HemaSphere
Powered by EHA

EHA EUROPEAN HEMATOLOGY ASSOCIATION

Letter

OPEN ACCESS

Prognostic Value of Baseline and Interim Positron Emission Tomography Markers in Diffuse Large B-cell Lymphoma Patients: A Real-world Perspective

Stefan Hatzl[1], Peter Kalmar[2], Florian Posch[3], Jakob M. Riedl[3], Christine Beham-Schmid[4], Katharina Prochazka[1], Hildegard Greinix[1], Thomas Schwarz[5], Christian Gstettner[5], Peter Neumeister[1], Eduard Schulz[1]

Correspondence: Peter Neumeister (peter.neumeister@medunigraz.at); Eduard Schulz (eduard.schulz@medunigraz.at).

Diffuse large B-cell lymphoma (DLBCL) is treated with 6–8 cycles of R-CHOP (rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone) given every 14 or 21 days (R-CHOP14 or R-CHOP21).[1] ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG-PET-CT) is the international standard for disease staging and is frequently used for response assessment.[2] In contrast to Hodgkin lymphoma, optimal PET parameters and imaging time points to prognosticate progression-free survival (PFS) as well as overall survival (OS) of patients with DLBCL have not been established. Here, we present our experience of using PET-CT for prognostication of treatment response.

To achieve a tailored therapy, there is a need to reliably identify high-risk disease at diagnosis, early treatment failure at interim, and cure at end of treatment imaging. Prognostic models based solely on clinical features such as the National Comprehensive Cancer Network International Prognostic Index (NCCN-IPI) are widely used at diagnosis and were validated in several studies comprising thousands of patients.[3,4] International guidelines recommend PET-CT for initial staging and end of treatment response assessment, whereas interim PET-CT after 2 (i-PET2) or 4 (i-PET4) cycles of immunochemotherapy reflects chemosensitivity and has been shown to predict the outcome.[5–7] Yet, PET-guided treatment escalation has been unsuccessful so far due to lack of more effective therapeutic options.[1,5,7] Measurement of baseline total metabolic tumor volume (TMTV) and change of maximum standardized uptake value (ΔSUV_{max}) at i-PET2 are current approaches under investigation for predicting treatment response.[1,5,6,8,9] TMTV and ΔSUV_{max} were reported to be superior to the widely used Deauville five-point scale (DS).[10]

Recently, baseline TMTV cutoff >220 cm^3 has been proposed as an adverse prognostic factor from a dataset of the prospective phase 3 REMARC study including 301 patients aged 60–80 years who had been treated with 6–8 cycles of R-CHOP14 or R-CHOP21.[8] High baseline TMTV, alone or complemented by Eastern Cooperative Oncology Group performance status, showed a strong association with inferior PFS and OS in patients who achieved partial (PR) or complete remission (CR), and outperformed established clinical prognostic models such as the NCCN-IPI.

In addition, the prospective PET-CT substudy of the CALGB 50303 trial, which included 158 patients treated with 6 cycles of R-CHOP21 or dose-adjusted EPOCH-R (etoposide plus the same agents of R-CHOP), demonstrated that the change of ΔSUV_{max} at i-PET2 was significantly associated with OS.[9] In this trial, patients were 20–82 years old, and two-third had "very good" or "good" prognosis according to revised IPI (R-IPI). Unexpectedly, ΔSUV_{max} and DS were not associated with PFS.[9,11]

Our aim was to study TMTV, ΔSUV_{max}, DS, and NCCN-IPI in DLBCL patients uniformly treated with the most commonly used regimen R-CHOP21. Hence, we conducted a retrospective analysis of 144 patients with DLBCL who were diagnosed and treated at our institution during January 1, 2015, and December 31, 2019. This study was approved by the ethics committee of the Medical University of Graz. We used stringent inclusion criteria: patients with human immunodeficiency virus infection (n = 2), transformed low-grade lymphoma (n = 23), primary testicular lymphoma (n = 3), primary central nervous system lymphoma (n = 4), or no interim imaging (n = 5) were excluded.

All patients underwent PET-CT using 2 similar systems (Discovery MI; GE Healthcare, Chicago, IL; Biograph mCT; Siemens Healthineers, Erlangen, Germany). The patients fasted for at least 6 hours before intravenous administration of ^{18}F-FDG to ensure a serum glucose level below 140 mg/dL. ^{18}F-FDG was administered according to EANM procedure guidelines.[12] Analysis of imaging data was performed by 2 nuclear medicine physicians blinded to patient outcome. SUV_{max} and TMTV were computed using the semiautomatic software Hermes Hybrid 3D

[1]Division of Hematology, Department of Internal Medicine, Medical University of Graz, Austria
[2]Division of Vascular and Interventional Radiology, Department of Radiology, Medical University of Graz, Austria
[3]Division of Oncology, Department of Internal Medicine, Medical University of Graz, Austria
[4]Institute of Pathology, Medical University of Graz, Austria
[5]Department of Radiology and Division of Nuclear Medicine, Medical University of Graz, Austria
SH and PK contributed equally to this study.
Supplemental digital content is available for this article.

HemaSphere (2021) 5:8(e621).
http://dx.doi.org/10.1097/HS9.0000000000000621.
Received: 23 April 2021 / Accepted: 17 June 2021

Tumorfinder (Hermes Medical Solutions, Stockholm, Sweden). Definitions of ΔSUV_{max}, TMTV, and visual DS were as previously described.[5,9,10]

All statistical analyses were performed with Stata (Windows version 15.0; Stata Corp., Houston, TX). Median follow-up was estimated with a reverse Kaplan–Meier estimator. PET-markers were dichotomized into binary variables at the 75th percentile of their distribution. The 75% percentile cutoff was chosen empirically before data analysis to get an unbiased approach for survival estimation since other cutoffs had not been validated. Primary endpoint for all time-to-event analyses was 5-year PFS defined as the time from diagnosis to disease progression, death, or censoring alive at a maximum follow-up of 5 years, whatever came first. Secondary endpoint was 2-year PFS.

Patient characteristics are summarized in Supplemental Digital Table 1, http://links.lww.com/HS/A178. Patients had a median age of 67 (interquartile range [IQR], 53–74) years and a median R-IPI score of 3 points (ie, "poor" prognosis). Treatment consisted of 6–8 cycles of R-CHOP21 at the discretion of the treating physician. During a median follow-up of 2.4 (IQR, 1.4–3.6) years, we observed 19 (18%) primary disease progressions during first-line therapy, 10 (9%) relapses after initial CR or PR, 16 (15%) deaths related to DLBCL, and 4 deaths from other causes. The estimated 2- and 5-year PFS and OS of the whole cohort were 69% (95% confidence interval [CI], 59–77) and 80% (95% CI, 82–88), and 63% (95% CI, 51–73) and 75% (95% CI, 63–84), respectively (Supplemental Digital Table 2, http://links.lww.com/HS/A178), which is comparable with published results.[5,7–9]

We correlated TMTV and NCCN-IPI at diagnosis with PFS to assess for their prognostic impact (Table 1). Higher baseline TMTV predicted for higher risk of progression and death. Five-year PFS estimates were 41% (95% CI, 19–62) in patients with baseline TMTV above the 75th percentile of its distribution (ie, >177 cm^3 in our population), and 69% (95% CI, 55–80) in patients below this cutoff, respectively (log-rank $P = 0.006$, Supplemental Digital Figure 1A, http://links.lww.com/HS/A178). Consequently, patients with a baseline TMTV >220 cm^3 proposed as an optimal cutoff for PFS prognostication by Vercellino et al[8] showed also inferior PFS in our cohort (HR 3.54, 95% CI, 1.73–7.22; $P = 0.001$; Figure 1A). Harrell's concordance index (with higher c-indices indicating better discrimination) and Aikaike's information criterion (with lower values indicating better model fit) were comparable between TMTV >220 cm^3 and NCCN-IPI. Baseline SUV_{max} was not associated with PFS outcomes (Table 1, Supplemental Digital Figure 1B and C, http://links.lww.com/HS/A178).

We sought to establish the best model of PFS prediction using different i-PET markers. Interim PET imaging was uniformly performed after the fourth cycle of R-CHOP (i-PET4) because our center had previously delivered up to eight cycles of R-CHOP. In our dataset, TMTV, SUV_{max}, and DS comparably predicted PFS with a high discriminatory performance (Harrell's c-indices between 0.76 and 0.80; Table 1). Five-year PFS were 38% and 17% in patients with interim TMTV or interim SUV_{max} above the 75th percentile of these variables' distributions (TMTV: 11 cm^3, SUV: 7 units), and 73% and 78% in patients below these thresholds, respectively (both $P < 0.0001$, Figure 1B and C). The 2-year PFS results are summarized in Supplemental Digital Table 3, http://links.lww.com/HS/A178.

The interim DS discriminated particularly well between patients who did or did not experience disease progression or death (Table 1; Figure 1D). In a subanalysis, 5-year cumulative incidences of primary disease progression or relapse were 32% and 76% in patients with DS of 1–3 points (n = 67) and 4–5 points (n = 34), respectively (Gray's test $P < 0.0001$, Figure 1E). Complete metabolic response (CMR), defined as lack of measurable metabolic tumor tissue and SUV_{max} being indistinguishable from surrounding background activity (n = 60 patients), was also a strong PFS predictor with 5-year PFS of 83% and 31% in patients with and without a CMR at i-PET4 (log-rank $P < 0.0001$, Figure 1F). In a multivariable model of PFS including CMR at i-PET4, double expressor lymphoma biology,[13] cell of origin,[14] and NCCN-IPI, only CMR prevailed as an independent predictor of PFS (Table 2). Furthermore, we were able to replicate the cutoff ΔSUV_{max} <70% at i-PET4 as an

Table 1.

Distribution of PET Markers and Their Association With Clinical Outcome

Variable	N-Missing	Measured Value [IQR], or Absolute Count (%)	Univariable		Harrell's c for PFS	AIC for PFS
			HR for PFS	95% CI (P)		
Markers at baseline						
SUV_{max} (per doubling)	1	31 [22–43]	1.16	0.80–1.69 (0.442)	0.50	287
TMTV (per doubling)	0	86 [26–188]	1.34	1.11–1.61 (0.002)	0.67	279
TMTV > 220 cm^3	0	20 (19%)	3.54	1.73–7.22 (0.001)	0.63	278
NCCN-IPI (points)	0	3 [2–5]	1.34	1.08–1.67 (0.008)	0.60	281
Interim markers						
SUV_{max} (per doubling)	6	0 [0–7]	1.71	1.44–2.03 (<0.001)	0.80	219
TMTV (per doubling)	6	0 [0–11]	1.42	1.25–1.60 (0.001)	0.78	231
DS (per point increase)	6	1 [1–4]	2.18	1.66–2.87 (<0.001)	0.80	219
DS	6	/	/	/	0.76	229
1–2 points	/	60 (59%)	Ref.	Ref.	/	/
3 points	/	7 (7%)	1.22	0.15–9.79 (0.849)	/	/
4–5 points	/	34 (34%)	8.86	3.86–20.36 (<0.001)	/	/
Change between baseline and interim						
$\Delta_{abs}SUV_{max}$ (per 10 SUV units increase)	7	–25 [–40 to –13]	1.33	1.11–1.58 (0.002)	0.68	248
$\Delta_{rel}SUV_{max}$ (per doubling)	7	–100 [–100 to –57]	4.33	2.80–6.68 (<0.001)	0.79	226
ΔSUV_{max} < 70%	6	28 (28%)	10.23	4.73–22.17 (<0.001)	0.77	223
Δ_{abs}TMTV (per 100 cm^3 increase)	6	–70 [–160 to –25]	0.87	0.79–0.96 (0.004)	0.65	253
Δ_{rel}TMTV (per doubling)	6	–100 [–100 to –88]	1.10	1.03–1.17 (0.004)	0.75	254

Univariable modeling of PFS functions was performed with Cox proportional hazards models. The discriminative potential and model fit of PET-markers toward PFS was quantified with Harrell's concordance index and Aikaike's information criterion. N-missing denotes the number of patients without shown variables.

/ = not applicable; Δ = change; 95%CI = 95% confidence interval; Abs = absolute; AIC = Aikaike's information criterion; HR = hazard ratio; IQR = interquartile range; *P* = Wald test *P* value; PFS = progression-free survival; rel = relative; SUV = standardized uptake value; TMTV = total metabolic tumor volume.

adverse prognostic parameter (Table 1), which has previously been described in DLBCL patients treated with R-ACVBP or R-CHOP14.[5] However, although variables representing absolute or relative changes from baseline to interim were highly prognostic for PFS, they were not superior to interim markers alone in terms of discriminatory performance in our population (Table 1, Supplemental Digital Figure 1D–H, http://links.lww.com/HS/A178).

Our study confirms the prognostic value of NCCN-IPI, baseline TMTV, ΔSUV_{max} and DS in a "real-world" cohort of newly diagnosed DLBCL. The strength of our study stems from the use of a relatively large contemporary cohort characterized by representative patient age distribution and uniform therapy. Limitations mainly pertain to single-center analysis and retrospective nature.

At diagnosis, NCCN-IPI and TMTV cutoff of >220 cm^3 significantly predicted for inferior survival allowing to stratify patients for study purposes and to counsel patients on their likely disease course. After 4 cycles of R-CHOP21, visual DS, CMR, and ΔSUV_{max} <70% significantly predicted relapse or progression with high discrimination in our population. This finding is in line with the report of the GAINED trial by Le Gouill et

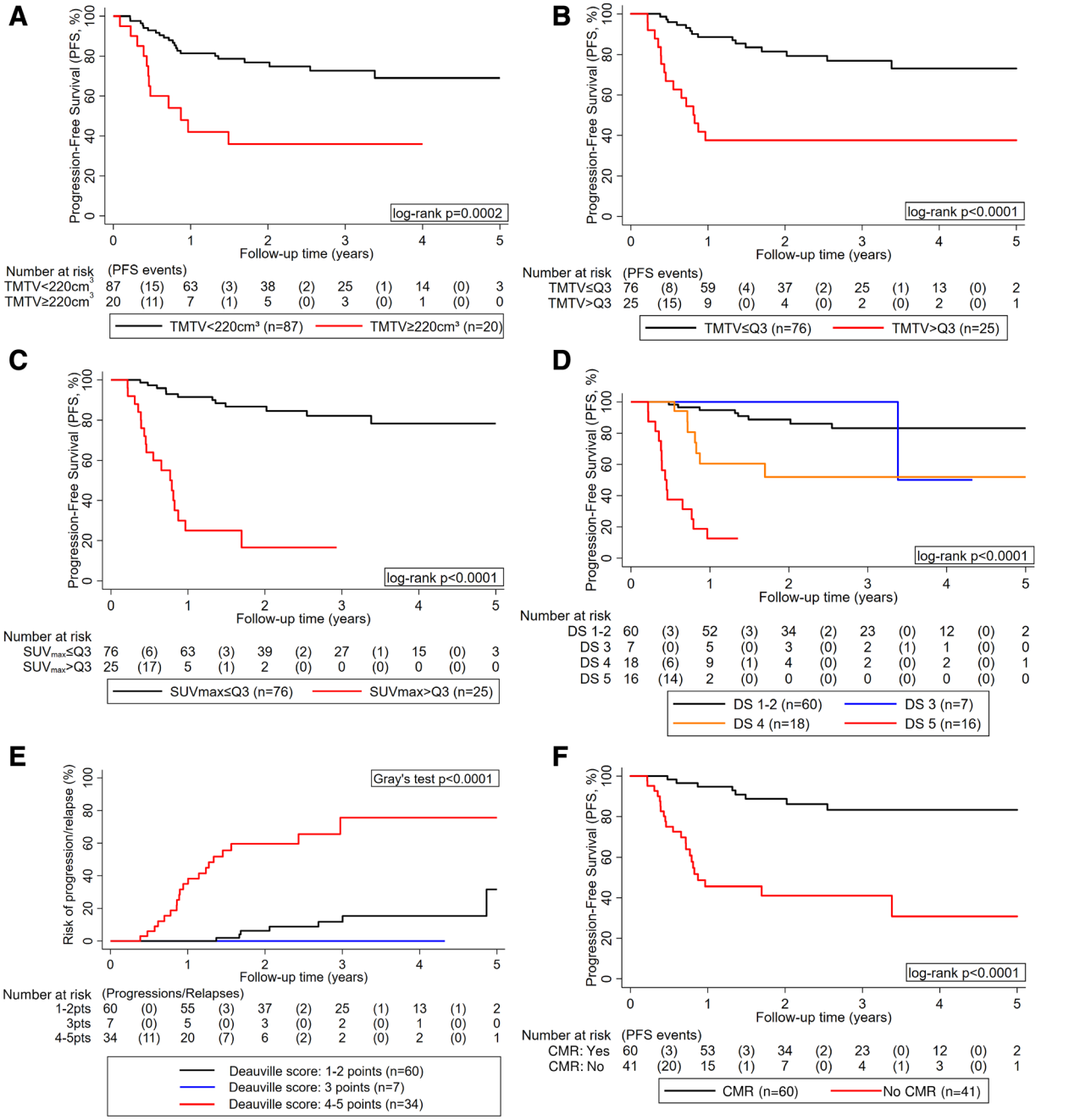

Figure 1. PET markers and clinical outcome. (A) TMTV at baseline using the cutoff 220 cm^3. (B) TMTV at interim. (C) SUV_{max} at interim. (D and E) DS at interim. (F) CMR at interim. PFS was estimated with Kaplan–Meier estimators, and compared between groups using log-rank tests. Cumulative incidences of disease progression for DS categories were calculated with competing risk estimators. The numbers below the x-axis represent a risk table, with the raw numbers indicating the number of patients at risk for a PFS event at the start of each interval and the numbers in round brackets indicating the number of patients who developed a PFS event within an interval, respectively. CMR = complete metabolic response; DS = Deauville 5-point scale; Q3 = third quartile; SUV_{max} = maximum standardized uptake value; TMTV = total metabolic tumor volume.

Table 2.

A Multivariable Model of Any Metabolic Activity at Interim PET for Prediction of PFS

Variable	Adjusted Hazard Ratio	95% CI	*P*
CMR	9.88	3.94–24.79	<0.001
DEL biology: 0 points	ref.	ref.	ref.
DEL biology: 1 point	5.21	1.18–22.92	0.029
DEL biology: 2 points	3.85	0.82–18.16	0.089
COO: ABC	0.94	0.45–1.98	0.868
NCCN-IPI (per 1 point increase)	1.06	0.81–1.39	0.674

Multivariable modeling of PFS functions was performed with Cox proportional hazards models. 95% CI = 95% confidence interval; ABC = activated B-Cell-like; CMR = complete metabolic response; COO = cell of origin (Choi's immunohistochemical algorithm[14]); DEL = immunohistochemic double expressor lymphoma status (point-based system according to Green et al[13]); NCCN-IPI = National Comprehensive Cancer Network International Prognostic Index; *P* = Wald test *P* value; ref. = reference category.

al[15] but contrasts with other studies probably reflecting different time points of i-PET (after 4 vs. 2 cycles) as well as distribution of R-IPI risk groups ("poor" versus "good" prognosis).[1,9] At i-PET4, there is likely less residual therapy-induced inflammation compared to i-PET2,[7] which might explain the better discrimination by DS in our study. Whereas FDG-avidity at i-PET2 likely reflects a mix of residual cancer cells and inflammation, i-PET4 specifically identifies resistant/refractory lymphoma.

In conclusion, our study corroborates findings that PET-CT provides important prognostic information which, after appropriate standardization and availability of more effective agents, might guide therapies in DLBCL. Patients with baseline TMTV >220 cm^3 at diagnosis and DS ≥4 or ΔSUV_{max} <70% at i-PET4 have an exceedingly high risk of relapse and should preferably be assigned to alternative treatments in prospective clinical trials.

Disclosures

The authors have no conflicts of interest to disclose.

References

1. Schmitz C, Hüttmann A, Müller SP, et al. Dynamic risk assessment based on positron emission tomography scanning in diffuse large B-cell lymphoma: post-hoc analysis from the PETAL trial. *Eur J Cancer.* 2020;124:25–36.
2. Cheson BD, Fisher RI, Barrington SF, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol.* 2014;32:3059–3068.
3. Zhou Z, Sehn LH, Rademaker AW, et al. An enhanced International Prognostic Index (NCCN-IPI) for patients with diffuse large B-cell lymphoma treated in the rituximab era. *Blood.* 2014;123:837–842.
4. Melchardt T, Troppan K, Weiss L, et al. A modified scoring of the NCCN-IPI is more accurate in the elderly and is improved by albumin and β2 -microglobulin. *Br J Haematol.* 2015;168:239–245.
5. Casasnovas RO, Ysebaert L, Thieblemont C, et al. FDG-PET-driven consolidation strategy in diffuse large B-cell lymphoma: final results of a randomized phase 2 study. *Blood.* 2017;130:1315–1326.
6. Mamot C, Klingbiel D, Hitz F, et al. Final results of a prospective evaluation of the predictive value of interim positron emission tomography in patients with diffuse large B-cell lymphoma treated with R-CHOP-14 (SAKK 38/07). *J Clin Oncol.* 2015;33:2523–2529.
7. Dührsen U, Müller S, Hertenstein B, et al. Positron emission tomography-guided therapy of aggressive non-Hodgkin lymphomas (PETAL): a multicenter, randomized phase III trial. *J Clin Oncol.* 2018;36:2024–2034.
8. Vercellino L, Cottereau AS, Casasnovas O, et al. High total metabolic tumor volume at baseline predicts survival independent of response to therapy. *Blood.* 2020;135:1396–1405.
9. Schöder H, Polley MC, Knopp MV, et al. Prognostic value of interim FDG-PET in diffuse large cell lymphoma: results from the CALGB 50303 clinical trial. *Blood.* 2020;135:2224–2234.
10. Meignan M, Gallamini A, Meignan M, et al. Report on the First International Workshop on Interim-PET-scan in lymphoma. *Leuk Lymphoma.* 2009;50:1257–1260.
11. Safar V, Dupuis J, Itti E, et al. Interim [18F]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.
12. Boellaard R, Delgado-Bolton R, Oyen WJ, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging.* 2015;42:328–354.
13. Green TM, Young KH, Visco C, et al. Immunohistochemical double-hit score is a strong predictor of outcome in patients with diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. *J Clin Oncol.* 2012;30:3460–3467.
14. Choi WW, Weisenburger DD, Greiner TC, et al. A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. *Clin Cancer Res.* 2009;15:5494–5502.
15. Le Gouill S, Ghesquières H, Oberic L, et al. Obinutuzumab vs rituximab for advanced DLBCL: a PET-guided and randomized phase 3 study by LYSA. *Blood.* 2021;137:2307–2320.