Blood First Edition Paper, prepublished online February 1, 2018; DOI 10.1182/blood-2017-07-795476

Title page

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

Running head

Baseline metabolic tumor volume and early response in HL

Authors

Anne Ségolène Cottereau[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], Loic Chartier[13], Thierry Vander Borght[14], Massimo Federico[5] , Martin Hutchings[15] , Umberto Ricardi[16], Marc Andre[17], Michel Meignan[18].

1Nuclear Medicine, Hôpital Cochin, Paris, France;

2Santa Maria Nuova Hospital-IRCCS Reggio Emilia, Italy;

3Dep. Of Clinical Physiology, Nuclear Medicine, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark;

4Hematology department, and INSERM 1231, University Hospital, Dijon, France;

5Department of Diagnostic, Clinical and Public Health Medicine, University of Modena and Reggio Emilia, Modena, Italy;

6LYSARC LYSA IM Hôpital Henri Mondor, Créteil, France;

7Department of Nuclear Medicine, CLCC François Baclesse, Caen, France;

8Nuclear Medicine Unit, Ospedale Santo Stefano, Prato, Italy;

9Hematology, Hopital Saint-Louis, APHP, Paris, France;

10Hematology, Radboud University Medical Center, Nijmegen, Netherlands;

11 Hematology, Hôpital Cochin, Paris, France.

12European Organisation for Research and Treatment of Cancer, Brussels, Belgium;

13 Biostatistical department, LYSARC, Centre Hospitalier Lyon Sud, Lyon, France

14 Nuclear Medicine Division, CHU UCL NAMUR, Yvoir, Belgium;

15 Department of Hematology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark;

16 Department of Oncology, University of Turin, Torino, Italy;

17 Hematology Department, CHU UCL NAMUR, Yvoir, Belgium

18 LYSA Imaging, Creteil, France.

This study has been partly presented orally during the ASH 2016 meeting

First Author

Anne-Ségolène Cottereau

Nuclear Medicine Department, Hôpital Cochin

27 rue du Faubourg Saint Jacques, Paris, France

Mail: annesegolene.cottereau@aphp.fr

Tel: +33 1 56 01 70 00

word counts for text : 3818

word counts for abstract: 242

figure/table count : 2 + 3

reference count: 34

2 Key Points

Baseline metabolic tumor volume is a strong prognostic factor in early stage HL.

Baseline metabolic tumor volume impacts the early response to treatment and combined with early PET improves risk stratification.

ABSTRACT

We tested baseline PET/CT as a measure of total tumor burden in order to better identify high risk patients in early-stage Hodgkin's lymphoma (HL). Stage I-II HL patients enrolled in the standard arm (combined modality treatment) of the H10 trial (NCT00433433) with available baseline PET and interim PET (iPET2) after two cycles of doxorubicine, bleomycin, vinblastine, 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. The prognostic value of TMTV was evaluated and compared to 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 PFS and 12 OS events. TMTV was prognosticator of PFS ($p<0.0001$) and OS ($p=0.0001$) with an 86% and 84% specificity respectively. The 5y-PFS and OS were 71% and 83% in the high TMTV (>147cm^3) group (n=46) vs. 92% and 98% in the low TMTV group (≤147cm^3). In multivariable analysis including iPET2, TMTV was the only baseline prognosticator compared to the current staging systems proposed by EORTC/GELA, GHSG, or NCCN groups. TMTV and iPET2 were independently prognostic and combined identified four risk groups: low (TMTV≤147+DS1-3; 5y-PFS 95%), low-intermediate (TMTV>147+DS1-3; 5y-PFS 81.6%), high-intermediate (TMTV≤147+DS4-5; 5y-PFS 50%) and high (TMTV>147+DS4-5; 5y-PFS 25%). TMTV improves baseline risk stratification of early stage HL patients compared to current staging systems and the predictive value of early PET response as well.

INTRODUCTION

In early stage Hodgkin's 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. Consequently the definition of the unfavorable 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 in advanced HL treatment escalation based on early PET response, also included stage IIA with adverse features considered as early unfavorable stage for the EORTC. Therefore, improving the ability of identifying 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 and aimed to give an indirect, fractional, appraisal of it. The first attempt to approach the total tumor burden 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 and recently total tumor burden measurement with CT has been proposed[7]. However in a retrospective analysis of 1173 early stage HL treated homogenously in the HD10 and HD11 trials[8], the GHSG showed that the best risk models included not only a large tumor burden but also a systemic inflammation assessed by an elevated ESR. Indeed, clinical and pathological features of HL depend on the interaction between tumor and micro environmental cells which maintains an intense inflammatory reaction. Today, unlike CT, functional imaging using 18F-FDG 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, helpful to stratify 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 (FIL)). TMTV was compared to clinic-biological prognostic factors used in the different classification systems and to 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 two 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 (CMT) 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, age 15 to 70 years. Both favorable (F) and unfavorable (U) patients according to EORTC/LYSA criteria entered[2] (U: at least one of the following criteria: age ≥50 years or ≥ 4 nodal areas or mediastinal-thoracic-ratio ≥ 0.35 or no B symptoms and erythrocyte sedimentation rate [ESR] ≥ 50 or B symptoms and ESR ≥30, F: all others). The standard combined modality treatment (CMT) consisted in doxorubicin, bleomycin, vinblastine and dacarbazine (ABVD) chemotherapy, 3 cycles for F or 4 cycles for U, followed by 30Gy involved-node radiotherapy (INRT)[17]. All these patients had an early PET after two cycles of ABVD (iPET2) with no impact on therapy. Baseline PET was recommended but not mandatory. Only 0.2% of patients of 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 GHSG [3] (high risk group includes at least one of the following factors: mediastinal mass ratio >0.33; B symptoms with ESR ≥30; No B symptoms with ESR ≥50, ≥ 3 lymph node sites involved, extranodal involvement) and NCCN scoring system (high risk group includes at least one of the following factors: bulky mediastinal

disease (mediastinal mass ratio >0.33) or bulky disease greater than 10 cm, B symptoms, ESR ≥ 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 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 (DICOM) format, were collected for functional parameter measurements. Analysis of imaging data was performed by three nuclear medicine physicians (ASC, AV, AL) blinded to patient outcome who analyzed each a randomized third of the population. TMTV was computed on semiautomatic user's free software Beth Israel Fiji[20], which can be uploaded at 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 ($TMTV = \Sigma MTV_L$). The 41% maximum SUV (SUVmax) threshold method was used for MTV_L computation, as recommended by the European Association of Nuclear Medicine[21] 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 SUVmax, we first drew a VOI engulfing this bulky region. If the volume determined on the basis of the SUVmax of the whole region had left out nodes with SUV lower than 41%, additional VOI 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. PET after two cycles of ABVD (iPET2), initially prospectively scored according to International Harmonization Project criteria in the H10 trial, was re-analyzed based on the Deauville 5-point scale (5-DS)[23,24], with score 4-5 for positivity (FDG uptake higher than the liver).

Statistical

Two different approaches, X-tile analysis[25] and receiver operating characteristic (ROC) 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 by using the log-rank test. Characteristics of populations were compared by using Chi2, Fisher or Mann-Whitney tests. A backward stepwise Cox model with all significant baseline univariate predictors and iPET2 was performed. Because EORTC, GHSG and NCCN are correlated to one another, 4 separate models were analyzed testing TMTV, iPET2 and A) all the individual factors with $p<0.15$ in univariate analysis B) EORTC C) GHSG D) NCCN classifications. Independent variables were combined for survival prediction. A stratified Cox model was used to account for differences between the two treatment regimens (3 or 4 cycles). Differences between results of comparative tests were considered significant if the two-sided P value was <0.05. 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 for ASC/AV, ASC/AL and AL/AV. 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

Out 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 to compute quantitative parameters were excluded. A total of 258 patients (101 F and 157 U) were suitable for TMTV calculation and 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 three 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 favorable group, the 5-

years PFS and OS were 95% and 100% respectively versus 84% and 92% in the unfavorable group.

Baseline PET parameters

Median TMTV was 67 cm^3 (interquartile range [IQR], 32 to 114 cm^3). A significant difference was observed between the U and the F groups, with a median TMTV of 87 cm^3 contrasting with 48 cm^3 ($p<0.0001$). 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 group was also excellent for ASC/AV, $\kappa=1$, ASC/AL $\kappa=0.9$, AV/AL $\kappa=0.9$.

From the results of X-tile and the training validation procedure the best TMTV cut-off was 147 cm^3 for PFS and OS (Supplemental 3). The presence of a high TMTV (> 147 cm^3) was significantly associated with a shorter PFS and OS ($p<0.0001$, HR=5.2 and $p=0.0001$, 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% versus 92% and 98% for patients with a lower TMTV (figure 1). The prognostic value of TMTV was not impacted 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 end of the 3-year period without a prior event, it was 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 cut off value for PFS and OS prediction could be found**.**

Individual baseline clinico-biological factors and EORTC, GHSG and NCCN staging systems (Table 2)

Age was not associated with either PFS or OS, whereas the presence of B symptoms and ≥4 lymph nodes sites involved were prognostic for both PFS and OS. M/T ratio≥0.35 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=0.075) (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 IIB patients with M/T>0.33 or extranodal disease which would have been included in advanced stage by GHSG has a significant worse PFS and OS (p=0.0002 and p=0.015 respectively)

In a sub analysis of EORTC unfavorable patients, TMTV maintained its prognostic significance for both PFS and OS (p=0.0001 HR= 4.2 and p=0.0035 HR=4.0 respectively). Patients with a small volume (74%), despite belonging to the U group, had a 5y 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). IPET2 was predictive of PFS and OS (p<0.0001 HR=12 and p<0.0001 HR=13.6 respectively). Positive iPET2 patients (n=21) had a 5year-PFS and OS of 38% and 68% versus 92.6% and 98% respectively. The frequency of iPET2 positivity was significantly higher in patients with high TMTV (Table 1) but 62% of positive iPET2 patients had low TMTV and 83% of high TMTV patients had a negative iPET2.

Multivariable analysis including baseline parameters and iPET2 (Table 3)

On multivariable analysis including TMTV, iPET2 and individual risk factors (model A) or EORTC classification (model B) or 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 impact of the treatment arm was observed.

These two independent parameters TMTV and iPET2 were combined. TMTV stratified iPET2 response patients in 4 risk categories (p<0.0001 for PFS and OS, Figure 2), with an increased percentage of PFS events (from 4%, 18%, 38%, 75%) as well as OS events (2%, 5%, 8%, 62%). It identifies in the negative iPET2 group (n=237, 92%) a subset of patients with a poorer

prognosis (p = 0.0009, HR=4.6 for PFS and p = 0.2 for OS): patients with a negative iPET2 and a baseline high TMTV had a 5y-PFS of 82% vs 95% for low TMTV patients and a 5y-OS of 95% vs 98%. In positive iPET2 patients (n=21), a high TMTV significantly (p = 0.026, HR=3.4 for PFS and p=0.002, HR=12.9 for OS, Figure 2) individualized a subgroup of patients with a dismal outcome (n=8) with 5-y PFS of 25% vs. 50% and 5y-OS of 50% vs92% for positive iPET2 with a small volume.

Moreover, the positive iPET2 patients with a small TMTV, despite a shorter PFS (5y PFS of 50%), retrieved a similar OS to the iPET2 negative groups (5y-OS of 92% vs 95% and 98% for the two iPET2 negative groups), Figure 2.

DISCUSSION

This study clearly shows the independent prognostic value of baseline TMTV in early stage HL patients, homogenously 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 unfavorable patients to a low risk group. In multivariable analysis, including early PET response, TMTV was the only baseline prognosticator compared to the current staging systems proposed by EORTC/GELA, GHSG, or NCCN groups. Moreover, this baseline parameter stratifies negative and positive iPET2 patients in two 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% of early stage reported a median TMTV equal to 117 cm^3, Casasnovas [26] in advanced stage a median of 200 cm^3 and recently Moskowicz[27] in relapsed /refractory patients a median of 50 cm^3. The median TMTV value reported by Song[14] in early stage HL is higher than our median (142.6 vs 67cm^3) 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 cm^3, 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 a good interobserver agreement[22]. Relative methods of TMTV measurement are not or almost not

influenced by the variations of SUV values due to 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, size of the area involved. Indeed, they are indirect and inaccurate surrogates for tumor burden. 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 axial plane. The use of the coronal plane has also been recently proposed[30]. The first demonstration of the strong prognostic value of total tumor burden (TB) came from Specht[6] with an interesting attempt to estimate Total tumor volume, based on the categorization of lesions size (by physical examination), 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 resulted as 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 heterogeneous admixture of non-neoplastic inflammatory and accessory cells with less than 1–2% of Reed–Sternberg cells. Therefore, PET/CT could be more appropriate to estimate the TB, by providing evaluation of the functionally active volume of the tumor rather than the whole visible mass of tumor tissue with CT scan[7]. The functional activity would better reflect the immunological disorder, i.e 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 five and seven times more risk to experience 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 unfavorable patients.

Due to its predictive value interim PET has been used in several recent trials including early stage HL patients to guide therapeutic strategy[17,32]. The data of the RAPID[32] and H10 [17]trials, showed that even if negative PET patients 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 unfavorable group; and 99% in favorable group) or, slightly higher, after chemotherapy alone (90.8% in RAPID trial; 89.6% and 87% for H10) who need to be identified by other factors.

In that way, the prognostic role of microenvironment cells has been recently highlighted by Agostinelli and colleagues[33] in a series of 208 HL patients treated with ABVD, including 61% in stage 1 and 2. The expression of CD68 and PD1 in micro environment 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, 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 significant reduced outcome of iPET2 negative patients with a large volume could require considering a different treatment approaches including BEACOPPesc. On the other hand, TMTV also significantly stratified two small subsets of positive iPET2 patients. Patients with a large volume had a very dismal outcome contrasting with patient with a small volume, who despite a high risk of PFS, had a similar OS to the negative iPET2 groups, suggesting that 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 don't seem to have benefited from second line treatment, and might require early innovative therapeutic approaches. In order to get more information, we plan to re-analyse the experimental arms of the H10 study to further investigate the prognostic value of volume in iPET2 negative patients who did not received 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 tumor burden. The role of baseline TMTV to improve the

predictive value of PET response assessment has already recently been demonstrated in relapse/ refractory HL[27].

While the proposed model combining TMTV and iPET2 deserves to be validated in another independent data set, the results of the present study points 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. Consequently, baseline TMTV should be taken into account for risk assessment in early stage HL patients avoiding assigning in the same group, patients with different levels of volume. The combination of TMTV and PET/CT response after two cycles assess with Deauville score improves the predictive value of interim PET and, if these data are confirmed, may help to design new response-adapted therapeutic strategies.

Acknowledgments of research support : 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 des Hopitaux de Paris (France), Societe Française de Medecine Nucleaire et Imagerie Moleculaire (France), Associazione Angela Serra (Italy), van Vlissingen Lymfoom Fonds (the Netherlands), and Chugai Pharmaceutical (Japan).

Authorship Contributions

Conception and design: Anne-Ségolène Cottereau, Michel Meignan, Annibale Versari, Marc André, John Raemaekers, Massimo Federico, René Olivier Casasnovas, Martin Hutchings

Collection and assembly of data: Michel Meignan, Anne-Ségolène Cottereau, Annibale Versari ,Marc André, John Raemakers, Massimo Federico, René Olivier Casasnovas, Monica Bellei, Romain Ricci, Catherine Fortpied, Tiana Ravelohivary, Antonio Castagnoli, Pauline Brice, Umberto Ricardi, Bénédicte Deau, Emelie Van Zele

Data analysis and interpretation: Anne-Ségolène Cottereau, Michel Meignan, Annibale Versari, Annika Loft, Stéphane Bardet, Thierry Van der Borght, Loïc Chartier, Catherine Fortpied, Marc André, René Olivier Casasnovas, Massimo Federico, Bénédicte Deau, John Raemakers, Martin Hutchings.

Manuscript writing: All authors

Final approval of manuscript: All authors

Conflict-of-interest disclosure:

The authors declare no conflict of interest for the present study.

REFERENCES

1. Eichenauer DA, Engert A, Andre M, et al. Hodgkin's lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2014;25 Suppl 3:iii70-75.
2. Ferme C, Eghbali H, Meerwaldt JH, et al. Chemotherapy plus involved-field radiation in early-stage Hodgkin's disease. *N Engl J Med*. 2007;357(19):1916-1927.
3. 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.
4. Hoppe RT, Advani RH, Ai WZ, et al. Hodgkin lymphoma, version 2.2015. *J Natl Compr Canc Netw*. 2015;13(5):554-586.
5. 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.
6. 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.
7. Gobbi PG. Tumor burden in Hodgkin's lymphoma: much more than the best prognostic factor. *Crit Rev Oncol Hematol*. 2014;90(1):17-23.
8. 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.
9. 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.
10. 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.
11. Cottereau 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.
12. Cottereau 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.
13. Meignan M, Cottereau 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.
14. 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.
15. 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.
16. Raemaekers JM, Andre 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.
17. Andre MP, 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:JCO2016686394.

18. 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.
19. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol*. 2007;25(5):579-586.
20. 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.
21. 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.
22. Meignan M, Sasanelli 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.
23. 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.
24. 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.
25. 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.
26. Casasnova RO KS, Tal I, Cottereau AS, Edeline V, Brice P, Bouabdallah R, Salles G, Stamatoullas A, Dupuis J, Reman T, Gastinne T, Joly B, Bouabdallah K, Nicolas-Virelizier E, Andre M, Mounier N, Ferme C, Meignan M, Berriolo-Riedinger A; . Baseline total metabolic volume (TMTV) to predict the outcome of patients with advanced Hodgkin lymphoma (HL) enrolled in the AHL2011 LYSA trial. *J Clin Oncol* 2016;34, 2016 (suppl; abstr 7509).
27. Moskowitz AJ, Schoder H, Gavane S, et al. Prognostic significance of baseline metabolic tumor volume in relapsed and refractory Hodgkin lymphoma. *Blood*. 2017.
28. Meignan M. Quantitative FDG-PET: a new biomarker in PMBCL. *Blood*. 2015;126(8):924-926.
29. 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.
30. 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.
31. 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.
32. 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.
33. 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.
34. Trotman J FA, Federico M, Stevens L, Kirkwood A, Clifton-Hadley L, Patrick P, Berkahn L, D'Amore F, Enblad G, Luminari S, Radford J, Barrington S.F. and Johnson P. Response-Adjusted therapy for advanced Hodgkin lymphoma (RATHL trial): longuer follow up confirms efficacy of de-escalation after a negatve interim PET scan (CRUK/07/033). *Hematological Oncology*. 2017;35(Issue S2, June 2017):65–67.

Tables

Table 1. Patient characteristics

Characteristics	Total Population n=258	Low tumor burden TMTV ≤ 147 cm^3 n=212	High tumor burden TMTV > 147 cm^3 n=46	P value
Median age, (range) (years)	31 (15-71)	32 (15-71)	27 (17-63)	0.009
Age ≥ 50 years (%)	34 (13%)	31 (15%)	3 (7%)	0.22
male, *n* (%)	129 (50%)	104 (49%)	25 (54%)	0.62
Nodular sclerosis histology *n* (%)	207 (80%)*	168 (79%)	37 (80%)	0.68
Ann Arbor Stage II, *n* (%)	198 (77%)	157 (74%)	41 (89%)	0.03
B symptoms, *n* (%)	85 (33%)	60 (28%)	25 (54%)	0.001
median ESR mm/h (IQR)	26	23 (12-50)	49 (26-72)	0.0001
≥ 4 involved sites, *n* (%)	28 (11%)	17 (8%)	11 (24%)	0.004
Bulk mediastinum (M/T ≥ 0.35	62 (24%)	34 (16%)	29 (63%)	< 0.0001
Unfavorable EORTC	157 (61%)	117 (55%)	40 (87%)	0.0001
Unfavorable GSHG	177 (69%)	133 (63%)	44 (96%)	<0.0001
Unfavorable NCCN	164 (64%)	121 (57%)	43 (93%)	<0.0001
Positive iPET2 (DS 4-5)	21 (8%)	13 (6%)	8 (17%)	0.028

*data not available for 2 patients

Table 2. Univariate analysis for baseline prognostic factors of survival

Prognostic factor	Progression-free survival		Overall survival	
	HR (95% CI)	P value	HR (95% CI)	P value
Male	1.3 (0.6-2.7)	0.53	0.7 (0.2-2.3)	0.60
Age ≥ 50 years	1.8 (0.6-5.6)	0.18	2.2 (0.4-11.4)	0.23
mixed cellularity	0.8 (0.3-2.2)	0.68	1.1 (0.2-4.8)	0.92
presence of B symptoms	2.9 (1.3-6.7)	0.0035	4.6 (1.3-15.6)	0.006
≥ 4 involved sites	2.7 (0.7- 9.5)	0.028	4.7 (0.7-31.5)	0.006
Bulk mediastinum (M/T ≥ 0.35)	0.9-5.7	0.026	2.3 (0.6-8.9)	0.13
TMTV >147 cm^3	5.2 (1.8-14.7)	<0.0001	7.2 (1.6-33.4)	0.0001
Unfavorable EORTC	5.7 (2.7-12.3)	0.0013	NR	0.0039
Unfavorable GHSG	2.8 (1.3-6.3)	0.046	5.3 (1.6-17.6)	0.075
Unfavorable NCCN	3.6 (1.7-7.9)	0.011	6.7 (2.1-21.6)	0.034
Positive iPET2 (DS 4-5)	12 (2.3-63.7)	<0.0001	13.2 (1.4-128.3)	<0.0001

Favorable (F), Unfavorable (U).

Table 3. Multivariate analysis testing total metabolic tumor volume (TMTV), with interim PET response after two cycles (iPET2) and individual baseline factors, EORTC, GHSG, NCCN staging systems. *: all variables integrated in the Cox model; final model: with significant factors after performing the backward stepwise Cox model.

	PFS*			PFS (final model)			OS*			OS (final model)		
TMTV tested with:	HR	95% CI	*P*	HR	95% CI	*P*	HR	95% CI	*P*	HR	95% CI	*P*
A.individual factors												
TMTV>147 cm^3	3.9	1.6-9.5	0.0032	4.4	2.0-9.5	0.0002	3.7	0.9-14.6	0.066	5.5	1.7-17.9	0.0043
IPET 2	11.0	4.8-25.1	<0.0001	10.9	4.9-24.4	<0.0001	11.3	3.2-39.9	0.0002	11.1	3.4-36.4	<0.0001
B symptoms	2.1	0.9-4.8	0.076				2.6	0.7-9.5	0.16			
≥ 4 involved sites	2.0	0.8-5.2	0.16				3.4	0.9-12.3	0.065			
M/T ≥ 0.35	0.8	0.3-2.0	0.65				0.6	0.2-2.4	0.51			
B.EORTC												
TMTV>147 cm^3	3.5	1.6-7.8	0.0016	4.4	2.0-9.5	0.0002	3.9	1.2-12.4	0.024	5.5	1.7-17.9	0.0043
IPET2	9.2	4.1-20.6	<0.0001	10.9	4.9-24.4	<0.0001	8.8	2.7-28.7	0.0003	11.1	3.4-36.4	<0.0001
UnfavorableEORTC	3.2	0.9-11.1	0.067				-	-	0.9			
C.GHSG												
TMTV>147 cm^3	4.1	1.8-9.3	0.0006	4.4	2.0-9.5	0.0002	4.8	1.4-16.3	0.0115	5.5	1.7-17.9	0.0043
IPET2	10.6	4.7-23.9	<0.0001	10.9	4.9-24.4	<0.0001	10.4	3.1-34.2	0.0001	11.1	3.4-36.4	<0.0001
UnfavorableGHSG	1.3	0.4-4.0	0.69				2.0	0.2-17.2	0.55			
D.NCCN												
TMTV>147 cm^3	3.7	1.7-8.4	0.00014	4.4	2.0-9.5	0.0002	4.3	1.3-14.6	0.0182	5.5	1.7-17.9	0.0043
IPET2	10.2	4.5-22.8	<0.0001	10.9	4.9-24.4	<0.0001	10.2	3.1-33.4	0.0001	11.1	3.4-36.4	<0.0001
UnfavorableNCCN	1.8	0.6-5.7	0.30				2.9	0.3-25.0	0.34			

Figure legends

Figure 1.

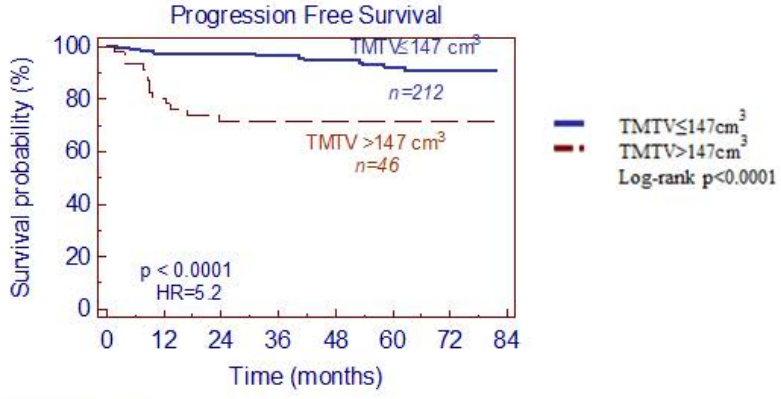
Progression-free survival (PFS) and overall survival (OS) according to high total metabolic tumor volume (TMTV> 147 cm^3) or low TMTV (TMTV≤ 147 cm^3).

Figure 2.

Progression-free survival (PFS) and overall survival (OS) according to total metabolic tumor volume (TMTV>147 cm^3 or ≤ 147) and early PET response after two cycles (negative iPET2 for Deauville score 1-3, positive iPET2 for Deauville score 4-5).

Figure 1

A

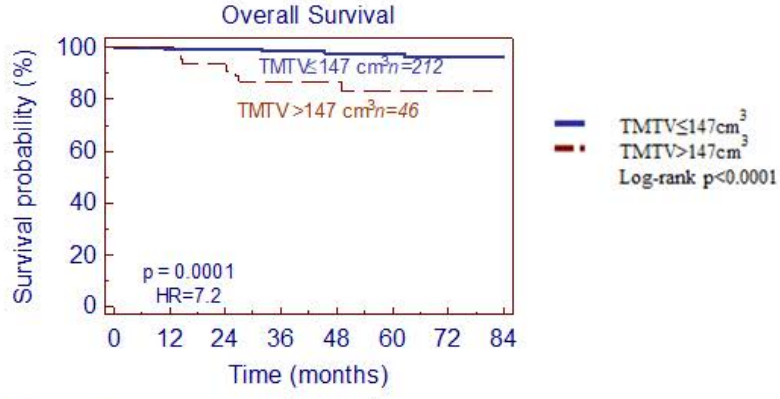


Number at risk

Group: 0	212	205	203	187	144	81	15	0
Group: 1	46	37	33	29	17	13	4	0

	Events	censored	5y PFS (95% CI) months	Log-rank p value
TMTV≤147cm³ (n=212)	14 (7%)	198 (93%)	92 % (90-94)	
TMTV>147cm³ (n=46)	13 (28%)	33 (72%)	71 % (65-77)	< 0.0001

B



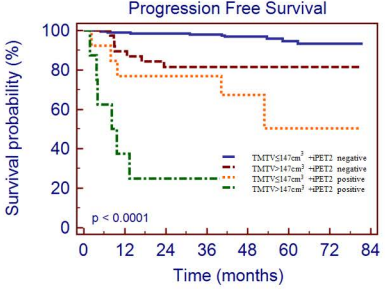
Number at risk

Group: 0	212	209	208	192	148	85	17	1
Group: 1	46	46	43	36	21	16	6	0

	Events	censored	5y PFS (95% CI) months	Log-rank p value
TMTV≤147cm³ (n=212)	5 (2%)	207 (98%)	98 % (97-99)	
TMTV>147cm³ (n=46)	7 (15%)	39 (85%)	83 % (77-89)	< 0.0001

Figure 2

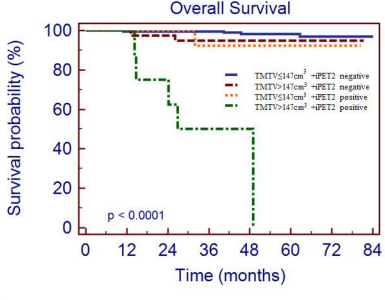
A



Number at risk

TMTV≤147cm³ +iPET2 negative	199	195	193	178	137	79	14	0
TMTV>147cm³ +iPET2 negative	38	34	31	27	17	13	4	0
TMTV≤147cm³ +iPET2 positive	13	10	10	9	7	2	1	0
TMTV>147cm³ +iPET2 positive	8	3	2	2	0	0	0	0

B



Number at risk

TMTV≤147cm³ +iPET2 negative	199	196	195	181	139	82	16	1
TMTV>147cm³ +iPET2 negative	38	38	37	32	20	16	6	0
TMTV≤147cm³ +iPET2 positive	13	13	13	11	9	3	1	0
TMTV>147cm³ +iPET2 positive	8	8	6	4	1	0	0	0

Prepublished online February 1, 2018;
doi:10.1182/blood-2017-07-795476

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

Anne Ségolène Cottereau, Annibale Versari, Annika Loft, Olivier Casasnovas, Monica Bellei, Romain Ricci, Stéphane Bardet, Antonio Castagnoli, Pauline Brice, John Raemaekers, Bénédicte Deau, Catherine Fortpied, Tiana Raveloarivahy, Emelie Van Zele, Loic Chartier, Thierry Vander Borght, Massimo Federico, Martin Hutchings, Umberto Ricardi, Marc Andre and Michel Meignan