



LYMPHOID NEOPLASIA

Prognostic value of baseline metabolic tumor volume in early-stage Hodgkin lymphoma in the standard arm of the H10 trial

Anne-Ségoène Cottreau,¹ Annibale Versari,² Annika Loft,³ Olivier Casasnovas,⁴ Monica Bellei,⁵ Romain Ricci,⁶ Stéphane Bardet,⁷ Antonio Castagnoli,⁸ Pauline Brice,⁹ John Raemaekers,¹⁰ Bénédicte Deau,¹¹ Catherine Fortpied,¹² Tiana Raveloarivahy,¹² Emelie Van Zele,⁶ Loïc Chartier,¹³ Thierry Vander Borcht,¹⁴ Massimo Federico,⁵ Martin Hutchings,¹⁵ Umberto Ricardi,¹⁶ Marc Andre,¹⁷ and Michel Meignan^{18,19}

¹Department of Nuclear Medicine, Cochin Hospital, Assistance Publique–Hôpitaux de Paris (AP-HP), Paris Descartes University, Paris, France; ²Santa Maria Nuova Hospital–IRCCS Reggio Emilia, Reggio Emilia, Italy; ³Department of Clinical Physiology, Nuclear Medicine, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark; ⁴Hematology Department, INSERM 1231, University Hospital, Dijon, France; ⁵Department of Diagnostic, Clinical and Public Health Medicine, University of Modena and Reggio Emilia, Modena, Italy; ⁶Lymphoma Academic Research Organization (LYSARC) Lymphoma Study Association Imaging, Hôpital Henri Mondor, Créteil, France; ⁷Department of Nuclear Medicine, Centre de Lutte Contre le Cancer François Baclesse, Caen, France; ⁸Nuclear Medicine Unit, Ospedale Santo Stefano, Prato, Italy; ⁹Hematology, Hôpital Saint-Louis, AP-HP, Paris, France; ¹⁰Hematology, Radboud University Medical Center, Nijmegen, The Netherlands; ¹¹Hematology, Hôpital Cochin, Paris, France; ¹²European Organisation for Research and Treatment of Cancer, Brussels, Belgium; ¹³Biostatistical Department, LYSARC, Centre Hospitalier Lyon Sud, Lyon, France; ¹⁴Nuclear Medicine Division, Centre Hospitalo-Universitaire Université Catholique de Louvain, Namur, Yvoir, Belgium; ¹⁵Department of Hematology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark; ¹⁶Department of Oncology, University of Turin, Turin, Italy; ¹⁷Hematology Department, Centre Hospitalo-Universitaire Université Catholique de Louvain, Namur, Yvoir, Belgium; ¹⁸Lymphoma Study Association Imaging, Hôpital Henri Mondor, Créteil, France; and ¹⁹Université Paris Est, Créteil, France

KEY POINTS

- Baseline metabolic tumor volume is a strong prognostic factor in early-stage HL.
- Baseline metabolic tumor volume affects the early response to treatment and, combined with early PET, improves risk stratification.

We tested baseline positron emission tomography (PET)/computed tomography (CT) as a measure of total tumor burden to better identify high-risk patients with early-stage Hodgkin lymphoma (HL). Patients with stage I-II HL enrolled in the standard arm (combined modality treatment) of the H10 trial (NCT00433433) with available baseline PET and interim PET (iPET2) after 2 cycles of doxorubicin, bleomycin, vinblastine, and dacarbazine were included. Total metabolic tumor volume (TMTV) was measured on baseline PET. iPET2 findings were reported negative (DS1-3) or positive (DS4-5) with the Deauville scale (DS). The prognostic value of TMTV was evaluated and compared with baseline characteristics, staging classifications, and iPET2. A total of 258 patients were eligible: 101 favorable and 157 unfavorable. The median follow-up was 55 months, with 27 progression-free survival (PFS) and 12 overall survival (OS) events. TMTV was a prognosticator of PFS ($P < .0001$) and OS ($P = .0001$), with 86% and 84% specificity, respectively. Five-year PFS and OS were 71% and 83% in the high-TMTV ($>147 \text{ cm}^3$) group ($n = 46$), respectively, vs 92% and 98% in the low-TMTV group ($\leq 147 \text{ cm}^3$). In multi-

variable analysis including iPET2, TMTV was the only baseline prognosticator compared with the current staging systems proposed by the European Organization for Research and Treatment of Cancer/Groupe d'Etude des Lymphomes de l'Adulte, German Hodgkin Study Group, or National Comprehensive Cancer Network. TMTV and iPET2 were independently prognostic and, combined, identified 4 risk groups: low (TMTV ≤ 147 +DS1-3; 5-year PFS, 95%), low-intermediate (TMTV >147 +DS1-3; 5-year PFS, 81.6%), high-intermediate (TMTV ≤ 147 +DS4-5; 5-year PFS, 50%), and high (TMTV >147 +DS4-5; 5-year PFS, 25%). TMTV improves baseline risk stratification of patients with early-stage HL compared with current staging systems and the predictive value of early PET response as well. (*Blood*. 2018;131(13):1456-1463)

Introduction

In early-stage Hodgkin lymphoma (HL), many factors have been shown to be of prognostic significance; notably, bulky disease, number of regions involved, B-symptoms, erythrocyte sedimentation rate (ESR), and advanced age.¹ These factors are diversely integrated in the different staging systems developed by the European Organization for Research and Treatment of Cancer (EORTC),² the German Hodgkin Study Group (GHSG),³ and the National Comprehensive Cancer Network (NCCN),⁴ leading

to different risk categories. As a consequence, the definition of the unfavorable (U) risk group changes in the different prospective trials, resulting in clinical difficulties when comparing final results. For example, the Response-Adjusted Therapy for Hodgkin Lymphoma (RATHL) trial,⁵ investigating treatment escalation based on early positron emission tomography (PET) response in advanced HL, also included stage IIA with adverse features considered early unfavorable stage for the EORTC. Therefore, improving the ability to identify high-risk patients in

early-stage HL is needed, and a single prognostic scoring system would simplify the staging.

Almost all the different prognostic factors adopted so far were clearly surrogates of tumor burden (TB) and aimed to give an indirect, fractional appraisal of it. The first attempt to approach the total TB was made by Specht et al in the prospective Danish National Hodgkin Study,⁶ with an index combining the tumor size of each involved region with the number of involved regions. Recently, total TB measurement with computed tomography (CT) has been proposed.⁷ However, in a retrospective analysis of 1173 patients with early-stage HL treated homogeneously in the HD10 and HD11 trials,⁸ the GHSG showed that the best risk models included not only a large TB but also a systemic inflammation assessed by an elevated ESR. Indeed, clinical and pathological features of HL depend on the interaction between tumor and microenvironmental cells that maintains an intense inflammatory reaction. Today, unlike CT, functional imaging using ¹⁸F-fluorodeoxyglucose–PET provides the possibility to measure the total metabolic tumor volume (TMTV), related both to the tumor size and the activity of tumor and microenvironment cells. For these reasons, baseline TMTV could be a new risk factor that is helpful in stratifying early-stage patients. Recent studies have reported that a high TMTV predicted a lower survival in various non-Hodgkin lymphoma subtypes,^{9–13} but only few retrospective studies have confirmed this promising role in early HL.^{14,15}

Therefore, we investigated the prognostic value of baseline TMTV in a prospective series of early-stage HL from the standard arm of the H10 Intergroup trial (EORTC, Lymphoma Study Association [LYSA] formerly GELA [Groupe d'Etude des Lymphomes de l'Adulte], and the Fondazione Italiana Linfomi). TMTV was compared with clinic-biological prognostic factors used in the different classification systems and with early PET response (iPET2), which is now proposed as a tool for guiding therapy in HL.^{16,17}

Materials and methods

Study design and participants

We enrolled patients from the H10 Intergroup trial (NCT00433433), a randomized trial to evaluate treatment adaptation on the basis of early PET response after 2 cycles (iPET2). The study was approved by the scientific and ethical committees, and all patients gave written informed consent. In the current study, we selected only patients from the standard arm, who received a standard combined modality treatment, regardless of iPET2 result, and included by LYSA centers. Their imaging data were centralized in a dedicated platform during the trial.¹⁸ Patients had supradiaphragmatic stage I and II HL and were aged 15 to 70 years. Both favorable (F) and U patients, according to EORTC/LYSA criteria, entered² (U: at least 1 of the following criteria: age \geq 50 years or \geq 4 nodal areas or mediastinal-thoracic ratio \geq 0.35 or no B symptoms and ESR \geq 50 or B symptoms and ESR \geq 30, F: all others). The standard combined modality treatment consisted of doxorubicin, bleomycin, vinblastine, and dacarbazine (ABVD) chemotherapy, 3 cycles for F or 4 cycles for U, followed by 30 Gy involved-node radiotherapy.¹⁷ All these patients had an early PET after 2 cycles of ABVD (iPET2) with no effect on therapy. Baseline PET was recommended but not mandatory. Only 0.2% of patients in the entire trial had progression before iPET2. Only patients whose TMTV could be computed

from baseline PETscan were eligible for this analysis. Patients were also classified as low risk or high risk according to GHSG³ (the high-risk group includes at least 1 of the following factors: mediastinal mass ratio >0.33 , B symptoms with ESR \geq 30, no B symptoms with ESR \geq 50, \geq 3 lymph node sites involved, extranodal involvement) and the NCCN scoring system (the high-risk group includes at least 1 of the following factors: bulky mediastinal disease [mediastinal mass ratio >0.33] or bulky disease greater than 10 cm, B symptoms, ESR \geq 50, and more than 3 sites of disease). Progression-free survival (PFS) is defined as the time from entry onto a study until lymphoma progression or death as a result of any cause. Overall survival (OS) is defined as the time from entry onto the clinical trial until death as a result of any cause.¹⁹

Procedures

Baseline PET image data, in anonymized Digital Imaging and Communications in Medicine format, were collected for functional parameter measurements. Analysis of imaging data was performed by 3 nuclear medicine physicians (A.-S.C., A.V., and A.L.) blinded to patient outcome, who each analyzed a randomized third of the population. TMTV was computed using the free semiautomatic software Beth Israel Fijj²⁰ (<http://petctviewer.org>). Lesions were identified by visual assessment, with PET images scaled to a fixed standardized uptake value (SUV) display and color table. TMTV was obtained by summing the metabolic volumes of all local (L) nodal and extranodal lesions ($TMTV = \sum MTV_L$). The 41% maximum SUV (SUV_{max}) threshold method was used for MTV_L computation, as recommended by the European Association of Nuclear Medicine²¹ and published in various lymphoma subtypes.^{11–13,15,22} A volume of interest (VOI) was set around each lesion (node or organ involvement), as previously described.²² To avoid the underestimation of volume of bulky regions made of contiguous lymph nodes with different SUV_{max}, we first drew a VOI engulfing this bulky region. If the volume determined on the basis of the SUV_{max} of the whole region had left out nodes with SUV lower than 41%, additional VOIs were drawn within the initial VOI targeting these nodes. An example can be found at <http://www.petctviewer.org/index.php/feature/quantification> and in Supplemental 1, available on the Blood Web site. PET after 2 cycles of ABVD (iPET2), initially prospectively scored according to International Harmonization Project criteria in the H10 trial, was re-analyzed on the basis of the Deauville 5-point scale (5-DS),^{23,24} with scores 4 to 5 for positivity (fluorodeoxyglucose uptake higher than the liver).

Statistical

Two different approaches, X-tile analysis²⁵ and receiver operating characteristic analysis, were used to define the optimal cutoff of TMTV for survival prediction. This cutoff was validated by using a training/validation method. A random sample of two thirds of the patients was determined by X-tile as the training cohort, and the remaining one third used as the validation cohort. Survival functions were calculated by using Kaplan-Meier estimates, and comparison between categories was made using the log-rank test. Characteristics of populations were compared by using χ^2 , Fisher, or Mann-Whitney *U* tests. A backward stepwise Cox model with all significant baseline univariate predictors and iPET2 was performed. Because EORTC, GHSG, and NCCN are correlated with one another, 4 separate models were analyzed testing TMTV, iPET2, and all the individual factors with $P < .15$ in univariate analysis, EORTC, GHSG, and NCCN classifications. Independent variables were combined for survival prediction. A stratified Cox model was used to account for

Table 1. Patient characteristics

Characteristics	Total population (N = 258)	Low TB TMTV ≤147 cm ³ (n = 212)	High TB TMTV >147 cm ³ (n = 46)	P
Median age (range), y	31 (15-71)	32 (15-71)	27 (17-63)	.009
Age ≥50 y (%)	34 (13)	31 (15)	3 (7)	.22
Male, n (%)	129 (50)	104 (49)	25 (54)	.62
Nodular sclerosis histology n (%)	207 (80)*	168 (79)	37 (80)	.68
Ann Arbor stage II, n (%)	198 (77)	157 (74)	41 (89)	.03
B symptoms, n (%)	85 (33)	60 (28)	25 (54)	.001
Median ESR (interquartile range), mm/h	26	23 (12-50)	49 (26-72)	.0001
≥4 Involved sites, n (%)	28 (11)	17 (8)	11 (24)	.004
Bulk mediastinum (M/T ≥ 0.35), n (%)	62 (24)	34 (16)	29 (63)	<.0001
U EORTC, n (%)	157 (61)	117 (55)	40 (87)	.0001
U GSHG, n (%)	177 (69)	133 (63)	44 (96)	<.0001
U NCCN, n (%)	164 (64)	121 (57)	43 (93)	<.0001
Positive iPET2 (DS 4-5), n (%)	21 (8)	13 (6)	8 (17)	.028

M/T, mass/thoracic ratio.

*Data not available for 2 patients.

differences between the 2 treatment regimens (3 or 4 cycles). Differences between results of comparative tests were considered significant if the 2-sided *P* value was <.05. Reproducibility between the reviewers was tested on a subset of 25 patients. The Lin concordance correlation coefficient ρ_c , the Pearson ρ , and interobserver agreement to classify TMTV in the high-risk or low-risk group were measured in a sample of 25 patients by A.-S.C./A.V., A.-S.C./A.L., and A.L./A.V. Statistical analyses used SAS 9.2, X-tile 3.6.1 software (Yale University, New Haven, CT), and MedCalc software (MedCalc Software, Ostend, Belgium).

Results

Of 549 patients from the standard arm recruited by LYSA centers, 294 baseline PET scans were sent to the imaging platform during the trial. Low-quality examinations with no possibility of computing quantitative parameters were excluded. A total of 258 patients (101 F and 157 U) were suitable for TMTV calculation and were included in this study (Supplemental 2: consort diagram).

Baseline characteristics of the current population (Table 1) did not differ significantly from the whole trial standard arm population (Supplemental 2). With a median follow-up of 55 months from registration, there were 27 PFS events (only 3 in the F group) and 12 OS events (none in the F group). The 5-year PFS was 88% and 5-year OS was 95%, and they did not differ significantly from those of the group not included in this study. In the F group, the 5-year PFS and OS were 95% and 100%, respectively, vs 84% and 92% in the U group.

Baseline PET parameters

Median TMTV was 67 cm³ (interquartile range, 32-114 cm³). A significant difference was observed between the U and the

F groups, with a median TMTV of 87 cm³ contrasting with 48 cm³ (*P* < .0001). TMTV calculation was highly reproducible (ρ_c from 0.98 to 0.99). Interobserver agreement to classify TMTV in the high-risk or low-risk groups was also excellent for A.-S.C./A.V. (κ = 1), A.-S.C./A.L. (κ = 0.9), and A.V./A.L. (κ = 0.9).

From the results of X-tile and the training validation procedure, the best TMTV cutoff was 147 cm³ for PFS and OS (Supplemental 3). The presence of a high TMTV (>147 cm³) was significantly associated with a shorter PFS and OS (*P* < .0001 [hazard ratio (HR), 5.2] and *P* = .0001 [HR, 7.2], respectively). The 46 patients with a high TMTV had a significantly worse outcome with a 5-year PFS and OS of 71% and 83% vs 92% and 98% for patients with a lower TMTV (Figure 1). The prognostic value of TMTV was not affected by the stratification on the treatment arm. Sensitivity and specificity of the TMTV cutoff were 48% and 86% for PFS and 58% and 84% for OS, respectively. Within 3 years, excluding all patients censored before the end of the 3-year period without a prior event, they were, respectively, 65% and 86.57% for PFS and 75% and 84.21% for OS.

Patient characteristics stratified according to high or low TMTV values are given in Table 1. A high TMTV was associated with the extension of the disease, with significantly bulkier mediastinum, more nodal involved areas, stage II, B symptoms, and higher ESR.

Regarding baseline SUVmax, no significant cutoff value for PFS and OS prediction could be found.

Individual baseline clinicobiological factors and EORTC, GSHG, and NCCN staging systems

Age was not associated with either PFS or OS, whereas the presence of B symptoms and 4 or more lymph nodes sites

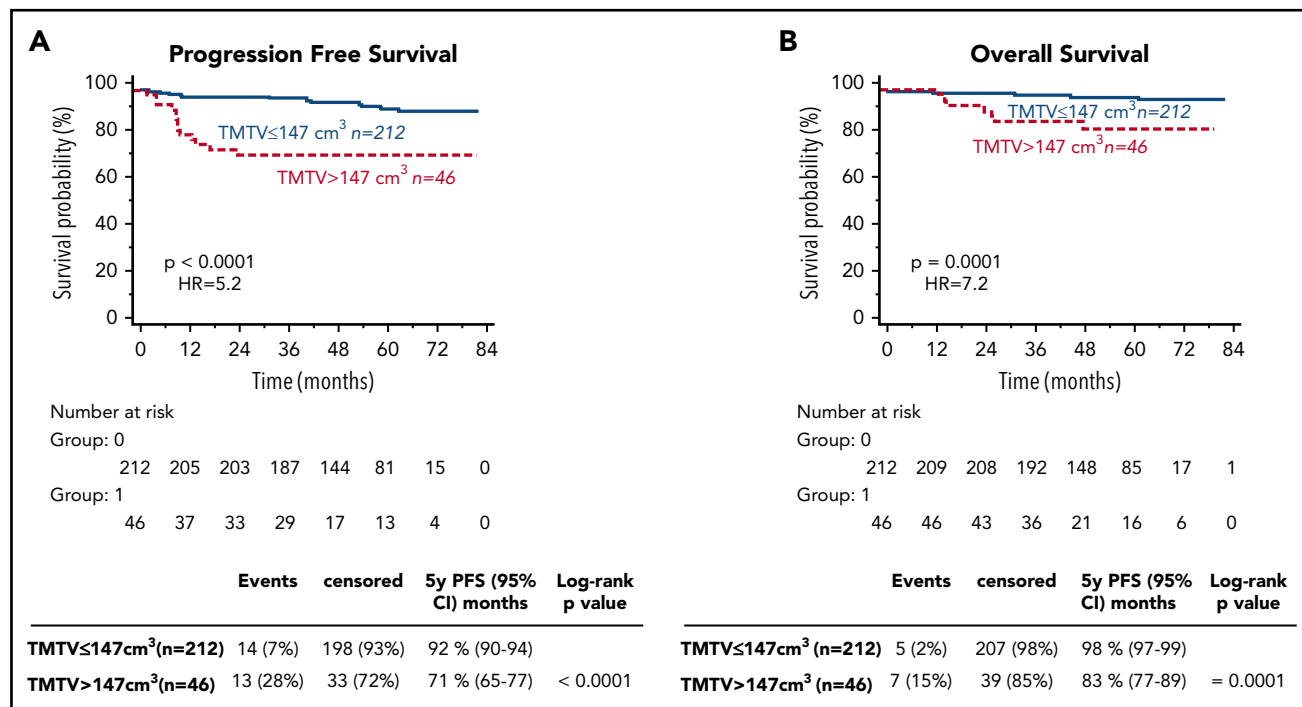


Figure 1. PFS and OS according to high total metabolic tumor volume (TMTV >147 cm³) or low TMTV (TMTV ≤147 cm³).

involved were prognostic for both PFS and OS. M/T ratio of 0.35 or higher was prognostic only for PFS (Table 2). All staging systems were predictors of PFS and OS except GHSG for OS, which did not reach significance ($P = .075$; Table 2). The numbers of high-risk patients were, respectively, 157, 164, and 177 for EORTC, NCCN, and GHSG classifications. However, the group of 34 patients with IIb with M/T higher than 0.33 or extranodal disease that would have been included in advanced stage by GHSG has significantly worse PFS and OS ($P = .0002$ and $P = .015$, respectively).

In a subanalysis of EORTC U patients, TMTV maintained its prognostic significance for both PFS and OS ($P = .0001$ [HR, 4.2] and $P = .0035$ [HR, 4.0], respectively). Patients with a small volume (74%), despite belonging to the U group, had a 5-year PFS and OS of 90% and 96%, contrasting with only 67% and 80% for patients with a large TMTV

Interim PET assessment

Interim PET2 reported with Deauville criteria was positive in 8% of the cases (3% in the F group and 11% in U group; Table 1).

Table 2. Univariate analysis for baseline prognostic factors of survival

Prognostic factor	PFS		OS	
	HR (95% CI)	P	HR (95% CI)	P
Male	1.3 (0.6-2.7)	.53	0.7 (0.2-2.3)	.60
Age ≥50 y	1.8 (0.6-5.6)	.18	2.2 (0.4-11.4)	.23
Mixed cellularity	0.8 (0.3-2.2)	.68	1.1 (0.2-4.8)	.92
Presence of B symptoms	2.9 (1.3-6.7)	.0035	4.6 (1.3-15.6)	.006
≥4 Involved sites	2.7 (0.7- 9.5)	.028	4.7 (0.7-31.5)	.006
Bulk mediastinum (M/T ≥ 0.35)	0.9-5.7	.026	2.3 (0.6-8.9)	.13
TMTV >147 cm³	5.2 (1.8-14.7)	<.0001	7.2 (1.6-33.4)	.0001
U EORTC	5.7 (2.7-12.3)	.0013	NR	.0039
U GHSG	2.8 (1.3-6.3)	.046	5.3 (1.6-17.6)	.075
U NCCN	3.6 (1.7-7.9)	.011	6.7 (2.1-21.6)	.034
Positive iPET2 (DS 4-5)	12 (2.3-63.7)	<.0001	13.2 (1.4-128.3)	<.0001

NR, not reached.

Table 3. Multivariate analysis testing TMTV, with interim PET response after 2 cycles (iPET2) and individual baseline factors, EORTC, GHSG, NCCN staging systems

TMTV tested with	PFS*			PFS (final model)			OS*			OS (final model)		
	HR	95% CI	P	HR	95% CI	P	HR	95% CI	P	HR	95% CI	P
Individual factors												
TMTV >147 cm ³	3.9	1.6-9.5	.0032	4.4	2.0-9.5	.0002	3.7	0.9-14.6	.066	5.5	1.7-17.9	.0043
iPET2	11.0	4.8-25.1	<.0001	10.9	4.9-24.4	<.0001	11.3	3.2-39.9	.0002	11.1	3.4-36.4	<.0001
B symptoms	2.1	0.9-4.8	.076				2.6	0.7-9.5	.16			
≥4 involved sites	2.0	0.8-5.2	.16				3.4	0.9-12.3	.065			
M/T ≥ 0.35	0.8	0.3-2.0	.65				0.6	0.2-2.4	.51			
EORTC												
TMTV >147 cm ³	3.5	1.6-7.8	.0016	4.4	2.0-9.5	.0002	3.9	1.2-12.4	.024	5.5	1.7-17.9	.0043
iPET2	9.2	4.1-20.6	<.0001	10.9	4.9-24.4	<.0001	8.8	2.7-28.7	.0003	11.1	3.4-36.4	<.0001
U EORTC	3.2	0.9-11.1	.067				—	—	.9			
GHSG												
TMTV >147 cm ³	4.1	1.8-9.3	.0006	4.4	2.0-9.5	.0002	4.8	1.4-16.3	.0115	5.5	1.7-17.9	.0043
iPET2	10.6	4.7-23.9	<.0001	10.9	4.9-24.4	<.0001	10.4	3.1-34.2	.0001	11.1	3.4-36.4	<.0001
U GHSG	1.3	0.4-4.0	.69				2.0	0.2-17.2	.55			
NCCN												
TMTV >147 cm ³	3.7	1.7-8.4	.00014	4.4	2.0-9.5	.0002	4.3	1.3-14.6	.0182	5.5	1.7-17.9	.0043
iPET2	10.2	4.5-22.8	<.0001	10.9	4.9-24.4	<.0001	10.2	3.1-33.4	.0001	11.1	3.4-36.4	<.0001
U NCCN	1.8	0.6-5.7	.30				2.9	0.3-25.0	.34			

*All variables integrated in the Cox model; final model: with significant factors after performing the backward stepwise Cox model.

iPET2 was predictive of PFS and OS ($P < .0001$ [HR, 12] and $P < .0001$ [HR, 13.6], respectively). Patients with positive iPET2 ($n = 21$) had a 5-year PFS and OS of 38% and 68% vs 92.6% and 98%, respectively. The frequency of iPET2 positivity was significantly higher in patients with high TMTV (Table 1), but 62% of patients with positive iPET2 had low TMTV and 83% of patients with high TMTV had a negative iPET2.

Multivariable analysis including baseline parameters and iPET2

On multivariable analysis including TMTV, iPET2, and individual risk factors (model A), EORTC classification (model B), GHSG (model C), or NCCN (model D), only TMTV and iPET2 retained statistical significance for both PFS and OS (Table 3). On a stratified Cox model, no effect of the treatment arm was observed.

These 2 independent parameters TMTV and iPET2 were combined. TMTV stratified iPET2 response patients in 4 risk categories ($P < .0001$ for PFS and OS; Figure 2), with an increased percentage of PFS events (from 4%, 18%, 38%, and 75%), as well as OS events (2%, 5%, 8%, and 62%). It identifies in the negative iPET2 group ($n = 237$, 92%) a subset of patients with a poorer prognosis ($P = .0009$ [HR, 4.6] for PFS and $P = .2$ for OS): Patients with a negative iPET2 and a baseline high TMTV had a 5-year PFS of 82% vs 95% for patients with low TMTV and a 5-year OS of 95% vs 98%. In patients with positive iPET2 ($n = 21$), a high TMTV significantly ($P = .026$ [HR, 3.4] for PFS and $P = .002$ [HR, 12.9] for OS; Figure 2) individualized a subgroup of patients with a dismal outcome ($n = 8$) with 5-year PFS of 25% vs 50% and 5-year OS of 50% vs 92% for positive iPET2 with a small volume.

Moreover, the patients with positive iPET2 with a small TMTV, despite a shorter PFS (5-year PFS of 50%), retrieved a similar OS

to the iPET2-negative groups (5-year OS of 92% vs 95% and 98% for the 2 iPET2-negative groups), Figure 2.

Discussion

This study clearly shows the independent prognostic value of baseline TMTV in patients with early-stage HL, homogeneously treated with combined modality treatment. A high TMTV discriminated high-risk patients in the whole group. The presence of a small TMTV reclassified more than 70% of EORTC/GELA U patients to a low-risk group. In multivariable analysis, including early PET response, TMTV was the only baseline prognosticator compared with the current staging systems proposed by EORTC/GELA, GHSG, or NCCN groups. Moreover, this baseline parameter stratifies negative and positive iPET2 patients in 2 different risk groups.

The TMTV values found in this study are in accordance to what is expected in early-stage HL from the data already published. Using the same method, Kanoun and colleagues,¹⁵ in a mixed population of HL with 37% early-stage, reported a median TMTV equal to 117 cm³; Casasnovas et al,²⁶ in advanced stage, reported a median TMTV of 200 cm³; and recently, Moskowitz et al,²⁷ in relapsed/refractory patients, reported a median TMTV of 50 cm³. The median TMTV value reported by Song et al¹⁴ in early-stage HL is higher than our median (142.6 vs 67 cm³), which is explained by the difference in methods. As previously shown, both methods could be prognostic.²⁰ The threshold defined in the present study for early-stage disease, 147 cm³, is reliable, as supported by the results of the training validation methods. Indeed, the same threshold was found in the training set and validated in the validation set. Moreover, the 41% SUVmax method used in this study showed good interobserver agreement.²² Relative methods

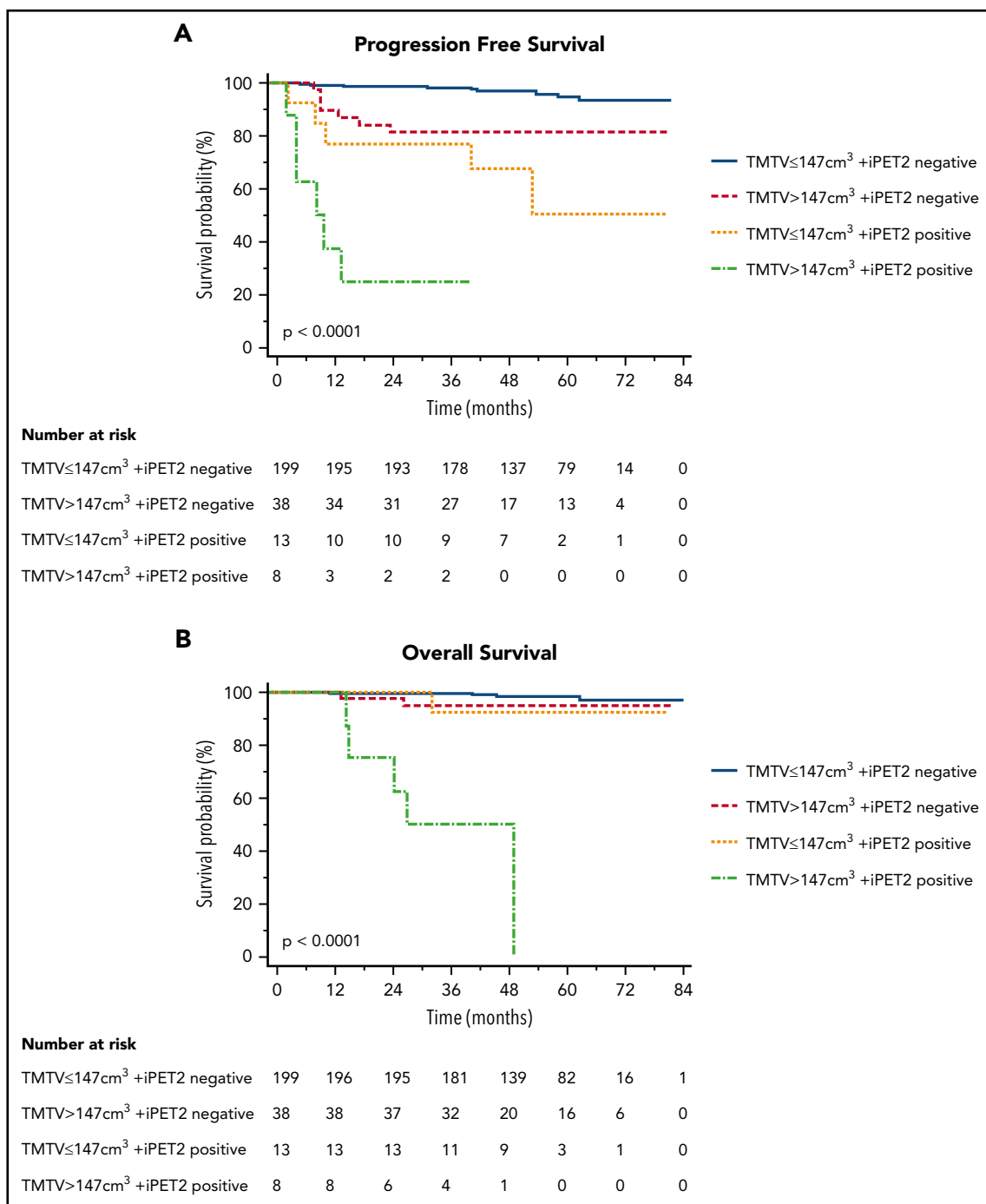


Figure 2. PFS and OS according to total metabolic tumor volume (TMTV $> 147 \text{ cm}^3$ or $\leq 147 \text{ cm}^3$) and early PET response after 2 cycles (negative iPET2 for DS1-3, positive iPET2 for DS4-5).

of TMTV measurement are not or almost not influenced by the variations of SUV values because of technical parameters. The 41% SUVmax method has been used to show the prognostic value of TMTV in different subtypes of lymphoma,^{9,11,13,15,27,28} and recently to measure drug delivery.²⁹ Even if the processing time is short for early-stage patients, it can be long in diffuse disease, but automatic methods of volume determination are under development to allow a routine use for all stages.

Most of the parameters included in the risk assessment systems currently available for stratifying early stage HL are variably

correlated to disease extent, number, and size of the area involved. Indeed, they are indirect and inaccurate surrogates for TB. For instance, the measure of tumor bulk, first evaluated by the M/T ratio on chest radiographs and limited to the mediastinum, has been in some classification replaced by the size of the largest mediastinal mass on CT scan, in an axial plane. The use of the coronal plane has also been recently proposed.³⁰ The first demonstration of the strong prognostic value of total TB came from Specht et al,⁶ with an interesting attempt to estimate total tumor volume, based on the categorization of lesion size (by physical examination) and mediastinal and hilar involvement (from

chest X-rays), and then by adding together the grades of all involved sites. Even if complex and observer-dependent, the approximate computation of total TB was superior to all other known prognostic factors. The superiority of total TB over every other single prognostic factor and composite prognostic score was confirmed 10 years later by Gobbi et al,³¹ through the evaluation of the TB on CT scan. However, it is well known that in HL, the neoplastic component resides in the heterogeneous admixture of nonneoplastic inflammatory and accessory cells with less than 1% to 2% of Reed-Sternberg cells. Therefore, PET/CT could be more appropriate for estimating TB by providing evaluation of the functionally active volume of the tumor, rather than the whole visible mass of tumor tissue, as with CT scan.⁷ The functional activity would better reflect the immunological disorder; that is, the infiltrating microenvironment cells. Quantification of the activity of this crucial component appeared therefore better than the simple volumetric measure. Indeed, the current exploratory study illustrates that TMTV is the only baseline prognosticator in multivariable analysis when tested with iPET2. Patients with a high TMTV had, respectively, 5 and 7 times more risk of experiencing a disease relapse or progression or to die than patients with a low TMTV. Therefore, at the condition that an external validation further confirms our results, TMTV could be proposed instead of the other current staging systems to select U patients.

Because of its predictive value, interim PET has been used in several recent trials including patients with early-stage HL to guide therapeutic strategy.^{17,32} The data of the RAPID³² and H10¹⁷ trials showed that even if patients with negative PET had a very good prognosis, there is still a small proportion of treatment failures in this group either with combined modality treatment (3-year PFS of 94.6% in the RAPID trial; 5-year PFS of 92.1% in the H10 U group and 99% in the F group) or, slightly higher, after chemotherapy alone (90.8% in the RAPID trial; 89.6% and 87% for H10 U and F groups, respectively) that need to be identified by other factors.

In that way, the prognostic role of microenvironment cells has been recently highlighted by Agostinelli et al³³ in a series of 208 patients with HL treated with ABVD, including 61% in stages 1 and 2. The expression of CD68 and PD1 in microenvironment cells, and STAT1 negativity in Hodgkin Reed-Sternberg cells, identified a subset of iPET2-negative patients with a 3-year PFS significantly lower than that of the remaining iPET2-negative population, at 64% vs 95%.

In our study, the baseline TMTV appears as a new tool to better stratify early PET response. Low-risk patients are identified by a small volume and a negative iPET2. Their treatment modalities could be discussed. In addition, the observation of a significantly reduced outcome of iPET2-negative patients with a large volume could require considering different treatment approaches including BEACOPPesc (escalated bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone). In contrast, TMTV also significantly stratified 2 small subsets of patients with positive iPET2. Patients with a large volume had a very dismal outcome contrasting with patients with a small volume, who, despite a high risk for PFS, had a similar OS as the negative iPET2 groups, suggesting they have benefited from second-line treatment. These patients could be those who benefit from the BEACOPP escalation proposed for PET-positive patients.^{17,34} Instead, large-volume iPET2-positive patients do not seem to have benefited from second-line treatment, and might require early innovative therapeutic approaches. To get more information,

we plan to reanalyze the experimental arms of the H10 study to further investigate the prognostic value of volume in iPET2-negative patients who did not receive radiotherapy, and in those patients who have been escalated to BEACOPP on the basis of positive PET. Indeed, although limited by the small number of patients included in some of the risk groups individualized by TMTV, these data suggest that interim PET response should be discussed in the light of the initial TB. The role of baseline TMTV in improving the predictive value of PET response assessment has already recently been demonstrated in relapse/refractory HL.²⁷

Although the proposed model combining TMTV and iPET2 deserves to be validated in another independent data set, the results of the present study point out the outstanding prognostic value of TMTV, an imaging biomarker available at diagnosis, measurable in early-stage HL and superior to the clinical and biological parameters already used. As a consequence, baseline TMTV should be taken into account for risk assessment in patients with early-stage HL, avoiding assigning in the same group patients with different levels of volume. The combination of TMTV and PET/CT response after 2 cycles assessed with Deauville score improves the predictive value of interim PET and, if these data are confirmed, may help design new response-adapted therapeutic strategies.

Acknowledgments

Research support provided by the European Organisation for Research and Treatment of Cancer (Belgium), La Ligue Nationale Contre le Cancer from France through the EORTC Cancer Research Fund, Lymphoma Study Association (France), Fondazione Italiana Linfomi (Italy), Fondation Belge contre le Cancer (Belgium), Dutch Cancer Society (The Netherlands), Institut National du Cancer (France), Assistance Publique-Hôpitaux de Paris (France), Societe Française de Medecine Nucleaire et Imagerie Moleculaire (France), Associazione Angela Serra (Italy), van Vlissingen Lymfom Fonds (The Netherlands), and Chugai Pharmaceutical (Japan).

Authorship

Contribution: A.-S.C., A.V., O.C., J.R., M.F., M.H., M.A., and M.M. conceived and designed the study; A.-S.C., A.V., O.C., M.B., R.R., A.C., P.B., J.R., B.D., C.F., T.R., E.V.Z., M.F., U.R., M.A., and M.M. collected and assembled the data; A.-S.C., A.V., A.L., O.C., S.B., J.R., B.D., C.F., L.C., T.V.B., M.F., M.H., M.A., and M.M. analyzed and interpreted the data; and all authors contributed to writing the manuscript and approved the final version.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Anne-Ségolène Cottreau, Nuclear Medicine Department, Hôpital Cochin, 27 rue du Faubourg Saint Jacques, 75014 Paris, France; e-mail: annesegolene.cottreau@aphp.fr.

Footnotes

Submitted 9 July 2017; accepted 23 January 2018. Prepublished online as *Blood* First Edition paper, 1 February 2018; DOI 10.1182/blood-2017-07-795476.

Presented orally at the 58th annual meeting of the American Society of Hematology, San Diego, CA, 3-6 December 2016.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

REFERENCES

- Eichenauer DA, Engert A, André M, et al; ESMO Guidelines Working Group. Hodgkin's lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2014;25(Suppl 3):iii70-iii75.
- Fermé C, Eghbali H, Meerwaldt JH, et al; EORTC-GELA H8 Trial. Chemotherapy plus involved-field radiation in early-stage Hodgkin's disease. *N Engl J Med*. 2007; 357(19):1916-1927.
- Sieber M, Engert A, Diehl V. Treatment of Hodgkin's disease: results and current concepts of the German Hodgkin's Lymphoma Study Group. *Ann Oncol*. 2000;11(Suppl 1): 81-85.
- Hoppe RT, Advani RH, Ai WZ, et al; National comprehensive cancer network. Hodgkin lymphoma, version 2.2015. *J Natl Compr Canc Netw*. 2015;13(5):554-586.
- 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(25):2419-2429.
- Specht L, Nordentoft AM, Cold S, Clausen NT, Nissen NI. Tumor burden as the most important prognostic factor in early stage Hodgkin's disease. Relations to other prognostic factors and implications for choice of treatment. *Cancer*. 1988;61(8):1719-1727.
- Gobbi PG. Tumor burden in Hodgkin's lymphoma: much more than the best prognostic factor. *Crit Rev Oncol Hematol*. 2014;90(1): 17-23.
- Klimm B, Goergen H, Fuchs M, et al. Impact of risk factors on outcomes in early-stage Hodgkin's lymphoma: an analysis of international staging definitions. *Ann Oncol*. 2013;24(12):3070-3076.
- 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.
- Mikhaeel NG, Smith D, Dunn JT, et al. Combination of baseline metabolic tumour volume and early response on PET/CT improves progression-free survival prediction in DLBCL. *Eur J Nucl Med Mol Imaging*. 2016; 43(7):1209-1219.
- Cottreau AS, Becker S, Broussais F, et al. Prognostic value of baseline total metabolic tumor volume (TMTV0) measured on FDG-PET/CT in patients with peripheral T-cell lymphoma (PTCL). *Ann Oncol*. 2016;27(4): 719-724.
- Cottreau AS, Lanic H, Mareschal S, et al. Molecular profile and FDG-PET/CT total metabolic tumor volume improve risk classification at diagnosis for patients with diffuse large B-cell lymphoma. *Clin Cancer Res*. 2016; 22(15):3801-3809.
- Meignan M, Cottreau AS, Versari A, et al. Baseline metabolic tumor volume predicts outcome in high-tumor-burden follicular lymphoma: a pooled analysis of three multicenter studies. *J Clin Oncol*. 2016;34(30): 3618-3626.
- Song MK, Chung JS, Lee JJ, et al. Metabolic tumor volume by positron emission tomography/computed tomography as a clinical parameter to determine therapeutic modality for early stage Hodgkin's lymphoma. *Cancer Sci*. 2013;104(12):1656-1661.
- Kanoun S, Rossi C, Berriolo-Riedinger A, et al. Baseline metabolic tumour volume is an independent prognostic factor in Hodgkin lymphoma. *Eur J Nucl Med Mol Imaging*. 2014;41(9):1735-1743.
- Raemaekers JM, André MP, Federico M, et al. Omitting radiotherapy in early positron emission tomography-negative stage I/II Hodgkin lymphoma is associated with an increased risk of early relapse: Clinical results of the preplanned interim analysis of the randomized EORTC/LYSA/FIL H10 trial. *J Clin Oncol*. 2014;32(12):1188-1194.
- André MPE, Girinsky T, Federico M, et al. Early positron emission tomography response-adapted treatment in stage i and ii Hodgkin lymphoma: final results of the randomized EORTC/LYSA/FIL H10 trial. *J Clin Oncol*. 2017;35(16):1786-1794.
- Meignan M, Itti E, Bardet S, et al. Development and application of a real-time on-line blinded independent central review of interim PET scans to determine treatment allocation in lymphoma trials. *J Clin Oncol*. 2009;27(16):2739-2741.
- Cheson BD, Pfistner B, Juweid ME, et al; International Harmonization Project on Lymphoma. Revised response criteria for malignant lymphoma. *J Clin Oncol*. 2007; 25(5):579-586.
- Kanoun S, Tal I, Berriolo-Riedinger A, et al. Influence of software tool and methodological aspects of total metabolic tumor volume calculation on baseline [18F]FDG PET to predict survival in Hodgkin lymphoma. *PLoS One*. 2015;10(10):e0140830.
- Boellaard R, O'Doherty MJ, Weber WA, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imaging*. 2010;37(1): 181-200.
- Meignan M, Sasaneli M, Casasnovas RO, et al. Metabolic tumour volumes measured at staging in lymphoma: methodological evaluation on phantom experiments and patients. *Eur J Nucl Med Mol Imaging*. 2014;41(6): 1113-1122.
- Meignan M, Gallamini A, Meignan M, Gallamini A, Haioun C. Report on the First International Workshop on Interim-PET-Scan in Lymphoma. *Leuk Lymphoma*. 2009;50(8): 1257-1260.
- 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.
- Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res*. 2004;10(21): 7252-7259.
- Casasnovas RO, Kanoun S, Tal I, Cottreau AS, et al. Baseline total metabolic volume (TMTV) to predict the outcome of patients with advanced Hodgkin lymphoma (HL) enrolled in the AHL2011 LYSA trial [abstract]. *J Clin Oncol* 2016;34(15 suppl). Abstract 7509.
- 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.
- Meignan M. Quantitative FDG-PET: a new biomarker in PMBCL. *Blood*. 2015;126(8): 924-926.
- Tout M, Casasnovas O, Meignan M, et al. Rituximab exposure is influenced by baseline metabolic tumor volume and predicts outcome of DLBCL patients: a Lymphoma Study Association report. *Blood*. 2017;129(19): 2616-2623.
- Kumar A, Burger IA, Zhang Z, et al. Definition of bulky disease in early stage Hodgkin lymphoma in computed tomography era: prognostic significance of measurements in the coronal and transverse planes. *Haematologica*. 2016;101(10):1237-1243.
- Gobbi PG, Ghirardelli ML, Solcia M, et al. Image-aided estimate of tumor burden in Hodgkin's disease: evidence of its primary prognostic importance. *J Clin Oncol*. 2001; 19(5):1388-1394.
- Radford J, Illidge T, Counsell N, et al. Results of a trial of PET-directed therapy for early-stage Hodgkin's lymphoma. *N Engl J Med*. 2015;372(17):1598-1607.
- Agostinelli C, Gallamini A, Stracqualursi L, et al. The combined role of biomarkers and interim PET scan in prediction of treatment outcome in classical Hodgkin's lymphoma: a retrospective, European, multicentre cohort study. *Lancet Haematol*. 2016;3(10): e467-e479.
- Trotman JFA, Federico M, Stevens L, et al. Response-adjusted therapy for advanced Hodgkin lymphoma (RATHL trial): longer follow up confirms efficacy of de-escalation after a negative interim PET scan (CRUK/07/033). *Hematol Oncol*. 2017;35(Suppl S2):65-67.