HemaSphere

EHA EUROPEAN HEMATOLOGY ASSOCIATION

Article
Open Access

Interim Positron Emission Tomography During Frontline Chemoimmunotherapy for Follicular Lymphoma

Reid W. Merryman[1], Laure Michaud[2], Robert Redd[3], Patrizia Mondello[4], Hyesun Park[5], Gabriela Spilberg[5], Matthew Robertson[5], Eleanor Taranto[6], Gulrayz Ahmed[7], Matthew Chase[8], Erin Jeter[1], Inhye E. Ahn[1], Jennifer R. Brown[1], Jennifer Crombie[1], Matthew S. Davids[1], David C. Fisher[1], Eric Jacobsen[1], Caron A. Jacobson[1], Austin I. Kim[1], Ann S. LaCasce[1], Samuel Y. Ng[1], Oreofe O. Odejide[1], Erin M. Parry[1], Gilles Salles[4], Andrew D. Zelenetz[4], Philippe Armand[1], Heiko Schöder[2], Heather Jacene[5]

Correspondence: Reid W. Merryman (Reid_merryman@dfci.harvard.edu).

ABSTRACT

While most patients with follicular lymphoma (FL) have excellent outcomes with frontline chemoimmunotherapy (CIT), a subset of patients will experience early progression, which is associated with poor subsequent outcomes. Novel biomarkers are needed to identify high-risk patients earlier. We hypothesized that interim positron emission tomography (PET) would predict progression-free survival (PFS) in this population. We retrospectively identified 128 patients with grade 1–3A FL who had an interim PET after 2–4 cycles of frontline CIT at 2 academic centers. PET scans were analyzed using Deauville score (DS) and change in maximum standardized uptake value (ΔSUVmax). Interim PET DS was a significant predictor of PFS ($P < 0.003$). Patients with a DS of 3 had outcomes similar to those of patients with a DS of 4, so were categorized as PET-positive for additional analyses. Interim PET remained a strong predictor of PFS (DS 3-5, hazard ratio [HR] 2.4, $P = 0.006$) in a multivariable analysis and was also an early predictor of both a positive end-of-treatment PET ($P < 0.001$) and progression of disease within 24 months (POD24) ($P = 0.006$). An optimal ΔSUVmax cutoff of 75% was selected using the bootstrap method. ΔSUVmax <75% was also a significant predictor of PFS on univariable and multivariable analyses (HR 2.8, $P < 0.003$). In a separate cohort of 50 patients with high-grade FL, interim PET interpreted using either DS ($P < 0.001$) or ΔSUVmax75% ($P = 0.034$) was also a significant predictor of inferior PFS. In conclusion, interim PET is an independent predictor of PFS and may be useful as a tool for response-adapted treatment strategies in FL.

INTRODUCTION

While most patients with follicular lymphoma (FL) have excellent outcomes with frontline chemoimmunotherapy (CIT), a subset of patients will experience early progression, which is associated with poor subsequent outcomes.[1,2] Clinical (FLIPI, FLIPI2, PRIMA-PI)[3–5] and clinicobiologic (m7FLIPI, BioFLIPI)[6,7] prognostic tools used prior to treatment initiation do not currently allow prospective identification of patients likely to have early progression, at least for patients receiving bendamustine-based CIT.[8,9] Other prognostic tools (PRIMA23, T effector cell signature) suggest that distinct biologic predictors may be needed for individual chemotherapy backbones.[10,11] Earlier identification of patients with high-risk FL could allow for earlier change in therapy and for investigation of novel treatments with the goal of forestalling early progression.

Among patients with FL, fluoro-[^{18}F]-deoxy-2-D-glucose (FDG) positron emission tomography (PET) performed at the end of CIT induction can identify patients at significantly higher risk of relapse[12]; however, such patients have already been exposed to the full course of CIT. If interim FDG-PET were similarly prognostic, it would permit earlier identification of patients likely to have a poor outcome and potentially an earlier change in therapy. To date, the prognostic impact of interim PET in FL has not been clearly established, unlike in

[1]Division of Hematologic Malignancies, Dana-Farber Cancer Institute, Boston, MA, USA
[2]Molecular Imaging and Therapy Service, Memorial Sloan Kettering Cancer Center, New York, NY, USA
[3]Department of Data Sciences, Dana-Farber Cancer Institute, Boston, MA, USA
[4]Lymphoma Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY and Weill Cornell Medicine, New York, NY, USA
[5]Department of Imaging, Dana-Farber Cancer Institute/Department of Radiology, Brigham and Women's Hospital, Boston, MA, USA
[6]Division of Hematology and Oncology, Hospital of the University of Pennsylvania
[7]Medical College of Wisconsin, Milwaukee, WI, USA
[8]Divisions of Hematology and Medical Oncology, Beth Israel Deaconess Medical Center, Boston, MA, USA
RWM and LM have contributed equally to this work.
Supplemental digital content is available for this article.

HemaSphere (2023) 7:2(e826).
http://dx.doi.org/10.1097/HS9.0000000000000826.
Received: August 10, 2022 / Accepted: December 5, 2022

other lymphomas such as classical Hodgkin lymphoma (cHL) or diffuse large B-cell lymphoma (DLBCL).[13–17] In FL, studies examining the prognostic value of interim PET scans have come to differing conclusions with some suggesting that a positive interim PET is associated with inferior PFS,[18,19] and others concluding that interim PET scans are not prognostic.[20–22] Nearly all these studies used older PET response criteria (eg, International Harmonization Project [IHP] criteria) and all were limited by small sample size. The Deauville scoring (DS) system is the current standard for PET response assessment in FL and quantitative criteria, such as ΔSUVmax, have demonstrated the ability to identify the highest risk patients in other lymphoma subtypes. We hypothesized that an interim PET scan (interpreted using DS or ΔSUVmax) could identify a subset of patients with FL with a high risk of relapse following frontline CIT.

METHODS

Patients

Patients were retrospectively identified at Dana-Farber Cancer Institute (DFCI) and Memorial Sloan-Kettering Cancer Center (MSKCC). Eligible patients met the following criteria: diagnosis of grade 1–3B FL (with a primary analysis restricted to patients with FL 1–3A and a separate analysis for patients with high-grade FL), initiation of first-line CIT between January 1, 2005, and March 1, 2019, performance of an interim PET scan after 2 to 4 cycles of CIT, and PET imaging available for review. Prior radiation (given with either palliative or curative intent) was permitted. Eligible patients were identified using pharmacy and PET databases at participating centers. Baseline characteristics, treatment information, and cancer-related outcomes were collected by review of electronic medical records. The study was approved by Institutional Review Boards at both centers and the requirement for informed consent was waived for this retrospective study. All research was conducted in accordance with the Declaration of Helsinki.

Analysis of PET

Interim PET scans and (when available) baseline and end-of-treatment (EOT) PET scans were reviewed by expert nuclear medicine radiologists (HJ, MR, HS, LM, HP) blinded to patient outcomes. Interim and EOT PET scan were assigned a DS of 1–5 according to the definitions in the Lugano classification.[23] Visual assessment with SUVmax of tumor versus blood pool or liver was used to aid the analysis in borderline cases. SUVmax of the most FDG-avid FL lesion was recorded on each PET scan. For PET scans with no areas of increased FDG uptake (eg, DS of 1), SUVmax was recorded as 1. ΔSUVmax was calculated by dividing the difference in SUVmax between the baseline and interim PET scans by the SUVmax of the baseline PET scan. To examine the inter-reader variability of DS assignments, PET scans of patients treated at DFCI were reviewed and DS assigned independently by 3 nuclear medicine radiologists (HJ, HP, MR). Cases with ≥1 discordant DS were re-reviewed in a joint reading session and a consensus DS assigned. In all cases, consensus was reached, and the consensus score was used for all analyses.

Statistics

PFS and overall survival (OS) were estimated using the Kaplan–Meier (KM) method with Greenwood's formula for variance estimation, and differences in survival between groups were assessed using the log-rank test. PFS was defined as the time from initiation of CIT to death from any cause, relapse, or progression, with patients censored at the last time seen alive and progression-free. OS was defined as the time from initiation of CIT to death from any cause, with patients censored at the last time seen alive. Median follow-up time was estimated using the reverse KM method. Descriptive statistics were used to summarize variables of interest, and association with binary outcomes were assessed with Wilcoxon rank-sum, Fisher exact, or Cochran-Armitage tests for continuous, nominal, or ordinal variables, respectively. Uni- and multivariable Cox regressions were used to evaluate associations between prognostic factors and PFS or OS. The final Cox model was selected using a penalized maximum likelihood model (LASSO), and a k-fold cross-validation was performed to select a subset of the predictive variables. Finally, a stepwise forward/backward model selection by Akaike information criterion was used to determine the predictive variables for inclusion in the model. Hazard ratios (HRs), 95% confidence intervals (CIs), and Wald *P* values were reported for covariates. Recursive partitioning was performed to establish a cut point for ΔSUVmax to distinguish patients by superior and inferior PFS outcomes. This was repeated for 10,000 iterations using bootstrap resampling, and the median of the 10,000 cut points was assessed in the Cox regressions. Baseline, interim, and EOT PET results were each assessed by three independent reviewers, and the inter-rater agreement was evaluated for each time point using the Fleiss kappa statistic. All analyses were performed using R v4.0.2 with packages *survival* (v3.2-11) for time-to-event analyses and *irr* (v0.84.1) for inter-rater agreement statistics.

RESULTS

Baseline characteristics

Baseline characteristics of the 128 patients with grade 1–3A FL are summarized in Table 1. The median age at diagnosis was 55 years (range 27–83). One-hundred patients (78%) had grade 1–2 FL and 28 (22%) had grade 3A. Most patients (84%) had advanced stage disease and 14% had B symptoms at diagnosis. Low, intermediate, and high FLIPI scores were observed for 27%, 41%, and 32% of patients, respectively. Most patients were treated with either RCHOP (rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone) (50%) or BR (bendamustine, rituximab) (44%). Five patients (4%) received radiation before systemic therapy.

One-hundred twenty one patients (95%) patients received 6 or more cycles of CIT, while 7 patients (5%) received only 4 cycles due to: cytopenias (n = 3), infectious complications (n = 1), physician decision (n = 1), progression (which occurred after the interim PET but before completion of 6 cycles of CIT), and patient decision (n = 1). An interim PET scan was performed after 2 cycles for 16 patients (12%), 3 cycles for 106 patients (83%), and 4 cycles for 6 patients (5%). In all cases, patients received at least 1 additional cycle of CIT after their interim PET scan, confirming that treatment was not changed based on the interim PET scan result. Following CIT, 5 patients (4%) received consolidative radiation therapy (RT) and 40 (31%) received maintenance rituximab (median: 10 doses, range 1–24).

Compared to patients in the DFCI cohort, those in the MSKCC cohort were more likely to receive RCHOP ($P < 0.001$), have a high FLIPI score ($P < 0.037$), have bulkier disease ($P = 0.021$), and have a higher baseline SUVmax ($P = 0.002$). Performance of an interim PET after 2 cycles was more common in the MSKCC cohort (26% vs 7%), while performance of an interim PET after 3 cycles was more common in the DFCI cohort (89% vs 68%).

PET analysis (DS and ΔSUVmax)

DSs for interim and EOT PET scans are summarized in Table 2. For interim PET scans, a DS of 1, 2, 3, 4, and 5 were assigned to 13 patients (10%), 77 patients (60%), 16 patients (12%), 19 patients (15%), and 3 patients (2%), respectively. Patients in the DFCI cohort were more likely to have a DS 1–2 on interim PET compared to the MSKCC cohort (77% vs 52%, $P = 0.006$). A DS of 1–2 was more common when an interim

Table 1

Baseline Characteristics

	Total	Center		
		DFCI	MSKCC	
	n = 128 (%)	n = 90 (70)	n = 38 (30)	*P* value
Age at diagnosis				
Median (range)	55 (27–83)	55 (27–82)	56 (35–83)	*0.52*[a]
Sex				
Female	62 (48)	40 (44)	22 (58)	*0.18*[b]
Grade				
1–2	100 (78)	73 (81)	27 (71)	*0.24*[b]
3a	28 (22)	17 (19)	11 (29)	
Stage				
1	7 (5)	5 (6)	2 (5)	*0.35*[c]
2	14 (11)	11 (12)	3 (8)	
3	36 (28)	27 (30)	9 (24)	
4	71 (55)	47 (52)	24 (63)	
B symptoms				
Yes	18 (14)	14 (16)	4 (11)	*0.58*[b]
FLIPI score				
Low (0–1)	35 (27)	29 (32)	6 (16)	*0.037*[b]
Intermediate (2)	52 (41)	38 (42)	14 (37)	
High (3–5)	41 (32)	23 (26)	18 (47)	
Maximum disease bulk (cm)				
Median (range)	7.90 (2.04–21.04)	7.30 (2.20–19.10)	9.41 (2.04–21.04)	*0.021*[a]
Baseline SUVmax				
Median (range)	12.80 (1.30–51.35)	11.50 (1.30–27.40)	15.92 (6.23–51.35)	*0.002*[a]
Radiation treatment before systemic therapy				
Yes	5 (4)	5 (6)	–	*0.32*[b]
No	123 (96)	85 (94)	38 (100)	
Immunochemotherapy regimen				
RCHOP	64 (50)	37 (41)	27 (71)	*< 0.001*[b]
BR	56 (44)	49 (54)	7 (18)	
RCVP	3 (2)	1 (1)	2(5)	
Other[d]	5 (4)	3 (3)	2 (5)	
Interim PET after				
2	16 (12)	6 (7)	10 (26)	*0.017*[c]
3	106 (83)	80 (89)	26 (68)	
4	6 (5)	4 (4)	2 (5)	
Consolidative radiation therapy after immunochemotherapy				
Yes	5 (4)	2 (2)	3 (8)	*0.15*[b]
Received rituximab maintenance				
Yes	40 (31)	26 (29)	14 (37)	*0.41*[b]
Doses of rituximab maintenance				
Median (range)	10 (1–24)	11 (1–13)	9 (1–24)	*0.54*[b]

[a]Wilcoxon rank-sum test.
[b]Fisher exact test.
[c]Cochran-Armitage test.
[d]Other regimens included: BO (n = 2), RCHOP->RCDOP (n = 1), RCHP (n = 1), and FCR (n = 1).
BR = bendamustine, rituximab; DFCI = Dana-Farber Cancer Institute; MSKCC = Memorial Sloan-Kettering Cancer Center; PET = positron emission tomography; RCHOP = rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone; SUVmax = maximum standardized uptake value.

PET was performed after more cycles of CIT (44% after 2 cycles, 74% after 3 cycles, and 83% after 4 cycles).

An EOT PET scan was available for review for 112 patients (88%). All 16 patients without an available EOT PET had a DS of 1–3 on interim PET and 13 (81%) underwent repeat imaging after completing CIT using another imaging modality (CT, n = 12; MRI, n = 1). All 13 patients were in a remission on their EOT assessment (12 CR, 1 PR). Among patients with an available EOT PET, DS of 1, 2, 3, 4, and 5 were assigned to 23 patients (21%), 67 patients (60%), 8 patients (7%), 6 patients (5%), and 8 patients (7%), respectively.

To assess inter-reader variability of DS score assignments, PET scans of patients in the DFCI cohort were independently assessed by 3 readers. The probability that 2 random readers agreed on the five-point DS of a random PET scan was higher for baseline PET scans (84%) (kappa = 0.60) compared to interim (53%) (kappa 0.30) or EOT PET scans (52%) (kappa 0.33). Using a 3-group categorization (D1–2 vs D3–4 vs D5), concordance was 84% for baseline PET (kappa 0.60), 65% for interim PET (kappa 0.40), and 79% for EOT PET (kappa 0.48).

ΔSUVmax could be calculated for 107 patients who had an available baseline PET for review. The median SUVmax was 12.8 (range 1.3–51.4) at baseline and 2.0 (range 0.5–19.5) on interim PET. The median ΔSUVmax of 82.1% for the entire cohort and was similar for patients treated at DFCI (81.2%) and MSKCC (84.4%). ΔSUVmax increased when an iPET was obtained after more cycles of CIT (median 78.4% after 2 cycles, 83.2% after 3 cycles, 90.9% after 4 cycles).

Table 2

Deauville Score for Interim and End-of-treatment PET Scans

Interim PET DS	DFCI	MSKCC	Combined
1	3 (3%)	10 (26%)	13 (10%)
2	67 (74%)	10 (26%)	77 (60%)
3	7 (8%)	9 (24%)	16 (12%)
4	12 (13%)	7 (18%)	19 (15%)
5	1 (1%)	2 (5%)	3 (2%)
EOT PET DS			
1	12 (13%)	11 (29%)	23 (18%)
2	63 (70%)	4 (11%)	67 (52%)
3	4 (4%)	4 (11%)	8 (6%)
4	2 (2%)	4 (11%)	6 (5%)
5	4 (4%)	4 (11%)	8 (6%)
Unavailable for review	5 (6%)	11 (29%)	16 (12%)

Patients in the DFCI cohort were more likely to have a DS 1-2 on interim PET compared to the MSKCC cohort (77% vs 52%, $P = 0.006$) and on EOT PET (83% vs 40%, $P < 0.001$).
DFCI = Dana-Farber Cancer Institute; DS = Deauville score; EOT = end-of-treatment; MSKCC = Memorial Sloan-Kettering Cancer Center; PET = positron emission tomography; SUVmax = maximum standardized uptake value.

Prognostic value of interim and EOT PET scans

With a median follow-up of 59 months (range 5–221), the 5-year PFS and OS were 60% (95% CI 52%-71%) and 93% (95% CI 88%-98%), respectively. Forty-four patients relapsed and 41 underwent a confirmatory biopsy which revealed FL for 33 patients, DLBCL for 7 patients, and cHL for 1 patient. Twenty-five had progression of disease within 24 months of CIT (POD24), whereas 19 patients relapsed more than 24 months after CIT.

Interim PET DS was a significant predictor of PFS ($P = 0.003$ for 5 categories) (Figure 1A). PFS curves were overlapping for patients with an interim PET DS of 1 and 2, and also for patients with an interim PET DS of 3 and 4, so patients with a DS of 3 were categorized as PET-positive for subsequent analyses (Figure 1B). Compared to patients with an interim PET DS of 1–2, patients with a DS of 3–4 (HR 2.0, $P = 0.026$) or a DS of 5 (HR 9.5, $P = 0.003$) had inferior PFS. A pattern of inferior PFS was seen for patients with a positive interim PET (interpreted using DS) across key patient subgroups (i.e. treatment center, baseline SUVmax, chemotherapy regimen), but the association did not reach significance across all patient subgroups (Suppl. Figure S1). Interim PET (interpreted using DS) was not associated with OS ($P = 0.60$) (Figure 2A).

ΔSUVmax (calculated between baseline and interim PET scans) was also a significant predictor of PFS. Patients with a ΔSUVmax <82% (the median value in this cohort) had 5-year PFS of 47% (95% CI 33%-68%) compared to 75% (95% CI 63%-88%) for patients with a ΔSUVmax ≥ 82% (HR 2.2, $P = 0.017$). A ΔSUVmax cutoff of 75% was selected by a recursive partitioning univariable PFS model with bootstrap resampling. Patients with a ΔSUVmax <75% had 5-year PFS of 38% (95% CI 22%-66%) compared to 71% (95% CI 60%-83%) for patients with a ΔSUVmax ≥ 75% (HR 3.2, $P < 0.001$) (Figure 1C). Neither ΔSUVmax75% ($P = 0.55$) (Figure 2B) nor ΔSUVmax82% ($P = 0.90$) showed a significant association with OS.

EOT PET DS was a significant predictor of PFS ($P < 0.001$ for 5 categories). The 8 patients with a DS of 3 on EOT PET had favorable outcomes (Suppl. Figure S2), so were grouped with DS 1–2 patients (similar to previously published EOT PET analyses).[17] Compared to patients with an EOT PET DS of 1–3, those with a DS of 4–5 had inferior PFS (HR 2.6, $P = 0.002$). A positive EOT PET was observed more frequently among patients with an interim PET DS of 3–4 (10/30 patients; 33%) or a DS of 5 (2/3 patients, 67%) compared to those with an interim PET DS of 1–2 (2/79 patients; 3%) ($P < 0.001$). Both interim PET and EOT PET were significant predictors of POD24. Among patients with sufficient follow-up to determine POD24 status, rates of POD24 were 15% (10/66), 36% (10/28), and 100% (3/3) for patients with an interim PET DS of 1–2, 3–4, and 5, respectively ($P = 0.006$). POD24 was seen in 62% (9/14) of patients with a positive (DS 4–5) EOT PET versus 16% (13/82) with a negative (DS 1–3) EOT PET ($P < 0.001$).

Multivariable analyses

In a multivariable analysis including key clinical variables (FLIPI score, CIT regimen, maintenance rituximab, bulk, age, sex, grade, and baseline SUVmax), an interim PET DS of 3–5 (HR 2.4, $P = 0.006$) remained a significant predictor of PFS. In addition, a high baseline FLIPI score was associated with a trend toward inferior PFS (HR 1.7, $P = 0.076$) and maintenance rituximab was associated with a trend toward improved PFS (HR 0.5, $P = 0.070$) (Table 3). ΔSUV75% was also a significant predictor of PFS (HR 2.8, $P = 0.003$) in a similar multivariable analysis (Suppl. Table S1). When EOT PET was included as a variable in the multivariable analyses, EOT PET was a significant predictor, but interim PET (assessed using either DS or ΔSUV) was not, suggesting that interim PET functions primarily as an earlier (but not independent) indicator of chemotherapy resistance. In addition, we identified an interaction between interim PET status and maintenance rituximab. Maintenance rituximab was associated with a significant PFS benefit among interim PET-positive (DS 3–5) patients (HR 0.34, $P = 0.031$), but not interim PET-negative (DS 1–2) patients (HR 0.86, $P = 0.74$). To account for patients with progression on their EOT PET (who would be unlikely to receive maintenance rituximab), we repeated the analysis among patients with a positive interim PET and a negative EOT PET (n=21) and found a consistent trend toward improved PFS among patients receiving maintenance rituximab (HR 0.22, $P = 0.061$) (Suppl. Figure S3).

Interim PET assessed using either DS ($P = 0.37$) or ΔSUVmax75% ($P = 0.73$) was not a significant predictor of OS on multivariable analyses (Suppl. Tables S2 and S3).

High-grade FL cohort

Clinical data and PET images were also analyzed for 50 patients with high-grade FL, but these patients were evaluated separately due to key treatment differences in this patient population. Baseline characteristics are summarized in Suppl. Table S4. The median age at diagnosis was 60 years (range 26–86). Thirty-three patients (66%) had grade 3B, 12 patients (24%) had grade 3 NOS, and grade was unknown in 5 patients (10%). Most patients (78%) had advanced stage disease. Low, intermediate, and high FLIPI scores were observed for 26%, 36%, and 38% of patients, respectively. Nearly, all patients were treated with RCHOP (94%). Deauville scores for interim and EOT PET scans are summarized in Suppl. Table S5.

With a median follow-up of 77 months (range 9–184), the 5-year PFS and OS for this cohort were 66% (95% CI 53%-83%) and 83% (95% CI 71%-96%), respectively. Ten patients relapsed, including 6 with POD24. In this cohort of patients, interim PET was also a significant predictor of PFS when interpreted using either DS or deltaSUV75%. Compared to patients with an interim PET DS of 1–2, patients with a DS of 3–5 (HR 7.5, $P < 0.001$) had inferior PFS. Similarly, patients with a ΔSUVmax <75% had inferior PFS compared to those with a ΔSUVmax ≥ 75% (HR 3.2, $P = 0.034$) (Figure 3). Interim PET scans interpreted using DS ($P < 0.001$) was associated with inferior OS, while there was a trend toward inferior OS for patients with a positive interim PET interpreted ΔSUVmax75% ($P = 0.11$).

DISCUSSION

To our knowledge, this multicenter study is the largest analysis to date of the prognostic value of interim PET scans in the

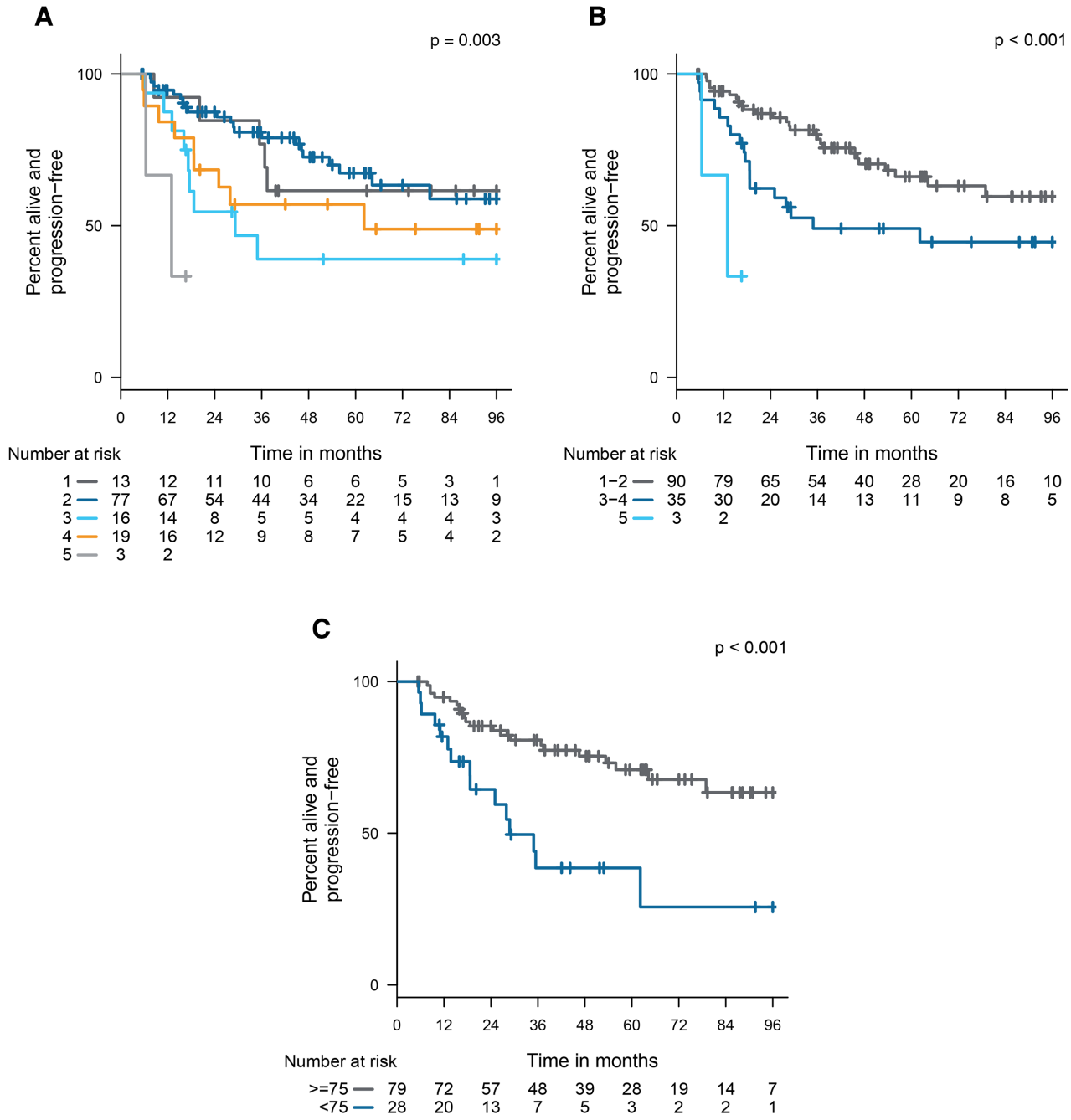

Figure 1. PFS according to interim PET, assessed using (A) DS 5 categories, (B) DS 1–2 vs 3–4 vs 5, (C) ΔSUVmax 75% among patients with grade 1–3A follicular lymphoma. DS = Deauville score; PET = positron emission tomography; PFS = progression-free survival; SUVmax = maximum standardized uptake value.

frontline therapy of FL and suggests that an interim PET scan is a useful predictor of early treatment failure. Indeed, in this cohort, a DS of 3–5 on an interim PET scan (observed in ~30% of grade 1–3A FL patients) could identify patients who were more likely to have residual FDG-avid disease on an EOT PET scan and early treatment failure. Only 15% of patients with a DS of 1–2 on an interim PET scan had POD24 compared to 36% for patients with a DS of 3–4 and 100% for patients with a DS of 5. These results suggest that an interim PET scan could be studied as a tool to develop response-adapted treatment strategies in FL, as has been done in cHL.[14–16] When similar DS categories are merged, our analysis of inter-reader variability demonstrates moderate reproducibility among nuclear medicine radiologists, which is similar to prior studies assessing interim PET scans in other lymphoma subtypes, including cHL[24,25] and DLBCL.[26] Slightly more inter-reader variability was seen in interpretation of interim PET scans (kappa 0.40) in our study compared to EOT PET scans in our study (kappa 0.48) or in the GALLIUM EOT PET analysis (kappa 0.49), suggesting that the benefits of an earlier response assessment may need to be weighed against slightly lower reproducibility.

Immune-based therapies (eg, CAR T cell therapy, bispecific antibodies, etc) would be an attractive alternative for interim PET-positive patients who are likely to do poorly with chemotherapy-based treatments. Outside of a clinical trial, our results suggests that it may be particularly advantageous to use maintenance rituximab among patients with a positive interim PET. While acknowledging the potential for selection bias, we

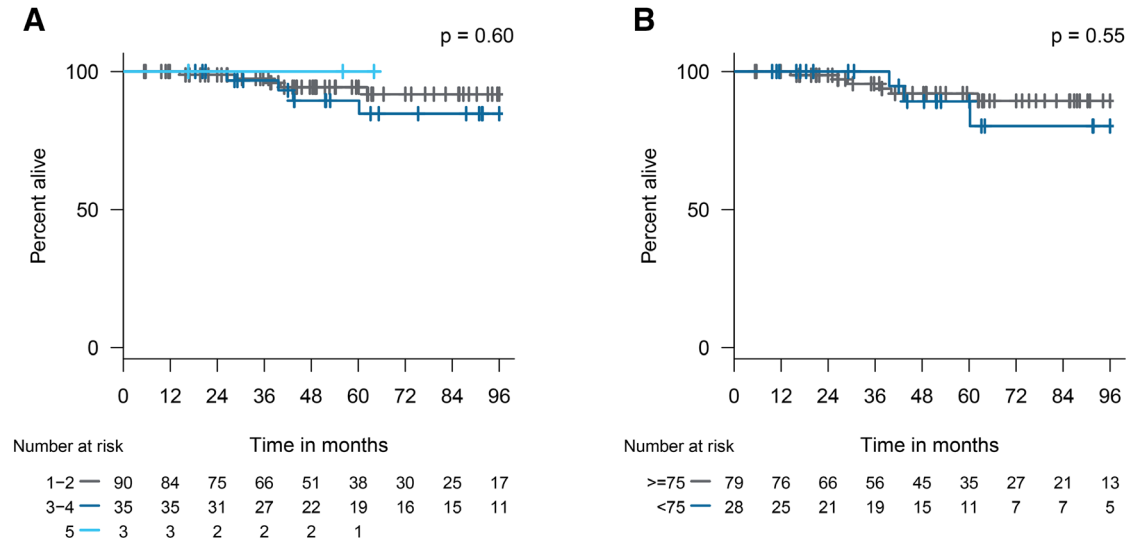

Figure 2. OS according to interim PET scan, assessed using (A) DS 1–2 vs 3–4 vs 5, (B) ΔSUVmax 75% among patients with grade 1–3A follicular lymphoma. OS = overall survival; PET = positron emission tomography; SUVmax = maximum standardized uptake value.

Table 3

Multivariable Analyses for PFS

	Univariable HR	Univariable *P* value	Multivariable HR	*P* value
Interim PET DS				
1–2				
3–5	**2.1**	**0.010**	**2.4**	**0.0061**
FLIPI high	**2.0**	**0.020**	1.7	0.076
CIT regimen				
RCHOP				
BR	1.3	0.35		
Other	1.4	0.54		
Maintenance rituximab	0.7	0.22	0.5	0.070
Bulk (>8 cm)	1.1	0.72		
Age (10-yr inc)	1.1	0.43		
Male	1.6	0.10		
Baseline SUVmax				
SUVmax ≤ 13	0.6	0.19		
SUVmax >13				

BR = bendamustine, rituximab; CIT = chemoimmunotherapy; DS = Deauville score; HR = hazard ratio; PET = positron emission tomography; PFS = progression-free survival; RCHOP = rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone; SUVmax = maximum standardized uptake value.

observed a significant improvement in PFS for interim PET-positive patients who received maintenance rituximab (such that their outcomes were similar to those of patients with a negative interim PET). These results are in contrast to the prospective, response-adapted FOLL12 trial which found a benefit for rituximab maintenance even among low-risk patients who achieved a CR on an EOT PET and had no minimal residual disease after completing induction.[27]

Interestingly, our results suggest that the interpretation of interim PET scans in FL may be different compared to other lymphoma subtypes and even possibly different compared to that of EOT scans in FL. Large studies examining the prognostic value of EOT PET scans using DS in FL defined a positive result as a DS of 4–5 and reported significantly worse outcomes for PET-positive patients; however, outcomes of patients with a DS of 3 on EOT PET are not reported separately.[12,28] In our study, patients who had a DS of 3 on an interim PET scan had outcomes that were similar to patients with a DS of 4 and worse than those with a DS of 1–2. This finding, which is based on a relatively small number of DS 3 interim PET scans (n = 16), would benefit from validation in subsequent studies. This apparent difference could reflect the lower starting SUVmax of FL lesions, which may impart a worse prognosis for residual FDG avidity above that of the mediastinum.

While DS is the recommended scale for PET assessment in lymphoma, it is possible that more objective methods could improve reproducibility. To our knowledge, this is the first study to assess ΔSUVmax for the evaluation of interim PET in FL. Even though FL is characterized by lower baseline SUVs compared to DLBCL, ΔSUVmax was also a significant predictor of PFS in our study. Using a bootstrap method, we identified 75% as an optimal ΔSUVmax cutoff in our cohort. This cutoff is more stringent than the most common threshold (66%) used among DLBCL patients receiving frontline CIT,[17,29,30] and together with the DS analysis, suggests that "deeper" metabolic remissions may be necessary for FL patients to achieve the most favorable outcomes. We found ΔSUVmax could identify high-risk patients even among those with only moderate FDG uptake on baseline PET scans; however, this assessment tool has a limited value for the very small number of FL patients who have an SUVmax of less than 4 on baseline PET (n = 2 in this study), in whom it was not possible to achieve a ΔSUVmax >75% given the minimum SUV of 1 used in our study. The optimal threshold of ΔSUVmax 75% should also be validated in future studies.

Our study has several important limitations. While the study population reflects the clinical heterogeneity of FL, it may not be representative of the general population of patients with FL. Interim PET scans were not performed universally at the two participating centers, but instead were acquired based on insurance coverage and physician preference. It is possible that patients with high-risk features either at baseline or during treatment were more likely to be selected for evaluation by an interim PET scan. Even so, we found that an interim PET scan remains a significant predictor of PFS for both low- and high-grade FL. We acknowledge other important sources of heterogeneity in our study cohorts, including use of rituximab maintenance, generation of PET scanner used, and timing of interim PET. Most patients had an interim PET scan after 3 cycles of CIT, however a subset of patients had interim scans after either 2 or 4 cycles. The prognostic value of interim PET scan appeared to be similar for these patients. Finally, we included a smaller cohort

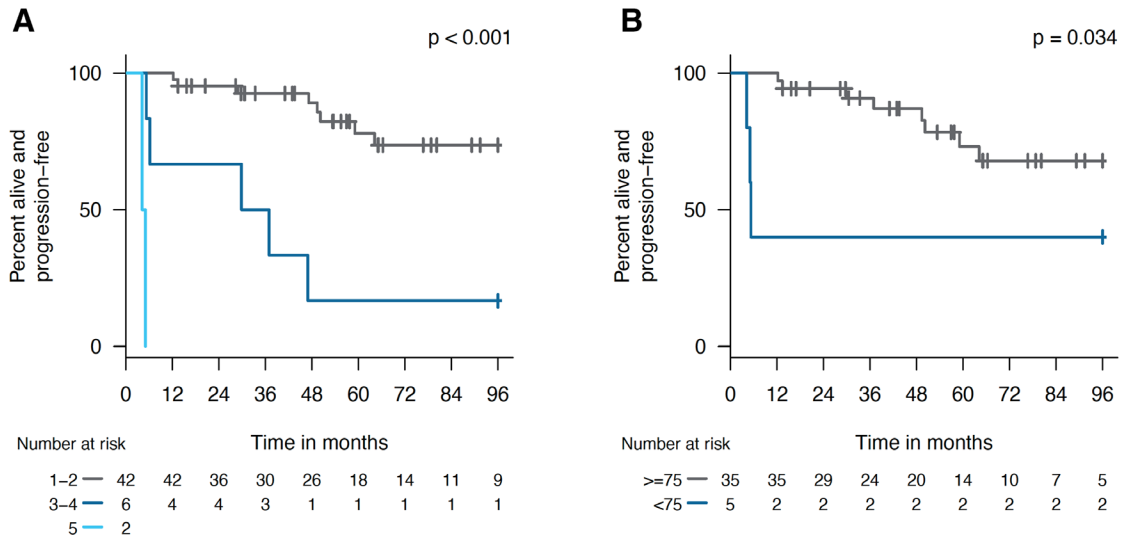

Figure 3. PFS according to interim PET, assessed using (A) DS 1-2 vs 3-4 vs 5 and (B) ΔSUVmax 75% among patients with high-grade follicular lymphoma. DS = Deauville score; PET = positron emission tomography; PFS = progression-free survival; SUVmax = maximum standardized uptake value.

of patients with high-grade FL, who were analyzed separately. In this cohort, we found a similar prognostic value for interim PET scans, but given the small number of patients, these results should be interpreted with caution.

In conclusion, an interim PET scan appears to be a useful biomarker for patients with FL receiving frontline CIT. A DS of 3–5 on an interim PET is an independent predictor of inferior PFS. Moreover, in this cohort, a positive interim PET could predict early progression, while providing response-driven prognostic information earlier in a patient's treatment course. These results suggest that interim PET should be investigated as a tool for response-adapted treatment strategies in FL.

ACKNOWLEDGMENTS

PA and RWM would like to acknowledge support from the Harold and Virginia Lash Grant Program. PA is supported by a Scholar Award from the Leukemia and Lymphoma Society. RWM would like to acknowledge support from an American Society for Transplantation and Cellular Therapy New Investigator Award, a Lymphoma Research Foundation (LRF) Clinical Investigator Career Development Award, and the LRF Lymphoma Clinical Research Mentoring Program.

AUTHOR CONTRIBUTIONS

RWM designed the research, performed research, analyzed the data, and wrote the paper; RR analyzed the data and reviewed the paper; GS, ET, GA, MC, EJ, IEA, JRB, JLC, MSD, DCF, ASF, EJ, CAJ, AIK, ASL, BLL, SYN, OOO, EMP, AZ and GS performed research and reviewed the paper. PM, HP, PA, HS, LM, and HJ designed the research, performed research, analyzed the data, and reviewed the paper.

DISCLOSURES

RWM: Consulting: Genmab, Adaptive Biotechnologies, Bristol Myers Squibb, Abbvie, Intellia, Epizyme. Research funding: Bristol Myers Squibb, Merck, Genentech/Roche, Genmab. JRB: Consulting: Abbvie, Acerta/AstraZeneca, BeiGene, Bristol-Myers Squibb/Juno/Celgene, Catapult, Eli Lilly, Genentech/Roche, Hutchmed, iOnctura, Janssen, MEI Pharma, Pharmacyclics. Research support: BeiGene, Gilead, Loxo/Lilly, MEI Pharma, SecuraBio, Sun, TG Therapeutics. JLC: Consulting: Incyte, Karyopharm, Kite. Research Funding: Bayer, Abbvie, Roche, Merck. MSD: Consulting: AbbVie, Adaptive Biosciences, Ascentage Pharma, AstraZeneca, BeiGene, BMS, Celgene, Eli Lilly, Genentech, Janssen, MEI Pharma, Novartis, Takeda, TG Therapeutics, Verastem, Zentalis. Research support: Ascentage Pharma, AstraZeneca, BMS, Genentech, MEI Pharma, Novartis, Pharmacyclics, Surface Oncology, TG Therapeutics, Verastem. EJ: Consulting: Syros, Takeda. Research funding: Acerta, Janssen, Novartis, Pharmacyclics. CAJ: Consulting: Kite/Gilead, Novartis, BMS/Celgene, bluebird bio, Epizyme, Ipsen, Instil Bio, ImmPACT Bio, Caribou Bio, Morphosys, Miltenyi, Abintus Bio. Research funding: Kite/Gilead and Pfizer. ASL: Consulting: Research to Practice, Seattle Genetics. GS: Consulting: Abbvie, Beigene, BMS/Celgene, Epizyme, Genentech/Roche, Genmab, Incyte, Janssen, Kite/Gilead, Loxo, Milteniy, Molecular Partners, Morphosys, Nordic Nanovector, Novartis, Rapt, Takeda; Debiopharm/ Velosbio/ Ipsen. Honararia: Abbvie, Bayer, Incyte, Kite/Gilead, Morphosys, Novartis, Regeneron. Shareholder: Owkin. ADZ: Consultant: BMS/Celgene/JUNO, Genentech/Roche, Gilead/Kite; BeiGene; Pharmacyclics/Abbvie, Jansen, Amgen, AstraZeneca, Novartis, MEI Pharma. Research Support/P.I.: Genentech/Roche, MEI Pharma, BeiGene; Abbvie. Scientific Advisory Board: Lymphoma Research Foundation, Adaptive Biotechnologies. PA: Consultancy: Merck, BMS, Pfizer, Affimed, Adaptive, Infinity, ADC Therapeutics, Celgene, Morphosys, Daiichi Sankyo, Miltenyi, Tessa, GenMab, C4, Enterome, Regeneron, Epizyme, Astra Zeneca, Genentech, Xencor. Research funding: Kite. Research funding (institutional): Merck, BMS, Affimed, Adaptive, Tensha, Otsuka, Sigma Tau, Genentech/Roche, IGM. Honoraria: Merck, BMS. HJ: Consulting: Advanced Accelerator Applications, Spectrum Dynamics Medical. Honoraria: Blue Earth Diagnostics. Research Funding: Siemens Healthcare, Inc, GTx, Inc, Blue Earth Diagnostics. Royalties: Cambridge University Press. All the other authors have no conflicts of interest to disclose.

REFERENCES

1. Casulo C, Byrtek M, Dawson KL, et al. Early relapse of follicular lymphoma after rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone defines patients at high risk for death: an analysis from the National LymphoCare Study. *J Clin Oncol*. 2015;33:2516–2522.
2. Jurinovic V, Kridel R, Staiger AM, et al. Clinicogenetic risk models predict early progression of follicular lymphoma after first-line immunochemotherapy. *Blood*. 2016;128:1112–1120.
3. Solal-Celigny P, Roy P, Colombat P, et al. Follicular lymphoma international prognostic index. *Blood*. 2004;104:1258–1265.
4. Federico M, Bellei M, Marcheselli L, et al. Follicular lymphoma international prognostic index 2: a new prognostic index for follicular lymphoma developed by the international follicular lymphoma prognostic factor project. *J Clin Oncol*. 2009;27:4555–4562.
5. Kimby E, Lockmer S, Holte H, et al. The simplified follicular lymphoma PRIMA-prognostic index is useful in patients with first-line chemo-free rituximab-based therapy. *Br J Haematol*. 2020;191:738–747.
6. Pastore A, Jurinovic V, Kridel R, et al. Integration of gene mutations in risk prognostication for patients receiving first-line immunochemotherapy for follicular lymphoma: a retrospective analysis of a prospective clinical trial and validation in a population-based registry. *Lancet Oncol*. 2015;16:1111–1122.
7. Mondello P, Fama A, Larson MC, et al. Lack of intrafollicular memory CD4 + T cells is predictive of early clinical failure in newly diagnosed follicular lymphoma. *Blood Cancer J*. 2021;11:130.

8. Mir F, Mattiello F, Grigg A, et al. Follicular Lymphoma Evaluation Index (FLEX): a new clinical prognostic model that is superior to existing risk scores for predicting progression-free survival and early treatment failure after frontline immunochemotherapy. *Am J Hematol*. 2020;95:1503–1510.
9. Jurinovic V, Passerini V, Oestergaard MZ, et al. Evaluation of the m7-FLIPI in patients with follicular lymphoma treated within the gallium trial: EZH2 mutation status may be a predictive marker for differential efficacy of chemotherapy. *Blood*. 2019;134(Supplement_1):122–122.
10. Huet S, Tesson B, Jais JP, et al. A gene-expression profiling score for prediction of outcome in patients with follicular lymphoma: a retrospective training and validation analysis in three international cohorts. *Lancet Oncol*. 2018;19:549–561.
11. Bolen CR, Mattiello F, Herold M, et al. Treatment dependence of prognostic gene expression signatures in de novo follicular lymphoma. *Blood*. 2021;137:2704–2707.
12. Trotman J, Barrington SF, Belada D, et al; PET investigators from the GALLIUM study. Prognostic value of end-of-induction PET response after first-line immunochemotherapy for follicular lymphoma (GALLIUM): secondary analysis of a randomised, phase 3 trial. *Lancet Oncol*. 2018;19:1530–1542.
13. 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*. 2007;25:3746–3752.
14. Johnson P, Federico M, Kirkwood A, et al. Adapted treatment guided by interim PET-CT scan in advanced Hodgkin's lymphoma. *N Engl J Med*. 2016;374:2419–2429.
15. Borchmann P, Goergen H, Kobe C, et al. PET-guided treatment in patients with advanced-stage Hodgkin's lymphoma (HD18): final results of an open-label, international, randomised phase 3 trial by the German Hodgkin Study Group. *Lancet*. 2017;390:2790–2802.
16. Casasnovas R-O, Bouabdallah R, Brice P, et al. PET-adapted treatment for newly diagnosed advanced Hodgkin lymphoma (AHL2011): a randomised, multicentre, non-inferiority, phase 3 study. *Lancet Oncol*. 2019;20:202–215.
17. Dührsen U, Müller S, Hertenstein B, et al; PETAL Trial Investigators. Positron emission tomography-guided therapy of aggressive non-Hodgkin lymphomas (PETAL): a multicenter, randomized phase III trial. *J Clin Oncol*. 2018;36:2024–2034.
18. Boo SH, O JH, Kwon SJ, et al. Predictive value of interim and end-of-therapy 18F-FDG PET/CT in patients with follicular lymphoma. *Nucl Med Mol Imaging (2010)*. 2019;53:263–269.
19. Dupuis J, Berriolo-Riedinger A, Julian A, et al. Impact of [^{18} F] fluorodeoxyglucose positron emission tomography response evaluation in patients with high–tumor burden follicular lymphoma treated with immunochemotherapy: a prospective study from the Groupe d'Etudes des Lymphomes de l'Adulte and GOELAMS. *J Clin Oncol*. 2012;30:4317–4322.
20. Bishu S, Quigley JM, Bishu SR, et al. Predictive value and diagnostic accuracy of F-18-fluoro-deoxy-glucose positron emission tomography treated grade 1 and 2 follicular lymphoma. *Leuk Lymphoma*. 2007;48:1548–1555.
21. Lu Z, Lin M, Downe P, et al. The prognostic value of mid- and post-treatment [(18)F]fluorodeoxyglucose (FDG) positron emission tomography (PET) in indolent follicular lymphoma. *Ann Nucl Med*. 2014;28:805–811.
22. Zhou Y, Zhao Z, Li J, et al. Prognostic values of baseline, interim and end-of therapy18F-FDG PET/CT in patients with follicular lymphoma. *Cancer Manag Res*. 2019;11:6871–6885.
23. Cheson BD, Fisher RI, Barrington SF, et al; Alliance, Australasian Leukaemia and Lymphoma Group. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol*. 2014;32:3059–3068.
24. Kluge R, Chavdarova L, Hoffmann M, et al. Inter-reader reliability of early FDG-PET/CT response assessment using the Deauville scale after 2 cycles of intensive chemotherapy (OEPA) in Hodgkin's lymphoma. *PLoS One*. 2016;11:e0149072.
25. Oki Y, Chuang H, Chasen B, et al. The prognostic value of interim positron emission tomography scan in patients with classical Hodgkin lymphoma. *Br J Haematol*. 2014;165:112–116.
26. Horning SJ, Juweid ME, Schöder 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.
27. Luminari S, Manni M, Galimberti S, et al. Response-adapted postinduction strategy in patients with advanced-stage follicular lymphoma: the FOLL12 study. *J Clin Oncol*. 2022;40:729–739.
28. Trotman J, Luminari S, Boussetta S, et al. Prognostic value of PET-CT after first-line therapy in patients with follicular lymphoma: a pooled analysis of central scan review in three multicentre studies. *Lancet Haematol*. 2014;1:e17–e27.
29. Casasnovas R-O, Meignan M, Berriolo-Riedinger A, et al. SUVmax reduction improves early prognosis value of interim positron emission tomography scans in diffuse large B-cell lymphoma. *Blood*. 2011;118:37–43.
30. 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.