blood®

Regular Article

LYMPHOID NEOPLASIA

International prognostic score for asymptomatic early-stage chronic lymphocytic leukemia

Adalgisa Condoluci,[1,2,*] Lodovico Terzi di Bergamo,[1,*] Petra Langerbeins,[3,*] Manuela A. Hoechstetter,[4,*] Carmen D. Herling,[3] Lorenzo De Paoli,[5] Julio Delgado,[6] Kari G. Rabe,[7] Massimo Gentile,[8] Michael Doubek,[9,10] Francesca R. Mauro,[11] Giorgia Chiodin,[12] Mattias Mattsson,[13] Jasmin Bahlo,[3] Giovanna Cutrona,[14] Jana Kotaskova,[9,10] Clara Deambrogi,[5] Karin E. Smedby,[15] Valeria Spina,[1] Alessio Bruscaggin,[1] Wei Wu,[1,16] Riccardo Moia,[5] Elena Bianchi,[2] Bernhard Gerber,[2] Emanuele Zucca,[2] Silke Gillessen,[2] Michele Ghielmini,[2] Franco Cavalli,[2] Georg Stussi,[2] Mark A. Hess,[17] Tycho S. Baumann,[6] Antonino Neri,[18] Manlio Ferrarini,[19] Richard Rosenquist,[15] Francesco Forconi,[12,20] Robin Foà,[11] Sarka Pospisilova,[9,10] Fortunato Morabito,[21,22] Stephan Stilgenbauer,[23] Hartmut Döhner,[23] Sameer A. Parikh,[7,†] William G. Wierda,[17,†] Emili Montserrat,[6,24,†] Gianluca Gaidano,[5,†] Michael Hallek,[3,†] and Davide Rossi[1,2,†]

[1]Institute of Oncology Research, Università della Svizzera Italiana, Bellinzona, Switzerland; [2]Oncology Institute of Southern Switzerland, Bellinzona, Switzerland; [3]Center of Integrated Oncology Cologne Bonn and German CLL Study Group, University of Cologne, Cologne, Germany; [4]Hospital Munich-Schwabing, German CLL Study Group, Munich, Germany; [5]Division of Hematology, Department of Translational Medicine, University of Eastern Piedmont, Novara, Italy; [6]Hospital Clinic, Barcelona, Spain; [7]Mayo Clinic, Rochester, MN; [8]Azienda Ospedaliera di Cosenza, Cosenza, Italy; [9]Interní Hematologická a Onkologická Klinika (IHOK), University Hospital Brno and Medical Faculty, Masaryk University, Brno, Czech Republic; [10]Central European Institute of Technology, Masaryk University, Brno, Czech Republic; [11]Division of Hematology, Sapienza University, Rome, Italy; [12]Cancer Sciences Division, University of Southampton, Southampton, United Kingdom; [13]Uppsala University Hospital, Uppsala, Sweden; [14]IRCCS Ospedale Policlinico San Martino, Genoa, Italy; [15]Karolinska Institute, Stockholm, Sweden; [16]The First Affiliated Hospital of Nanjing Medical University, Jiangsu Province Hospital, Nanjing, China; [17]University of Texas MD Anderson Cancer Center, Houston, TX; [18]Department of Oncology and Hemato-Oncology, University of Milan, Milan, Italy; [19]Department of Experimental Medicine, University of Genoa, Genoa, Italy; [20]Southampton University Hospital Trust, Southampton, United Kingdom; [21]Biotechnology Research Unit Aprigliano, Cosenza, Italy; [22]Augusta Victoria Hospital, Jerusalem, Israel; [23]Department of Internal Medicine III, University Hospital of Ulm, Ulm, Germany; and [24]Institute of Hematology and Oncology, University of Barcelona, Barcelona, Spain

KEY POINTS

- IPS-E is a simple and robust prognostic model for early-stage CLL.

- IPS-E can be helpful in patients' counseling and design of clinical trials.

Most patients with chronic lymphocytic leukemia (CLL) are diagnosed with early-stage disease and managed with active surveillance. The individual course of patients with early-stage CLL is heterogeneous, and their probability of needing treatment is hardly anticipated at diagnosis. We aimed at developing an international prognostic score to predict time to first treatment (TTFT) in patients with CLL with early, asymptomatic disease (International Prognostic Score for Early-stage CLL [IPS-E]). Individual patient data from 11 international cohorts of patients with early-stage CLL (n = 4933) were analyzed to build and validate the prognostic score. Three covariates were consistently and independently correlated with TTFT: unmutated immunoglobulin heavy variable gene (IGHV), absolute lymphocyte count higher than 15 × 10^9/L, and presence of palpable lymph nodes. The IPS-E was the sum of the covariates (1 point each), and separated low-risk (score 0), intermediate-risk (score 1), and high-risk (score 2-3) patients showing a distinct TTFT. The score accuracy was validated in 9 cohorts staged by the Binet system and 1 cohort staged by the Rai system. The C-index was 0.74 in the training series and 0.70 in the aggregate of validation series. By meta-analysis of the training and validation cohorts, the 5-year cumulative risk for treatment start was 8.4%, 28.4%, and 61.2% among low-risk, intermediate-risk, and high-risk patients, respectively. The IPS-E is a simple and robust prognostic model that predicts the likelihood of treatment requirement in patients with early-stage CLL. The IPS-E can be useful in clinical management and in the design of early intervention clinical trials. (*Blood*. 2020; 135(21):1859-1869)

Introduction

Chronic lymphocytic leukemia (CLL) is characterized by the relentless accumulation of monoclonal B lymphocytes with a distinct immunophenotype (ie, CD5, CD19, CD20, CD23) in peripheral blood, bone marrow, and lymphoid organs. CLL is the most frequent form of leukemia in Western countries, where 0.6% of the population is diagnosed with this B-cell tumor during their lifetime.[1] In most instances, CLL is diagnosed in general practice, and approximately 70% of patients present in an early phase of the disease, with no anemia, no thrombocytopenia, and no significantly enlarged lymph nodes or spleen.[2]

According to international guidelines, patients with asymptomatic early-stage CLL should not be treated until disease

progression occurs, and active surveillance remains the standard for management.[3-5] Indeed, different clinical trials enrolling patients with early-stage CLL failed to show a survival benefit of early intervention strategies based on chemo-chemoimmunotherapy.[6-10] However, patients with early-stage CLL have a variable clinical course. Some require treatment soon after diagnosis because of development of cytopenia or bulky lymphadenopathy, whereas others show a stable or a slowly progressive disease not requiring treatment of decades.[11] Because of this, the management of patients with early-stage CLL is challenging and shaped by uncertainty. Of note, it is estimated that more than 400 000 individuals in Europe and the United States diagnosed with CLL are followed by active surveillance.[1,12] Unfortunately, there is no a specific, simple, accurate, and widely accepted prognostic model to predict the likelihood of disease progression, and hence need for therapy, in patients with asymptomatic early-stage CLL.

The management of early, asymptomatic CLL may change in the future if a survival benefit is proven by early intervention with novel agents in patients who are at risk for impending progression to symptomatic disease requiring treatment.[13] In this new scenario, upfront identification of high-risk patients is warranted. The objective of this study was to construct a simple, robust, and validated predictor of disease progression and the need for intervention in patients with asymptomatic early-stage CLL.

Methods

Study design

This is a multicenter, international, retrospective, observational study in which already-existing and coded health-related data were further used (ClinicalTrials.gov identifier: NCT03436524). We obtained the individual patient data sets from 11 cohorts, including 8 series that were previously used to generate or validate scoring systems in CLL (Tables 1 and 2).[10,14-19] Inclusion criteria were: presentation after 1996 confirmed by flow cytometry[3,4]; Binet stage A at diagnosis, as defined by blood cell count and physical examination[3,4]; and active surveillance as initial management, defined by no treatment requirement within the first 3 months after diagnosis. The Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK) criteria were followed throughout the study.[20] Patients provided informed consent in accordance with local institutional review board requirements and the Declaration of Helsinki. The ethics committee approved the study (BASEC 2018-00341). Early-stage CLL was defined according to the Binet system in 10 of 11 cohorts. Data of individual patients with early (Binet stage A) disease managed with active surveillance policy were collected from the CLL1 (ClinicalTrials.gov identifier: NCT00262782; N = 547)[9,18] and CLL7 (ClinicalTrials.gov identifier: NCT00275054; N = 339) trials of the German CLL Study Group, and the O-CLL1 GISL trial (ClinicalTrials.gov identifier: NCT00917540; N = 312).[15,19] The M.D. Anderson Cancer Center (N = 1225),[14] the University of Eastern Piedmont (UEP; N = 333),[17] the Barcelona (N = 355),[19] the Brno University Hospital (Brno; N = 269),[19] the Southampton (N = 226), and the Sapienza University (N = 223) cohorts are institutional consecutive series of newly presented and prospectively observed patients with early (Binet stage A) disease. Binet stage A patients' data of the SCAN cohort (N = 223) were collected from the SCALE Scandinavian population-based

Table 1. Characteristics of the training cohort

Variable	UEP (N = 333)		
	N	%	Missing
Age >65 years	208	62.4	—
Male sex	178	53.4	—
Palpable lymph nodes	68	20.4	—
Palpable spleen	20	6.0	—
Lymphocytes >15 × 10⁹/L	65	19.5	—
Hb <13 g/dL (male)/<12 g/dL (female)	39	11.7	—
Platelets <150 × 10⁹/L	46	13.8	—
B2M >3.5 mg/L	32	9.6	—
FISH normal	112	33.6	—
Del(13q)	168	50.4	—
Trisomy 12	59	17.7	—
Del(11q)	18	5.4	—
Del(17p)	19	5.7	—
Unmutated IGHV	92	27.6	—
TP53 mutation	24	7.2	—
ATM mutation	24	8.3	44
MYD88 mutation	18	5.4	4
NOTCH1 mutation	35	10.6	4
SF3B1 mutation	17	5.1	4
CLL-IPI			
Low risk	182	54.7	
Intermediate risk	94	28.2	
High risk	45	13.5	
Very high risk	12	3.6	

B2M, β-2-microglobulin; Del, deletion; Hb, hemoglobin.

case-control study.[16] The Mayo Clinic cohort (N = 881) is an institutional series of patients with early stage 0, I, and II disease defined according to the Rai system.

Data analysis

We consecutively performed both uni- and multivariable analyses using the full data set of the training cohort (UEP). Subsequently, we analyzed the external validation data sets to confirm the findings from the full-analysis data set. The study end point was time to first treatment (TTFT), defined as the time between presentation and start of first treatment of CLL because of progression to symptomatic disease according to the National Cancer Institute-Working Group/International Workshop on Chronic Lymphocytic Leukemia guidelines (patients without a documented event were censored at the date of last observation or death).[3,4] Nineteen baseline biomarkers, assessed within 1 month from initial presentation, were initially considered as

CONDOLUCI et al

Table 2. Characteristics of the validation cohorts

Variable	MDACC (N = 1225)			CLL1 (N = 547)			Barcelona (N = 355)			CLL7 (N = 339)			O-CLL1 (N = 312)			Brno (N = 269)			Southampton (N = 226)			SCAN (N = 223)			SU (N = 223)			Mayo Clinic (N = 881)		
	N	%	Miss.	N	%	Miss.	N	%	Miss.	N	%	Miss.	N	%	Miss.	N	%	Miss.	N	%	Miss.	N	%	Miss.	N	%	Miss.	N	%	Miss.
Age >65 years	374	30.5	—	154	28.1	—	166	46.7	—	98	28.9	—	80	25.6	—	113	42.0	—	120	53.0	—	95	42.6	—	28	12.7	3	380	43.1	—
Male sex	727	59.3	—	335	61.2	—	202	56.9	—	214	63.1	—	190	60.8	—	167	62.0	—	134	59.2	—	131	58.7	—	124	55.6	—	591	67.1	—
Palpable lymph nodes	536	43.7	—	122	22.3	—	81	22.8	—	115	33.9	—	62	19.8	—	119	44.2	—	93	41.1	—	94	42.1	—	114	51.1	—	424	48.1	—
Palpable spleen	68	5.5	—	29	5.5	24	6	1.6	—	16	5.1	28	21	6.7	—	2	0.7	—	4	1.7	—	7	3.2	5	29	13.0	—	62	7.0	—
Lymphocytes >15 × 10⁹/L	609	49.7	—	280	51.1	—	91	25.6	—	141	41.6		139	44.5	—	199	73.9	—	78	34.5	—	127	56.9	—	75	33.6	—	353	40.1	2
B2M >3.5 mg/L	116	9.7	41	44	8.0	—	39	11.3	11	7	2.1	16	3	1.3	82	31	11.9	10	31	20.6	76	16	7.3	5	8	3.6	1	80	9.7	53
Trisomy 12	196	16.0	—	49	9.1	12	55	16.5	23	28	8.3	3	26	9.2	30	31	11.5	1	20	10.9	44	12	6.9	51	20	10.4	31	151	17.8	33
Del(17p)	69	5.6	—	15	2.8	12	14	4.1	15	7	2.0	3	6	2.1	30	18	6.6	—	7	3.7	38	3	1.7	51	8	4.1	31	40	4.8	33
Unmutated IGHV	480	39.1	—	157	28.7	—	138	38.8	—	74	21.8	—	106	33.9	—	136	50.5	—	70	30.9	—	56	25.1	—	70	31.3	—	387	44.7	15
CLL-IPI			41			14			26			43			93			10			99			10			19			76
Low risk	566	47.8		330	61.9		174	52.9		201	67.9		136	62.1		98	37.8		57	44.9		134	62.9		129	63.2		349	43.4	
Intermediate risk	422	35.7		138	25.9		93	28.3		76	25.7		72	32.9		91	35.1		41	32.3		54	25.4		49	24.1		290	36.0	
High risk	160	13.5		58	10.9		46	14.0		18	6.1		10	4.6		45	17.4		27	21.2		22	10.3		19	9.3		142	17.6	
Very high risk	36	3.0		7	1.3		16	4.8		1	0.3		1	0.4		25	9.7		2	1.6		3	1.4		7	3.4		24	3.0	

Miss., missing.

covariates for construction of the prognostic index. These covariates were clinical characteristics (age, sex, ≥1 palpable lymph node with a diameter ≥1 cm and palpable spleen by investigator's physical examination),[21] laboratory values (hemoglobin level, platelet count, absolute lymphocyte count, and β-2-microglobulin), cytogenetic abnormalities as assessed by fluorescence in situ hybridization [FISH; del(17p), del(11q), trisomy 12, del(13q)], and gene mutations (immunoglobulin heavy variable gene [IGHV], ATM, MYD88, NOTCH1, SF3B1, and TP53). FISH was performed for assessing the t(11;14) in those cases in which a diagnosis of mantle cell lymphoma could be considered; patients harboring such translocations were excluded from the study. Continuous variables were categorized by published thresholds (age, hemoglobin, platelet count, β-2-microglobulin, IGHV identity)[18,22-25] or by identifying the best cutoff through recursive partitioning (absolute lymphocyte count; supplemental Figure 1, available on the Blood Web site). Blood cell count was assessed by local laboratories. β-2-microglobulin was assessed by local laboratories with the exception of the CLL1, CLL7, and SCAN cohorts for which the information was provided centrally. FISH cytogenetics were assessed locally, with the exception of the CLL1, CLL7, and O-CLL1 cohorts, for which the information was provided centrally and scored according to the laboratory cutoff. IGHV and TP53 mutations were assessed by Sanger sequencing locally, with the exception of CLL1, CLL7, and O-CLL1 cohorts, for which the information was provided centrally. Mutations of ATM, MYD88, NOTCH1, and SF3B1 were assessed in the training cohort by targeted deep next-generation sequencing, using an allele frequency of at least 10% for defining a positive result.[26] For the validation data sets, only variables composing the final model were included in the analyses. In the training and validation cohorts, minimal missing data were documented among covariates used for development and validation of the prognostic score. Therefore, no missing data imputation procedures were used. All statistical tests were 2-sided. Statistical significance was defined as P < .05. The analysis was performed with the Statistical Package for the Social Sciences software v.22.0 (Chicago, IL) and with R statistical package v3.4.1 (www.r-project.org).

Development and validation of a score for TTFT prognostication

Survival analysis was performed by the Kaplan-Meier method and comparison between strata, using the log-rank test.[27] Sensitivity analyses of TTFT were performed by considering death without treatment as a competing risk and by using Gray's test for comparisons between strata.[28,29] The adjusted association between exposure variables and TTFT was estimated by Cox regression.[30] In the training cohort, Cox regression included exposure variables showing an univariable association with TTFT with a Bonferroni corrected P < .05 to account for multiple testing. Backward elimination using likelihood ratio statistics with selection criterion P < .05 was used to derive the final Cox model to be validated. An elastic net penalized regression model was fit to the training set to exclude potential variable loss in the backward selection procedure.[31] The mixing parameter was selected using 10-fold cross-validation. The ideal model was selected as that corresponding to the largest value of the tuning parameter, such that the error was within 1 standard error of the minimum mean cross-validated error. The proportional hazard assumption was assessed by plotting the smoothed Schoenfeld

residuals against time.[32] The stability of the Cox model was externally validated across the 9 Binet A validation cohorts. Only variables independently and significantly associated with TTFT in more than 50% of the validation cohorts were used to construct the final score. We assigned a weighted-risk score to each factor based on the regression parameters from the Cox regression analysis. The score was defined as the sum of single-risk parameters.[18] We stratified patients in different risk groups by recursive partitioning of the score.[33] Two major steps were used to derive the best decision tree: growing an initial tree under the following constraints and stopping rules: i) split criteria of P < .01 according to the log-rank test, ii) more than 20 patients in a node to be considered for splitting, and iii) more than 10 patients in a terminal node; and applying a pruning algorithm based on the complexity parameter (cp > 0.01). Model discrimination was computed using the Harrel's C-index.[34] The 9 Binet A validation cohorts were meta-analyzed to interpolate the validation series in a single plot. TTFT and corresponding 95% confidence intervals (95% CIs) were estimated in each of the Binet A validation cohorts at 6-month intervals. At each point, a random-effect model was used to provide a meta-analytic estimate of TTFT and corresponding 95% CIs across the Binet A validation cohorts. Survival curves were plotted by fitting a cubic smoothing spline to the estimates of TTFT and the respective 95% CIs, using a smoothing parameter of 0.6. A fixed-effect model was used for the meta-analysis of the C-index across the Binet A cohorts, the 1- and 5-year cumulative risk for treatment need, and number of events per 100 person years, and bootstrapping was used to calculate 95% CI. Heterogeneity was assessed by using the I[2] statistic and Cochran's Q test.

Results

Individual patient data from 11 cohorts of patients with early-stage CLL initially managed with active surveillance were collected, accounting for a total of 4933 patients (Tables 1 and 2). At the time of treatment, Binet stage and/or indication for therapy were consistent across the study cohorts and were similar to those reported in first-line clinical trials (supplemental Table 1), thus mitigating the risk for biases resulting from heterogeneity in treatment initiation triggers.

By univariable analysis in the UEP training series, which reflects patients managed in daily practice (N = 333; median follow-up, 7.2 years; number of treatment events, 95), baseline factors associated with an increased risk for treatment need (ie, TTFT) with a Bonferroni corrected P value < .05 were palpable lymph nodes, absolute lymphocyte count (>15 × 10⁹/L), mild anemia (hemoglobin between 10 and 13 g/dL in men and 10 and 12 g/dL in women), unmutated IGHV genes, and trisomy 12 (Figure 1A). Sensitivity analysis accounting for death without treatment as competing risk confirmed the univariable associations between baseline variables and TTFT (supplemental Table 2). By multivariable analysis in the training series, 4 variables independently associated with TTFT were identified, including unmutated IGHV genes, absolute lymphocyte count higher than 15 × 10⁹/L, palpable lymph nodes, and trisomy 12 (Figure 1B). The elastic net approach confirmed the stability of the selected variables, as the coefficients for unmutated IGHV genes, absolute lymphocyte count higher than 15 × 10⁹/L, palpable lymph nodes, and trisomy 12 were not penalized and therefore not shrunk to zero. Unmutated IGHV genes, absolute

Figure 1. Univariable and multivariable analysis for time to first treatment. (A) Forest plot of the hazard ratio for the 19 covariates assessed for association with TTFT by univariable analysis. Solid boxes indicate the hazard ratio, horizontal lines indicate the 95% CIs. The red dot line marks hazard ratio of 1. The bar graph on the left shows the multiplicity adjusted -log10 *P* value by Bonferroni. The red solid line marks the .05 significance level. (B) Forest plot of the hazard ratio for the 5 covariates assessed for association with TTFT by multivariable analysis. Solid boxes indicate the hazard ratio, horizontal lines indicate the 95% CIs. The red dot line marks hazard ratio of 1. The bar graph on the left shows the -log10 *P* value. The red solid line marks the .05 significance level. (C-D) The bar plot shows the percentage of covariate selection in the final multivariable model across the 9 Binet A validation cohorts. The red line indicates the threshold selected as inclusion criteria in the final model.
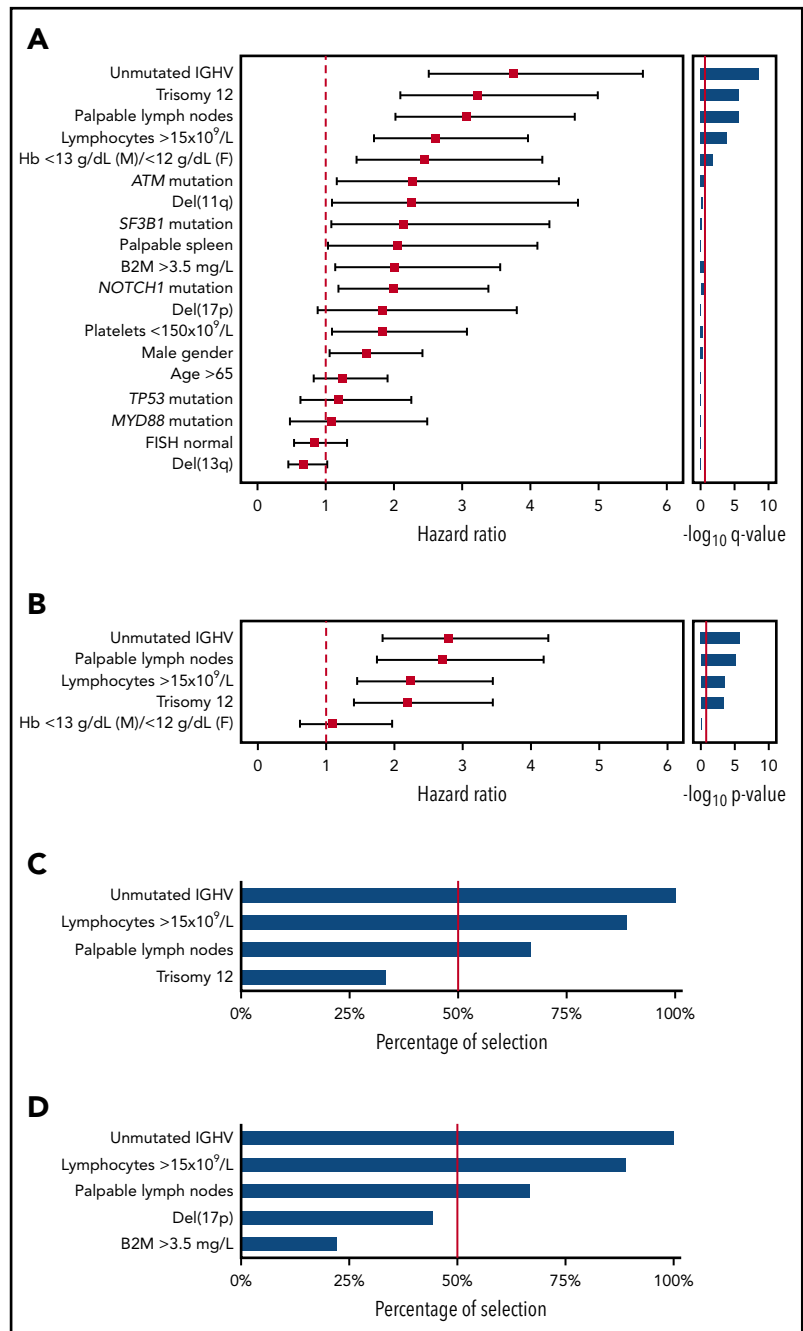
lymphocyte count higher than 15 × 10⁹/L, and palpable lymph nodes, but not trisomy 12, were found to be the most stable biomarkers after iterating the multivariable analysis in each single-validation Binet A cohort (Figure 1C). Across the validation cohorts, unmutated IGHV genes, absolute lymphocyte count higher than 15 × 10⁹/L, and palpable lymph nodes maintained the highest percentage of selection in the final model, as well as after including in the multivariable analyses del(17p) and elevated β-2-microglobulin, which are part of the CLL-IPI score (Figure 1D).[18]

After pruning the model from the less consistent covariates, unmutated IGHV genes, absolute lymphocyte count higher than 15 × 10⁹/L, and palpable lymph nodes were used to build the score and received a weight of 1 because of similar regression coefficients (Figure 1B). The hierarchical order of the sum of the scores in prognosticating TTFT was established by recursive partitioning analysis. This approach allowed us to establish the final international prognostic score to classify patients with asymptomatic CLL in early-stage disease (International Prognostic Score for Early-stage CLL [IPS-E]) according to the risk for treatment need. Patients with a score of 0 had the longest TTFT, and patients with a score of 2 or 3 had the shortest TTFT, whereas patients with a score of 1 had an intermediate TTFT (supplemental Figure 2).

Patients from the training cohort were segregated into 3 distinct risk categories according to the IPS-E (Figure 2A): low (score 0),
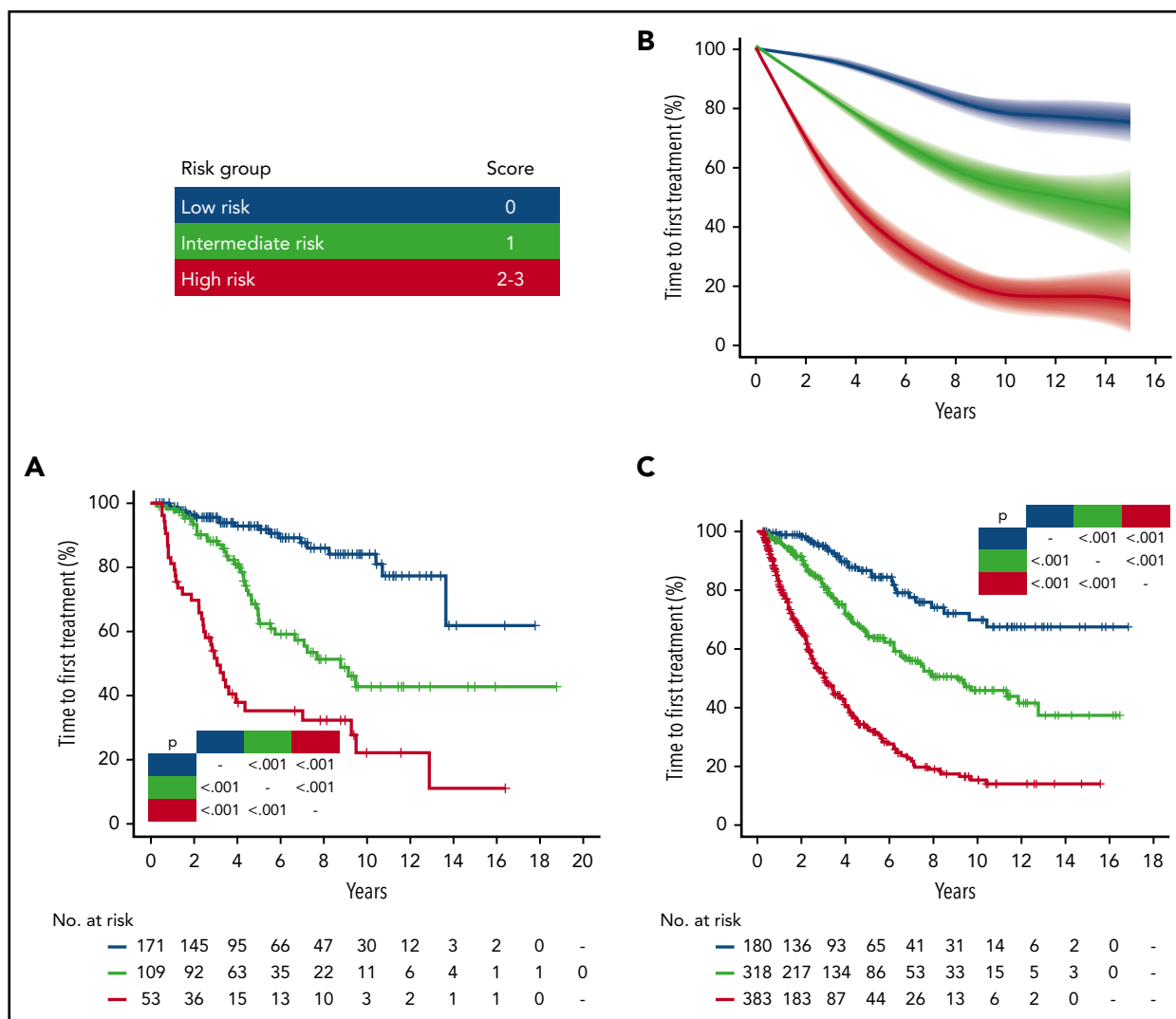
Figure 2. IPS-E stratified TTFT in patients with early-stage CLL managed with active surveillance. (A) Kaplan-Meier curves of TTFT stratified by IPS-E in the UEP discovery cohort. (B) Meta-analytic estimate of TTFT by IPS-E and the corresponding variability across the 9 Binet A validation cohorts. The bold line shows the cubic spline fitted on the meta-analytic estimate of the cumulative proportion of TTFT at each point. The shadow shows the cubic splines fitted on the meta-analytic estimate of the 95% CI of the cumulative proportion of TTFT at each point. (C) Kaplan-Meier curves of TTFT stratified by IPS-E in the Mayo Clinic validation cohort. Blue, low risk; green, intermediate risk; red, high risk by IPS-E. Multiplicity corrected *P* values by pairwise log-rank tests are shown.
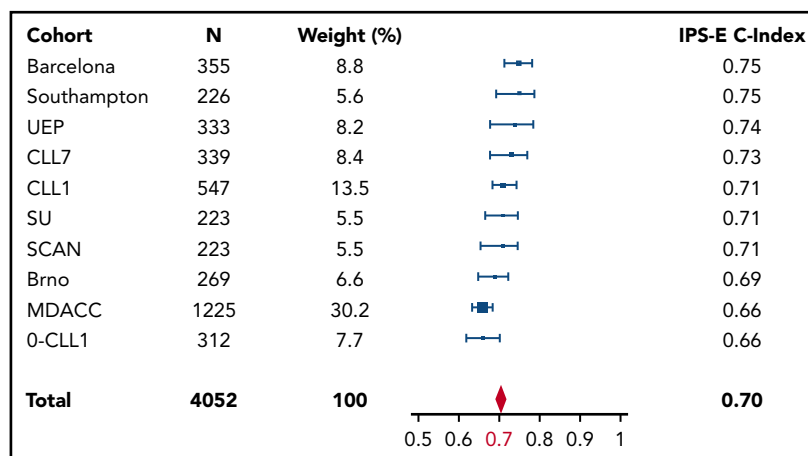
intermediate (score 1), and high (score 2-3) risk. Sensitivity analysis accounting for death before treatment as competing risk consistently confirmed that IPS-E discriminates 3 risk groups with significantly different TTFT (supplemental Figure 3A). The ability of the IPS-E in discriminating TTFT (C-index) was 0.74 (Figure 3).

We validated the IPS-E, using 9 different independent series of patients with Binet stage A CLL. Some heterogeneity across the study cohorts allowed us to robustly confirm the external validity and generalizability of IPS-E irrespective of the case mix of the population. Heterogeneity in the baseline features reflected the source of the cohort; namely, population-based (SCAN: N = 223; median follow-up, 12.9 years; number of treatment events, 104) vs primary care institutions (Barcelona: N = 355; median follow-up, 10.3 years; number of treatment events, 155; Southampton: N = 226; median follow-up, 6.6 years; number of treatment events, 80; Sapienza University: N = 223; median follow-up, 11.3 years; number of treatment events, 117) vs referral centers (M.D. Anderson Cancer Center: N = 1225; median

follow-up, 5.2 years; number of treatment events, 357; Brno: N = 269; median follow-up, 7.7 years; number of treatment events, 135) vs clinical trials (CLL1: N = 547; median follow-up, 8.3 years; number of treatment events, 229; CLL7: N = 339; median follow-up, 4.2 years; number of treatment events, 75; O-CLL-1: N = 312; median follow-up, 7.5 years; number of treatment events, 136). The IPS-E was confirmed and the 3 risk groups with significantly different TTFT were reproduced in the independent validation series both by primary analysis (Figure 2B; supplemental Figure 4A-I) and by sensitivity analysis, accounting for death before treatment as competing risk (supplemental Figure 3B-J). By meta-analysis across the cohorts, the IPS-E C-index for TTFT was 0.70 (Figure 3), whereas the C-index of the sole IGHV mutation status was 0.67 (95% CI, 0.66-0.68; supplemental Figure 5). Approaching the same meta-analytic assessment across the study cohorts, IPS-E showed a slightly higher discrimination capacity of TTFT compared with CLL-IPI (C-index, 0.68; 95% CI, 0.67-0.70; supplemental Figure 6).

Figure 3. Meta-analysis of the estimate of the IPS-E discrimination capacity of TTFT across the study cohorts. Forest plot of the C-index of the 10 Binet A cohorts according to the IPS-E. Solid boxes indicate C-index in each study with dimensions proportional to sample size weights, horizontal lines indicate bootstrap 95% CIs. The red rhombus represents the overall estimate by the meta-analysis. MDACC, M.D. Anderson Cancer Center; SU, Sapienza University.

Cohort	N	Weight (%)		IPS-E C-Index
Barcelona	355	8.8		0.75
Southampton	226	5.6		0.75
UEP	333	8.2		0.74
CLL7	339	8.4		0.73
CLL1	547	13.5		0.71
SU	223	5.5		0.71
SCAN	223	5.5		0.71
Brno	269	6.6		0.69
MDACC	1225	30.2		0.66
0-CLL1	312	7.7		0.66
Total	**4052**	**100**		**0.70**

0.5 0.6 0.7 0.8 0.9 1

The training and validation series were meta-analyzed to provide a more precise estimate of the risk for treatment need in each risk group (Figure 4). Low-risk patients overall corresponded to 29.5% of cases across the 10 Binet stage A CLL cohorts. Their cumulative risk for treatment need (not accounting for death resulting from competing risk) was less than 0.1% after 1 year of surveillance, and 8.4% after 5 years. The risk for treatment was 2.0 events per 100 person years among low-risk patients. Intermediate-risk patients overall corresponded to 36.2% of cases across the 10 Binet stage A CLL cohorts. Their cumulative risk for treatment need (not accounting for death resulting from competing risk) was 3.1% after 1 year of surveillance, and 28.4% after 5 years. The need for therapy was 6.1 events per 100 person years among intermediate-risk patients. High-risk patients overall corresponded to 34.3% of cases across the 10 validation cohorts. Their cumulative risk for treatment need (not accounting for death resulting from competing risk) was 14.1% after 1 year of surveillance and 61.2% after 5 years. The risk for treatment was 16.1 events per 100 person years among high-risk patients. IPS-E performed equivalently over sequential periods, which is consistent with the lack of updates of treatment initiation criteria in the guidelines[3-5,21] and the steadiness of the threshold for commencement of therapy adopted over time by the physicians (supplemental Figure 7).

Two different clinical staging systems are used to stratify patients with CLL; namely, Binet staging, mostly used in European countries, and Rai staging, mostly used in the United States. We aimed at validating the IPS-E also in patients with early-stage CLL defined according to the Rai system (stage 0, I, and II) and managed with active surveillance (Mayo Clinic cohort: N = 881). The 3 IPS-E risk groups were well segregated both in the primary analysis (Figure 2C) and in the sensitivity analysis, accounting for death before treatment as competing risk (supplemental Figure 3K), thus confirming the generalizability of IPS-E also in patients with CLL staged according to the Rai system.

Discussion

The majority of patients with CLL are diagnosed in asymptomatic, early phases of the disease and are followed with no therapy.[1,12] The clinical evolution of these patients is heterogeneous and difficult to predict. Although some patients display a quiescent disease never requiring therapy, others present active disease shortly after diagnosis and require intervention. Despite the general improvement on the outcome of patients with CLL, the relative survival of early-stage patients (ie, Binet stage A) has not significantly changed during the last decades.[2] The management of these patients constitutes a major challenge and should be made on the basis of the best possible evidence and a risk-tailored policy.

There are several prognostic models that can be used to separate patients with different outcome within the whole population of subjects with CLL.[14,18,19,35-40] However, in most of these models, the outcome of patients with early-stage CLL is either not analyzed nor investigated as a unique subgroup, and asymptomatic patients are not separately investigated. Moreover, in most such models, overall survival is the main end point.[18,19] The uniqueness of patients with CLL diagnosed with asymptomatic early disease and the challenges posed by their management make advisable a specific prognostic model with TTFT as a main end point.

An important general caveat is that outcome indicators derived from patients requiring therapy do not necessarily apply to patients not requiring intervention. Among patients with CLL managed with active surveillance, the IGHV unmutated status is the biomarker with the strongest effect on TTFT prognostication. In contrast, other recurrent molecular features associated with inferior overall survival such as TP53 abnormalities, ATM abnormalities, and NOTCH1 mutations did not enter into the IPS-E model. Patients harboring TP53 aberrations, but mutated IGHV genes, may have a prolonged TTFT and a relatively benign clinical course under observation.[41-43] Patients having ATM abnormalities or NOTCH1 mutations are highly enriched with unmutated IGHV genes. However, a large proportion of IGHV unmutated patients does not harbor ATM abnormalities or NOTCH1 mutation.[41-46] Therefore, IGHV unmutated status not only captures ATM and NOTCH1 abnormalities but also expands its prognostic information to cases lacking these unfavorable genetic lesions.

The IPS-E presented in this study is a robust prognostic tool based on routine clinical and laboratory variables that informs at the time of diagnosis about the probability that a given patient with CLL in early-stage disease progresses and needs treatment. The cumulative risk for need of treatment after 1 and 5 years of observation was 14.1% and 61.2%, respectively, for an IPS-E
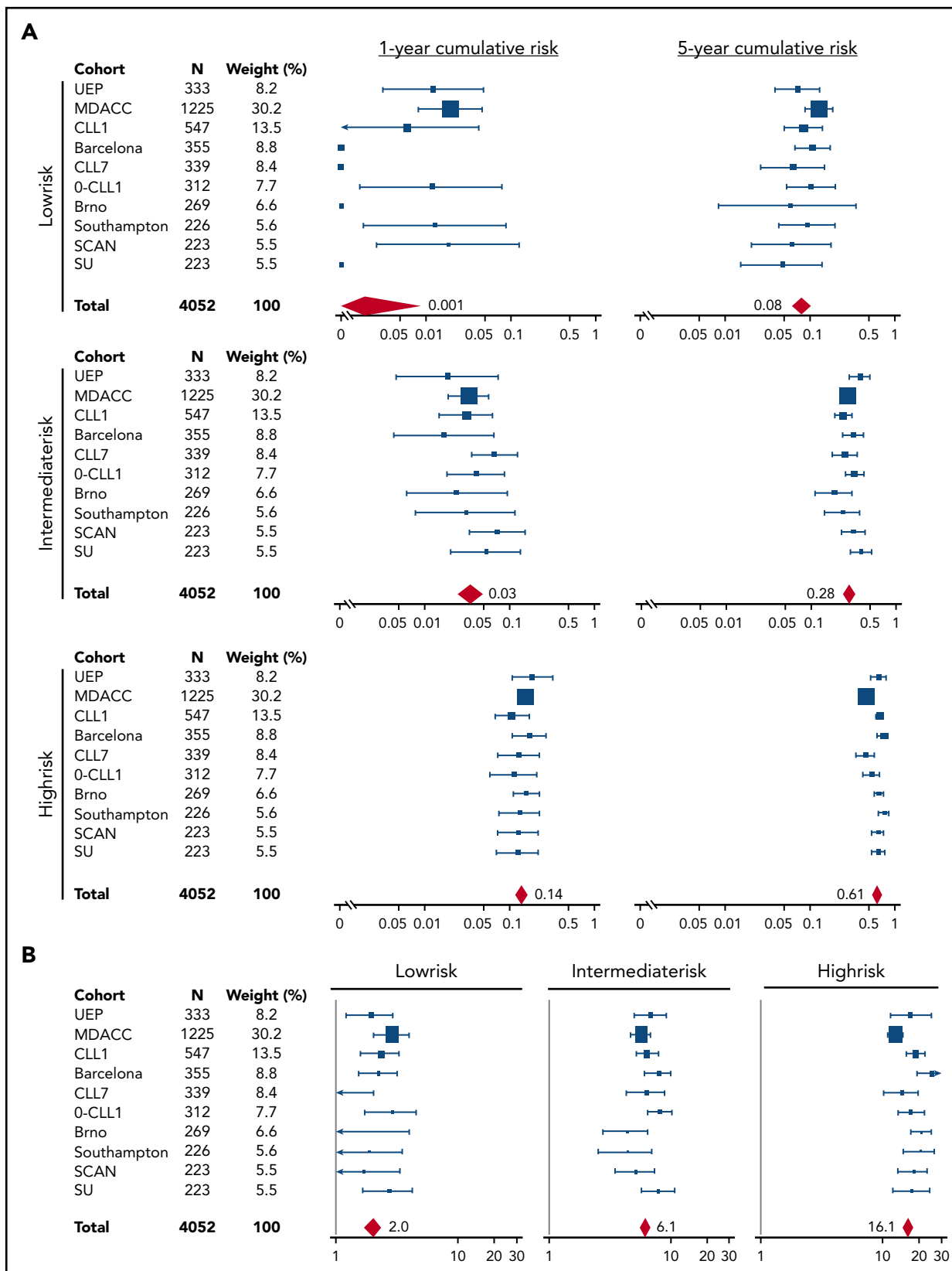
Figure 4. Meta-analysis of the estimates of the IPS-E cumulative incidence of treatment need across the study cohorts. Forest plots of the (A) 1- and 5-year cumulative incidence of treatment need across the 10 Binet A cohorts according to the IPS-E. (B) Forest plots of the number of treatment events per 100 person-years across the 10 Binet A cohorts according to the IPS-E. Solid boxes indicate the estimate in each study with dimensions proportional to sample size weights; horizontal lines indicate 95% CIs. The red rhombus represents the overall estimate by the meta-analysis.

high-risk patient, whereas it was 2.1% and 28.4%, for intermediate-risk patients and less than 0.1% and 8.4% for low-risk patients. Against this backdrop, IPS-E high-risk patients should be followed more closely than low- and intermediate-risk patients because of the likelihood of requiring treatment sooner.

IPS-E calculation is simple, being the sum of 3 variables. Compared with other scores, IPS-E calculation needs the assessment of only 1 molecular variable; namely, the IGHV mutation status, the testing of which is broadly available and standardized. IGHV status has been recognized by the current guidelines as a predictive biomarker for treatment tailoring.[5,22,47] Our results further support the indication for testing IGHV mutations in patients with CLL. Moreover, as IGHV status never changes during the course of disease, it might be evaluated at the time of first diagnosis to provide an estimate of TTFT to the patient and to the treating physician. The results of our study also support that, as recommended by guidelines, FISH analysis and molecular testing for *TP53* has no clinical utility when performed at CLL presentation in early-stage asymptomatic patients. *TP53* abnormalities have a predictive role at time of therapy, but no prognostic role in an early-stage setting when determining TTFT. Also, they can change during the course of disease. Therefore, *TP53* status may not be routinely evaluated in early-stage asymptomatic patients lacking treatment indication (in line with the iwCLL criteria and CLL guidelines).[3-5,21] CLL can exhibit diverse growth patterns, including a continued stable state, never requiring therapy after diagnosis.[48,49] However, a short lymphocyte doubling time is rarely an indication to start therapy by itself, as shown in previous reports[50] and in this study, and therefore was not included as a variable for developing IPS-E.

The simplicity of IPS-E should facilitate its translation to the clinic. The IPS-E could help physicians, medical care providers, and health authorities in their strategies and resources allocation. Likewise, such a prognostic model could be useful to design clinical trials for patients with high-risk, early-stage CLL, an issue that has gained momentum because of the availability of effective small molecules with manageable toxicity.[51,52]

The development of the IPS-E followed an extensively used approach for the generation of prognostic scores in hematology.[18,53-56] However, the risk for biases related to the timing of scheduled evaluations and premature censoring cannot be discarded. In contrast, although the IPS-E has an outcome discrimination capacity that is generally considered as acceptable[57] and similar to that of other CLL prognostication scores now used in clinical practice and for clinical trial design,[18] it can be eventually improved. Therefore, the IPS-E warrants prospective evaluation and can be regarded as a building block to which newly discovered independent outcome predictors for patients with early-stage CLL could be added.

Acknowledgments

Authorship

Contribution: A.C. and L.T.d.B. performed statistical analysis, interpreted data, and wrote the manuscript; P.L., M. A. Hoechstetter, C.D.H., L.D.P., J.D., K.G.R., M. Gentile, M.D., F.R.M., G. Chiodin, M.M., J.B., J.K., K.E.S., R.M., M. A. Hess, and T.S.B. provided clinical data and contributed to manuscript revision; G. Cutrona, C.D., V.S., A.B., and A.N. contributed to molecular studies and manuscript revision; W.W., E.B., B.G., E.Z., S.G., M. Ghielmini, F.C., G.S., M.F., R.R., F.F., R.F., S.P., F.M., S.S., and H.D. contributed to manuscript revision; S.A.P., W.G.W., E.M., G.G., and M.H. provided key scientific insights and contributed to study design, data interpretation, and manuscript revision; and D.R. designed the study, interpreted data, and wrote the manuscript.

Conflict-of-interest disclosure: D.R. received honoraria from AbbVie, AstraZeneca, Gilead, Janssen, and Roche, and research grants from AbbVie, Gilead, Janssen, and Cellestia. M. Hallek received research support and honoraria from Roche, Gilead, Mundipharma, Janssen, Celgene, Pharmacyclics, and AbbVie. G.G. received honoraria from AbbVie, Janssen, AstraZeneca, and Sunesis. G.S. received honoraria from Celgene and Gilead. S.A.P. received research funding (to the institution) from Pharmacyclics, MorphoSys, Janssen, AstraZeneca, and Ascentage Pharma for clinical studies and honoraria from Pharmacyclics, AstraZeneca, Genentech, Gilead, and AbbVie (all to the institution). R.R. received honoraria from Illumina, Roche, Janssen, and AbbVie. R.F. received honoraria from AbbVie, Janssen, Amgen, Novartis, and AstraZeneca. F.F. received honoraria from Janssen-Cilag, and AbbVie, and a research grant from Gilead. C.D.H. received travel support and research funding from Roche. G. Cutrona received a research grant from Gilead (fellowship program 2017). The remaining authors declare no competing financial interests.

ORCID profiles: K.G.R., 0000-0002-7313-1875; G. Chiodin, 0000-0002-1456-8997; G. Cutrona, 0000-0002-3335-1101; B.G., 0000-0001-5418-0949; E.Z., 0000-0002-5522-6109; A.N., 0000-0001-9047-5912; F.F., 0000-0002-2211-1831; S.P., 0000-0001-7136-2680; F.M., 0000-0002-2585-7073; S.A.P., 0000-0002-3221-7314.

Correspondence: Davide Rossi, Department of Hematology, Oncology Institute of Southern Switzerland, Via Ospedale, 6500 Bellinzona, Switzerland; e-mail: davide.rossi@eoc.ch.

Footnotes

REFERENCES

1. Teras LR, DeSantis CE, Cerhan JR, Morton LM, Jemal A, Flowers CR. 2016 US lymphoid malignancy statistics by World Health Organization subtypes. *CA Cancer J Clin*. 2016; 66(6):443-459.

2. Abrisqueta P, Pereira A, Rozman C, et al. Improving survival in patients with chronic lymphocytic leukemia (1980-2008): the Hospital Clinic of Barcelona experience. *Blood*. 2009;114(10):2044-2050.

3. Cheson BD, Bennett JM, Grever M, et al. National Cancer Institute-sponsored Working Group guidelines for chronic lymphocytic leukemia: revised guidelines for diagnosis and treatment. *Blood*. 1996;87(12):4990-4997.

4. Hallek M, Cheson BD, Catovsky D, et al; International Workshop on Chronic Lymphocytic Leukemia. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood*. 2008;111(12): 5446-5456.

5. Eichhorst B, Robak T, Montserrat E, et al. Chronic lymphocytic leukaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2015; 26(Suppl 5):v78-v84.

6. Shustik C, Mick R, Silver R, Sawitsky A, Rai K, Shapiro L. Treatment of early chronic lymphocytic leukemia: intermittent chlorambucil versus observation. *Hematol Oncol*. 1988;6(1): 7-12.

7. Dighiero G, Maloum K, Desablens B, et al; French Cooperative Group on Chronic Lymphocytic Leukemia. Chlorambucil in indolent chronic lymphocytic leukemia. *N Engl J Med*. 1998;338(21):1506-1514.

8. Schweighofer CD, Cymbalista F, Müller C, et al. Early versus deferred treatment with combined fludarabine, cyclophosphamide and rituximab (FCR) improves event-free survival in patients with high-risk Binet stage A chronic lymphocytic leukemia – first results of a randomized German-French Cooperative Phase III Trial. *Blood*. 2013;122(21):524.

9. Hoechstetter MA, Busch R, Eichhorst B, et al. Early, risk-adapted treatment with fludarabine in Binet stage A chronic lymphocytic leukemia patients: results of the CLL1 trial of the German CLL study group. *Leukemia*. 2017;31(12): 2833-2837.

10. Chemotherapeutic options in chronic lymphocytic leukemia: a meta-analysis of the randomized trials. CLL Trialists' Collaborative Group. *J Natl Cancer Inst*. 1999;91(10): 861-868.

11. Hallek M, Shanafelt TD, Eichhorst B. Chronic lymphocytic leukaemia. *Lancet*. 2018; 391(10129):1524-1537.

12. Sant M, Allemani C, Tereanu C, et al; HAEMACARE Working Group. Incidence of hematologic malignancies in Europe by morphologic subtype: results of the HAEMACARE project. *Blood*. 2010;116(19): 3724-3734.

13. Langerbeins P, Bahlo J, Rhein C, et al. Ibrutinib versus placebo in patients with asymptomatic, treatment-naïve early stage CLL: primary endpoint results of the phase 3 double-blind randomized CLL12 trial. *Hematol Oncol*. 2019;37(S2):38-40.

14. Wierda WG, O'Brien S, Wang X, et al. Multivariable model for time to first treatment in patients with chronic lymphocytic leukemia. *J Clin Oncol*. 2011;29(31):4088-4095.

15. Molica S, Giannarelli D, Gentile M, et al. External validation on a prospective basis of a nomogram for predicting the time to first treatment in patients with chronic lymphocytic leukemia. *Cancer*. 2013;119(6):1177-1185.

16. Mansouri L, Cahill N, Gunnarsson R, et al. NOTCH1 and SF3B1 mutations can be added to the hierarchical prognostic classification in chronic lymphocytic leukemia. *Leukemia*. 2013;27(2):512-514.

17. Rossi D, Rasi S, Spina V, et al. Integrated mutational and cytogenetic analysis identifies new prognostic subgroups in chronic lymphocytic leukemia. *Blood*. 2013;121(8): 1403-1412.

18. International CLL-IPI working group. An international prognostic index for patients with chronic lymphocytic leukaemia (CLL-IPI): a meta-analysis of individual patient data. *Lancet Oncol*. 2016;17(6):779-790.

19. Delgado J, Doubek M, Baumann T, et al. Chronic lymphocytic leukemia: A prognostic model comprising only two biomarkers (IGHV mutational status and FISH cytogenetics) separates patients with different outcome and simplifies the CLL-IPI. *Am J Hematol*. 2017; 92(4):375-380.

20. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM; Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics. Reporting recommendations for tumor marker prognostic studies (REMARK). *J Natl Cancer Inst*. 2005; 97(16):1180-1184.

21. Hallek M, Cheson BD, Catovsky D, et al. iwCLL guidelines for diagnosis, indications for treatment, response assessment, and supportive management of CLL. *Blood*. 2018; 131(25):2745-2760.

22. Bain BJ. Ethnic and sex differences in the total and differential white cell count and platelet count. *J Clin Pathol*. 1996;49(8):664-666.

23. Damle RN, Wasil T, Fais F, et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood*. 1999;94(6):1840-1847.

24. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood*. 1999; 94(6):1848-1854.

25. Beutler E, Waalen J. The definition of anemia: what is the lower limit of normal of the blood hemoglobin concentration? *Blood*. 2006; 107(5):1747-1750.

26. Rossi D, Condoluci A, Spina V, Gaidano G. Methods for measuring ctDNA in lymphomas. *Methods Mol Biol*. 2019;1881:253-265.

27. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53(282):457-481.

28. Thabane L, Mbuagbaw L, Zhang S, et al. A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. *BMC Med Res Methodol*. 2013;13(1):92.

29. Dignam JJ, Kocherginsky MN. Choice and interpretation of statistical tests used when competing risks are present. *J Clin Oncol*. 2008;26(24):4027-4034.

30. Cox DR. Regression models and life tables. *J R Stat Assoc*. 1972;34:187-220.

31. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for Cox's proportional hazards model via coordinate descent. *J Stat Softw*. 2011;39(5):1-13.

32. Schoenfeld D. Partial residuals for the proportional hazard regression model. *Biometrika*. 1982;69(1):239-241.

33. Ciampi A, Negassa A, Lou Z. Tree-structured prediction for censored survival data and the Cox model. *J Clin Epidemiol*. 1995;48(5): 675-689.

34. Harrell FE Jr. Regression modeling strategies. New York, NY: Springer-Verlag; 2001.

35. Rai KR, Sawitsky A, Cronkite EP, Chanana AD, Levy RN, Pasternack BS. Clinical staging of chronic lymphocytic leukemia. *Blood*. 1975; 46(2):219-234.

36. Wierda WG, O'Brien S, Wang X, et al. Prognostic nomogram and index for overall survival in previously untreated patients with chronic lymphocytic leukemia. *Blood*. 2007; 109(11):4679-4685.

37. Pflug N, Bahlo J, Shanafelt TD, et al. Development of a comprehensive prognostic index for patients with chronic lymphocytic leukemia. *Blood*. 2014;124(1):49-62.

38. Gentile M, Shanafelt TD, Cutrona G, et al. A progression-risk score to predict treatment-free survival for early stage chronic lymphocytic leukemia patients. *Leukemia*. 2016;30(6): 1440-1443.

39. Molica S, Giannarelli D, Mirabelli R, et al. Reliability of six prognostic models to predict time-to-first-treatment in patients with chronic lymphocytic leukaemia in early phase. *Am J Hematol*. 2017;92(6):E91-E93.

40. Montserrat E, Gale RP. Predicting the outcome of patients with chronic lymphocytic leukemia: Progress and uncertainty. *Cancer*. 2019;125(21):3699-3705.

41. Best OG, Gardiner AC, Davis ZA, et al. A subset of Binet stage A CLL patients with TP53 abnormalities and mutated IGHV genes have stable disease. *Leukemia*. 2009;23(1): 212-214.

42. Tam CS, Shanafelt TD, Wierda WG, et al. De novo deletion 17p13.1 chronic lymphocytic leukemia shows significant clinical heterogeneity: the M. D. Anderson and Mayo Clinic experience. *Blood*. 2009;114(5):957-964.

43. Hu B, Patel KP, Chen HC, et al. Association of gene mutations with time-to-first treatment in 384 treatment-naive chronic lymphocytic leukaemia patients. *Br J Haematol*. 2019; 187(3):307-318.

44. Nadeu F, Delgado J, Royo C, et al. Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in

chronic lymphocytic leukemia. *Blood*. 2016; 127(17):2122-2130.

45. Montserrat E, Bauman T, Delgado J. Present and future of personalized medicine in CLL. *Best Pract Res Clin Haematol*. 2016;29(1): 100-110.

46. Baliakas P, Hadzidimitriou A, Sutton LA, et al; European Research Initiative on CLL (ERIC). Recurrent mutations refine prognosis in chronic lymphocytic leukemia. *Leukemia*. 2015;29(2):329-336.

47. Wierda WG, Byrd JC, Abramson JS, et al. NCCN Guidelines Insights: Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma, Version 2.2019. *J Natl Compr Canc Netw*. 2019;17(1):12-20.

48. Gruber M, Bozic I, Leshchiner I, et al. Growth dynamics in naturally progressing chronic lymphocytic leukaemia. *Nature*. 2019; 570(7762):474-479.

49. Montserrat E, Sánchez-Bisonó J, Viñolas N, Rozman C. Lymphocyte doubling time in chronic lymphocytic leukaemia: analysis of its prognostic significance. *Br J Haematol*. 1986;62(3):567-575.

50. Mozas P, Rivas-Delgado A, Baumann T, et al. Analysis of criteria for treatment initiation in patients with progressive chronic lymphocytic leukemia. *Blood Cancer J*. 2018;8(1):10.

51. Sharma S, Rai KR. Chronic lymphocytic leukemia (CLL) treatment: So many choices, such great options. *Cancer*. 2019;125(9): 1432-1440.

52. Hallek M, Fürstenau M. How to approach CLL in clinical practice. *Hematol Oncol*. 2019; 37(Suppl 1):38-42.

53. 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.

54. Solal-Céligny P, Roy P, Colombat P, et al. Follicular lymphoma international prognostic index. *Blood*. 2004;104(5): 1258-1265.

55. 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.

56. Hoster E, Dreyling M, Klapper W, et al; European Mantle Cell Lymphoma Network. A new prognostic index (MIPI) for patients with advanced-stage mantle cell lymphoma. *Blood*. 2008;111(2): 558-565.

57. Hosmer DW, Lemeshow S. Applied logistic regression. 2nd ed. New York: John Wiley & Sons; 2000.