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.

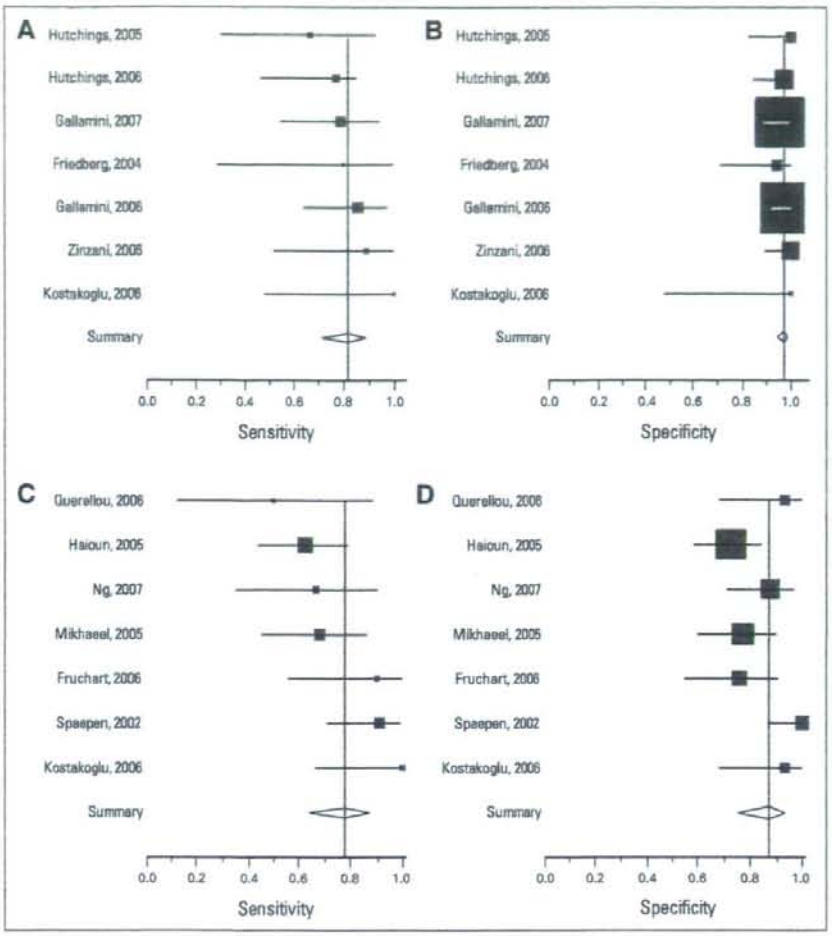

Fig 1. Sensitivity and specificity for (A, B) advanced-stage Hodgkin's lymphoma and (C, D) diffuse large B-cell lymphoma. The size of the square plotting is proportional to the number of patients with treatment failure for sensitivity and in remission for specificity. The horizontal lines are the 95%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

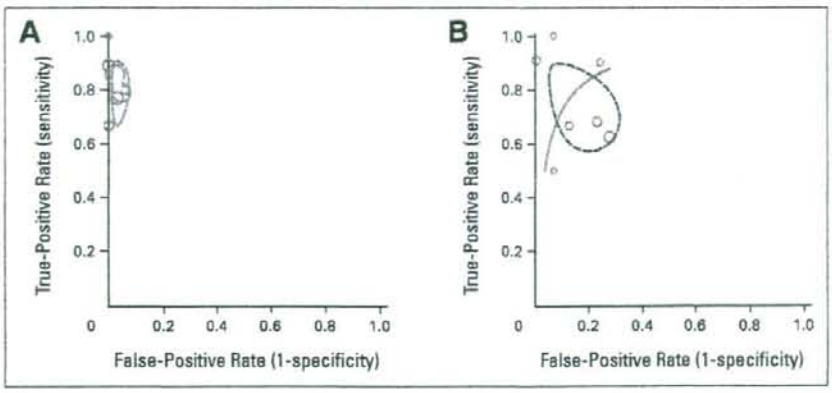

Fig 2. Receiver operating characteristic (ROC) plotting for (A) advanced-stage Hodgkin's lymphoma and (B) diffuse large B-cell lymphoma. Individual study estimates of sensitivity and 1 − specificity are shown (open circles). Summary ROC curve is presented only for DLBCL. Closed square represents the 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, Tsborska 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. Speeper 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 13:1358-1363, 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

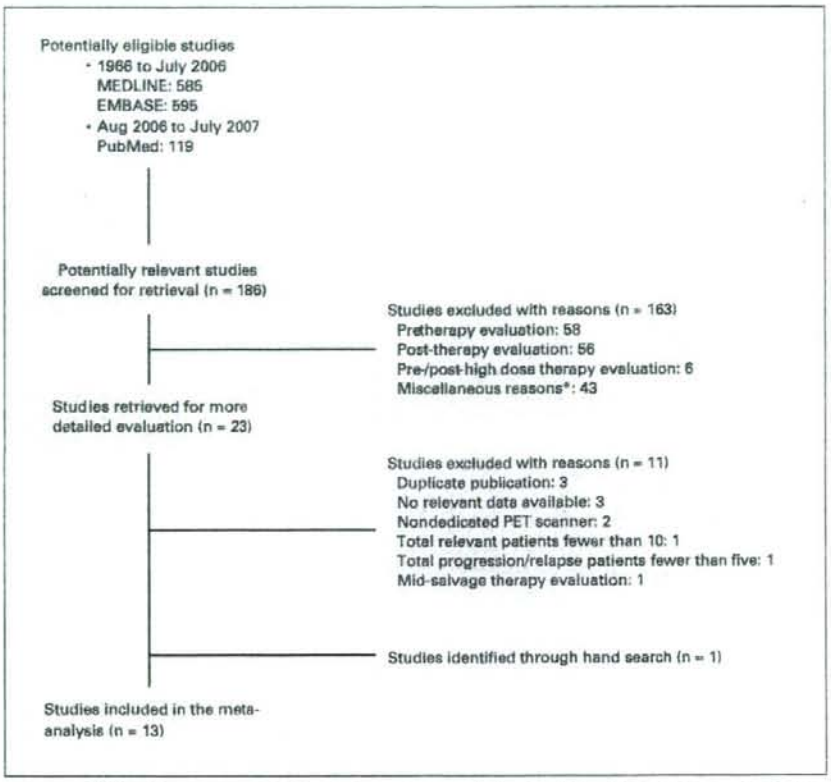

Fig A1. Article selection process. (*) Miscellaneous reasons include studies on staging evaluation at multiple different timings and contexts (n = 10), studies focusing on lymphoma involvement in a specific organ or anatomic region (n = 7), review articles (n = 5), studies focusing exclusively on positive 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).

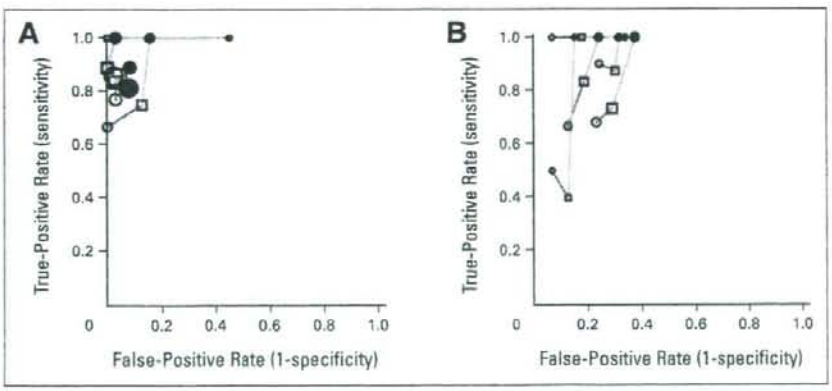

Fig A2. Receiver operating characteristic plotting for (A) advanced-stage Hodgkin's lymphoma and (B) diffuse large B-cell lymphoma. Individual study estimates of sensitivity and 1 − specificity are shown based on the duration of clinical follow-up: all treatment failures (open circles), early treatment 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-quantitative 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		Minimal Residual Uptake	Reading Condition: Availability of Pretherapy PET to Interpreters of Interim PET	Interpreter	
			Positive	Negative			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	2: Reporting of Inclusion Criteria	3: Appropriate Reference Standard	4: Avoidance of Partial Verification Bias	5: Avoidance of Differential Verification Bias	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	Not
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	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
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
Quereliou et al (Quereliou 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.

日本臨牀　67巻 増刊号1（2009年1月28日発行）別刷

がん薬物療法学

—基礎・臨床研究のアップデート—

VII. 抗悪性腫瘍薬の臨床試験—行政とのかかわり—

適応外医薬品・未承認薬を用いた臨床試験

堀田知光

適応外医薬品・未承認薬を用いた臨床試験

Clinical trials by off-label or unapproved drug use

堀田知光

Key words：抗癌剤，未承認薬，適応外使用，臨床試験

はじめに

海外で承認されているが，我が国ではどの疾患に対してもまだ適応承認のない医薬品を国内「未承認薬」といい，既承認の効能・効果および用法・用量によらない薬剤使用を「適応外使用」という．我が国におけるがんの臨床試験は海外の先進諸国に比べて周回遅れであるといわれる．海外で承認されている抗腫瘍薬が，我が国では治験の着手，進捗ないし承認が遅れているために新規薬剤を組み込んだ臨床試験の実施が困難であり，またエビデンスに基づく標準的治療確立のための臨床試験には承認された効能・効果の適応のみでなく新たな用法や用量の組み合わせの工夫が避けられないが，現状では保険上の制約のために実施は困難になっている．こうした状況は我が国から新たなエビデンスを発信するための障壁として国際的競争力の低下を招いているばかりでなく，何よりも我が国のがん患者に不利な状況を作り出してきており，未承認薬や適応外医薬品を用いた研究者主導の臨床試験を可能にする法的な枠組みの必要性が指摘されてきた．このような背景において平成20年度から「高度医療評価制度」が導入された．この制度は一定の条件のもとに未承認薬や適応外医薬品を用いた臨床試験を保険診療と併用可能にするものである．本制度を適切に活用することにより我が国において魅力的で質の高い臨床試験が活性化されることが期待される．

1. 適応外使用はどこが問題なのか

a. 保険診療は「療担規則」によって規制されている

適応外医薬品を用いた臨床試験を保険診療で行うことはこれまでは高度先進医療もしくは薬事法(昭和35年法律145号)[1]に基づく治験(医師主導を含む)以外ではできなかった．その根拠は「保険医療機関及び保険医療養担当規則」(療担規則)(昭和32年4月30日，厚生省令第15号，最終改正：平成20年3月5日)[2]の第2章「保険医の診療方針等」における第18条に「保険医は，特殊な療法又は新しい療法については，厚生労働大臣の定めるものの他行ってはならない．ただし，特定認定保険機関において行う高度先進医療である療養についてはこの限りでない」と明記されている．これが保険医療における「特殊医療もしくは研究的治療」の禁止条項と呼ばれるものである．また同第19条には「保険医は，厚生労働大臣の定める医薬品以外の薬物を患者に施用し，又は処方してはならない．ただし，薬事法第2条第16項に規定する治験に係る診療において，当該治験の対象とされる薬物を使用する場合その他 厚生労働大臣が定める場合については，この限りでない」とされている．これがいわゆる「混合診療」の原則禁止条項と呼ばれている規則である．これによって治

Tomomitsu Hotta: Nagoya Medical Center, National Hospital Organization 国立病院機構名古屋医療センター

0047-1852/09/¥40/頁/JCLS

験以外での未承認薬や適応外医薬品の使用は，混合診療の禁止に抵触し，原則として全額自己負担になるという解釈である．

b.「療担規則」の解釈に新たな動き

適応外医薬品を用いた臨床試験が療担規則に違反するかどうかについて国の解釈を示す事例が平成19年にあった．国内で行われた高血圧症，高脂血症または糖尿病を有する高齢者に対するアスピリンの一次予防投与のリスク/ベネフィットを評価する大規模臨床試験が療担規則に違反するのではないかとする国会議員からの質問[3]に対して内閣総理大臣が答弁書[4]という形で表明したものである．これによると薬剤費を患者負担にしていないので療担規則第5条に違反せず，我が国においては効能効果の適応がないが海外で承認されており，アスピリンの本来的な薬理作用であることを踏まえると療担規則第18条の特殊療法の禁止条項に当たらない．さらに，すでに薬価収載医薬品であるので，その使用について療担規則第19条に違反するものではないとして，研究対象薬投与以外の診療について「保険請求を行うことは可能である」と表明している．それならば適応外医薬品を用いた臨床試験はかなりの自由度をもって実施できることになる．しかし，適応外使用と混合診療問題についてはどこまでが保険診療との併用が可能なのかについての線引きが曖昧で，明確な法的な枠組みの必要性が指摘されていた．

2. 高度医療評価制度の導入

このような状況で，厚生労働省は平成20年度より未承認薬や適応外医薬品を用いた臨床試験を保険との併用を可能とする枠組みを「高度医療評価制度」[5]として創設した．制度は「薬事法の承認が得られていない医薬品・医療機器の使用を伴う先進的な医療技術については，一般的な治療法ではないとの理由から原則として保険との併用が認められていないが，今般，これらの医療技術のうち，一定の要件の下に行われるものについて，先進医療の一類型として保険診療との併用を認め，薬事法による申請等に繋がる科学的評価可能なデータ収集の迅速化を図ること」を創設の趣旨としている．

a. 対象となる医療技術

(1) 薬事法上の承認を受けていない医薬品（未承認薬）の使用を伴う医療技術

(2) 薬事法上の承認を受けて製造販売されている医薬品を，承認事項に含まれない用量，用法または他の効能または効果等を目的とした使用（適応外使用）を伴う医療技術

b. 医療機関の要件

(1) 特定機能病院または，①緊急時の対応および②医療安全対策の体制がとられている保険医療機関であること

(2) 臨床研究に関する倫理指針（平成16年厚労省告示第459号）[6]に適合する臨床研究実施体制を有すること

(3) 高度医療として実施される医療技術において使用する医薬品の管理体制，入手方法が適切であること

(4) 実施医療機関の長が院内で行われる全ての高度医療について実施責任医師，研究内容等を把握できる体制を確保すること

c. 高度医療に係る要件

(1) 国内外の使用実績や有用性を示す文献等により，安全性および有効性の確保が期待できる科学的な根拠を有する医療技術であること

(2) 試験計画が臨床研究に関する倫理指針に基づいた内容であること

これには結果責任と補償の内容，治療の内容，合併症や副作用の可能性および費用等について，事前に患者やその家族に説明し文書により同意を得ることが明記されており，試験記録の適切な保管や管理によりデータの信頼性が確保されていることとされ，次の体制の確保に努めることが求められている．

・データマネジメント体制

・多施設共同研究の場合のモニタリング体制

このような要件を満たした医療技術について，厚労省医政局研究開発振興課に事前相談をした上で申請書と必要な添付書類（実施体制，実施計画，宣誓書，概要，費用の積算根拠，同意文書および同意文書の雛形，先進医療届出書，当該技術の内容を記述した論文および有効性を評

価した原著論文：いずれも査読のある学術雑誌でそれぞれ一編以上)を厚労省医政局長に提出することとされている．実施に当たって実施状況について公表するとともに，定期的に厚労省に報告することが定められている．予期しない重篤な有害事象や不具合が生じた場合には，速やかに必要な対応を行うとともに実施医療機関に周知し，対応状況・結果の公表と厚労省への報告が必要である．また，厚労省が事前通告なく行う実地調査に応じることが求められている．

d. 未承認薬・適応外医薬品の入手

さて，現実的に懸案となる未承認薬や適応外医薬品の入手の件については，①実施責任医師の指示の下での自家製造(委託製造を含む)，あるいは②実施責任医師の指示による個人輸入のいずれかとされている．しかし，これらは未承認薬に対する入手方法であって，適応外医薬品の入手については明記されていない．通知文書の不備か，意図的かは判じかねるが，適応外使用については一般的には，次の方法が想定される．

1) 保険にレセプト請求

療担規則の「研究的治療の禁止」条項に抵触することが明らかであるので，査定・返還の対象となる．このような不都合があるため，高度医療評価制度ができた経緯からすると，上記原則は容易には変わらないと予想される．

2) 研究費購入もしくは個人負担

研究代表者が一括購入して実施施設に配布する方法がある．この対応がもっとも適切であると考えられる．しかし，多施設共同の大規模試験で薬剤費が高額の場合は対応しきれない可能性がある．

3) 企業からの提供

企業の理解が得られれば，研究者側および患者側にとって負担が少なく現実的な対応といえる．しかし，公正取引規約上の配慮と科学研究における利益相反(Conflict of Interest: COI)への配慮が必要である．

公正取引上では日本製薬工業協会は自主基準[7]としてプロモーションコードに「試用医薬品の提供」の項があるが，医薬品情報提供の一手段として用いるもので，必要最小限にとどめることが定められていること，「対照薬の提供及び譲受に関する申し合わせ」において企業が新薬開発に際して行う比較試験を想定した基準があるものの，医師主導の臨床試験への薬剤提供に関する記述は見あたらない．著者が得ている情報では以下の条件を満たす場合には薬剤の無償提供が可能とする対応が取られた例がある．その場合の条件は以下であるという．

- 当該医薬品を用いた治療によって患者の利益が大きいこと
- 厳正な臨床試験登録患者に限定
- 厳重な薬剤管理
- 供与薬剤を保険診療請求しない
- 有害事象の報告
- 医師の責任に基づく厳正な臨床試験体制の確立

製薬協もしくは公正取引協議会として医師主導の臨床試験における薬剤の無償提供のガイドラインを示すことを強く求めたい．

e. 利益相反(COI)の管理

企業から薬剤提供を受ける場合の研究者側の問題として，臨床試験の科学性と倫理性に疑問をもたれないように臨床研究に関する倫理指針に準拠して，透明性の確保に配慮する必要がある．厚労科学研究費で行う臨床試験では「厚生労働科学研究における利益相反(Conflict of Interest: COI)の管理に関する指針」(平成20年3月31日科発第0331001号)[8]により科学研究の公正性，信頼性を確保するために，利害関係者とのかかわりについて適正に対応することが求められている．経済的な利益関係の対象として給与，謝金，株式，受託研究費や寄付等に限定せず，何らかの金銭的価値をもつものが含まれるので，薬価がついた製販後の医薬品もこれに該当すると理解される．もとより，本指針は，意欲ある研究者が安心して研究に取り組めるよう環境を整備する趣旨で策定されたものであり，研究をバイアスから保護し，社会から研究の客観性と公平性に疑問をもたれることのないように定められたものであるので，適切に対応することが研究者保護につながるものと理解すべき

である．

f. 補償問題

さて，もう一つの問題は，適応外医薬品によって健康被害を生じた場合に誰がどのように補償するのかという問題である．治験の場合には企業は「治験保険」に加入しており，承認薬による適正な適応と用法・用量で治療を行って生じた健康被害については医薬品医療機器総合機構の医薬品健康被害救済制度が機能している．しかし，適応外使用や研究費購入，無償提供薬での健康被害にはこのような補償はない．現状では抗癌剤については補償できないことを同意説明文書に明記して同意を得ることで対応するしかないが，研究者が行う臨床試験研究に関しての環境整備がこの面においても必要である．補償問題については厚労省において新たな補償保険制度の導入が検討されつつある．

おわりに

平成20年度から導入された高度医療評価制度によって，未承認薬や適応外医薬品を用いた臨床試験が保険診療と併用することが公に可能となったことにより我が国の臨床試験を取り巻く環境に大きな変化をもたらし，我が国から臨床試験のエビデンスが発信されることが期待されている．しかし，この制度は当該臨床試験で得られた成果が保険承認につながることをめざすものであるので，実施に当たってはGCP省令そのものではないにしてもこれに準じたレベルでの対応が求められている．そのために現在改訂作業が進められている「臨床研究に関する倫理指針」に準拠することが求められている．今後は未承認薬や適応外医薬品を用いた臨床試験はこの制度以外での実施は難しくなるものと予想される．本制度を適正に活用し，我が国の臨床試験の活性化とレベルアップが期待される．

■文　献

1）厚生労働省：「薬事法」［http://www.houko.com/00/01/S35/145.HTM］
2）厚生労働省：「保険医療機関及び保険医療養担当規則」（昭和32年4月30日，厚生省令第15号，最終改正：平成20年3月5日）［http://law.e-gov.go.jp/htmldata/S32/S32F03601000015.html］
3）郡　和子：いわゆる混合診療問題及び未承認薬剤の授受に関する質問主意書［http://www.shugiin.go.jp/itdb_shitsumon.nsf/html/shitsumon/a166380.htm?OpenDocument］
4）内閣総理大臣　阿倍晋三：衆議院議員郡和子君提出いわゆる混合診療問題及び未承認薬剤の授受に関する質問に対する答弁書［http://www.shugiin.go.jp/itdb_shitsumon.nsf/html/shitsumon/b166380.htm?OpenDocument］
5）厚生労働省医政局長：「高度医療に係る申請等の取扱い及び取扱い上の留意事項について」（平成20年3月31日 医政発第0331022号）［http://www.mhlw.go.jp/topics/2008/04/tp0402-1.html］
6）厚生労働省：臨床研究に関する倫理指針（平成15年7月30日制定，平成16年12月28日全部改正）［http://www.imcj.go.jp/rinri/index.html］
7）日本製薬工業協会ホームページ：自主基準［http://www.jpma.or.jp/about/basis/guide/control.html］
8）厚生科学審議会科学技術部会：第6回厚生労働科学研究における利益相反に関する検討会（平成20年1月22日）［http://www.mhlw.go.jp/shingi/2008/01/s0122-2.html］

VOLUME 27 · NUMBER 3 · JANUARY 20 2009

JOURNAL OF CLINICAL ONCOLOGY

SPECIAL ARTICLE

Definition, Prognostic Factors, Treatment, and Response Criteria of Adult T-Cell Leukemia-Lymphoma: A Proposal From an International Consensus Meeting

Kunihiro Tsukasaki, Olivier Hermine, Ali Bazarbachi, Lee Ratner, Juan Carlos Ramos, William Harrington Jr, Deirdre O'Mahony, John E. Janik, Achiléa L. Bittencourt, Graham P. Taylor, Kazunari Yamaguchi, Atae Utsunomiya, Kensei Tobinai, and Toshiki Watanabe

From the Nagasaki University, Nagasaki, Japan; Hospital Necker, University Paris V Rene Descartes and CNRS UMR 8147, Paris, France; Department of Internal Medicine, American University of Beirut, Beirut, Lebanon; Washington University, St Louis, MO; National Cancer Institute, Bethesda, MD; Federal University of Bahia, Bahia, Brazil; Imperial College London, London, United Kingdom; National Institute of Infectious Diseases; National Cancer Center Hospital; Tokyo University, Tokyo; and Imamura Bun-in Hospital, Kagoshima, Japan.

Submitted May 18, 2008; accepted September 17, 2008; published online ahead of print at www.jco.org on December 8, 2008.

Supported in part by the intramural Research Program of the National Cancer Institute, National Institutes of Health.

Presented in part at the 13th International Conference on Human Retrovirology: HTLV, May 22-25, 2007, Hakone, Japan, and the 49th Annual Meeting of the American Society of Hematology, December 8-11, 2007, Atlanta, GA.

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

Corresponding author: Kunihiro Tsukasaki, MD, PhD, Department of Molecular Medicine and Hematology, Molecular Medicine Unit, Atomic Bomb Disease Institute, Nagasaki University Graduate School of Biomedical Science, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan; e-mail: tsukasak@net.nagasaki-u.ac.jp.

The Acknowledgment is included in the full-text version of this article, available online at www.jco.org. It is not included in the PDF version (via Adobe® Reader®).

0732-183X/09/2703-453/$20.00

DOI: 10.1200/JCO.2008.18.2428

ABSTRACT

Adult T-cell leukemia-lymphoma (ATL) is a distinct peripheral T-lymphocytic malignancy associated with a retrovirus designated human T-cell lymphotropic virus type I (HTLV-1). The diversity in clinical features and prognosis of patients with this disease has led to its subclassification into the following four categories: acute, lymphoma, chronic, and smoldering types. The chronic and smoldering subtypes are considered indolent and are usually managed with watchful waiting until disease progression, analogous to the management of some patients with chronic lymphoid leukemia (CLL) or other indolent histology lymphomas. Patients with aggressive ATL generally have a poor prognosis because of multidrug resistance of malignant cells, a large tumor burden with multiorgan failure, hypercalcemia, and/or frequent infectious complications as a result of a profound T-cell immunodeficiency. Under the sponsorship of the 13th International Conference on Human Retrovirology: HTLV, a group of ATL researchers joined to form a consensus statement based on established data to define prognostic factors, clinical subclassifications, and treatment strategies. A set of response criteria specific for ATL reflecting a combination of those for lymphoma and CLL was proposed. Clinical subclassification is useful but is limited because of the diverse prognosis among each subtype. Molecular abnormalities within the host genome, such as tumor suppressor genes, may account for these diversities. A treatment strategy based on the clinical subclassification and prognostic factors is suggested, including watchful waiting approach, chemotherapy, antiviral therapy, allogeneic hematopoietic stem-cell transplantation (alloHSCT), and targeted therapies.

J Clin Oncol 27:453-459. © 2008 by American Society of Clinical Oncology

DEFINITION

Adult T-cell leukemia-lymphoma (ATL) is a distinct peripheral T-lymphocytic malignancy associated with a retrovirus designated human T-cell leukemia virus type 1 or human T-cell lymphotropic virus type 1 (HTLV-1).[1-3] We recommend following the WHO classification of ATL published in 2001.[4]

PROGNOSTIC FACTORS

Major prognostic indicators[5-8] for ATL have been elucidated in 854 patients; advanced performance status (PS), high lactic dehydrogenase (LDH) level, age ≥ 40 years, more than three involved lesions, and hypercalcemia[5] are prognostic factors that have been identified by multivariate analysis. These factors were used to construct a risk model.[5] Additional factors associated with poor prognosis include thrombocytopenia,[9] eosinophilia,[10] bone marrow involvement,[11] high interleukin-5 serum level,[12] C-C chemokine receptor 4 expression,[13] lung resistance–related protein,[14] *p53* mutation,[15] and *p16* deletion.[9] For the chronic type of ATL, high LDH, high blood urea nitrogen, and low albumin levels have been identified as poor prognostic factors by multivariate analysis.[6] Univariate analysis has revealed that neutrophilia,[11] *p16* deletion,[9] and chromosomal deletion detected by comparative genomic hybridization[16] are associated with poor prognosis in chronic ATL. In contrast, chronic lymphoid leukemia (CLL)–like morphology of ATL cells was associated with longer transformation-free survival of chronic ATL.[17] Primary cutaneous tumoral type, although generally included among smoldering ATL, was a poor prognostic factor by univariate analyses.[18] A combination of these and more novel prognostic factors may be superior to elucidate better risk ATL groups for stratification of treatment decision than the Shimoyama criteria, which stratify

