References

[1] Tsujimoto Y, Cossman J, Jaffe E, et al. Involvement of the bcl-2 gene in human follicular lymphoma. Science (Washington DC) 1985;228:1440–3.
[2] Luthra R, McBride JA, Cabanillas F, et al. Novel 5′ exonuclease-based real-time PCR assay for the detection of t(14;18)(q32;q21) in patients with follicular lymphoma. Am J Pathol 1998;153:63–8.
[3] Olsson K, Gerard CJ, Zehnder J, et al. Real time t(11;14) and t(14;18) PCR assays provide sensitive and quantitative assessment of minimal residual disease (MRD). Leukemia 1999;13:1833–42.
[4] Hirt C, Dölken G. Quantitative detection of t(14;18)-positive cells in patients with follicular lymphoma before and after autologous bone marrow transplantation. Bone Marrow Transplant 2000;25:419–26.
[5] Voso MT, Pantel G, Weis M, et al. In vivo depletion of B cells using a combination of high-dose cytosine arabinoside/mitoxantrone and rituximab for autografting in patients with non-Hodgkin's lymphoma. Br J Haematol 2000;109:729–35.
[6] Volpe G, Vitolo U, Carbone A, et al. Molecular heterogenity of B-lineage diffuse large cell lymphoma. Genes Chromosomes Cancer 1996;16:21–30.
[7] Yashima A, Chihaya M, Tarusawa M, et al. Consensus strategy of real-time quantitative polymerase chain reaction for detection of minimal residual disease in non-Hodgkin lymphoma. Leuk Res 2003;27:925–34.
[8] Donovan JW, Ladetto M, Zou G, et al. Immunoglobulin heavy-chain consensus probes for real-time PCR quantification of residual disease in acute lymphoblastic leukemia. Blood 2000;95:2651–8.
[9] Verhagen OJ, Willemse MJ, Breunis WB, et al. Application of germline IGH probes in real-time quantitative PCR for the detection of minimal residual disease in acute lymphoblastic leukemia. Leukemia 2000;14:1426–35.
[10] Uchiyama M, Maesawa C, Yashima A, et al. Development of consensus fluorogenically labeled probes of the immunoglobulin heavy-chain gene for detecting minimal residual disease in B-cell non-Hodgkin lymphomas. Cancer Sci 2003;94:877–85.
[11] Nagai H, Kinoshita T, Ichikawa A, et al. Malignant lymphoma and tumor suppressor genes. J Clin Exp Hematopathol 2002;42:11–23.
[12] Kinoshita T. Epigenetic inactivation of tumor suppressor genes in hematologic malignancies. Int J Hematol 2004;80:108–19.
[13] Lo YM, Wong IHN, Zhang J, et al. Quantitative analysis of aberrant p16 methylation using real-time quantitative methylation-specific polymerase chain reaction. Cancer Res 1999;59:3899–903.
[14] Wong IH, Zhang J, Lai PB, et al. Quantitative analysis of tumor-derived methylated p16INK4a sequences in plasma, serum, and blood cells of hepatocellular carcinoma patients. Clin Cancer Res 2003;9:1047–52.
[15] Hoque MO, Topaloglu O, Begum S, et al. Quantitative methylation-specific polymerase chain reaction gene patterns in urine sediment distinguish prostate cancer patients from control subjects. J Clin Oncol 2005;23:6569–75.
[16] Hoque MO, Feng Q, Toure P, et al. Detection of aberrant methylation of four genes in plasma DNA for the detection of breast cancer. J Clin Oncol 2006;24:4262–9.
[17] Agrawal S, Unterberg M, Koschmieder S, et al. DNA methylation of tumor suppressor genes in clinical remission predicts the relapse risk in acute myeloid leukemia. Cancer Res 2007;67:1370–7.
[18] Theillet C, Lidereau R, Escot C, et al. Loss of a c-H-ras-1 allele and aggressive human primary breast carcinomas. Cancer Res 1986;46:4776–81.
[19] Fujimori M, Tokino T, Hino O, et al. Allelotype study of primary hepatocellular carcinoma. Cancer Res 1991;51:89–93.
[20] Fearon ER, Feinberg AP, Hamilton SR, et al. Loss of genes on the short arm of chromosome 11 in bladder cancer. Nature 1985;318:377–80.
[21] Li Y, Nagai H, Ohno T, et al. Aberrant DNA methylation of $p57^{KIP2}$ gene in the promoter region in lymphoid malignancies of B-cell phenotype. Blood 2002;100:2572–7.
[22] Shen L, Toyota M, Kondo Y, et al. Aberrant DNA methylation of *p57KIP2* identifies a cell-cycle regulatory pathway with prognostic impact in adult acute lymphocytic leukemia. Blood 2003;101:4131–6.
[23] Pateras IS, Apostolopoulou K, Koutsami M, et al. Downregulation of the KIP family members $p27^{KIP1}$ and $p57^{KIP2}$ by SKP2 and the role of methylation in $p57^{KIP2}$ inactivation in nonsmall cell lung cancer. Int J Cancer 2006;119:2546–56.
[24] Brakensiek K, Länger F, Kreipe H, et al. Absence of $p21^{CIP1}$, $p27^{KIP1}$ and $p57^{KIP2}$ methylation in MDS and AML. Leukemia Res 2005;29:1357–60.
[25] Saiki RK, Gelfand DH, Stoffel S, et al. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science (Washington DC) 1988;239:487–91.
[26] International non-Hodgkin's lymphoma prognostic factors project: a predictive model for aggressive non-Hodgkin's lymphoma. N Engl J Med 1993;329:987–94.
[27] Kramer MH, Hermans J, Wijburg E, et al. Clinical relevance of bcl2, bcl6, and myc rearrangements in diffuse large B-cell lymphoma. Blood 1998;92:3152–62.

Published Ahead of Print on March 9, 2009 as 10.1200/JCO.2008.16.0861
The latest version is at http://jco.ascopubs.org/cgi/doi/10.1200/JCO.2008.16.0861

JOURNAL OF CLINICAL ONCOLOGY

REVIEW ARTICLE

Fluorine-18-Fluorodeoxyglucose Positron Emission Tomography for Interim Response Assessment of Advanced-Stage Hodgkin's Lymphoma and Diffuse Large B-Cell Lymphoma: A Systematic Review

Teruhiko Terasawa, Joseph Lau, Stéphane Bardet, Olivier Couturier, Tomomitsu Hotta, Martin Hutchings, Takashi Nihashi, and Hirokazu Nagai

From the Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA; Clinical Research Center for Blood Diseases, National Hospital Organization Nagoya Medical Center; Department of Radiology, Nagoya University Graduate School of Medicine, Nagoya, Japan; Department of Nuclear Medicine, François Baclesse Center, Caen; Department of Nuclear Medicine, University of Angers, Angers, France; and the Department of Oncology and Haematology, Copenhagen University Hospital, Copenhagen, Denmark.

Submitted January 7, 2008; accepted December 18, 2008; published online ahead of print at www.jco.org on March 9, 2009.

Supported by Banyu Life Science Foundation International (H19) and the Ministry of Health, Labor, and Welfare, Japan (15-2).

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Teruhiko Terasawa, MD, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, 800 Washington St, #63, Boston, MA 02111; e-mail: tterasawa@tuftsmedicalcenter.org.

0732-183X/09/2799-1/$20.00

DOI: 10.1200/JCO.2008.16.0861

A B S T R A C T

Purpose
To systematically review the prognostic accuracy of fluorine-18-fluorodeoxyglucose positron emission tomography (FDG-PET) for interim response assessment of patients with untreated advanced-stage Hodgkin's lymphoma (HL) or diffuse large B-cell lymphoma (DLBCL).

Methods
MEDLINE, EMBASE, SCOPUS, and Biologic Abstracts were searched for relevant studies. Two assessors independently reviewed studies for inclusion and extracted data. Relevant unpublished data were requested from the investigators if unavailable from publications. A meta-analysis of the prognostic accuracy was performed.

Results
Thirteen studies involving 360 advanced-stage HL patients and 311 DLBCL patients met our inclusion criteria. Advanced-stage HL studies included few unfavorable-risk patients. DLBCL studies were heterogeneous. FDG-PET had an overall sensitivity of 0.81 (95% CI, 0.72 to 0.89) and a specificity of 0.97 (95% CI, 0.94 to 0.99) for advanced-stage HL, and a sensitivity of 0.78 (95% CI, 0.64 to 0.87) and a specificity of 0.87 (95% CI, 0.75 to 0.93) for DLBCL. Meta-regression and subgroup analyses did not identify factors that affect prognostic accuracy.

Conclusion
For low- to intermediate-risk advanced-stage HL, FDG-PET performed after a few cycles of standard chemotherapy seems to be a reliable prognostic test to identify poor responders, warranting prospective studies to assess PET-based treatment strategies. For DLBCL, no reliable conclusions can be drawn due to heterogeneity. Interim PET remains an unproven test for routine clinical practice. Its use should be reserved for research settings where treatment regimens and imaging conditions are standardized.

J Clin Oncol 27. © 2009 by American Society of Clinical Oncology

INTRODUCTION

Malignant lymphoma is the fifth most commonly diagnosed cancer in the United States.[1] With advances in treatments, Hodgkin's lymphoma (HL) and diffuse large B-cell lymphoma (DLBCL) are potentially curable lymphomas.[2,3] However, challenges remain especially in the treatment for high-risk patients,[4,5] since more than half of these patients do not achieve long-term survival with currently available standard first-line chemotherapy. A possible treatment involves intensive and toxic polychemotherapy for advanced-stage HL[6] or first-line high-dose chemotherapy with stem-cell support for DLBCL,[7] depending on individual risk of treatment failure. Therefore, better identification of poor responders to first-line therapy is important to advance risk-adapted treatment strategies.

Fluorine-18-fluorodeoxyglucose positron emission tomography (FDG-PET) is a functional imaging test that has become widely used in the management of both HL and non-Hodgkin's lymphoma (NHL).[8] Studies that assessed FDG-PET as a prognostic tool performed during chemotherapy have reported the ability to predict poor outcomes.[8] However, the studies used different design, conduct, and reporting, making interpretation of the results difficult. In particular, inclusion of heterogeneous populations with different categories of disease (eg, limited-stage *v* advanced-stage HL or DLBCL

Table 1. Studies of PET for Interim Response Assessment of Malignant Lymphoma Included in the Systematic Review

Study	Year	Country	Study Design	No. of Involved Institutions	Start of Follow-Up Period	Follow-Up (months)		Pretherapy Scan to Confirm FDG Avidity (%)
						Median	Range	
Advanced-stage HL + DLBCL								
Kostakoglu et al[32]	2006	USA	Retrospective	1	Start of therapy	21†	3-47	100
Advanced-stage HL								
Friedberg et al[33]	2004	USA	Prospective	3	Pre-therapy PET	24†	10-32	100
Hutchings et al[37]	2005	UK	Retrospective	1	Diagnosis of lymphoma	40†	6-125	100
Gallamini et al[30]	2006	Italy	Prospective	11	Diagnosis of lymphoma	20‖	2-46	100
Hutchings et al[13]	2006	Denmark	Prospective	3	Diagnosis of lymphoma	22	6-40	100
Zinzani et al[14]	2006	Italy	Prospective	1	NR	18	12-27	100
Gallamini et al[29]	2007	Italy + Denmark	Prospective	14	Diagnosis of lymphoma	26†	4-62	100
DLBCL								
Spaepen et al[34]	2002	Belgium	Prospective	1	End of therapy	36††	19-51	97†
Haioun et al[31]	2005	France	Prospective	4	Study enrollment	24†	NR	100
Mikhaeel et al[12]	2005	UK	Retrospective	1	Diagnosis of lymphoma	24†	NR	100
Fruchart et al[35]	2006	France	Prospective	1	Start of therapy	19	2-35	100
Querellou et al[38]	2006	France	Retrospective	1	Start of therapy	15†¶	9-28	100
Ng et al[36]	2007	Australia	Retrospective	1	Start of therapy	28	2-81	Partial

(continued on following page)

v other aggressive NHLs) clearly affects the clinical applicability of the study results because each category has different clinical profiles (eg, treatment strategies, response, and prognosis). In this systematic review, we assessed the prognostic accuracy of FDG-PET performed during first-line therapy to predict disease progression or relapse in patients with advanced-stage HL and DLBCL, paying particular attention to the clinical applicability of the reported results.

METHODS

Data Sources and Searches

We searched Ovid MEDLINE and EMBASE from 1966 through July 2006,[9] and PubMed from August 2006 through July 2007 without language restriction. The search strategy can be found in online-only Appendix Table A1. This search was augmented by searches of SCOPUS and Biologic Abstracts. We also examined the reference lists of eligible studies, review articles, and textbooks.

Study Selection

Two reviewers (T.T., H.N.) screened abstracts and determined eligibility. Full-text articles were reviewed when abstracts did not provide sufficient information for determination. We included studies that evaluated FDG-PET performed between the first and the fourth cycle of first-line chemotherapy for patients with advanced-stage HL or DLBCL. We included both prospective and retrospective studies, and we considered clinical follow-up with or without pathologic confirmation to be a reference standard. We included studies that evaluated at least 10 patients and included at least five patients who progressed during chemotherapy or relapsed through clinical follow-up. We accepted studies in which patients received high-dose chemotherapy followed by autologous stem cell transplantation as long as it was administered as a part of primary therapy or consolidation therapy after standard induction chemotherapy. We excluded abstracts, editorials, comments, letters, and review articles. We excluded studies that enrolled patients with HIV-associated or post-transplant lymphoproliferative disorders.

Many studies did not meet all the inclusion criteria, but did partially include a relevant patient population. For these studies, we contacted the authors for relevant individual patient or subgroup data. When there was no response after 4 weeks, another correspondence was sent. When there was no response after the third communication attempt, we considered the request rejected.

Data Extraction and Quality Assessment

Two independent, board-certified hematologists (T.T., H.N.) abstracted relevant data. We extracted patients' demographic and clinical characteristics including the International Prognostic Scores (IPS) for advanced-stage HL[4] or the International Prognostic Indexes (IPI) for DLBCL,[5] therapeutic interventions, interim PET results, and final clinical outcomes. We subdivided the treatment failures into three categories based on the relative timing to the completion of first-line therapy: during therapy, after 1 year from diagnosis or the start of therapy, and in between. When the timing of completion of first-line therapy was unclear, we arbitrarily considered the treatment period to be 6 months. We also extracted the number of cases in remission but censored from follow-up within 1 year from the start of therapy (early censoring). One nuclear medicine specialist (T.N.) evaluated the technical specification and quality of PET procedures using recommended guidelines.[10] Reviewers were not blinded to the name of the journal. Inconsistencies between reviewers were either clarified by the authors or resolved by consensus.

To evaluate the quality, applicability, and reporting of the studies, we used QUADAS, a recently proposed tool to assess the quality of studies of diagnostic accuracy included in a systematic review.[11] Details on how we scored each item can be found in online-only Appendix Table A2. We assessed only published data and did not use unpublished data because the latter was not available from all the studies.

Data Synthesis and Statistical Analysis

For each study, we constructed a 2×2 contingency table consisting of true positive (TP), false positive (FP), false negative (FN), and true negative (TN), where all patients were categorized according to whether they were PET positive or negative, and whether they experienced treatment failure. In the main analysis, we employed the entire clinical follow-up as the reference standard. In sensitivity analysis, we categorized patients using shorter clinical follow-up as the alternative reference standard to focus on very early treatment failures (only during therapy or < 6 months), or early treatment failures (< 12 months). We counted patients in remission during the specified follow-up period as no treatment failure even if they eventually experienced treatment failure thereafter. We counted early censorings as no treatment failure in the main analysis. In sensitivity analysis to explore a worst-case scenario, early censorings were excluded from the analysis, and then counted as FP if they had negative PET results and were lost to follow-up early without treatment

Table 1. Studies of PET for Interim Response Assessment of Malignant Lymphoma Included in the Systematic Review (continued)

Study	No. of Chemotherapy Cycles Before PET Scan	Duration Between Chemotherapy and PET Scan (days)	No. of Total Participants*	Women		Age (years)	
				No.	%	Median	Range
Advanced-stage HL + DLBCL							
Kostakoglu et al[32]	1	8-15 for HL, 15-22‡ for DLBCL	34§	23†	49	48.2†‖	18-76
Advanced-stage HL							
Friedberg et al[33]	3¶	NR	22	NR†	36	NR†	18-60
Hutchings et al[37]	2 or 3	8-15	28	42†	49	36.7†	15-73
Gallamini et al[30]	2	11.6‖	108	57	53	32.6‖	14-79
Hutchings et al[13]	2	8-15	46	28†	36	36	18-74
Zinzani et al[14]	2	NR	40	21	53	32	14-48
Gallamini et al[29]	2	NR	106#	127†	49	32†	14-79
DLBCL							
Spaepen et al[34]	3 or 4#	14†† or 21‡	47	18†	26	40†	3-78
Haioun et al[31]	2	13-14†† or 20-21‡	83	34†	38	53†	17-78
Mikhaeel et al[12]	2 or 3	NR	57	56†	46	55†	20-84
Fruchart et al[35]	2 or 3	12†† or 18‡	35	13†	33	56†	24-77
Querellou et al[38]	2, 3, or 4¶‡‡	15-21‡	21	NR†	33	NR†	17-75
Ng et al[36]	2, 3, or 4§§	12-14†† or 19-21‡	44	21	48	60	27-83

Abbreviations: FDG, fluorodeoxyglucose; ACVBP, doxorubicin, cyclophosphamide, vindesine, bleomycin, prednisone; CHOP, cyclophosphamide, doxorubicin, vincristine, prednisone; DLBCL, diffuse large B-cell lymphoma; HL, Hodgkin's lymphoma; NR, not reported; PET, positron emission tomography; R, rituximab.
*Only advanced-stage HL or DLBCL patients were included in this systematic review.
†Data abstracted from total participants of original report, not exclusively for relevant patient population.
‡For tri-weekly cycle chemotherapy [eg, (R-)CHOP].
§Including 10 advanced-stage HL patients and 24 DLBCL patients.
‖Mean.
¶Patients underwent PET at the midpoint of the whole chemotherapy cycles (the end of the second cycle for 4-cycle chemotherapy regimens, the third cycle for 6-cycle regimens, and the fourth cycle for 8-cycle regimens).
#Only patients not included in the previous reports[13,30] were left.
**Only patients in long-term remission.
††For bi-weekly cycle chemotherapy [eg, (R-)ACVBP].
‡‡Eleven patients underwent PET at the end of the fourth cycle.
§§Eleven patients underwent PET at the end of the fourth cycle.

failure. Three studies reported intermediate PET results as minimal residual uptake (MRU).[12–14] We considered this category negative scan in the main analysis because this was how investigators analyzed the results. In sensitivity analyses, MRU results were excluded from analysis, considered positive, considered positive in the case of treatment failure and negative in the case of continuing remission (best-case scenario), and considered negative in the case of treatment failure and positive in the case of continuing remission (worst-case scenario).

We calculated sensitivity, specificity, and likelihood ratios (LRs) for each study. For the estimation of 95% CI, we used the binominal Wilson method for sensitivity and specificity, and normal approximation for LRs. Then we combined summary statistics, 95% confidence regions of summary sensitivity and specificity, and summary receiver operating characteristic (ROC) curves by the hierarchical SROC method,[15] which takes into account both within-study and between-studies variation. We fitted the model by using maximum likelihood estimation implemented in the GLLAMM algorithm[16] in STATA (version 9.2; Stata Corp, College Station, TX), and depicted the summary ROC curves and confidence regions for summary sensitivity and specificity.[17] We estimated the Q* statistic,[15] the point on the curve where sensitivity equals specificity, as global measures for the summary ROC.

To explore heterogeneity, we performed subgroup analyses by visual assessment of ROC plots and univariate meta-regression analyses. In the meta-regression, we incorporated study design or clinical characteristics as covariates into the bivariate model using Meta-Analyst (Tufts Medical Center, Boston, MA). Our preplanned analyses included characteristics of study design (prospective *v* retrospective), whether studies included more than 10 patients with treatment failure, rates of treatment failure, adoption of combined FDG-PET and computed tomography (FDG-PET/CT), the mean number of chemotherapy cycles before PET, timing of PET scan after the administration of chemotherapy, percentage of high or high-intermediate risk for DLBCL, and percentage of rituximab (R) use for DLBCL. We also performed posthoc analyses on the use of high-dose chemotherapy. Two-sided *P* values lower than .05 were considered to be statistically significant.

RESULTS

Search Results

Online-only Appendix Figure A1 summarizes the search results. We retrieved 23 full reports for further review and contacted nine authors for additional data. We excluded three studies that presented the same participants as previous reports,[18–20] three studies that did not provide information to calculate prognostic accuracy,[21–23] two studies that adopted nondedicated PET scanner,[24,25] one study with fewer than 10 relevant participants,[26] one study with fewer than five patients who progressed or relapsed,[27] and one study that evaluated patients during salvage therapy.[28] One study[29] presented updated results combining previous reports from two independent groups[13,30] together with 106 newly evaluated patients from both groups. In this report, we included only the added subpopulation as an independent study. Three studies reported FDG-PET results at completion of second cycle and fourth cycle of chemotherapy.[13,14,31] We abstracted data only on the second cycle in these studies. One study evaluated

Table 2. Patient Characteristics of Studies of Positron Emission Tomography for Interim Response Assessment of Malignant Lymphoma

Study	Year	No. of Participants Included	Clinical Staging*	Staging Before Therapy (No.)	Standard Prognostic Scores (No.)	Therapy	Use of Rituximab (%)
Advanced-stage HL			Inclusion criteria of advanced-stage		International Prognostic Scores		
Friedberg et al[33]	2004	22	IIB-IVB, any stage with bulky disease		NR	ABVD × 6 or MOPP/ABVD × 6 ± radiotherapy	—
Hutchings et al[37]	2005	28	IIB-IVB, any stage with bulky disease		NR	ABVD × 6 to 8 ± radiotherapy	—
Gallamini et al[30]	2006	108	IIB-IVB, IIA with adverse prognostic factors†		0 pts: 28, 1 pt: 34, 2 pts: 29, 3 pts: 10, 4 pts: 3, ≥ 5 pts: 4	ABVD × 6 or COPP/EBV/CAD × 6 ± radiotherapy	—
Hutchings et al[13]	2006	46	IIB-IVB		Median 3 pts	ABVD × 6 to 8 or comparable anthracycline-containing regimen ± radiotherapy	—
Kostakoglu et al[32]	2006	10	III-IV, any stage with bulky disease‡		0 pts: 3, 1 pt: 2, 2 pts: 4, 4 pts: 1	ABVD × 6	—
Zinzani et al[14]	2006	40	IIB-IVB		NR	ABVD × 6	—
Gallamini et al[29]	2007	106	IIB-IVB, IIA with adverse prognostic factors†		0 pts: 38, 1 pt: 70, 2 pts: 87, 3 pts: 42, 4 pts: 13, ≥ 5 pts: 10§	ABVD × 6, ABVD-like regimen × 6, or COPP/EBV/CAD × 6 ± radiotherapy	—
DLBCL					International Prognostic Indexes		
Spaepen et al[34]	2002	47		IA: 1, IIA: 15, IIB: 6, IIIA: 14, IIIB: 2, IVA: 14, IVB: 20§	L: 26, L-I: 22, H-I: 17, H: 17§	CHOP × 8, biweekly CHOP × 6, CHVmPBV × 8, or COP/COPADM/CYM × 6‖	0
Haioun et al[31]	2005	83		I-II: 8, III-IV: 82§	L: 14, L-I: 23, H-I: 30, H: 23§	(R-)CHOP × 8, R-ACVBP × 4¶, or ACVBP × 4 or ACE × 4#	45
Mikhaeel et al[12]	2005	57		I: 21, II: 14, III: 9, IV: 13	NR	(R-)CHOP × 6 or PMitCEBO × 6**	16
Fruchart et al[35]	2006	35		I-II: 13, III-IV: 27§	L: 13, L-I: 2, H-I or H: 15§	(R-)CHOP × 8 or (R-)ACVBP × 4††	74
Kostakoglu et al[32]	2006	24		I: 2, II: 11, III: 10, IV: 1	L: 16, L-I: 8	R-CHOP × 6 to 8	100
Querellou et al[38]	2006	21		I: 3, II: 2, III: 4, IV: 15§	L: 8, L-I: 5, H-I: 6, H: 5§	(R-)CHOP × 8, R-COP × 6, or (R-)CEEP × 4‡‡	90
Ng et al[36]	2007	44		I: 16, II: 9, III: 5, IV: 14	L: 17, L-I: 9, H-I: 12, H: 1, NA: 5	(R-)CHOP or CHOP-like regimen × 6 to 8, (R-)Hyper-CVAD × 8, or biweekly (R-) CHOP × 6 ± radiotherapy§§	40

Abbreviations: ABVD, doxorubicin, bleomycin, vinblastine, dacarbazine; ACE, doxorubicin, cyclophosphamide, etoposide; ACVBP, doxorubicin, cyclophosphamide, vindesine, bleomycin, prednisone; pts, patients; CAD, lomustine, doxorubicin, vindesine; CEEP, cyclophosphamide, epirubicin, vindesine, prednisone; CHOP, cyclophosphamide, doxorubicin, vincristine, prednisone; CHVmPBV, cyclophosphamide, doxorubicin, teniposide, prednisone, bleomycin, vincristine; COP, cyclophosphamide, vincristine, prednisone; COPADM, cyclophosphamide, vincristine, prednisolone, doxorubicin, high-dose methotrexate; COPP, cyclophosphamide, vincristine, procarbazine, prednisone; CVAD, cyclophosphamide, vincristine, doxorubicin, dexamethasone; CYM, cytrabine, high-dose methotrexate; DLBCL, diffuse large B-cell lymphoma; EBV, epirubicin, bleomycin, vinblastine; H, high risk; H-I, high-intermediate risk; HL, Hodgkin's lymphoma; L, low risk; L-I, low-intermediate risk; MOPP, nitrogen mustard, vincristine, procarbazine, prednizone; NR, not reported; PMitCEBO, cyclophosphamide, mitoxantrone, etoposide, prednisolone, vincristine, bleomycin; R, rituximab.

*According to the Ann Arbor staging system.

†> 3 nodal cites, subdiaphragmatic involvement, bulky disease, erythrocyte sedimentation rate > 40 mm/hour.

‡Selected post hoc because of no information on B symptoms.

§Abstracted from total participants of original report, not exclusively for relevant patient population.

‖Some underwent high-dose chemotherapy followed by autologous stem-cell transplantation as consolidation therapy.

¶All received an eight-cycle biweekly consolidation therapy consisting high-dose methotrexate, etoposide, ifosfamide, and cytarabine after the ACVBP regimen.

#All underwent high-dose chemotherapy followed by autologous stem-cell transplantation with or without rituximab maintenance therapy.

**A portion of patients (n = 16) with limited-stage disease underwent 2 to 4 cycles of (R-)CHOP followed by involved field radiation therapy instead of full course (R-)CHOP.

††Patents with one age-adjusted international prognostic risk factor received an eight-cycle consolidation therapy, and patients with two or three factors underwent high-dose chemotherapy followed by autologous stem-cell transplantation.

‡‡All underwent high-dose chemotherapy followed by autologous stem-cell transplantation.

§§A portion of patients (n = 13) with limited-stage disease underwent 2 to 4 cycles of (R-)CHOP or similar regimens followed by involved field radiation therapy instead of full-cycle chemotherapy.

PET at varied timing ranging from the first to fifth cycle.[36] We contacted the investigators for individual patient data, and excluded one patient who underwent PET at the fifth cycle. We found one study[14] through hand searching of the reference lists. As a result, we included 13 studies: eight studies[13,14,29,30,32-35] that met all eligibility criteria and five studies[12,31,36-38] with unpublished data available through contacting the authors (Table 1).[12-14,29-36,38]

Study Characteristics

Thirteen included studies had 360 advanced-stage HL patients and 311 DLBCL patients (Table 1). Eight reports were prospective single- or multi-institutional studies enrolling adults or adolescents. Only one study evaluated both adults and children.[34] Most of the patients in the HL studies underwent PET after receiving two cycles of first-line chemotherapy, while the number of cycles before the PET scan varied in DLBCL studies. In three DLBCL studies, 25% to 52% of included patients underwent PET after the fourth cycle.[34,36,38] One study evaluated PET after one cycle.[32] In general, participants underwent PET during the second week of intended chemotherapy cycle for biweekly chemotherapies (eg, doxorubicin, bleomycin, vinblastine, dacarbazine [ABVD] or (R-) doxorubicin, cyclophosphamide, vindesine, bleomycin, prednisone [ACVBP]) and during the third week for triweekly regimens (eg, (R-) cyclophosphamide, doxorubicin, vincristine, prednisone [CHOP]). Four studies performed CT for a portion of patients at the same timing as interim PET but they did not perform direct comparison between the two tests.[13,30,36,38]

For advanced-stage HL studies, fewer than 10% of included patients had unfavorable risk by standard prognostic tool (IPS > 3 points; Table 2). Progression or relapse rates were between 20% and 30% except for one study of 50%.[32] All studies adopted currently available standard first-line chemotherapy: six to eight cycles of ABVD or comparable regimens with or without radiotherapy. For DLBCL studies, the percentage of patients with unfavorable prognosis (high-intermediate to high risk by IPI) ranged from 0% to 59%, with progression or relapse rates of 27% to 47%. Full course (R-) CHOP and (R-) ACVBP were the two most widely adopted regimens. Two studies employed abbreviated course of (R-) CHOP or comparable regimens followed by involved-field radiation for patients with limited-stage disease.[12,36] No patients received rituximab in one study.[34] In four studies, some patients received consolidation auto-transplant after induction chemotherapy.[31,34,35,38]

Concerning imaging techniques and technologies, included studies generally followed guidelines by the Society of Nuclear Medicine (Table 3). One study exclusively adopted combined PET/CT scanner.[38] In five studies, some patients underwent combined PET/CT while the others were evaluated with stand-alone dedicated PET scanner.[13,29,30,32,36] All but one study[34] adopted attenuation correction for image reconstruction.

In general, multiple experienced nuclear medicine physicians interpreted PET results with pretherapy baseline scan as reference. All studies adopted qualitative positive and negative diagnostic criteria with various definitions (online-only Appendix Table A3). Only two studies clearly reported the referential backgrounds to define positive lesion. Five studies defined MRU criterion,[12-14,29,37] which was eventually reported as negative in three studies.[13,14,29] No study reported between-observer variability.

Quality Assessment of Published Studies

Only two studies[13,35] reported all items of the QUADAS tool (online-only Appendix Table A4). Reporting was especially limited in

Table 3. Technical Specification of PET for Interim Response Assessment of Malignant Lymphoma

			Procedure				
Study	Year	Preparation: Measurement of Blood Glucose	Type of PET Scanner	Time of Scan After Injection (minutes)	Attenuation Correction	Image Reconstruction Method	Administered Activity (MBq)
Advanced-stage HL + DLBCL							
Kostakoglu et al[32]	2006	Yes	PET-CT or dedicated	60	Yes	OSEM	370-444
Advanced-stage HL							
Friedberg et al[33]	2004	Yes	Dedicated	50	Yes	OSEM	370
Hutchings et al[37]	2005	Yes	Dedicated	60	Yes	NR	350
Gallamini et al[30]	2006	Yes	PET-CT or dedicated	60	Yes	OSEM or RAMLA	370/70, 259/70, 2*
Hutchings et al[13]	2006	NR	PET-CT or dedicated	45-90	Yes	OSEM	400
Zinzani et al[14]	2006	NR	Dedicated	70-90	Yes	NR	6†
Gallamini et al[29]	2007	Yes	PET-CT or dedicated	60	Yes	OSEM or RAMLA	370/70, 259/70, 2*
DLBCL							
Spaepen et al[34]	2002	Yes	Dedicated	60	No	OSEM	370-555
Haioun et al[31]	2005	Yes	Dedicated	60	Yes	OSEM	2†
Mikhaeel et al[12]	2005	NR	Dedicated	60	Yes	NR	350
Fruchart et al[35]	2006	NR	Dedicated	60	Yes	OSEM	2.5†
Querellou et al[38]	2006	Yes	PET-CT	73 ± 15‡	Yes	OSEM	5.0-7.6†
Ng et al[36]	2007	Yes	PET-CT or dedicated	60-70	Yes	OSEM	5†

Abbreviations: CT, computed tomography; DLBCL, diffuse large B-cell lymphoma; HL, Hodgkin's lymphoma; NR, not reported; OSEM, ordered subsets expectation maximization; PET, positron emission tomography; RAMLA, row-action maximum likelihood algorithm; SUV, standard uptake value.

*Three hundred seventy MBq/70 kg at the centers that used a GE scanner, 259 MBq/70 kg at the centers that used a Philips scanner, and 2 MBq/body weight kg at the centers that used a C-PET scanner.

†Administred activity was reported as the amount per body weight MBq/kg; eg, 360 MBq was administered to a 60 kg patient for 6 MBq/kg).

‡Mean ± standard deviation.

three retrospective studies.[14,36,37] Physicians' knowledge of interim PET results may affect their assessment of patients' response as well as their treatment decisions, introducing biases.[39] Only three prospective studies[13,29,35] explicitly adopted blinding of clinicians to interim PET results to deal with these biases. In three prospective studies,[14,30,31] although they did not explicitly report the use of blinding, interim PET was not utilized to alter the preplanned treatment strategies. In two retrospective studies,[32,38] interim PET results had no effect on the treatment decisions. Because the assessment of treatment failure is not always objective, the absence of blinding can still potentially influence the way treating physicians judge the final clinical outcome in favor of interim PET, especially when the outcome is equivocal.[11,39] Although all the studies adopted the standard guidelines on response assessment[40,41] as the reference standard, they did not specify minimum follow-up period or situations where pathological confirmation was required. Four studies[29,32,33,38] employed post-therapy or follow-up PET to complement post-therapy response assessment. Because post-therapy response assessment with PET is still imperfect,[9] the applied reference standard could overestimate prognostic accuracy.[39]

Sensitivity, Specificity, LRs, and Summary ROC Curves

For advanced-stage HL, studies reported sensitivity from 0.67 to 1.00 and consistently high specificity from 0.94 to 1.00 for interim FDG-PET (Table 4; Fig 1). Summary estimates were 0.81 for sensitivity (95% CI, 0.72 to 0.89), 0.97 for specificity (95% CI, 0.94 to 0.99), 28.4 for positive LR (95% CI, 14.2 to 56.7), and 0.19 for negative LR (95% CI, 0.12 to 0.30). We did not estimate summary ROC curves because data points were closely clustered together with limited variations, a situation in which the hierarchical model could not produce reliable estimates (Fig 2).

DLBCL studies reported wide-ranging sensitivity (0.50 to 1.0) and specificity (0.73 to 1.00) values for interim FDG-PET (Table 4; Fig 1). Combined estimates had a sensitivity of 0.78 (95% CI, 0.64 to 0.87), a specificity of 0.87 (95% CI, 0.75 to 0.93), a positive LR of 5.9 (95% CI, 2.8 to 12.3), and a negative LR of 0.26 (95% CI, 0.15 to 0.46). The Q* statistic for the summary ROC curve was 0.82 (Fig 2).

In sensitivity analyses, the summary prognostic accuracy was stable for both advanced-stage HL and DLBCL regardless of how MRU results or early-censored cases without treatment failure were counted (results not shown). Regarding alternative reference standards based on the duration of clinical follow-up, subgroup data were available for five advanced-stage HL studies (n = 232)[13,14,30,32,37] and five DLBCL studies (n = 181)[12,32,35,36,38] (online-only Appendix Table A5). All DLBCL studies had improvement in sensitivity with loss of specificity when only progression during first-line therapy was counted by the alternative reference standard. A similar tendency was

Table 4. Study Results of Positron Emission Tomography for Interim Response Assessment of Malignant Lymphoma

Study	Year	Total No.	Progression or Relapse (%)	TP	FN	FP	TN	Sensitivity	95% CI	Specificity	95% CI	Positive Likelihood Ratio	95% CI	Negative Likelihood Ratio	95% CI
Advanced-stage HL															
Friedberg et al[33]	2004	22	23	4	1	1	16	0.80	0.28 to 1.00	0.94	0.71 to 1.00	13.6	1.9 to 95.7	0.21	0.04 to 1.23
Hutchings et al[37]	2005	28	32	6	3	0	19	0.67	0.30 to 0.93	1.00	0.82 to 1.00	26.0	1.6 to 416.8	0.36*	0.15 to 0.84
Gallamini et al[30]	2006	108	19	18	3	2	85	0.86	0.64 to 0.97	0.98	0.92 to 1.00	37.3	9.4 to 148.4	0.15	0.05 to 0.42
Hutchings et al[13]	2006	46	28	10	3	1	32	0.77	0.46 to 0.95	0.97	0.84 to 1.00	25.4	3.6 to 178.9	0.24	0.09 to 0.64
Kostakoglu et al[32]	2006	10	50	5	0	0	5	1.00	0.48 to 1.00	1.00	0.48 to 1.00	11.0	0.8 to 158.0	0.09	0.01 to 1.31
Zinzani et al[14]	2006	40	23	8	1	0	31	0.89	0.52 to 1.00	1.00	0.89 to 1.00	54.4	3.4 to 861.6	0.15†	0.04 to 0.67
Gallamini et al[29]	2007	106	20	15	4	4	83	0.79	0.54 to 0.94	0.95	0.89 to 0.99	17.2	6.4 to 46.0	0.22	0.09 to 0.53
DLBCL															
Spaepen et al[34]	2002	47	47	20	2	0	25	0.91	0.71 to 0.99	1.00	0.86 to 1.00	46.3	3.0 to 724.1	0.11	0.03 to 0.36
Haioun et al[31]	2005	83	39	20	12	14	37	0.63	0.44 to 0.79	0.73	0.58 to 0.84	2.3	1.4 to 3.8	0.52	0.32 to 0.83
Mikhaeel et al[12]	2005	57	38	15	7	8	27	0.68	0.45 to 0.86	0.77	0.60 to 0.90	3.0	1.5 to 5.8	0.41‡	0.22 to 0.78
Fruchart et al[35]	2006	35	29	9	1	6	19	0.90	0.56 to 1.00	0.76	0.55 to 0.91	3.8	1.8 to 7.8	0.13	0.02 to 0.86
Kostakoglu et al[32]	2006	24	38	9	0	1	14	1.00	0.66 to 1.00	0.93	0.68 to 1.00	10.1	2.2 to 46.8	0.06	0.00 to 0.83
Querellou et al[38]	2006	21	29	3	3	1	14	0.50	0.12 to 0.88	0.93	0.68 to 1.00	7.5	1.0 to 58.6	0.54	0.24 to 1.21
Ng et al[36]	2007	45	27	8	4	4	28	0.67	0.35 to 0.90	0.88	0.71 to 0.97	5.3	2.0 to 14.5	0.38	0.17 to 0.86

Abbreviations: DLBCL, diffuse large B-cell lymphoma; FN, false negative; FP, false positive; MRU, minimal residual uptake; TN, true negative; TP, true positive.
*The likelihood ratios for a MRU and a negative scan were 0.35 (95% CI, 0.05 to 2.5) and 0.33 (95% CI, 0.09 to 1.1), respectively, if these two categories were estimated separately.
†The likelihood ratios for a MRU and a negative scan were 1.1 (95% CI, 0.14 to 9.7) and 0.06 (95% CI, 0.00 to 0.84), respectively, if these two categories were estimated separately.
‡The likelihood ratios for a MRU and a negative scan were 0.96 (95% CI, 0.25 to 3.6) and 0.29 (95% CI, 0.12 to 0.73), respectively, if these two categories were estimated separately.

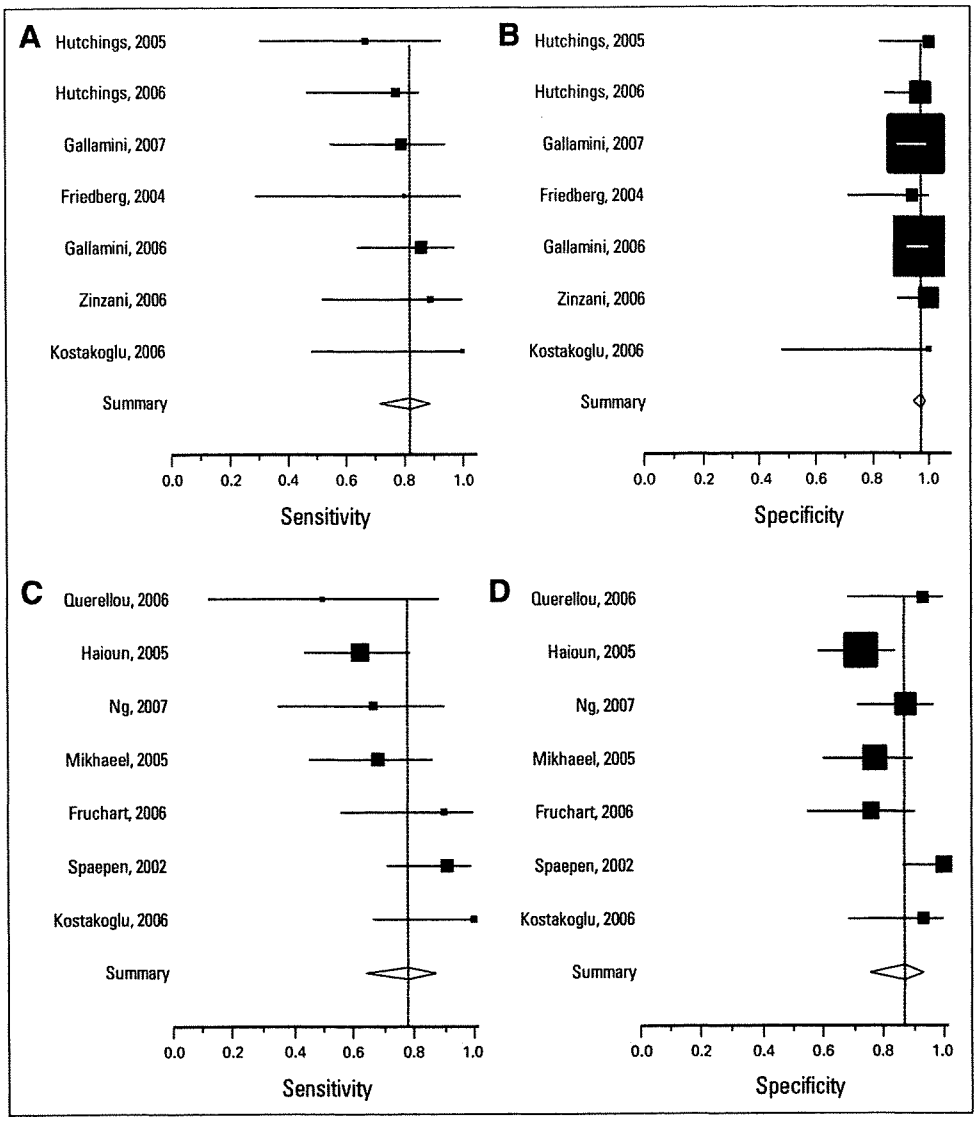

Fig 1. Sensitivity and specificity for (A, B) advanced-stage Hodgkin's lymphoma and (C, D) diffuse large B-cell lymphoma. The size of the square plotting is proportional to the number of patients with treatment failure for sensitivity and in remission for specificity. The horizontal lines are the 95% CIs. The vertical lines represent the summary estimates.

observed in all but one[30] advanced-stage HL studies (online-only Appendix Fig A2).

Subgroup Analyses and Meta-Regression Analyses

We did not perform subgroup analyses for advanced-stage HL because there were too few data points and there was little variation of the results across studies (Fig 1). Visual assessment of the ROC plots of DLBCL studies did not identify meaningful subgroups (data not shown). Meta-regression analyses on both advanced-stage HL and DLBCL did not find any clinical or test characteristics to explain the observed variability (data not shown).

DISCUSSION

This systematic review of interim response assessment of FDG-PET for patients with untreated advanced-stage HL showed that studies consistently reported high specificity and positive LRs. Although study quality was limited in some studies, as demographic and clinical characteristics of included patients were reasonably comparable over the studies, our results should generally be applicable to adult and adolescent patients with low- to intermediate-risk (IPS 0 to 3) receiving standard full course ABVD or comparable regimens. Because the summary positive LR is very high, positive PET results after a few cycles of chemotherapy would probably have an excellent ability to predict poor responders. Patients with negative PET, which predicts good response during the therapy, still have a moderate risk of post-treatment relapse since the summary negative LR is 0.19.[42]

The reported sensitivity and specificity of DLBCL studies of interim FDG-PET varied. This review also identified considerable clinical heterogeneity in these studies. For example, studies included patients with varied risk of treatment failure and adopted various therapeutic interventions. Also, studies were heterogeneous in how PET was used, such as the number of chemotherapy cycles before PET

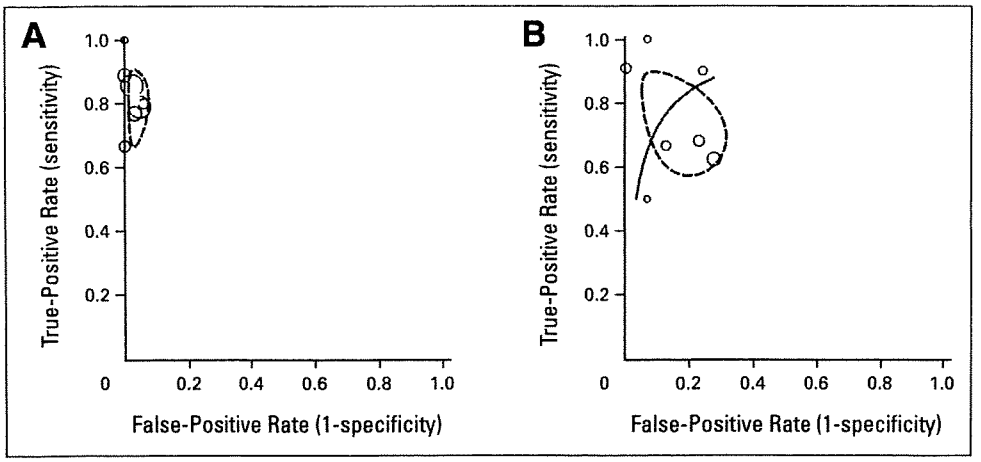

Fig 2. Receiver operating characteristic (ROC) plotting for (A) advanced-stage Hodgkin's lymphoma and (B) diffuse large B-cell lymphoma. Individual study estimates of sensitivity and 1 – specificity are shown (open circles). Summary ROC curve is presented only for DLBCL. Closed square represents the summary estimates. Dashed boundary represents the 95% confidence region for the summary sensitivity and specificity.

and the timing of scanning during the chemotherapy cycle.[43] Thus, our summary estimates should be interpreted carefully. Although we performed subgroup analyses and meta-regression analyses, we could not identify characteristics to explain the variability.

This study has several important limitations. Because only 13 studies with pertinent data were included in the meta-analysis, it may lack the power to detect clinically meaningful factors. In sensitivity analyses, fewer studies were available; therefore, the results may be less reliable. Although we did not independently estimate the summary LR for a MRU result, this distinct category may carry a worse prognosis than a clearly negative scan as reported.[12,14] Also, our results are likely subject to overestimation due to methodologic limitations in original studies, such as the absence of blinding of interim PET results to clinicians to assess final clinical outcomes.[11] Further, because of lack of data, we did not address the comparison between FDG-PET and CT or FDG-PET/CT and PET alone[38]; this review cannot answer whether PET is better than CT or whether the combined modality is superior to stand alone PET. In addition, this review did not specifically focus on limited-stage lymphoma; thus our results cannot answer the clinical question of whether early-interim PET can reliably identify good responders with localized disease. Finally, although three advanced-stage HL studies[13,29,30] and one DLBCL study[31] reported interim FDG-PET scan as a statistically significant independent prognostic factor in addition to IPS and IPI, respectively, we did not directly address this issue. For advanced-stage HL, because the included studies had few poor-risk (IPS 4 to 7) patients, our results may be less applicable to high-risk populations.

Interim PET should remain at this time as a test to be evaluated as part of clinical research where treatment regimens and imaging conditions are standardized; thus it should not be employed in the routine setting. This review supports conducting prospective trials for advanced-stage HL patients especially with low- to intermediate-risk (IPS 0 to 3) that incorporate early altering treatment to more intensive approach on the basis of positive FDG-PET results. For DLBCL, there is insufficient data to support similar trials. Additional prospective prognostic accuracy studies in the setting of conventional strategy would be needed to elucidate subgroups and timings of interim PET to better identify poor responders. Also, outside of study protocols where treatment strategies are explicitly defined on the basis of scan results, biopsy should be considered for positive PET findings if they are used to prompt a change in patient management. This is especially relevant if there is discrepancy between the scan results and other clinical data. Although biopsy cannot provide quantitative information as to how much residual tumor exists, it still is the most reliable way to confirm the presence of disease.

Authors' Disclosures of Potential Conflicts of Interest

Although all authors completed the disclosure declaration, the following author(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a "U" are those for which no compensation was received; those relationships marked with a "C" were compensated. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Employment or Leadership Position: None **Consultant or Advisory Role:** None **Stock Ownership:** None **Honoraria:** None **Research Funding:** Teruhiko Terasawa, Nihon Medi-Physics Co Ltd **Expert Testimony:** None **Other Remuneration:** None

Author Contributions

Conception and design: Teruhiko Terasawa, Joseph Lau, Tomomitsu Hotta, Takashi Nihashi, Hirokazu Nagai
Administrative support: Tomomitsu Hotta, Hirokazu Nagai
Provision of study materials or patients: Stéphane Bardet, Olivier Couturier, Martin Hutchings
Collection and assembly of data: Teruhiko Terasawa, Takashi Nihashi, Hirokazu Nagai
Data analysis and interpretation: Teruhiko Terasawa, Joseph Lau, Stéphane Bardet, Olivier Couturier, Martin Hutchings, Takashi Nihashi
Manuscript writing: Teruhiko Terasawa, Joseph Lau, Stéphane Bardet
Final approval of manuscript: Teruhiko Terasawa, Joseph Lau, Stéphane Bardet, Olivier Couturier, Tomomitsu Hotta, Martin Hutchings, Takashi Nihashi, Hirokazu Nagai

1. Jemal A, Siegel R, Ward E, et al: Cancer statistics, 2007. CA Cancer J Clin 57:43-66, 2007
2. Connors JM: State-of-the-art therapeutics: Hodgkin's lymphoma. J Clin Oncol 23:6400-6408, 2005
3. Coiffier B: State-of-the-art therapeutics: Diffuse large B-cell lymphoma. J Clin Oncol 23:6387-6393, 2005
4. Hasenclever D, Diehl V: A prognostic score for advanced Hodgkin's disease: International Prognostic Factors Project on advanced Hodgkin's disease. N Engl J Med 339:1506-1514, 1998
5. The International Non-Hodgkin's Lymphoma Prognostic Factors Project: A predictive model for aggressive non-Hodgkin's lymphoma. N Engl J Med 329:987-994, 1993
6. Diehl V, Franklin J, Pfreundschuh M, et al: Standard and increased-dose BEACOPP chemotherapy compared with COPP-ABVD for advanced Hodgkin's disease. N Engl J Med 348:2386-2395, 2003
7. Greb A, Bohlius J, Trelle S, et al: High-dose chemotherapy with autologous stem cell support in first-line treatment of aggressive non-Hodgkin lymphoma: Results of a comprehensive meta-analysis. Cancer Treat Rev 33:338-346, 2007
8. Seam P, Juweid ME, Cheson BD: The role of FDG-PET scans in patients with lymphoma. Blood 110:3507-3516, 2007
9. Terasawa T, Nihashi T, Hotta T, et al: Fluorine-18-Fluorodeoxyglucose positron emission tomography for post-therapy assessment of Hodgkin's lymphoma and aggressive non-Hodgkin's lymphoma: A systematic review. J Nucl Med 49:13-21, 2008
10. Schelbert HR, Hoh CK, Royal HD, et al: Society of Nuclear Medicine Procedure Guideline for Tumor Imaging Using F-18 FDG version 2.0. 1999. http://interactive.snm.org/index.cfm?PageID=772&RPID=969
11. Whiting P, Rutjes AW, Reitsma JB, et al: The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 3:25, 2003
12. Mikhaeel NG, Hutchings M, Fields PA, et al: FDG-PET after two to three cycles of chemotherapy predicts progression-free and overall survival in high-grade non-Hodgkin lymphoma. Ann Oncol 16:1514-1523, 2005
13. Hutchings M, Loft A, Hansen M, et al: FDG-PET after two cycles of chemotherapy predicts treatment failure and progression-free survival in Hodgkin lymphoma. Blood 107:52-59, 2006
14. Zinzani PL, Tani M, Fanti S, et al: Early positron emission tomography (PET) restaging: A predictive final response in Hodgkin's disease patients. Ann Oncol 17:1296-1300, 2006
15. Gatsonis C, Paliwal P: Meta-analysis of diagnostic and screening test accuracy evaluations: Methodologic primer. Am J Roentgenol 187:271-281, 2006
16. Rabe-Hesketh S, Skrondal A, Pickles A: GLLAMM Manual. U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 160, 2004. http://www.bepress.com/ucbbiostat/paper160/
17. Harbord RM, Deeks JJ, Egger M, et al: A unification of models for meta-analysis of diagnostic accuracy studies. Biostatistics 8:239-251, 2007
18. Kostakoglu L, Coleman M, Leonard JP, et al: PET predicts prognosis after 1 cycle of chemotherapy in aggressive lymphoma and Hodgkin's disease. J Nucl Med 43:1018-1027, 2002
19. Itti E, Haioun C, Rahmouni A, et al: Aggressive large cell lymphoma: Early assessment of therapeutic efficacy by positron emission tomography. Medecine Nucleaire 28:327-332, 2004
20. Dupuis J, Gaulard P, Hemery F, et al: Respective prognostic values of germinal center phenotype and early (18)fluorodeoxyglucose-positron emission tomography scanning in previously untreated patients with diffuse large B-cell lymphoma. Haematologica 92:778-783, 2007
21. Romer W, Hanauske AR, Ziegler S, et al: Positron emission tomography in non-Hodgkin's lymphoma: Assessment of chemotherapy with fluorodeoxyglucose. Blood 91:4464-4471, 1998
22. Slaby J, Belohlavek O, Taborska K, et al: Predictive features of positron emission tomography after two cycles of induction therapy in malignant lymphoma. Cas Lek Cesk 141:312-315, 2002
23. Zijlstra JM, Hoekstra OS, Raijmakers PG, et al: 18FDG positron emission tomography versus 67Ga scintigraphy as prognostic test during chemotherapy for non-Hodgkin's lymphoma. Br J Haematol 123:454-462, 2003
24. Hoekstra OS, van Lingen A, Ossenkoppele GJ, et al: Early response monitoring in malignant lymphoma using fluorine-18 fluorodeoxyglucose single-photon emission tomography. Eur J Nucl Med 20:1214-1217, 1993
25. Hoekstra OS, Ossenkoppele GJ, Golding R, et al: Early treatment response in malignant lymphoma, as determined by planar fluorine-18-fluorodeoxyglucose scintigraphy. J Nucl Med 34: 1706-1710, 1993
26. Torizuka T, Nakamura F, Kanno T, et al: Early therapy monitoring with FDG-PET in aggressive non-Hodgkin's lymphoma and Hodgkin's lymphoma. Eur J Nucl Med Mol Imaging 31:22-28, 2004
27. Jerusalem G, Beguin Y, Fassotte MF, et al: Persistent tumor 18F-FDG uptake after a few cycles of polychemotherapy is predictive of treatment failure in non-Hodgkin's lymphoma. Haematologica 85: 613-618, 2000
28. Schot B, van Imhoff G, Pruim J, et al: Predictive value of early 18F-fluoro-deoxyglucose positron emission tomography in chemosensitive relapsed lymphoma. Br J Haematol 123:282-287, 2003
29. Gallamini A, Hutchings M, Rigacci L, 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 25:3746-3752, 2007
30. Gallamini A, Rigacci L, Merli F, et al: The predictive value of positron emission tomography scanning performed after two courses of standard therapy on treatment outcome in advanced stage Hodgkin's disease. Haematologica 91:475-481, 2006
31. Haioun C, Itti E, Rahmouni A, 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 106:1376-1381, 2005
32. Kostakoglu L, Goldsmith SJ, Leonard JP, et al: FDG-PET after 1 cycle of therapy predicts outcome in diffuse large cell lymphoma and classic Hodgkin disease. Cancer 107:2678-2687, 2006
33. Friedberg JW, Fischman A, Neuberg D, et al: FDG-PET is superior to gallium scintigraphy in staging and more sensitive in the follow-up of patients with de novo Hodgkin lymphoma: A blinded comparison. Leuk Lymphoma 45:85-92, 2004
34. Spaepen K, Stroobants S, Dupont P, et al: Early restaging positron emission tomography with (18)F-fluorodeoxyglucose predicts outcome in patients with aggressive non-Hodgkin's lymphoma. Ann Oncol 13:1356-1363, 2002
35. Fruchart C, Reman O, Le Stang N, et al: Prognostic value of early 18 fluorodeoxyglucose positron emission tomography and gallium-67 scintigraphy in aggressive lymphoma: A prospective comparative study. Leuk Lymphoma 47:2547-2557, 2006
36. Ng AP, Wirth A, Seymour JF, 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 48:596-600, 2007
37. Hutchings M, Mikhaeel NG, Fields PA, et al: Prognostic value of interim FDG-PET after two or three cycles of chemotherapy in Hodgkin lymphoma. Ann Oncol 16:1160-1168, 2005
38. Querellou S, Valette F, Bodet-Milin C, et al: FDG-PET/CT predicts outcome in patients with aggressive non-Hodgkin's lymphoma and Hodgkin's disease. Ann Hematol 85:759-767, 2006
39. Whiting P, Rutjes AW, Reitsma JB, et al: Sources of variation and bias in studies of diagnostic accuracy: A systematic review. Ann Intern Med 140:189-202, 2004
40. Cheson BD, Horning SJ, Coiffier B, et al: Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas: NCI Sponsored International Working Group. J Clin Oncol 17:1244-1253, 1999
41. Lister TA, Crowther D, Sutcliffe SB, et al: Report of a committee convened to discuss the evaluation and staging of patients with Hodgkin's disease: Cotswolds meeting. J Clin Oncol 7:1630-1636, 1989
42. Grimes DA, Schulz KF: Refining clinical diagnosis with likelihood ratios. Lancet 365:1500-1505, 2005
43. Juweid ME, Stroobants S, Hoekstra OS, 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 25:571-578, 2007

Acknowledgment

We thank Tatsuo Torizuka, MD, Jerusalem Guy, MD, Lale Kostakoglu, MD, Corinne Haioun, MD, Ashley Ng, MD, Andrew Wirth, MD, and Rodney Hicks, MD, for providing data on their original work; Roger Harbord, PhD, for providing statistical programs; and Christopher Schmid, PhD, for assisting with statistical analysis.

Appendix

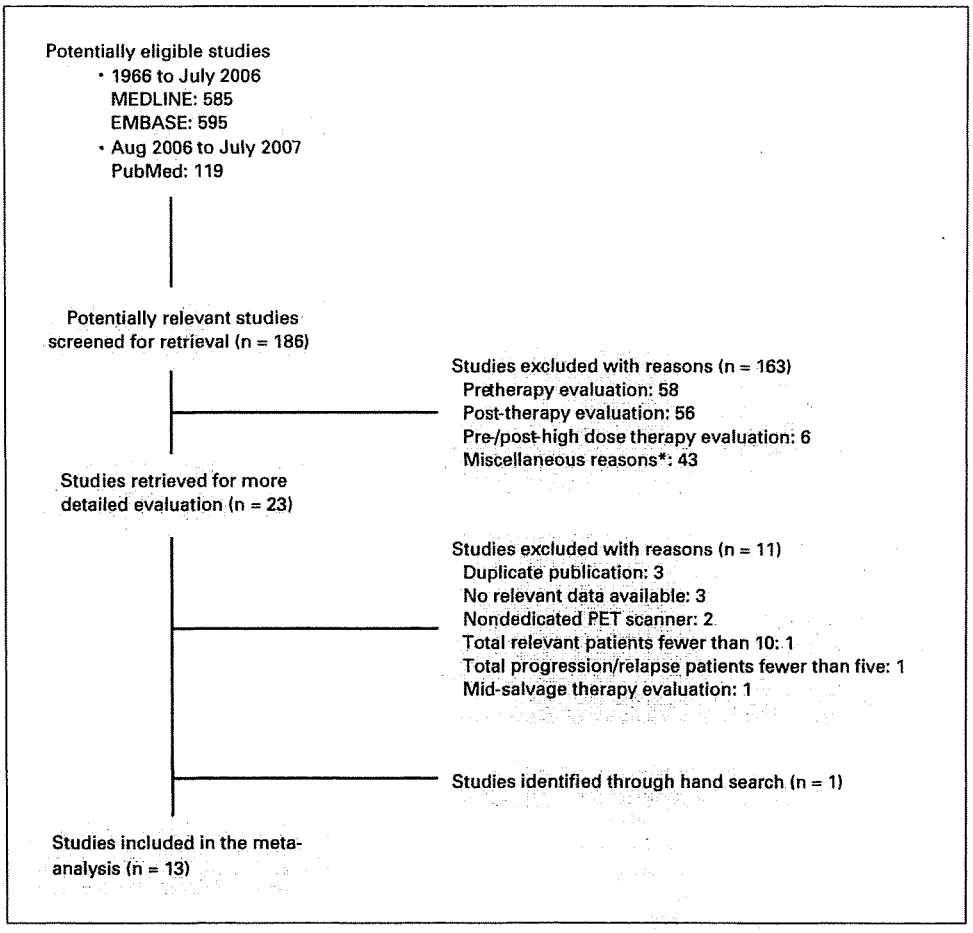

Fig A1. Article selection process. (*) Miscellaneous reasons include studies on staging evaluation at multiple different timings and contexts (n = 10), studies focusing on lymphoma involvement in a specific organ or anatomic region (n = 7), review articles (n = 5), studies focusing exclusively on positive positron emission tomography (PET) findings (n = 3), letters or comments (n = 3), studies on post-therapy follow-up (n = 3), studies on glucose metabolism (n = 3), and others (n = 9).

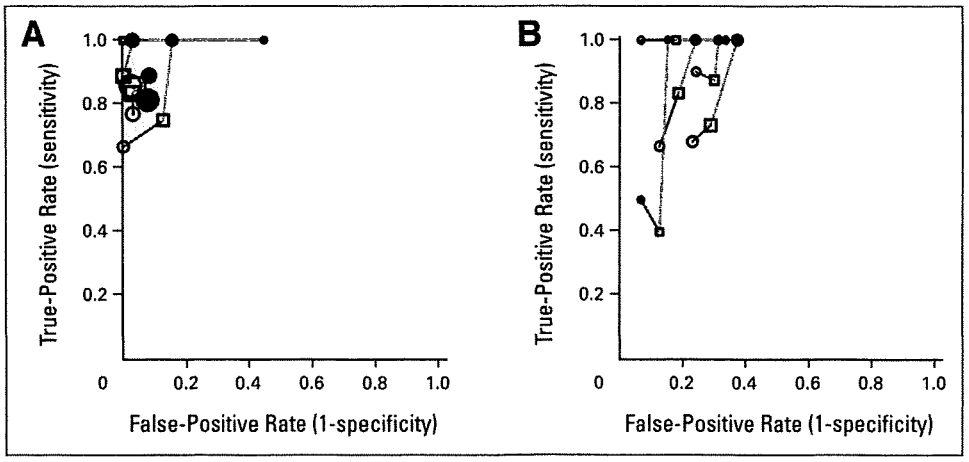

Fig A2. Receiver operating characteristic plotting for (A) advanced-stage Hodgkin's lymphoma and (B) diffuse large B-cell lymphoma. Individual study estimates of sensitivity and 1 − specificity are shown based on the duration of clinical follow-up: all treatment failures (open circles), early treatment failures (less than 12 months; open squares), very early treatment failures (less than 6 months; closed circles).

Table A1. Search Strategy

Ovid MEDLINE Search	EMBASE Search	PubMed Search
#1 exp Tomography, emission computed/	#1 exp computer assisted emission tomography/or exp positron emission tomography/or exp whole body tomography/	#1 tomography, emission computed [MeSH terms]
#2 positron emission tomography.ti,ab,rw,sh.	#2 positron emission tomography.ti,ab,hw,tn,mf.	#2 positron emission tomography
#3 pet$.ti,ab,rw,sh.	#3 pet$.ti,ab,hw,tn,mf.	#3 pet
#4 animal not (human and animal) sh.	#4 (animal not (human and animal)).ti,ab,hw,tn,mf.	#4 pet*
#5 #3 not #4	#5 3 not 4	#5 #1 OR #2 OR #3 OR #4
#6 #1 or #2 or #5	#6 1 or 2 or 5	#6 deoxyglucose [MeSH Terms]
#7 exp Deoxyglucose/	#7 exp Deoxyglucose/	#7 deoxyglucose
#8 deoxyglucose.ti,ab,rw,sh.	#8 deoxyglucose.ti,ab,hw,tn,mf.	#8 deoxy-glucose
#9 deoxy-glucose.ti,ab,rw,sh.	#9 deoxy-glucose.ti,ab,hw,tn,mf.	#9 fluorodeoxyglucose
#10 fluorodeoxyglucose.ti,ab,rw,sh.	#10 fluorodeoxyglucose.ti,ab,hw,tn,mf.	#10 18fluorodeoxyglucose
#11 18fluorodeoxyglucose.ti,ab,rw,sh.	#11 18fluorodeoxyglucose.ti,ab,hw,tn,mf.	#11 fludeoxyglucose
#12 fludeoxyglucose.ti,ab,rw,sh.	#12 fludeoxyglucose.ti,ab,hw,tn,mf.	#12 fdg
#13 fdg$.ti,ab,rw,sh.	#13 fdg$.ti,ab,hw,tn,mf.	#13 fdg*
#14 18fdg.ti,ab,rw,sh.	#14 18fdg.ti,ab,hw,tn,mf.	#14 18fdg
#15 f-18-dg.ti,ab,rw,sh.	#15 f-18-dg.ti,ab,hw,tn,mf.	#15 f-18-dg
#16 fluoro-2-deoxy-*d*-glucose.ti,ab,rw,sh.	#16 fluoro-2-deoxy-*d*-glucose.ti,ab,hw,tn,mf.	#16 fluoro-2-deoxy-*d*-glucose
#17 2fluoro-2deoxyglucose.ti,ab,rw,sh.	#17 2fluoro-2deoxyglucose.ti,ab,hw,tn,mf.	#17 2fluoro-2deoxyglucose
#18 fluoro-*d*-glucose.ti,ab,rw,sh.	#18 fluoro-*d*-glucose.ti,ab,hw,tn,mf.	#18 fluoro-*d*-glucose
#19 or/#8-#18	#19 or/#8-18	#19 #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18
#20 #7 or #19	#20 #7 or #19	#20 lymphoma
#21 #6 and #20	#21 #6 and #20	#21 lymphom*
#22 exp sensitivity-and-specificity or predict$ or diagnos$ or di.fs. or du.fs. or accura$	#22 sensitiv$ or detect$ or accura$ or specific$ or reliab$ or positive or negative diagnos$ or di.fs.	#22 Hodgkin*
#23 #21 and #22	#23 #21 and #22	#23 #20 OR #21 OR #22
#24 exp Lymphoma/	#24 exp Lymphoma/	#24 #5 AND #19 AND #23
#25 lymphoma.ti,ab,rw,sh.	#25 lymphoma.ti,ab,rw,sh.	
#26 lymphom$.ti,ab,rw,sh.	#26 lymphom$.ti,ab,rw,sh.	
#27 hodgkin$.ti,ab,rw,sh.	#27 hodgkin$.ti,ab,rw,sh.	
#28 or/#24-#27	#28 or/#24-#27	
#29 #23 and #28	#29 #23 and #28	

Table A2. Quality Assessment of Studies of Positron Emission Tomography for Interim Response Assessment of Malignant Lymphoma

Item No.	Bias or Issue Addressed	Question	How Scored
1	Avoidance of spectrum bias	Was the spectrum of patients representative of the patients who will receive the test in practice?	Scored as "yes" if patients were enrolled onto a study prospectively and consecutively based on the predefined inclusion criteria
2	Provision of a clear definition of the inclusion (and exclusion) criteria	Were selection criteria clearly described?	Scored as "yes" if a study reported clear inclusion criteria
3	Appropriateness of the reference standard	Is the reference standard likely to correctly classify the target condition?	Scored as "yes" if a study employed clinical follow-up with or without biopsy as the reference standard*
4	Avoidance of partial verification bias	Did the whole sample or a random selection of the sample receive verification using a reference standard of diagnosis?	Scored as "yes" if the whole patients of a study received disease verification through clinical follow-up with or without biopsy
5	Avoidance of differential verification bias	Did patients receive the same reference standard regardless of the index test result?	Scored as "yes" if the whole patients of a study received disease verification through clinical follow-up with or without biopsy regardless of the interim PET results
6	Avoidance of incorporation bias	Was the reference standard independent of the index test (i.e., the index test did not form part of the reference standard)?	Scored as "yes" as long as the ultimate diagnosis was made through predefined reference standard (ie, conventional response assessment with or without biopsy during clinical follow-up for disease progression or relapse, or sufficiently long follow-up for continuing remission) even if interim PET results were available to clinicians; scored as "no" only if the mid-therapy results were specifically used to determine the final clinical outcome
7	Replicativeness of the index test	Was the execution of the index test described in sufficient detail to permit replication of the test?	Scored as "yes" if a study reported sufficient details on the procedure and diagnostic criteria of interim PET
8	Replicativeness of the reference standard	Was the execution of the reference standard described in sufficient detail to permit its replication?	Scored as "yes" if a study evaluated disease status and followed up patients following the recommended standard guidelines*
9	Avoidance of test review bias	Were the index test results interpreted without knowledge of the results of the reference standard?	Scored as "yes" if interim PET was interpreted without knowledge of the clinical information on patients; scored as "no" if PET interpreters read the scan results in the presence of any clinical data including conventional imaging tests, laboratory test, and physical examinations obtained after the initiation of treatment, which could have contained the information on the response assessment or disease status of patients
10	Avoidance of diagnosis review bias	Were the reference standard results interpreted without knowledge of the results of the index test?	Scored as "yes" if the clinicians treated and followed up patients without knowledge of interim PET results
11	Availability of clinical data to test interpreters	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	Scored as "yes" if interpreters read interim PET scan in the presence of clinical information excluding baseline pre-therapy PET scan
12	Reporting of uninterpretable or intermediate results	Were uninterpretable/intermediate test results reported?	Scored as "yes" if a study reported the number of patients with minimal residual uptake
13	Provision of the information on withdrawals from a study	Were withdrawals from the study explained?	Scored as "yes" if a study clearly reported the number of patients satisfying the inclusion criteria that did or did not undergo interim PET and/or clinical follow-up with or without biopsy; scored as "no" if a study included exclusively patients who underwent interim PET and did not report the information on patients who satisfied the same inclusion criteria but did not undergo interim PET

Abbreviation: PET, positron emission tomography.
*Alternative more stringent criterion was also employed as follows: Scored as "yes" if a study explicitly stated that investigators assessed disease status according to the standard guidelines (Cheson BD, Horning SJ, Coiffier B, et al: J Clin Oncol 17:1244-1253, 1999; Lister TA, Crowther D, Sutcliffe SB, et al: J Clin Oncol 7:1630-1636, 1989) and followed up patients in remission including negative biopsy of lesions suspected of treatment failure (progression or relapse) for at least 1 year.

Table A3. Diagnostic Criteria and Interpreters of PET for Interim Response Assessment of Malignant Lymphoma

Study	Year	Interpretation Method	Qualitative Diagnostic Criteria			Reading Condition: Availability of Pretherapy PET to Interpreters of Interim PET	Interpreter	
			Positive	Negative	Minimal Residual Uptake		No.	Experience
Advanced-stage HL + DLBCL								
Kostakoglu et al (Kostakoglu L, Goldsmith SJ, Leonard JP, et al: Cancer 107:2678-2687, 2006)	2006	Qualitative analysis	Presence of FDG uptake that exceeded the uptake seen on the contralateral site or in the background in a location incompatible with normal anatomy or physiologic variants	No pathologically increased FDG uptake at any site compared with the uptake on the contralateral site or the background	Not specified	Yes	2	Expert
		Semi-quantitative analysis (complemental): SUV_{max} measured for only measurable nodal sites (lesion-based ROC analysis)	Increased FDG uptake in contralateral and asymmetrical sites compared with background activity to be compared with: general, the highest activity excluding pathological and physiologic sites of uptake; head and neck, within the jugular vessels; chest, in the mediastinum around the aortic arch region; abdomen/pelvis, in the mesentery or abdominal vessels, whichever had the higher activity					
Advanced-stage HL								
Friedberg et al (Friedberg JW, Fischman A, Neuberg D, et al: Leuk Lymphoma 45:85-92, 2004)	2004	Qualitative analysis	Nodal involvement: FDG avidity above mediastinal blood pool activity	Not specified	Not specified	No	2	Expert
Hutchings et al (Hutchings M, Mikhaeel NG, Fields PA, et al: Ann Oncol 16: 1160-1168, 2005)	2005	Qualitative analysis	Increased uptake suspicious for malignant disease, which does not have a benign explanation	No evidence of disease	Low-grade uptake of FDG (just above background) in a focus within an area of previously noted disease, not likely representing malignancy	Yes*	2	Expert
Gallamini et al (Gallamini A, Rigacci L, Merli F, et al: Haematologica 91:475-481, 2006)	2006	Qualitative analysis Semi-quantitative analysis (complemental): SUV_{max} measured for regions of interest (patient-based % of SUV_{max} reduction from baseline)	Presence of a focal concentration of FDG outside the areas of physiological uptake, with a value increased relative to background	No pathological FDG uptake at any site, including all sites of previously increased pathological uptake	Not specified	Yes	2	Expert

(continued on following page)

Table A3. Diagnostic Criteria and Interpreters of PET for Interim Response Assessment of Malignant Lymphoma (continued)

Study	Year	Interpretation Method	Qualitative Diagnostic Criteria			Reading Condition: Availability of Pretherapy PET to Interpreters of Interim PET	Interpreter	
			Positive	Negative	Minimal Residual Uptake		No.	Experience
Hutchings et al (Hutchings M, Loft A, Hansen M, et al: Blood 107:52-59, 2006)	2006	Qualitative analysis Semi-quantiitative analysis (complemental): SUV_{max} measured for regions of interest (patient-based distribution of SUV_{max})	Focal FDG concentration outside the physiological uptake areas, with clearly increased activity relative to the background	No pathologic FDG uptake at any site, including all sites of previously increased pathologic uptake	Low-grade FDG-uptake with avidity smaller than, equal to, or only slightly higher than the uptake in the mediastinal blood pool structures	Yes	2	Expert
Zinzani et al (Zinzani PL, Tani M, Fanti S, et al: Ann Oncol 17:1296-1300, 2006)	2006	Qualitative analysis	Areas of focal uptake other than the sites of known accumulation, including the kidney, bladder, and gastrointestinal tract	No evidence of disease Skeletal areas showing symmetric joint uptake, especially within the shoulder (considered arthritis)	Low-grade uptake of FDG (just above background) in a focus within an area of previously noted disease	Yes	3	Expert
Gallamini et al (Gallamini A, Hutchings M, Rigacci L, et al: J Clin Oncol 25: 3746-3752, 2007)	2007	Qualitative analysis Semi-quantitative analysis (complemental): SUV_{max} was measured for regions of interest	Presence of a focal FDG concentration outside the physiological uptake areas, with clearly increased activity relative to the background	No pathologic FDG uptake at any site, including all sites of previously increased pathologic uptake	Low-grade FDG-uptake with avidity smaller than, equal to, or only slightly higher than the uptake in the mediastinal blood pool structures A SUV of 2.0 to 3.5	Yes	2	Expert
DLBCL								
Spaepen et al (Spaepen K, Stroobants S, Dupont P, et al: Ann Oncol 13: 1356-1363, 2002)	2002	Qualitative analysis	Any focal or diffuse area of increased activity in a location incompatible with normal anatomy and suspect for residual disease and/or new localizations	No evidence of disease	Not specified	Unclear	2	Expert
Haioun et al (Haioun C, Itti E, Rahmouni A, et al: Blood 106: 1376-1381, 2005)	2005	Qualitative analysis†	At least one residual with a low extent and moderate intensity of abnormal FDG uptake Two or more residual sites with any extent and intensity of abnormal FDG uptake	No residual abnormal FDG uptake A unique residual site with a low extent and low intensity of FDG uptake, with all the other previously hypermetabolic sites extinguished	Not specified	Yes	2	Expert
Mikhaeel et al (Mikhaeel NG, Hutchings M, Fields PA, et al: Ann Oncol 16: 1514-1523, 2005)	2005	Qualitative analysis	Persistence or appearance of new areas of increased uptake, thought to be lymphoma-related	Disappearance of all abnormal disease-related uptake	Low-grade uptake of FDG in a focus within an area of previously noted disease, likely to represent inflammation, where small volume malignancy could not be excluded	Yes*	2	Expert

(continued on following page)

Table A3. Diagnostic Criteria and Interpreters of PET for Interim Response Assessment of Malignant Lymphoma (continued)

Study	Year	Interpretation Method	Qualitative Diagnostic Criteria			Reading Condition: Availability of Pretherapy PET to Interpreters of Interim PET	Interpreter	
			Positive	Negative	Minimal Residual Uptake		No.	Experience
Fruchart et al (Fruchart C, Reman O, Le Stang N, et al: Leuk Lymphoma 47:2547-2557, 2006)	2006	Qualitative analysis	At least one site of residual uptake	No significant residual uptake in suspected sites of lymphoma before treatment	Not specified	Yes	1	Expert
Querellou et al (Querellou S, Valette F, Bodet-Milin C, et al: Ann Hematol 85: 759-767, 2006)	2006	Qualitative analysis	Any focus of increased FDG uptake over background not located in areas of normal FDG uptake and/or FDG excretion Any focal or diffuse area of increased activity in a location suspect for residual disease	No evidence of disease, i.e., no abnormal residual uptake in previously involved sites resulting in a complete normalization	Not specified	Yes	2	Expert
Ng et al (Ng AP, Wirth A, Seymour JF, et al: Leuk Lymphoma 48: 596-600, 2007)	2007	Qualitative analysis	Increased FDG-avidity above a baseline level, as subjectively characterized by FDG-avidity of the liver parenchyma, in a region of lymphoma, documented clinically or radiologically, at diagnosis	No residual uptake in suspected sites of lymphoma before treatment	Not specified	Yes*	3	Expert

Abbreviations: CT, computed tomography; DLBCL, diffuse large B-cell lymphoma; FDG, fluorodeoxyglucose; HL, Hodgkin's lymphoma; PET, positron emission tomography; ROC, receiver operating characteristics; SUV_{MAX}, maximum standard uptake value in region(s) of interest.
*Mid-therapy PET was interpreted without pre-therapy baseline scan in some patients.
†Only visual interpretations were taken into account.

Table A4. Quality Assessment of Studies of Positron Emission Tomography for Interim Response Assessment of Malignant Lymphoma

Study	Year	QUADAS 1: Avoidance of Spectrum Bias	QUADAS 2: Reporting of Inclusion Criteria	QUADAS 3: Appropriate Reference Standard	QUADAS 4: Avoidance of Partial Verification Bias	QUADAS 5: Avoidance of Differential Verification Bias	QUADAS 6: Avoidance of Incorporation Bias
Advanced-stage HL + DLBCL							
Kostakoglu et al (Kostakoglu L, Goldsmith SJ, Leonard JP, et al: Cancer 107:2678-2687, 2006)	2006	No	Yes	Yes	Yes	Yes	Yes
Advanced-stage HL							
Friedberg et al (Friedberg JW, Fischman A, Neuberg D, et al: Leuk Lymphoma 45:85-92, 2004)	2004	Yes	Yes	Yes*	Yes	Yes	Yes
Hutchings et al (Hutchings M, Mikhaeel NG, Fields PA, et al: Ann Oncol 16:1160-1168, 2005)	2005	No	Yes	Yes*	Yes	Yes	Yes
Gallamini et al (Gallamini A, Rigacci L, Merli F, et al: Haematologica 91:475-481, 2006)	2006	Yes	Yes	Yes*	Yes	Yes	No‡
Hutchings et al (Hutchings M, Loft A, Hansen M, et al: Blood 107:52-59, 2006)	2006	Yes	Yes	Yes†	Yes	Yes	Yes
Zinzani et al (Zinzani PL, Tani M, Fanti S, et al: Ann Oncol 17:1296-1300, 2006)	2006	Yes	Yes	Yes	Yes	Yes	Yes
Gallamini et al (Gallamini A, Hutchings M, Rigacci L, et al: J Clin Oncol 25:3746-3752, 2007)	2007	Yes	Yes	Yes†	Yes	Yes	Yes
DLBCL							
Spaepen et al (Spaepen K, Stroobants S, Dupont P, et al: Ann Oncol 13:1356-1363, 2002)	2002	Yes	Yes	Yes*	Yes	Yes	Yes
Haioun et al (Haioun C, Itti E, Rahmouni A, et al: Blood 106:1376-1381, 2005)	2005	Yes	Yes	Yes*	Yes	Yes	Yes
Mikhaeel et al (Mikhaeel NG, Hutchings M, Fields PA, et al: Ann Oncol 16:1514-1523, 2005)	2005	No	Yes	Yes*	Yes	Yes	Yes
Fruchart et al (Fruchart C, Reman O, Le Stang N, et al: Leuk Lymphoma 47:2547-2557, 2006)	2006	Yes	Yes	Yes*	Yes	Yes	Yes
Querellou et al (Querellou S, Valette F, Bodet-Milin C, et al: Ann Hematol 85:759-767, 2006)	2006	No	Yes	Yes*	Yes	Yes	Yes
Ng et al (Ng AP, Wirth A, Seymour JF, et al: Leuk Lymphoma 48:596-600, 2007)	2007	No	Yes	Yes*	Yes	Yes	Yes

(continued on following page)

Table A4. Quality Assessment of Studies of Positron Emission Tomography for Interim Response Assessment of Malignant Lymphoma (continued)

Study	QUADAS						
	7: Replicativeness of Index Test	8: Replicativeness of Reference Standard	9: Avoidance of Test Review Bias	10: Avoidance of Diagnosis Review Bias	11: Availability of Clinical Data to Test Interpreters	12: Reporting of Uninterpretable or Intermediate Results	13: Reporting of Withdrawals
Advanced-stage HL + DLBCL							
Kostakoglu et al (Kostakoglu L, Goldsmith SJ, Leonard JP, et al: Cancer 107:2678-2687, 2006)	No	Yes	Yes	No¶	No	No	No
Advanced-stage HL							
Friedberg et al (Friedberg JW, Fischman A, Neuberg D, et al: Leuk Lymphoma 45:85-92, 2004)	Yes	Yes	Yes	No	No	No	Unclear
Hutchings et al (Hutchings M, Mikhaeel NG, Fields PA, et al: Ann Oncol 16:1160-1168, 2005)	Yes	No§	Unclear	Unclear‖	Unclear	Yes	No
Gallamini et al (Gallamini A, Rigacci L, Merli F, et al: Haematologica 91:475-481, 2006)	Yes	Yes	No	No‡¶	Yes	No	Unclear
Hutchings et al (Hutchings M, Loft A, Hansen M, et al: Blood 107:52-59, 2006)	Yes	Yes	Yes	Yes#	No	No	Yes
Zinzani et al (Zinzani PL, Tani M, Fanti S, et al: Ann Oncol 17:1296-1300, 2006)	Yes	Yes	Unclear	Unclear¶	Unclear	Yes	Unclear
Gallamini et al (Gallamini A, Hutchings M, Rigacci L, et al: J Clin Oncol 25:3746-3752, 2007)	Yes	Yes	No	Yes#	Yes	No	Unclear
DLBCL							
Spaepen et al (Spaepen K, Stroobants S, Dupont P, et al: Ann Oncol 13:1356-1363, 2002)	Yes	Yes	Yes	Unclear	No	No	Unclear
Haioun et al (Haioun C, Itti E, Rahmouni A, et al: Blood 106:1376-1381, 2005)	Yes	Yes	Yes	Unclear¶	No	No	Unclear
Mikhaeel et al (Mikhaeel NG, Hutchings M, Fields PA, et al: Ann Oncol 16:1514-1523, 2005)	Yes	No§	Yes	Unclear‖	No	Yes	No
Fruchart et al (Fruchart C, Reman O, Le Stang N, et al: Leuk Lymphoma 47:2547-2557, 2006)	Yes	Yes	No	Yes#	Yes	No	Yes
Querellou et al (Querellou S, Valette F, Bodet-Milin C, et al: Ann Hematol 85:759-767, 2006)	Yes	Yes	Yes	Unclear‖¶	No	No	Unclear
Ng et al (Ng AP, Wirth A, Seymour JF, et al: Leuk Lymphoma 48:596-600, 2007)	No	No§	Unclear	Unclear‖	Unclear	No	No

Abbreviations: DLBCL, diffuse large B-cell lymphoma; HL, Hodgkin's lymphoma; QUADAS, quality assessment tool of diagnostic accuracy studies.

*Scored as "unclear" if alternative criterion was applied; none of these studies specified minimum follow-up period for continuous remission or reported the data on censoring within a year.

†Scored as "no" if alternative criterion was applied; all these studies explicitly reported at least one patient without treatment failure censored within a year.

‡One interim PET scan result was used to determine final clinical outcome because biopsy could not be performed.

§All these studies did not explicitly report the use of the standard guidelines; however, they actually employed them per unpublished data.

‖All these studies did not explicitly report the blinding of clinicians to interim PET scan; however, the scan results were timely made available to treating physicians per unpublished data.

¶Interim PET scan was not used at least to alter the preplanned treatment strategy including adjuvant involved-field radiation or high-dose chemotherapy although the blinding of treating physicians to the results was either unclear or unemployed.

#Interim PET scan results were explicitly excluded from clinical data with which treating physicians made clinical decision.

Table A5. Study Results of PET for Interim Response Assessment of Malignant Lymphoma

Study	Year	Total No.	Crude Cummulative Incidence of Treatment Failure (%)			Crude Incidence of Loss to Follow-Up ≤ 1 Year (%)	Treatment Failure (No.)									Continuing Remission (No.)			No. of Censored Patients ≤ 1 Year		
							PET Positive			PET MRU			PET Negative								
								Post-Therapy*			Post-Therapy*			Post-Therapy*							
			During Therapy	≤ 1 Year	Entire Follow-Up		During Therapy	Early	Late	During Therapy	Early	Late	During Therapy	Early	Late	PET Positive	PET MRU	PET Negative	PET Positive	PET MRU	PET Negative
Advanced-stage HL																					
Friedberg et al (Friedberg JW, Fischman A, Neuberg D, et al: Leuk Lymphoma 45:85-92, 2004)	2004	22	NR	NR	23	NR	NR	NR	4†	—	—	—	NR	NR	1†	1	—	16	NR	—	NR
Hutchings et al (Hutchings M, Mikhaeel NG, Fields PA, et al: Ann Oncol 16: 1160-1168, 2005)	2005	28	7	14	32	11	2	1	3	0	1	0	0	0	2	0	6	13	0	1	2
Gallamini et al (Gallamini A, Rigacci L, Merli F, et al: Haematologica 91:475-481, 2006)	2006	108	15	19	19	NR	13	4	1	—	—	—	3	0	0	2	—	85	NR	—	NR
Hutchings et al (Hutchings M, Loft A, Hansen M, et al: Blood 107:52-59, 2006)	2006	46	20	26	28	9	8	2	0	—	—	—	1	1	1	1	—	32	0	—	4
Kostakoglu et al (Kostakoglu L, Goldsmith SJ, Leonard JP, et al: Cancer 107: 2678-2687, 2006)	2006	10	10	50	50	0	1	4	0	—	—	—	0	0	0	0	—	5	0	—	0
Zinzani et al (Zinzani PL, Tani M, Fanti S, et al: Ann Oncol 17:1296-1300, 2006)	2006	40	18	23	23	0	7	1	0	0	1	0	0	0	0	0	3	28	0	0	0
Gallamini et al (Gallamini A, Hutchings M, Rigacci L, et al: J Clin Oncol 25: 3746-3752, 2007)	2007	106	—	10	18	NR	NR	9†	6	—	—	—	NR	2†	2	4	—	83	NR	—	NR
DLBCL																					
Spaepen et al (Spaepen K, Stroobants S, Dupont P, et al: Ann Oncol 13: 1356-1363, 2002)	2002	47	NR	NR	47	NR	NR	NR	20†	—	—	—	NR	NR	2†	0	—	25	NR	—	NR
Haioun et al (Haioun C, Itti E, Rahmouni A, et al: Blood 106: 1376-1381, 2005)	2005	83	NR	NR	39	NR	NR	NR	20†	—	—	—	NR	NR	12†	14	—	37	NR	—	NR

(continued on following page)

Table A5. Study Results of PET for Interim Response Assessment of Malignant Lymphoma (continued)

Study	Year	Total No.	Crude Cummulative Incidence of Treatment Failure (%)			Crude Incidence of Loss to Follow-Up	Treatment Failure (No.)									Continuing Remission (No.)			No. of Censored Patients ≤ 1 Year		
							PET Positive			PET MRU			PET Negative								
								Post-Therapy*			Post-Therapy*			Post-Therapy*							
			During Therapy	≤ 1 Year	Entire Follow-Up	≤ 1 Year (%)	During Therapy	Early	Late	During Therapy	Early	Late	During Therapy	Early	Late	PET Positive	PET MRU	PET Negative	PET Positive	PET MRU	PET Negative
Mikhaeel et al (Mikhaeel NG, Hutchings M, Fields PA, et al: Ann Oncol 16: 1514-1523, 2005)	2005	57	5	26	38	14	3	8	4	0	2	1	0	2	2	8	5	22	2	1	5
Fruchart et al (Fruchart C, Reman O, Le Stang N, et al: Leuk Lymphoma 47:2547-2557, 2006)	2006	35	17	23	29	3	6	1	2	—	—	—	0	1	0	6	—	19	0	—	1
Kostakoglu et al (Kostakoglu L, Goldsmith SJ, Leonard JP, et al: Cancer 107: 2678-2687, 2006)	2006	24	13	29	38	0	3	4	2	—	—	—	0	0	0	1	—	14	0	—	0
Querellou et al (Querellou S, Valette F, Bodet-Milin C, et al: Ann Hematol 85:759-767, 2006)	2006	21	5	24	29	5	1	1	1	—	—	—	0	3	0	1	—	14	0	—	1
Ng et al (Ng AP, Wirth A, Seymour JF, et al: Leuk Lymphoma 48: 596-600, 2007)	2007	44	5	14	27	2	2	3	3	—	—	—	0	1	3	4	—	28	0	—	2

Abbreviations: DLBCL, diffuse large B-cell lymphoma; HL, Hodgkin's lymphoma; MRU, minimal residual uptake; NR, not reported; PET, positron emission tomography.

*Treatment failures after completing first-line therapy were divided into two groups: early progression or relapse (within a year from the start of therapy) and late relapse (after a year).

†Data were reported as the total number of treatment failures through the entire follow-up period.

‡Data were reported as the total number of treatment failures within a year from the start of first-line therapy.