

LYMPHOID NEOPLASIA

A simplified scoring system in de novo follicular lymphoma treated initially with immunochemotherapy

Emmanuel Bachy,¹ Matthew J. Maurer,² Thomas M. Habermann,³ Bénédicte Gelas-Dore,⁴ Delphine Maucourt-Boulch,^{5,6} Jane A. Estell,^{7,8} Eric Van den Neste,⁹ Réda Bouabdallah,¹⁰ Emmanuel Gyan,¹¹ Andrew L. Feldman,¹² Joan Bargay,¹³ Alain Delmer,¹⁴ Susan L. Slager,² Maria Gomes da Silva,¹⁵ Olivier Fitoussi,¹⁶ David Belada,^{17,18} Hervé Maisonneuve,¹⁹ Tanin Intragumtornchai,²⁰ Stephen M. Ansell,³ Thierry Lamy,²¹ Peggy Dartigues,²² Brian K. Link,²³ John F. Seymour,²⁴ James R. Cerhan,² and Gilles Salles¹

¹Department of Hematology, Hospices Civils de Lyon and Université Claude Bernard, Université de Lyon, INSERM102, Centre National de la Recherche Scientifique (CNRS)5286, Pierre-Bénite, France; ²Department of Health Sciences Research, Mayo Clinic, Rochester, MN; ³Division of Hematology, Mayo Clinic, Rochester, MN; ⁴Department of Biostatistics, Lysarc, Lyon, France; ⁵Hospices Civils de Lyon, Service de Biostatistique et Bioinformatique & Université Lyon 1, Lyon, France; ⁶CNRS, Unité Mixte de Recherche 5558, Laboratoire de Biométrie et Biologie Evolutive, Equipe Biostatistique-Santé, Villeurbanne, France; ⁷Department of Haematology, Concord Repatriation General Hospital, Concord, Sydney, NSW, Australia; ⁸Sydney Medical School, University of Sydney, Sydney, NSW, Australia; ⁹Hematology Department, Cliniques Universitaires UCL Saint-Luc, Brussels, Belgium; ¹⁰Institut Paoli Calmettes, Marseille, France; ¹¹Department of Hematology, Centre Hospitalier Universitaire (CHU) de Tours, Tours, France; ¹²Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN; ¹³Hospital Son Llàtzer, Palma de Mallorca, Spain; ¹⁴CHU de Reims and Université Reims Champagne Ardenne, Reims, France; ¹⁵Portuguese Institute of Oncology, Lisbon, Portugal; ¹⁶Polyclinique Bordeaux Nord Aquitaine, Bordeaux, France; ¹⁷4th Department of Internal Medicine—Hematology, University Hospital, Hradec Kralove, Czech Republic; ¹⁸Faculty of Medicine, Charles University in Prague, Hradec Kralove, Czech Republic; ¹⁹Centre Hospitalier de la Roche sur Yon, Roche sur Yon, France; ²⁰Chulalongkorn University, Bangkok, Thailand; ²¹Centre Hospitalier Pontchaillou, Rennes, France; ²²Institut Gustave Roussy, Villejuif, France; ²³Department of Internal Medicine, University of Iowa College of Medicine, Iowa City, IA; and ²⁴Peter MacCallum Cancer Centre and University of Melbourne, Melbourne, VIC, Australia

KEY POINTS

- In FL, no prognostic index has been built based solely on a large cohort of patients treated with initial immunochemotherapy.
- The PRIMA-PI is a simplified index based on β_2m and bone marrow involvement for patients treated with immunochemotherapy.

In follicular lymphoma (FL), no prognostic index has been built based solely on a cohort of patients treated with initial immunochemotherapy. There is currently a need to define parsimonious clinical models for trial stratification and to add on biomolecular factors. Here, we confirmed the validity of both the follicular lymphoma international prognostic index (FLIPI) and the FLIPI2 in the large prospective PRIMA trial cohort of 1135 patients treated with initial R-chemotherapy \pm R maintenance. Furthermore, we developed a new prognostic tool comprising only 2 simple parameters (bone marrow involvement and β_2m -microglobulin [β_2m]) to predict progression-free survival (PFS). The final simplified score, called the PRIMA-PI (PRIMA-prognostic index), comprised 3 risk categories: high ($\beta_2m > 3$ mg/L), low ($\beta_2m \leq 3$ mg/L without bone marrow involvement), and intermediate ($\beta_2m \leq 3$ mg/L with bone marrow involvement). Five-year PFS rates were 69%, 55%, and 37% in the low-, intermediate-, and high-risk groups, respectively ($P < .0001$). In addition, achieving event-free survival (EFS) or not at 24 months (EFS24) was a strong posttreatment prognostic parameter for subsequent overall survival, and the PRIMA-PI was correlated

with EFS24. The results were confirmed in a pooled external validation cohort of 479 patients from the FL2000 LYSA trial and the University of Iowa/Mayo Clinic Lymphoma Specialized Program of Research Excellence Molecular Epidemiology Resource. Five-year EFS in the validation cohort was 77%, 57%, and 44% in the PRIMA-PI low-, intermediate-, and high-risk groups, respectively ($P < .0001$). The PRIMA-PI is a novel and easy-to-compute prognostic index for patients initially treated with immunochemotherapy. This could serve as a basis for building more sophisticated and integrated biomolecular scores. (*Blood*. 2018;132(1):49-58)

Introduction

Follicular lymphoma (FL) is 1 of the most common non-Hodgkin lymphoma, accounting for about 20% to 30% of all cases.^{1,2} The course of the disease is characterized by the responsiveness to initial therapy followed by repeated relapses and/or transformation to high-grade non-Hodgkin lymphoma. Treatment options differ widely according to disease stage (limited vs disseminated disease) or the presence of a high tumor burden criterion.³

Immunochemotherapy consisting of an anti-CD20 monoclonal antibody in association with an alkylating agent, a vinca alkaloid with or without anthracyclines,⁴⁻⁹ or bendamustine^{10,11} is widely accepted as a standard of care for stage III/IV FL presenting with at least 1 high tumor burden criterion.^{12,13}

Numerous individual parameters were shown to have prognostic significance in the disease related to the patient (age, sex), the disease itself (stage, bone marrow involvement, serum lactate

Table 1. Patient characteristics in the training and validation cohorts

	Training (PRIMA) (n = 1135), N (%)	Validation (FL2000+MER) (n = 479), N (%)
Age >60 y	402 (35)	225 (47)
Male sex	590 (52)	261 (54)
ECOG >1	49 (4)	24 (5)
β_2m >3 mg/L	341 (30)	163 (34)
Stage III-IV	1026 (90)	388 (81)
Nodal sites involvement >4	639 (56)	211 (44)
Bone marrow involvement	635 (56)	238 (50)
Extranodal sites involvement (other than bone marrow)	598 (53)	169 (35)
LDH > UNL	378 (33)	139 (29)
Hemoglobin <12 g/dL	239 (21)	91 (19)
B symptoms*	363 (32)	
LoDLIN >6 cm*	508 (45)	
Effusion syndrome*	150 (13)	
Compression syndrome*	208 (18)	
Circulating malignant cells*	92 (9)	
Platelets <150 \times 10 ⁹ /L*	181 (16)	
Albumin <40 g/L*	301 (33)	
Induction treatment		
R-CHOP	840 (74)	121 (25)
R-CVP	253 (22)	68 (14)
R-FCM	42 (4)	0 (0)
R-bendamustine	0 (0)	104 (22)
R-CHVP+IFN	0 (0)	175 (37)
Others	0 (0)	11 (2)
Maintenance†		
Rituximab	489 (50)	95 (20)
IFN	0 (0)	175 (36)
None	485 (50)	209 (44)

Missing data: for the training cohort: β_2m (n = 85), bone marrow involvement (n = 34), LDH (n = 5), LoDLIN (n = 7), effusion syndrome (n = 23), circulating malignant cells (n = 122), platelets (n = 1), albumin (n = 223); for the validation cohort: ECOG (n = 4), β_2m (n = 11), stage (n = 8), nodal sites (n = 21), bone marrow involvement (n = 24), extranodal sites (n = 9), LDH (n = 26), hemoglobin (n = 16).

IFN, interferon; LoDLIN, longest diameter of the largest involved node; UNL, upper normal limit.

*Data not extracted in the validation cohort.

†Only responding patients could be eligible for maintenance therapy.

dehydrogenase [LDH], β_2 -microglobulin [β_2m]), or the consequences of the disease (performance status [PS], systemic symptoms). To date, several indices have been proposed to describe the heterogeneity of the disease and refine prognosis. A first multi-institutional score was proposed in 2000 by the Intergruppo Italiano Linfomi,¹⁴ followed by the follicular lymphoma international prognostic index (FLIPI) in 2004¹⁵ and the more recently published FLIPI2 in 2009.¹⁶ A simplified scoring system based on LDH and β_2m levels was also proposed by Press

and colleagues in 2013.¹⁷ During the last few years, new biomolecular scores have been developed, such as the m7-FLIPI, taking into account bioclinical prognostic parameters (FLIPI, PS) and mutational status in a set of defined genes (EZH2, FOXO1, EP300, CREBBP, CARD11, MEF2B, ARID1A).¹⁸

However, determination of the number of nodal sites is usually cumbersome and error-prone in routine practice for FLIPI assessment.¹⁵ The FLIPI2 circumvented this fastidious computation by

assessing tumor bulk through the use of longest diameter of the largest involved node.¹⁶ However, the FLIPI2 has not supplanted the FLIPI for FL prognostication in routine practice because of inconsistent superiority in validation cohorts.^{17,19,20}

Because the FLIPI was built on a cohort of patients treated without immunotherapy, and as only 59% of patients received rituximab as part of frontline therapy for the construction of the FLIPI2,^{15,16} the first objective of this study was to assess and compare the prognostic value of these previously published indexes in the PRIMA cohort of patients homogeneously treated with a rituximab-containing induction regimen and followed by rituximab maintenance for half of them. The second objective was to develop and validate a new simplified scoring system for progression-free survival (PFS) in patients with FL homogeneously treated with immunochemotherapy.

Patients and methods

Study population

The randomized, open-label PRIMA study enrolled 1217 patients with de novo FL from 223 centers in 25 countries, and 1193 patients received induction treatment. The final population considered in the study comprised 1135 patients with histologically confirmed grade 1, 2, or 3a FL. Patients achieving at least a partial response after frontline therapy with physician-selected R-CHOP (rituximab plus cyclophosphamide, doxorubicin, vincristine and prednisone), R-CVP (rituximab plus cyclophosphamide, vincristine and prednisone), or R-FCM (rituximab plus fludarabine, cyclophosphamide and mitoxantrone) were randomly assigned between 2-year rituximab maintenance (every 8 weeks) or observation.²¹ Patients with grade 1, 2, or 3a FL were eligible for induction if they were older than 18 years and presented with at least 1 high-burden criterion among the following: any nodal or extranodal tumor mass with a diameter >7 cm, involvement of three nodal sites with a diameter >3 cm, the presence of systemic symptoms, substantial splenic enlargement, presence of any compression syndrome (ureteral, orbital, gastrointestinal) or serous effusion (irrespective of cell content), or elevated serum levels of LDH or β_2m . Patients with responding disease (either complete or partial response) were randomly assigned in a 1:1 ratio to receive 2 years of rituximab maintenance therapy (375 mg/m² IV every 8 weeks) or observation, but all enrolled patients before the induction phase were considered in the present study.

The protocol was approved by local or national ethics committees according to the laws of each country, and the study was in accordance with the Declaration of Helsinki. All patients gave written informed consent before registration. The PRIMA study was registered on the National Institutes of Health website, number NCT00140582. The trial demonstrated a significant improvement of the primary endpoint (ie, PFS from randomization) in the rituximab maintenance group.²¹ Three-year PFS was 74.9% (95% confidence interval [CI], 70.9-78.9) in the rituximab maintenance group and 57.6% (95% CI, 53.2-62.0) in the observation group (stratified log rank, $P < .0001$).

External validation cohorts of patients

One hundred seventy-five patients with high-tumor-burden FL from the rituximab-containing arm of treatment (R-CHVP+interferon) from the FL2000 trial^{8,21} and 304 patients prospectively

enrolled in the Molecular Epidemiology Resource (MER) of the University of Iowa/Mayo Clinic Lymphoma Specialized Program of Research Excellence treated with initial immunochemotherapy served as a pooled validation cohort. For the FL2000 cohort, patients were required to have Ann Arbor stage II-IV disease with at least 1 criterion of high tumor burden. Patients in the R-CHVP+interferon group received 6 monthly courses consisting of rituximab (375 mg/m²), cyclophosphamide (600 mg/m² IV on day 1), doxorubicin (25 mg/m² IV on day 1), etoposide (100 mg/m² IV on day 1), and prednisolone (40 mg/m² orally from days 1 to 5) combined with interferon- α 2a (4.5 or 3 million units per injection 3 times a week for patients aged ≤ 70 or >70 years, respectively), followed by 18 months of interferon- α 2a maintenance. For the MER cohort, briefly, since September 2002, enrollment was offered to consecutive, newly diagnosed patients with lymphoma (within 9 months) who were evaluated at the University of Iowa or Mayo Clinic Rochester and who were aged 18 years or older, had no history of HIV infection, and were residents of the United States.²² All diagnoses were confirmed by study hematopathologists. Baseline clinical, laboratory, and treatment data were abstracted from medical records, using a standard protocol. Patient management including treatments was per treating physician. All participants were systematically contacted every 6 months for the first 3 years, and then annually thereafter. Inclusion criteria for this analysis were initial diagnosis of grade 1-3a FL and being enrolled from 1 September 2002 to 31 December 2012. Patients with primary cutaneous follicle center lymphoma, FL grade 3b subtype, and a composite diagnosis including another (nonfollicular) lymphoma subtype, posttransplant lymphoproliferative disorder, or evidence of clinical or pathological transformation at the time of initial FL diagnosis were excluded. This study was reviewed and approved by the human subject institutional review board at the Mayo Clinic and the University of Iowa, and written informed consent was obtained from all participants.

Statistical methods

In the present study, PFS was measured from the date of treatment initiation to the date of death from any cause, disease relapse or progression, or the date of last contact. Overall survival (OS) was calculated from the date of treatment initiation until the date of death from any cause or the date of last contact. PFS was considered as the primary efficacy endpoint for model building. Event-free survival (EFS) was the outcome endpoint used in the MER cohort and was therefore used for the validation combined cohort. EFS was calculated from the date of treatment initiation until the date of death from any cause, disease relapse or progression, unplanned retreatment of lymphoma after initial management, or the date of last contact. EFS24 was defined as EFS status 24 months from diagnosis, as previously described.^{23,24} OS from failing EFS24 (ie, early progression) was defined as time from EFS24 failure to death or last follow-up. OS from achieving EFS24 was defined as time from achieving EFS24 to death or last follow-up.

Variables considered for model building included age, sex, PS, B symptoms, stage, number of nodal and extranodal sites involved, LDH, hemoglobin, longest diameter of the largest lymph node, presence of effusion and compression syndrome, circulating lymphoma cells, platelet count, serum albumin, bone marrow involvement, and β_2m . Serum albumin and Eastern Cooperative Oncology Group (ECOG) score could not be eligible because of, respectively, a high number of missing values ($n = 223$) and a low

number of patients with a score above 1 ($n = 49$). For model development, categorical variables were used (bone marrow involvement, sex, B symptoms, effusion syndrome, or the presence of circulating malignant cells). All continuous variables were dichotomized according to thresholds determined using Cox regression cubic splines analysis for PFS (supplemental Figure 1, available on the *Blood* Web site).²⁵ For score system building, a subsampling analysis was performed on 100 subsets of a randomly selected subpopulation consisting of 90% of the whole cohort of patients. For each sample, stepwise multivariate Cox model was used with a P level of entry set up at 0.1 and a level of stay at 0.05. To determine a simplified and parsimonious prognostic model, the most significant and frequently selected covariates in the subsampling analysis were used in conditional inference trees²⁶ to identify 3 patient-groups experiencing differential outcome in terms of PFS. The minimum P value for which a split was implemented was set up at $P = .001$ (no pruning was performed), and the maximum depth of trees was fixed at 2.²⁷

For comparison with the FLIPI, each model's discrimination was computed using the log-rank χ^2 value, the Concordance Probability Estimates (CPE) by Harrell's c-index,²⁸ and the Net Reclassification Improvement (NRI; see supplemental Statistical Methods for further details).²⁹

The PRIMA-PI was externally validated on the independent combined cohort of patients from the FL2000 trial and the MER cohort. As patient characteristics (age, sex, stage, and β_2m) were comparable, the 2 cohorts were pooled to provide an independent data set with meaningful size. Models' performances were compared as previously described for the PRIMA cohort.

All statistical tests were 2-sided. A P value $<.05$ was considered statistically significant. Statistical analyses were performed using SAS version 9.2 for PC (SAS Institute, Cary, NC) and R version 3.2.3.³⁰

Results

Patient characteristics

Patient characteristics of the PRIMA data set are summarized in Table 1 and in supplemental Table 1 for the FL2000^{8,21} and the MER cohorts. Briefly, the final population of the PRIMA trial with a confirmed FL grade 1-3a histology and who received an induction treatment was 1135 patients. Thirty-five percent of patients were older than 60 years, 52% were men, and most of them had disseminated stage III-IV disease (90%). β_2m , which was locally measured in each center, was above 3 mg/L in 30% of patients. Most patients received an induction regimen containing R-CHOP (74%), and only 22% and 4% received R-CVP and R-FCM, respectively. As per protocol design, 50% of the 974 randomized patients (ie, who were in response after induction) received rituximab maintenance. In the pooled external validation cohort, 47% of patients were older than 60 years, 54% were men, 81% had stage III/IV disease, 34% had a β_2m higher than 3 mg/L and 50% had bone marrow involvement. Treatment was R-CHVP followed by interferon maintenance in 37% of the cohort (corresponding to the population of the R-containing group of treatment of the FL2000 trial), R-CHOP in 25%, and R-bendamustine in 22%. Twenty percent of patients in the pooled validation cohort received rituximab maintenance.

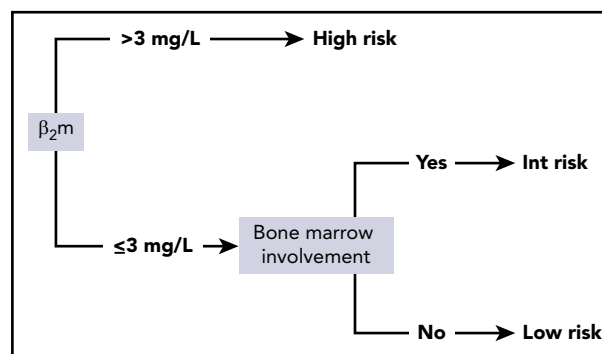


Figure 1. PRIMA-PI score calculation.

Simplified prognostic score development in the PRIMA cohort

The best cut-points were found identical to usual clinical thresholds for most parameters: 12 g/dL for hemoglobin level, $150 \times 10^9/L$ for platelets, 40 g/L for albumin, 60 mm for diameter of the largest tumor size, upper normal limit for LDH, more than 1 extranodal site, or a PS above 1. In this high-burden population requiring immediate treatment with immunochemotherapy, best cutoff for β_2m was 3 mg/L (and not the upper normal limit, as for the FLIPI2), and a limit of 6 involved nodal areas was more discriminative compared with the 4 nodal areas threshold in the FLIPI score (supplemental Figure 1). Because computation of the number of nodal areas was already cumbersome with a cut-point at 4, the cutoff was yet left unchanged. The most frequently selected variables in the subsampling analysis for prognostic model determination were β_2m , hemoglobin level, sex, LDH level, and bone marrow involvement. Effusion syndrome and number of nodal areas were selected in approximately half of models, whereas age, B symptoms, Ann Arbor stage, and extranodal involvement were retained in less than half of final models. Other parameters were not maintained in any of the models in subsampling analysis.

Given those results, the 5 variables (sex, LDH, bone marrow involvement, hemoglobin, and β_2m) were considered for conditional inference trees. A 2-variable model score using only β_2m and bone marrow involvement was found to be optimal. The final simplified score herein referred to as the PRIMA-PI comprised 3 risk categories: high ($\beta_2m > 3$ mg/L), low ($\beta_2m \leq 3$ mg/L without bone marrow involvement), and intermediate ($\beta_2m \leq 3$ mg/L with bone marrow involvement; Figure 1).

Scoring systems in the PRIMA cohort

PFS curves according to the PRIMA-PI, the FLIPI, the FLIPI2, and the β_2m +LDH score are presented in Figure 2A-D. All scores were able to discriminate subgroups of patients with significantly different prognosis ($P < .0001$ for all). PFS according to the 4 different scoring systems in the observation and rituximab maintenance groups of treatment of the PRIMA study are presented in supplemental Figure 2A-H.

Risk categories were evenly distributed by the PRIMA-PI (34%, 34%, and 32% for the low-, intermediate-, and high-risk groups, respectively). Five-year PFS was 69% (95% CI, 64%-73%), 55% (95% CI, 49%-60%), and 37% (95% CI, 32%-42%) for the PRIMA-PI in the low-, intermediate-, and high-risk groups, respectively (Table 2). As expected, because it was built on the

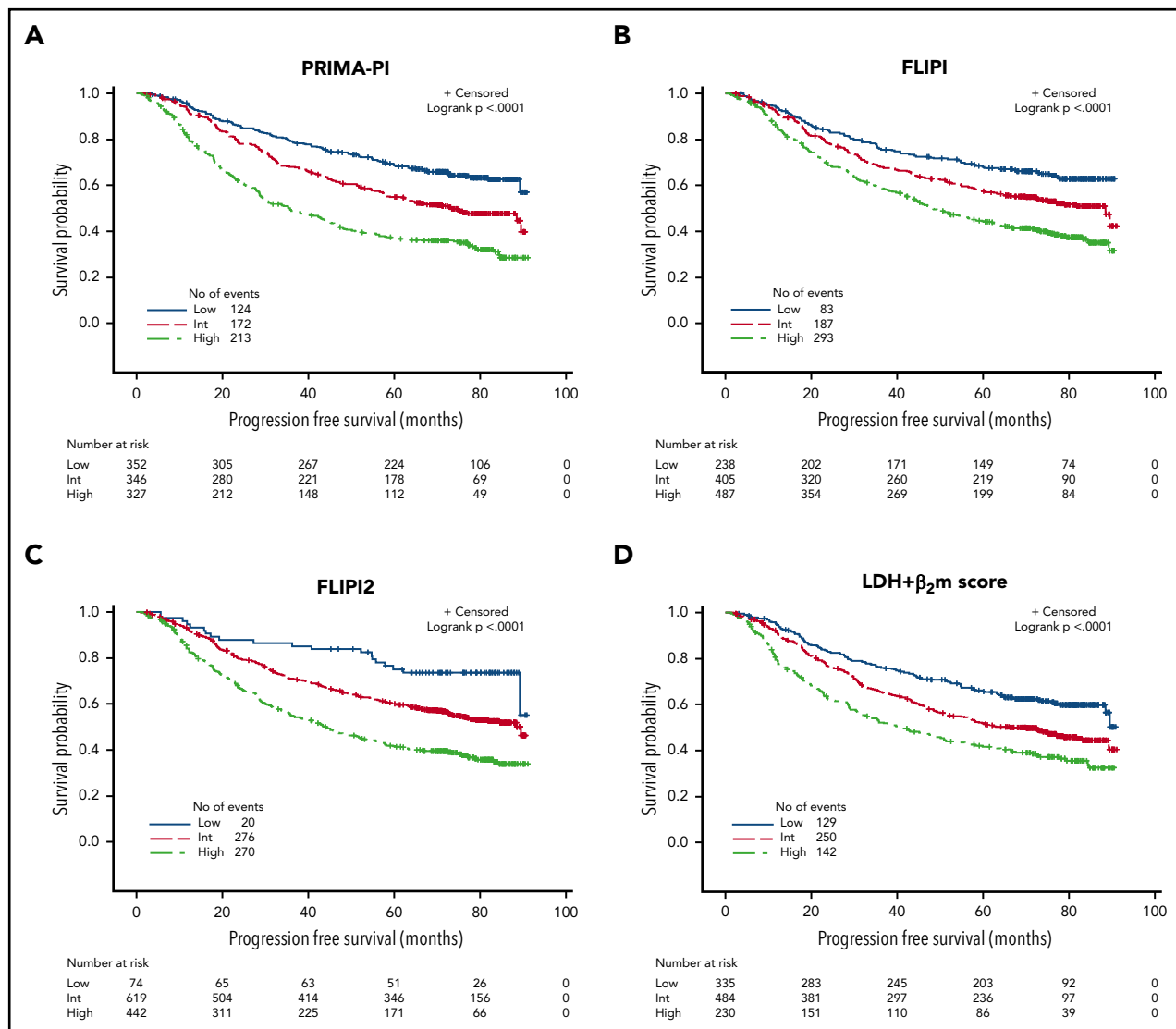


Figure 2. PFS in the training cohort. (A) PRIMA-PI. (B) FLIPI. (C) FLIPI2. (D) LDH + β_2m score.

PRIMA cohort, the PRIMA-PI displayed high model performances (Table 3). When comparing the FLIPI, the FLIPI2, and the β_2m +LDH score, all presented with robust performances, as assessed by log-rank χ^2 , CPE, and positive NRI (Table 3). The FLIPI2 identified a low number of patients ($n = 74$) falling into the low-risk category (ie, with zero risk factor) with an excellent prognosis (Figure 2C).

EFS24 in the PRIMA cohort

Approximately 25% of patients experienced early disease progression within 24 months of therapy initiation. We confirmed that EFS24 was highly correlated with subsequent OS duration. For patients without disease progression before 24 months, 5-year OS from the 24-month landmark was 92% (95% CI, 80%-94%) compared with 63% (95% CI, 57%-70%) from the risk-defining event for patients who progressed in the first 24 months ($P < .0001$; Figure 3).

The FLIPI was shown to be correlated with the risk for progression before 24 months after treatment initiation.²³ According to the FLIPI, proportions of patients experiencing EFS24 (ie, early treatment failure) were indeed 16%, 21%, and 31% in the low-, intermediate-, and high-risk categories, respectively ($P = 1.36 \times 10^{-5}$).

According to the PRIMA-PI, proportions of patients experiencing EFS24 were 14%, 21%, and 38% in the low-, intermediate-, and high-risk categories, respectively ($P = 1.41 \times 10^{-12}$; supplemental Table 2). Both the FLIPI and the PRIMA-PI were therefore strongly associated with early progression.

OS in the PRIMA cohort

Because the primary endpoint for building the FLIPI was OS compared with PFS for the PRIMA-PI, we next assessed whether the PRIMA-PI could segregate patient outcome in terms of OS. Because of the very small number of deaths in the low- and intermediate-risk categories, both for the PRIMA-PI and the FLIPI, the 2 categories were merged. OS in the PRIMA cohort was 93% at 5 years in the low-/intermediate-risk category vs 84% in the high-risk category for the FLIPI and 93% vs 81% for the PRIMA-PI ($P < .0001$ for both; supplemental Figure 3A-B). Because age is part of the FLIPI but not of the PRIMA-PI, we also considered lymphoma-specific survival. Both the FLIPI and the PRIMA-PI were significantly associated with lymphoma-specific survival ($P < .0001$ for both; supplemental Figure 3C-D).

External validation in the pooled FL2000 and MER cohorts

Because the PRIMA-PI was built on the PRIMA cohort and was therefore associated with overfitting, the model was tested in a pooled external and independent validation cohort (N = 479) (Table 1; supplemental Table 1). In this validation cohort, the PRIMA-PI remained at least as discriminatory for EFS as the FLIPI (Table 3; Figure 4A and B). Five-year EFS in the validation cohort was 77% (95% CI, 69%-83%), 57% (95% CI, 48%-66%), and 44% (95% CI, 35%-52%) in the PRIMA-PI low-, intermediate-, and high-risk groups, respectively ($P < .0001$).

Because the precise longest diameter of the largest involved lymph node was not available in the MER cohort, the FLIPI2 could not be computed and the PRIMA-PI could therefore not be compared with the FLIPI2 in this validation cohort. Regarding the β_{2m} +LDH scoring system developed by Press et al.,¹⁷ the PRIMA-PI demonstrated better performance (Table 3; Figure 4C). With regard to EFS24, the PRIMA-PI also showed good discrimination power in this validation data set (supplemental Table 3).

Discussion

In FL, several prognostic indexes have been proposed during the last 30 years. Developed in the early 1990s, the international prognostic index (IPI) for aggressive lymphomas has shown modest performance in discriminating patients' prognosis in FL.³¹ Most patients with FL fall into the low- or the intermediate-risk categories of the IPI, precluding balanced risk group comparisons.³² More recently, the Italian Lymphoma Intergroup Index¹⁴ in 2000, followed by the Follicular Lymphoma International Prognostic Index¹⁵ in 2004, have refined prognostic assessment in FL based on 6 and 5 bioclinical parameters, respectively. Finally, the FLIPI2 was proposed in 2009, based on a large and, for the first time, prospective data collection by Federico et al.¹⁶ β_{2m} , which was already considered a significant prognostic parameter in FL in previously published series,³³⁻³⁵ could be incorporated into the model and further improved patients' segregation, whereas it could not be considered for the FLIPI because of a large number of missing data. In an attempt to find a surrogate for the number of nodal sites involved, which is the most laborious and non-reproducible parameter to assess in routine practice in the FLIPI, it was shown that the tumor burden could be approximated by the longest diameter of the largest involved node. However, over the course of almost 10 years, the FLIPI has still been the most widely used prognostic index in FL, mainly because of its robustness in discriminating fairly balanced risk groups in populations treated with highly diverse regimens. More recently, efforts have been made to develop simplified scoring systems like the LDH+ β_{2m} score, developed by Press et al.,¹⁷ defining 3 risk categories according to these 2 simple parameters (LDH and/or β_{2m} above the upper limit of the normal).

To date, no score has been specifically developed on the basis of a cohort of patients solely treated with rituximab-containing immunochemotherapy. In the FLIPI2, for instance, only 59% of patients received a systemic therapy containing rituximab. Here, we took advantage of the large randomized phase 3 PRIMA trial with prospectively collected bioclinical data to validate previously developed scoring systems and to define a new prognostic tool in patients treated with initial immunochemotherapy.

Table 2. Five-year progression- and event-free survival according to prognostic scoring systems in the training and validation cohorts

	Training cohort (PRIMA)						Validation cohort (FL2000+MER)							
	FLIPI		PRIMA-PI		FLIPI2		β_{2m} +LDH		FLIPI		PRIMA-PI		β_{2m} +LDH	
	N (%)	5-y PFS % (95% CI)	N (%)	5-y PFS % (95% CI)	N (%)	5-y PFS % (95% CI)	N (%)	5-y PFS % (95% CI)	N (%)	5-y EFS % (95% CI)	N (%)	5-y EFS % (95% CI)	N (%)	5-y EFS % (95% CI)
Low	238 (21)	68 (62-74)	352 (34)	69 (64-73)	74 (6)	75 (63-83)	335 (32)	65 (60-70)	126 (28)	76 (66-83)	168 (36)	77 (69-83)	230 (52)	69 (62-75)
Intermediate	405 (36)	58 (53-62)	346 (34)	55 (49-60)	619 (55)	60 (56-63)	484 (46)	51 (47-56)	159 (35)	64 (55-72)	136 (29)	57 (48-66)	144 (32)	51 (42-60)
High	487 (43)	44 (38-48)	327 (32)	37 (32-42)	442 (39)	41 (36-46)	230 (22)	41 (34-47)	165 (37)	45 (37-53)	158 (34)	44 (35-52)	71 (16)	44 (31-56)

Table 3. Model performance for PFS in the training cohort and for EFS in the validation cohort

	Training cohort (PRIMA)				Validation cohort (FL2000+MER)*		
	FLIPI	PRIMA-PI	FLIPI2	β_2m +LDH	FLIPI	PRIMA-PI	β_2m +LDH
Log-rank χ^2	45.64	81.96	54.12	41.23	27.17	29.47	19.39
CPE (\pm SE)	0.577 (\pm 0.011)	0.604 (\pm 0.011)	0.577 (\pm 0.010)	0.572 (\pm 0.011)	0.605 (\pm 0.020)	0.606 (\pm 0.019)	0.577 (\pm 0.017)
NRI (\pm SE)	ref	+36.7% (\pm 6.1%)	+24.7% (\pm 5.8%)	-1.3% (\pm 5.9%)	ref	+3.5% (\pm 8.7%)	-7.1% (\pm 9.8%)

*The FLIPI2 could not be calculated in the validation cohort because of the unavailable diameter of the longest diameter of the largest involved lymph node with a cutoff set up at 6 cm in the MER cohort.

ref, reference value; SE, standard error.

An important finding of the present study is the confirmation of the prognostic value of the FLIPI, the FLIPI2, and the aforementioned LDH+ β_2m scoring system in the PRIMA cohort of patients. When developing a new scoring system, rather than the sum of different individual prognostic factors, the PRIMA-PI was devised using conditional inference trees. The technique allows for a maximal separation of populations with different prognoses, using the minimum number of significant parameters, and has been already fruitfully applied in multiple myeloma for building the International Staging System.³⁶ The main advantage of the PRIMA-PI is that it is easy to calculate in routine clinical practice and does not require the cumbersome and error-prone computation of the lymph node areas. Furthermore, based on model performance comparison in the validation data set, it appears to be as discriminant as the FLIPI and more performant than the LDH+ β_2m score for this specific population of patients treated upfront with immunochemotherapy. One caveat of the study is that it could not be compared with the FLIPI2 in the validation data set because of the unavailability of the longest diameter of the largest involved lymph node with a cutoff set up at 6 cm. Therefore, further validation will be warranted to confirm that the PRIMA-PI can be considered as discriminant as the FLIPI and the FLIPI2 for patients treated with immunochemotherapy.

Because only 22% of the patients in the validation cohort were treated with R-bendamustine, the PRIMA-PI also warrants further validation in other trial and real-world cohorts using this agent. In addition, recent data providing evidence of a PFS advantage associated with the use of obinutuzumab in first-line therapy in FL underlines the need to validate these scoring systems when this new anti-CD20 is used in combination with chemotherapy.¹³ Finally, novel chemotherapy-free combinations entering the field will probably challenge conventional clinico-biological parameters, and score performances will need to be further confirmed.

Altogether, most of these different scoring systems demonstrated the key prognostic value of β_2m that should be part of the pretreatment workup in FL. We found that a 3 mg/L cutoff value was able to identify a subpopulation with a particularly pejorative outcome, even when testing was performed in different laboratories using different methods.

Bone marrow involvement might not be the most convenient parameter to evaluate in the disease, especially in routine practice. With the broader use of positron emission tomography/computed tomography (PET-CT) in the clinical staging and response assessment of FL,^{37,38} the role and the utility of the bone

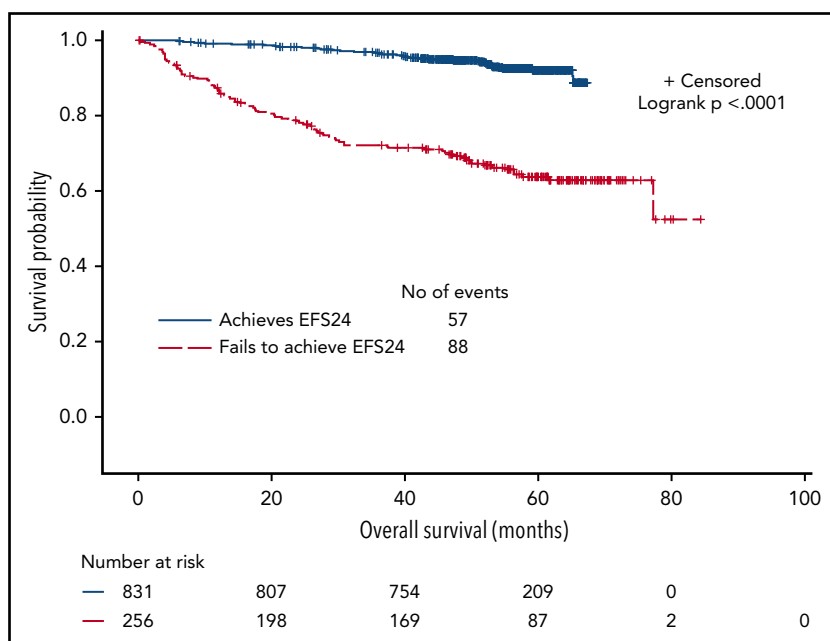


Figure 3. OS from achieving EFS24 or from EFS24 event in the training cohort.

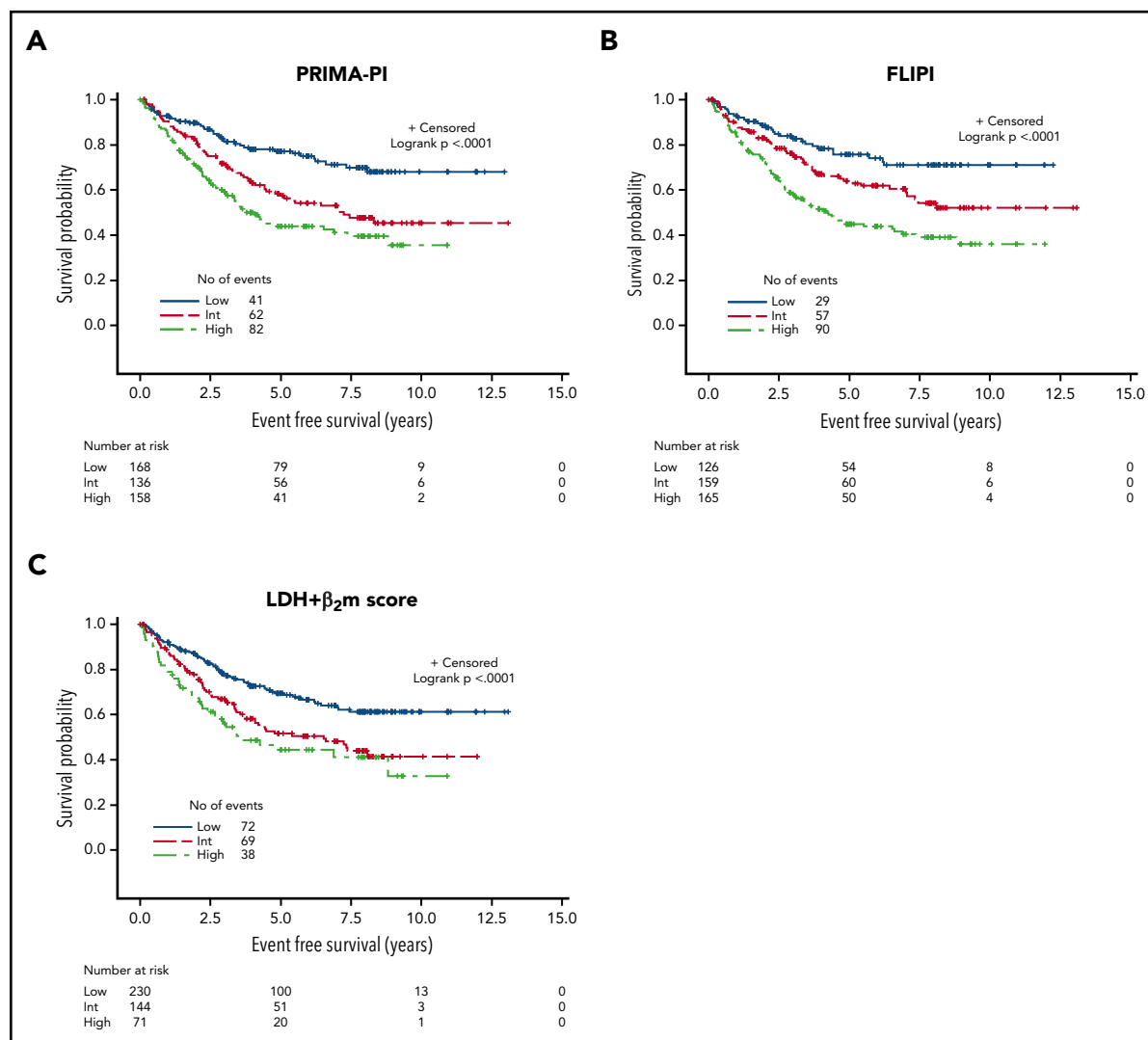


Figure 4. EFS in the external validation cohort. (A) PRIMA-PI. (B) FLIPI. (C) LDH + β_2m score.

marrow biopsy tends to be challenged. However, the parameter could not be substituted for any other variable while keeping with the objective of being the most parsimonious, despite extensive statistical analyses (not shown). In case of a high tumor burden disease or in case of a disease requiring immediate immunochemotherapy (where the PRIMA-PI has been built and validated), most clinicians still perform a bone marrow biopsy. Recent reports suggest that the PET scanner could be useful for evaluating bone marrow infiltration,³⁹ and whether PET-CT results could be used for PRIMA-PI calculation warrants further investigation.

Recently, postinduction variables have been validated as strong predictive parameters for subsequent outcome. Best response during the first 2 years,⁴⁰ complete response at 30 months,⁴¹ and achievement of EFS24 or PFS24 for patients treated with immunochemotherapy,^{23,24} have demonstrated robust prediction of subsequent OS. In the present study, the PRIMA-PI was correlated to EFS24, highlighting its capacity to identify patients with an adverse prognosis with only 2 parameters.

New advances in biology have demonstrated the role of mutational status of selected genes for outcome prediction.¹⁸ Mutations

in *EP300*, *FOXO1*, *CREBBP*, and *CARD11* have been associated with an inferior outcome as opposed to mutations in *MEF2B*, *ARID1A*, and *EZH2*, found in patients with a more prolonged failure-free survival.¹⁸ Along with the FLIPI and the ECOG status, mutation patterns have thus been incorporated into the m7-FLIPI for more accurate outcome prediction. As molecular tools as well as posttreatment parameters will be progressively incorporated, parsimony in a clinical model backbone is desirable. Because of its simplicity, bio-molecular scores based on the PRIMA-PI might then be easier to compute than scores based on the FLIPI or the FLIPI2.

In conclusion, we confirmed that all previously described scoring systems in FL are still robust predictors of outcome in patients treated with initial immunochemotherapy with or without maintenance therapy, but that a simplified scoring system, the PRIMA-PI, might have at least as good discriminatory capacity with only 2 parameters for convenient use in routine clinical practice.

Acknowledgments

E.B. has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie

(grant agreement 661106); the Philippe Foundation; and the Fulbright Program (Monahan Foundation). The MER was supported by the National Institutes of Health, National Cancer Institute grants P50 CA97274 and U01 CA195568.

Authorship

Contribution: E.B. and G.S. treated patients, designed and performed research, analyzed data, and wrote the paper; M.J.M., B.G.-D., D.M.-B., and J.R.C. designed and performed research, analyzed data, and wrote the paper; T.M.H., J.A.E., E.V.d.N., R.B., E.G., J.B., A.D., M.G.d.S., O.F., D.B., H.M., T.I., S.M.A., T.L., B.K.L., and J.F.S. treated patients and wrote the manuscript; and A.L.F., S.L.S., and P.D. performed research and wrote the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profile: E.B., 0000-0003-2694-7510.

Correspondence: Gilles Salles, Hematology Department, Centre Hospitalier Lyon Sud, 165 Chemin du Grand Revoyet, 69495 Pierre Bénite Cedex, France; e-mail: gilles.salles@chu-lyon.fr.

Footnotes

Submitted 14 November 2017; accepted 8 April 2018. Prepublished online as *Blood* First Edition paper, 17 April 2018; DOI 10.1182/blood-2017-11-816405.

The online version of this article contains a data supplement.

There is a *Blood* Commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

REFERENCES

1. Teras LR, DeSantis CE, Cerhan JR, Morton LM, Jemal A, Flowers CR. 2016 US lymphoid malignancy statistics by World Health Organization subtypes [published online ahead of print 12 September 2016]. *CA Cancer J Clin*. doi:10.3322/caac.21357.
2. Link BK, Maurer MJ, Nowakowski GS, et al. Rates and outcomes of follicular lymphoma transformation in the immunochemotherapy era: a report from the University of Iowa/MayoClinic Specialized Program of Research Excellence Molecular Epidemiology Resource. *J Clin Oncol*. 2013;31(26):3272-3278.
3. Brice P, Bastion Y, Lepage E, et al. Comparison in low-tumor-burden follicular lymphomas between an initial no-treatment policy, prednimustine, or interferon alfa: a randomized study from the Groupe d'Etude des Lymphomes Folliculaires. *J Clin Oncol*. 1997;15(3):1110-1117.
4. Herold M, Haas A, Srock S, et al; East German Study Group Hematology and Oncology Study. Rituximab added to first-line mitoxantrone, chlorambucil, and prednisolone chemotherapy followed by interferon maintenance prolongs survival in patients with advanced follicular lymphoma: an East German Study Group Hematology and Oncology Study. *J Clin Oncol*. 2007;25(15):1986-1992.
5. Hiddemann W, Kneba M, Dreyling M, et al. Frontline therapy with rituximab added to the combination of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) significantly improves the outcome for patients with advanced-stage follicular lymphoma compared with therapy with CHOP alone: results of a prospective randomized study of the German Low-Grade Lymphoma Study Group. *Blood*. 2005;106(12):3725-3732.
6. Marcus R, Imrie K, Belch A, et al. CVP chemotherapy plus rituximab compared with CVP as first-line treatment for advanced follicular lymphoma. *Blood*. 2005;105(4):1417-1423.
7. Marcus R, Imrie K, Solal-Celigny P, et al. Phase III study of R-CVP compared with cyclophosphamide, vincristine, and prednisone alone in patients with previously untreated advanced follicular lymphoma. *J Clin Oncol*. 2008;26(28):4579-4586.
8. Salles G, Mounier N, de Guibert S, et al. Rituximab combined with chemotherapy and interferon in follicular lymphoma patients: results of the GELA-GOELAMS FL2000 study. *Blood*. 2008;112(13):4824-4831.
9. Hochster H, Weller E, Gascoyne RD, et al. Maintenance rituximab after cyclophosphamide, vincristine, and prednisone prolongs progression-free survival in advanced indolent lymphoma: results of the randomized phase III ECOG1496 Study. *J Clin Oncol*. 2009;27(10):1607-1614.
10. Rummel MJ, Al-Batran SE, Kim SZ, et al. Bendamustine plus rituximab is effective and has a favorable toxicity profile in the treatment of mantle cell and low-grade non-Hodgkin's lymphoma. *J Clin Oncol*. 2005;23(15):3383-3389.
11. Rummel MJ, Niederle N, Maschmeyer G, et al; Study group indolent Lymphomas (StiL). Bendamustine plus rituximab versus CHOP plus rituximab as first-line treatment for patients with indolent and mantle-cell lymphomas: an open-label, multicentre, randomised, phase 3 non-inferiority trial. *Lancet*. 2013;381(9873):1203-1210.
12. Salles G, Ghesquières H, Bachy E. What is the best first-line treatment strategy for patients with indolent lymphomas? *Am Soc Clin Oncol Educ Book*. 2012:488-493.
13. Marcus R, Davies A, Ando K, et al. Obinutuzumab for the first-line treatment of follicular lymphoma. *N Engl J Med*. 2017;377(14):1331-1344.
14. Federico M, Vitolo U, Zinzani PL, et al. Prognosis of follicular lymphoma: a predictive model based on a retrospective analysis of 987 cases. Intergroupo Italiano Linfomi. *Blood*. 2000;95(3):783-789.
15. Solal-Celigny P, Roy P, Colombat P, et al. Follicular lymphoma international prognostic index. *Blood*. 2004;104(5):1258-1265.
16. Federico M, Bellei M, Marcheselli L, et al. Follicular lymphoma international prognostic index 2: a new prognostic index for follicular lymphoma developed by the international follicular lymphoma prognostic factor project. *J Clin Oncol*. 2009;27(27):4555-4562.
17. Press OW, Unger JM, Rimsza LM, et al. A comparative analysis of prognostic factor models for follicular lymphoma based on a phase III trial of CHOP-rituximab versus CHOP + 131iodine-tositumomab. *Clin Cancer Res*. 2013;19(23):6624-6632.
18. Pastore A, Jurinovic V, Kridel R, et al. 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(9):1111-1122.
19. Mercadal S, Pomares H, Sancho JM, et al. Clinico-biological features, treatment and survival of 457 patients with histological Grades 3A and 1-2 follicular lymphoma mostly treated with immunochemotherapy. *Br J Haematol*. 2016;172(3):470-473.
20. Arcaini L, Merli M, Passamonti F, et al. Validation of follicular lymphoma international prognostic index 2 (FLIPI2) score in an independent series of follicular lymphoma patients. *Br J Haematol*. 2010;149(3):455-457.
21. Salles G, Seymour JF, Offner F, et al. Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (PRIMA): a phase 3, randomised controlled trial. *Lancet*. 2011;377(9759):42-51.
22. Cerhan JR, Link BK, Habermann TM, et al. Cohort Profile: The Lymphoma Specialized Program of Research Excellence (SPORE) Molecular Epidemiology Resource (MER) Cohort Study. *Int J Epidemiol*. 2017;46(6):1753-1754i.
23. Casulo C, Byrtek M, Dawson KL, et al. early relapse of follicular lymphoma after rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone defines patients at high risk for death: an analysis from the National LymphoCare Study. *J Clin Oncol*. 2015;33(23):2516-2522.
24. Maurer MJ, Bachy E, Ghesquières H, et al. Early event status informs subsequent outcome in newly diagnosed follicular lymphoma. *Am J Hematol*. 2016;91(11):1096-1101.
25. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model.

- Statistics for Biology and Health. New York, NY: Springer; 2000.
26. Segal MR. Extending the elements of tree-structured regression. *Stat Methods Med Res.* 1995;4(3):219-236.
 27. Hothorn T, Zeileis A. partykit: a modular toolkit for recursive partytioning in R. *J Mach Learn Res.* 2015;16:3905-3909.
 28. Lin Y, Chappell R, Gönen M. A systematic selection method for the development of cancer staging systems. *Stat Methods Med Res.* 2016;25(4):1438-1451.
 29. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med.* 2011;30(1): 11-21.
 30. R Core Team. (2015). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>.
 31. International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. *N Engl J Med.* 1993;329(14):987-994.
 32. López-Guillermo A, Montserrat E, Bosch F, Terol MJ, Campo E, Rozman C. Applicability of the International Index for aggressive lymphomas to patients with low-grade lymphoma. *J Clin Oncol.* 1994;12(7):1343-1348.
 33. Federico M, Guglielmi C, Luminari S, et al. Prognostic relevance of serum beta2 microglobulin in patients with follicular lymphoma treated with anthracycline-containing regimens. A GISL study. *Haematologica.* 2007; 92(11):1482-1488.
 34. Montoto S, López-Guillermo A, Ferrer A, et al. Survival after progression in patients with follicular lymphoma: analysis of prognostic factors. *Ann Oncol.* 2002;13(4): 523-530.
 35. Litam P, Swan F, Cabanillas F, et al. Prognostic value of serum beta-2 microglobulin in low-grade lymphoma. *Ann Intern Med.* 1991; 114(10):855-860.
 36. Greipp PR, San Miguel J, Durie BG, et al. International staging system for multiple myeloma. *J Clin Oncol.* 2005;23(15):3412-3420.
 37. Meignan M, Cottreau AS, Versari A, et al. Baseline metabolic tumor volume predicts outcome in high-tumor-burden follicular lymphoma: a pooled analysis of three multicenter studies. *J Clin Oncol.* 2016; JCO669440.
 38. Trotman J, Fournier M, Lamy T, et al. Positron emission tomography-computed tomography (PET-CT) after induction therapy is highly predictive of patient outcome in follicular lymphoma: analysis of PET-CT in a subset of PRIMA trial participants. *J Clin Oncol.* 2011; 29(23):3194-3200.
 39. Perry C, Lerman H, Joffe E, et al. The value of PET/CT in detecting bone marrow involvement in patients with follicular lymphoma. *Medicine (Baltimore).* 2016;95(9): e2910.
 40. Bachy E, Brice P, Delarue R, et al. Long-term follow-up of patients with newly diagnosed follicular lymphoma in the prirituximab era: effect of response quality on survival—A study from the groupe d'étude des lymphomes de l'adulte. *J Clin Oncol.* 2010; 28(5):822-829.
 41. Shi Q, Flowers CR, Hiddemann W, et al. Thirty-month complete response as a surrogate end point in first-line follicular lymphoma therapy: an individual patient-level analysis of multiple randomized trials. *J Clin Oncol.* 2017;35(5): 552-560.