Received: 24 May 2023 | Revised: 14 October 2023 | Accepted: 19 October 2023

DOI: 10.1002/hon.3235

WILEY

ORIGINAL ARTICLE

A low total metabolic tumor volume independently predicts for a longer time to first treatment in initially observed, low tumor burden follicular lymphoma

Pablo Mozas[1,2,3] | Sebastian Casanueva-Eliceiry[4] | Andrea Rivero[1,2] | Ángel Serna[5] | Marc Simó[6] | Sonia Rodríguez[7] | Alfredo Rivas-Delgado[1,2] | Ferran Nadeu[2,8] | Juan Gonzalo Correa[1] | Juan Antonio Piñeyroa[1] | Amanda Isabel Pérez-Valencia[1] | Katia Guinetti-Ortiz[1] | Marta Gómez-Hernando[1] | Eva Giné[1,2,8] | Julio Delgado[1,2,3,8] | Neus Villamor[2,8,9] | Elías Campo[2,3,8,9] | Laura Magnano[1,2,8] | Pau Abrisqueta[5] | Xavier Setoain[3,4,10] | Armando López-Guillermo[1,2,3,8]

[1]Department of Hematology, Hospital Clínic de Barcelona, Barcelona, Spain

[2]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain

[3]Departament de Medicina, Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona, Barcelona, Spain

[4]Department of Nuclear Medicine, Hospital Clínic de Barcelona, Barcelona, Spain

[5]Department of Hematology, Hospital Universitari Vall d'Hebron, Barcelona, Spain

[6]Department of Nuclear Medicine, Hospital Universitari Vall d'Hebron, Barcelona, Spain

[7]Department of Radiology, Hospital Clínic de Barcelona, Barcelona, Spain

[8]Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain

[9]Hematopathology Unit, Department of Pathology, Hospital Clínic de Barcelona, Barcelona, Spain

[10]Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Madrid, Spain

Correspondence
Pablo Mozas, Department of Hematology, Hospital Clínic, Barcelona, Spain.
Email: mozas@clinic.cat

Funding information
Instituto de Salud Carlos III; European Regional Development Fund, Grant/Award Numbers: PI19/00925, PI19/00887

Abstract

Watchful waiting is an acceptable management strategy for advanced-stage, low tumor burden (LTB) patients with follicular lymphoma (FL). However, the prediction of how long this treatment-free observation period will last remains imperfect. We explored whether total metabolic tumor volume (TMTV) and other positron emission tomography parameters were predictive of time to first treatment (TTFT). We analyzed 97 grade 1–3A advanced-stage LTB FL patients and found that a high TMTV was associated with other tumor burden features at diagnosis. Patients with a TMTV above our established cutoff of 50 mL had a significantly shorter median duration of observation (2.6 vs. 8.8 years; $p = 0.001$). At 5 years, 77% of patients

Pablo Mozas and Sebastian Casanueva-Eliceiry co-first authors.

Xavier Setoain and Armando López-Guillermo co-senior authors.

with a high TMTV and 46% of patients with a low TMTV required treatment. In the multivariable analysis, a high TMTV was the only independent factor predicting TTFT (hazard ratio = 2.09; $p = 0.017$). Overall, TMTV is a strong predictor of the duration of observation in LTB FL patients. Upon validation of our cutoff in external series and standardization of the methodology, the TMTV could become an additional factor to consider deferring or initiating treatment in otherwise LTB patients.

KEYWORDS
follicular lymphoma, positron emission tomography-computed tomography, time to first treatment, total metabolic tumor volume, watchful waiting

1 | INTRODUCTION

Follicular lymphoma (FL), the most common indolent B-cell lymphoma, is considered an incurable malignancy. Although some patients relapse early (POD24)[1] or develop histological transformation (HT),[2] both of which dramatically worsen patients' outcomes, most individuals experience a protracted disease history, with long remissions and recurring relapses, and a median overall survival (OS) now exceeding 20 years.[3] It is also acknowledged that around 10% of FL patients can experience spontaneous regression.[4]

The chronic nature and prolonged survival of FL have led to the development of older and more recent studies evaluating the benefit of active treatment versus a watchful waiting (WW) approach in low tumor burden (LTB disease).[4–7] Comparable outcomes between both strategies (in terms of OS and the risk of HT) have made it customary in many countries to conservatively manage asymptomatic patients. In a population characterized by its advanced age and the presence of comorbidity,[8] sparing therapy and thus toxicity seems more than reasonable. Individuals with localized disease can be treated with anti-CD20 immunotherapy and/or radiation, while advanced-stage cases are generally divided into those with and without high tumor burden features, which are an indication for starting treatment.[9]

Various high tumor burden criteria have been proposed,[10,11] being the one by the *Groupe d'Étude des Lymphomes Folliculaires* (GELF)[12] among the most widespread. Requirements for WW include the patient's will to undergo such a strategy (some individuals have a poor tolerance to having cancer and not receiving any therapy), absence of lymphoma-related symptoms, of large lymphoid masses, and of lymphoma-related organ dysfunction (including bone marrow (BM)). GELF criteria only consider CT-derived morphological parameters measured in a single plane. Due to CT limitations to define tumor limits, difficult-to-measure lesions such as those located in the spleen, BM, and pleura are not included in the assessment.

For LTB patients undergoing observation, the median time to first treatment (TTFT) has been set at around 3 years.[4,7,13,14] However, the identification of specific factors predicting treatment initiation in these individuals remains elusive. In a recent single-center study,[15] the Follicular Lymphoma International Prognostic Index (FLIPI) score, Ki67 index, and the proportion of $CD4^+$ and $FOXP3^+$ cells were predictive of WW discontinuation. Of note, the FLIPI score only evaluates tumor burden by CT (enlarged lymph nodes in a single plane), without considering the size of non-measurable splenic and other extranodal lesions.

Follicular lymphoma is considered to be fluorodeoxyglucose (FDG)-avid, and [^{18}F]-FDG positron emission tomography/computed tomography (PET/CT) imaging is recommended before first and subsequent lines of therapy.[16] Besides its usefulness in staging,[17] for guiding biopsy toward the most active lesion, and for assessing response, semiquantitative PET calculations allow for a whole-body volumetric tumor burden measurement that predicts progression-free survival (PFS)[18] and has been incorporated into novel prognostic indexes.[19] In the setting of initially observed LTB FL, two small Chinese ($n = 38$)[20] and Italian ($n = 54$)[21] single-center studies demonstrated that the maximum standardized uptake value (SUVmax), total lesion glycolysis (TLG) and total metabolic tumor volume (TMTV) predicted outcomes, with the caveat that the TTFT cutoff was obtained by receiver operating curve (ROC) analysis, thus treating it as a categorical variable (need of treatment within 2 years from diagnosis).

The aim of our study was to further explore the potential of semiquantitative PET/CT parameters to predict TTFT (performing a time-to-event analysis) in a larger bicentric cohort of LTB, initially observed FL patients.

2 | METHODS

2.1 | Patients

We retrospectively identified 97 grade 1–3A FL patients (43 females, 54 males; median age, 59 years, range 32–84) consecutively diagnosed at two institutions (Hospital Clínic de Barcelona and Hospital Universitari Vall d'Hebron) between June 2006 and September 2020. All patients underwent PET/CT staging within 3 months of FL diagnosis, were considered to have LTB disease (i.e., not fulfilling criteria for initiating treatment as per GELF,[12] i.e., no bulky masses, no involvement of ≥3 nodal sites, each with a diameter >3 cm, no systemic symptoms, no symptomatic splenomegaly, no compression syndrome, no tumor effusions, no overt leukemic involvement, no disease-related cytopenia) and were observed without treatment for

a minimum of 3 months. Most patients (95/97, 98%) underwent a staging BM biopsy at diagnosis. Monitoring of disease progression was based on physical examination and laboratory analyses performed every 3–4 months, and this policy was similar across participating centers and treating physicians. Patients may eventually have or have not received treatment during follow-up, which in all cases responded to the development of high tumor burden disease. Patients with the following entities were not included: grade 3B FL, primary gastrointestinal or cutaneous FL, and composite lymphoma (FL/DLBCL). Baseline characteristics, treatment, and outcomes were evaluated and compared according to PET/CT parameters. The study was approved by Hospital Clínic de Barcelona Institutional Review Board (HCB/2021/0415) and followed the Declaration of Helsinki.

2.2 | PET/CT parameters

2.2.1 | [^{18}F]-FDG PET/CT protocol

Baseline PET/CT images were acquired from the cranial vault to mid-thigh approximately 60 min after intravenous administration of 3.7 MBq/kg of [^{18}F]-FDG, by means of hybrid PET/CT equipment (Biograph mCT TrueV, Siemens Medical Solutions USA, Inc.), including 5–6 beds (2 min per bed). CT images were enhanced by both oral and intravenous iodinated contrast. All patients underwent low-dose CT for attenuation correction. Further details concerning PET/CT and image analysis are provided in the Supplementary Material.

2.2.2 | Image analysis

Using the same software for all patients, images were visually assessed by two nuclear medicine physicians, and by an independent specialist in conflicting cases, reaching a final image interpretation consensus in all cases. Image analysis was blinded to outcome. The segmentation of the tumor contours was semiautomatically performed by the MIM Software version 7.2.1 (Cleveland, OH). Segmentation threshold was established at a SUV $\geq$2.5 (Figure 1). All included contours that did not correspond to tumor activity were then manually removed (i.e., physiological uptake or concomitant inflammatory/infectious processes). In the BM, only focal tracer uptakes were considered pathological. Both focal and diffuse splenic uptakes were included (>150% of the liver background). Volumetric parameters such as TMTV (defined as the sum of the metabolic volumes of all lesions), TLG (defined as the sum of individual MTV multiplied by its mean SUV) and SUVmax were then obtained.

2.3 | Clinical endpoints

The main endpoint of the study was TTFT, defined as the interval between FL diagnosis and the initiation of frontline therapy. Although this parameter is influenced by several clinical factors, we consider it a relevant outcome in indolent malignancies, such as chronic lymphocytic leukemia or FL, in which WW is a common strategy. It is an indirect but eloquent measure of quality of life, since it is a period of time in which the patient is free from disease- and treatment-related complications. Besides, unlike PFS and OS, TTFT remains independent from the specific therapeutic strategy, which makes of it a good indicator of the natural history of disease.

Response criteria to frontline treatment were the standard.[22] PFS was calculated from frontline treatment to relapse or death of any cause. Early progressors (POD24) were patients who relapsed within 24 months of initial treatment. Overall survival was calculated from diagnosis to last follow-up or death from any cause. Survival from treatment was calculated from frontline treatment initiation to last follow-up or death from any cause.

2.4 | Statistical analysis

The method of maximally selected rank statistics (*maxstat* package, R software, Vienna, Austria), was used to calculate the best TMTV, TLG and SUVmax cutoffs to predict TTFT. The χ^2 or Fisher's exact test were used to compare categorical variables. For TTFT, where a possible competing event exists, the primary event was the initiation of treatment and the competing event was death during WW. Cumulative incidence was then calculated (*cmprsk* R package) and Gray's test[23] was used for comparisons between both groups. For the estimation of hazard ratios in the uni- and multivariable analyses, Cox and Fine-Gray regression models were used. For the calculation of odds ratios in uni- and multivariable analyses, logistic regression was employed. We plotted Kaplan-Meier survival curves and used the log-rank test to explore PFS and OS differences based on the TMTV. Statistical significance was defined as a p value < 0.05.

3 | RESULTS

3.1 | Time to first treatment

After a median follow-up of 6.9 years [95% confidence interval, 6.0–7.8], 66 patients (68%) eventually initiated treatment (Table 1). For all patients, the median TTFT was 3.1 years (95% CI: 2.6–3.6) and the 5-year probability of initiating frontline therapy was 66% (Figure 2A). The reason for starting treatment was lymphoma progression (high tumor burden disease) in all cases, including lymph node growth (49 cases; 80%), development of B symptoms (5 cases; 8%), HT (4 cases; 7%), and other causes [3 cases (5%): Increase in lactate dehydrogenase (LDH) levels, bone involvement, cytopenia].

3.2 | Baseline features and volumetric parameters

Total metabolic tumor volume, TLG and SUVmax were determined for all patients, including three cases in which no tumor mass was

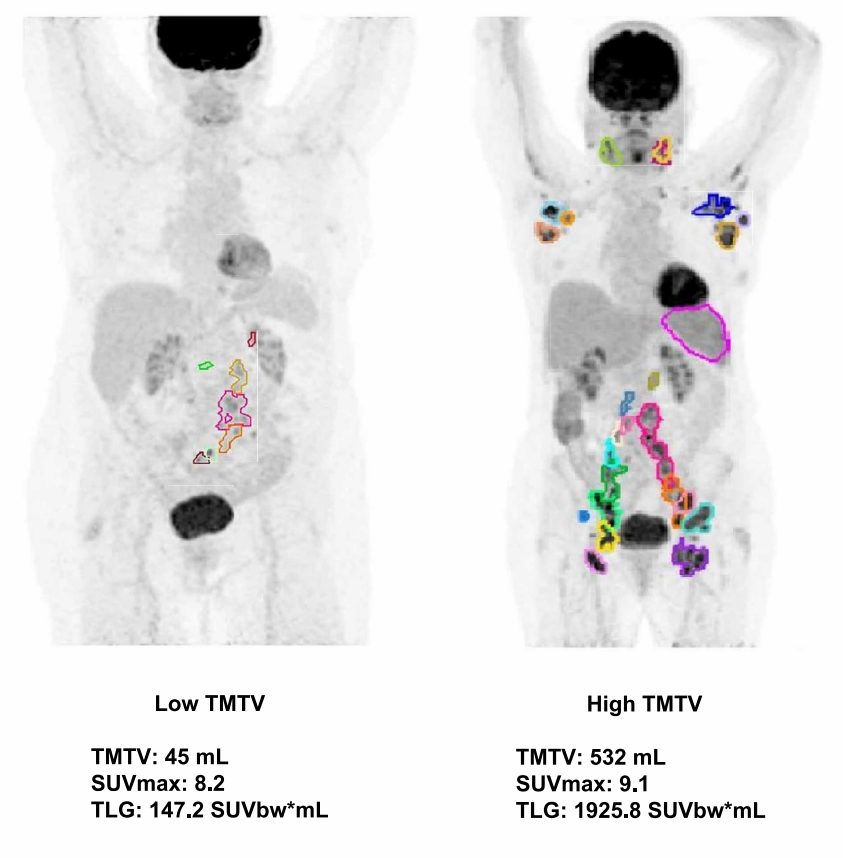

FIGURE 1 Illustrative calculation of the total metabolic tumor volume (TMTV) (semiautomatic segmentation using an standardized uptake value (SUV) ≥2.5 threshold uptake)] in a low tumor burden (LTB), $TMTV^{lo}$ patient (left) and in a LTB, $TMTV^{hi}$ patient (right).

TABLE 1 Time to first treatment (TTFT), frontline therapy modalities, and reasons for initiating therapy, globally and according to the total metabolic tumor volume (TMTV).

		TMTV		
	All patients	$TMTV^{lo}$ (<50 mL)	$TMTV^{hi}$ (>50 mL)	*P*
Median time to first treatment, *y* (95% CI)	3.1 (2.6–3.6)	8.8 (2.5–15.1)	2.6 (1.8–3.4)	**0.001**
Probability of initiating treatment, % at 5 years (95% CI)	66 (54–75)	46 (27–64)	77 (62–86)	
Never treated during follow-up, *n* (%)	31 (32)	17 (51)	14 (22)	**0.005**
Treated during follow-up, *n* (%)	66 (68)	16 (49)	50 (78)	
Immunochemotherapy	51	12	39	0.47
Chemo-free regimens	12	4	8	
Others	3	0	3	
Reason for initiating treatment, *n* (%)[a]				
Lymph node growth	49 (80)	11 (79)	38 (81)	0.46
B symptoms	5 (8)	1 (7)	4 (9)	
Histological transformation	4 (7)	2 (14)	2 (4)	
Others	3 (5)	0	3 (6)	

Note: Statistically significant associations are highlighted in bold.

Abbreviations: CI, confidence interval; FL, follicular lymphoma; TMTV, total metabolic tumor volume.

[a]Only calculated for treated patients.

detectable at the time of imaging. The median TMTV for the entire series was 138.08 mL (range, 0–3027.82), the median SUVmax was 8.85 (range, 0–24.47), and the median TLG was 470.39 SUVbw*mL (range, 0–15,378).

Since TTFT was the endpoint of the study, we dichotomized the radiomic parameters according to their ability to predict TTFT. For TMTV, a cutoff of 53.17 mL was obtained (Supplementary Figure S1). We performed a 100,000-sample bootstrap validation of the *maxstat*-obtained cutoff (mean, 54.99189). This cutoff was then rounded to 50 mL for the sake of practicality and to ensure external validity by avoiding overfitting. This same approach was used for SUVmax and TLG and the resulting cutoff values were 5 (unitless parameter) and 500 SUVbw*mL, respectively. We evaluated whether there was any association between the distribution of TMTV, SUVmax and TLG and the center of origin (Supplementary Table S1), and whether TMTV differed according to the participating center, and we found no significant differences.

Sixty-four patients (66%) had a TMTV above the established cutoff of 50 mL (TMTV^{hi}, Table 2). The distribution of patients according to TMTV, SUVmax and TLG is depicted in Figure 3. SUVmax was the parameter identifying a highest percentage of patients at risk (88% of patients had a SUVmax >5), while 49% of cases had a TLG >500 SUVbw*mL.

As expected, TMTV^{hi} patients had a more advanced stage, more frequent BM involvement by biopsy, more extensive nodal and extranodal disease and higher β_2-microglobulin levels. No differences were seen with regard to age, sex, histological grade, LDH or hemoglobin levels. TMTV^{hi} patients showed a trend toward a higher-risk FLIPI score, although this difference was not statistically significant. Twenty-four patients exhibited splenic involvement: 18 diffuse, 4 focal

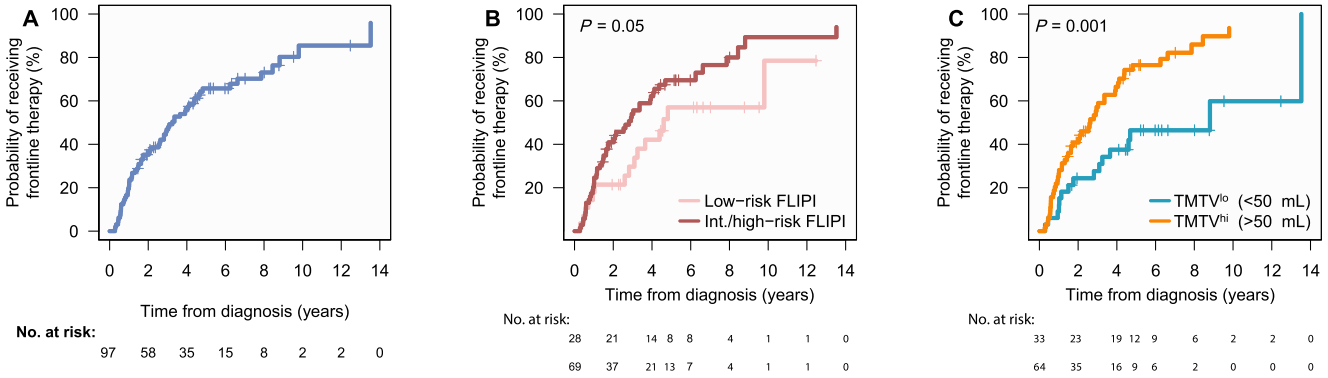

FIGURE 2 Probability of receiving frontline therapy for all patients of the series (A), according to the total metabolic tumor volume (TMTV, B) and to the follicular lymphoma (FL) International Prognostic Index (FLIPI, C).

TABLE 2 Baseline features of the 97 patients with initially observed, low tumor burden (LTB) follicular lymphoma (FL), globally and according to the total metabolic tumor volume (TMTV).

		TMTV		
	All patients (*N* = 97)	**TMTV^{lo} (<50 mL) *n* = 33 (34%)**	**TMTV^{hi} >50 mL *n* = 64 (66%)**	***P***
Age >60 years, *n* (%)	45 (46)	18 (55)	27 (42)	0.25
Female sex, *n* (%)	43 (44)	14 (42)	29 (45)	0.79
Histological grade 1–2, *n* (%)	65 (81)	20 (74)	45 (85)	0.24
Ann-Arbor stage III-IV, *n* (%)[a]	84 (87)	25 (76)	59 (92)	**0.02**
Bone marrow involvement, *n* (%)	62 (65)	15 (48)	47 (73)	**0.02**
≥2 extranodal sites, *n* (%)	26 (27)	3 (10)	23 (36)	**0.007**
>4 lymph node areas, *n* (%)	46 (48)	3 (9)	43 (67)	**<0.001**
Elevated serum LDH, *n* (%)	14 (14)	4 (12)	10 (16)	0.64
Elevated β2-microglobulin, *n* (%)	31 (33)	6 (19)	25 (40)	**0.04**
Hemoglobin <120 g/L, *n* (%)[b]	9 (9)	3 (9)	6 (9)	0.96
High-risk FLIPI[24] score, *n* (%)	23 (24)	4 (12)	19 (30)	0.054
High-risk FLIPI2[25] score, *n* (%)	17 (18)	3 (9)	14 (22)	0.117
High-risk PRIMA-PI[26] score, *n* (%)	17 (18)	3 (9)	14 (22)	0.123

Note: Statistically significant associations are highlighted in bold.

Abbreviations: FLIPI, Follicular Lymphoma International Prognostic Index; LDH, lactate dehydrogenase; PRIMA-PI, PRIMA Prognostic Index; TMTV, total metabolic tumor volume.

[a]The remaining patients had stage II disease.

[b]Anemia, if present, was considered to be non-lymphoma-related.

and 2 mixed (diffuse and focal). Information concerning the evaluation of splenic involvement is displayed in Supplementary Table S2.

Time to first treatment was significantly shorter for TMTV^{hi} (2.6 years, CI: 1.8–3.4) as compared with TMTV^{lo} patients (8.8 years, CI: 2.5–15.1; $p = 0.001$, Table 1 and Figure 2B), with a 5-year probability of initiating treatment of 77% and 46% for TMTV^{hi} and TMTV^{lo} patients, respectively. This difference was also seen in the proportion of patients receiving treatment during follow-up (78 vs. 49% for TMTV^{hi} and TMTV^{lo}, respectively; $p = 0.005$). Of note, the reasons to start therapy were comparable between both groups (lymph node growth in 79% and 81% of TMTV^{lo} and TMTV^{hi} patients, respectively).

To assess baseline features predicting the duration of observation, we built univariable Cox regression models for TTFT (Table 3) and found that an older age was predictive of a lower probability of initiating frontline therapy (HR = 0.59; $p = 0.033$). Factors predicting for a higher likelihood of starting treatment were the presence of ≥2 extranodal sites (HR = 1.69; $p = 0.027$), an intermediate/high-risk FLIPI score (HR = 1.77; $p = 0.05$, Figure 2C), and a TMTV^{hi} (HR = 2.48; $p = 0.0017$).

Considering the statistically significant factors from the univariable analyses, a multivariable model for TTFT was built, excluding the variables included in the FLIPI score, to avoid redundancy. In a model with 94 cases and 66 events, also including extranodal involvement and the FLIPI score, a TMTV^{hi} was the only factor retaining statistical significance (HR = 2.09, CI: 1.14–3.82; $p = 0.017$; Table 3; Supplementary Figure S2). We also investigated the factors predicting the categorical event of needing treatment within five years of diagnosis by means of logistic regression and confirmed that, among 12 clinically relevant variables, TMTV was the only one predicting this endpoint [OR = 4.29 (CI: 1.52–12.66); $p = 0.0067$] (Supplementary Table S3).

We then explored the potential of predicting TTFT of other two imaging parameters, categorized according to previously obtained cutoffs. Both a SUVmax >5 [HR = 3.21 (CI: 1.34–7.70); $p = 0.0089$] and a TLG >500 SUVbw*mL [HR = 2.17 (CI: 1.32–3.57); $p = 0.0022$] anticipated a shorter duration of WW (Supplementary Figure S3), although only SUVmax retained statistical significance in a multivariable model also including the FLIPI score and extranodal

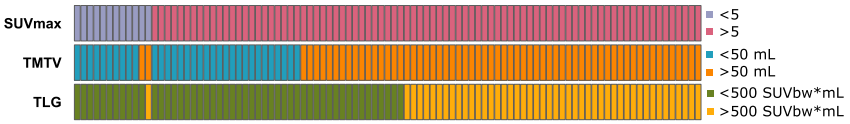

FIGURE 3 Distribution of the patients of the series according to their maximum standardized uptake volume (SUVmax), total metabolic tumor volume (TMTV) and total lesion glycolysis (TLG).

TABLE 3 Univariable and multivariable analyses for the cumulative incidence of receiving frontline therapy using Fine-Gray competing risk regression.

		Cumulative incidence of receiving frontline therapy (94 cases, 66 events)			
		Univariate analysis		Multivariable analysis	
Parameter	Risk category	HR	*P*	HR (CI)	*P*
Sex	Male	1.07 (0.66–1.74)	0.78	NI*	
Age	>60 years	0.59 (0.36–0.96)	**0.033**	NI§	
Histological grade	3A	0.96 (0.46–2.0)	0.92	NI*	
Ann-Arbor stage	III–IV	1.77 (0.80–3.91)	0.16	NI*§	
No. of nodal areas	>4	1.56 (0.97–2.52)	0.069	NI*§	
Extranodal involvement	≥2 sites	1.69 (1.06–2.71)	**0.027**	1.42 (0.88–2.29)	0.15
BM involvement	Present	1.53 (0.89–2.61)	0.12	NI*	
FLIPI score	Int./high risk	1.77 (1.00–3.15)	**0.05**	1.12 (0.60–2.11)	0.72
Hemoglobin	<120 g/L	1.29 (0.65–2.57)	0.47	NI*§	
LDH	Above ULN	0.78 (0.35–1.74)	0.55	NI*§	
B2M	Above ULN	0.99 (0.60–1.64)	0.97	NI*	
TMTV	TMTV^{hi} (>50 mL)	2.48 (1.41–4.37)	**0.0017**	2.09 (1.14–3.82)	**0.017**

Note: Statistically significant findings are highlighted in bold. NI, not included in the multivariable model due to absence of statistical significance in the univariate analysis (*) or to avoid redundancy with the FLIPI score (§).

Abbreviations: B2M, β2-microglobulin; BM, bone marrow; CI, confidence interval; FLIPI, Follicular Lymphoma International Prognostic Index; HR, hazard ratio; LDH, lactate dehydrogenase; TMTV, total metabolic tumor volume; ULN, upper limit of normal.

involvement (HR = 3.40, CI: 1.42–8.15; $p = 0.0062$, Supplementary Table S4), while TLG did not ($p = 0.06$, Supplementary Table S5).

3.3 | Frontline treatment, response, and survival

Most patients (68%) were treated with immunochemotherapy (R-CHOP, R-bendamustine, R-CVP). For the entire cohort, the proportion of patients achieving a complete response after frontline treatment was 69%, without significant differences based on the initial TMTV (Supplementary Table S6). Fifteen patients (16%) died during follow-up, three of which had not received treatment for FL. Five-year PFS and OS estimates were 67% and 91%, respectively. Although a trend toward a lower PFS and OS was seen for patients with a $TMTV^{hi}$ at diagnosis, differences did not reach statistical significance. As expected, the TMTV could not predict survival from treatment (Supplementary Figure S4).

4 | DISCUSSION

Due to the incurable nature and prolonged survival of FL, observation is an acceptable strategy for most patients with advanced-stage, LTB disease.[9] Several motivations lie behind the interest of predicting the duration of WW, such as the psychological tolerance of younger patients. Although some factors and indexes (FLIPI,[24] FLIPI2[25]) anticipate TTFT,[27] predictions remain imperfect. Semiquantitative PET/CT parameters are strong predictors of survival in FL patients in need of treatment,[19,28] but efforts to apply them to LTB patients have been scarce. We evaluated the potential of TMTV, TLG and SUVmax to foresee the duration of observation in 97 patients from two Spanish institutions who did not require treatment[12] at the time of diagnosis.

Two thirds of patients in our cohort had a high TMTV (>50 mL), which was associated with tumor burden features and more extensive nodal and extranodal disease. With a median follow-up of almost 7 years, the median TTFT was 3.1 years, which is in line with previous studies.[4] The main finding of our research was that the median duration of observation was significantly shorter for $TMTV^{hi}$ (2.6 years) as compared with $TMTV^{lo}$ patients (8.8 years). In the multivariable analysis, we found that $TMTV^{hi}$ was the only factor predicting for a shorter TTFT (HR = 2.09), while extranodal involvement and the FLIPI score did not.

Long-term data of the randomized trial comparing single-agent rituximab with WW in LTB patients were recently presented.[7] With a median follow-up of 12.3 years, rituximab monotherapy was highly effective at prolonging time to next treatment, and outcomes with subsequent lines of treatment were not inferior compared with that of patients undergoing initial observation. Our data could help identify a subset of asymptomatic patients (LTB, high TMTV) who could benefit most from single-agent rituximab, although this hypothesis remains to be proven in the setting of prospective clinical trials.

We also analyzed SUVmax and TLG, and found that they can both predict TTFT. Total metabolic tumor volume and SUVmax have been postulated as parameters reflecting different cell compartments. While TMTV best reflects the malignant B-cell burden, intratumoral T cells influence SUVmax, and this can be dependent on the treatment regimen.[29] Due to the small number of patients with an SUVmax <5, the absence of independent impact of TLG on TTFT, the more consolidated role of TMTV in other settings in FL[19] and the contradicting results regarding TLG and SUVmax in previous studies,[30] we focused our analysis on the impact of TMTV.

Two previous small series[20,21] have used PET parameters for TTFT prediction. In both of them, however, cutoffs for such variables were calculated using receiver operating characteristic (ROC) curves, treating the need of therapy as a categorical variable. We believe that using a time-to-event analysis is more correct, since TTFT constitutes a dynamic clinical endpoint. Besides, the Kaplan-Meier method is not entirely appropriate for assessing TTFT, since it disregards cases who died without having received treatment. This can in turn be solved by the calculation of the cumulative incidence of initiating treatment, with competing risks of death.[31]

In the Leccisotti study[21] the median TMTV was much lower than in ours (7.1 vs. 138.08 mL), as was the TMTV cutoff (14 vs. 50 mL), which might be explained by different inclusion criteria (i.e., more stringent criteria to undertake a WW approach) and different segmentation methods (PERCIST instead of SUVmax ≥2.5). That study also showed that TMTV and TLG predicted TTFT independently of FLIPI. However, the presence of extranodal disease, which we consider an important factor guiding treatment initiation, is not accounted for by the FLIPI score nor was it included in the multivariable analysis. A combined FLIPI and TMTV risk-stratification tool was also proposed by the authors. As much as we believe in the potential of radiomics to improve prognostication, we find the proposal of a new score too daring at this time, due to the small cohort size and lack of validation series.

In contrast with other clinical endpoints such as PFS or OS, TTFT (the duration of WW) has the peculiarity of deriving from a clinical decision-making process. The interpretation of the so-called high tumor burden features[10–12] is subject to significant variability among clinicians. Besides, other factors that are not accounted for by those criteria, such as patient preferences, age, or comorbidities, are integrated in the decision of starting treatment or continuing observation. These facts can lead to initially puzzling observations in our cohort, such as an older age (>60 years) being predictive of a longer TTFT (HR = 0.59). This is in all likelihood explained by a greater reluctance to administer therapy to an older, more comorbid individual, and not by a more indolent biological behavior.

The diversity of PET parameters, segmentation algorithms, thresholds and manual contouring methods can be overwhelming. Besides, relevant TMTV cutoffs may significantly differ in various histologies (FL and DLBCL) and clinical situations. For instance, we found a TMTV of 50 mL to be predictive of TTFT in LTB patients, while a TMTV >510 mL anticipated a shorter PFS in the Meignan study,[19] in which all patients had a high tumor burden. We believe

that the definitive incorporation of radiomics into lymphoma prognostication calls for international standardization and a solid methodological consensus for each clinical scenario.

One of the considerations regarding PET parameters is whether they substantially improve the prognostic information provided by CT scan data alone. Some risk scores include the extension of nodal involvement measured by CT, in the form of the number[24] or size[19,25,32] of involved lymph node areas. Indeed, the GELF criteria only consider lymph node size to recommend treatment. In our view, tumor volume measurement using PET/CT has clear advantages over morphological imaging techniques, especially in the case of lesions that are not measurable by CT (e.g., spleen infiltration without splenomegaly, and bone, pleural or peritoneal infiltration), where delineating tumor contours becomes challenging, due to the contiguity with vascular, nervous and muscular structures (Supplementary Figure S4).

We have to acknowledge several limitations of our work. First, the number of patients is modest. Second, the retrospective nature of the study makes it vulnerable to inherent flaws. Third, due to the lack of validation cohort, TMTV cutoff definition might be subject to overfitting. Fourth, the clinical application of PET calculations in clinical practice might not be straightforward, since it is time-consuming. Lastly, although decisions were taken in a multidisciplinary team, using similar criteria for the past 20 years, we cannot deny the subjectivity of deciding when to initiate treatment. Despite all that, our data arise from a well-annotated clinical database, with a mature follow-up, employing semiquantitative PET measurements performed by two independent nuclear medicine physicians and robust statistical methods for cutoff calculation and TTFT analysis.

It could be argued that by limiting the patients included in our study to those not fulfilling the GELF criteria, which are themselves a measure of tumor burden and were empirically defined, might diminish the relevance of our conclusions. However, we focused on this subset of cases in order to identify patients without any of the classical high tumor burden features who might not benefit from WW for a long time. As mentioned before, the decision of initiating treatment derives not only from tumor burden features, but also personal factors from the physician and the patient.

In our exploratory study, we found that a high TMTV is a strong independent predictor of the duration of WW in initially observed, LTB FL patients. Although we failed to find a TMTV threshold identifying a subset of patients with an extremely low long-term probability of requiring treatment, we did recognize a third of LTB FL patients with a low TMTV who had a median treatment-free survival beyond 8 years. Upon the validation of our cutoff and the standardization of segmentation methods, the information provided by PET/CT could become an additional factor to consider deferring or initiating treatment (such as single agent rituximab) in asymptomatic patients.

AUTHOR CONTRIBUTIONS

Pablo Mozas, Sebastian Casanueva-Eliceiry, Xavier Setoain and Armando López-Guillermo conceived the study, performed the analyses and wrote the manuscript. Sebastian Casanueva-Eliceiry, Marc Simó, Sonia Rodríguez and Xavier Setoain performed the PET and CT measurements. Andrea Rivero, Ángel Serna, Alfredo Rivas-Delgado, Ferran Nadeu, Juan Gonzalo Correa, Juan Antonio Piñeyroa, Amanda Isabel Pérez-Valencia, Katia Guinetti-Ortiz, Marta Gómez-Hernando, Eva Giné, Julio Delgado, Neus Villamor, Elías Campo, Laura Magnano, Pau Abrisqueta contributed to data collection. All authors reviewed and approved the final version of the manuscript.

ACKNOWLEDGMENTS

This study was supported by Fondo de Investigaciones Sanitarias, Instituto de Salud Carlos III and European Regional Development Fund "Una manera de hacer Europa" [grant numbers: PI19/00925 (LM), PI19/00887 (ALG and EG)].

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

All clinical and imaging data are available upon reasonable request.

ORCID

Pablo Mozas https://orcid.org/0000-0001-9528-4971
Sebastian Casanueva-Eliceiry https://orcid.org/0000-0002-5539-5850

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/hon.3235.

REFERENCES

1. Casulo C, Byrtek M, Dawson KL, et al. Early relapse of follicular lymphoma after rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone defines patients at high risk for death: an analysis from the National LymphoCare Study. *J Clin Oncol*. 2015;33(23):2516-2522. https://doi.org/10.1200/JCO.2014.59.7534
2. Alonso-Álvarez S, Magnano L, Alcoceba M, et al. Risk of, and survival following, histological transformation in follicular lymphoma in the rituximab era. A retrospective multicentre study by the Spanish GELTAMO group. *Br J Haematol*. 2017;178(5):699-708. https://doi.org/10.1111/bjh.14831
3. Bachy E, Seymour JF, Feugier P, et al. Sustained progression-free survival benefit of rituximab maintenance in patients with follicular lymphoma: long-term results of the PRIMA study. *J Clin Oncol*. 2019;37(31):2815-2824. https://doi.org/10.1200/JCO.19.01073
4. Ardeshna KM, Qian W, Smith P, et al. Rituximab versus a watch-and-wait approach in patients with advanced-stage, asymptomatic, non-bulky follicular lymphoma: an open-label randomised phase 3 trial. *Lancet Oncol*. 2014;15(4):424-435. https://doi.org/10.1016/S1470-2045(14)70027-0
5. Ardeshna KM, Smith P, Norton A, et al. Long-term effect of a watch and wait policy versus immediate systemic treatment for asymptomatic advanced-stage non-Hodgkin lymphoma: a randomised controlled trial. *Lancet*. 2003;362(9383):516-522. https://doi.org/10.1016/S0140-6736(03)14110-4
6. Solal-Céligny P, Bellei M, Marcheselli L, et al. Watchful waiting in low-tumor burden follicular lymphoma in the rituximab era: results of an F2-study database. *J Clin Oncol*. 2012;30(31):3848-3853. https://doi.org/10.1200/JCO.2010.33.4474

7. Northend M, Wilson W, Clifton-Hadley L, et al. Long term follow-up of international randomised phase 3 study of rituximab versus a watch and wait approach for patients with asymptomatic, low tumour burden follicular lymphoma shows rituximab is highly effective at delaying time to new treatment without. In: *ASH*; 2022. Accessed 3 December 2022. https://ash.confex.com/ash/2022/webprogram/Paper156790.html
8. Mozas P, Rivero A, Rivas-Delgado A, et al. Age and comorbidity are determining factors in the overall and relative survival of patients with follicular lymphoma. *Ann Hematol*. 2021;100(5):1231-1239. https://doi.org/10.1007/s00277-021-04470-7
9. Dreyling M, Ghielmini M, Rule S, et al. Newly diagnosed and relapsed follicular lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2021;32(3):298-308. https://doi.org/10.1016/j.annonc.2020.11.008
10. Martinelli G, Schmitz SFH, Utiger U, et al. Long-term follow-up of patients with follicular lymphoma receiving single-agent rituximab at two different schedules in trial SAKK 35/98. *J Clin Oncol*. 2010;28(29):4480-4484. https://doi.org/10.1200/JCO.2010.28.4786
11. McNamara C, Davies J, Dyer M, et al. Guidelines on the investigation and management of follicular lymphoma. *Br J Haematol*. 2012;156(4):446-467. https://doi.org/10.1111/j.1365-2141.2011.08969.x
12. Brice P, Bastion Y, Lepage E, et al. Comparison in low-tumor-burden follicular lymphomas between an initial no-treatment policy, prednimustine, or interferon alfa: a randomized study from the Groupe d'Etude des Lymphomes Folliculaires. *J Clin Oncol*. 1997;15(3):1110-1117. https://doi.org/10.1200/JCO.1997.15.3.1110
13. El-Galaly TC, Bilgrau AE, de Nully Brown P, et al. A population-based study of prognosis in advanced stage follicular lymphoma managed by watch and wait. *Br J Haematol*. 2015;169(3):435-444. https://doi.org/10.1111/bjh.13316
14. Nastoupil LJ, Sinha R, Byrtek M, et al. Outcomes following watchful waiting for stage II-IV follicular lymphoma patients in the modern era. *Br J Haematol*. 2016;172(5):724-734. https://doi.org/10.1111/bjh.13895
15. Yuda S, Miyagi Maeshima A, Taniguchi H, et al. Clinicopathological factors and tumor microenvironment markers predicting watch-and-wait discontinuation in 82 patients with follicular lymphoma. *Eur J Haematol*. 2021;107(1):157-165. https://doi.org/10.1111/ejh.13637
16. Cheson BD, Fisher RI, Barrington SF, et al. Recommendations for initial evaluation, staging, and response assessment of hodgkin and non-hodgkin lymphoma: the lugano classification. *J Clin Oncol*. 2014;32(27):3059-3067. https://doi.org/10.1200/JCO.2013.54.8800
17. Luminari S, Biasoli I, Arcaini L, et al. The use of FDG-PET in the initial staging of 142 patients with follicular lymphoma: a retrospective study from the FOLL05 randomized trial of the Fondazione Italiana Linfomi. *Ann Oncol*. 2013;24(8):2108-2112. https://doi.org/10.1093/annonc/mdt137
18. Delfau-Larue MH, Van Der Gucht A, Dupuis J, et al. Total metabolic tumor volume, circulating tumor cells, cell-free DNA: distinct prognostic value in follicular lymphoma. *Blood Adv*. 2018;2(7):807-816. https://doi.org/10.1182/bloodadvances.2017015164
19. 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;34(30):3618-3626. https://doi.org/10.1200/JCO.2016.66.9440
20. Yang Q, Luo Y, Zhang Y, Zhang W, Zhou D, Li F. Baseline [18F]FDG PET/CT may predict the outcome of newly diagnosed follicular lymphoma in patients managed with initial "watch-and-wait" approach. *Eur Radiol*. 2022;32(8):5568-5576. https://doi.org/10.1007/s00330-022-08624-7
21. Leccisotti L, Maccora D, Malafronte R, et al. Predicting time to treatment in follicular lymphoma on watchful waiting using baseline metabolic tumour burden. *J Cancer Res Clin Oncol*. 2022;149(7):1-9. Published online July 2. https://doi.org/10.1007/s00432-022-04138-3
22. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol*. 2007;25(5):579-586. https://doi.org/10.1200/JCO.2006.09.2403
23. Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat*. 2007;16(3):1141-1154. https://doi.org/10.1214/aos/1176350951
24. Solal-Céligny P, Roy P, Colombat P, et al. Follicular lymphoma international prognostic index. *Blood*. 2004;104(5):1258-1265. https://doi.org/10.1182/blood-2003-12-4434
25. Federico M, Bellei M, Marcheselli L, et al. Follicular lymphoma international prognostic index 2: a new prognostic index for follicular lymphoma developed by the international follicular lymphoma prognostic factor project. *J Clin Oncol*. 2009;27(27):4555-4562. https://doi.org/10.1200/JCO.2008.21.3991
26. Bachy E, Maurer MJ, Habermann TM, et al. A simplified scoring system in de novo follicular lymphoma treated initially with immunochemotherapy. *Blood*. 2018;132(1):49-58. https://doi.org/10.1182/blood-2017-11-816405
27. Mozas P, Rivero A, Rivas-Delgado A, et al. Prognostic ability of five clinical risk scores in follicular lymphoma: a single-center evaluation. *Hematol Oncol*. 2021;39(5):639-649. https://doi.org/10.1002/hon.2922
28. Cottereau AS, Versari A, Luminari S, et al. Prognostic model for high-tumor-burden follicular lymphoma integrating baseline and end-induction PET: a LYSA/FIL study. *Blood*. 2018;131(22):2449-2453. https://doi.org/10.1182/blood-2017-11-816298
29. Nath K, Law SC, Sabdia MB, et al. Intratumoral T cells have a differential impact on FDG-PET parameters in follicular lymphoma. *Blood Adv*. 2021;5(12):2644-2649. https://doi.org/10.1182/bloodadvances.2020004051
30. Cottereau AS, Versari A, Chartier L, et al. Low Suvmax measured on baseline FDG-PET/CT and elevated β2 microglobulin are negative predictors of outcome in high tumor burden follicular lymphoma treated by immunochemotherapy: a pooled analysis of three prospective studies. *Blood*. 2016;128(22):1101. https://doi.org/10.1182/blood.v128.22.1101.1101
31. Delgado J, Pereira A, Villamor N, López-Guillermo A, Rozman C. Survival analysis in hematologic malignancies: recommendations for clinicians. *Haematologica*. 2014;99(9):1410-1420. https://doi.org/10.3324/haematol.2013.100784
32. Mir F, Mattiello F, Grigg A, et al. Follicular Lymphoma Evaluation Index (FLEX): a new clinical prognostic model that is superior to existing risk scores for predicting progression-free survival and early treatment failure after frontline immunochemotherapy. *Am J Hematol*. 2020;95(12):1503-1510. https://doi.org/10.1002/ajh.25973

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Mozas P, Casanueva-Eliceiry S, Rivero A, et al. A low total metabolic tumor volume independently predicts for a longer time to first treatment in initially observed, low tumor burden follicular lymphoma. *Hematol Oncol*. 2024;e3235. https://doi.org/10.1002/hon.3235