

Positron emission tomography–computed tomography predictors of progression after DA-R-EPOCH for PMBCL

Chelsea C. Pinnix,¹ Andrea K. Ng,² Bouthaina S. Dabaja,¹ Sarah A. Milgrom,¹ Jillian R. Gunther,¹ C. David Fuller,¹ Grace L. Smith,¹ Zeinab Abou Yehia,¹ Wei Qiao,³ Christine F. Wogan,¹ Mani Akhtari,¹ Osama Mawlawi,⁴ L. Jeffrey Medeiros,⁵ Hubert H. Chuang,⁶ William Martin-Doyle,⁷ Philippe Armand,⁷ Ann S. LaCasce,⁷ Yasuhiro Oki,⁸ Michelle Fanale,⁸ Jason Westin,⁸ Sattva Neelapu,⁸ and Loretta Nastoupil⁸

¹Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX; ²Department of Radiation Oncology, Dana-Farber Cancer Institute, Boston, MA; ³Department of Biostatistics, ⁴Department of Imaging Physics, ⁵Department of Hematopathology, and ⁶Department of Nuclear Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX; ⁷Department of Hematologic Oncology, Dana-Farber Cancer Institute, Boston, MA; and ⁸Department of Lymphoma/Myeloma, The University of Texas MD Anderson Cancer Center, Houston, TX

Key Points

- Biomarkers are needed to identify patients with PMBCL who will not be cured after single-modality therapy with R-EPOCH.
- Volume-based and metabolic variables on pre- and post-chemotherapy PET-CT seem to identify patients who progress after R-EPOCH alone.

Dose-adjusted rituximab plus etoposide, prednisone, vincristine, cyclophosphamide, and doxorubicin (DA-R-EPOCH) has produced good outcomes in primary mediastinal B-cell lymphoma (PMBCL), but predictors of resistance to this treatment are unclear. We investigated whether [¹⁸F]fluorodeoxyglucose positron emission tomography–computed tomography (PET-CT) findings could identify patients with PMBCL who would not respond completely to DA-R-EPOCH. We performed a retrospective analysis of 65 patients with newly diagnosed stage I to IV PMBCL treated at 2 tertiary cancer centers who had PET-CT scans available before and after frontline therapy with DA-R-EPOCH. Pretreatment variables assessed included metabolic tumor volume (MTV) and total lesion glycolysis (TLG). Optimal cutoff points for progression-free survival (PFS) were determined by a machine learning approach. Univariate and multivariable models were constructed to assess associations between radiographic variables and PFS. At a median follow-up of 36.6 months (95% confidence interval, 28.1–45.1), 2-year PFS and overall survival rates for the 65 patients were 81.4% and 98.4%, respectively. Machine learning–derived thresholds for baseline MTV and TLG were associated with inferior PFS (elevated MTV: hazard ratio [HR], 11.5; *P* = .019; elevated TLG: HR, 8.99; *P* = .005); other pretreatment clinical factors, including International Prognostic Index and bulky (>10 cm) disease, were not. On multivariable analysis, only TLG retained statistical significance (*P* = .049). Univariate analysis of posttreatment variables revealed that residual CT tumor volume, maximum standardized uptake value, and Deauville score were associated with PFS; a Deauville score of 5 remained significant on multivariable analysis (*P* = .006). A model combining baseline TLG and end-of-therapy Deauville score identified patients at increased risk of progression.

Introduction

Primary mediastinal B-cell lymphoma (PMBCL) is a subtype of diffuse large B-cell lymphoma (DLBCL) that typically affects young patients. Until recently, doxorubicin-based chemotherapy followed by consolidative radiation therapy (RT) had been considered the standard of care.^{1–3} To improve outcomes and reduce the

risk of late toxicity, the National Cancer Institute (NCI) undertook a phase 2 trial of dose-adjusted rituximab plus etoposide prednisone, vincristine, cyclophosphamide, and doxorubicin (DA-R-EPOCH) without RT^{4,5} and found that both progression-free survival (PFS) and overall survival (OS) rates were >90%.

Despite these favorable outcomes, a minority of patients have chemotherapy-refractory disease that is difficult to salvage with additional chemotherapy, autologous stem-cell transplantation (SCT), or RT.^{6,7} Robust prognostic markers are needed to identify these patients requiring alternative therapies. The International Prognostic Index (IPI) is often of limited value.^{8,9} The International Extranodal Lymphoma Study Group (IELSG) investigated the prognostic significance of functional [¹⁸F]fluorodeoxyglucose ([¹⁸F]FDG) positron emission tomography–computed tomography (PET-CT) variables among 103 patients with PMBCL enrolled in the IELSG-26 trial.¹⁰ Those patients were treated with rituximab and non-EPOCH doxorubicin-based regimens, and 93 received consolidative RT. Metabolic activity characterized on pretreatment PET-CT imaging as metabolic tumor volume (MTV) and total lesion glycolysis (TLG) were powerful predictors of PFS and OS.

Posttreatment PET-CT findings are also prognostic in many lymphomas, including Hodgkin lymphoma and PMBCL.⁹ The Lugano classification recommends assessing response with the Deauville 5-point scale, which quantifies uptake in residual tumor masses relative to the mediastinum and liver blood pools.^{11,12} Several studies have demonstrated the impressive negative predictive value (NPV) of the Lugano response classification (D1-3) for patients with PMBCL.^{3,7,9} However, because posttreatment inflammation is common, the positive predictive value (PPV) of the Lugano classification may be limited, particularly for patients with Deauville 4 responses. Distinguishing between refractory disease and posttreatment inflammation is critically important.^{6,13}

We sought here to determine if findings on PET-CT scans obtained before and after treatment could robustly identify patients who would not be cured after DA-R-EPOCH for PMBCL.

Methods

Two groups of patients were evaluated for this study, 1 consisting of 49 consecutive patients with newly diagnosed PMBCL evaluated at MD Anderson Cancer Center (cohort 1) from 2009 to 2016, and the other consisting of 16 such patients treated at Dana-Farber Cancer Institute (cohort 2) from 2012 to 2016. All patients had baseline and postchemotherapy PET-CT scans available for review, and all diagnoses were confirmed by a hematopathologist. All patients received DA-EPOCH. After approval by both institutions' institutional review boards, we reviewed demographic, clinical, radiographic, and treatment-related factors in accordance with the Declaration of Helsinki. When radiation was required for consolidation, involved-site RT targeting was used as recommended by the International Lymphoma Radiation Oncology Group.¹⁴

PET-CT imaging

Baseline PET-CT scans had been obtained before chemotherapy under conditions that were similar at both institutions. After patients had fasted for at least 4 to 6 hours, blood glucose was measured and confirmed to be <140 mg/dL (<200 mg/dL for patients with diabetes) before injection of 333 to 407 MBq (9 to 11 mCi) of [¹⁸F]FDG. Emission scans were acquired at 2 to 3 minutes per field of view in the

3-dimensional mode after a 60-minute uptake time (± 10 minutes). CT noncontrast images were acquired in helical mode with 3.75-mm slices from the skull base through the midhigh. Commercially available iterative algorithms were used for image reconstruction.

PET-CT and imaging variables

Prechemotherapy PET images were collected and transferred to commercially available software (MIMVista version 6.4.9; MIMVista Corporation, Cleveland, OH). An isocontour threshold method was used to automatically demarcate FDG-avid disease by a single blinded observer based on 25% and 41% of the maximum standardized uptake value (SUV). The mean and maximum SUV (SUV_{max}), MTV, and TLG were determined. The 25% and 41% threshold levels were compared because both methods have been validated for patients with lymphoma.^{10,15} Although we confirmed that both methods were comparable, the 25% threshold was used given previous work on this topic in PMBCL.^{10,16} With a threshold of 25% of SUV_{max}, scans from 20 (31%) of 65 patients required manual modification, because the automatically determined contours initially included physiologic osseous (n = 2), laryngeal (n = 1), or cardiac activity with (n = 15) or without (n = 2) bone activity. All avid disease sites were considered in the threshold approach for MTV and TLG automatic delineation. Among cohort 1 patients, disease below the diaphragm was present in 2 patients and contributed to <2% of the total MTV and TLG in each patient; in cohort 2, extramediastinal disease was present in 6 patients and contributed to 12.2% of the total MTV in 1 patient, 8.2% of the total MTV in 1 patient, and <1% of the total MTV in 4 patients. The maximum dimension of mediastinal disease in the axial plane was also recorded.

For postchemotherapy images, residual abnormal soft tissue masses were contoured by a single blinded observer. Posttreatment scans were compared with pretreatment scans to confirm that all soft tissue delineated had been previously involved. SUV_{max} was determined from the posttreatment scans, which were also assessed according to the Lugano 5-point scale.¹²

Statistical methods

PFS and OS were estimated by the Kaplan-Meier method.¹⁷ PFS was defined from the date of diagnosis to disease relapse, progression, or death from any cause. Patients without relapse or progression were censored at the time of last follow-up. OS time was defined from the date of diagnosis to death from any cause.

Regarding PET variables, receiver operating characteristic curves and corresponding area under the curve analyses were used to assess the performance of each radiographic variable in relation to PFS. To identify radiographic variables associated with increased risk of disease progression (PFS), we used a machine learning approach with multivariable bootstrap resample recursive partitioning analysis using 10 000 replicates.¹⁸ This approach is well suited to situations where the goal is to identify ≥ 1 threshold continuous variables associated with a binary variable (progression/relapse) in the setting of several predictor variables. In this case, recursive partitioning was performed using the Martingale residuals to account for the right censored nature of PFS.¹⁹ To define a specific threshold correlated with PFS within the candidate PET variables, a decision tree–based partitioning with 30% verification holdback and a minimum split size of 10% per split/partition was used for the pretreatment variables (MTV and TLG) and the posttreatment variables (Deauville score, post SUV_{max}, and CT residual) and optimized on Martingale

Table 1. Patient and treatment characteristics

Characteristic	N (%)			P
	All patients	MD Anderson group	Dana-Farber group	
All patients				
Total	65	49	16	
Age, y				
Median	35	36	34.5	
Range	19-65	19-65	21-55	
Female sex	34 (52)	26 (53)	8 (50)	.831
Disease stage				.011
I	24 (37)	19 (39)	5 (31)	
II	32 (49)	27 (55)	5 (31)	
III	4 (6)	2 (4)	2 (12)	
IV	5 (8)	1 (2)	4 (25)	
B symptoms	17 (26)	15 (31)	2 (12)	.152
Performance status				.001
0-1	63 (97)	47 (96)	16 (100)	
2	2 (3)	2 (4)		
Bulky disease				
Median axial dimension, cm	10.8	10.6	11.3	.325
Range	4-18	4-18	6-18	
Maximum axial dimension \geq 10 cm	43 (66)	29 (59)	14 (87)	.038
Serum LDH, IU/L				.880
>ULN	41 (63)	31 (63)	10 (62)	
IPI score				.005
0	22 (34)	17 (35)	5 (31)	
1	33 (51)	28 (57)	5 (31)	
2	5 (8)	3 (6)	2 (12)	
3	4 (6)	0	4 (25)	
Unknown	1 (1)	1 (2)		
DA-R-EPOCH				.333
N of cycles				
Median	6	6	6	
Range	4-7	4-7	4-6	
6 cycles	62 (95)	47 (96)	15 (94)	
RT	13 (20)	10 (20)	3 (19)	.886
Consolidative	3 (5)	2 (4)	1 (6)	
Salvage	10 (15)	8 (16)	2 (12)	
Dose, Gy				.336
Median	42	39.6	42	
Range	30-49	30.6-49	30-44	
Salvage chemotherapy	10 (15)	8 (16)	2 (12)	.650
SCT	10 (15)	8 (16)	2 (12)	.713
Autologous	6 (9)	4 (8)	2 (12)	
Allogeneic	4 (6)	4 (8)		
Follow-up time, mo				.0003
Median	36.6	41.2	25.7	
95% CI	28.1-45.1	35.6-46.9	22.4-29.0	

LDH, lactate dehydrogenase; ULN, upper limit of normal.

Table 1. (continued)

Characteristic	N (%)			P
	All patients	MD Anderson group	Dana-Farber group	
Patients undergoing salvage therapy				
Total	12	10	2	
Age, y				
Median	34	34.5	28.5	
Range	21-50	21-50	22-35	
Female sex	4 (33)	4 (40)	0	
Bulky disease >10 cm	11 (92)	9 (90)	2 (100)	
Biopsy findings				
Positive	6 (50)	6 (60)		
Not done	6 (50)	4 (40)	2 (100)	
Salvage chemotherapy	10 (83)	8 (80)	2 (100)	
Salvage RT	10 (83)	8 (80)	2 (100)	
Dose, Gy				
Median	43.6	43.6	42	
Range	31-49	31-49	40-44	
SCT				
Autologous	6 (50)	4 (50)	2 (100)	
Allogeneic	4 (33)	4 (50)		

LDH, lactate dehydrogenase; ULN, upper limit of normal.

residuals.²⁰ Receiver operating characteristic analysis was performed after completion of all partitions. Post hoc K-fold cross validation (n = 10) was conducted to assess overfitting. This process was used to define optimal thresholds for each radiographic variable.

Between the institutional groups, categorical variables were compared by Fisher's exact tests, with log-rank tests used to evaluate associations between categorical variables and PFS. Univariate Cox proportional hazards regression was used to evaluate associations between prognostic factors and PFS. Multivariable Cox proportional hazard models were used to determine the effects of dichotomized pretreatment and posttreatment variables on PFS. Selection of variables for multivariable analysis was based on the number of events, results of the univariate analysis ($P < .1$), and clinical interest, with selection of covariates that minimized overfitting. For multivariate exploration, the number of variables was limited to no more than 2. Two-sided P values of $<.05$ were considered significant. Hazard ratios (HRs) and corresponding 95% confidence intervals (CIs) are reported. Bayesian information criterion (BIC) were used for model comparison, with a lower BIC indicating enhanced model performance and parsimony.^{21,22} Sensitivity, specificity, NPV, and PPV were calculated based on standard definitions.^{23,24} Statistical analyses were performed with commercially available software (JMP v12Pro; SAS Institute, Cary, NC; IBM SPSS 22.0, Chicago, IL). Graphs were constructed with GraphPad Prism (GraphPad Software Inc).

Results

Baseline clinical and treatment characteristics of the 65 patients are listed in Table 1. All patients were treated with DA-EPOCH, and 95% received 6 cycles (1 patient with CD20⁺ PMBCL did not receive rituximab).²⁵ For all patients, the median follow-up time was 36.6 months (95% CI, 28.1-45.1), and the 2-year PFS and OS rates

were 81.4% and 98.4%, respectively. For cohort 1, the median follow-up time was 41 months (95% CI, 34-48 months), and the 2-year PFS and OS rates were 79.4% and 97.9%, respectively (Figure 1). The median follow-up time was shorter in cohort 2 (27 months; 95% CI, 23-31 months; $P < .001$), but the 2-year PFS and OS rates were no different from those of cohort 1 (87.5%; $P = .5017$ and 100%, $P = .5721$, respectively; Figure 1).

Treatment failure

Disease progression or relapse occurred in 12 patients, 10 in cohort 1 and 2 in cohort 2. The median time to relapse for cohort 1 was 5.2 months (range, 3.6-17.6 months); 6 patients underwent biopsy, which was positive for PMBCL in 5 and for Hodgkin lymphoma in 1 (Table 1). In 2 of those 6 patients, the initial biopsy at relapse was negative for disease, but high clinical suspicion led to confirmation on a second biopsy. For the patient with Hodgkin lymphoma at biopsy (relapse at 17.6 months), the pretreatment biopsy had been diagnosed as PMBCL with no indication of gray zone or Hodgkin lymphoma. The 4 patients in cohort 1 without a biopsy at relapse had posttreatment PET-CT that revealed Deauville 5 in 1 patient and Deauville 4 in 3; corresponding SUV_{max} values were 15.7, 4.7, 5.4, and 5.2. All 4 of these patients had increasing activity within a residual mediastinal mass. The patient with a Deauville 5 response received RT but experienced out-of-field relapse shortly thereafter followed by successful autologous SCT. One patient had continued disease progression after salvage chemotherapy but achieved remission after autologous SCT and RT. The remaining 2 patients who did not undergo biopsy had CT evidence of mediastinal progression in addition to new FDG-avid mediastinal foci. Among all 49 patients in cohort 1, 1 died after unsuccessful salvage chemotherapy, autologous SCT, and RT.

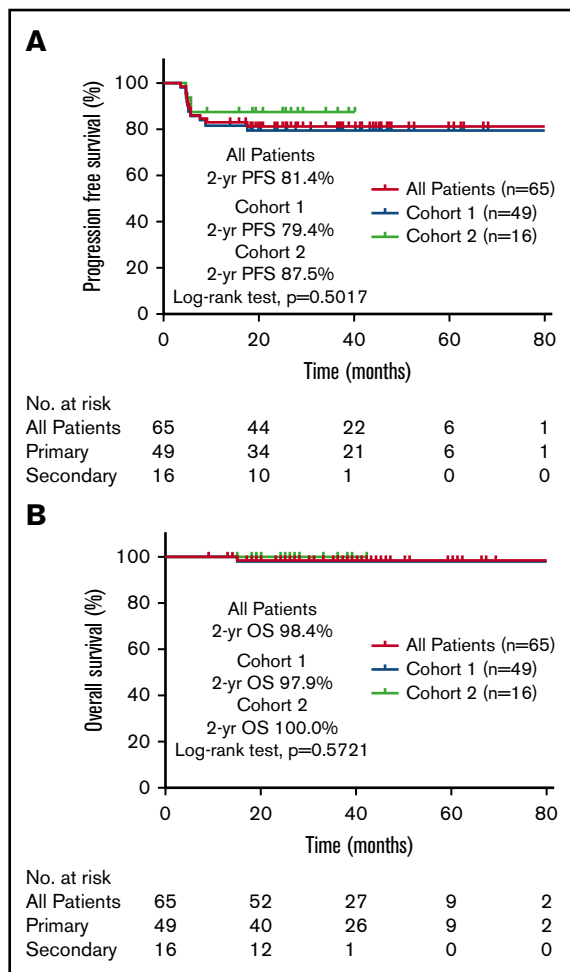


Figure 1. Survival outcomes after DA-R-EPOCH for PMBCL. PFS (A) and OS (B) after treatment with DA-R-EPOCH among patients with PMBCL treated at 2 tertiary cancer centers (cohorts 1 and 2).

Of the 2 patients with relapse in cohort 2, 1 occurred at 4.7 months and the other at 5.8 months after diagnosis; neither had biopsy at that time. One patient continued to have refractory disease after 2 lines of salvage chemotherapy but responded to salvage RT and then underwent SCT. The other patient received salvage RT followed by SCT. Both patients were alive without disease at the time of last follow-up.

Pretreatment variables

Baseline and postchemotherapy radiographic variables are listed in Table 2. Of the baseline radiographic factors evaluated (SUV_{max}, MTV, and TLG), the area under the curve was most significant for baseline TLG (0.756; $P = .006$), which corresponds to sensitivity and specificity rates of 83% and 70%, respectively, for a TLG threshold >3941.4 g. Machine learning–derived thresholds for MTV and TLG successfully classified patients as being at low and high risk of progression (Figure 2A-B).

Posttreatment variables

Kaplan-Meier estimates according to dichotomized postchemotherapy variables revealed statistically significant associations of

high CT residual mass volume, SUV_{max}, and Deauville score with inferior PFS (Figure 2C-E). Patients with Deauville 1 to 3 had a 2-year PFS rate of 100%, compared with 51% for those with Deauville 4 to 5 ($P < .001$; Figure 2E). Alternatively, when Deauville 5 was classified as positive, the PPV improved from 48% to 78%, but at the expense of sensitivity (which decreased from 100% to 58%; Table 2). Patients with a Deauville 5 response after DA-R-EPOCH had a 2-year PFS of 22% (Figure 2F).

Univariate and multivariable analyses

Univariate Cox regression analysis revealed increased risk of progression with high MTV (>323.6 ; HR, 11.528; $P = .019$) and high TLG (>3941.4 ; HR, 8.989; $P = .005$; Table 3). Bulky disease (>10 cm) and an IPI >1 were not associated with PFS ($P = .082$ and $P = .732$, respectively). All 3 dichotomized posttherapy variables evaluated (CT residual mass volume, SUV_{max}, and Deauville score) were associated with an increased risk of progression on univariate analysis. On multivariable analysis of pretreatment variables, in a model that included bulky disease with either MTV_{high} or TLG_{high}, TLG retained statistical significance (HR, 7.879; $P = .049$; Table 3). On the basis of BIC statistics, the pretreatment multivariable model that included disease bulk and elevated TLG was superior to the model with bulk and elevated MTV. In the posttreatment multivariable analysis, combining elevated CT residual volume with Deauville score of 1 to 4 vs 5, an end-of-therapy Deauville 5 response was associated with inferior PFS (HR, 9.525; $P = .006$; Table 3). BIC statistics, however, suggest that elevated CT residual volume combined with Deauville 1 to 3 vs 4 to 5 is the superior model (BIC = 78.7). In a final multivariable model analysis combining pretherapy TLG_{high} with end-of-therapy Deauville score, in the model with elevated TLG and Deauville score of 5, both variables were significant (TLG_{high}: HR, 5.046; 95% CI, 1.016-25.053; $P = .048$; Deauville 5: HR, 8.578; 95% CI, 2.470-29.798; $P = .001$). BIC statistics, however, suggest that the combination of elevated TLG with Deauville 4 to 5 may be the superior model (BIC = 77.2). The combination of elevated TLG and Deauville score of 5 had a sensitivity of 50%, specificity of 98%, PPV of 86%, and NPV of 90% (Table 2). However, if elevated TLG was combined with end-of-therapy Deauville 4 to 5, the sensitivity improved to 83% at the expense of PPV (59%; Table 2). The 2-year PFS rate for patients with elevated TLG and Deauville score of 4 to 5 was 39% (Figure 2G); for patients with high TLG and Deauville score of 5, it was 14% (Figure 2H).

Discussion

In this series, we investigated baseline and posttreatment functional PET-CT variables in 65 patients with PMBCL treated with DA-R-EPOCH at 2 institutions. Using machine learning approaches to identify potential thresholds for radiographic variables, we found that baseline PET-CT variables were more informative and more strongly associated with PFS than the clinical variables in our data set, including IPI. Our results add to the growing body of evidence supporting the prognostic utility of baseline MTV and TLG. Our findings, coupled with those of previous studies, suggest these radiographic variables can be powerful biomarkers that could improve outcomes and risk-adapted strategies for patients with PMBCL.^{10,15,16}

Other studies have documented the value of MTV and TLG in lymphoma.^{10,15,26} Among patients with relapsed/refractory Hodgkin lymphoma, baseline MTV predicted outcome and improved the predictive power of pre-SCT PET-CT.¹⁵ In IELSG-26, the only variable (including clinical and PET factors) associated with PFS

Table 2. Imaging variables (N = 66)

Variable	Median	IQR	AUC (95% CI)	P	Cutoff value	2-y PFS rate, % (low vs high)	P	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Baseline SUV _{max}	20.6	16.9-25.6	0.522 (0.354-0.690)	.813							
Baseline MTV	323.9	194.0-514.1	0.717 (0.591-0.843)	.020	323.6	96.8 vs 67.1	.003	91.67 (61.52-99.79)	56.6 (42.28-70.16)	32.35 (25.18-40.47)	96.77 (81.90-99.50)
Baseline TLG	3470.5	1913.4-5082.3	0.756 (0.631-0.882)	.006	3941.4	94.9 vs 59.9	.001	83.3 (51.59-97.91)	69.81 (55.66-81.66)	38.46 (27.86-50.28)	94.87 (83.76-98.52)
CT residual mass	33.6	11.3-64.1	0.854 (0.748-0.960)	.0001	44.05	92.4 vs 62.5	.002	75.0 (42.81-94.51)	71.7 (57.65-83.21)	37.50 (25.93-50.70)	92.88 (82.41-97.16)
Posttreatment SUV _{max}	2.83	2.5-4.4	0.929 (0.840-1.0)	<.0001	3.98	97.9 vs 34.3	<.001	91.67 (61.52-99.79)	8.68 (76.97-95.73)	64.71 (45.85-79.88)	97.92 (87.77-99.68)
Deauville score	3	2-4	0.930 (0.868-0.992)	<.0001	>3 (1-3 vs 4-5)	100 vs 51.3	<.001	100 (73.54-100.0)	75.47 (61.72-86.24)	48.0 (36.53-59.68)	100
TLG _{high} and Deauville 4 or 5					>4 (1-4 vs 5)	90.9 vs 22.2	<.001	58.33 (27.67-84.83)	96.23 (87.02-99.54)	77.78 (45.30-93.67)	91.07 (83.90-95.23)
TLG _{high} and Deauville 5						95.8 vs 39.2	<.001	83.33 (51.59-97.91)	86.79 (74.66-94.52)	58.82 (40.65-74.87)	95.83 (86.60-98.79)
						89.5 vs 14.3	<.001	50.0 (21.09-78.91)	98.11 (89.93-99.5)	85.71 (44.26-97.84)	89.66 (83.1-93.86)

AUC, area under the curve.

and OS in multivariable analysis was TLG.¹⁰ Similarly, in this series, TLG was the most significant baseline PET-CT variable associated with inferior PFS, underscoring the robustness of this variable.

Outcomes for patients with relapsed/refractory PMBCL are poor and often inferior to those for patients with nodal diffuse large B-cell lymphoma.^{6,27} Ceriani et al¹⁰ proposed using baseline TLG to select patients for frontline high-dose chemotherapy and autologous stem-cell rescue. In a report of patients with PMBCL in IELSG-26 treated with non-EPOCH chemoimmunotherapy followed by consolidative RT, the combination of baseline TLG and end-of-therapy Deauville score of 4 or 5 was associated with inferior PFS in a multivariable model.¹⁶ Our results from patients with PMBCL treated with DA-R-EPOCH generally without consolidative RT also suggest that combining pretreatment and posttreatment functional PET variables may improve discrimination in identifying candidates for treatment escalation. Indeed, our patients with elevated baseline TLG and post-DA-R-EPOCH Deauville 5 had a 2-year PFS rate of only 14%, and this combination had an impressive PPV of 86% and NPV of 90%. Consolidative RT can be successful as salvage therapy for patients with persistent disease after chemotherapy,^{3,4} but emerging evidence suggests that increased disease burden is challenging to control with radiation alone,⁷ and delaying salvage therapy could be fatal. This situation is particularly challenging in that postchemotherapy FDG-avid lesions could represent only evolving thymic inflammatory and treatment changes. A combination of pretreatment and posttreatment variables could be used to counsel at-risk patients on the potential adverse effects of additional therapy vs continued observation. More aggressive biopsy strategies can also be undertaken to establish the diagnosis of residual disease. Also, immune checkpoint and anti-CD19 chimeric antigen receptor T-cell therapies have shown promise in PMBCL, and therefore, use of functional PET variables may facilitate early introduction of immune therapy.^{28,29}

Patients with a Deauville 4 to 5 response after DA-R-EPOCH pose a therapeutic dilemma. Several studies have indicated that in PMBCL, PET-CT positivity should be based on uptake greater than the liver blood pool.^{3,7,9} In an update of the NCI study,³⁰ of 76 patients with an end-of-chemotherapy PET-CT scan, 25 (33%) were positive (Deauville 4 or 5). Five (20%) of those patients had residual disease, correlating to a 5-year event-free survival rate of 80% for Deauville 4 to 5 patients, leading the authors to conclude that end-of-therapy PET-CT did not accurately identify patients with residual disease. In the current study, 25 (38%) of 65 patients had Deauville 4 to 5 after DA-R-EPOCH, and 12 relapsed, corresponding to a 2-year PFS rate of 51%; that rate was even lower for patients with Deauville 5 (22%). Outcomes for patients with Deauville 5 in the NCI study are unclear. In the current study, no patients with Deauville 1 to 3 after DA-R-EPOCH experienced relapse, suggesting that observation after DA-R-EPOCH may be appropriate for such patients. This suggestion is further supported by findings from a recent multi-institutional study of 156 adult and pediatric patients with PMBCL treated with DA-R-EPOCH, in which the 3-year event-free survival rate was 95% for patients with Deauville 1 to 3 after systemic therapy vs 55% for those with Deauville 4 to 5.⁵

Our study has several limitations, first among them being its retrospective nature. Second, 6 of 12 patients in this study did not undergo biopsy before salvage therapy; however, most cases involved either progression after firstline salvage treatment or evidence of progression on CT, strongly suggestive of residual disease. In addition, with a relatively few number events in our cohort, there is a

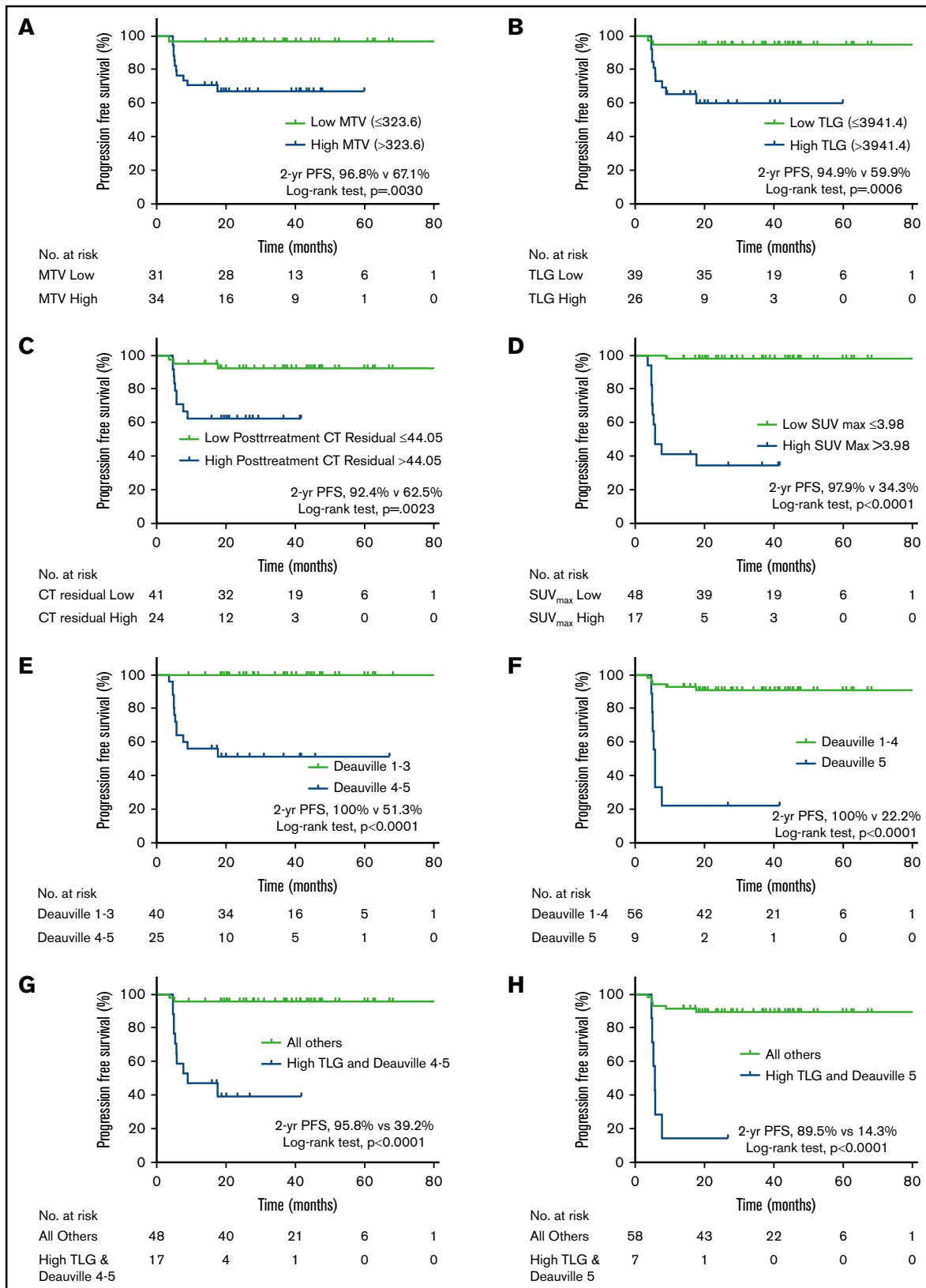


Figure 2. Effect of pretreatment and posttreatment radiographic parameters on PFS. Elevated pretreatment MTV (A) and TLG (B) were associated with inferior PFS. Posttherapy increased CT residual mass volume (C), SUV_{max} (D), and end-of-therapy Deauville score (E-F) predicted inferior PFS. The combination of pretreatment TLG and posttreatment Deauville score (G-H) were also associated with decreased PFS.

Table 3. Univariate and multivariate models of factors associated with PFS after treatment of PMBCL (N = 65)

Variable	Univariate		Multivariate model 1		Multivariate model 2	
	HR (95% CI)	P	HR (95% CI)	P	HR (95% CI)	P
Pre-DA-R-EPOCH						
Age	0.977 (0.923-1.033)	.407				
Female	0.417 (0.125-1.386)	.154				
B symptoms	0.926 (0.251-3.419)	.908				
Bulky disease (>10 cm)	6.17 (0.80-47.81)	.082	2.107 (0.233-19.042)	.507	1.282 (0.080-20.494)	.861
IPI score (0-1 vs 2)	1.304 (0.286-5.954)	.732				
MTV _{high}	11.528 (1.486-89.442)	.019	8.345 (0.924-75.387)	.059		
TLG _{high}	8.989 (1.96-41.228)	.005			7.879 (1.005-61.754)	.049
BIC			95.5		94.8	
Post-DA-R-EPOCH						
CT residual _{high}	5.974 (1.61-22.117)	.007	1.530 (0.413-5.673)	.525	1.861 (0.311-11.14)	.496
Posttreatment SUV _{max high}	49.45 (6.334-386.037)	.00019				
Deauville 1-3 vs 4-5	187.329 (0.874-40 150.184)	.056	1.7 × 10 ⁵ (0-2.0 × 10 ⁸²)	.894		
Deauville 1-4 vs 5	14.442 (4.418 – 47.209)	.00001			9.525 (1.928-47.066)	.006
BIC			78.7		88.0	
Pre- and Post-DA-R-EPOCH						
TLG _{high}			5.046 (1.016-25.053)	.048	2.703 (0.588-12.422)	.201
Deauville 1-3 vs 4-5					1.3 × 10 ⁵ (0-8.2 × 10 ⁸¹)	.896
Deauville 1-4 vs 5			8.578 (2.470-29.798)	.001		
BIC			83.6		77.2	

potential risk for insufficient power to determine an association of TLG and MTV with PFS; however, the comparative strength of the association of these variables supports the robust predictive value of these functional PET parameters. Also, our machine learning-derived thresholds for MTV and TLG are exploratory rather than confirmatory, because we identified these cut points from a single data set combining patients from 2 institutions; the number of patients in cohort 2 (n = 16) was too small to allow the cohort 1 results to be validated with an independent cohort. Although we used bootstrapping methods to account for this shortcoming, we cannot ensure that our findings apply to other patients with PMBCL treated with DA-R-EPOCH. However, in the current report, we independently identified the importance of pretreatment TLG and MTV, as did the IELSG-26 investigators for patients treated with non-EPOCH immunochemotherapy-based treatment, strongly suggesting that these variables do have clinical value. Interestingly, the thresholds identified in the IELSG study (MTV = 703 and TLG = 5814) and those in the current study (MTV = 323.6 and TLG = 3941.4) are different despite the same methods of data acquisition (25% threshold method). When we recalculated our PFS outcomes according to the IELSG-derived thresholds, the results were not significant (data not shown). In addition to potential differences in PET-CT acquisition, another probable confounder is that most patients in IELSG-26 received consolidative RT, as opposed to ~5% of the patients in the current study. Because consolidative radiation to sites of bulky disease is known to reduce rates of local relapse,³¹ the threshold for discriminating between patients at high and low risk of relapse would logically be lower for patients who do not receive RT. Therefore, using threshold values to identify high-risk patient populations must be done cautiously in the absence of verified, prospective data.^{32,33} In the future,

a standard definition of a functional PET variable will be essential, as has been the case for assessing disease response with the 5-point Deauville scale. Radiographic biomarkers should be routinely included in prospective trials (as exploratory end points) with rigorous PET-CT quality assurance programs to help identify clinically robust cut points.

Acknowledgments

This work was supported in part by the National Institutes of Health (NIH), National Cancer Institute Cancer Center Support Core Grant CA 016672 to the University of Texas MD Anderson Cancer Center. C.D.F. is a Sabin Family Foundation Fellow. C.D.F. receives funding and salary support from the National Institutes of Health (NIH), including: the National Institute for Dental and Craniofacial Research Award (1R01DE025248-01/R56DE025248-01); a National Science Foundation (NSF), Division of Mathematical Sciences, Joint NIH/NSF Initiative on Quantitative Approaches to Biomedical Big Data (QuBBD) Grant (NSF 1557679); the NIH Big Data to Knowledge (BD2K) Program of the National Cancer Institute (NCI) Early Stage Development of Technologies in Biomedical Computing, Informatics, and Big Data Science Award (1R01CA214825-01); NCI Early Phase Clinical Trials in Imaging and Image-Guided Interventions Program (1R01CA218148-01); an NIH/NCI Cancer Center Support Grant (CCSG) Pilot Research Program Award from the UT MD Anderson CCSG Radiation Oncology and Cancer Imaging Program (P30CA016672); and an NIH/NCI Head and Neck Specialized Programs of Research Excellence (SPORE) Developmental Research Program Award (P50 CA097007-10). No other funding was received for design, completion, or analysis of this study.

Authorship

Contribution: C.C.P. designed the research, collected the data, analyzed the data, and wrote the paper; A.K.N., S.A.M., J.R.G., B.S.D., and L.N. designed the research, collected the data, and wrote the paper; C.D.F. designed the research, analyzed the data, and wrote the paper; W.Q. analyzed the data and wrote the paper; G.L.S. designed the research and wrote the paper; M.A. and O.M. designed the research and collected the data; Z.A.Y. collected the data; L.J.M., H.H.C., W.M.-D., P.A., A.S.L., Y.O., M.F., J.W., and

S.N. designed the research and wrote the paper; and C.F.W. wrote the paper.

Conflict-of-interest disclosure: C.D.F. has received direct industry grant support and travel funding from Elekta AB. The remaining authors declare no competing financial interests.

Correspondence: Chelsea C. Pinnix, Department of Radiation Oncology, Unit 97, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030; e-mail: ccpinnix@mdanderson.org.

References

1. De Sanctis V, Finolezzi E, Osti MF, et al. MACOP-B and involved-field radiotherapy is an effective and safe therapy for primary mediastinal large B cell lymphoma. *Int J Radiat Oncol Biol Phys*. 2008;72(4):1154-1160.
2. Vassilakopoulos TP, Pangalis GA, Katsigiannis A, et al. Rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone with or without radiotherapy in primary mediastinal large B-cell lymphoma: the emerging standard of care. *Oncologist*. 2012;17(2):239-249.
3. Pinnix CC, Dabaja B, Ahmed MA, et al. Single-institution experience in the treatment of primary mediastinal B cell lymphoma treated with immunochemotherapy in the setting of response assessment by 18fluorodeoxyglucose positron emission tomography. *Int J Radiat Oncol Biol Phys*. 2015;92(1):113-121.
4. Dunleavy K, Pittaluga S, Maeda LS, et al. Dose-adjusted EPOCH-rituximab therapy in primary mediastinal B-cell lymphoma. *N Engl J Med*. 2013;368(15):1408-1416.
5. Giulino-Roth L, O'Donohue T, Chen Z, et al. Outcomes of adults and children with primary mediastinal B-cell lymphoma treated with dose-adjusted EPOCH-R. *Br J Haematol*. 2017;179(5):739-747.
6. Kuruvilla J, Pintilie M, Tsang R, Nagy T, Keating A, Crump M. Salvage chemotherapy and autologous stem cell transplantation are inferior for relapsed or refractory primary mediastinal large B-cell lymphoma compared with diffuse large B-cell lymphoma. *Leuk Lymphoma*. 2008;49(7):1329-1336.
7. Filippi AR, Piva C, Levis M, et al. Prognostic role of pre-radiation therapy (18)F-fluorodeoxyglucose positron emission tomography for primary mediastinal B-cell lymphomas treated with R-CHOP or R-CHOP-like chemotherapy plus radiation. *Int J Radiat Oncol Biol Phys*. 2016;95(4):1239-1243.
8. International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. *N Engl J Med*. 1993;329(14):987-994.
9. Martelli M, Ceriani L, Zucca E, et al. [18F]fluorodeoxyglucose positron emission tomography predicts survival after chemoimmunotherapy for primary mediastinal large B-cell lymphoma: results of the International Extranodal Lymphoma Study Group IELSG-26 Study. *J Clin Oncol*. 2014;32(17):1769-1775.
10. Ceriani L, Martelli M, Zinzani PL, et al. Utility of baseline 18FDG-PET/CT functional parameters in defining prognosis of primary mediastinal (thymic) large B-cell lymphoma. *Blood*. 2015;126(8):950-956.
11. Barrington SF, Mikhaeel NG, Kostakoglu L, et al. Role of imaging in the staging and response assessment of lymphoma: consensus of the International Conference on Malignant Lymphomas Imaging Working Group. *J Clin Oncol*. 2014;32(27):3048-3058.
12. Cheson BD, Fisher RI, Barrington SF, et al; United Kingdom National Cancer Research Institute. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol*. 2014;32(27):3059-3068.
13. Ceriani L, Martelli M, Gospodarowicz MK, et al. Positron emission tomography/computed tomography assessment after immunochemotherapy and irradiation using the Lugano classification criteria in the IELSG-26 study of primary mediastinal B-cell lymphoma. *Int J Radiat Oncol Biol Phys*. 2017;97(1):42-49.
14. Illidge T, Specht L, Yahalom J, et al; International Lymphoma Radiation Oncology Group. Modern radiation therapy for nodal non-Hodgkin lymphoma-target definition and dose guidelines from the International Lymphoma Radiation Oncology Group. *Int J Radiat Oncol Biol Phys*. 2014;89(1):49-58.
15. Moskowitz AJ, Schöder H, Gavane S, et al. Prognostic significance of baseline metabolic tumor volume in relapsed and refractory Hodgkin lymphoma. *Blood*. 2017;130(20):2196-2203.
16. Ceriani L, Martelli M, Conconi A, et al. Prognostic models for primary mediastinal (thymic) B-cell lymphoma derived from 18-FDG PET/CT quantitative parameters in the International Extranodal Lymphoma Study Group (IELSG) 26 study. *Br J Haematol*. 2017;178(4):588-591.
17. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53(282):471-481.
18. Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol Methods*. 2009;14(4):323-348.
19. Therneau TM, Grambsch PM, Fleming TR. Martingale-based residuals for survival models. *Biometrika*. 1990;77(1):147-160.
20. Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Ann Behav Med*. 2003;26(3):172-181.
21. Volinsky CT, Raftery AE. Bayesian information criterion for censored survival models. *Biometrics*. 2000;56(1):256-262.
22. Raftery AE. Bayesian model selection in social research. *Sociol Methodol*. 1995;25:111-163.

23. Altman DG, Bland JM. Diagnostic tests 2: predictive values. *BMJ*. 1994;309(6947):102.
24. Altman DG, Bland JM. Diagnostic tests. 1: Sensitivity and specificity. *BMJ*. 1994;308(6943):1552.
25. Oschlies I, Burkhardt B, Salaverria I, et al. Clinical, pathological and genetic features of primary mediastinal large B-cell lymphomas and mediastinal gray zone lymphomas in children. *Haematologica*. 2011;96(2):262-268.
26. Sasanelli M, Meignan M, Haioun C, et al. Pretherapy metabolic tumour volume is an independent predictor of outcome in patients with diffuse large B-cell lymphoma. *Eur J Nucl Med Mol Imaging*. 2014;41(11):2017-2022.
27. Lazzarino M, Orlandi E, Paulli M, et al. Treatment outcome and prognostic factors for primary mediastinal (thymic) B-cell lymphoma: a multicenter study of 106 patients. *J Clin Oncol*. 1997;15(4):1646-1653.
28. Zinzani PL, Ribrag V, Moskowitz CH, et al. Safety and tolerability of pembrolizumab in patients with relapsed/refractory primary mediastinal large B-cell lymphoma. *Blood*. 2017;130(3):267-270.
29. Neelapu SS, Locke FL, Bartlett NL, et al. Axicabtagene ciloleucel CAR T-cell therapy in refractory large B-cell lymphoma. *N Engl J Med*. 2017;377(26):2531-2544.
30. Melani CJ, Advani R, Chen CC, et al. DA-EPOCH-R in primary mediastinal B-cell lymphoma; analysis of end of therapy FDG-PET and outcome [abstract]. *Blood*. 2016;128(22). Abstract 1116.
31. Held G, Murawski N, Ziepert M, et al. Role of radiotherapy to bulky disease in elderly patients with aggressive B-cell lymphoma. *J Clin Oncol*. 2014;32(11):1112-1118.
32. Schöder H, Moskowitz C. Metabolic tumor volume in lymphoma: hype or hope? *J Clin Oncol*. 2016;34(30):3591-3594.
33. Adams HJ, Kwee TC. Overestimated value of baseline total metabolic tumor volume at ¹⁸F-labeled fluorodeoxyglucose positron emission tomography in follicular lymphoma. *J Clin Oncol*. 2017;35(8):918-919.