RESEARCH ARTICLE

Prognostic value of baseline metabolic tumor volume and total lesion glycolysis in patients with lymphoma: A meta-analysis

Baoping Guo◉, **Xiaohong Tan**◉, **Qing Ke**◉, **Hong Cen**◉ *

Department of Chemotherapy, Guangxi Cancer Hospital and of Guangxi Medical University Affiliated Cancer Hospital, Nanning, Guangxi, People's Republic of China

◉ These authors contributed equally to this work.
* cenhong01@163.com

Abstract

Whether baseline metabolic tumor volume (TMTV) and total lesion glycolysis (TLG) measured by FDG-PET/CT affected prognosis of patients with lymphoma was controversial. We searched PubMed, EMBASE and Cochrane to identify studies assessing the effect of baseline TMTV and TLG on the survival of lymphoma patients. Pooled hazard ratios (HR) for overall survival (OS) and progression-free survival (PFS) were calculated, along with 95% confidence intervals (CI). Twenty-seven eligible studies including 2,729 patients were analysed. Patients with high baseline TMTV showed a worse prognosis with an HR of 3.05 (95% CI 2.55–3.64, p<0.00001) for PFS and an HR of 3.07 (95% CI 2.47–3.82, p<0.00001) for OS. Patients with high baseline TLG also showed a worse prognosis with an HR of 3.44 (95% CI 2.37–5.01, p<0.00001) for PFS and an HR of 3.08 (95% CI 1.84–5.16, p<0.00001) for OS. A high baseline TMTV was significantly associated with worse survival in DLBCL patients treated with R-CHOP (OS, pooled HR = 3.52; PFS, pooled HR = 2.93). A high baseline TLG was significantly associated with worse survival in DLBCL patients treated with R-CHOP (OS, pooled HR = 3.06; PFS, pooled HR = 2.93). The negative effect of high baseline TMTV on PFS was demonstrated in HL (pooled HR = 3.89). A high baseline TMTV was significantly associated with worse survival in ENKL patients (OS, pooled HR = 2.24; PFS, pooled HR = 3.25). A high baseline TLG was significantly associated with worse survival in ENKL patients (OS, pooled HR = 2.58; PFS, pooled HR = 2.99). High baseline TMTV or TLG predict significantly worse PFS and OS in patients with lymphoma. Future studies are warranted to explore whether TMTV or TLG could be integrated into various prognostic models for clinical decision making.

Introduction

Lymphoma continues to be the most common form of hematological malignancy worldwide [1,2]. Lymphoma is a heterogeneous group of biologically and clinically distinct neoplasms and have been historically divided into two distinct categories: non-Hodgkin lymphoma

(NHL) and Hodgkin lymphoma (HL) [3]. Although major progress has been made in the treatment of patients with lymphoma, many still fail to achieve a response or subsequently relapse [4]. These patients are not easily identified by existing pretreatment prognostic indexes such as the IPI (international prognostic index), IPS (international prognostic score for Hodgkin lymphoma [HL]), FLIPI (prognostic score for follicular lymphoma), MIPI (prognostic score for mantle cell lymphoma), and PIT (prognostic index for peripheral T cell lymphoma) or by conventional computed tomography (CT)–based response assessment [5–9]. Therefore, there is an urgent need for new prognostic and predictive markers which permit an accurate and early identify high-risk patient categories.

[18]Fluorine fluorodeoxyglucose-positron emission tomography/computed tomography (FDG-PET/CT) has been recognized by the 2014 International Conference on Malignant Lymphoma imaging consensus guidelines as the standard imaging modality to evaluate glucose metabolism in fluorodeoxyglucose (FDG)-avid lymphoma tumors [10,11]. Its value for prognosis prediction at interim and end treatment has been recently investigated [12–14]. Many studies have also showed that quantitative volumetric parameters derived from baseline ^{18}F-FDG PET such as total metabolic tumor volume (TMTV) or total lesion glycolysis (TLG) could predict outcome in diffuse large B-cell lymphoma [15–17], in follicular lymphoma [18], in peripheral T-cell lymphoma [19], in extranodal natural killer/T-cell lymphoma [20] and in Hodgkin lymphoma [21,22]. However, those studies evaluating the prognostic values of pretherapy TMTV and TLG in in patients with various lymphoma subtypes showed inconclusive and contradictory results [23].

Therefore, the purpose of this meta-analysis was to evaluate the prognostic value of baseline TMTV or TLG by PET/CTin patients with lymphoma, in order to provide more evidence of their clinical value as prognostic biomarkers.

Materials and methods

Inclusion and exclusion criteria

A computerized search of PubMed, Embase and Cochrane was conducted to find relevant studies published prior to May 01, 2018. The following search terms were used: ("lymphoma"[-MeSH Terms] OR lymphom*[All Fields] OR lymphoproliferative [All Fields] OR hodgkin* [All Fields] OR non-hodgkin* [All Fields]) AND ("Tomography, emission-computed"[MeSH Terms] OR ("positron emission tomograpy"[MeSH Terms]) OR (computed [All Fields] AND tomograph*[All Fields])) AND (prognos* OR predict* OR surviv* OR overall survival* OR recurrence* OR progress*). All searches were limited to human studies.

Eligible studies met the following criteria: (i) They were observational studies (retrospective or prospective) or clinical trials, (ii) the studies were limited to lymphoma, (iii) ^{18}F-FDG PET was used as an initial imaging tool, (iv) patients had not undergone chemotherapy, immune-chemotherapy or radiotherapy before the ^{18}F-FDG PET scan, (v) the volume of the lymphoma was measured, (vi) the survival data was reported. Studies were excluded if: (i) they were case reports, case series, review articles, editorials, letters or comments; (ii) the patient survival data was unavailable or insufficient to perform the meta-analysis, (iii) the data included was specifically for HIV-associated lymphoma, pediatric lymphoma, primary central nervous system lymphoma, primary testicular lymphoma, or primary mediastinal large B-cell lymphoma, or (iv) they included overlapping patients and data. Two reviewers (B.P. Guo and Q. Ke) independently selected the literature using a standardized protocol. Disagreements were resolved by discussion.

Quality assessment

The methodological quality of the primary manuscripts was independently evaluated by two reviewers (B.P. Guo and X.H. Tan) by means of the Newcastle-Ottawa-Scale (NOS) [24], which is used for the quality assessment of cohort and case-control studies. The NOS comprises three quality parameters: selection (0–4 points), comparability (0–2 points), and outcome assessment (0–3 points). Studies with scores of six or more were considered to be of high quality. Any disparities between investigators were resolved by discussion. Study quality was not an exclusion criterion.

Data extraction

Data extraction was carried out by B.P Guo and independently confirmed by the other authors (X.H. Tan and H. Cen). The collected data included the following: Study characteristics: first author, year of publication, country, study design, imaging modalities, type of lymphoma, number of patients, treatment, tumor volume parameters (maximum threshold for PET volume auto-segmentation, and median MTV/TLG), MTV/TLG cut-off values, determination of MTV cut-off, median follow-up, and endpoints. Extracted data were entered onto a standardized Excel file (Microsoft Corporation). Discrepancies were resolved by discussion with coauthors.

We chose OS and PFS as endpoints for our meta-analysis. Overall survival is defined as the length of time from either the date of diagnosis or the date of recruitment in a study to the moment of death as a result of any cause. PFS is defined as the length of time from either the date of diagnosis or the date of recruitment in a study until lymphoma progression or death as a result of any cause.

Statistical analysis

The impact of MTV or TLG on survival was measured by estimating the effect size of the hazard ratios (HR). Pooled HRs of more than 1.00 indicated poor survival in the group with high MTV or TLG values when compared with the group with low values. For studies in which the HRs and CIs were not available, we used the method proposed by Parmar et al. [25] to derive estimates from survival curves. The point estimate of the HR was considered statistically significant at the $p < 0.05$ level if the 95% CI did not include the value 1.

Heterogeneity was assessed by means of Cochran Q and I^2 statistics. I-square (I^2) values of <30%, 30%-50%, 50%-75% and >75% were defined as low, moderate, substantial and considerable heterogeneity, respectively [26]. If heterogeneity existed between primary studies, a random effects model was used. Otherwise, a fixed effects model was used for the meta-analysis. Publication bias was assessed by visual inspection of the funnel plot and also by means of the Begg and Egger tests [27,28]. A p-value less than 0.05 indicates the existence of publication bias. All statistical analyses were performed by using RevMan 5.3 (Nordic Cochrane Centre).

Results

Literature search procedure

The PRISMA statement flowchart shows the process of literature screening and selection, as well as the reasons for exclusion (Fig 1). Our initial search yielded 5045 articles. After removing duplicates and screening the titles and abstracts, 45 articles were reviewed in more detail. Of the 45 full-text studies, 18 were excluded for the following reasons: 7 studies had incomplete or unavailable data and other topics; 5 reports were reviews and editorials; 1 study involved pediatric lymphoma; 3 studies included primary extranodal lymphoma and relapsed or
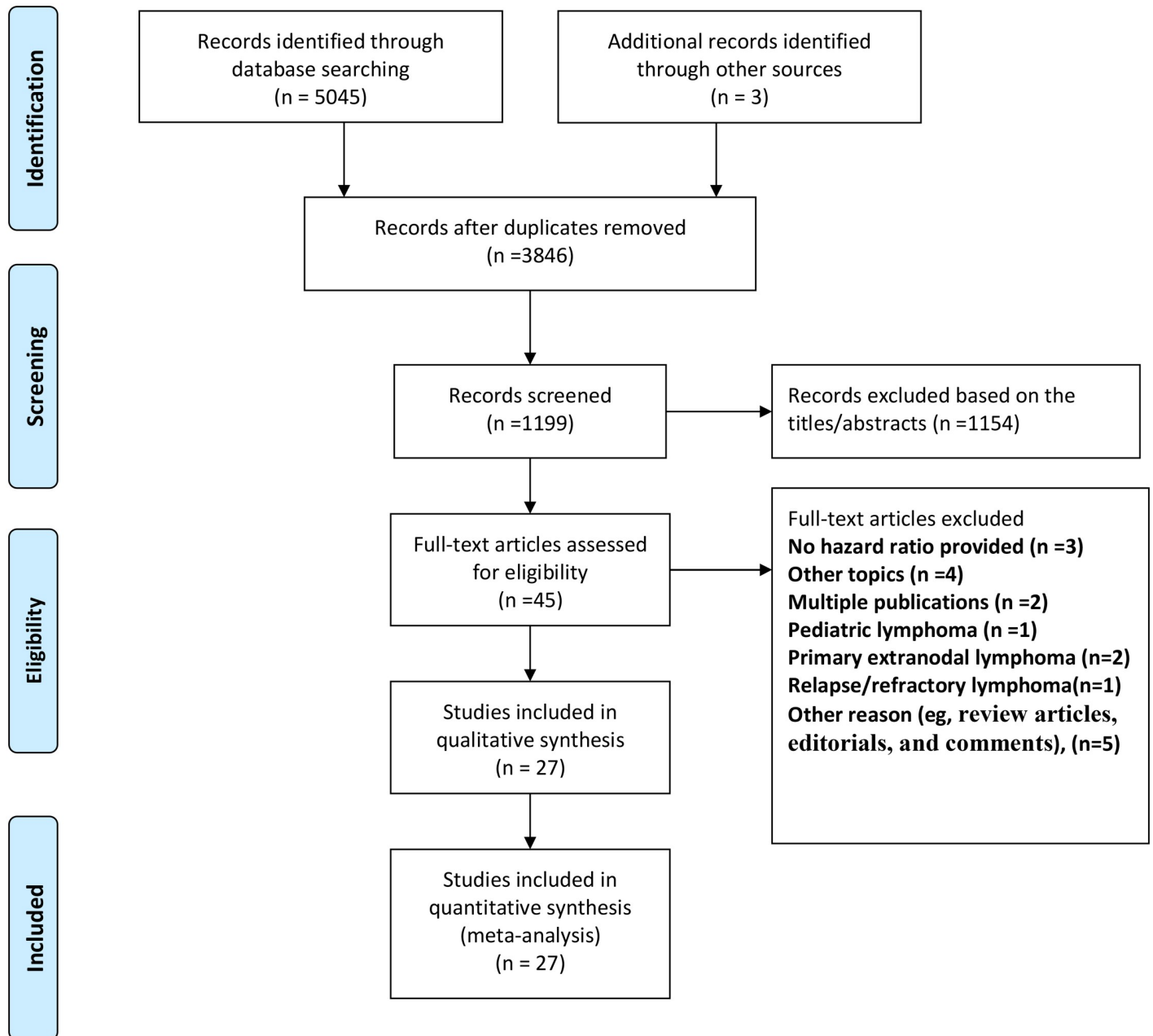
Fig 1. Flow diagram of the systematic review and meta-analysis process.

refractory lymphoma; 2 studies were multiple publications. After reviewing the full texts, 27 studies were finally selected as potentially appropriate for inclusion in the meta-analysis [15–22,29–47].

Study characteristics

The 27 observational studies fulfilled the inclusion criteria, were published between 2012 and 2018 and are summarized in Table 1. Twenty-five studies were retrospective observational studies, and three studies were prospective multicenter trials. Seventeen studies included

Table 1. Characteristics of studies included in the meta-analysis.

Study	Year	Country	Study Design	NOS	Type of Lymphoma	Patients (No.)	Treatment regimen	Threshold (%)	Median MTV (cm³)	Median TLG	MTV (cm³)	TLG	Determination of MTV Cut off	Endpoints	Median follow-up
Song et al. [a]	2012	Korea	R	8	DLBCL	169	R-CHOP	≥2.5	198.1	NR	220	NR	ROC analysis	PFS/OS	36 months
Manohar et al.	2012	India	R	6	NHL[§]	51	R-CHOP-like	Background-level[†]	957	5356	416	3340	ROC analysis	PFS/OS	12 months
Kim et al.	2013	Korea	R	8	DLBCL	140	R-CHOP	Various[‡]	NR	415.5	NR	415.5	N/A	PFS/OS	28.5 months
Oh et al.	2013	Korea	R	6	DLBCL	181	R-CHOP	≥2.5	156.89	NR	65.975	NR	ROC analysis	PFS/OS	NR
Song et al.	2013	Korea	R	7	HL	127	ABVD	≥2.5	142.6	NR	198	NR	ROC analysis	PFS/OS	45.8 months
Kim et al.	2013	Korea	R	5	ENKTL	20	CTx follow RT/Only CTx	NR	10.7	46.9	14.4	52.7	ROC analysis	PFS/OS	26.3 months
Esfahani et al.	2013	USA	R	7	DLBCL	20	R-CHOP	50	379.16	704.77	379.16	704.77	ROC analysis	PFS	mean 51.35 months
Sasanelli et al.	2014	France	R	9	DLBCL	114	R-CHOP/R-ACVBP	41	315	2974	550	4,576	ROC analysis	PFS/OS	39 months
Gallicchio et al.	2014	Italy	R	5	DLBCL	52	R-CHOP-like	42	43	596.9	16.1	589.5	ROC analysis	EFS/OS	18 months
Kim et al.	2014	Korea	R	5	DLBCL	96	R-CHOP	≥2.5	130.7	NR	130.7	NR	ROC analysis	EFS/OS	27.8 months
Adams et al.	2015	Netherland	R	9	DLBCL	73	R-CHOP	40	272	2955.4	272	2955.4	ROC analysis	PFS/OS	2.7years
Schoder et al.	2015	USA	P	6	DLBCL	65	R-CHOP	Various[‡]	226	NR	NR	NR	ROC analysis	PFS/OS	51 months
Kanoun et al.	2015	France	R	8	HL	59	CTx ± RT	Various[‡]	160	NR	313	NR	ROC analysis	PFS/OS	39 months
Mikhaeel et al.	2016	UK	R	8	DLBCL	147	R-CHOP	41	595	4669.5	396	4541	ROC analysis	PFS	3.8 years
Zhou et al.	2016	China	R	8	DLBCL	91	R-CHOP	Background-level[†]	50.7	497.3	PFS: 70 OS: 78	PFS: 827 OS: 726	ROC analysis	PFS/OS	30 months
Song et al. [b]	2016	Korea	R	6	DLBCL	107	R-CHOP	≥2.5	526.8	NR	601.2	NR	ROC analysis	PFS/OS	40.8 months
Cottereau et al.	2016	France	R	8	PTCL	108	CHOP-like/ACVBP	41	224	1155	230	1068	ROC analysis	PFS/OS	23 months
Meignan et al.	2016	France	R	9	FL 1-3a	185	R-CHOP/R-CVP/R-FM	41	297	NR	510	NR	X-tile analysis	PFS/OS	64 months
Chang et al.	2017	China	R	6	ENKTL	52	DDGP/SMILE	40	11.2	46.4	16.1	44.7	ROC analysis	PFS/OS	19 months
Chang et al.	2017	Taiwan	R	7	DLBCL	118	R-CHOP	≥2.5	550.4*	3533.2*	165.4	1204.9	ROC analysis	PFS/OS	28.7 months

(Continued)

Table 1. (Continued)

Study	Year	Country	Study Design	NOS	Type of Lymphoma	Patients (No.)	Treatment regimen	Tumor Volume Parameters(MTV/TLG)			Cut-off values		Determination of MTV Cut off	Endpoints	Median follow-up
								Threshold (%)	Median MTV (cm³)	Median TLG	MTV (cm³)	TLG			
Kesavan et al.	2017	Australia	P	7	FL	68	Iodine-131-rituximab	41	510	NR	510	NR	NR	TTNT/OS	59 months
Song et al.	2017	Korea	R	5	ENKTL	100	CTx follow RT/CTx	≥2.5	36.2	NR	94.2	NR	ROC analysis	PFS/OS	NR
Cottereau et al.	2018	France	P	9	HL	258	ABVD	41	67	332	147	495	X-tile and ROC analysis	PFS/OS	55 months
Ding et al.	2018	China	R	7	DLBCL	72	R-CHOP	40	139.48	1413.77	67.71	1413.77	ROC analysis	PFS/OS	45 months
Toledano et al.	2018	France	R	8	DLBCL	114	R-CHOP/R-CHOP-like	41	275.8	NR	261.4	1325.80	ROC analysis	PFS/OS	40 months
Delfau-Larue et al.	2018	France	R	8	FL	133	R-CHOP/R-CHOP-like/Chemo-free	41	354	NR	510	NR	X-tile and ROC analysis	PFS/OS	48 months
Pak et al.	2018	Korea	R	7	ENKTL	36	NR	40	NR	NR	7	45.8	NR	RFS/OS	20.6 months

PET/CT, positron emission tomography/computed tomography; MTV, metabolic tumor volume; TLG, total lesion glycolysis; NOS, the Newcastle-Ottawa-Scale; DLBCL, diffuse large B cell lymphoma; HL, hodgkin lymphoma; NHL, non-hodgkin's lymphoma; ENKTL, extranodal natural killer/T cell lymphoma; PTCL, peripheral T-cell lymphoma; FL, follicular Lymphoma; R, retrospective; P, prospective; NR, not reported; PFS, progression-free surviva; TTNT:time-to-next-treatment; IQR, interquartile range; ROC, receiver operator curve; N/A, not applicable; R-CHOP: rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone; ABVD, doxorubicin,bleomycin, vinblastine, dacarbazine; CTx, chemotherapy; RT, radiotherapy; R-ACVBP: rituximab, doxorubicine, vindesine, cyclophosphamide, bleomycin, prednisolone; R-ICE, rituximab, etoposide, ifosfamide, carboplatin; R-CVP, rituximab, cyclophosphamide, vincristine and prednisolone; R-FM, rituximab, fludarabine and mitoxantrone; DDGP, dexamethasone, cisplatin, gemcitabline and pegaspargase; SMILE, dexamethasone, methotrexate, ifosfamide, L-asparaginase and etoposide;

‡ tested various proposed thresholds, including 41%;

† MTV was measured by setting the tumor marginal threshold of liver SUVmean plus 3SDs. SUVmean in liver was calculated in a standard-sized ROI of 3cm in diameter;

§ Of the 51 patients, 39 (77%) had DLBCL, 8 had anaplastic large T-cell lymphoma and 4 had high-grade peripheral T-cell lymphoma.

ᵃ, In staged II and III patients without extranodal site involvement; b, In patients with bone marrow involvement of lymphoma.

*The mean values of total MTV and TLG were 550.4 ± 678.3 cm3 and 3533.2 ± 4394.1 cm3 respectively.

patients with diffuse large B cell lymphoma, three included patients with follicular lymphoma, one included patients with peripheral T-cell lymphoma, four studies included patients with extranodal natural killer/T cell lymphoma, and three studies included patients with Hodgkin lymphoma. The total sample size was 2729. Either TMTV or TLG was measured in 12 studies, and both were measured in 15 studies.

Three thresholding methods for the auto segmentation of PET volumes exist. A fixed SUV of 2.5 was used in 7 studies, the percentage of SUVmax (40%, 41%, 42% or 50%) was used in 18 studies, and in 2 studies the MTV was measured by setting the tumor margin threshold as the liver SUVmean plus 3SDs. In each study, patients were divided into 2 groups (high and low volume) based on the cut-off values. The MTV and TLG cut-off values were determined by means of receiver operating curve (ROC) and X-tile analyses. Receiver operating characteristics (ROCs) was used in 19 studies, receiver-operating characteristics and X-tile analysis in 3 studies, X-tile analysis in 1 study, and 4 studies did not provide cut-off information. The MTV cut-off values ranged between 10.7 and 595 cm^3 and the TLG values ranged from 46.4 to 5356. The study quality assessed by means of the NOS was fair, with a median quality score of 8 (range 5–9).

Prognostic value of baseline metabolic tumor volume

Twenty-one studies [15–19,21,22,29,30–35,37,39,43–47] on adult lymphoma were included for the analysis of the relationship between TMTV and PFS. The results of our meta-analysis showed that a high TMTV was associated with a shorter PFS than a low TMTV, with a pooled HR of 3.05 (95% CI, 2.55–3.64, $p<0.00001$). No significant heterogeneity was found across studies ($I^2 = 0.0\%$, $P = 0.78$). Twenty-one studies [15,17–20,22,29–35,38–40,43–47] reported data on TMTV and OS in adult lymphoma. The meta-analysis results demonstrated that a high TMTV was associated with shorter OS than a low TMTV, with a pooled HR of 3.07 (95% CI, 2.47–3.82, $p<0.00001$). No significant heterogeneity was found across studies ($I^2 = 0.0\%$, $p = 0.60$; Fig 2).

Next, we examined the relationship between TMTV and the clinical outcome of different types of lymphomas. A meta-analysis of thirteen studies of DLBCL patients showed poorer PFS and OS in patients with high TMTV than in those with low TMTV, with pooled HRs for PFS and OS of 2.93 (95%CI, 2.29–3.73, $p<0.001$; heterogeneity: $I^2 = 0.0\%$, $p = 0.481$) and 3.52 (95%CI, 2.67–4.64, $p<0.001$; heterogeneity: $I^2 = 0.0\%$, $p = 0.806$), respectively. Three studies examined the prognostic significance of high TMTV in FL patients, and the pooled HRs for PFS and OS were 2.55 (95%CI, 1.65–3.92, $p<0.001$; heterogeneity: $I^2 = 0.0\%$, $p = 0.622$) and 2.89 (95%CI, 1.04–7.99, $p<0.001$; heterogeneity: $I^2 = 0.0\%$, $p = 0.419$), respectively. Four studies assessed the prognostic significance of high TMTV in ENKL patients, and the pooled HRs for PFS and OS were 3.25 (95%CI, 1.75–6.07, $p<0.001$; heterogeneity: $I^2 = 10.7\%$, $p = 0.340$) and 2.24 (95%CI, 1.23–4.08, $p = 0.008$; heterogeneity: $I^2 = 24\%$, $p = 0.267$), respectively. Two studies investigated the prognostic significance of high TMTV in HL patients, and the pooled HR for PFS was 3.89 (95%CI, 2.19–6.90, $p<0.001$; heterogeneity: $I^2 = 0.0\%$, $p = 0.646$) (Table 2). Notably, only one study provided relevant data on the correlation between TMTV and PTCL outcome and only one study provided data on the correlation between TMTV and OS in HL patients; therefore, the pooled analysis could not be performed.

We also conducted subgroup analyses stratified by data collection method, sample size, and different threshold values. The subgroup analysis of retrospectively collected data showed pooled HRs for PFS and OS of 3.20 (95%CI, 2.64–4.01, $p<0.001$; heterogeneity: $I^2 = 0.0\%$, $p = 0.774$) and 3.21 (95%CI, 2.54–4.07, $p<0.001$; heterogeneity: $I^2 = 0.0\%$, $p = 0.743$), respectively. The prospectively collected data showed pooled HRs for OS and PFS of 2.81 (95%CI,
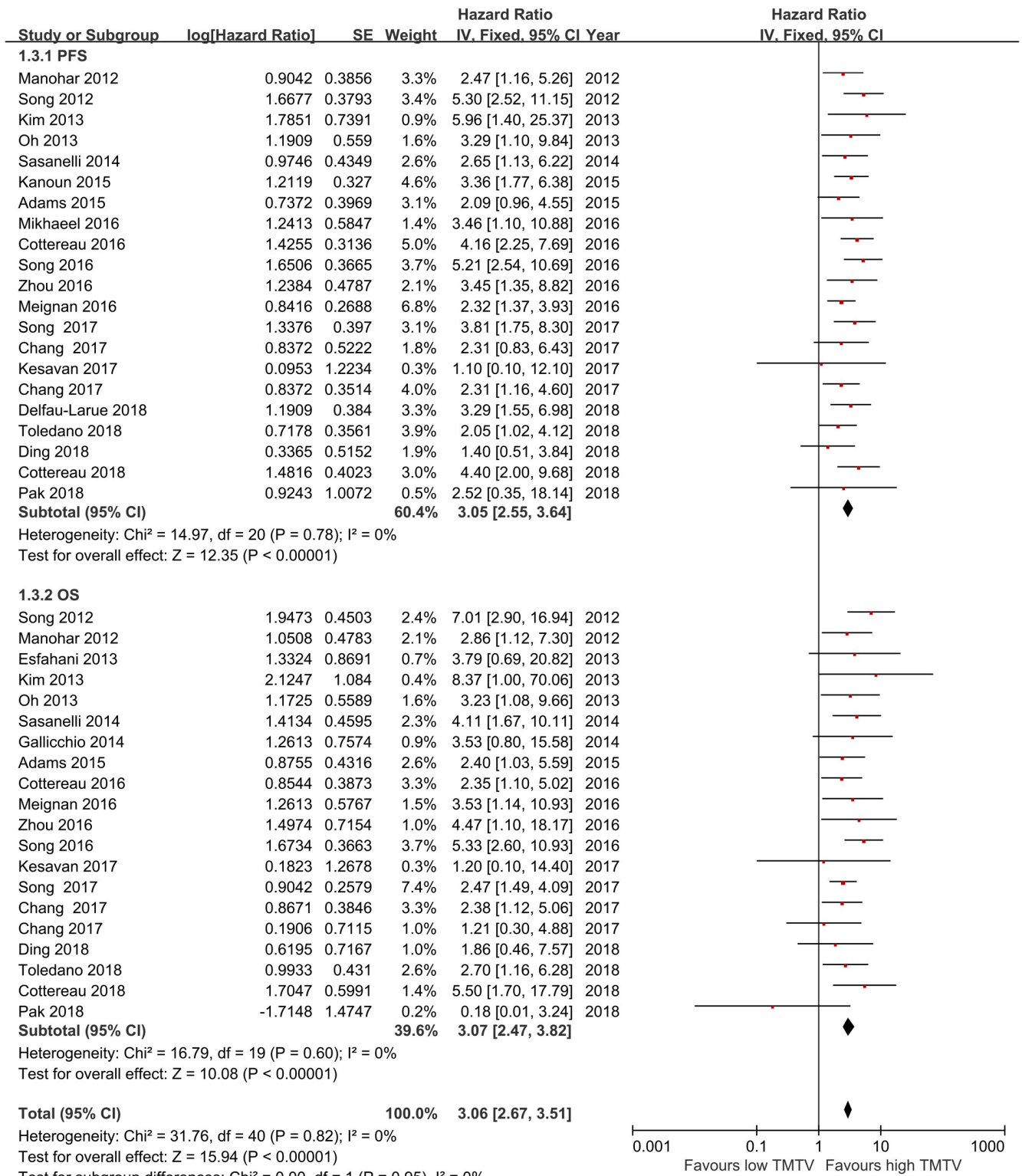
Study or Subgroup	log[Hazard Ratio]	SE	Weight	Hazard Ratio IV, Fixed, 95% CI	Year
1.3.1 PFS					
Manohar 2012	0.9042	0.3856	3.3%	2.47 [1.16, 5.26]	2012
Song 2012	1.6677	0.3793	3.4%	5.30 [2.52, 11.15]	2012
Kim 2013	1.7851	0.7391	0.9%	5.96 [1.40, 25.37]	2013
Oh 2013	1.1909	0.559	1.6%	3.29 [1.10, 9.84]	2013
Sasanelli 2014	0.9746	0.4349	2.6%	2.65 [1.13, 6.22]	2014
Kanoun 2015	1.2119	0.327	4.6%	3.36 [1.77, 6.38]	2015
Adams 2015	0.7372	0.3969	3.1%	2.09 [0.96, 4.55]	2015
Mikhaeel 2016	1.2413	0.5847	1.4%	3.46 [1.10, 10.88]	2016
Cottereau 2016	1.4255	0.3136	5.0%	4.16 [2.25, 7.69]	2016
Song 2016	1.6506	0.3665	3.7%	5.21 [2.54, 10.69]	2016
Zhou 2016	1.2384	0.4787	2.1%	3.45 [1.35, 8.82]	2016
Meignan 2016	0.8416	0.2688	6.8%	2.32 [1.37, 3.93]	2016
Song 2017	1.3376	0.397	3.1%	3.81 [1.75, 8.30]	2017
Chang 2017	0.8372	0.5222	1.8%	2.31 [0.83, 6.43]	2017
Kesavan 2017	0.0953	1.2234	0.3%	1.10 [0.10, 12.10]	2017
Chang 2017	0.8372	0.3514	4.0%	2.31 [1.16, 4.60]	2017
Delfau-Larue 2018	1.1909	0.384	3.3%	3.29 [1.55, 6.98]	2018
Toledano 2018	0.7178	0.3561	3.9%	2.05 [1.02, 4.12]	2018
Ding 2018	0.3365	0.5152	1.9%	1.40 [0.51, 3.84]	2018
Cottereau 2018	1.4816	0.4023	3.0%	4.40 [2.00, 9.68]	2018
Pak 2018	0.9243	1.0072	0.5%	2.52 [0.35, 18.14]	2018
Subtotal (95% CI)			**60.4%**	**3.05 [2.55, 3.64]**	
Heterogeneity: Chi² = 14.97, df = 20 (P = 0.78); I² = 0%					
Test for overall effect: Z = 12.35 (P < 0.00001)					
1.3.2 OS					
Song 2012	1.9473	0.4503	2.4%	7.01 [2.90, 16.94]	2012
Manohar 2012	1.0508	0.4783	2.1%	2.86 [1.12, 7.30]	2012
Esfahani 2013	1.3324	0.8691	0.7%	3.79 [0.69, 20.82]	2013
Kim 2013	2.1247	1.084	0.4%	8.37 [1.00, 70.06]	2013
Oh 2013	1.1725	0.5589	1.6%	3.23 [1.08, 9.66]	2013
Sasanelli 2014	1.4134	0.4595	2.3%	4.11 [1.67, 10.11]	2014
Gallicchio 2014	1.2613	0.7574	0.9%	3.53 [0.80, 15.58]	2014
Adams 2015	0.8755	0.4316	2.6%	2.40 [1.03, 5.59]	2015
Cottereau 2016	0.8544	0.3873	3.3%	2.35 [1.10, 5.02]	2016
Meignan 2016	1.2613	0.5767	1.5%	3.53 [1.14, 10.93]	2016
Zhou 2016	1.4974	0.7154	1.0%	4.47 [1.10, 18.17]	2016
Song 2016	1.6734	0.3663	3.7%	5.33 [2.60, 10.93]	2016
Kesavan 2017	0.1823	1.2678	0.3%	1.20 [0.10, 14.40]	2017
Song 2017	0.9042	0.2579	7.4%	2.47 [1.49, 4.09]	2017
Chang 2017	0.8671	0.3846	3.3%	2.38 [1.12, 5.06]	2017
Chang 2017	0.1906	0.7115	1.0%	1.21 [0.30, 4.88]	2017
Ding 2018	0.6195	0.7167	1.0%	1.86 [0.46, 7.57]	2018
Toledano 2018	0.9933	0.431	2.6%	2.70 [1.16, 6.28]	2018
Cottereau 2018	1.7047	0.5991	1.4%	5.50 [1.70, 17.79]	2018
Pak 2018	-1.7148	1.4747	0.2%	0.18 [0.01, 3.24]	2018
Subtotal (95% CI)			**39.6%**	**3.07 [2.47, 3.82]**	
Heterogeneity: Chi² = 16.79, df = 19 (P = 0.60); I² = 0%					
Test for overall effect: Z = 10.08 (P < 0.00001)					
Total (95% CI)			**100.0%**	**3.06 [2.67, 3.51]**	
Heterogeneity: Chi² = 31.76, df = 40 (P = 0.82); I² = 0%					
Test for overall effect: Z = 15.94 (P < 0.00001)					
Test for subgroup differences: Chi² = 0.00, df = 1 (P = 0.95), I² = 0%					

Hazard Ratio IV, Fixed, 95% CI

0.001 0.1 1 10 1000
Favours low TMTV Favours high TMTV

Fig 2. Meta-analysis of the hazard ratios for PFS, and OS for high TMTV vs low TMTV. Hazard ratios and 95% confidence intervals from individual studies are depicted as squares and horizontal lines, respectively. The pooled estimate is shown as a diamond shape, where the center represents the pooled HR and the horizontal borders represent the 95% CI. Hazard ratios are defined as high TMTV vs low TMTV, therefore a hazard ratio >1 represents a higher risk of death or progression associated with high TMTV. TMTV = total metabolic tumor volume, OS = overall survival, PFS = progression-free survival, CI = confidence interval.

https://doi.org/10.1371/journal.pone.0210224.g002

Table 2. Pooled hazard ratios for PFS and OS according to TMTV.

Study selection	N	PFS				N of cohorts	OS			
		Random-effects model		Heterogeneity			Random-effects model		Heterogeneity	
		Pooled HR (95% CI)	P value	I^2 (%)	P_H value		Pooled HR (95% CI)	P value	I^2 (%)	P_H value
Type of Lymphoma										
DLBCL	12	2.93(2.29–3.73)	<0.001	0	0.481	13	3.52(2.67–4.64)	<0.001	0	0.806
FL	3	2.55(1.65–3.92)	<0.001	0	0.622	2	2.89(1.04–7.99)	<0.001	0	0.419
PTCL	1	4.16(2.25–7.68)	<0.0001	-	-	1	2.35(1.10–5.04)	0.028	-	-
ENKTL	4	3.25(1.75–6.07)	<0.001	10.7	0.34	4	2.24(1.23–4.08)	0.008	24	0.267
HL	2	3.89(2.19–6.90)	<0.001	0	0.646	1	3.90(1.60–9.50)	0.0032	-	-
Study Design										
Retrospective	19	3.20(2.64–4.01)	<0.001	0	0.774	17	3.21(2.54–4.07)	<0.001	0	0.743
Prospective	3	2.81(1.72–4.59)	<0.001	11.2	0.342	3	1.91(0.51–7.15)	0.337	23.1	0.273
Sample size										
≥100	11	3.51(2.78–4.42)	<0.001	0	0.667	10	3.59(2.70–4.76)	<0.001	0	0.531
<100	7	2.82(1.99–3.99)	<0.001	0	0.55	9	2.45(1.65–3.65)	<0.001	0	0.809
Threshold										
≥2.5	5	3.93(2.76–5.60)	<0.001	0	0.553	5	3.65(2.38–5.61)	<0.001	28.9	0.229
41	6	2.56(1.83–3.58)	<0.001	0	0.554	5	3.56(2.22–5.70)	<0.001	0	0.772
40	3	2.19(1.21–3.98)	0.01	0	0.98	3	1.80(0.84–3.83)	0.129	5	0.349
other	4	3.20(2.02–5.05)	<0.001	0	0.753	4	3.69(1.89–7.22)	<0.001	0	0.817

N: number of studies; HR: hazard ratio; 95% CI: 95% confidence interval; P_H: p values of Q test for heterogeneity test; OS: Overall survival; PFS: Progression free survival.

1.72–4.59, p<0.001; heterogeneity: I^2 = 11.2%, p = 0.342) and 1.91 (95%CI, 0.51–7.15, p = 0.337; heterogeneity: I^2 = 23.1%, p = 0.273), respectively. The subgroup analysis performed on the basis of sample size showed that the negative predictive value of high TMTV on PFS and OS was present both in samples with sizes ≥100 (PFS, HR: 3.51, 95%CI: 2.78–4.42, p<0.001; OS, HR: 3.59, 95%CI: 2.70–4.76, p<0.001) and <100 (PFS, HR: 2.82, 95%CI: 1.99–3.99, p< 0.001; OS, HR: 2.45, 95%CI:1.65–3.65, p<0.001). Lastly, the subgroup analyses by thresholds (≥ 2.5, 40%, 41%, and others) demonstrated that a high TMTV was associated with shorter PFS and OS than a low TMTV (Table 2).

Prognostic value of baseline total lesion glycolysis

Ten studies [15–17,20,22,29,32,37,43,46] were included for the analysis of the relationship between TLG and PFS in adult lymphoma. The results of our meta-analysis showed that high TLG values were associated with shorter PFS than low TLG values, with a pooled HR of 3.44 (95%CI, 2.37–5.01, p<0.00001). There was moderate heterogeneity across the studies, but it did not reach statistical significance (I^2 = 41.0%, p = 0.08). Twelve studies [15,17,20,22,29,32,34,37,38,40,43,46] reported data on TLG and OS in adult lymphoma. Because significant heterogeneity was found across the studies (I^2 = 62.0%, P = 0.002), a pooled HR of 3.08 (95%CI, 1.84–5.16, p = 0.0001) was calculated on the basis of a random-effects model (Fig 3).

Next, we examined the prognostic significance of TLG on different types of lymphomas. A meta-analysis of seven and eight studies involving DLBCL patients showed poorer PFS and OS in those with high TMTV values than in those with low TMTV values, with pooled HRs for OS and PFS of 3.06 (95%CI, 1.52–6.18, p = 0.002; heterogeneity: I^2 = 67.3%, p = 0.003) and 2.93
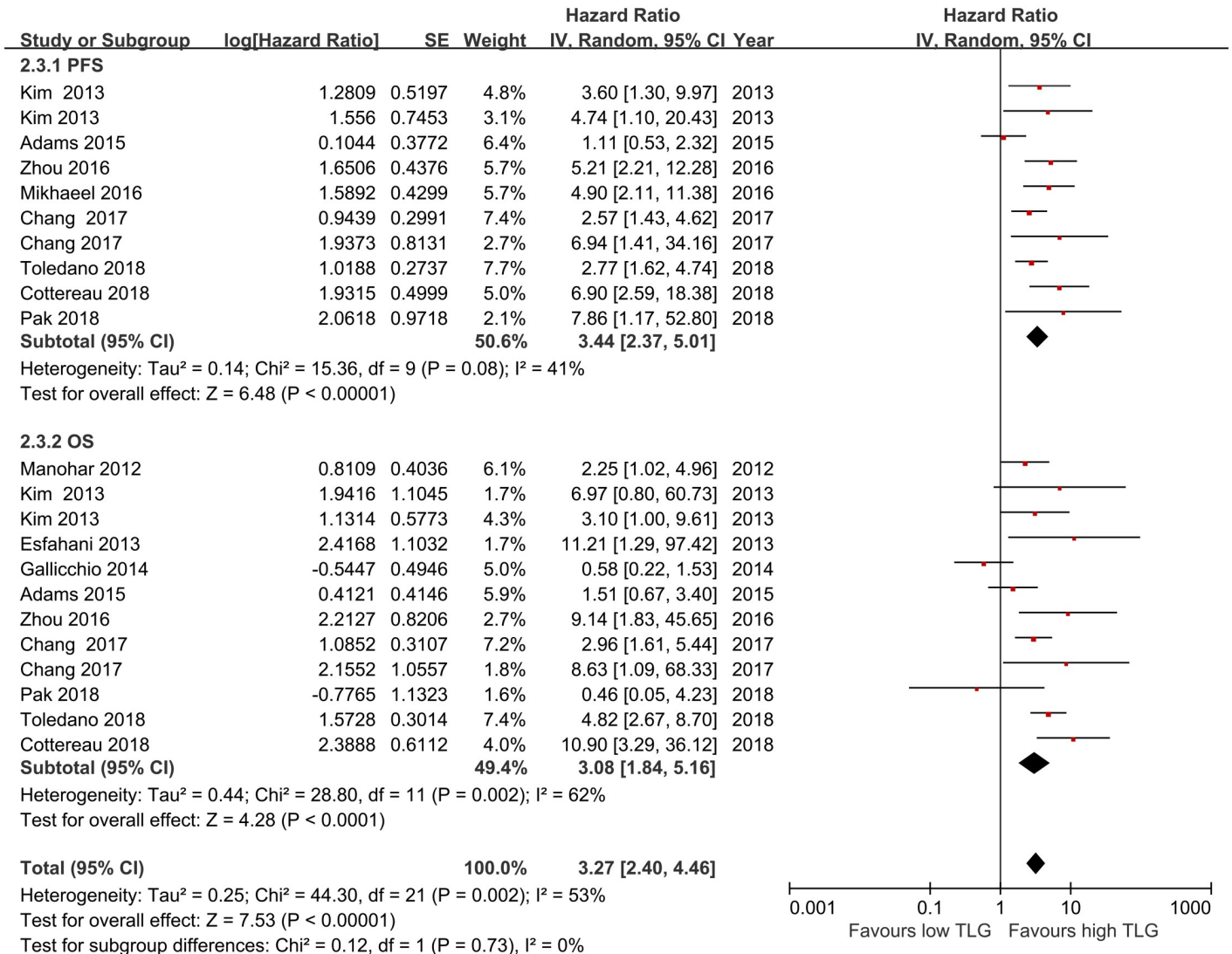
Study or Subgroup	log[Hazard Ratio]	SE	Weight	Hazard Ratio IV, Random, 95% CI	Year	Hazard Ratio IV, Random, 95% CI
2.3.1 PFS						
Kim 2013	1.2809	0.5197	4.8%	3.60 [1.30, 9.97]	2013	
Kim 2013	1.556	0.7453	3.1%	4.74 [1.10, 20.43]	2013	
Adams 2015	0.1044	0.3772	6.4%	1.11 [0.53, 2.32]	2015	
Zhou 2016	1.6506	0.4376	5.7%	5.21 [2.21, 12.28]	2016	
Mikhaeel 2016	1.5892	0.4299	5.7%	4.90 [2.11, 11.38]	2016	
Chang 2017	0.9439	0.2991	7.4%	2.57 [1.43, 4.62]	2017	
Chang 2017	1.9373	0.8131	2.7%	6.94 [1.41, 34.16]	2017	
Toledano 2018	1.0188	0.2737	7.7%	2.77 [1.62, 4.74]	2018	
Cottereau 2018	1.9315	0.4999	5.0%	6.90 [2.59, 18.38]	2018	
Pak 2018	2.0618	0.9718	2.1%	7.86 [1.17, 52.80]	2018	
Subtotal (95% CI)			50.6%	3.44 [2.37, 5.01]		

Heterogeneity: Tau² = 0.14; Chi² = 15.36, df = 9 (P = 0.08); I² = 41%
Test for overall effect: Z = 6.48 (P < 0.00001)

2.3.2 OS						
Manohar 2012	0.8109	0.4036	6.1%	2.25 [1.02, 4.96]	2012	
Kim 2013	1.9416	1.1045	1.7%	6.97 [0.80, 60.73]	2013	
Kim 2013	1.1314	0.5773	4.3%	3.10 [1.00, 9.61]	2013	
Esfahani 2013	2.4168	1.1032	1.7%	11.21 [1.29, 97.42]	2013	
Gallicchio 2014	-0.5447	0.4946	5.0%	0.58 [0.22, 1.53]	2014	
Adams 2015	0.4121	0.4146	5.9%	1.51 [0.67, 3.40]	2015	
Zhou 2016	2.2127	0.8206	2.7%	9.14 [1.83, 45.65]	2016	
Chang 2017	1.0852	0.3107	7.2%	2.96 [1.61, 5.44]	2017	
Chang 2017	2.1552	1.0557	1.8%	8.63 [1.09, 68.33]	2017	
Pak 2018	-0.7765	1.1323	1.6%	0.46 [0.05, 4.23]	2018	
Toledano 2018	1.5728	0.3014	7.4%	4.82 [2.67, 8.70]	2018	
Cottereau 2018	2.3888	0.6112	4.0%	10.90 [3.29, 36.12]	2018	
Subtotal (95% CI)			49.4%	3.08 [1.84, 5.16]		

Heterogeneity: Tau² = 0.44; Chi² = 28.80, df = 11 (P = 0.002); I² = 62%
Test for overall effect: Z = 4.28 (P < 0.0001)

Total (95% CI)			100.0%	3.27 [2.40, 4.46]		

Heterogeneity: Tau² = 0.25; Chi² = 44.30, df = 21 (P = 0.002); I² = 53%
Test for overall effect: Z = 7.53 (P < 0.00001)
Test for subgroup differences: Chi² = 0.12, df = 1 (P = 0.73), I² = 0%

0.001 0.1 1 10 1000
Favours low TLG   Favours high TLG

Fig 3. Meta-analysis of the hazard ratios for PFS and OS for high TLG vs low TLG. Hazard ratios and 95% confidence intervals for death or progression associated with high vs low TLG. TLG = total lesion glycolysis, OS = overall survival, PFS = progression-free survival, CI = confidence interval.

(95%CI, 1.89–4.53, $p<0.001$; heterogeneity: $I^2 = 49.5\%$, $p = 0.065$), respectively. Three studies examined the prognostic significance of high TLG values in ENKL patients, and the pooled HRs for PFS and OS were 2.99 (95%CI, 1.83–4.89, $p<0.001$; heterogeneity: $I^2 = 0.0\%$, $p = 0.503$) and 2.58 (95%CI, 1.33–5.01, $p = 0.005$; heterogeneity: $I^2 = 19.0\%$, $p = 0.291$), respectively (Table 3). Only one study provided relevant data on the correlation between TLG and clinical outcome in PTCL patients and only one study provided data on the correlation between TLG and outcome in HL patients; therefore, the pooled analysis could not be performed.

We also conducted subgroup analyses stratified by data collection method, sample size, and different threshold values. The subgroup analysis of retrospectively collected data showed pooled HRs for OS and PFS of 2.28 (95%CI, 1.40–3.71, $p = 0.001$; heterogeneity: $I^2 = 50.1\%$, $p = 0.029$) and 2.97 (95%CI, 2.03–4.35, $p<0.001$; heterogeneity: $I^2 = 35.7\%$, $p = 0.122$), respectively. Only one study involved prospectively collected data, so the pooled analysis could not

Table 3. Pooled hazard ratios for PFS and OS according to TLG.

Study selection	N	PFS				N	OS			
		Random-effects model		Heterogeneity			Random-effects model		Heterogeneity	
		Pooled HR (95% CI)	P value	I^2 (%)	P_H value		Pooled HR (95% CI)	P value	I^2 (%)	P_H value
Type of Lymphoma										
DLBCL	7	2.93(1.89–4.53)	<0.001	49.5	0.065	8	3.06(1.52–6.18)	0.002	67.3	0.003
PTCL	-	-	-	-	-	-	-	-	-	-
ENKTL	3	2.99(1.83–4.89)	<0.001	0	0.503	3	2.58(1.33–5.01)	0.005	19	0.291
HL	1	6.90(2.59–18.36)	0.0001	-	-	1	10.9(3.29–36.23)	0.0001	-	-
Study Design										
Retrospective	10	2.97(2.03–4.35)	<0.001	35.7	0.122	11	2.28(1.40–3.71)	0.001	50.1	0.029
Prospective	1	6.90(2.59–18.36)	0.0001	-	-	1	10.9(3.29–36.23)	0.0001	-	-
Sample size										
≥100	5	4.60(2.83–7.47)	<0.001	0	0.59	4	3.54(1.56–8.06)	0.003	62.6	0.045
<100	6	2.73(1.53–4.86)	0.001	48.7	0.083	9	2.07(1.21–3.55)	0.008	54.9	0.023
Threshold										
≥2.5	1	6.94(1.41–34.12)	0.017	-	-	1	8.63(1.09–68.34)	0.041	-	-
41	3	4.64(2.44–8.85)	<0.001	13.9	0.313	1	10.9(3.29–36.23)	0.0001	-	-
40	3	2.16(0.94–4.96)	0.07	60.6	0.079	3	1.90(0.91–3.98)	0.086	43.7	0.169
other	3	4.52(2.48–8.21)	<0.001	0	0.86	5	3.53(1.98–6.28)	<0.001	3.1	0.389

N: number of studies; HR: hazard ratio; 95% CI: 95% confidence interval; P_H: p values of Q test for heterogeneity test; OS: Overall survival; PFS: Progression free survival.

https://doi.org/10.1371/journal.pone.0210224.t003

be performed. The subgroup analysis conducted by sample size showed that the negative predictive value of high TMTV on PFS and OS was present both in samples with sizes ≥100 (PFS, HR: 4.60, 95%CI: 2.83–7.47, $p<0.001$; OS, HR: 3.54, 95%CI: 1.56–8.06, $p<0.001$) and <100 (PFS, HR: 2.73, 95%CI: 1.53–4.86, $p = 0.001$; OS, HR: 2.07, 95%CI:1.21–3.55, $p = 0.008$). A meta-analysis of three studies with a 40% threshold showed there was a trend for a correlation between high TLG values and lymphoma prognosis, with pooled HRs for OS and PFS of 1.90 (95%CI, 0.91–3.98, $p = 0.086$; heterogeneity: $I^2 = 43.7\%$, $p = 0.169$) and 2.16 (95%CI, 0.94–4.96, $p = 0.07$; heterogeneity: $I^2 = 60.6\%$, $p = 0.079$), respectively. The meta-analysis of studies using other thresholds (including the liver SUVmean plus 3SDs, 50%) showed a negative correlation between high TLG values and lymphoma prognosis, with pooled HRs for OS and PFS of 3.53 (95%CI, 1.98–6.28, $p<0.001$; heterogeneity: $I^2 = 3.1\%$, $p = 0.389$) and 4.52 (95%CI, 2.48–8.21, $p<0.001$; heterogeneity: $I^2 = 0.0\%$, $p = 0.860$), respectively. The subgroup analysis selecting a 41% threshold showed a negative predictive value of high TLG on PFS, with a pooled HR for PFS of 4.64 (95%CI, 2.44–8.85, $p<0.001$; heterogeneity: $I^2 = 13.9\%$, $p = 0.313$) (Table 3). Only one study used a threshold ≥2.5, so the pooled analysis could not be performed.

Publication bias

Inspection of the funnel plot and formal statistical tests (**TMTV:** Egger test, **p = 0.931**; Begg test, **p = 0.867**; **TLG:** Egger test, **p = 0.200**; Begg test, **p = 0.236**; Fig 4) showed no evidence of
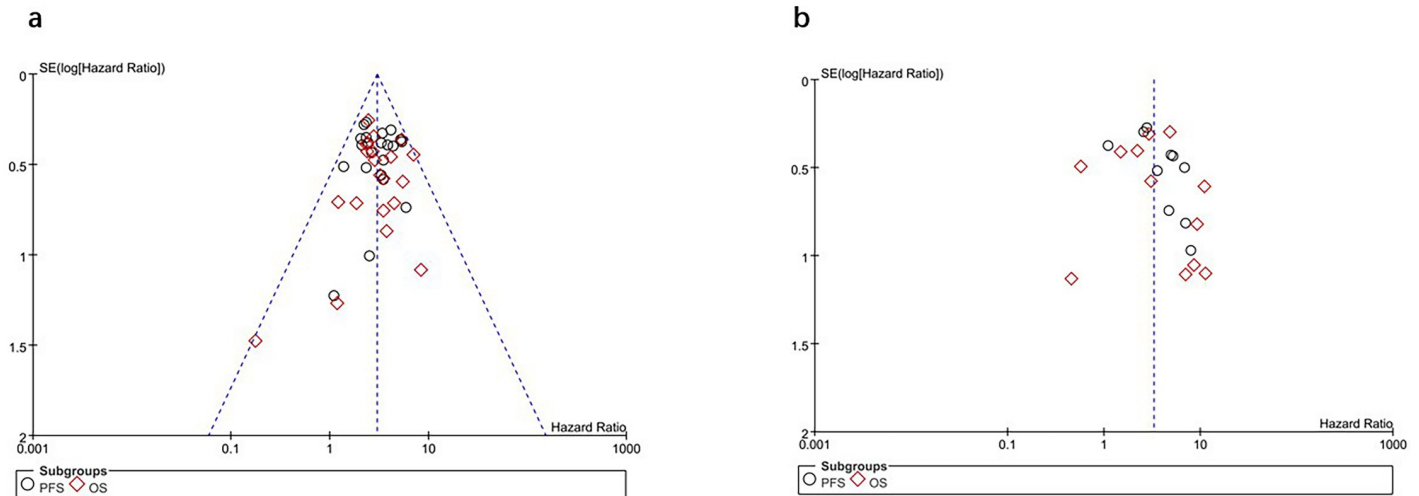
Fig 4. Assessment of publication bias using Funnel plot analysis. (a) Funnel plot of hazard ratio for overall survival and progression-free survival for high TMTV (horizontal axis) and the standard error (SE) for the hazard ratio (vertical axis). (b) Funnel plot of hazard ratio for overall survival and progression-free survival for high TLG (horizontal axis) and the standard error (SE) for the hazard ratio (vertical axis). Each study is represented by one circle. The vertical line represents the pooled effect estimate.

publication bias in the meta-analysis of the prognostic significance of baseline metabolic tumor volume and total lesion glycolysis in adult lymphoma.

Discussion

Main findings

This meta-analysis comprehensively and systematically reviewed the current available literature and found that: (1) A high baseline TMTV significantly predicted poor OS and shorter PFS in adult lymphoma patients ($p < 0.00001$ and $p < 0.00001$, respectively); (2) A high baseline TMTV was significantly associated with reduced survival in DLBCL patients treated with R-CHOP and predicted poor OS and PFS for different types of lymphomas, such as FL, ENKL and HL. The evidence supporting this association was consistent in most subgroup analyses (retrospective data collection, ethnicity, sample size, and different thresholds). The analysis of prospectively collected data and studies using a 40% threshold suggested a trend towards poor OS; however, these results were not statistically significant; (3) A high baseline TLG significantly predicted poor OS and shorter PFS in adult lymphoma patients ($p < 0.00001$ and $p = 0.005$, respectively); (4) A high baseline TLG was significantly associated with reduced survival in DLBCL patients treated with R-CHOP and predicted poor OS and PFS in different types of lymphomas, such as FL, ENKL and HL. The evidence of this association was consistent in most subgroup analyses (data collection method, ethnicity, sample size, and different thresholds).

Metabolic tumor volumes can be segmented by using various methods, such as a fixed SUV threshold, a percentage (based on the percentage of maximum uptake in the lesion), a threshold adjusted to the tumor-to-background ratio, or a gradient [23]. Reproducibility is the key for reliable volumetric tumor segmentation. Different TMTV measurement methods have been used in various types of lymphoma, each with specific advantages and disadvantages. A method based on the 41% SUVmax threshold is recommended by the European Association of Nuclear Medicine (EANM) for TMTV measurement of solid tumors. It has been developed in patients with HL and DLBCL, showing good reproducibility [48]. Different thresholding

methods were used for PET volume auto-segmentation in the studies included herein; however, a threshold of 41% or 40% of the SUVmax was widely used. In our study, we conducted subgroup analyses stratified by different thresholds ($\geq$2.5, 41%, and others). The results demonstrated that high TMTV or TLG values were associated with shorter PFS and OS. Subgroup stratification based on a threshold of 40% of the SUVmax showed that a high TMTV was a negative predictor of PFS; however, it did not significantly predict poor PFS and OS in the case of TLG.

Baseline MTV by PET/CT, is a promising prognostic indicator in patients with lymphoma, which is better than using size-defined bulk [16, 33]. TLG, which is the MTV multiplied by the mean SUV in the volume, is also prognostic [37], but appears no better than MTV in prediction of survival in lymphoma [16].

Several retrospective studies have shown that metabolic tumor volume (MTV) is a strong predictor of prognosis irrespective of the method [19,21,49]. However, cut-offs used to divide patients into high and low risk groups by MTV are highly dependent on the patient population and the method used. A fixed 41% SUVmax relative thresholding method has been applied successfully in different subtypes of lymphoma, but probably overestimated the volume of lesions with low SUVmax, particularly for smaller VOIs [19,21,48]. The 2.5 method could include the volume of nontumor regions located between small distant nodes with high uptake [50]. The 2.5 method probably overestimated MTV in approximately 12% of patients who had low FDG uptake in the liver or liver involvement by lymphoma [49]. Furthermore, the negative and positive predictive values of the 41% method have been shown to be superior to other methods, which results in excellent outcome prediction in other subtypes of lymphoma [21]. Generally, current evidence showed that metabolic tumor volume values were significantly influenced by the choice of the method used for determination of volume. However, no significant differences were found in term of prognosis [21]. In clinical practice, a consensus on the most accurate method or an optimal cut-off to define the MTV for specific lymphoma subtypes will be required, which will require validation in multicenter prospective trials.

Several methods for autosegmentation of PET volumes exist (e g, threshold-based, gradient-based, statistical, and texture-based methods) [51]. All methods have strengths and limitations. Reproducibility is the key for tumor segmentation in routine practice [23]. There is no universally accepted reproducible and practical method for tumor segmentation. Recently, Yu et al. reported a new semi-automatic approach that applies first an anatomical multi-atlas segmentation on the CT images to remove the organs having hyper uptake value on PET images. Using a CRFs (Conditional Random Fields) model, the rate of good detection of lymphoma is 100% in 11 patients [52]. Meanwhile, this new semi-automatic approach has the best dice index for the real lymphoma regions. However, this new methodology will require prospective validation in sufficiently large patient cohorts.

Among the included studies, there were mainly two different approaches to define the optimal TMTV cut-off value as a predictor of survival: X-tile analysis, receiver operating curve (ROC) analysis, or both. X-tile analysis is the primary approach for reliable cut-point determination. This method creates separate training and validation data sets, improving the robustness of the analysis [53]. In the studies included in this meta-analysis, ROC was widely used. This method defines the optimal cut-off point as the value whose sensitivity and specificity are closest to the value of the area under the ROC curve, and for which the absolute value of the difference between the sensitivity and specificity values is minimal. This method is recommended for finding the true cut-off point [54]. Meignan et al. [18] used another restricted cubic spline to define the optimal TMTV cut-off point. Splines are used to model the relationship between TMTV as a continuous variable and survival time, but their contribution to optimal cut-off point definition is minimal. Subgroup analyses based on different MTV cut-off

values demonstrated that patients with a high TMTV had shorter PFS and OS than those with a low TMTV.

A major strength of this meta-analysis is that it complied with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [55]. In addition, we extracted the maximum information from the included studies by a thorough qualitative review and quantitative meta-analysis.

Our study also has some limitations. Firstly, nearly all of the included studies were retrospective, which may result in confounding and detection bias. Secondly, patients with different types of lymphoma were treated with different therapeutic regimes. Thirdly, PET scans were performed using scanners of different generations, which may potentially affect the calculation of the SUV and therefore, of TMTV and TLG as well. Similarly, the FDG uptake times were difficult to standardize. Based on all of the above, the clinical heterogeneity of the included studies could be an issue. Finally, our meta-analysis was based on data from published trials, and we did not obtain individual patient data.

Conclusions

Our meta-analysis suggests that high baseline metabolic tumor volumes or total lesion glycolysis measured by FDG-PET/CT predict significantly worse overall survival and progression-free survival in patients with lymphoma. Therefore, TMTV and TLG may serve as new prognostic biomarkers. In view of our findings, future clinical trials with patients with different types of lymphoma are warranted to determine whether these novel findings can be integrated into various prognostic models, with the goal of achieving better risk stratification and treatment selection.

Supporting information

S1 Text. PRISMA checklist.
(DOC)

S1 Table. Risk of bias assessment.
(DOCX)

S2 Table. Characteristics of studies included in the meta-analysis.
(DOCX)

S3 Table. Excluded studies.
(DOCX)

S1 Appendix. Search strategy.
(DOCX)

Author Contributions

Formal analysis: Qing Ke, Hong Cen.

Methodology: Qing Ke, Hong Cen.

Project administration: Hong Cen.

Software: Qing Ke.

Writing – original draft: Baoping Guo.

Writing – review & editing: Baoping Guo, Xiaohong Tan, Hong Cen.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin. 2018; 68(1):7–30. https://doi.org/10.3322/caac.21442 PMID: 29313949

2. Linet MS, Brown LM, Mbulaiteye SM, Check D, Ostroumova E, Landgren A, et al. International long-term trends and recent patterns in the incidence of leukemias and lymphomas among children and adolescents ages 0–19 years. Int J Cancer. 2016; 138(8):1862–74. https://doi.org/10.1002/ijc.29924

3. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, et al. WHO classification of tumours of haematopoietic and lymphoid tissues. 4th eds. Lyon, France: International Agency for Research on Cancer press; 2017.

4. Cheson BD. International staging and response criteria for lymphomas. In: O'Brien S, Vose JM, Kantarjian HM. Management of Hematological Malignancies. New York, NY: Cambridge University Press; 2010. pp.277–58.

5. Vaidya R, Witzig TE. Prognostic factors for diffuse large B-cell lymphoma in the R(X)CHOP era. Ann Oncol. 2014; 25(11):2124–33. https://doi.org/10.1093/annonc/mdu109 PMID: 24625454

6. Moccia AA, Donaldson J, Chhanabhai M, Hoskins PJ, Klasa RJ, Savage KJ, et al. International Prognostic Score in advanced-stage Hodgkin's lymphoma: altered utility in the modern era. J Clin Oncol. 2012; 30(27):3383–8. https://doi.org/10.1200/JCO.2011.41.0910 PMID: 22869887

7. Solal-Celigny P, Roy P, Colombat P, White J, Armitage JO, Arranz-Saez R, et al. Follicular lymphoma international prognostic index. Blood. 2004; 104(5):1258–65. https://doi.org/10.1182/blood-2003-12-4434 PMID: 15126323

8. Hoster E, Dreyling M, Klapper W, Gisselbrecht C, van Hoof A, Kluin-Nelemans HC, et al. A new prognostic index (MIPI) for patients with advanced-stage mantle cell lymphoma. Blood. 2008; 111(2):558–65. https://doi.org/10.1182/blood-2007-06-095331 PMID: 17962512

9. Gallamini A, Stelitano C, Calvi R, Bellei M, Mattei D, Vitolo U, et al. Peripheral T-cell lymphoma unspecified (PTCL-U): a new prognostic model from a retrospective multicentric clinical study. Blood. 2004; 103(7):2474–9. https://doi.org/10.1182/blood-2003-09-3080 PMID: 14645001

10. Barrington SF, Mikhaeel NG, Kostakoglu L, Meignan M, Hutchings M, Mueller SP, 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):3048–58. https://doi.org/10.1200/JCO.2013.53.5229 PMID: 25113771

11. Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, 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):3059–68. https://doi.org/10.1200/JCO.2013.54.8800 PMID: 25113753

12. Adams HJ, Nievelstein RA, Kwee TC. Prognostic value of complete remission status at end-of-treatment FDG-PET in R-CHOP-treated diffuse large B-cell lymphoma: systematic review and meta-analysis. Br J Haematol. 2015; 170(2):185–91. https://doi.org/10.1111/bjh.13420 PMID: 25833790

13. El-Galaly TC, Pedersen MB, Hutchings M, Mylam KJ, Madsen J, Gang AO, et al. Utility of interim and end-of-treatment PET/CT in peripheral T-cell lymphomas: A review of 124 patients. Am J Hematol. 2015; 90(11):975–80. https://doi.org/10.1002/ajh.24128 PMID: 26201505

14. Adams HJ, Nievelstein RA, Kwee TC. Prognostic value of interim FDG-PET in Hodgkin lymphoma: systematic review and meta-analysis. Br J Haematol. 2015; 170(3):356–66. https://doi.org/10.1111/bjh.13441 PMID: 25868485

15. Chang CC, Cho SF, Chuang YW, Lin CY, Chang SM, Hsu WL, et al. Prognostic significance of total metabolic tumor volume on (18)F-fluorodeoxyglucose positron emission tomography/ computed tomography in patients with diffuse large B-cell lymphoma receiving rituximab-containing chemotherapy. Oncotarget. 2017; 8(59):99587–600. https://doi.org/10.18632/oncotarget.20447

16. Mikhaeel NG, Smith D, Dunn JT, Phillips M, Moller H, Fields PA, 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):1209–19. https://doi.org/10.1007/s00259-016-3315-7 PMID: 26902371

17. Adams HJ, de Klerk JM, Fijnheer R, Heggelman BG, Dubois SV, Nievelstein RA, et al. Prognostic superiority of the National Comprehensive Cancer Network International Prognostic Index over pretreatment whole-body volumetric-metabolic FDG-PET/CT metrics in diffuse large B-cell lymphoma. Eur J Haematol. 2015; 94(6):532–9. https://doi.org/10.1111/ejh.12467 PMID: 25311082

18. 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):3618–26. https://doi.org/10.1200/JCO.2016.66.9440 PMID: 27551111

19. Cottereau AS, Becker S, Broussais F, Casasnovas O, Kanoun S, Roques M, 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):719–24. https://doi.org/10.1093/annonc/mdw011 PMID: 26787236

20. Kim CY, Hong CM, Kim DH, Son SH, Jeong SY, Lee SW, et al. Prognostic value of whole-body meta-bolic tumour volume and total lesion glycolysis measured on (1)(8)F-FDG PET/CT in patients with extra-nodal NK/T-cell lymphoma. Eur J Nucl Med Mol Imaging. 2013; 40(9):1321–9. https://doi.org/10.1007/s00259-013-2443-6

21. Kanoun S, Tal I, Berriolo-Riedinger A, Rossi C, Riedinger JM, Vrigneaud JM, 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. https://doi.org/10.1371/journal.pone.0140830 PMID: 26473950

22. Cottereau AS, Versari A, Loft A, Casasnovas O, Bellei M, Ricci R, et al. Prognostic value of baseline metabolic tumor volume in early-stage Hodgkin lymphoma in the standard arm of the H10 trial. Blood. 2018; 131(13):1456–63. https://doi.org/10.1182/blood-2017-07-795476 PMID: 29437590

23. Kostakoglu L, Chauvie S. Metabolic Tumor Volume Metrics in Lymphoma. Semin Nucl Med. 2018; 48 (1):50–66. https://doi.org/10.1053/j.semnuclmed.2017.09.005 PMID: 29195618

24. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonran-domized studies in meta-analyses. Eur J Epidemiol. 2010; 25(9):603–5. https://doi.org/10.1007/s10654-010-9491-z PMID: 20652370

25. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med. 1998; 17(24):2815–34. PMID: 9921604

26. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7. Rating the quality of evidence–;inconsistency. J Clin Epidemiol. 2011; 64(12):1294–302. https://doi.org/10.1016/j.jclinepi.2011.03.017

27. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Bio-metrics. 1994; 50(4):1088–101. PMID: 7786990

28. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997; 315(7109):629–34. PMID: 9310563

29. Pak K, Kim BS, Kim K, Kim IJ, Jun S, Jeong YJ, et al. Prognostic significance of standardized uptake value on F18-FDG PET/CT in patients with extranodal nasal type NK/T cell lymphoma: A multicenter, retrospective analysis. Am J Otolaryngol. 2018; 39(1):1–5. https://doi.org/10.1016/j.amjoto.2017.10.009 PMID: 29056243

30. Song MK, Chung JS, Yhim HY, Lim SN, Kim SJ, Han YH, et al. Tumor necrosis and complete resection has significant impacts on survival in patients with limited-stage upper aerodigestive tract NK/T cell lym-phoma. Oncotarget. 2017; 8(45):79337–46. https://doi.org/10.18632/oncotarget.18107 PMID: 29108312

31. Kesavan M, Boucek J, MacDonald W, McQuillan A, Turner JH. Imaging of Early Response to Predict Prognosis in the First-Line Management of Follicular Non-Hodgkin Lymphoma with Iodine-131-Rituxi-mab Radioimmunotherapy. Diagnostics (Basel). 2017; 7(2).

32. Chang Y, Fu X, Sun Z, Xie X, Wang R, Li Z, et al. Utility of baseline, interim and end-of-treatment (18)F-FDG PET/CT in extranodal natural killer/T-cell lymphoma patients treated with L-asparaginase/pegas-pargase. Sci Rep. 2017; 7:41057. https://doi.org/10.1038/srep41057

33. Song MK, Chung JS, Shin HJ, Lee SM, Lee SE, Lee HS, et al. Clinical significance of metabolic tumor volume by PET/CT in stages II and III of diffuse large B cell lymphoma without extranodal site involve-ment. Ann Hematol. 2012; 91(5):697–703. https://doi.org/10.1007/s00277-011-1357-2 PMID: 22071570

34. Manohar K, Mittal BR, Bhattacharya A, Malhotra P, Varma S. Prognostic value of quantitative parame-ters derived on initial staging 18F-fluorodeoxyglucose positron emission tomography/computed tomog-raphy in patients with high-grade non-Hodgkin's lymphoma. Nucl Med Commun. 2012; 33(9):974–81. https://doi.org/10.1097/MNM.0b013e32835673ec PMID: 22811258

35. Oh MY, Chung JS, Song MK, Shin HJ, Lee HS, Lee SM, et al. Prognostic value of Waldeyer's ring involvement of diffuse large B-cell lymphoma treated with R-CHOP. Int J Hematol. 2013; 97(3):397–402. https://doi.org/10.1007/s12185-013-1282-3 PMID: 23400410

36. Song MK, Chung JS, Lee JJ, Jeong SY, Lee SM, Hong JS, 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):1656–61. https://doi.org/10.1111/cas.12282 PMID: 24033666

37. Kim TM, Paeng JC, Chun IK, Keam B, Jeon YK, Lee SH, et al. Total lesion glycolysis in positron emis-sion tomography is a better predictor of outcome than the International Prognostic Index for patients

with diffuse large B cell lymphoma. Cancer. 2013; 119(6):1195–202. https://doi.org/10.1002/cncr.27855 PMID: 23212736

38. Esfahani SA, Heidari P, Halpern EF, Hochberg EP, Palmer EL, Mahmood U. Baseline total lesion gly-colysis measured with (18)F-FDG PET/CT as a predictor of progression-free survival in diffuse large B-cell lymphoma: a pilot study. Am J Nucl Med Mol Imaging. 2013; 3(3):272–81. PMID: 23638338

39. Sasanelli M, Meignan M, Haioun C, Berriolo-Riedinger A, Casasnovas RO, Biggi A, et al. Pretherapy metabolic tumour volume is an independent predictor of outcome in patients with diffuse large B-cell lymphoma. Eur J Nucl Med Mol Imaging. 2014; 41(11):2017–22. https://doi.org/10.1007/s00259-014-2822-7 PMID: 24902639

40. Gallicchio R, Mansueto G, Simeon V, Nardelli A, Guariglia R, Capacchione D, et al. F-18 FDG PET/CT quantization parameters as predictors of outcome in patients with diffuse large B-cell lymphoma. Eur J Haematol. 2014; 92(5):382–9. https://doi.org/10.1111/ejh.12268 PMID: 24428392

41. Kim J, Hong J, Kim SG, Hwang KH, Kim M, Ahn HK, et al. Prognostic Value of Metabolic Tumor Volume Estimated by (18) F-FDG Positron Emission Tomography/Computed Tomography in Patients with Dif-fuse Large B-Cell Lymphoma of Stage II or III Disease. Nucl Med Mol Imaging. 2014; 48(3):187–95. https://doi.org/10.1007/s13139-014-0280-6 PMID: 25177375

42. Schoder H, Zelenetz AD, Hamlin P, Gavane S, Horwitz S, Matasar M, et al. Prospective Study of 3'-Deoxy-3'-18F-Fluorothymidine PET for Early Interim Response Assessment in Advanced-Stage B-Cell Lymphoma. J Nucl Med. 2016; 57(5):728–34. https://doi.org/10.2967/jnumed.115.166769

43. Zhou M, Chen Y, Huang H, Zhou X, Liu J, Huang G. Prognostic value of total lesion glycolysis of base-line 18F-fluorodeoxyglucose positron emission tomography/computed tomography in diffuse large B-cell lymphoma. Oncotarget. 2016; 7(50):83544–53. https://doi.org/10.18632/oncotarget.13180 PMID: 27835875

44. Song MK, Yang DH, Lee GW, Lim SN, Shin S, Pak KJ, et al. High total metabolic tumor volume in PET/CT predicts worse prognosis in diffuse large B cell lymphoma patients with bone marrow involvement in rituximab era. Leuk Res. 2016; 42:1–6. https://doi.org/10.1016/j.leukres.2016.01.010 PMID: 26851438

45. Ding CY, Guo Z, Sun J, Yang WP, Li TR. Prognostic value of pretreatment (18)F-FDG PET-CT for patients with advanced diffuse large B-cell lymphoma. Zhonghua Zhong Liu Za Zhi. 2018; 40(7):528–33. https://doi.org/10.3760/cma.j.issn.0253-3766.2018.07.009 PMID: 30060362

46. Toledano MN, Desbordes P, Banjar A, Gardin I, Vera P, Ruminy P, et al. Combination of baseline FDG PET/CT total metabolic tumour volume and gene expression profile have a robust predictive value in patients with diffuse large B-cell lymphoma.Eur J Nucl Med Mol Imaging. 2018; 45(5):680–88. https://doi.org/10.1007/s00259-017-3907-x PMID: 29344718

47. Delfau-Larue MH, van der Gucht A, Dupuis J, Jais JP, Nel I, Beldi-Ferchiou A, et al.Total metabolic tumor volume, circulating tumor cells, cell-free DNA: distinct prognostic value in follicular lymphoma. Blood Adv. 2018; 2(7):807–16. https://doi.org/10.1182/bloodadvances.2017015164 PMID: 29636326

48. Meignan M, Sasanelli M, Casasnovas RO, Luminari S, Fioroni F, Coriani C, et al. Metabolic tumour vol-umes measured at staging in lymphoma: methodological evaluation on phantom experiments and patients. Eur J Nucl Med Mol Imaging. 2014; 41(6):1113–22. https://doi.org/10.1007/s00259-014-2705-y PMID: 24570094

49. Ilyas H, Mikhaeel NG, Dunn JT, Rahman F, Møller H, Smith D, et al. Defining the optimal method for measuring baseline metabolic tumour volume in diffuse large B cell lymphoma. Eur J Nucl Med Mol Imaging. 2018; 45(7):1142–54. https://doi.org/10.1007/s00259-018-3953-z PMID: 29460024

50. Soret M, Bacharach SL, Buvat I: Partial-volume effect in PET tumor imaging. J Nucl Med. 2007; 48 (6):932–45. https://doi.org/10.2967/jnumed.106.035774

51. Schöder H, Moskowitz C. Metabolic Tumor Volume in Lymphoma: Hype or Hope? J Clin Oncol. 2016; 34(30):3591–94. https://doi.org/10.1200/JCO.2016.69.3747 PMID: 27601547

52. Yu Y, Decazes P, Lapuyade-Lahorgue J, Gardin I, Vera P, Ruan S. Semi-automatic lymphoma detec-tion and segmentation using fully conditional random fields. Comput Med Imaging Graph. 2018; 70:1–7. https://doi.org/10.1016/j.compmedimag.2018.09.001 PMID: 30253305

53. 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):7252–9. https://doi.org/10.1158/1078-0432.CCR-04-0713 PMID: 15534099

54. Unal I. Defining an Optimal Cut-Point Value in ROC Analysis: An Alternative Approach. Comput Math Methods Med. 2017; 2017:3762651. https://doi.org/10.1155/2017/3762651 PMID: 28642804

55. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al.The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009; 6(7):e1000100. https://doi.org/10.1371/journal.pmed.1000100 PMID: 19621070