Am J Cancer Res 2022;12(8):3857-3869
www.ajcr.us /ISSN:2156-6976/ajcr0144588

Original Article

Improving the prognostic ability of PET/CT SUVmax to identify follicular lymphoma with early treatment failure

Xin Wan[1], Wei Guo[1], Xingtong Wang[1], Jia Li[1], Yangzhi Zhao[1], Xiaomeng Feng[1], Ken H Young[2], Ou Bai[1]

[1]Department of Hematology, The First Hospital of Jilin University, Jilin, China; [2]Division of Hematopathology, Department of Pathology, Duke University Medical Center and Duke Cancer Institute, Durham, NC, USA

Received June 6, 2022; Accepted June 27, 2022; Epub August 15, 2022; Published August 30, 2022

Abstract: Follicular lymphoma (FL) has a high degree of heterogeneity both clinically and molecularly. Early treatment failure (ETF), progression or relapse within 24 months of frontline immunochemotherapy is associated with a poor prognosis in FL. However, the clinical utility of ETF at diagnosis is limited. The maximum standardized uptake value (SUVmax) is a metabolic parameter for positron emission tomography/computed tomography (PET/CT); nevertheless, the relationship between SUVmax and ETF remains unclear. Thus, identifying early biomarkers that incorporate SUVmax and other clinical correlative variables could be helpful in identifying patients at high risk of ETF. A nomogram consisted of three independent variables, including SUVmax ≥ 12, beta-2 microglobulin > 3 mg/L, and Ki67 > 40%, was established to predict ETF in 127 patients with grade 1, 2, or 3a FL from the First Hospital of Jilin University (training cohort) and was validated using data from the Duke University Medical Center (validation cohort, n=95). The nomogram demonstrated prognostic accuracy in predicting ETF (sensitivity 70.8% and specificity 83.5% in the training cohort; sensitivity 84.2% and specificity 68.4% in the validation cohort). The patients were stratified into three groups: low-, intermediate-, and high-risk. In the training cohort, the corresponding 5-year progression-free survival (PFS) rates were 81.7%, 73.4%, and 34.9%, and the 5-year overall survival (OS) rates were 97.4%, 87.4%, and 62.3%, respectively. In the validation cohort, the 5-year PFS rates were 77.7%, 52.9%, and 34.8%, and the 5-year OS rates were 96.4%, 94.1%, and 73.7%, respectively. This was the first study to use a nomogram with SUVmax to predict ETF in FL to identify a subset of patients who might benefit from individualized targeted therapy.

Keywords: Follicular lymphoma, PET/CT, SUVmax, early treatment failure, nomogram

Introduction

Follicular lymphoma (FL), the second most common non-Hodgkin lymphoma derived from mature B cells [1], is classified into grades 1, 2, 3a, and 3b. The clinical course of FL1-3a mostly resembles indolent lymphoma, whereas that of FL3b is more similar to diffuse large B-cell lymphoma (DLBCL). Over the past few decades, immunochemotherapy and rituximab maintenance have significantly improved the prognosis of FL, with an estimated median overall survival (OS) of 18-20 years [2]. Cancer cure in clinic refers to the maintenance of complete remission for five years. Patients with high-risk factors often develop drug resistance and repeated failure; early detection and active treatment may help reduce the risk of the development of recurrence [3, 4]. Approximately 20% of patients with FL experience early treatment failure (ETF), progression or relapse within 24 months of initial frontline immunochemotherapy [5]. According to Casulo *et al*. [6], early progression is correlated with an inferior 5-year OS of 50% in patients who initially underwent R-CHOP treatment (rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone) compared with 90% in the control group. The results of the GALLIUM study, which included patients treated with obinutuzumab-based regimens, showed that patients who experienced progression within 24 months had a markedly increased risk of death with a hazard ratio (HR) of 26 [7]. However, ETF cannot be predicted at the time of FL diagnosis, and the current risk assessment criteria, such as the Follicular Lymphoma International Prognostic Index (FLIPI, FLIPI2), cannot accurately identify ETF. Thus, identifying patients in whom first-line treatment may fail, leading to early progression, and predicting outcomes at FL diagnosis remains a significant challenge.

In recent years, 18F-fluorodeoxyglucose (FDG) positron emission tomography/computed tomography (PET/CT) has been used for staging, restaging, and evaluating the response in patients with lymphoma. The total metabolic tumor volume (TMTV) and pre-treatment maximum standardized uptake value (SUVmax) are both strong predictors of FL outcomes. The SUVmax of PET/CT is a semi-quantitative scale and is associated with the prognosis of FL. Strati [8] reported that SUVmax > 18 was associated with inferior 8-year OS in patients with FL treated with either R-CHOP (70% vs. 90%; *P*=0.02) or non-anthracycline-based frontline regimens (50% vs. 85%; *P*=0.001). Very recently, a retrospective study reported that pretreatment SUVmax > 14.5, instead of TMTV, was associated with poorer progression-free survival (PFS) than baseline SUVmax ≤ 14.5, and 2-year PFS was 54% versus 86%. This indicated that SUVmax was associated with early progression in patients with FL 1-3a [9]. However, the relationship between the PET/CT SUVmax and ETF remains unclear.

In addition, numerous independent parameters, such as lactate dehydrogenase (LDH), beta-2 microglobulin (β2MG), Ki67, and advanced stage, were previously shown to have prognostic significance in FL [10, 11]. In this retrospective study, we identified SUVmax ≥ 12, β2MG > 3 mg/L, and Ki67 > 40% as independent prognostic risk factors for ETF in two independent cohorts. Furthermore, we constructed a nomogram and compared its ability with current risk assessment criteria, such as FLIPI, FLIPI2, and PRIMA-Prognostic Index (PRIMA-PI), to predict FL prognosis in patients with a high risk of ETF at the diagnosis stage.

Methods

Patient population

We analyzed clinical data from two independent cohorts of patients with symptomatic, bulky (≥ 6 cm), or advanced-stage FL considered ineligible for curative radiotherapy. The inclusion criteria were as follows: (a) patients with newly diagnosed FL 1-3a, (b) FDG-avid lesions on baseline 18F-FDG PET/CT, (c) older than 18 years, and (d) received immunochemotherapy as initial treatment. The exclusion criteria were as follows: (a) newly diagnosed FL 1-3a with transformation, (b) secondary malignant disease, (c) those who may benefit from curative radiotherapy, (d) incomplete information, and (e) severe organ dysfunction due to lymphoma.

In the training cohort, we retrospectively investigated the records of 296 patients diagnosed with FL for the first time at the First Hospital of Jilin University from January 2012 to December 2021, according to the 2008 or 2016 revised World Health Organization classification of hematological malignancies. As shown in **Figure 1A**, 127 patients were enrolled in this study, with 120 receiving an anthracycline-based regimen with anti-CD20 monoclonal antibody (115 rituximab and 5 obinutuzumab), and 7 receiving bendamustine with rituximab (BR).

Briefly, in the validation cohort from the Duke University Medical Center, 95 patients were enrolled (**Figure 1B**). Thirteen patients received anthracycline-based regimens with rituximab and 82 received BR. The dose of drugs was based on NCCN guidelines of FL in these two independent cohorts. This study was conducted in accordance with the principles of the Declaration of Helsinki and protocol approved by the institutional review board of Universities.

Clinical and metabolic variables

Clinical information was obtained by reviewing the patients' medical charts, including sex, age at diagnosis, pathologic grade, Ann Arbor stage, Eastern Cooperative Oncology Group (ECOG), serum LDH and β2MG, bone marrow (BM) involvement on biopsy, FLIPI, FLIPI2, SUVmax of PET/CT, Ki67, and other laboratory data. Patients received treatment or observation according to published criteria [12, 13]. The primary endpoint in this study was ETF, defined as disease relapse or progression within 2-years of starting frontline immunochemotherapy.

Baseline PET/CT was performed before immunochemotherapy, and two nuclear physicians who were unaware of the results of any other imaging tests and clinical data reviewed the images.

Statistical analyses

The SUVmax cut-off value was determined using a receiver operating characteristic (ROC) curve assay. The chi-square test was used to compare categorical variables. Logistic regres-

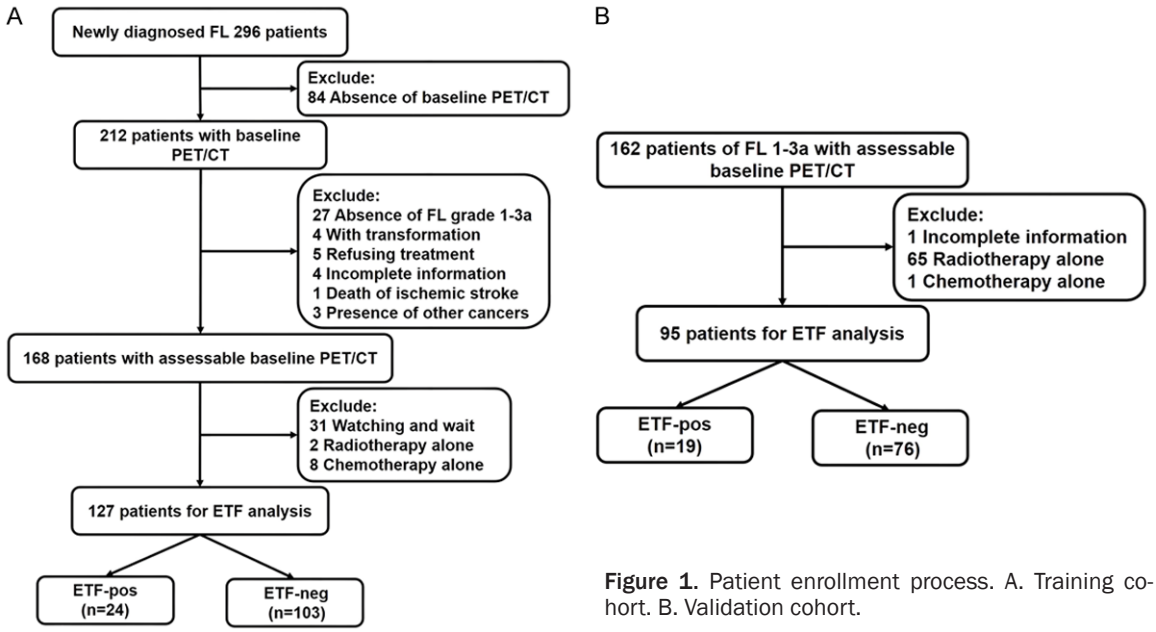


Figure 1. Patient enrollment process. A. Training cohort. B. Validation cohort.

sion was used for multivariate analyses, and factors ($P < 0.05$) on univariate analysis were included. A nomogram was constructed according to the significant prognostic factors using the "rms" package in the R software. The performance of the nomogram was evaluated using C-statistics and ROC analyses. The area under the ROC curve (AUC) was used to assess the accuracy of this nomogram. Differences between the Kaplan-Meier curves of PFS and OS were evaluated using the nonparametric log-rank test. *P* value < 0.05 was considered statistically significant. SPSS version 25.0 or R version 4.1.3 was used for all statistical analyses.

Results

Baseline patient characteristics

The patients' baseline characteristics are shown in **Table 1**. The median age at diagnosis in the training and validation cohorts was 50 years (range, 28-80 years) and 60 years, respectively. A total of 33.9% (43/127) in the training cohort and 17.9% (17/95) in the validation cohort had FL 3a. More patients had advanced-stage disease (100/127, 78.7% in the training cohort; 89/95, 93.7% in the validation cohort). Furthermore, ETF occurred in 24 (18.9%, ETF-pos group) and 19 (20.0%) patients in the training and validation cohorts, respectively.

Optimal cutoff value determination

The optimal threshold of baseline SUVmax was determined in the training cohort. ROC curve analysis showed an AUC of 0.617 (0.508-0.725), with an optimal cutoff value of SUVmax at 12.0 for PFS prediction (sensitivity 0.55, specificity 0.71, *P*=0.046). In the ETF-pos group, 14 patients (58.3%) had an SUVmax ≥ 12 in the training cohort and eight (42.1%) in the validation cohort (**Table 1**). In addition, we found that a Ki67 index of 40% would be more optimal than other cutoff values. β2MG concentration > 3 mg/L was consistent with a previous study [14].

Survival analyses

With a median follow-up time of 41 months in the training cohort and 84 months in the validation cohort, 14 of 127 (11.0%) and 11 of 95 (11.6%) patients died, respectively. In the training and validation cohorts, the 5-year PFS rates were 66.2% and 62.8%, respectively. The 5-year OS rates were 84.1% and 90.5%, respectively (**Figure 2A-D**). The median OS of FL patients in the ETF-pos group was poor: 34 months in the training cohort and 63 months in the validation cohort. The 5-year OS was significantly lower in the ETF-pos group (42.3% and 62.7%, respectively) than in the ETF-neg group (96.3% and 97.4%, respectively; *P* < 0.0001) (**Figure 2E**, **2F**).

Table 1. Baseline characteristics and univariate analyses of patients with FL1-3a

Characteristics	Training cohort					Validation cohort				
	ETF-neg		ETF-pos		*P* value	ETF-neg		ETF-neg		*P* value
	N=103	%	N=24	%		N=76	%	N=19	%	
Gender					0.066					0.837
Male	43	41.7	15	62.5		38	50.0	10	52.6	
Female	60	58.3	9	37.5		38	50.0	9	47.4	
Age					0.424					0.757
≤ 60	78	75.7	20	83.3		43	56.6	10	52.6	
> 60	25	24.3	4	16.7		33	43.4	9	47.4	
Histologic grade					0.675					0.082
FL 1-2	69	67.0	15	62.5		65	85.5	13	68.4	
FL 3a	34	33.0	9	37.5		11	14.5	6	31.6	
Ann Arbor stage					0.023					0.206
I-II	26	25.2	1	4.2		6	7.9	0	0	
III-IV	77	74.8	23	95.8		70	92.1	19	100	
ECOG					0.009					0.040
0-1	98	95.1	19	79.2		64	84.2	12	63.2	
> 1	5	4.9	5	20.8		12	15.8	7	36.8	
B symptom					0.196					0.885
No	74	71.8	14	58.3		65	85.5	16	84.2	
Yes	29	28.2	10	41.7		11	14.5	3	15.8	
LDH					0.003					0.179
Normal	79	76.7	11	45.8		31	40.8	11	57.9	
Elevate	24	23.3	13	54.2		45	59.2	8	42.1	
β2MG (mg/L)					< 0.0001					0.005
≤ 3	76	73.8	8	33.3		63	82.9	10	52.6	
> 3	27	26.2	16	66.7		13	17.1	9	47.4	
Hemoglobin level g/dl					0.478					0.359
> 12	80	77.7	17	70.8		63	82.9	14	73.7	
≤ 12	23	22.3	7	29.2		13	17.1	5	26.3	
BMB					0.120					0.051
Negative	65	63.1	11	45.8		54	71.1	9	47.4	
Positive	38	36.9	13	54.2		22	28.9	10	52.6	
Number of extranodal sites					0.237					0.034
0-1	65	63.1	12	50.0		27	35.5	2	10.5	
≥ 2	38	36.9	12	50.0		49	64.5	17	89.5	
Size of largest node > 6 cm					0.755					0.834
No	74	71.8	18	75.0		30	39.5	8	42.1	
Yes	29	28.2	6	25.0		46	60.5	11	57.9	
SUVmax					0.009					0.007
< 12	72	69.9	10	41.7		65	85.5	11	57.9	
≥ 12	31	30.1	14	58.3		11	14.5	8	42.1	
Ki67					0.015					< 0.0001
≤ 40%	74	71.8	11	45.8		57	75.0	6	31.6	
> 40%	29	28.2	13	54.2		19	25%	13	68.4	

β2MG: beta-2 microglobulin, BMB: bone marrow biopsy; ECOG: Eastern Cooperative Oncology Group; LDH: lactate dehydrogenase; SUVmax: maximum standardized uptake value.

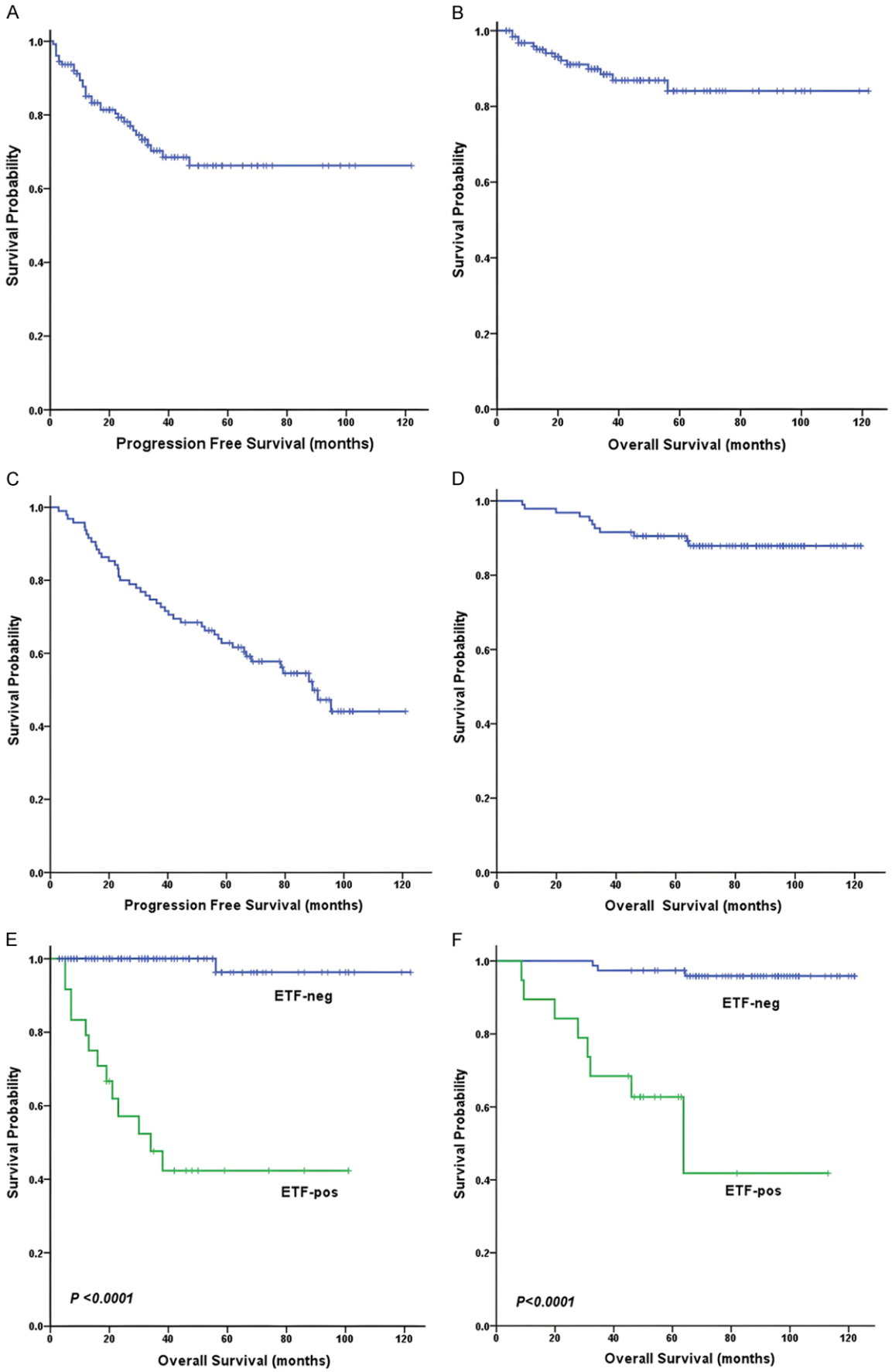

A
Survival Probability
Progression Free Survival (months)
B
Survival Probability
Overall Survival (months)
C
Survival Probability
Progression Free Survival (months)
D
Survival Probability
Overall Survival (months)
E
ETF-neg
ETF-pos
P <0.0001
Survival Probability
Overall Survival (months)
F
ETF-neg
ETF-pos
P<0.0001
Survival Probability
Overall Survival (months)

Figure 2. Survival analyses. PFS (A) and OS (B) of patients in the training cohort. PFS (C) and OS (D) in the validation cohort. Kaplan-Meier curves of OS according to early treatment failure in the training cohort (E) and in the validation cohort (F).

Table 2. Multivariate analyses of ETF in the training and validation cohorts

	Training cohort				Validation cohort		
	HR	95% CI	*P* value		HR	95% CI	*P* value
SUVmax ≥ 12	3.282	1.140-9.447	0.028	SUVmax ≥ 12	8.028	1.651-39.023	0.01
β2MG > 3 mg/L	4.061	1.272-12.972	0.018	β2MG > 3 mg/L	6.506	1.199-35.304	0.030
Ki67 > 40%	3.234	1.079-9.688	0.036	Ki67 > 40%	5.734	1.438-22.869	0.013
ECOG > 1	1.836	0.376-8.961	0.453	Extranodal sites ≥ 2	24.024	2.622-220.128	0.005
Elevated LDH	1.692	0.545-5.258	0.363	ECOG > 1	0.348	0.053-2.272	0.270
Advanced stage	5.124	0.599-43.830	0.136				

SUVmax: maximum standardized uptake value; β2MG: beta-2 microglobulin; ECOG: Eastern Cooperative Oncology Group; LDH: lactate dehydrogenase.

Identification of independent risk factors

Univariate analyses revealed that SUVmax ≥ 12, Ki67 > 40%, β2MG > 3 mg/L, and ECOG > 1 were significantly correlated with the occurrence of ETF in these two independent cohorts. In addition, elevated LDH levels and advanced stage were associated with ETF in the training cohort, and extranodal sites ≥ 2 was associated with ETF in the validation cohort (**Table 1**; $P < 0.05$). In the multivariate analyses, the common independent predictive factors for ETF were SUVmax ≥ 12, β2MG > 3 mg/L, and Ki67 > 40% in these two independent cohorts (**Table 2**; $P < 0.05$). There was no correlation between the grades of FL and ETF, but more patients with FL 3a had an SUVmax ≥ 12 than those with FL1-2 (48.8% vs. 28.6% in the training cohort; 52.9% vs. 12.8% in the validation cohort; $P < 0.05$). However, FLIPI1 and FLIPI2 were not independent prognostic factors for ETF, nor were the Ann Arbor stage or other clinical factors.

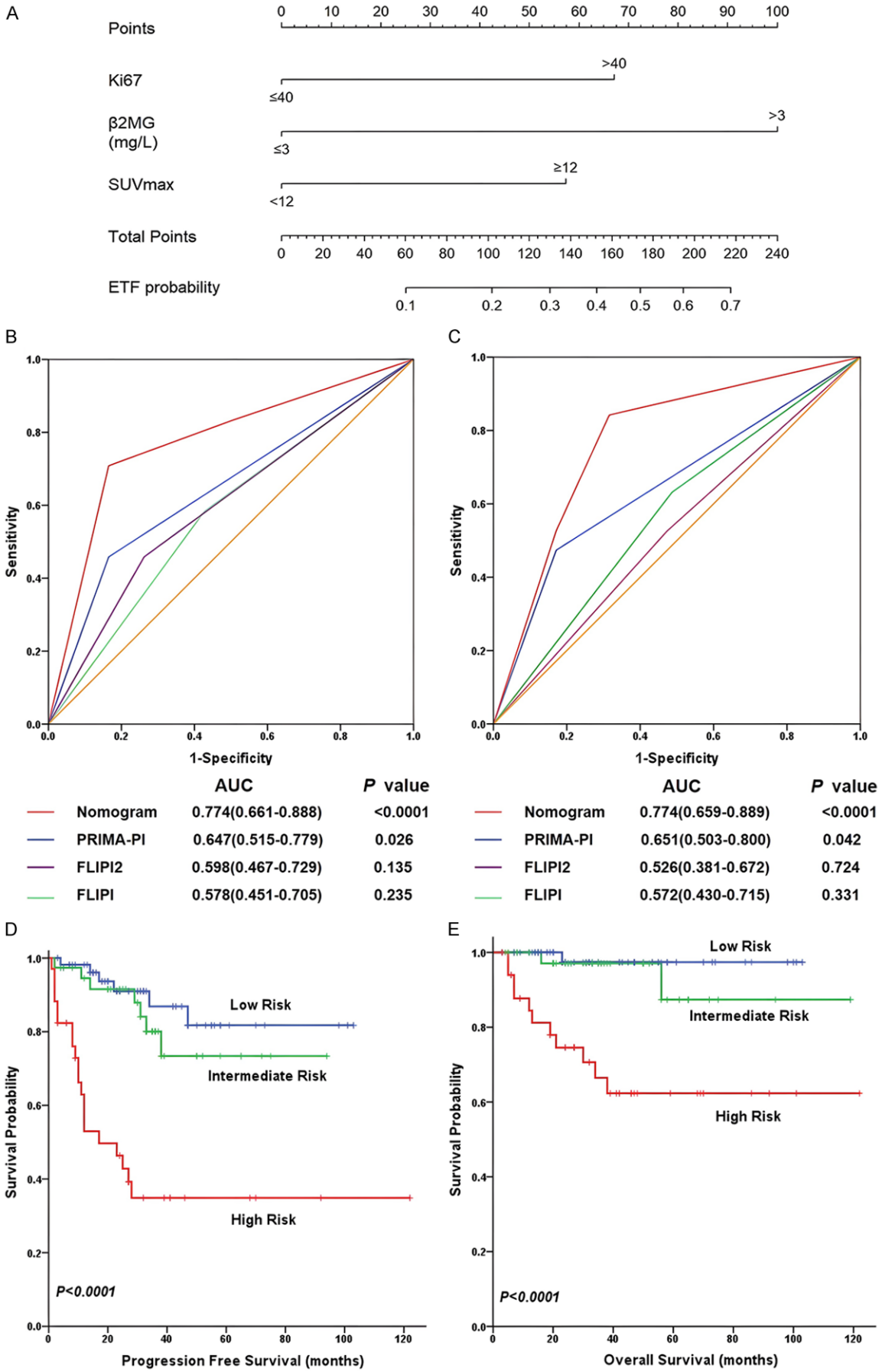
Performance of a predictive nomogram to identify ETF

Subsequently, we constructed a prognostic nomogram to predict ETF based on three independent predictive factors of ETF via multivariate analysis in the training cohort and confirmed it in the validation cohort (**Figure 3A**). The C-statistic for the nomogram was 0.832. ROC analysis showed an AUC of 0.774 in these two independent cohorts, indicating that this nomogram significantly improved prognostic accuracy (sensitivity 70.8%, specificity of 83.5% in the training cohort, sensitivity 84.2%, specificity 68.4% in the validation cohort) compared with other evaluation tools, such as PRIMA-PI, FLIPI, and FLIPI2 (**Figure 3B**, **3C**).

According to the prognostic nomogram score, patients were categorized into three subgroups (score: 0-57.5, 67.5-100, and > 100, referred to as low-, intermediate-, and high-risk groups, respectively), with significant differences in prognosis among each subgroup. In the training cohort, the 5-year PFS rates for patients were 81.7%, 73.4%, and 34.9%, respectively ($P < 0.0001$), and the 5-year OS rates were 97.4%, 87.4%, and 62.3%, respectively ($P < 0.0001$) (**Figure 3D**, **3E**). In the validation cohort, the 5-year PFS rates were 77.7%, 52.9%, and 34.8%, respectively ($P=0.001$), and the 5-year OS rates were 96.4%, 94.1%, and 73.7%, respectively ($P=0.028$) (**Figure 3F**, **3G**). These results indicate that our nomogram can successfully identify patients with FL who may experience ETF with poor prognosis.

Discussion

FL is a mostly incurable disease with a heterogeneous clinical course. Therefore, there are many efficacious treatment modalities currently available, ranging from watching and waiting to radiotherapy, chemotherapy, immunochemotherapy, monoclonal antibodies, autologous stem cell transplantation, and novel immunomodulatory agents. The introduction of rituximab significantly improves the OS of patients with FL [15]. In the GALLIUM study, Marcus et al. compared rituximab-combined chemothera-


A
Points
0 10 20 30 40 50 60 70 80 90 100
Ki67
≤40
>40
β2MG (mg/L)
≤3
>3
SUVmax
<12
≥12
Total Points
0 20 40 60 80 100 120 140 160 180 200 220 240
ETF probability
0.1 0.2 0.3 0.4 0.5 0.6 0.7
B
Sensitivity
1-Specificity
AUC
P value
Nomogram 0.774(0.661-0.888) <0.0001
PRIMA-PI 0.647(0.515-0.779) 0.026
FLIPI2 0.598(0.467-0.729) 0.135
FLIPI 0.578(0.451-0.705) 0.235
C
Sensitivity
1-Specificity
AUC
P value
Nomogram 0.774(0.659-0.889) <0.0001
PRIMA-PI 0.651(0.503-0.800) 0.042
FLIPI2 0.526(0.381-0.672) 0.724
FLIPI 0.572(0.430-0.715) 0.331
D
Survival Probability
Low Risk
Intermediate Risk
High Risk
P<0.0001
Progression Free Survival (months)
E
Survival Probability
Low Risk
Intermediate Risk
High Risk
P<0.0001
Overall Survival (months)

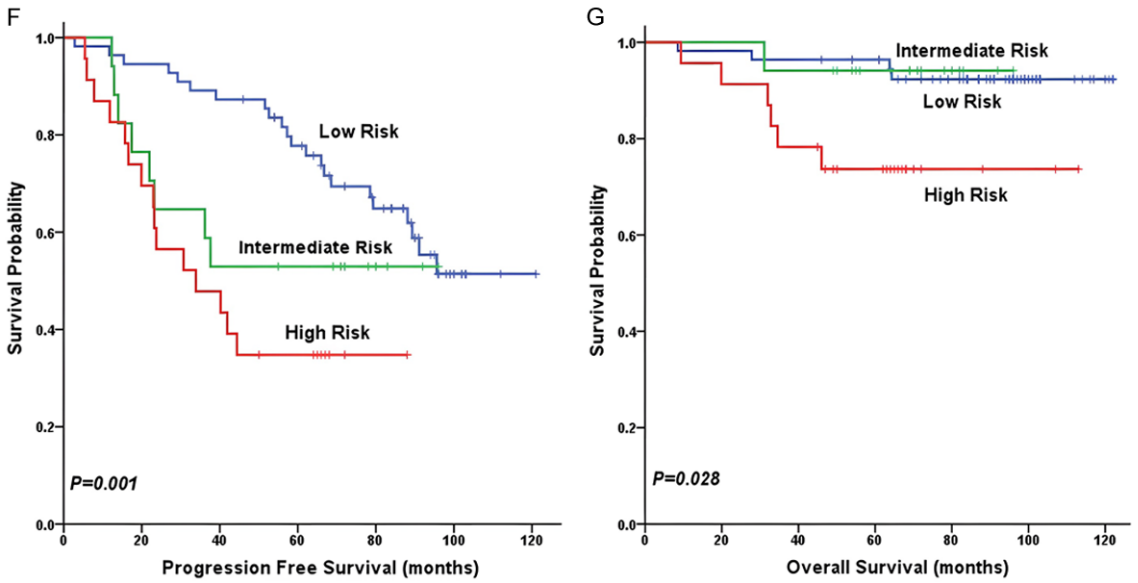


Figure 3. Nomogram to predict ETF of FL patients. (A) Nomogram to predict ETF of FL patients according to SUVmax ≥ 12, β2MG > 3 mg/L and Ki67 > 40% in the training cohort. ROC curves to assess the predictive accuracy of the nomogram, compared with the FLIPI, FLIPI2 and PRIMA-PI in the training cohort (B) and the validation cohort (C). Kaplan-Meier curves of PFS (D) and OS (E) in three risk groups according to nomogram in the training cohort; PFS (F) and OS (G) in the validation cohort.

py with obinutuzumab-combined chemotherapy in 1202 patients with previously untreated advanced-stage FL. The 3-year PFS was estimated to be 73.3% in the rituximab group and 80.0% in the obinutuzumab group [16]. In this study, > 80% of the patients survived for > 5 years. The 3- and 5-year PFS rates were 70.3% and 66.2%, respectively, which were shorter than those in clinical trials, reflecting the heterogeneity of the cohort in the clinical setting. Despite significant advances in FL treatment, a subgroup of approximately 20% of patients still experience a short remission and markedly inferior outcome because of ETF [17, 18]. ETF is defined as a group of very high-risk patients and is an accurate predictor of poor OS. Patients who developed ETF had a substantially increased risk of death within five years, and the OS was significantly worse than that in patients without early progression. In the National LymphoCare Study conducted by Casulo *et al*. [6], 19% of FL patients who received R-CHOP as the first-line treatment experienced early progression and a lower 5-year OS rate (50% vs. 90%). Similarly, 18.9% and 20% of patients in these two independent cohorts experienced ETF, respectively, with a lower 5-year OS than that in the ETF-neg group (42.3% vs. 96.3% and 62.7% vs. 97.4%, respectively). However, the likelihood of developing ETF remains unknown at the time of diagnosis. Thus, identifying patients who may experience early progression following first-line treatment and predicting outcomes at diagnosis of FL poses a major challenge. Currently, existing prognostic models still have some limitations in predicting ETF, such as the FLIPI, FLIPI2, or PRIMA-PI (**Table 3**). Jurinovic *et al*. [19] reported data from two independent cohorts and showed that m7-FLIPI had the highest accuracy in predicting early progression, with a sensitivity of approximately 61% and specificity of 79%, compared with FLIPI (sensitivity 78%, specificity 56%) and POD24-PI (sensitivity 78%, specificity 67%). In addition, another gene model, 23 gene-expression profiling (GEP), was used to predict early progression with a sensitivity of approximately 43% and specificity of 79% [20]. However, the m7-FLIPI and 23 GEP scores are not widely available in clinical practice because of the cost and requirement of infrastructure. In addition, Mir et al. constructed the FL Evaluation Index (FLEX) score to predict ETF, with a sensitivity of approximately 60% and specificity of 68% compared with FLIPI, FLIPI2 (sensitivity 53%, specificity 59%), and PRIMA-PI (sensitivity 69%, specificity 47%) [21]. Nevertheless, FLEX is a

Table 3. Prognostic Models in Follicular Lymphoma and Correlation With early progression

Prognostic Index	Prognostic Factors	Risk groups	Outcomes	Notes	Prediction of early progression	Reference
FLIPI	• Age > 60 y • Ann-Arbor III/IV • > 4 nodal regions • Hb < 12 g/dl • LDH elevate	LR (0-1 points, 36%) IR (2 points, 37%) HR (≥ 3 points, 27%)	10 y OS 71% 10 y OS 51% 10 y OS 36%	• Endpoint: OS • Established in the pre-rituximab era • ECOG, β2MG not included • Nodal areas cumbersome to count • Can't guide treatment	Sensitivity 78% Specificity 56%	[19, 33]
FLIPI2	• Age > 60 y • Hb < 12 g/dl • β2MG elevate • BMI • LodLIN ≥ 6 cm	LR (0 points, 20%) IR (1-2 points, 53%) HR (≥ 3 points, 27%)	5 y: PFS 80%, OS 98% 5 y: PFS 51%, OS 88% 5 y: PFS 19%, OS 77%	• Endpoint: PFS • Prospective design • Rituximab era (only 59% received R) • Short follow-up (median 3.2 y) • Can't guide treatment	Sensitivity 53% Specificity 59%	[21, 34]
m7-FLIPI	• FLIPI score > 2 • ECOG PS > 1 • Mutational genes (EZH2, ARID1A, EP300, MEF2B, FOX01, CREBBP, CARD11)	LR (index < 0.8, 72%) HR (index ≥ 0.8, 28%)	5 y: FFS 77%, OS 90% 5 y: FFS 38%, OS 65%	• Endpoint: FFS • Combined genetic information • Individual mutations were not independently prognostic in the MVA • limiting its applicability to patients treated with different regimens • Costly	Sensitivity 61% Specificity 79%	[19, 35]
POD24-PI	• HR FLIPI • Mutational genes (EP300, FOX01, EZH2)	LR (index < 0.71, 58%) HR (index > 0.71, 42%)	5 y: FFS 77%, OS 91% 5 y: FFS 50%, OS 71%	• Endpoint: POD24 • Failed to improve the overall performance of m7-FLIPI • Costly	Sensitivity 78% Specificity 67%	[19]
TMTV with FLIPI2	• TMTV > 510 cm^3 • FLIPI2 score 3-5	LR (0 points, 53%) IR (1 points, 33%) HR (2 points, 14%)	5 y: PFS 69%, OS 99% 5 y: PFS 46%, OS 85% 5 y: PFS 20%, OS 87%	• Endpoint: PFS • Similar OS for IR and HR groups • No external validation	Not reported	[28]
PRIMA-PI	• β2MG > 3 mg/L • BMI	LR (β2MG ≤ 3 mg/L and no BMI, 34%) IR (β2MG ≤ 3 mg/L and BMI, 34%) HR (β2MG > 3 mg/L, 32%)	5 y: PFS 69%, OS 93% 5 y: PFS 55%, OS 93% 5 y: PFS 37%, OS 81%	• Endpoint: PFS • Simple model • Similar OS for BMI alone or not	Sensitivity 69% Specificity 47%	[14, 21]
TMTV with EOI PET	• TMTV > 510 cm^3 • EOI PET positive	LR (0 risk factors, 64%) IR (1 risk factor, 27%) HR (2 risk factors, 8%)	5 y: PFS 67% 5 y: PFS 33% 5 y: FFS 23%	• Can't guide initial therapy • Acquisition of TMTV is not quite convenient • No difference in 2 y PFS between the IR and HR groups	2 y: PFS 90.2% 2 y: PFS 61.4% 2 y: PFS 46.2%	[29]
23-GEP	• Expression levels of 23 genes	LR (index < 1.075, 65%) HR (index ≥ 1.075, 35%)	5 y: PFS 73% 5 y: PFS 26%	• Endpoint: PFS • Absence of clinical variables • highly dependent on the treatment regimen (BR better) • Costly	Sensitivity 43% Specificity 79%	[20]
FLEX	• Male sex • SPD in the highest quartile • Grade 3a • > 2 extranodal areas • ECOG > 1 • Hb < 12 g/dl • β2MG elevate • NK count < 100/μL • LDH elevate	LR (0-2 risk factors, 64%) HR (3-9 risk factors, 36%)	3 y: PFS 86%, OS 97% 3 y: PFS 68%, OS 87%	• Endpoint: PFS • more accurate to predict PFS and OS than FLIPI, FLIPI2, and PRIMA-PI • Short follow-up (median 57 months) • complex model (9 risk factors) • Questionable biological rationale of NK cell count inclusion	Sensitivity 60% Specificity 68%	[21]
BioFLIPI	• lack of intra-follicular CD4 expression • HR FLIPI	BioFLIPI 1 (23.6%) BioFLIPI 2 (35.1%) BioFLIPI 3 (30.8%) BioFLIPI 4 (10.5%)	2 y: EFS24: 84.6% 2 y: EFS24: 69.5% 2 y: EFS24: 61.4% 2 y: EFS24: 46.1%	• Endpoint: EFS 12/24 • Prospective design • Failed to improve the accuracy	Sensitivity 72% Specificity 56%	[22]

y, years; ECOG, Eastern Cooperative Oncology Group performance status; LDH, lactate dehydrogenase; BMI, bone marrow involvement; Hb, hemoglobin; β2MG, beta-2 microglobulin; LodLIN, longest diameter of the largest node; R, rituximab; BR, bendamustine with rituximab; OS, overall survival; FFS, failure-free survival; PFS, progression-free survival; EFS, early failure survival; POD24, progression of disease within 24 months; LR, low risk; IR, intermediate risk; HR, high risk; MVA, multivariate analysis; TMTV, total metabolic tumor volume; NK, natural killer; SPD, sum of the products of lesion diameters; EOI PET, end-of-induction treatment PET/CT.

complex model that requires nine risk factors, including the NK cell count, which is questionable in clinical settings. In addition, there are immuno-microenvironment-related prognostic scores for predicting early progression, such as BioFLIPI, including two factors (lack of intrafollicular CD4 expression and high-risk FLIPI), with a sensitivity of 72% and specificity of 56%. However, the accuracy of BioFLIPI is lower than that of FLIPI (sensitivity 60%, specificity 69%) and 23 GEP score (sensitivity 44%, specificity 82%) [22]. Therefore, ongoing studies are aimed at identifying more accurate clinical, metabolic, biological, and genetic factors that may predict the likelihood of early recurrence.

Studies have shown that FDG PET/CT is a more sensitive staging method for FL patients than CT [23, 24], and FDG PET/CT at the end of induction therapy can predict PFS, OS, and early progression [25-27]. SUVmax and TMTV, both critical parameters of PET/CT, have been applied to FL. Baseline TMTV > 510 cm^3 combined with FLIPI2 can identify patients at a high risk of early progression, with a 5-year PFS of 20% in the high-risk group [28]. In addition, Cottereau *et al*. [29] established a model that incorporates baseline TMTV and end of induction treatment PET/CT (EOI PET) to identify FL patients at a very high risk of early progression, in which the 5-year PFS of patients with TMTV > 510 cm^3 and positive EOI PET was only 23%. However, this model cannot predict patients' experience of ETF at diagnosis, and the acquisition of TMTV is not convenient in clinical settings. Furthermore, a recent retrospective study reported that baseline SUVmax instead of TMTV was associated with poorer prognosis in patients with FL1-3a [9].

The SUVmax in PET/CT, as a semi-quantitative scale, has been widely used to evaluate the prognosis of patients with FL. In previous retrospective studies, it was recommended to set the cutoff value of SUVmax as 10, 14, and 17 to identify FL patients with a higher transformation risk [30-32]. In addition, a retrospective study indicated that Ki67 ≥ 10% correlated to increased SUVmax, which implied that baseline SUVmax and Ki67 could define patients with a high risk of relapse or progression [9]. However, the relationship between the PET/CT SUVmax and ETF remains unclear. We proposed combining established clinical risk factors with SUVmax to identify high-risk patients who may experience an ETF. In this analysis, the independent predictive factors for ETF were SUVmax ≥ 12, β2MG > 3 mg/L, and Ki67 > 40%. A combination of these three adverse factors could classify the patients into three risk groups according to the nomogram, with patients in the high-risk group experiencing poor survival. In the training cohort, the 5-year OS rates were 97.4%, 87.4%, and 62.3% for the low-, intermediate-, and high-risk groups, respectively. Moreover, the 5-year OS rates were 96.4%, 94.1%, and 73.7% in the validation cohort, respectively. However, the prognosis between the low- and intermediate-risk groups was not statistically different; therefore, an extended follow-up may be required. This nomogram, based on easily available clinical data, is convenient, feasible, and has significant prognostic accuracy (sensitivity, 70.8%; specificity, 83.5% in the training cohort; sensitivity, 84.2%; specificity, 68.4% in the validation cohort) compared with other prognostic evaluation tools. Significantly, in this study, more patients (94.5%) received anthracycline-based immunochemotherapy in the training cohort, but 86.3% received bendamustine-based immunochemotherapy in the validation cohort, which may indicate that this nomogram was not restricted by the treatment regimens. In addition, it is worth noting that more patients with FL 3a had SUVmax ≥ 12 than those with FL1-2 in these two independent cohorts, which may indicate a limitation in the histological diagnosis of FL. To the best of our knowledge, this is the first nomogram to predict the probability of ETF in patients with FL. This integrated approach, rather than each parameter alone, can identify FL patients with a higher risk of early progression and treatment failure for whom new agents or intensive treatment should be considered at diagnosis.

In conclusion, ETF in patients with FL treated with immunochemotherapy marks a unique group of patients with a high risk of early death. In this study, SUVmax ≥ 12, β2MG > 3 mg/L, and Ki67 > 40% were independent predictive factors for ETF. Our study applied a nomogram to predict FL patients with a high risk of ETF at diagnosis, and it may represent a promising new tool for identifying a subset of patients who may benefit from individualized targeted therapy, because of the risk of early treatment

failure within two years. However, this study has some limitations. We will try to explore artificial intelligence or radiomics to further analyze PET/CT or CT data in detail and incorporate into the nomogram. Besides, this was a retrospective study, and the nomogram should be evaluated in prospective trials with a large sample size in the future.

Acknowledgements

This work was supported by Science and Technology Agency of Jilin province (20200-201591JC).

Disclosure of conflict of interest

None.

Address correspondence to: Ou Bai, Department of Hematology, The First Hospital of Jilin University, Jilin, China. Tel: 0431-88782054; Fax: 88786134; E-mail: baiou@jlu.edu.cn; Ken H Young, Division of Hematopathology, Department of Pathology, Duke University Medical Center and Duke Cancer Institute, Durham, NC, USA. Tel: 1-919-668-7568; Fax: 1-919-684-1856; E-mail: ken.young@duke.edu

References

[1] Becnel MR and Nastoupil LJ. Follicular lymphoma: past, present, and future. Curr Treat Options Oncol 2018; 19: 32.

[2] Tan D, Horning SJ, Hoppe RT, Levy R, Rosenberg SA, Sigal BM, Warnke RA, Natkunam Y, Han SS, Yuen A, Plevritis SK and Advani RH. Improvements in observed and relative survival in follicular grade 1-2 lymphoma during 4 decades: the Stanford University experience. Blood 2013; 122: 981-987.

[3] Wang S, Liu Y, Feng Y, Zhang J, Swinnen J, Li Y and Ni Y. A review on curability of cancers: more efforts for novel therapeutic options are needed. Cancers (Basel) 2019; 11: 1782.

[4] Maeda H and Khatami M. Analyses of repeated failures in cancer therapy for solid tumors: poor tumor-selective drug delivery, low therapeutic efficacy and unsustainable costs. Clin Transl Med 2018; 7: 11.

[5] Casulo C and Barr PM. How I treat early-relapsing follicular lymphoma. Blood 2019; 133: 1540-1547.

[6] Casulo C, Byrtek M, Dawson KL, Zhou X, Farber CM, Flowers CR, Hainsworth JD, Maurer MJ, Cerhan JR, Link BK, Zelenetz AD and Friedberg JW. 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: 2516-2522.

[7] Launonen A, Hiddemann W, Duenzinger U, Fingerle-Rowson G, Nielsen T and Marcus R. Early disease progression predicts poorer survival in patients with follicular lymphoma (FL) in the GALLIUM Study. Blood 2017; S1: 1490.

[8] Strati P, Ahmed MA, Fowler NH, Nastoupil LJ, Samaniego F, Fayad LE, Hagemeister FB, Romaguera JE, Rodriguez A, Wang M, Westin JR, Cheah C, Noorani M, Feng L, Davis RE and Neelapu SS. Pre-treatment maximum standardized uptake value predicts outcome after frontline therapy in patients with advanced stage follicular lymphoma. Haematologica 2020; 105: 1907-1913.

[9] Rossi C, Tosolini M, Gravelle P, Pericart S, Kanoun S, Evrard S, Gilhodes J, Franchini DM, Amara N, Syrykh C, Bories P, Oberic L, Ysebaert L, Martin L, Ramla S, Robert P, Tabouret-Viaud C, Casasnovas RO, Fournié JJ, Bezombes C and Laurent C. Baseline SUVmax is related to tumor cell proliferation and patient outcome in follicular lymphoma. Haematologica 2022; 107: 221-230.

[10] Yamamoto E, Tomita N, Sakata S, Tsuyama N, Takeuchi K, Nakajima Y, Miyashita K, Tachibana T, Takasaki H, Tanaka M, Hashimoto C, Koharazawa H, Fujimaki K, Taguchi J, Harano H, Motomura S and Ishigatsubo Y. MIB-1 labeling index as a prognostic factor for patients with follicular lymphoma treated with rituximab plus CHOP therapy. Cancer Sci 2013; 104: 1670-1674.

[11] Kedmi M, Hedvat CV, Maragulia J, Zhang Z and Zelenetz AD. Association of quantitative assessment of the intrafollicular proliferation index with outcome in follicular lymphoma. Br J Haematol 2014; 164: 646-652.

[12] Brice P, Bastion Y, Lepage E, Brousse N, Haïoun C, Moreau P, Straetmans N, Tilly H, Tabah I and Solal-Céligny P. 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. Groupe d'Etude des Lymphomes de l'Adulte. J Clin Oncol 1997; 15: 1110-1117.

[13] Nastoupil LJ, Sinha R, Byrtek M, Ziemiecki R, Zhou X, Taylor M, Friedberg JW, Link BK, Cerhan JR, Dawson K and Flowers CR. Outcomes following watchful waiting for stage II-IV follicular lymphoma patients in the modern era. Br J Haematol 2016; 172: 724-734.

[14] Bachy E, Maurer MJ, Habermann TM, Gelas-Dore B, Maucort-Boulch D, Estell JA, Van den Neste E, Bouabdallah R, Gyan E, Feldman AL,

Bargay J, Delmer A, Slager SL, Gomes da Silva M, Fitoussi O, Belada D, Maisonneuve H, Intragumtornchai T, Ansell SM, Lamy T, Dartigues P, Link BK, Seymour JF, Cerhan JR and Salles G. A simplified scoring system in de novo follicular lymphoma treated initially with immunochemotherapy. Blood 2018; 132: 49-58.

[15] Sarkozy C, Maurer MJ, Link BK, Ghesquieres H, Nicolas E, Thompson CA, Traverse-Glehen A, Feldman AL, Allmer C, Slager SL, Ansell SM, Habermann TM, Bachy E, Cerhan JR and Salles G. Cause of death in follicular lymphoma in the first decade of the rituximab era: a pooled analysis of French and US cohorts. J Clin Oncol 2019; 37: 144-152.

[16] Marcus R, Davies A, Ando K, Klapper W, Opat S, Owen C, Phillips E, Sangha R, Schlag R, Seymour JF, Townsend W, Trněný M, Wenger M, Fingerle-Rowson G, Rufibach K, Moore T, Herold M and Hiddemann W. Obinutuzumab for the first-line treatment of follicular lymphoma. N Engl J Med 2017; 377: 1331-1344.

[17] Bitansky G, Avigdor A, Vasilev E, Zlotnick M, Ribakovsky E, Benjamini O, Nagler A and Kedmi M. Progression of disease within 24 months of initial therapy (POD24) detected incidentally in imaging does not necessarily indicate worse outcome. Leuk Lymphoma 2020; 61: 2645-2651.

[18] Sortais C, Lok A, Tessoulin B, Gastinne T, Mahé B, Dubruille V, Blin N, Touzeau C, Moreau A, Bossard C, Peterlin P, Garnier A, Guillaume T, Le Bourgeois A, Chevallier P, Moreau P, Leux C and Le Gouill S. Progression of disease within 2 years (POD24) is a clinically relevant endpoint to identify high-risk follicular lymphoma patients in real life. Ann Hematol 2020; 99: 1595-1604.

[19] Jurinovic V, Kridel R, Staiger AM, Szczepanowski M, Horn H, Dreyling MH, Rosenwald A, Ott G, Klapper W, Zelenetz AD, Barr PM, Friedberg JW, Ansell S, Sehn LH, Connors JM, Gascoyne RD, Hiddemann W, Unterhalt M, Weinstock DM and Weigert O. Clinicogenetic risk models predict early progression of follicular lymphoma after first-line immunochemotherapy. Blood 2016; 128: 1112-1120.

[20] Huet S, Tesson B, Jais JP, Feldman AL, Magnano L, Thomas E, Traverse-Glehen A, Albaud B, Carrère M, Xerri L, Ansell SM, Baseggio L, Reyes C, Tarte K, Boyault S, Haioun C, Link BK, Feugier P, Lopez-Guillermo A, Tilly H, Brice P, Hayette S, Jardin F, Offner F, Sujobert P, Gentien D, Viari A, Campo E, Cerhan JR and Salles G. A gene-expression profiling score for prediction of outcome in patients with follicular lymphoma: a retrospective training and validation analysis in three international cohorts. Lancet Oncol 2018; 19: 549-561.

[21] Mir F, Mattiello F, Grigg A, Herold M, Hiddemann W, Marcus R, Seymour JF, Bolen CR, Knapp A, Nielsen T and Casulo C. 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: 1503-1510.

[22] Mondello P, Fama A, Larson MC, Feldman AL, Villasboas JC, Yang ZZ, Galkin I, Svelolkin V, Postovalova E, Bagaev A, Ovcharov P, Varlamova A, Huet S, Tesson B, McGrath KR, Slager S, Link BK, Syrbu S, Novak AJ, Habermann TM, Witzig TE, Nowakowski GS, Salles G, Cerhan JR and Ansell SM. Lack of intrafollicular memory CD4+T cells is predictive of early clinical failure in newly diagnosed follicular lymphoma. Blood Cancer J 2021; 11: 130.

[23] Le Dortz L, De Guibert S, Bayat S, Devillers A, Houot R, Rolland Y, Cuggia M, Le Jeune F, Bahri H, Barge ML, Lamy T and Garin E. Diagnostic and prognostic impact of 18F-FDG PET/CT in follicular lymphoma. Eur J Nucl Med Mol Imaging 2010; 37: 2307-2314.

[24] Luminari S, Biasoli I, Arcaini L, Versari A, Rusconi C, Merli F, Spina M, Ferreri AJ, Zinzani PL, Gallamini A, Mastronardi S, Boccomini C, Gaidano G, D'Arco AM, Di Raimondo F, Carella AM, Santoro A, Musto P and Federico M. 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: 2108-2112.

[25] Dupuis J, Berriolo-Riedinger A, Julian A, Brice P, Tychyj-Pinel C, Tilly H, Mounier N, Gallamini A, Feugier P, Soubeyran P, Colombat P, Laurent G, Berenger N, Casasnovas RO, Vera P, Paone G, Xerri L, Salles G, Haioun C and Meignan M. Impact of [(18)F]fluorodeoxyglucose positron emission tomography response evaluation in patients with high-tumor burden follicular lymphoma treated with immunochemotherapy: a prospective study from the Groupe d'Etudes des Lymphomes de l'Adulte and GOELAMS. J Clin Oncol 2012; 30: 4317-4322.

[26] Luminari S, Biasoli I, Versari A, Rattotti S, Bottelli C, Rusconi C, Merli F, Spina M, Ferreri AJ, Zinzani PL, Gallamini A, Franceschetto A, Boccomini C, Franceschetti S, Salvi F, Raimondo FD, Carella AM, Micol Q, Balzarotti M, Musto P and Federico M. The prognostic role of post-induction FDG-PET in patients with follicular lymphoma: a subset analysis from the FOLL05 trial of the Fondazione Italiana Linfomi (FIL). Ann Oncol 2014; 25: 442-447.

[27] Trotman J, Luminari S, Boussetta S, Versari A, Dupuis J, Tychyj C, Marcheselli L, Berriolo-

Riedinger A, Franceschetto A, Julian A, Ricard F, Guerra L, Haioun C, Biasoli I, Tilly H, Federico M, Salles G and Meignan M. Prognostic value of PET-CT after first-line therapy in patients with follicular lymphoma: a pooled analysis of central scan review in three multicentre studies. Lancet Haematol 2014; 1: e17-27.

[28] Meignan M, Cottereau AS, Versari A, Chartier L, Dupuis J, Boussetta S, Grassi I, Casasnovas RO, Haioun C, Tilly H, Tarantino V, Dubreuil J, Federico M, Salles G, Luminari S and Trotman J. Baseline metabolic tumor volume predicts outcome in high-tumor-burden follicular lymphoma: a pooled analysis of three multicenter Studies. J Clin Oncol 2016; 34: 3618-3626.

[29] Cottereau AS, Versari A, Luminari S, Dupuis J, Chartier L, Casasnovas RO, Berriolo-Riedinger A, Menga M, Haioun C, Tilly H, Tarantino V, Federico M, Salles G, Trotman J and Meignan M. Prognostic model for high-tumor-burden follicular lymphoma integrating baseline and end-induction PET: a LYSA/FIL study. Blood 2018; 131: 2449-2453.

[30] Schöder H, Noy A, Gönen M, Weng L, Green D, Erdi YE, Larson SM and Yeung HW. Intensity of 18fluorodeoxyglucose uptake in positron emission tomography distinguishes between indolent and aggressive non-Hodgkin's lymphoma. J Clin Oncol 2005; 23: 4643-4651.

[31] Bodet-Milin C, Kraeber-Bodéré F, Moreau P, Campion L, Dupas B and Le Gouill S. Investigation of FDG-PET/CT imaging to guide biopsies in the detection of histological transformation of indolent lymphoma. Haematologica 2008; 93: 471-472.

[32] Karam M, Feustel PJ, Vera CD and Nazeer T. Features of large cell transformation of indolent lymphomas as observed on sequential PET/CT. Nucl Med Commun 2011; 32: 177-185.

[33] Solal-Céligny P, Roy P, Colombat P, White J, Armitage JO, Arranz-Saez R, Au WY, Bellei M, Brice P, Caballero D, Coiffier B, Conde-Garcia E, Doyen C, Federico M, Fisher RI, Garcia-Conde JF, Guglielmi C, Hagenbeek A, Haïoun C, LeBlanc M, Lister AT, Lopez-Guillermo A, McLaughlin P, Milpied N, Morel P, Mounier N, Proctor SJ, Rohatiner A, Smith P, Soubeyran P, Tilly H, Vitolo U, Zinzani PL, Zucca E and Montserrat E. Follicular lymphoma international prognostic index. Blood 2004; 104: 1258-1265.

[34] Federico M, Bellei M, Marcheselli L, Luminari S, Lopez-Guillermo A, Vitolo U, Pro B, Pileri S, Pulsoni A, Soubeyran P, Cortelazzo S, Martinelli G, Martelli M, Rigacci L, Arcaini L, Di Raimondo F, Merli F, Sabattini E, McLaughlin P and Solal-Céligny P. 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: 4555-4562.

[35] Pastore A, Jurinovic V, Kridel R, Hoster E, Staiger AM, Szczepanowski M, Pott C, Kopp N, Murakami M, Horn H, Leich E, Moccia AA, Mottok A, Sunkavalli A, Van Hummelen P, Ducar M, Ennishi D, Shulha HP, Hother C, Connors JM, Sehn LH, Dreyling M, Neuberg D, Möller P, Feller AC, Hansmann ML, Stein H, Rosenwald A, Ott G, Klapper W, Unterhalt M, Hiddemann W, Gascoyne RD, Weinstock DM and Weigert O. Integration of gene mutations in risk prognostication for patients receiving first-line immunochemotherapy for follicular lymphoma: a retrospective analysis of a prospective clinical trial and validation in a population-based registry. Lancet Oncol 2015; 16: 1111-1122.