Prospective International Cohort Study Demonstrates Inability of Interim PET to Predict Treatment Failure in Diffuse Large B-Cell Lymphoma

Robert Carr[1], Stefano Fanti[2], Diana Paez[3], Juliano Cerci[4], Tamás Györke[5,6], Francisca Redondo[7], Tim P. Morris[8], Claudio Meneghetti[9], Chirayu Auewarakul[10], Reena Nair[11], Charity Gorospe[12], June-Key Chung[13], Isinsu Kuzu[14], Monica Celli[2], Sumeet Gujral[15], Rose Ann Padua[16], Maurizio Dondi[3], and the IAEA Lymphoma Study Group

[1]Department of Haematology, Guy's and St. Thomas' Hospital, King's College, London, United Kingdom; [2]Policlinico S. Orsola Malpighi, Universita de Bologna, Bologna, Italy; [3]Division of Human Health, Department of Nuclear Sciences and Application, International Atomic Energy Agency, Vienna, Austria; [4]Quanta Diagnóstico e Terapia, Curitiba, Sâo Paulo, Brazil; [5]ScanoMed Medical Diagnostic Ltd., Budapest, Hungary; [6]Department of Nuclear Medicine, Semmelweis University, Budapest, Hungary; [7]Oncologic Clinic, Fundación Arturo Lopez Perez, Santiago, Chile; [8]Medical Research Council Clinical Trials Unit, University College London, London, United Kingdom; [9]Hospital de Clinicas, Universidade de Sâo Paulo, Sâo Paulo, Brazil; [10]Chulabhorn Cancer Centre and Faculty of Medicine Siriraj Hospital, Bangkok, Thailand; [11]Department of Medical Oncology, Tata Memorial Hospital, Mumbai, India; [12]St Luke's Medical Centre, Manila, Philippines; [13]Seoul National University Hospital, Seoul, South Korea; [14]Department of Pathology, Ankara University School of Medicine, Ankara, Turkey; [15]Department of Pathology, Tata Memorial Hospital, Mumbai, India; and [16]Université Paris-Diderot, UMRS-940, Institut d'Hématologie, Hôpital Saint-Louis, Paris, France

The International Atomic Energy Agency sponsored a large, multinational, prospective study to further define PET for risk stratification of diffuse large B-cell lymphoma and to test the hypothesis that international biological diversity or diversity of healthcare systems may influence the kinetics of treatment response as assessed by interim PET (I-PET). **Methods:** Cancer centers in Brazil, Chile, Hungary, India, Italy, the Philippines, South Korea, and Thailand followed a common protocol based on treatment with R-CHOP (cyclophosphamide, hydroxyadriamycin, vincristine, prednisolone with rituximab), with I-PET after 2–3 cycles of chemotherapy and at the end of chemotherapy scored visually. **Results:** Two-year survivals for all 327 patients (median follow-up, 35 mo) were 79% (95% confidence interval [CI], 74%–83%) for event-free survival (EFS) and 86% (95% CI, 81%–89%) for overall survival (OS). Two hundred ten patients (64%) were I-PET–negative, and 117 (36%) were I-PET–positive. Two-year EFS was 90% (95% CI, 85%–93%) for I-PET–negative and 58% (95% CI, 48%–66%) for I-PET–positive, with a hazard ratio of 5.31 (95% CI, 3.29–8.56). Two-year OS was 93% (95% CI, 88%–96%) for I-PET–negative and 72% (95% CI, 63%–80%) for I-PET–positive, with a hazard ratio of 3.86 (95% CI, 2.12–7.03). On sequential monitoring, 192 of 312 (62%) patients had complete response at both I-PET and end-of-chemotherapy PET, with an EFS of 97% (95% CI, 92%–98%); 110 of these with favorable clinical indicators had an EFS of 98% (95% CI, 92%–100%). In contrast, the 107 I-PET–positive cases segregated into 2 groups: 58 (54%) achieved PET-negative complete remission at the end of chemotherapy (EFS, 86%; 95% CI, 73%–93%); 46% remained PET-positive (EFS, 35%; 95% CI, 22%–48%). Heterogeneity analysis found no significant difference between countries for outcomes stratified by I-PET. **Conclusion:** This large international cohort delivers 3 novel findings: treatment response assessed by I-PET is comparable across disparate healthcare systems, secondly a negative I-PET findings together with good clinical status identifies a group with an EFS of 98%, and thirdly a single I-PET scan does not differentiate chemoresistant lymphoma from complete response and cannot be used to guide risk-adapted therapy.

Key Words: diffuse large B-cell lymphoma; positron emission tomography; prospective observational study; risk stratification; risk-adapted therapy

Large studies of diffuse large B-cell lymphoma (DLBCL) in western populations have demonstrated event-free survivals (EFSs) with R-CHOP (cyclophosphamide, hydroxyadriamycin, vincristine, prednisolone with rituximab) of 79% at 3 y in adults aged younger than 60 y and 75% at 2 y in patients of all ages (1,2).

Much effort has been invested to prospectively discriminate between patients with a high probability of prolonged EFS and those who are unlikely to be cured by standard therapy so that their chance of cure may be increased by early intensification. Indicators including the International Prognostic Index (IPI), cell-of-origin tissue phenotype, and gene expression profiles have defined subgroups with predicted better or worse outcomes (3–6). Investigational high-throughput gene sequencing is revealing the biological heterogeneity of DLBCL, which may guide more targeted treatment in the future, but until now prospective personalization of each patient's treatment remained elusive (7,8).

Proof of principle that the speed of response to treatment as an indicator of tumor chemosensitivity and ultimate cure was first demonstrated to be a powerful predictor of individual outcome in childhood lymphoblastic leukemia (9). PET uses preferential accumulation within tumor cells of ^{18}F-FDG to measure glucose

metabolic activity as a surrogate for cell viability. The hope has been that the speed of metabolic response, as judged by reduction in intensity of ^{18}F-FDG uptake by the tumor early in treatment as a marker of chemosensitivity, might similarly identify rapidly responding cases with a high likelihood of cure and incomplete responders who would benefit from early treatment intensification.

A decade of studies has demonstrated that rapid response on an interim PET (I-PET) scan after 2–4 cycles of chemotherapy predicts a comparatively better outcome, but to a variable degree between studies, whereas no study has identified a I-PET–stratified patient group with a high enough probability of treatment failure to provide sound clinical basis for directing therapy (10–12). Most studies report single-center experience in Europe or North America where PET is well established.

At a time when PET scanning is increasing in developing countries, we questioned whether ethnic, economic, and environmental diversity might result in different disease biology or whether advanced disease at presentation, compared with the western world, might confound the utility of risk prediction by an I-PET scan early in treatment. If so, it would limit the global generalizability of data from predominantly Caucasian populations in high-income countries.

To address these issues, the International Atomic Energy Agency (IAEA) initiated a Coordinated Research Project to examine PET monitoring for risk stratification of DLBCL in 5 geographic regions to inform international practice. Coordinated Research Projects, enshrined under Article III of the IAEA's statute, facilitate the international development of the practical use of atomic energy for peaceful purposes and promote the bringing together of researchers in both developing and industrialized countries to solve a problem of common interest (13).

The primary aim of this prospective international cohort study was to define, with the greater precision afforded by a large cohort, whether the rate of response to treatment as assessed by a mid-treatment I-PET scan could achieve clinically useful prediction of outcomes at 2 y for individual patients. The secondary aim was to establish whether there was clinically important variation in PET-stratified outcomes between participating countries.

MATERIALS AND METHODS

The project was approved by the IAEA and protocol developed jointly at 2 investigator meetings in 2006 and 2008.

Eligibility Criteria and Treatment Protocol

Patients with DLBCL (age, ≥16 y) who had provided informed consent were recruited. Exclusions were cancer within the preceding 5 y, steroid therapy before the staging scan, and no ^{18}F-FDG–avid disease on baseline PET. Diagnosis was based on biopsy with immunohistochemistry and classification by World Health Organization criteria (14). All patients were staged by PET/CT, or PET and CT separately, and iliac crest marrow biopsy.

The treatment protocol was for 6 cycles of R-CHOP at 21-d intervals. To accommodate clinician preference and local practice, up to 8 cycles was permitted. Omission of rituximab was allowed in recognition that some eligible patients might otherwise be excluded for financial reasons. Scan results were reported to treating clinicians, but modification to planned treatment on the basis of the I-PET response was not permitted. Treatment escalation in response to a positive I-PET result was classified as treatment failure (see the "Classification of Events" section).

Radiotherapy

Consolidation radiotherapy, if planned as part of primary treatment (e.g., to sites of bulk disease or to specific sites of extranodal disease), was permitted, as directed by local practice. Preplanned radiotherapy was deemed consolidation only if given after a negative end-treatment PET and confirmation of complete response defined by international criteria (15).

PET Scheduling and Reporting

Scans were required at 3 time points: before treatment, mid treatment (I-PET), and end-chemotherapy (E-PET). The I-PET scan was recommended after 2 cycles of chemotherapy, at a maximum interval from the preceding treatment (median treatment to scan interval, 18 d [interquartile range, 17–21 d]). In recognition of technical and scheduling constraints, I-PET after 3 cycles was permitted and in exceptional circumstances after 4 cycles. The protocol stipulated a minimum of a 4- to 8-wk interval between final chemotherapy and E-PET.

I-PET scan reporting was based on visual assessment and classified into 4 categories: negative/CR (resolution of abnormal ^{18}F-FDG uptake at sites of disease identified on staging PET, with any residual ^{18}F-FDG uptake less than or equal to the mediastinal blood pool), complete response with minimal residual uptake (CR-MRU) (residual low-level ^{18}F-FDG uptake at disease sites greater than mediastinum but less than or equal to physiologic uptake in liver), positive (residual or increased ^{18}F-FDG uptake with intensity greater than liver at a site of known disease), and mixed response (reduction in ^{18}F-FDG uptake at some disease sites, with increased ^{18}F-FDG uptake at other existing or new sites). For outcome analysis, scans scored as CR-MRU were grouped with PET-negative; mixed response was classified as PET-positive.

The study commenced in 2008, before the Deauville 5-point classification was devised in
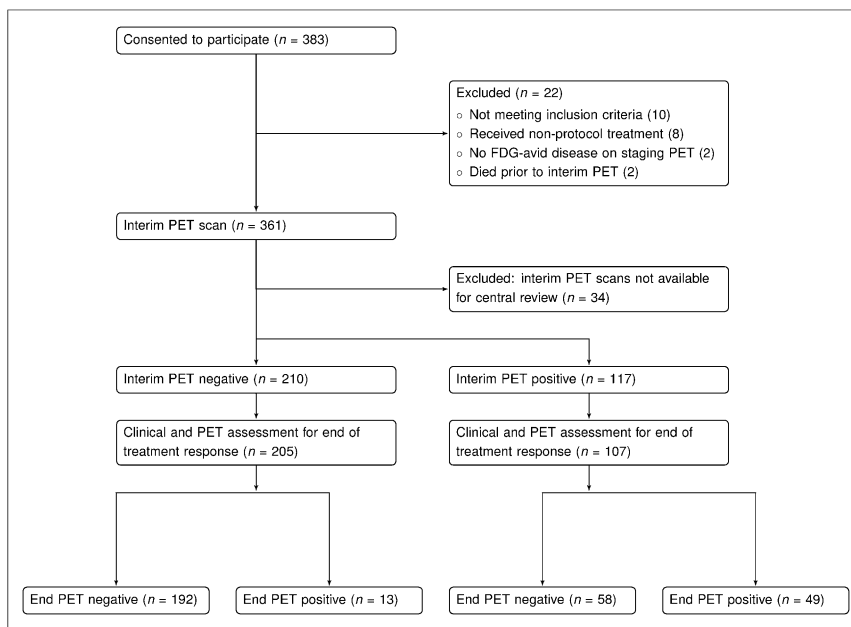

FIGURE 1. Consort diagram.

Consented to participate (n = 383)

Excluded (n = 22)
○ Not meeting inclusion criteria (10)
○ Received non-protocol treatment (8)
○ No FDG-avid disease on staging PET (2)
○ Died prior to interim PET (2)

Interim PET scan (n = 361)

Excluded: interim PET scans not available for central review (n = 34)

Interim PET negative (n = 210)

Interim PET positive (n = 117)

Clinical and PET assessment for end of treatment response (n = 205)

Clinical and PET assessment for end of treatment response (n = 107)

End PET negative (n = 192)

End PET positive (n = 13)

End PET negative (n = 58)

End PET positive (n = 49)

TABLE 1
Patient and Disease Characteristics

Characteristic	Brazil	Chile	Hungary	India	Italy	South Korea	Philippines	Thailand	Total
No. of patients	61	47	65	32	49	9	20	44	327
Sex (M)	29 (48)	27 (57)	35 (54)	22 (69)	23 (47)	6 (67)	8 (40)	23 (52)	173 (53)
Ethnicity									
Asian	0	0	0	32	0	9	20	44	105
Caucasian	0	47	65	0	48	0	0	0	160
Chinese	0	0	0	0	1	0	0	0	1
Mixed	61	0	0	0	0	0	0	0	61
Age at diagnosis (y)									
Median	54	59	56	53	55	56	52	55	55
Quartiles	45, 65	46, 65	43, 68	47, 57	43, 66	54, 60	41, 64	45, 63	44, 64
WHO/ECOG performance status									
0	31 (51)	23 (48)	28 (43)	15 (47)	35 (72)	6 (67)	0	8 (18)	146 (45)
1	29 (47)	12 (35)	23 (35)	16 (50)	7 (14)	3 (33)	15 (75)	25 (57)	130 (40)
2	1 (2)	4 (9)	11 (17)	1 (3)	5 (10)	0	5 (25)	9 (20)	36 (11)
3	0	4 (9)	2 (3)	0	1 (2)	0	0	2 (5)	9 (3)
4	0	4 (9)	1 (2)	0	1 (2)	0	0	0	6 (2)
IPI score									
0–1	27 (44)	19 (40)	33 (51)	14 (44)	20 (41)	2 (22)	6 (30)	14 (32)	135 (41)
2	17 (28)	7 (15)	10 (15)	13 (41)	13 (27)	3 (33)	6 (30)	12 (28)	81 (25)
3	14 (23)	10 (21)	11 (17)	4 (13)	10 (20)	3 (33)	5 (25)	11 (25)	68 (21)
4–5	3 (5)	11 (23)	11 (17)	1 (3)	6 (12)	1 (11)	3 (15)	7 (11)	43 (13)
Clinical stage									
I	4 (7)	7 (15)	7 (11)	1 (3)	1 (4)	0	0	1 (2)	20 (6)
II	20 (33)	9 (19)	21 (32)	12 (38)	7 (14)	4 (44)	7 (35)	15 (34)	96 (29)
III	8 (13)	12 (26)	8 (12)	11 (34)	9 (18)	2 (22)	7 (35)	10 (23)	65 (20)
IV	29 (48)	19 (40)	29 (45)	8 (25)	30 (63)	3 (33)	6 (30)	18 (41)	145 (44)
Extranodal sites ≥ 2	18 (30)	19 (40)	14 (22)	3 (9)	13 (27)	5 (56)	7 (35)	16 (36)	95 (29)
LDH > normal	31 (51)	20 (43)	30 (46)	22 (69)	20 (41)	7 (78)	10 (50)	26 (59)	166 (51)
Bulky disease > 5 cm	29 (48)	23 (49)	35 (54)	12 (38)	27 (55)	1 (11)	11 (55)	26 (59)	164 (50)

WHO/ECOG = World Health Organization/Eastern Cooperative Oncology Group.
Data in parentheses are percentages.

2009 (15). The scoring scheme used here was based on its antecedent, the validated 5-point London Criteria, modified to combine Deauville-equivalent scores 1 and 2 into a single negative category and scores 4 and 5 into a positive category (16). CR-MRU (equivalent to Deauville 3) was reported as a separate score, as at that time it was uncertain whether or not this level of ^{18}F-FDG uptake represented persistent disease at I-PET.

As recommended for multicenter PET-stratification studies, scans were reviewed by the 4 lead nuclear medicine physicians working together on a common platform at the final collaborator meeting (16, 17). Reviewers were masked to clinical details and patient outcomes. Classification of PET response was by consensus.

Classification of Events

Study events were relapse after complete remission; death from any cause; treatment escalation for progressive disease while on treatment; and disease progression or failure to achieve complete remission at end-chemotherapy based on the revised response criteria for PET, with confirmation by biopsy that residual or increased ^{18}F-FDG uptake was due to lymphoma (18). Each patient record was reviewed with the country chief investigator during the final collaborator meeting to ensure correct classification of events.

Research Regulation and Data Protection

Each country gained research ethics approval for the study protocol and patient information from the appropriate national or local Ethics Review Board. Fully informed consent was an inclusion criterion for recruitment. Signed consent forms were kept by the local investigators. To ensure confidentiality while sharing data internationally, cases were assigned a numeric code, and only 2 identifiers for data validation—initials and date of birth—were recorded in the central database (19).

Statistical Methods

Follow-up continued until 75% patients had reached 2 y or died. Cases lost to follow-up were censored at date of last known disease status. Survival was estimated using Kaplan–Meier methods, with the date of first treatment as origin.

TABLE 2
Treatment, Monitoring, and Outcomes

Parameter	Brazil	Chile	Hungary	India	Italy	South Korea	Philippines	Thailand	Total
Patients (n)	61	47	65	32	49	9	20	44	327
R-CHOP*	60 (98)	44 (94)	65 (100)	20 (63)	43 (88)	9 (100)	10 (50)	29 (66)	280 (86)
Chemotherapy cycles									
<6	6 (10)	10 (21)	5 (8)	3 (9)	1 (2)	4 (44)	0	2 (5)	31 (9)
6	19 (31)	25 (53)	18 (28)	29 (91)	43 (88)	3 (33)	11 (55)	14 (32)	162 (50)
>6	36 (59)	12 (26)	42 (65)	0	5 (10)	2 (22)	9 (45)	28 (64)	134 (41)
Consolidation radiotherapy									
Total	26	5	13	14	1	2	1	4	66 (20)
Bulky disease†	13	4	11	11	1	0	1	4	45
Nonbulky site	13	1	2	3	0	2	0	0	21
No. of patients with significant therapy delays or dose reductions	0	14 (30)	6 (9)	Data incomplete	8 (16)	4 (44)	2 (10)	data incomplete	41/259 (16) (68 not known)
I-PET timing after cycle									
2	59	44	55	27	2	7	16	41	251 (77)
3	2	3	10	5	44	2	4	3	73 (22)
4	0	0	0	0	3	0	0	0	3 (1)
I-PET–positive (n)	21 (34)	7 (15)	21 (32)	11 (34)	11 (22)	4 (44)	16 (80)	26 (60)	117 (36)
Outcome by country									
2-y EFS (95% CI)	80% (68–88)	89% (76–95)	80% (67–88)	74% (54–86)	76% (61–86)	78% (36–94)	74% (48–88)	68% (50–81)	79% (74–83)
2-y OS (95% CI)	86% (74–93)	91% (79–97)	88% (77–94)	81% (62–91)	87% (72–94)	78% (36–94)	82% (53–94)	79% (60–90)	86% (81–89)

*Five patients received chemotherapy other than R-CHOP/CHOP (rituximab-CNOP [cyclophosphamide, mitoxantrone, vincristine, prednisone], n = 4; MACOPB [methotrexate, cytarabine, cyclophosphamide, vincristine, prednisolone, bleomycin], n = 1).
†Bulky disease defined as > 5 cm.
Data in parentheses are percentages, except where indicated otherwise.

The prognostic ability of I-PET is estimated from a Cox proportional hazards model (20). Between-country heterogeneity in the prognostic value of I-PET was explored using a 2-stage meta-analysis. Taking the Cox model based on I-PET as the sole covariate for each country, the overall effect of I-PET was estimated between countries by random country effects calculated using generalized Q statistics (21). Secondly, each country was omitted in turn to identify how each contributed to any differences identified.

To investigate whether I-PET adds prognostic discrimination beyond that of established factors, a Cox model was fitted with IPI and age as covariates. Other variables (stage, performance status, extranodal sites, lactate dehydrogenase [LDH], bulky disease > 5 cm, rituximab treatment, and I-PET timing [2 or 3+ cycles]) were chosen by a process of backward elimination, using a P value of 0.1 as the elimination criteria. Finally I-PET and E-PET classification, as positive or negative, were included as independent variables. A multivariable model was developed using prognostic factors chosen from this model to identify risk categories with a significant degree of prognostic separation. Analyses and graphs were produced using Stata 12 (StataCorp).

RESULTS

Major cancer centers in 9 countries (São Paulo, Brazil; Santiago, Chile; Budapest and Debrecen, Hungary; Mumbai, India; Bologna, Italy; Seoul, South Korea; Manila, Philippines; Bangkok, Thailand; and Ankara, Turkey) from 5 United Nations–defined geographic regions participated in the study. Recruitment commenced in 2008 through September 2011. Of the 383 patients recruited, 56 were excluded, based on predefined eligibility criteria. Twenty-two did not meet recruitment criteria, and 34 could not be analyzed because the scans were not submitted for central review, leading to exclusion of all recruited Turkish cases (Fig. 1). Nine cases classified as primary mediastinal B-cell lymphoma were included. Of the 327 eligible cases, 52 were from low-middle income countries (India, Philippines), 170 from upper-middle (Brazil,
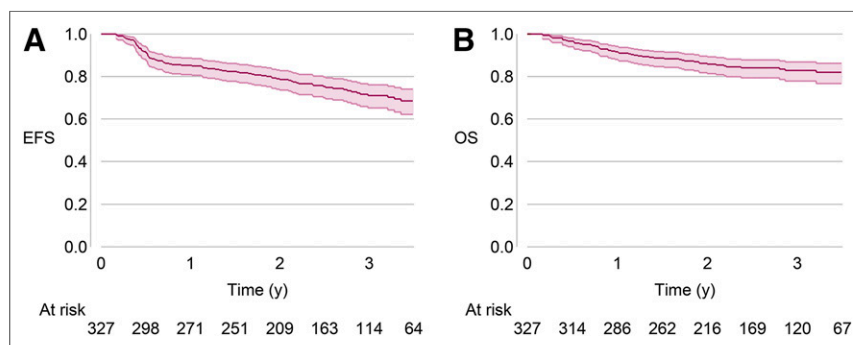

FIGURE 2. Kaplan–Meier plots of EFS (A) and OS (B) for entire eligible cohort. 95% CIs and number of cases at risk are shown; 2-y EFS was 79% (95% CI, 74%–83%) and OS was 86% (95% CI, 81%–89%).
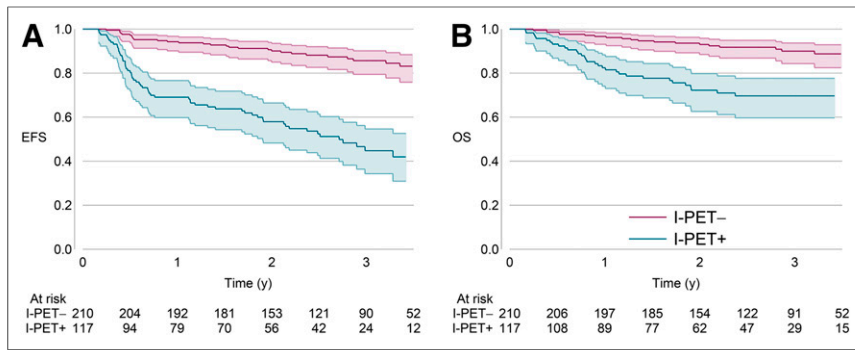
FIGURE 3. Kaplan–Meier plots of EFS (A) and OS (B) for cases stratified by I-PET. 95% CIs and number of cases at risk are shown. Two-year EFS: I-PET–negative, 90% (95% CI, 85%–93%); I-PET–positive, 58% (95% CI, 48%–66%); and HR, 5.31 (3.29–8.56). Two-year OS: I-PET–negative, 93% (95% CI, 88%–96%), I-PET–positive, 72% (95% CI, 63%–80%); and HR, 3.86 (95% CI, 2.12–7.03).

Hungary, Thailand), and 105 from high-income countries (Chile, Italy, South Korea).

Patient Characteristics

Table 1 provides patient characteristics by country: 35% of patients were older than 60 y. The 8 contributing centers demonstrated good compliance with the protocol; 280 (86%) received rituximab (Table 2). The number of chemotherapy cycles varied between countries, with more than 6 cycles given most commonly for advanced or bulky disease; 66 patients were given consolidation radiotherapy after a negative E-PET result and confirmation of CR. I-PET timing was consistent with the protocol, with 77% scanned after 2 chemotherapy cycles, 99% after 2 or 3 cycles, and 1% after 4 cycles.

Outcomes

At a median follow-up of 2 y 11 mo, the 2-y survival for all 327 cases was 79% (95% confidence interval [CI], 74%–83%) for EFS and 86% (95% CI, 81%–89%) for OS. Three-year survival was 71% (95% CI, 65%–76%) for EFS and 83% (95% CI, 78%–87%) for OS (Fig. 2). There were 93 events after I-PET, 51 deaths as the primary event and 42 treatment failures or relapses after remis-

sion. The 47 cases treated without rituximab had a 2-y EFS of 75% (95% CI, 60%–85%) versus 79% (95% CI, 74%–85%) with rituximab.

Stratified by I-PET. The I-PET scan was negative in 210 (64%) and positive in 117 (36%) patients. Two-year EFS as stratified by I-PET was 90% (95% CI, 85%–93%) for I-PET–negative patients and 58% (95% CI, 48%–66%) for I-PET–positive patients, with a hazard ratio (HR) of 5.31 (95% CI, 3.29–8.56). Two-year OS as stratified by I-PET was 93% (95% CI, 88%–96%) for I-PET–negative patients and 72% (95% CI, 63%–80%) for I-PET–positive patients, with an HR of 3.86 (95% CI, 2.12–7.03) (Fig. 3).

Stratified by Sequential I-PET and End-Treatment PET. Fifteen cases did not have a study E-PET due to death (*n* = 12) or early treatment escalation (*n* = 3). Analysis of the 312 cases with both I-PET and E-PET scans showed that most cases (96%) clustered into 3 prognostic groups (Fig. 4). The largest, 192 (62%) cases, had negative I-PET and E-PET demonstrating a rapid response with an excellent 2-y EFS of 97% (95% CI, 92%–98%) and an OS of 97% (95% CI, 93%–99%).

The second group, 58 (19%) cases, had positive I-PET but negative E-PET and were in clinical remission at the end of chemotherapy; this slow-response group had a 2-y EFS of 86% (95% CI, 73%–93%) and OS of 92% (95% CI, 79%–97%). HRs comparing rapid- and slow-response cases show slow responders to have approximately double the risk of an event by 2 y, compared with those with a negative I-PET (HR for EFS, 2.56 [95% CI, 1.08–6.11]; for OS, 1.83 [95% CI, 0.61–5.51]).

The third largest group, 49 (16%) cases, had positive I-PET and E-PET scans, a 2-y EFS of 35% (95% CI, 22%–48%), and had continuing relapses beyond 2 y. In 7 of these cases, the E-PET appeared to be false-positive, with no residual disease if biopsied or continued clinical remission without additional
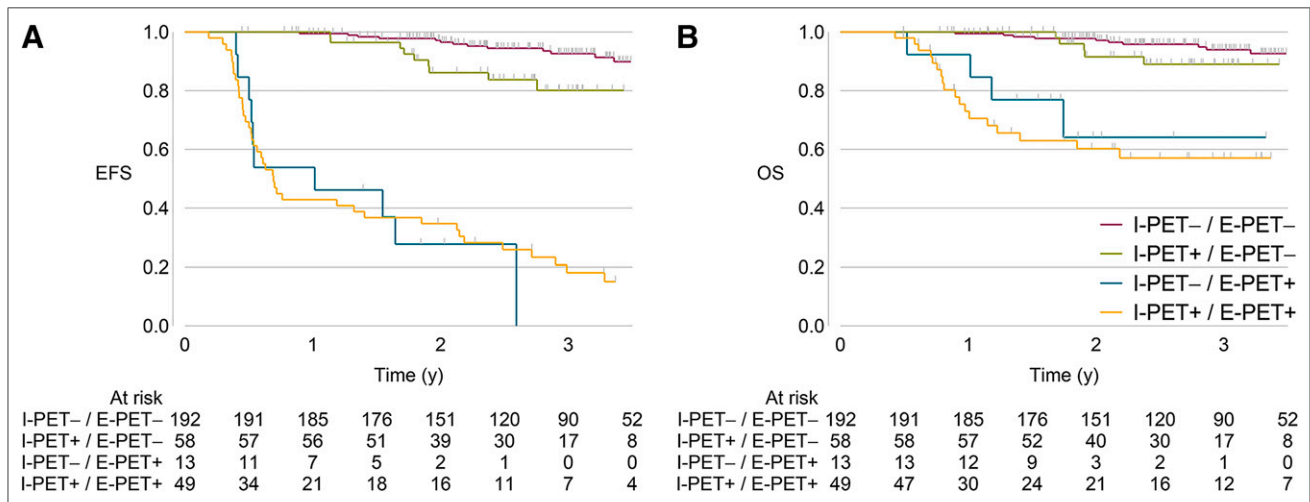

FIGURE 4. Kaplan–Meier plots of EFS (A) and OS (B) for cases stratified by both I-PET and E-PET. Number of cases at risk is shown. I-PET–negative/E-PET–negative: EFS, 97% (95% CI, 92%–98%), OS, 97% (95% CI, 93%–99%); I-PET–positive/E-PET–negative: EFS, 86% (95% CI, 73%–93%), OS, 92% (95% CI, 79%–97%); I-PET–negative/E-PET–positive: EFS, 28% (95% CI, 7%–54%), OS, 64% (95% CI, 28%–86%); and I-PET–positive/E-PET–positive: EFS, 35% (95% CI, 22%–48%), OS, 60% (95% CI, 44%–73%).
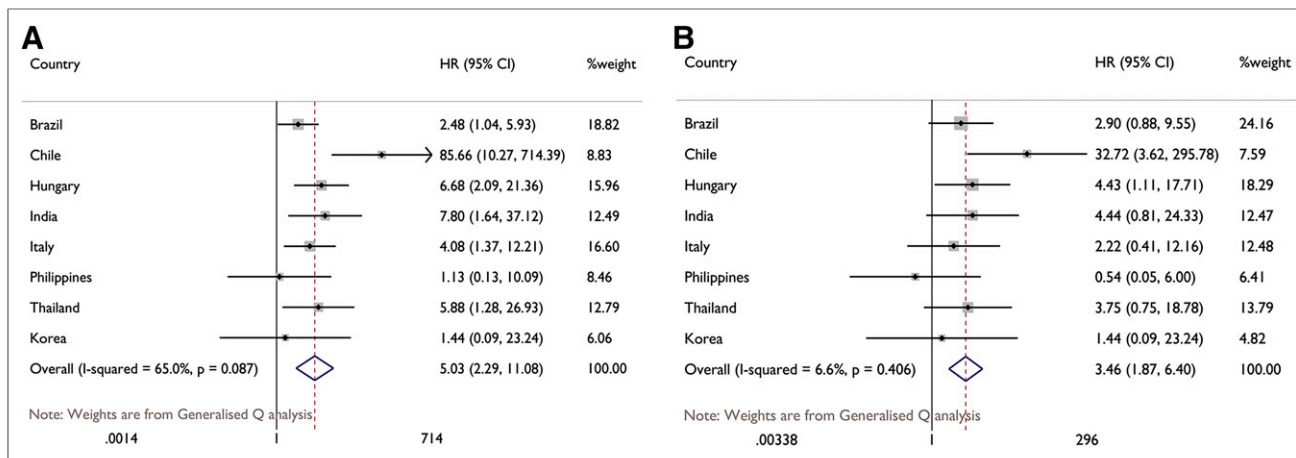
FIGURE 5. Between-country heterogeneity analysis of outcomes stratified by I-PET: EFS (A) and OS (B). Figures show forest plot of EFS/OS HR and 95% CI for I-PET–positive vs. I-PET–negative cases by country, combined meta-analysis HR, and estimate of heterogeneity, I^2. These plots show all countries. Additional analyses explored contribution of each country to any differences identified (data not shown).

chemotherapy; none of these 7 cases had consolidation radiotherapy. An additional 13 patients (4%) were I-PET–negative but E-PET–positive with ^{18}F-FDG uptake at a previous or new site (EFS, 28% [7%–54%]); 11 of these patients (85%) had biopsy-confirmed disease progression.

In the slow-response group, bulky disease was more common, 67% versus 50% in the cohort overall (Table 1). Consolidation radiotherapy (after confirmed CR) was given to 19% of patients with bulky disease at diagnosis, compared with 20% of the cohort overall (Table 2).

Between-Country Heterogeneity

Survivals stratified by I-PET were compared for consistency of direction and magnitude across countries. Initial analysis of I-PET–stratified EFS showed modest between-country heterogeneity ($P = 0.09$; $I^2 = 65\%$). With Chile omitted from the analysis (see the "Materials and Methods" section), there was complete absence of heterogeneity among the other 7 countries ($I^2 = 0\%$). Heterogeneity for OS across all countries was low and similarly not significant ($P = 0.4$; $I^2 = 6.6\%$) (Fig. 5). Chile's survival figures were characterized by the highest 2-y EFS (89%) and the lowest proportion of I-PET–positive patients (15%), all of whom had disease progression or died during treatment. These outcomes are sufficient to explain Chile's noncongruence in the heterogeneity analysis. It is of importance, in the context of this study, that Chile is a high-income country, with healthcare relatively well resourced, and the study patients were predominantly of European origin. We therefore found no relationship between ethnicity, geographic region, or economic status and outcomes stratified by I-PET.

Additional Risk Factors

Multivariate analysis including the entire study population was performed to investigate the relative influence on EFS of the IPI and its components, bulky disease, rituximab treatment, timing of I-PET (after 2 vs. 3 cycles), and I-PET response status. IPI, age, stage, extranodal disease in 2 or more sites, and bulky disease were not significant variables, nor were timing of the I-PET (Fig. 6) or the omission of rituximab in 14% of cases. Only I-PET status, performance status, and LDH reached significance: positive I-PET, HR of 4.32 (2.64–7.10); performance status ≥ 2, 1.79

(1.05–3.07); and abnormal LDH, 1.66 (1.03–2.67). I-PET–negative cases with a performance status of 0–1 and normal LDH ($n = 110$) had an exceptionally good 2-y EFS (98% [92.0%–100%]).

The multivariate analysis was repeated to include E-PET status on the 312 cases with end-treatment scans. In this analysis, E-PET was strongly predictive of outcome, with an HR of 14.3 (7.74–22.45), in contrast to I-PET (HR, 1.16 [0.63–2.16]). Performance status and LDH remained the only predictive clinical variables. However, high performance status and raised LDH together with a positive I-PET did not necessarily predict poor outcome; 39% of I-PET–positive cases had these adverse clinical characteristics and yet were in PET-negative remission at completion of chemotherapy, and 84% of these were alive in first remission at 2 y.

Outcome of I-PET CR-MRU

We examined, as a secondary analysis, the outcome of cases in which the I-PET was reported as CR-MRU ($n = 88$), compared with I-PET–negative with no residual ^{18}F-FDG uptake ($n = 122$) (Fig. 7). I-PET MRU and I-PET–negative cases had almost identical EFS and OS over time, justifying combining both as I-PET–negative in the analyses.
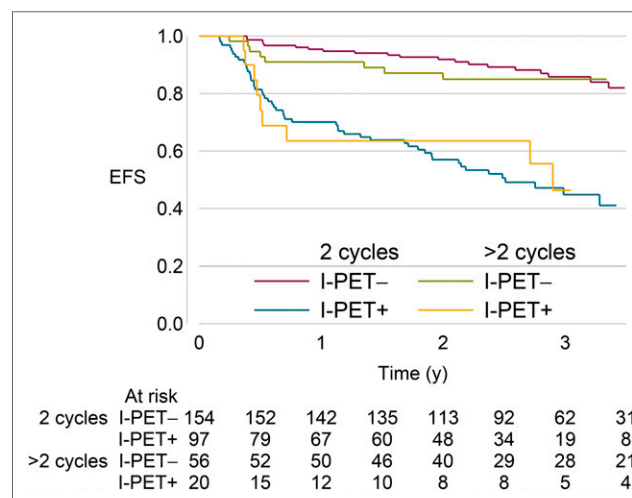

FIGURE 6. I-PET–stratified EFS for cases by timing of I-PET: after 2 ($n = 251$) vs. after 3 ($n = 73$) or 4 ($n = 3$) cycles of chemotherapy.
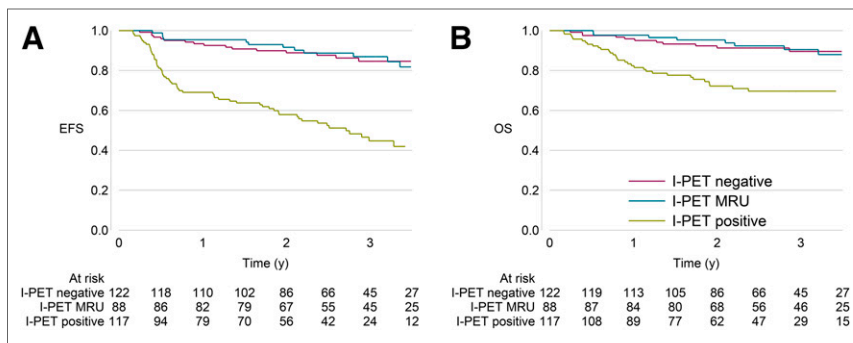
FIGURE 7. Outcome of cases with I-PET scans scored as minimal residual uptake. Kaplan–Meier plots compare EFS (A) and OS (B) of cases, with I-PET scans classified visually as negative, MRU, and positive.

DISCUSSION

This study was designed to define the international role of PET for risk stratification of DLBCL and to address the hypothesis that biological diversity or diversity of healthcare in different geographic regions may influence the kinetics of treatment response as assessed by I-PET.

In this cohort of 327 patients, EFSs were close to those reported from recent large European R-CHOP studies (*1,2*). Similarly, 2-y EFS and OS, stratified by I-PET response after 2–3 cycles chemotherapy, were comparable to recent studies from high-income countries, likewise using visual PET reporting and early I-PET scanning after 2–3 cycles (Table 3) (*22–26*).

Analysis for heterogeneity of OS and EFS stratified by I-PET found little difference between countries. Therefore, any population heterogeneity based on clinical risk factors at diagnosis was normalized to deliver remarkable consistency of outcome prediction across participating countries when stratified by I-PET response. Consequently, our hypothesis that international diversity

might limit the global applicability of western-generated PET data was not supported. Equally, the consistency of I-PET–stratified survivals across the different healthcare environments justifies pooling data from all countries for the study's primary analysis.

This large prospective study demonstrates a highly significant difference between EFS of the I-PET–positive and –negative cohorts. Sequential monitoring, by both I-PET and E-PET, identified 4 risk groups with greater and clinically important separation between cases with good and poor prognosis. Two groups had good and 2 poor EFS at 2 y, 97% and 86%, versus 35% and 28%, respectively. The large cohort size not only provided more precise survival estimates, it also enabled a more informative multivariable analysis of clinical risk factors than previously possible. We identified that a substantial subgroup with complete metabolic response at both I-PET and E-PET, coupled with a marker both of more favorable tumor biology (normal LDH) and of patient fitness (good performance status, though not age), had an excellent EFS of 98%.

More important was the revelation that of 107 I-PET–positive cases, more than half (54%) became PET-negative by the end of chemotherapy, and most of these slow responders had durable remissions. This finding not only explains the inability of a single positive I-PET scan to predict poor outcome, it also indicates that to intensify therapy on this basis would put a significant number of patients at risk of unjustified treatment-related toxicity.

A slow but complete response group has been previously noted in 2 much smaller studies. In one, 15 of 25 (60%) patients with positive I-PET achieved complete response by E-PET (*23*). In another highly selective retrospective database study, 35 of 55

TABLE 3
EFS or PFS, Stratified by I-PET, for Studies of DLBCL Without Risk-Adapted Therapy

| Author | Type of study | I-PET after *n* cycles (cycles %) | No. of subjects | Median follow-up (mo) | Study cohort PFS (y) | I-PET–positive (n) | 2-y PFS/EFS | | 3-y PFS/EFS | |
							I-PET–negative survival	I-PET–positive survival	I-PET–negative survival	I-PET–positive survival
Zinzani et al. (*26*)	Retrospective	3 cycles (or mid-therapy*)	91	50 m		38%	94%[†]	45%[†]		
Cashen et al. (*25*)	Prospective	2–3 cycles (2 in 94%)	50	34 m	74% (2 y)	48%	85% (72%–100%)[‡]	63% (46%–85%)[‡]		
Yoo et al. (*24*)	Retrospective	2–4 cycles (2–3 in 57%)	155	20 m	77% (3 y)	36%			84%	66%
Pregno et al. (*23*)	Prospective	2–4 cycles (2–3 in 76%)	88	26 m	77% (2 y)	28%	85%	72%		
Safar et al. (*22*)	Prospective	2 cycles	112	38 m	84% (3 y)	37%			84% (75%–94%)[‡]	47% (32%–62%)[‡]
IAEA Lymphoma Group	Prospective	2–4 cycles (2–3 in 99%)	327	35 m	79% (2 y)	36%	90 (85%–93%)[‡]	58% (48%–66%)[‡]	86% (79%–90%)[‡]	45% (34%–55%)[‡]

*Treatments included MACOP-B, R-VNCOP-B, R-CHOP21.
[†]2-y EFS by personal communication from Pier Luigi Zinzani (associated member of IAEA Lymphoma Study Group).
[‡]CIs are shown for studies when they were included in cited report.
Only studies using visual reporting criteria are included in this comparison. Studies primarily assessing after 4 cycles of chemotherapy were not included. Outcome, PFS, or EFS, as reported by each study.

(63%) achieved remission after a positive I-PET (*24*). These smaller studies provide reassurance that our findings are not a unique phenomenon. Our large prospective cohort provides more definitive evidence that an early I-PET scan cannot be used to guide early treatment escalation.

In some, slow response may represent less chemosensitive disease, reflected by the increased event rate of the I-PET–positive/E-PET–negative subgroup, compared with the rapid response I-PET–negative/E-PET–negative cases reported here. In others, the likely explanation is that persistent ^{18}F-FDG uptake in the tumor mass is due to inflammatory reaction within necrotic tumor rather than residual viable lymphoma (*27*). A positive I-PET due to inflammatory cells in the absence of tumor has previously been reported in a DLBCL study, which found cases with inflammatory cells only on biopsy had good outcomes (*28*). It has been suggested that this phenomenon is more common after rituximab (*29*).

Since the study commenced, there has been growing interest in standardized uptake value reduction (ΔSUV_{max}) at I-PET as a more sensitive method of separating good- from poor-outcome patients. Recent studies comparing visual with semiquantitative ΔSUV_{max} for predicting outcomes found ΔSUV_{max} to better predict progression-free survival (PFS)/EFS after 2 or 4 cycles than visual assessment, and when ΔSUV_{max} cutoff was optimized to the timing of I-PET, discrimination between those with and without residual lymphoma was further improved (*30–32*). Early response assessment using ΔSUV or volumetric analysis may evolve to a degree to which it could be justifiably used to guide risk-adapted therapy.

In the global healthcare context explored by this study, quantitative techniques may be less practicable because of the demanding conditions required for accurate and reproducible results. However, our study has demonstrated that multinational collaboration between developed and developing countries to test newer and more demanding methodologies is feasible and informative as well as providing mentoring and training for investigators.

Coordinated Research Projects provide limited financial support. This, and geographic logistics, prevented calibration of PET scanners to a common standard for this study, though there was central review of all PET scans. Central pathologic review was similarly not practicable, though diagnostic support was provided by 2 senior lymphoma pathologists. Despite these potential confounding factors, our outcomes are similar to those from the United States and Europe, with the strength that they reflect local practice.

At a time when health priorities in the developing world are shifting to noncommunicable disease, with calls to make cancer cure a global priority, it is important to demonstrate that data that underpin oncology practice in the developed world can be applied internationally (*33*). We found that geographic and population diversity did not influence I-PET–stratified outcomes.

CONCLUSION

A decade of investigation has sought to establish I-PET as a reliable indicator to guide early treatment intensification. This large international cohort delivers the strongest evidence that a positive I-PET result does not differentiate chemoresistant residual tumor from complete response, nor does it provide sound basis for early escalation of therapy in individuals with DLBCL.

APPENDIX
The International Atomic Energy Agency Lymphoma Study Group

Austria–Vienna (IAEA)
Diana Paez*[†], Maurizio Dondi*, Noura El-Haj

Brazil–Sâo Paulo
Juliano Cerci*[†], Claudio Meneghetti[†], Artur Coutinho, Jose Soares, Debora Levy, Segio Bydtowsky, Juliana Pereria, Renata Costa, Sheila Coelho

Chile–Santiago
Francisco Redondo[†], Paulette Conget, Eva Bustamante, Flavia Bruna

France–Paris
Rose Ann Padua*[†]

Hungary–Budapest, Debrecen
Tamás Masszi[†], Botond Timár, Ágota Szepesi, Andrea Sipos, Judith Demeter, Lajos Gergely, Ildikó Garai

India–Mumbai
Reena Nair[†], Sumeet Gujral[†], Venkatesh Rangarajan, Ranjan Basak

Italy–Bologna
Stefano Fanti*[†], Monica Celli[†], Pier Luigi Zinzani

South Korea–Seoul
June-Key Chung[†], Keon Wook Kang, Jae Seon Eo

Philippines–Manila
Charity Gorospe[†], Maejoy Vena Campo, Filipinas Natividad, Mark Piere Dinamay

Thailand–Bangkok
Chirayu Auewarakul[†], Narongrit Sritana

Turkey–Ankara
Hilal Ozdag, Isinsu Kuzu[†], Nilgun Tekin, Ozem Kucuk, Gulseren Aras, Melike Ozbilgin, Muhit Ozcan, Pervin Topcuolglu

U.K.–London
Robert Carr (Chief Scientific Investigator)*[†], Tim Morris[†]

*Study Management Group.
[†]Author.

REFERENCES

1. Pfreundschuh M, Trumper L, Osterborg A, et al. CHOP-like chemotherapy plus rituximab versus CHOP-like chemotherapy alone in young patients with

good prognosis diffuse large B-cell lymphoma: a randomised controlled trial by the MabThera International Trial (MinT) Group. *Lancet Oncol.* 2006;7:379–391.

2. Cunningham D, Hawkes EA, Jack A, et al. Rituximab plus cyclophosphamide, doxorubicin, vincristine and prednisolone in patients with newly diagnosed diffuse large B-cell lymphoma: a phase 3 comparison of intensification with 14-day versus 21-day cycles. *Lancet.* 2013;381:1817–1826.

3. The international Non-Hodgkin's Lymphoma Prognostic Factors Project: a predictive model for aggressive non-Hodgkin's Lymphoma. *N Engl J Med.* 1993; 329:987–994.

4. Hans CP, Weisenburger DD, Greineer TC, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using tissue microarray. *Blood.* 2004;103:275–282.

5. Rosenwald A, Wright G, Chan WC, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large B-cell lymphoma. *N Engl J Med.* 2002;346:1937–1947.

6. Lossos IS, Czerwinski DK, Alizadeh AA, et al. Prediction of survival in diffuse large B-cell lymphoma based on the expression of six genes. *N Engl J Med.* 2004;350:1828–1837.

7. Dave SS. Genomic stratification for the treatment of lymphomas. *Hematology* 2013;2013:331–334.

8. Sehn LH. Paramount prognostic factors that guide therapeutic strategies in diffuse large B-cell lymphoma. *Hematology.* 2012;2012:402–409.

9. van Dongen JJ, Seiu T, Panzer-Grumayer ER, et al. Prognostic value of minimal residual disease in acute lymphoblastic leukaemia in childhood. *Lancet.* 1998;352:1731–1738.

10. Jerusalem G, Beguin Y, Fassotte MF, et al. Persistent tumour ^{18}F-FDG uptake after a few cycles of polychemotherapy is predictive of treatment failure in non-Hodgkin's Lymphoma. *Haematologica.* 2000;85:613–618.

11. Mikhaeel NG, Timothy AR, O'Doherty MJ, et al. 18-FDG-PET as a prognostic indicator in the treatment of aggressive non-Hodgkin Lymphoma: comparison with CT. *Leuk Lymphoma.* 2000;39:543–553.

12. Cheson BD. Role of functional imaging in the management of lymphoma. *J Clin Oncol.* 2011;29:1844–1854.

13. International Atomic Energy Agency. CRP: Coordinated Research Projects. International Atomic Energy Agency website. http://www.iaea.org/monaco/page.php?page=2117. Accessed October 30, 2014.

14. Swerdlow SH, Campo E, Harris NL, et al. *World Health Organization Classification of Tumours of Haematopoietic and Lymphoid Tissues.* Lyon, France: International Agency for Research on Cancer; 2008.

15. Meignan M, Gallamini A, Haioun C. Report on the first international workshop on interim-PET scan in lymphoma. *Leuk Lymphoma.* 2009;50:1257–1260.

16. Barrington SF, Qian W, Somer EJ, et al. Concordance between four European centres of PET reporting criteria designed for use in multicentre trials in Hodgkin lymphoma. *Eur J Nucl Med Mol Imaging.* 2010;37:1824–1833.

17. Horning SJ, Juweid ME, Schoder H, et al. Interim positron emission tomography scans in diffuse large B-cell lymphoma: an independent expert nuclear medicine evaluation of the Eastern Cooperative Oncology Group E3404 study. *Blood.* 2010;115:775–777.

18. Cheson BD, Pfistner B, Juweid M, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol.* 2007;25:579–586.

19. Hrynaszkiewicz I, Norton ML, Vickers AJ, Altman DG. Preparing raw clinical data for publication: guidance for journal editors, authors, and peer reviewers. *BMJ.* 2010;340:c181.

20. Cox DR. Regression models and life tables. *J R Stat Soc (B).* 1972;34:187–220.

21. DerSimonian R, Kacker R. Random-effects models for meta-analysis of clinical trials: an update. *Contemp Clin Trials.* 2007;28:105–114.

22. Safar V, Dupuis J, Itti E, et al. Interim ^{18}F-fluorodeoxyglucose positron emission tomography scan in diffuse large B-cell lymphoma treated with anthracycline-based chemotherapy plus rituximab. *J Clin Oncol.* 2012;30: 184–190.

23. Pregno P, Chiappella A, Bello M, et al. Interim 18-FDG-PET/CT failed to predict outcome in diffuse large B-cell lymphoma patients treated at diagnosis with rituximab-CHOP. *Blood.* 2012;119:2066–2073.

24. Yoo C, Lee DH, Kim JE, et al. Limited role of interim PET/CT in patients with diffuse large B-cell lymphoma treated with R-CHOP. *Ann Hematol.* 2011; 90:797–802.

25. Cashen AF, Dehdashti F, Luo J, et al. ^{18}F-FDG PET/CT for early response assessment in diffuse large B-cell lymphoma: poor predictive value of international harmonisation project interpretation. *J Nucl Med.* 2011;52:386–392.

26. Zinzani PL, Gandolfi L, Broccoli A, et al. Mid-treatment ^{18}F-fluorodeoxyglucose positron emission tomography in aggressive non-Hodgkin lymphoma. *Cancer.* 2011;117:1010–1018.

27. Spaepen K, Stroobants S, Dupont P, et al. (^{18}F)FDG monitoring of tumour response to chemotherapy: does (^{18}F) FDG uptake correlate with the viable tumour cell fraction? *Eur J Nucl Med Mol Imaging.* 2003;30:682–688.

28. Moskowitz CH, Schoder H, Teruya-Feldstein J, et al. Risk-adapted dose-dense immunochemotherapy determined by interim FDG-PET in advance-stage diffuse large B-cell lymphoma. *J Clin Oncol.* 2010;28:1896–1903.

29. Han HS, Escalon MP, Hsiao B, et al. High incidence of false-positive PET scans in patients with aggressive non-Hodgkin's lymphoma treated with rituximab-containing regimens. *Ann Oncol.* 2009;20:309–318.

30. Itti E, Meignan M, Berriolo-Rieddinger A, et al. An international confirmatory study of the prognostic value of early PET/CT in diffuse large B-cell lymphoma: comparison between Deauville criteria and ΔSUVmax. *Eur J Nucl Med Mol Imaging.* 2013;40:1312–1320.

31. Nols N, Mounier N, Bouazza S, et al. Quantitative and qualitative analysis of metabolic response at interim positron emission tomography scan combined with International Prognostic Index is highly predictive of outcome in diffuse large B-cell lymphoma. *Leuk Lymphoma.* 2014;55:773–780.

32. Casasnovas R-O, Meignan M, Berriolo-Riedinger A, et al. SUVmax reduction improves early prognosis value of interim PET scans in diffuse large B-cell lymphoma. *Blood.* 2011;118:37–43.

33. Cavalli F. An appeal to world leaders: stop cancer now. *Lancet.* 2013;381:425–426.