

---

---

# Baseline Total Metabolic Tumor Volume Measured with Fixed or Different Adaptive Thresholding Methods Equally Predicts Outcome in Peripheral T Cell Lymphoma

Anne-Ségolène Cottureau<sup>1</sup>, Sebastien Hapdey<sup>2,3</sup>, Loic Chartier<sup>4</sup>, Romain Modzelewski<sup>2,3</sup>, Olivier Casasnovas<sup>5</sup>, Emmanuel Itti<sup>1</sup>, Herve Tilly<sup>6</sup>, Pierre Vera<sup>2,3</sup>, Michel A. Meignan<sup>1</sup>, and Stéphanie Becker<sup>2,3</sup>

<sup>1</sup>Nuclear Medicine Department, Hôpital Henri Mondor, University Paris-Est Créteil, Créteil, France; <sup>2</sup>Nuclear Medicine Department, Henri Becquerel Cancer Center and Rouen University Hospital, Rouen, France; <sup>3</sup>QuantIF-LITIS (EA [Equipe d'Accueil] 4108), Faculty of Medicine, University of Rouen, Rouen, France; <sup>4</sup>Department of Biostatistics (LYSARC), Centre Hospitalier Lyon Sud, Pierre Bénite, France; <sup>5</sup>Hematology Department, Hopital Le Bocage-CHU Dijon, France; and <sup>6</sup>Hematology Department, UMR918, Henri Becquerel Cancer Center and Rouen University Hospital, Rouen, France

The purpose of this study was to compare in a large series of peripheral T cell lymphoma, as a model of diffuse disease, the prognostic value of baseline total metabolic tumor volume (TMTV) measured on <sup>18</sup>F-FDG PET/CT with adaptive thresholding methods with TMTV measured with a fixed 41% SUV<sub>max</sub> threshold method. **Methods:** One hundred six patients with peripheral T cell lymphoma, staged with PET/CT, were enrolled from 5 Lymphoma Study Association centers. In this series, TMTV computed with the 41% SUV<sub>max</sub> threshold is a strong predictor of outcome. On a dedicated workstation, we measured the TMTV with 4 adaptive thresholding methods based on characteristic image parameters: Daisne (Da) modified, based on signal-to-background ratio; Nes-*tle* (Ns), based on tumor and background intensities; Fit, including a 3-dimensional geometric model based on spatial resolution (Fit); and Black (Bl), based on mean SUV<sub>max</sub>. The TMTV values obtained with each adaptive method were compared with those obtained with the 41% SUV<sub>max</sub> method. Their respective prognostic impacts on outcome prediction were compared using receiver-operating-characteristic (ROC) curve analysis and Kaplan-Meier survival curves. **Results:** The median value of TMTV<sub>41%</sub>, TMTV<sub>Da</sub>, TMTV<sub>Ns</sub>, TMTV<sub>Fit</sub>, and TMTV<sub>Bl</sub> were, respectively, 231 cm<sup>3</sup> (range, 5–3,824), 175 cm<sup>3</sup> (range, 8–3,510), 198 cm<sup>3</sup> (range, 3–3,934), 175 cm<sup>3</sup> (range, 8–3,512), and 333 cm<sup>3</sup> (range, 3–5,113). The intraclass correlation coefficients were excellent, from 0.972 to 0.988, for TMTV<sub>Da</sub>, TMTV<sub>Fit</sub>, and TMTV<sub>Ns</sub>, and less good for TMTV<sub>Bl</sub> (0.856). The mean differences obtained from the Bland-Altman plots were 48.5, 47.2, 19.5, and –253.3 cm<sup>3</sup>, respectively. Except for Black, there was no significant difference within the methods between the ROC curves (*P* > 0.4) for progression-free survival and overall survival. Survival curves with the ROC optimal cutoff for each method separated the same groups of low-risk (volume ≤ cutoff) from high-risk patients (volume > cutoff), with similar 2-y progression-free survival (range, 66%–72% vs. 26%–29%; hazard ratio, 3.7–4.1) and 2-y overall survival (79%–83% vs. 50%–53%; hazard ratio, 3.0–3.5). **Conclusion:** The prognostic value of TMTV remained quite similar whatever the methods, adaptive or 41% SUV<sub>max</sub>, supporting its use as a strong prognosticator

in lymphoma. However, for implementation of TMTV in clinical trials 1 single method easily applicable in a multicentric PET review must be selected and kept all along the trial.

**Key Words:** metabolic tumor volume; lymphoma; PTCL; adaptive thresholds

**J Nucl Med 2017; 58:276–281**

DOI: 10.2967/jnumed.116.180406

---

**P**ET/CT with <sup>18</sup>F-FDG has been recognized as the best imaging tool for staging and response assessment in FDG-avid lymphoma. The last International Conference on Malignant Lymphoma recommendations (1) encourage investigating the quantitative analysis of <sup>18</sup>F-FDG PET/CT at staging. In this regard, the measurement of the total metabolic tumor volume (TMTV), which gives an estimation of the total tumor burden, has gained special interest. Indeed, several series have shown that TMTV was predictive of outcome in different lymphoma subtypes: diffuse large B cell lymphoma (2,3), Hodgkin lymphoma (4), peripheral T cell lymphoma (PTCL) (5), and follicular lymphoma (6). In these studies, different methods of TMTV measurement were used; all were based on a fixed thresholding principle to determine the metabolic volume of local tumors. The threshold can be absolute: an SUV of 2.5 was generally chosen (2). It can be relative, using a percentage of the maximum uptake. A threshold of 41% of the SUV<sub>max</sub> within the lesion, recommended by the European Association of Nuclear Medicine for solid tumors (7), has been used in patients with Hodgkin lymphoma (4), diffuse large B cell lymphoma (3), and PTCL (5), with a good interobserver reproducibility. However, because lymphomas are heterogeneous disease with several tumor sites with a wide range of volumes, SUVs, and tumor background ratios, the adaptive segmentation methods might be of interest for TMTV measurement and could be proposed as an alternative to fixed thresholding methods. The principle of these adaptive methods developed for radiotherapy planning of solid tumors is to adapt the threshold following a fitting model according to 1 or 2 characteristic image parameters, such as the SUV or the contrast. In a previous study, we have demonstrated in a retrospective group of PTCL patients that TMTV measured with the 41% SUV<sub>max</sub> threshold method was a good predictor of outcome (5). The

Received Jun. 29, 2016; revision accepted Sep. 8, 2016.

For correspondence or reprints contact: Anne-Ségolène Cottureau, Nuclear Medicine Department, AP-HP, Groupe Henri Mondor Albert Chenevier, CHU Henri Mondor, 51 Avenue du Marechal de Lattre de Tassigny, 94010 Créteil, France.

E-mail: annesegolene.cottureau@aphp.fr

Published online Oct. 6, 2016.

COPYRIGHT © 2017 by the Society of Nuclear Medicine and Molecular Imaging.

aim of the present study was to compare in the same series, taken as a model of diffuse lymphoma, different adaptive thresholding methods with this fixed 41% method and to evaluate whether these adaptive thresholding methods were better predictors of outcome than a fixed relative threshold.

## MATERIALS AND METHODS

### Patients

A group of consecutive patients with PTCL newly diagnosed during 2006–2014, from 5 Lymphoma Study Association centers (Creteil, Dijon, Marseille, Rouen, and Liege), were included in this study. All were part of a previous study already published on the prognostic value of baseline TMTV (5). Patients underwent baseline PET/CT, with central data available for review. The median follow-up of this population was 23 mo. The 2-y progression-free survival (PFS) and overall survival (OS) of the population were 49% and 67%, respectively. The patient characteristics were a median age of 58 y; 91% of advanced stage, half of them with an international prognostic index greater than 2 and 45% with a prognostic index for PTCL greater than 1; and 80% treated with cyclophosphamide, doxorubicin, vincristine, prednisone (CHOP)/CHOP-like therapy and 20% with cyclophosphamide, doxorubicin, vindesine, bleomycin. The study was conducted in accordance with the precepts of the Helsinki Declaration and received approval by the Ethical Committee with a waiver of informed consent because of the retrospective nature of the study (5).

### PET Acquisition

All the centers adhered to guidelines of the European Association of Nuclear Medicine for patient preparation and PET/CT acquisition. All patients were instructed to fast for at least 6 h before the injection of 4–5 MBq per kilogram of  $^{18}\text{F}$  FDG, to ensure that serum glucose and endogenous serum insulin levels were low. Non-contrast-enhanced CT images were acquired before PET data acquisition. Whole-body PET was acquired sequentially using a dedicated PET/CT system. For the PET imaging, the emission data were acquired from the base of the skull to the proximal thigh with 3–3.5 min of acquisition per bed position. A Biograph Sensation 16 Hi-Rez (Siemens Medical Solution) or Gemini GXL or Gemini TOF (Philips) scanner was used by the 5 centers. All the devices used in this study followed a quality control program ensuring that the data were quantitatively correct (quarterly SUV verification). Four centers had obtained European Association of Nuclear Medicine Research Ltd. accreditation at the time of the study, and 1 was accredited according to the RTEP (Radiotherapy and PET) trial procedure (8). The similarity of performances of the different equipment was confirmed with the analysis of the recovery curves obtained from their NEMA phantoms in terms of volume and contrast.

### TMTV Measurement

The baseline  $^{18}\text{F}$ -FDG PET/CT was processed with a Planet Onco workstation (version 2.0; DOSISoft) located at the Henri Becquerel Center, Rouen.

TMTV was computed using the following steps. First, the volumetric regions of interest were placed around each lesion, avoiding physiologic uptake (urinary elimination, heart). The reproducibility of the region-of-interest setting has been evaluated, as previously published (5). Then the tumor volume was delineated with 5 thresholding methods: 1 fixed, 41%  $\text{SUV}_{\text{max}}$  considered thereafter as the reference (9), and 4 adaptive based on mathematic algorithms—Daisne modified by Vauclin et al. (TMTV<sub>Da</sub>), which iteratively adapts the threshold according to the local signal-to-background ratio (10); Fit (TMTV<sub>Fit</sub>), which fits the sphere image using a 3-dimensional geometric model based on the spatial resolution in the reconstructed

images and on a tumor shape derived from activity thresholding (11,12); Nestle (TMTV<sub>Ns</sub>) according to tumor and background intensities (13); and Black (TMTV<sub>Bl</sub>) according to the  $\text{SUV}_{\text{mean}}$  (14). The tumor  $\text{SUV}_{\text{max}}$  and the liver  $\text{SUV}_{\text{max}}$  were also reported.

### Statistical Analysis

Quantitative variables were expressed as median and ranges. The comparison of TMTV between methods was analyzed with the Friedman nonparametric test. When a significant difference was detected, a Wilcoxon signed-rank test was performed, with Bonferroni adjustment for multiple testing. The correlation between TMTV values from the 41%  $\text{SUV}_{\text{max}}$  method and those from adaptive methods (Daisne modified, Fitting, Nestle, and Black) was tested using Spearman coefficients. Agreement between methods were represented on Bland–Altman plots and quantified with the intraclass correlation coefficient (ICC) based on Shrout–Fleiss formulae.

For each method, receiver-operating-characteristic (ROC) curves were obtained to define the optimal TMTV cutoff for survival prediction. The prognostic relevance of each method to predict PFS and OS was pairwise comparison of these ROC curves. Survival functions were calculated with Kaplan–Meier estimates for each method using their optimal TMTV cutoff. The comparison between categories was made using the log-rank test and Cox proportional hazards models. The agreement between the dichotomization of patients in low- and high-TMTV groups obtained with the 41%  $\text{SUV}_{\text{max}}$  method and each adaptive method was tested with the Cohen  $\kappa$ -coefficient. Survival functions were also calculated for each method using the threshold of 230  $\text{cm}^3$  determined with the 41%  $\text{SUV}_{\text{max}}$  method for PFS and OS (5).

Because PET/CT was performed in several institution and adaptive methods were optimized on a Dosisoft workstation for the Rouen equipment (Biograph Sensation 16 Hi-Rez), we verified that the Spearman coefficients as well as the ICC based on Shrout–Fleiss formulae existing between methods were similar between the Rouen subset and the whole population. Statistical analyses used SAS 9.2 (SAS Institute) and Med Calc (MedCalc Software).

## RESULTS

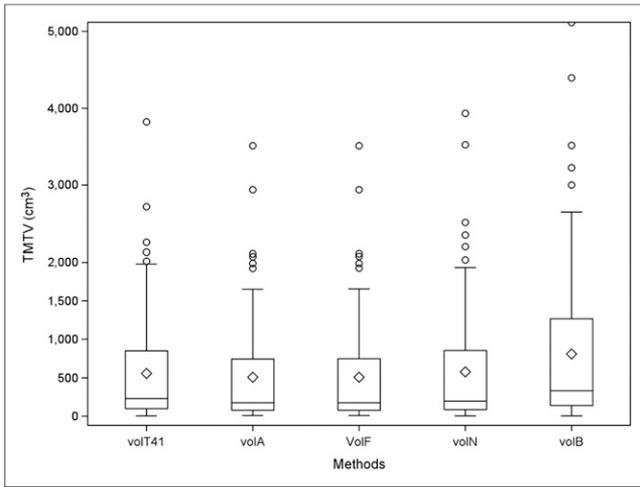
One hundred six patients with PTCL newly diagnosed and for whom PET/CT could be retrieved for analysis on Dosisoft software were included in the present study. Characteristics of this group were similar to the initial population.

### TMTV Measurement

The median values of TMTV<sub>41%</sub>, TMTV<sub>Da</sub>, TMTV<sub>Fit</sub>, TMTV<sub>Ns</sub>, and TMTV<sub>Bl</sub> were 231  $\text{cm}^3$  (range, 5–3,824), 175  $\text{cm}^3$  (range, 8–3,510), 175  $\text{cm}^3$  (range, 8–3,512), 198  $\text{cm}^3$  (range, 3–3,934), and 333  $\text{cm}^3$  (range, 3–5,113), respectively (Fig. 1). We observed no significant difference between Nestle and 41%  $\text{SUV}_{\text{max}}$  methods ( $P = 0.7$ ), but significant differences existed between the other adaptive methods and 41%  $\text{SUV}_{\text{max}}$  ( $P < 0.001$ ): from 24% for Fit and Daisne to 44% for Black. The median tumor  $\text{SUV}_{\text{max}}$  was 14, with a wide range of values (3.4–39.0). The tumor  $\text{SUV}_{\text{max}}$ -to-liver  $\text{SUV}_{\text{max}}$  ratio, taken as an index of tumor-to-background ratio, ranged from 1 to 15. The reproducibility of volume-of-interest setting was excellent, with a Lin concordance correlation coefficient of  $\rho = 0.995$  (95% confidence interval [CI], 0.992–0.997) (5).

### Comparison of TMTV Values Between 41% $\text{SUV}_{\text{max}}$ Method and Adaptive Methods

The correlation between TMTV values from the 41%  $\text{SUV}_{\text{max}}$  method and those from adaptive methods was excellent, with a



**FIGURE 1.** TMTV distribution according to each methodology, with median and interquartile range (box), mean (diamond), and outliers (circle).

Spearman coefficient of 0.99, 0.99, 0.97, and 0.96 for TMTV<sub>Da</sub>, TMTV<sub>Fit</sub>, TMTV<sub>Ns</sub>, and TMTV<sub>Bl</sub>, respectively.

The Bland–Altman plots are presented in Figure 2. The means of difference between TMTV<sub>41%</sub> and TMTV<sub>Da</sub>, TMTV<sub>Fit</sub>, TMTV<sub>Ns</sub>, and TMTV<sub>Bl</sub> were  $48.5 \pm 97.3$  SD (95% CI,  $-147$  to  $243$ ),  $47.2 \pm 96.7$  (95% CI,  $-146$  to  $241$ ),  $-19.5 \pm 170.6$  (95% CI,  $-360$  to  $322$ ), and  $-253.3 \pm 393.8$  (95% CI,  $534$  to  $-1,041$ ), respectively. A major overestimation was observed with TMTV<sub>Bl</sub> compared with TMTV<sub>41%</sub>. The differences between methods increased for the high-TMTV group, with a proportional error for TMTV values greater than  $500 \text{ cm}^3$ . Looking at TMTV under  $500 \text{ cm}^3$ , the mean differences between 41% SUV<sub>max</sub> and adaptive methods were really reduced:  $21.8 \pm 36.3$  for TMTV<sub>Da</sub>,  $21 \pm$

$34.8$  for TMTV<sub>Fit</sub>,  $0.4 \pm 54.5$  for TMTV<sub>Ns</sub>, and  $-105.8 \pm 124.5$  for TMTV<sub>Bl</sub>.

The ICC varied from 0.972 to 0.988 for TMTV<sub>Da</sub>, TMTV<sub>Fit</sub>, and TMTV<sub>Ns</sub> and was 0.856 for TMTV<sub>Bl</sub>. The coefficient of variation between TMTV<sub>41%</sub> and TMTV<sub>Da</sub>, TMTV<sub>Fit</sub>, TMTV<sub>Ns</sub>, and TMTV<sub>Bl</sub> was, respectively, 14%, 14%, 21%, and 48%.

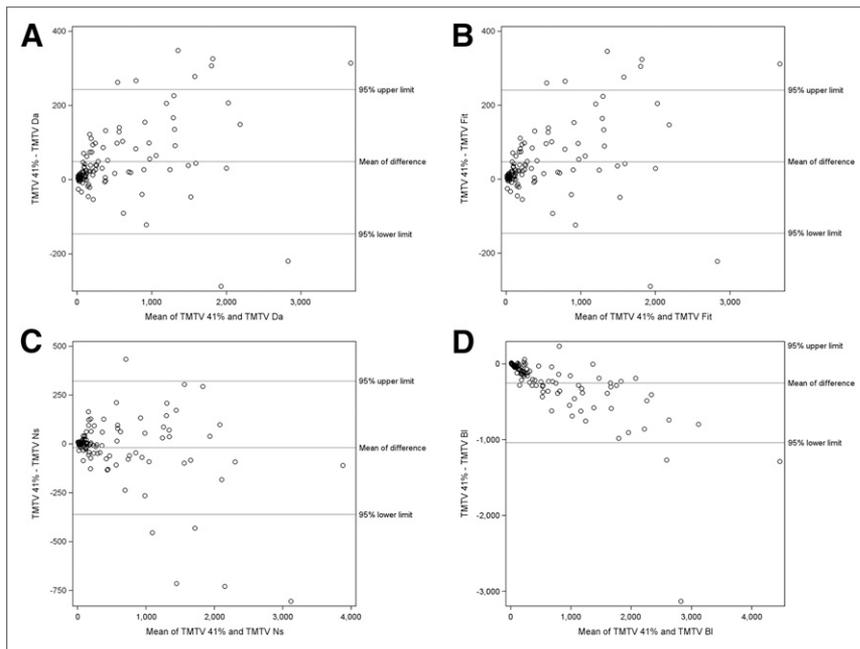
### Prognostic Value of Different Methods

The respective optimal cutoffs found with ROC analysis for TMTV<sub>41%</sub>, TMTV<sub>Da</sub>, TMTV<sub>Fit</sub>, TMTV<sub>Ns</sub>, and TMTV<sub>Bl</sub> were 230, 132, 147, 277, and  $345 \text{ cm}^3$  for PFS and 260, 132, 147, 191, and  $345 \text{ cm}^3$  for OS (Table 1). The respective area under the curve of TMTV<sub>41%</sub>, TMTV<sub>Da</sub>, TMTV<sub>Fit</sub>, TMTV<sub>Ns</sub>, and TMTV<sub>Bl</sub> varied from 0.68 to 0.71 for PFS and from 0.60 to 0.62 for OS. For PFS, the area under the curve obtained with TMTV<sub>Da</sub>, TMTV<sub>Fit</sub>, and TMTV<sub>Ns</sub> was not significantly different ( $P > 0.4$  for each pairwise comparison) from the area under the curve of TMTV<sub>41%</sub> (Fig. 3). A significant difference was observed for TMTV<sub>Bl</sub> on PFS ( $P = 0.02$ ). No significant difference was observed for OS.

TMTV was significantly associated with inferior PFS ( $P < 0.001$ ) and OS ( $P < 0.001$ ) whatever the method of computation, with no significant difference between them and similar hazard ratio (HR) (Table 1). The HR ranged from 3.7 (TMTV<sub>Bl</sub>) to 4.1 (TMTV<sub>41%</sub>) on PFS and from 3.0 (TMTV<sub>Bl</sub>) to 3.5 (TMTV<sub>Da</sub>) on OS. The 2-y PFS ranged from 66% to 72% for the low-TMTV groups versus 26%–29% for the high-TMTV groups and 2-y OS from 79% to 83% versus 50% to 53% (Fig. 4).

On the basis of the  $230\text{-cm}^3$  cutoff (threshold determined with the 41% SUV<sub>max</sub> method), the agreement between the 41% SUV<sub>max</sub> method and adaptive methods to dichotomize the population was almost perfect, with a  $\kappa$  of 0.87, 0.87, and 0.87 for TMTV<sub>Da</sub>, TMTV<sub>Fit</sub>, and TMTV<sub>Ns</sub>, respectively. As expected, because of the major overestimation, no agreement was observed with the Black method. When this single cutoff of  $230 \text{ cm}^3$  was applied to

the various methods, comparable PFS prediction was found ( $P < 0.0001$ , HR, 3.4, for Daisne modified and Fit;  $P < 0.0001$ , HR, 3.3, for Nestle; and  $P = 0.0056$ , HR, 2.4, for Black). This TMTV cutoff remained significant to predict OS except for Black ( $P = 0.058$ ). For the other adaptive methods, this TMTV cutoff was slightly less significant than when using a specific optimal cutoff for each method ( $P = 0.0055$ , HR, 2.4, for Daisne modified and Fit; and  $P = 0.0037$ , HR, 2.6, for Nestle).



**FIGURE 2.** Bland–Altman analysis comparing TMTV values of TMTV<sub>Da</sub> (A), TMTV<sub>Fit</sub> (B), TMTV<sub>Ns</sub> (C), and TMTV<sub>Bl</sub> (D) with TMTV<sub>41%</sub>. Mean bias and limits of agreements are represented by solids lines.

### Comparison Between Rouen and Other Centers

Among the 106 patients, 28 were included in Rouen and 78 in the other centers. The coefficient of variation observed in the Rouen center between TMTV<sub>41%</sub> and TMTV<sub>Da</sub>, TMTV<sub>Fit</sub>, TMTV<sub>Ns</sub>, and TMTV<sub>Bl</sub> was, respectively, 11.3%, 11.2%, 20.7%, and 40.1%, similar to those observed in the whole population. Correlations between methods were similar in the Rouen subgroup and in the other centers, with a Spearman coefficient from 0.96 to 0.99 for both subgroups. ICC

**TABLE 1**

ROC Optimal TMTV Cutoff, PFS, and OS Analyzed According to TMTV<sub>41%</sub>, TMTV<sub>Da</sub>, TMTV<sub>Fit</sub>, TMTV<sub>Ns</sub>, and TMTV<sub>Bl</sub>

Measurement	PFS			OS		
	Specific threshold (cm <sup>3</sup> )	P	HR	Specific threshold (cm <sup>3</sup> )	P	HR
TMTV <sub>41%</sub>	230	<0.0001	4.1 (2.3–7.3)	260	0.0005	3.1 (1.6–6.0)
TMTV <sub>Da</sub>	132	<0.0001	4.0 (2.3–7.0)	132	0.0007	3.5 (1.8–6.6)
TMTV <sub>Fit</sub>	147	<0.0001	4.0 (2.3–6.9)	147	0.0009	3.3 (1.7–6.2)
TMTV <sub>Ns</sub>	277	<0.0001	3.9 (2.2–7.0)	191	0.0007	3.2 (1.7–6.2)
TMTV <sub>Bl</sub>	345 cm <sup>3</sup>	<0.0001	3.7 (2.1–6.6)	345 cm <sup>3</sup>	0.0009	3.0 (1.6–5.8)

Data in parentheses are 95% CIs.

was also similar for TMTV<sub>Da</sub>, TMTV<sub>Fit</sub>, and TMTV<sub>Ns</sub> between both subgroups, from 0.974 to 0.993 in the Rouen center and from 0.972 to 0.985 in the other centers. Regarding the Black method, ICC observed in the other centers was slightly lower than in the Rouen center, with an ICC of 0.836 compared to 0.9.

**DISCUSSION**

The major result of this study showed that the prognostic value of baseline TMTV computed with several adaptive methods was similar to TMTV computed with 41% SUV<sub>max</sub> threshold method in a large series of PTCL patients taken as an example of diffuse tumor disease.

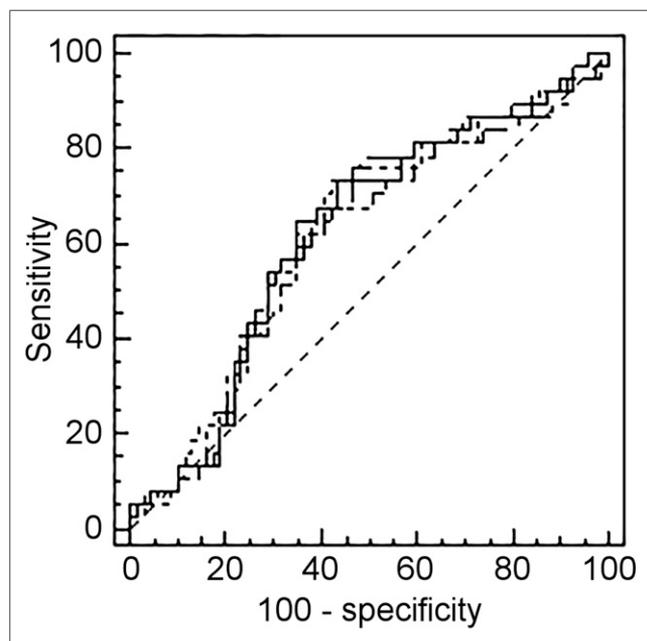
Retrospective studies have demonstrated that TMTV was a powerful predictor of outcome in different lymphoma subtypes. TMTV measurement at baseline is important because it could help in stratifying patients in different risk categories and has been suggested as a possible tool for early guiding of therapy.

However, until now in lymphoma different TMTV methodologies have been used: an absolute cutoff of SUV greater than 2.5 or a relative SUV<sub>max</sub> thresholding of the tumor sites. The absolute threshold using SUV greater than 2.5 is limited by the variability of SUVs, because of PET/CT devices, PET acquisition protocol, and reconstruction methods. In addition, because of partial-volume effect nontumor regions located between small distant nodes with high uptake could be included (15,16). Therefore, relative thresholds have been used in several lymphoma studies: a 41% SUV<sub>max</sub> cutoff as recommended by guidelines of the European Association of Nuclear Medicine for solid tumors was applied in diffuse large B cell lymphoma, Hodgkin lymphoma, follicular lymphoma, and PTCL, and a 25% SUV<sub>max</sub> was recently evaluated in primary mediastinal large B cell lymphoma (17).

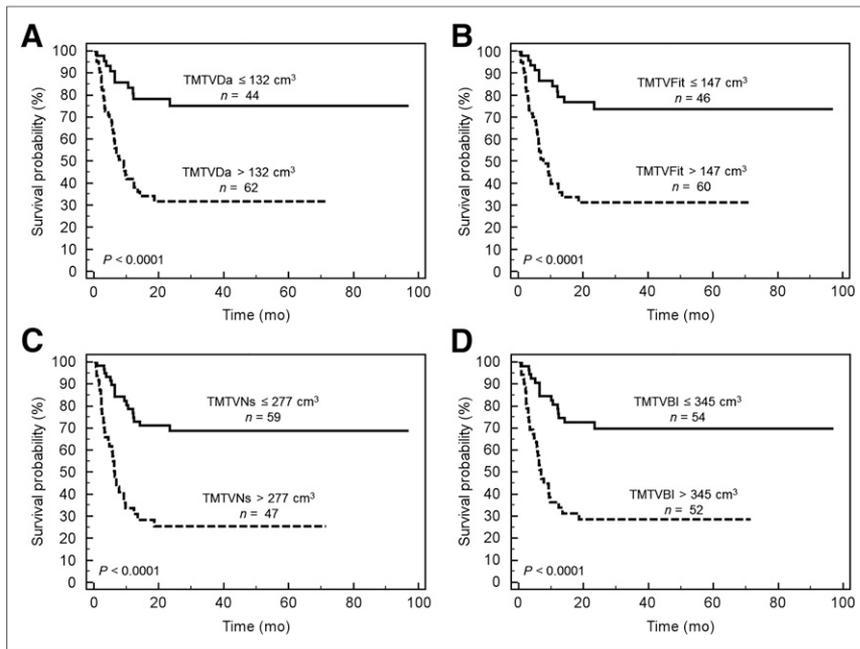
Lymphoma characteristics—that is, a disseminated disease with different size of lesion, different sites with nodal or extranodal lesions, and heterogeneous <sup>18</sup>F-FDG uptake (tumor-to-liver ratios varying from 1 to 15 in our series)—might limit the efficiency of both fixed-threshold methods. Conversely, adaptive methods may be more accurate and even easier to use in routine but they had not yet been tested on lymphoma.

In a previous study, we have shown in PTCL that baseline TMTV with a 230-cm<sup>3</sup> threshold was a good tool for outcome prediction and predicted PFS and OS much better than the currently used clinical index.

In the same series of patients, we observed that the ICC found between TMTV values obtained with the 3 adaptive methods (Daisne modified, Fit, and Nestle) and those from the 41% SUV<sub>max</sub> method were excellent. The optimal thresholds dichotomizing the population in low- and high-volume groups for each adaptive method were different, but despite these differences all these methods predicted PFS and OS with similar P and HR values for small and large volumes. The only slight incremental prognostic value compared with the 41% threshold method was observed for the Daisne modified method for OS prediction. Moreover, when the same threshold of TMTV obtained with the 41% SUV<sub>max</sub> method (230 cm<sup>3</sup>) was used for Daisne modified, Fit, and Nestle methods, the HRs obtained for PFS and OS prediction were comparable, supporting the fact that the values obtained with each method were really close. Only Black was out of range with a major overestimation. This is probably explained because Black is based on the SUV<sub>mean</sub> whereas the other methods are based on the SUV<sub>max</sub>.



**FIGURE 3.** PFS ROC curve comparison according to each methodology.



**FIGURE 4.** Kaplan–Meier estimates of PFS according to TMTV<sub>Da</sub> (A), TMTV<sub>Fit</sub> (B), TMTV<sub>Ns</sub> (C), and TMTV<sub>Bi</sub> (D).

A comparison between different methods of metabolic volume measurements has already been done in Hodgkin lymphoma. Kanoun et al. (18) have compared 41% SUV<sub>max</sub> threshold to a per-patient adapted threshold based on SUV<sub>max</sub> of the liver (>125% and >140% SUV<sub>max</sub> of the liver background) in a monocentric study. They found no significant difference between ROC curves and similar prediction of PFS and OS according to high volume. It emphasized the strong prognostic value of metabolic imaging because TMTV remained prognostic throughout the different methods used. The current study is the first, to our knowledge, comparing adaptive methods to a fixed-threshold method in a large series of patients with a diffuse subtype of lymphoma, PTCL. Our results further confirm the strong prognostic value of baseline TMTV and demonstrate that these different methods of TMTV measurement equally predict outcome.

In this study, we analyzed retrospective data acquired with different PET systems on a workstation already calibrated to one of them for adaptive methods. However, standardized phantom experiments confirmed the similarity of image characteristics between centers and patients' data comparison between adaptive techniques, and the 41% threshold was similar within centers. Indeed, TMTVs were all over 5 cm<sup>3</sup>, with a median of 231 cm<sup>3</sup>, an order of magnitude of volume where PET systems have similar detection capability.

Several quantitative measurements including TMTV and total lesion glycolysis have been done in ancillary studies of prospective trials based on quality-controlled PET (2,6,17). However, to our knowledge, no ongoing trials have been launched using the TMTV to guide therapy. Even if the prerequisite for this type of trial is quality control, as done using various existing control systems (7,19,20), it is anyway required for good PET clinical practice. The main problem is which TMTV technique measurement should be chosen because there is no established consensus.

Relative methods (SUV<sub>max</sub> thresholding or adaptive) have the advantage to minimize the errors linked to the use of different devices and the participation of different centers.

In addition, our results suggest that it is possible to conduct a prospective trial based on TMTV measurement provided a single relative method of TMTV measurement is used by all participating centers. The 41% TMTV threshold method is currently available in all commercial software and can be used in most of the PET/CT systems. It has demonstrated a good reproducibility among trained observers but requires accurate manual drawing of the volume of interest around each lesion. Adaptive methods might be an option if available in all the centers.

## CONCLUSION

Our results demonstrate that adaptive methods can be used with the same efficacy as the 41% TMTV method in PTCL and would open the way to automatic procedures

of volume computation. These conclusions should be confirmed for other types of diffuse aggressive lymphoma and new-generation devices.

## DISCLOSURE

No potential conflict of interest relevant to this article was reported.

## ACKNOWLEDGMENTS

We acknowledge the radiophysicists of all centers participating in the study: Claire Bernard, Sebastien Hapdey, Helene Masset, and Jean Marc Vrigneaud. This study was presented during the Annual Meeting of the Society of Nuclear Medicine and Molecular Imaging, San Diego, California, June 11–15, 2016.

## REFERENCES

- 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;32:3048–3058.
- 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;43:1209–1219.
- 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;22:3801–3809.
- 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;41:1735–1743.
- 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;27:719–724.
- 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*. August 22, 2016 [Epub ahead of print].

7. 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.
8. Vera P, Mezzani-Saillard S, Edet-Sanson A, et al. FDG PET during radiochemotherapy is predictive of outcome at 1 year in non-small-cell lung cancer patients: a prospective multicentre study (RTEP2). *Eur J Nucl Med Mol Imaging*. 2014;41:1057–1065.
9. 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;41:1113–1122.
10. Vauclin S, Doyeux K, Hapdey S, Edet-Sanson A, Vera P, Gardin I. Development of a generic thresholding algorithm for the delineation of <sup>18</sup>F-FDG-PET-positive tissue: application to the comparison of three thresholding models. *Phys Med Biol*. 2009;54:6901–6916.
11. Tylski PGN, Giraud P, Rosenwald J, Buvat I. Experimental comparison of three methods for estimating tumor volume in FDG PET [abstract]. *J Nucl Med*. 2007;48(suppl 2):43P.
12. Tylski P, Stute S, Grotus N, et al. Comparative assessment of methods for estimating tumor volume and standardized uptake value in <sup>18</sup>F-FDG PET. *J Nucl Med*. 2010;51:268–276.
13. Nestle U, Kremp S, Schaefer-Schuler A, et al. Comparison of different methods for delineation of <sup>18</sup>F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *J Nucl Med*. 2005;46:1342–1348.
14. Black QC, Grills IS, Kestin LL, et al. Defining a radiotherapy target with positron emission tomography. *Int J Radiat Oncol Biol Phys*. 2004;60:1272–1282.
15. Quak E, Le Roux PY, Hofman MS, et al. Harmonizing FDG PET quantification while maintaining optimal lesion detection: prospective multicentre validation in 517 oncology patients. *Eur J Nucl Med Mol Imaging*. 2015;42:2072–2082.
16. Quak E, Hovhannisyann N, Lasnon C, et al. The importance of harmonizing interim positron emission tomography in non-Hodgkin lymphoma: focus on the Deauville criteria. *Haematologica*. 2014;99:e84–e85.
17. Ceriani L, Martelli M, Zinzani PL, et al. Utility of baseline <sup>18</sup>F-FDG-PET/CT functional parameters in defining prognosis of primary mediastinal (thymic) large B-cell lymphoma. *Blood*. 2015;126:950–956.
18. Kanoun S, Tal I, Berriolo-Riedinger A, et al. Influence of software tool and methodological aspects of total metabolic tumor volume calculation on baseline [<sup>18</sup>F]FDG PET to predict survival in hodgkin lymphoma. *PLoS One*. 2015;10:e0140830.
19. Chauvie S, Bergesio F, Fioroni F, et al. The <sup>68</sup>Ge phantom-based FDG-PET site qualification program for clinical trials adopted by FIL (Italian Foundation on Lymphoma). *Phys Med*. 2016;32:651–656.
20. Palie O, Michel P, Menard JF, et al. The predictive value of treatment response using FDG PET performed on day 21 of chemoradiotherapy in patients with oesophageal squamous cell carcinoma: a prospective, multicentre study (RTEP3). *Eur J Nucl Med Mol Imaging*. 2013;40:1345–1355.