00 (1.00–1.00)
Hilar	0.90 (0.69–1.00)	0.84 (0.55–1.00)	1.00 (1.00–1.00)
Paraaortic	0.67 (0.36–1.00)	0.55 (0.15–0.94)	0.88 (0.64–1.00)
Pelvic	1.00 (1.00–1.00)	1.00 (1.00–1.00)	0.85 (0.56–1.00)
Mesentery	0.66 (0.03–1.00)	0.66 (0.03–1.00)	N.A.
Inguinal	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)
Total	0.83 (0.74–0.92)	0.78 (0.68–0.88)	0.92 (0.85–0.99)
Lymphatic organs			
Tonsil	0.66 (0.03–1.00)	0.79 (0.39–1.00)	0.79 (0.39–1.00)
Spleen	1.00 (1.00–1.00)	0.84 (0.55–1.00)	0.84 (0.55–1.00)
Extranodal sites			
BM/Bone	0.83 (0.60–1.00)	0.43 (0.06–0.79)	0.55 (0.21–0.89)
Lung	0.48 (−0.12–1.00)	0.36 (−0.20–0.92)	0.88 (0.64–1.00)
Liver	1.00 (1.00–1.00)	0.66 (0.03–1.00)	N.A.
S/B	N.A.	N.A.	N.A.

Data are ICC with 95% confidence intervals in parentheses.
*Interpretation without information
†Interpretation with information.
IC, infraclavicular; BM, bone marrow; S/B, stomach/bowel; N.A., not applicable (no findings from at least one observer).

correlation coefficients (ICCs) were not calculated for gastric or bowel regions because each observer evaluated all cases as negative. Some of the ICCs were not calculated, as the C2 evaluations for liver involvement were all negative. Suppl. Table 2 lists the positive rates for each region according to the three readers.

In 11 of 42 cases, at least one of the observers reported positive results in sites other than lymph nodes, tonsils, spleen, bone/bone marrow, lungs, liver, or stomach/bowel. Table 3 details 11 cases with positive findings in sites other than those listed above. The following organs were reported as positive by at least one observer: muscles (neck, arms, legs, or back), paranasal sinus, pleura, thymus, spinal cord, and testis. Intra- and inter-observer agreements for evaluations of these sites were generally low.

Representative images are presented in Figs. 1 and 2.

Discussion

In this study, substantial to almost perfect intra- and inter-observer agreements were observed in the evaluation of lymph node uptake, regardless of clinical information. Perfect agreements ($\kappa = 1.00$) were recorded in the assessment of inguinal lymph nodes surrounded mainly by fat tissue with low physiological uptake. In a past study on intra- and inter-observer variations in the PET/CT staging of malignant lymphoma, the agreement was lower in the evaluation of mesenteric lymph nodes ($\kappa = 0.61$–0.67) than in the other lymph node regions, owing to physiological uptake into the intestine (16). Similarly, we found that agreements were lower in para-aortic ($\kappa = 0.55$–1.00) and mesenteric ($\kappa = -0.02$–0.66) lymph node regions, which may be due to physiological uptake in the abdominal organs.

In contrast to lymph node regions, there are no definite criteria to assess tonsils, spleen, bone marrow, or the other extranodal sites in interim PET, though international recommendations stated that areas with high physiological uptake involved of initial disease are regarded negative if the uptake is lower than the surrounding normal tissue (6). In a previous report on PET staging, 78% of Hodgkin and non-Hodgkin lymphoma patients showed a concordance between increased bone marrow uptake and iliac-crest bone marrow biopsy (19). Khan et al. reported that routine PET/CT staging identified all clinically relevant bone marrow involvement in DLBCL (20), while another report on PET staging in HL demonstrated that diffusely increased bone marrow uptake showed no

significant correlation with bone marrow infiltration (21). This latter study also reported that splenic involvement in HL patients is more frequently associated with diffuse uptake than with focal uptake. Asymmetrical increased uptake in the tonsils or focal uptake in the spleen or bone marrow may be associated with lymphoma involvement, but these findings are not specific. Assessment of lymphoma involvement during chemotherapy is therefore a challenging issue, because the uptake caused by inflammation or increased hematopoiesis may overlap with the uptake in residual tumor. In this situation, clinical information regarding treatment induced activated bone marrow will probably help.

Table 3. Cases with positive findings in the other extranodal sites.

Case no.	A1*	A2†	B1*	B2†	C1*	C2†
2	R arm	R arm			R arm	
3	Neck					
5		R leg				
9		R arm		L arm		L arm
12	Pleura	Pleura				
16	Testis				Spinal cord	
20	Back					
23	Thymus					
26				Paranasal sinus		
29	R leg		R leg	R leg		
35			L arm			

*Interpretations without access to patient information.
†Interpretation with access to patient information.
R, right; L, left.

In this study, each reader indicated suspicious sites of lymphoma involvement in a variety of extranodal organs, and this obviously lowered observer agreements on a per-patient basis. Suzuki et al. reported significant inconsistencies between readers in the selection and measurement of lesions, and in the tracking of new lesions and attention to non-target lesions (22). In our study, two of three readers showed increase of the positive rate on a patient basis in the second interpretation, probably because readers paid more attention to target lesions suggested by the clinical information. However, observer agreements did not improve with clinical information. As lymphoma can involve almost all organs, the reader should be sensitive enough to pick up suspicious lesions, but false positive results caused by inflammation, physiological uptake, or co-existing FDG-avid lesions should be carefully avoided (23). Consensus reading involving more than two readers is a known method to improve specificity (15). Careful

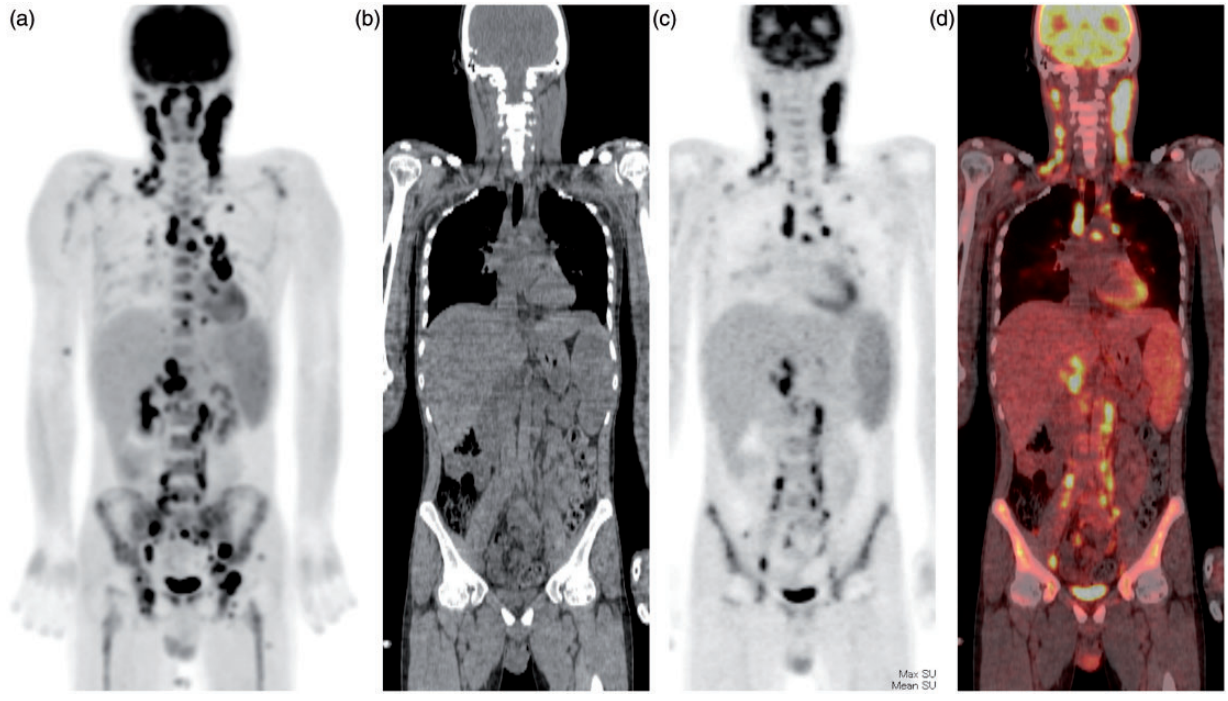

Fig. 1. A 26-year-old man with Hodgkin lymphoma. A maximum intensity projection image (a), and coronal images of CT (b), PET (c), and fusion (d) reveal multiple nodal swelling and splenomegaly with increased uptake after two cycles of chemotherapy. All readers interpreted nodal, spleen, and bone marrow involvement as positive, regardless of clinical information.

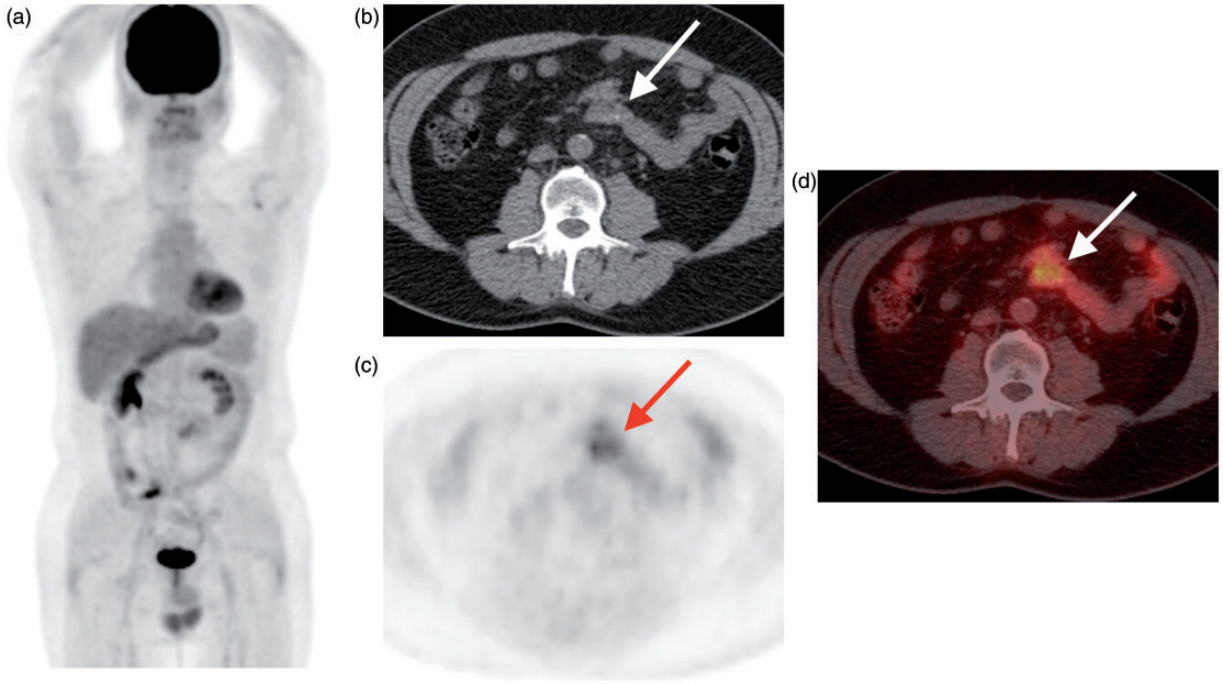

Fig. 2. A 59-year-old man with follicular lymphoma. A maximum intensity projection image (a), axial images of CT (b), PET (c), and fusion (d) demonstrate a mesenteric mass with moderate uptake (arrows) (SUVmax = 3.5) after two cycles of chemotherapy. All readers interpreted mesenteric region as negative without information; readers B and C interpreted it as positive after clinical information of primary site (mesentery) was provided.

interpretation of CT in addition to PET, would be another method to improve specificity. Interpretation with clinical information and previous investigations improve specificity further.

Once target lesion is identified by an expert reader, quantitative evaluation would also improve observer agreements. Barrington et al. reported that quantitative measures can be used to improve on visual analysis for response assessment in DLBCL. Semi-quantitative methods to extend Deauville 5-point scores were developed and higher inter-observer agreements are expected (24).

This study has several limitations. The number of patients and the number of readers included in this study were small, and further studies including more patients and readers are warranted. The proportions of classically FDG-avid lymphomas, such as HL, DLBCL, or FL, in the study population were relatively small, and that may have lowered observer agreements. There may have been a bias from memory that affected on the second interpretation, because of the short interval of two series of interpretations. Histopathological confirmation was not obtained from "positive" lesions, because additional therapy followed interim PET/CT in most cases.

In conclusion, further standardization of interpretation criteria for extranodal sites is required, because the intra- and inter-observer agreements for the interpretation of extranodal sites were relatively low. However, whether or not clinical information was provided, the intra- and inter-observer agreements between experienced readers were substantial to almost perfect for the interpretation of nodal regions in interim PET/CT in malignant lymphoma, although biases caused by memory at the first interpretation cannot be excluded. This finding is especially useful for multicenter trials.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

1. Cheson BD, Fisher RI, Barrington SF, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. J Clin Oncol 2014;32: 3059–3067.
2. Hutchings M, Barrington SF. PET/CT for therapy response assessment in lymphoma. J Nucl Med 2009; 50(Suppl): 21S–30S.

3. Seam P, Juweid ME, Cheson BD. The role of FDG-PET scans in patients with lymphoma. Blood 2007;110: 3507–3516.
4. Wahl RL, Jacene H, Kasamon Y, et al. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. J Nucl Med 2009;50(Suppl 1): 122S–150S.
5. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. J Clin Oncol 2007;25:579–586.
6. Barrington SF, Mikhaeel NG, Kostakoglu L, et al. Role of imaging in the staging and response assessment of lymphoma: consensus of the International Conference on Malignant Lymphomas Imaging Working Group. J Clin Oncol 2014;32:3048–3058.
7. Kostakoglu L, Cheson BD. Current role of FDG PET/CT in lymphoma. Eur J Nucl Med Mol Imaging 2014;41: 1004–1027.
8. Kostakoglu L, Gallamini A. Interim 18F-FDG PET in Hodgkin lymphoma: would PET-adapted clinical trials lead to a paradigm shift? J Nucl Med 2013;54:1082–1093.
9. Barrington SF, Kirkwood AA, Franceschetto A, et al. PET-CT for staging and early response: results from the Response Adapted Therapy in Advanced Hodgkin Lymphoma study. Blood 2016;127:1531–1538.
10. Adams HJ, Nievelstein RA, Kwee TC. Prognostic value of interim and end-of treatment FDG-PET in follicular lymphoma: systematic review. Ann Hematol 2016;95: 11–18.
11. El-Galaly TC, Pedersen MB, Hutchings M, et al. Utility of interim and end-of-treatment PET/CT in peripheral T-cell lymphomas: a review of 124 patients. Am J Hematol 2015;90:975–980.
12. Khong P-L, Huang B, Phin Lee EY, et al. Midtreatment 18F-FDG PET/CT scan for early response assessment of SMILE therapy in natural killer/T-cell lymphoma: a prospective study from a single center. J Nucl Med 2014;55: 911–916.
13. Li Y-J, Li Z-M, Xia X-Y, et al. Prognostic value of interim and posttherapy 18F-FDG PET/CT in patients with mature T-cell and natural killer cell lymphomas. J Nucl Med 2013;54:507–515.
14. Meignan M, Gallamini A, Haioun C, et al. Report on the Second International Workshop on interim positron emission tomography in lymphoma held in Menton, France, 8–9 April 2010. Leuk Lymphoma 2010;51: 2171–2180.
15. Barrington SF, Qian W, Somer EJ, et al. Concordance between four European centres of PET reporting criteria designed for use in multicenter trials in Hodgkin lymphoma. Eur J Nucl Med Mol Imaging 2010;37:1824–1833.
16. Hofman MS, Smeeton NC, Rankin SC, et al. Observer variation in interpreting 18F-FDG PET/CT findings for lymphoma staging. J Nucl Med 2009;50:1594–1597.
17. Stauss J, Franzius C, Pfluger T, et al. Guidelines for 18F-FDG PET and PET-CT imaging in pediatric oncology. Eur J Nucl Med Mol Imaging 2008;35:1581–1588.
18. Kundel HM, Polansky M. Measurement of observer agreement. Radiology 2003;228:303–308.
19. Carr R, Barrington SF, Madan B, et al. Detection of lymphoma in bone marrow by whole-body positron emission tomography. Blood 1998;91:3340–3346.
20. Khan AB, Barrington SF, Mikhaeel NG, et al. PET-CT staging of DLBCL accurately identifies and provides new insight into the clinical significance of bone marrow involvement. Blood 2013;122:61–67.
21. Salaun PY, Gastinne T, Bodet-Milin C, et al. Analysis of 18F-FDG PET diffuse bone marrow uptake and splenic uptake in staging of Hodgkin's lymphoma: a reflection of disease infiltration or just inflammation? Eur J Nucl Med Mol Imaging 2009;36:1813–1821.
22. Suzuki C, Torkzad MR, Jacobsson H, et al. Interobserver and intraobserver variability in the response evaluation of cancer therapy according to RECIST and WHO-criteria. Acta Oncol 2010;49:509–514.
23. Boellaard R. Need for standardization of 18F-FDG PET/CT for treatment response assessments. J Nucl Med 2011;52(Suppl 2): 93S–100S.
24. Hasenclever D, Kurch L, Mauz- Körholz C, et al. qPET – a quantitative extension of the Deauville scale to assess response in interim FDG-PET scans in lymphoma. Eur J Nucl Med Mol Imaging 2014;41:1301–1308.

Suppl. Table 1. Interobserver variations with or without access to patient information

	A1 vs. B1	B1 vs. C1	C1 vs. A1	A2 vs. B2	B2 vs. C2	C2 vs. A2
Patient	0.51 (0.24–0.77)	0.62 (0.39–0.85)	0.62 (0.34–0.85)	0.42 (0.15–0.70)	0.76 (0.57–0.96)	0.48 (0.21–0.74)
Nodal sites						
Cervical	0.74 (0.56–0.94)	0.59 (0.34–0.84)	0.66 (0.42–0.91)	0.77 (0.56–0.99)	0.86 (0.71–1.00)	0.88 (0.71–1.00)
IC	0.74 (0.39–1.00)	0.88 (0.66–1.00)	0.85 (0.56–1.00)	1.00 (1.00–1.00)	0.74 (0.39–1.00)	0.74 (0.39–1.00)
Axillary	0.66 (0.04–1.00)	N.A.	N.A.	1.00 (1.00–1.00)	N.A.	N.A.
Mediastinal	0.91 (0.73–1.00)	1.00 (1.00–1.00)	0.91 (0.73–1.00)	0.83 (0.60–1.00)	0.73 (0.44–1.00)	0.76 (0.50–1.00)
Hilar	0.78 (0.48–1.00)	1.00 (1.00–1.00)	0.77 (0.48–1.00)	0.73 (0.37–1.00)	0.84 (0.55–1.00)	0.88 (0.64–1.00)
Paraaortic	0.77 (0.47–1.00)	0.77 (0.47–1.00)	1.00 (1.00–1.00)	0.67 (0.36–1.00)	0.63 (0.24–1.00)	0.55 (0.16–0.94)
Pelvic	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	0.85 (0.56–1.00)	0.85 (0.56–1.00)
Mesentery	0.66 (0.03–1.00)	N.A.	N.A.	0.66 (0.03–1.00)	0.66(0.03–1.00)	–0.02(–0.06–0.01)
Inguinal	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)
Total	0.81 (0.72–0.90)	0.82 (0.72–0.91)	0.80 (0.70–0.90)	0.83 (0.73–0.92)	0.79 (0.68–0.90)	0.77 (0.66–0.88)
Lymphatic organs						
Tonsil	0.79 (0.39–1.00)	0.79 (0.39–1.00)	1.00 (1.00–1.00)	0.66 (0.03–1.00)	0.79 (0.39–1.00)	0.48 (–0.11–1.00)
Spleen	0.64 (0.19–1.00)	1.00 (1.00–1.00)	0.64 (0.19–1.00)	0.79 (0.39–1.00)	1.00 (1.00–1.00)	0.79 (0.39–1.00)
Extranodal sites						
BM/Bone	0.78 (0.48–1.00)	0.88 (0.64–1.00)	0.90 (0.69–1.00)	0.69 (0.40–0.97)	0.85 (0.64–1.00)	0.85 (0.64–1.00)
Lung	0.66 (0.03–1.00)	0.64 (0.19–1.00)	0.38 (–0.16–0.91)	0.64 (0.18–1.00)	0.73 (0.37–1.00)	0.73 (0.37–1.00)
Liver	0.66 (0.03–1.00)	–0.03 (–0.08–0.01)	–0.02 (–0.06–0.01)	1.00 (1.00–1.00)	N.A.	N.A.
S/B	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.

A1, B1, and C1: Interpretation without information. A2, B2, and C2: Interpretation with information. IC: Infraclavicular. BM: Bone marrow. S/B: Stomach or bowel. N.A.: Not applicable (no findings from at least one observer).
Data are intraclass correlation coefficients with 95% confidence intervals in parentheses.

Suppl. Table 2. Numbers and percentages of positive ratings

	A1	A2	B1	B2	C1	C2
Patient	17(40.5%)	24(57.1%)	17(40.5%)	22(52.4%)	21(50.0%)	21(50.0%)
Nodal sites						
Cervical (n=84)	13 (15.5%)	10 (11.9%)	15 (17.9%)	10 (11.9%)	7 (8.3%)	8 (9.5%)
IC (n=84)	3 (3.6%)	3 (3.6%)	5 (6.0%)	3 (3.6%)	4 (4.8%)	5 (6.0%)
Axillary (n=84)	2 (2.4%)	1 (1.2%)	1 (1.2%)	1 (1.2%)	0 (0.0%)	0 (0.0%)
Mediastinal (n=42)	6 (14.3%)	8 (19.0%)	7 (16.7%)	6 (14.3%)	7 (16.7%)	7 (16.7%)
Hilar (n=42)	6 (14.3%)	5 (11.9%)	4 (9.5%)	3 (7.1%)	4 (9.5%)	4 (9.5%)
Para-aortic (n=42)	5 (11.9%)	6 (14.3%)	5 (11.9%)	5 (11.9%)	5 (11.9%)	4 (9.5%)
Pelvic (n=84)	3 (3.6%)	3 (3.6%)	3 (3.6%)	3 (3.6%)	3 (3.6%)	4 (4.8%)
Mesentery (n=42)	2 (4.8%)	1 (2.4%)	1 (2.4%)	2 (4.8%)	0 (0.0%)	1 (2.4%)
Inguinal (n=84)	2 (2.4%)	2 (2.4%)	2 (2.4%)	2 (2.4%)	2 (2.4%)	2 (2.4%)
Total (n=588)	42 (7.1%)	39 (6.6%)	43 (7.3%)	35 (6.0%)	32 (5.4%)	35 (6.0%)
Lymphatic organs						
Tonsil (n=42)	2 (4.8%)	1 (2.4%)	3 (7.1%)	2 (4.8%)	2 (4.8%)	3 (7.1%)
Spleen (n=42)	2 (4.8%)	2 (4.8%)	4 (9.5%)	3 (7.1%)	4 (9.5%)	3 (7.1%)
Extranodal sites						
BM/Bone (n=42)	6 (14.3%)	8 (19.0%)	4 (9.5%)	8 (19.0%)	5 (11.9%)	8 (19.0%)
Lung (n=42)	1 (2.4%)	3 (7.1%)	2 (4.8%)	3 (7.1%)	4 (9.5%)	5 (11.9%)
Liver (n=42)	1 (2.4%)	1 (2.4%)	2 (4.8%)	1 (2.4%)	1 (2.4%)	0 (0.0%)
S/B (n=42)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Others (n=42)	7 (16.7%)	4 (9.5%)	2(4.8%)	3 (7.1%)	2 (4.8%)	1 (2.4%)

A1, B1, and C1: Interpretation without information. A2, B2, and C2: Interpretation with information. IC: Infraclavicular. BM: Bone marrow. S/B: Stomach or bowel.

Corrigendum

Acta Radiologica
0(0) 1

DOI: 10.1177/0284185118770126
journals.sagepub.com/home/acr

Arimoto MK, Nakamoto Y, Higashi T, et al. Intra- and inter-observer agreement in the visual interpretation of interim 18F-FDG PET/CT in malignant lymphoma: influence of clinical information. Acta Radiol. Epub ahead of print January 15. DOI: 10.1177/0284185117 751279.

In this article, 'Intra-class correlation coefficients' or 'ICCs' should have read 'Kappa values' on pages 3 and 4 and Supplementary Material page 1.