ARTICLE

A laboratory-based scoring system predicts early treatment in Rai 0 chronic lymphocytic leukemia

Jared A. Cohen,[1] Francesca Maria Rossi,[1] Antonella Zucchetto,[1] Riccardo Bomben,[1] Lodovico Terzi-di-Bergamo,[2] Kari G. Rabe,[3] Massimo Degan,[1] Agostino Steffan,[4] Jerry Polesel,[5] Enrico Santinelli,[6] Idanna Innocenti,[7] Giovanna Cutrona,[8] Giovanni D'Arena,[9] Gabriele Pozzato,[10] Francesco Zaja,[10] Annalisa Chiarenza,[11] Davide Rossi,[12,13] Francesco Di Raimondo,[11] Luca Laurenti,[7] Massimo Gentile,[14,15] Fortunato Morabito,[15,16] Antonino Neri,[17] Manlio Ferrarini,[18] Christopher D. Fegan,[19] Christopher J. Pepper,[19,20] Giovanni Del Poeta,[6] Sameer A. Parikh,[21] Neil E. Kay[21] and Valter Gattei[1]

[1]Clinical and Experimental Onco-Haematology Unit, Centro di Riferimento Oncologico, I.R.C.C.S., Aviano (PN), Italy; [2]Institute of Oncology Research, Bellinzona, Switzerland; [3]Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA; [4]Immunopathology and Cancer Biomarkers, Centro di RiferimentoOncologico, I.R.C.C.S., Aviano (PN), Italy; [5]Unit of Cancer Epidemiology, Centro di Riferimento Oncologico, I.R.C.C.S., Aviano (PN), Italy; [6]Division of Haematology, S. Eugenio Hospital and University of Tor Vergata, Rome, Italy; [7]Hematology Institute, Catholic University of the Sacred Heart, Fondazione Policlinico A. Gemelli, Rome, Italy; [8]UO Molecular Pathology, Ospedale Policlinico San Martino IRCCS, Genova, Italy; [9]Onco-Haematology Department, Centro di Riferimento Oncologico della Basilicata, I.R.C.C.S., Rionero in Vulture, Italy; [10]Department of Internal Medicine and Haematology, Maggiore General Hospital, University of Trieste, Trieste, Italy; [11]Division of Haematology, Ferrarotto Hospital, Catania, Italy; [12]Hematology, Oncology Institute of Southern Switzerland, Bellinzona, Switzerland; [13]Universita' della Svizzera Italiana, Lugano, Switzerland; [14]Hematology Unit, AO, Cosenza, Italy; [15]Biotechnology Research Unit, Aprigliano, Cosenza, Italy; [16]Hematogy Department and Bone Marrow Transplant Unit, Cancer Care Center, Augusta Victoria Hospital, East Jerusalem, Israel; and [17]Hematology Unit, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico and University of Milan, Milan, Italy; [18]Department of Experimental Medicine, University of Genova, Genova, Italy; [19]Division of Cancer and Genetics, Cardiff University, School of Medicine, Heath Park, Cardiff, UK; [20]University of Sussex, Brighton and Sussex Medical School, Brighton, UK and [21]Division of Hematology, Department of Medicine, Mayo Clinic, Rochester, MN, USA

Ferrata Storti Foundation

Haematologica 2020
Volume 105(6):1613-1620

ABSTRACT

We present a laboratory-based prognostic calculator (designated CRO score) to risk stratify treatment-free survival in early stage (Rai 0) chronic lymphocytic leukemia (CLL) developed using a training-validation model in a series of 1,879 cases from Italy, the United Kingdom and the United States. By means of regression analysis, we identified five prognostic variables with weighting as follows: deletion of the short arm of chromosome 17 and unmutated immunoglobulin heavy chain gene status, 2 points; deletion of the long arm of chromosome 11, trisomy of chromosome 12, and white blood cell count >32.0x10^3/microliter, 1 point. Low-, intermediate- and high-risk categories were established by recursive partitioning in a training cohort of 478 cases, and then validated in four independent cohorts of 144 / 395 / 540 / 322 cases, as well as in the composite validation cohort. Concordance indices were 0.75 in the training cohort and ranged from 0.63 to 0.74 in the four validation cohorts (0.69 in the composite validation cohort). These findings advocate potential application of our novel prognostic calculator to better stratify early-stage CLL, and aid case selection in risk-adapted treatment for early disease. Furthermore, they support immunocytogenetic analysis in Rai 0 CLL being performed at the time of diagnosis to aid prognosis and treatment, particularly in today's chemo-free era.

Correspondence:

JARED A. COHEN
jcoen781@gmail.com

VALTER GATTEI
vgattei@cro.it

Received: June 3, 2019.

Accepted: October 2, 2019.

Pre-published: October 3, 2019.

doi:10.3324/haematol.2019.228171

Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: www.haematologica.org/content/105/6/1613

Introduction

Clinical staging using the Binet and Rai classification systems provides a simple and inexpensive approach to assess prognosis in chronic lymphocytic leukemia (CLL).[1,2] However, most patients today are diagnosed in early stages of the disease (Binet A or Rai 0) when these prognosticators fail to provide adequate risk stratification.[3] Although similarly classified as early-stage CLL, Binet A and Rai 0 patients demonstrate heterogeneous clinical courses ranging from normal life expectancy in the absence of any treatment to unexpectedly short progression-free intervals rapidly requiring clinical intervention.[4]

To overcome the inherent weakness of clinical staging systems, other parameters have been sought and proposed by several studies as reliable prognosticators in CLL, including immunocytogenetic and molecular markers such as deletions of the short arm of chromosome 17 (del17p) and mutations of the *TP53* gene, deletions of the long arm of chromosome 11 (del11q), and trisomy of chromosome 12 (tris12), the immunoglobulin heavy chain (IGHV) gene mutational status as well as biochemical parameters such as beta-2-microglobulin (B2M) and thymidine kinase (TK) and cell surface receptors such as the integrin CD49d.[5-10]

Novel prognostic indices and model systems have been developed to integrate these markers into comprehensive scoring systems, such as the CLL International Prognostic Index (CLL-IPI), the German CLL Study Group (GCLLSG) index, and the MD Anderson Cancer Center (MDACC) score.[11-14] Although validations in the setting of early-stage CLL and/or treatment-free-survival (TFS) prediction have been undertaken, these indices were originally generated to predict overall survival operating across all stages of disease.[11,13,15,16]

Here we present a novel laboratory-based prognostic index specifically developed to predict TFS in Rai 0 CLL, thus allowing clinicians and researchers to uniformly and more accurately identify cases with higher risk for needing early treatment.

Methods

We applied a training-validation strategy using 1,879 cases of phenotypically confirmed Rai 0 CLL[17] collected in the context of an international effort from Italy, the United Kingdom and the United States (Figure 1). The training cohort included 478 Rai 0 cases identified from a consecutive series of Italian multicenter patients (1,201 cases) referred to a single center (Clinical and Experimental Onco-Hematology Unit of the Centro Riferimento Oncologico in Aviano, Italy) for immunocytogenetic analyses between 2006 and 2017. Four independent Rai 0 cohorts were used for external validation made up of three 'real world' cohorts from single centers, i) Gemelli Hospital in Rome, Italy (144 cases, Gemelli cohort), ii) Cardiff University Hospital in Wales, UK (395 cases, Cardiff cohort), iii) Mayo Clinic in Rochester, MN, USA (540 cases, Mayo cohort), and one investigational cohort from the multicenter O-CLL1-GISL Italian prospective observational study (O-CLL cohort; *clinicaltrial.gov identifier: 00917540*; 322 cases) (Figure 1). Cases of monoclonal B lymphocytosis were excluded, and the TFS was defined as time from diagnosis to treatment, according to the revised 2018 International Workshop on Chronic Lymphocytic Leukemia (IWCLL) guidelines.[18] Patient information was obtained from the participating centers in accordance with the Declaration of Helsinki and local ethics committee approvals (Approvals n. IRB-05-2010, LREC #02/4806, IRB-12-000969 and NCT00917540).

Deletions at chromosomes 13q14 (del13q), 11q23 (del11q), 17p13 (del17p), and trisomy 12 (tris12), and IGHV gene status were determined at the different participating centers, as reported previously.[5,19] Cytogenetic thresholds were set at 5% for del13q, del11q, and tris12, and 10% for del17p, and cases were categorized according to the hierarchical model proposed by Dohner *et al.*[7] The positive fluorescence *in situ* hybridization (FISH) threshold of 10% for del17p was selected in accordance with the European Research Initiative on CLL (ERIC) recommendations and previous clinical studies.[20-22] IGHV status was considered unmutated (UM) at ≥98%,[6] and CD49d positivity (only determined in the training cohort) was set at >30%, as reported previously.[19] Investigation of mutations for *TP53* (exons 2–11) and *NOTCH1* (exon 34) (determined in 304 of 478 cases in the training cohort) was performed

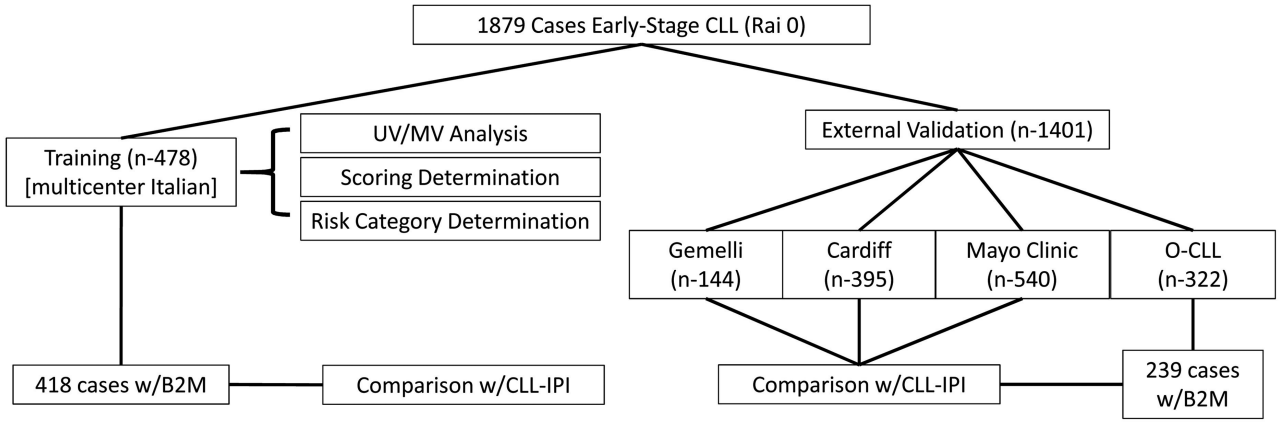

Figure 1. Study design. Training cohort: 478 cases included for univariable (UV) and multivariable (MV) analyses, scoring determination and risk category determination; 418 of 478 cases with beta-2-microglobulin (B2M) data (w/B2M) were employed for comparison with chronic lymphocytic leukemia International Prognostic Index (CLL-IPI) (w/CLL-IPI). Validation cohorts included 1,401 Rai 0 cases. The Italian prospective observational study (O-CLL) cohort had 239 of 322 cases available for comparison to CLL-IPI. In the remaining validation cohorts, all cases were available for comparison with CLL-IPI. n: number.

by a next-generation sequencing (NGS) approach, as previously reported.[5,23-26] In the case of *TP53* mutation, cases were defined as *TP53* mutated if the variant allele frequency was >10% according to ERIC recommendations.[27]

The Cox proportional hazards regression model was chosen to assess the independent effect of co-variables on TFS, with a stepwise procedure for selecting significant variables. All co-variables, apart from FISH categories,[7] were treated as dichotomous and evaluated at diagnosis. Independent variables were internally validated using bootstrapping procedures and weighted based on the proportion of their normalized hazard ratios (HR) rounded to the nearest whole integer (Table 1). Risk-categories were determined by recursive partitioning (*Online Supplementary Figure S1*), and Kaplan-Meier analyses were used to generate survival curves.

In the training cohort, five cases died without treatment and were censored at the date of death. A sensitivity analysis for competitive risk, conducted on the training cohort according to the Fine-Gray model,[28] reported no substantial modification in level of risk (*data not shown*).

The concordance index (C-index) was used to compare our model with the CLL-IPI[11] in 418 of 478 (training cohort), 144 of 144 (Gemelli cohort), 395 of 395 (Cardiff cohort), 540 of 540 (Mayo cohort), 239 of 322 (O-CLL cohort), and 1,318 of 1,401 (composite validation cohort) cases with available B2M data. In all cases, the statistical significance between C-indices was evaluated by applying the Student *t*-test and internally validated by applying a bootstrapping procedure. The Akaike information criterion (AIC) was also employed as an estimator of the relative quality of the model proposed in this study in comparison to the CLL-IPI as TFS predictors. When applicable, $P<0.05$ was considered statistically significant. Statistical calculations were made using MedCalc or the open source R package (*http://www.r-project.org/*) statistics software.

Results

Identification of the training cohort and construction of a scoring system

The TFS curves of a consecutive series of 1,201 cases from a single center, split according to Rai staging, are reported in Figure 2A. As expected, the median [95% confidence interval (CI)] TFS of the 478 Rai 0 cases was significantly longer at 124 months (m) (104-183 m) compared to that of Rai I-IV cases, with a median (95%CI) follow up of 62 m (57-68 m 95%CI). The baseline characteristics of this Rai 0 cohort are summarized in *Online Supplementary Table S1*. When Dohner's hierarchical model[7] was applied to this Rai 0 cohort, cases bearing either del17p, del11q or tris12 experienced the shortest TFS, with no difference among these three cytogenetic categories, while similarly longer TFS intervals were observed for del13q cases and cases lacking the four major chromosomal abnormalities (Figure 2B).

Therefore, the presence of these three chromosomal aberrations were included, along with white blood cell (WBC) counts, IGHV gene status, CD49d expression, gender and age in a univariate analysis. For the purposes of the present study, WBC counts were dichotomized according to the cut off of >32x10^3 cells/μL (32K), as established by a maximally selected log rank analysis carried out in the training cohort (*Online Supplementary Figure S2*). In patients with WBC counts ≤32K and >32K, median (95%CI, x10^3 cell/μL) counts were 15.3 (15.0-16.8) *versus* 54.7 (49.8-58.0), respectively, without clustering around the threshold value (*Online Supplementary Figure S3*).

Apart from age and gender, all of the tested variables associated with a shorter TFS, and five of them (del17p, del11q, tris12, WBC and IGHV gene status) emerged as independent predictors of short TFS by multivariate analysis (Table 1).

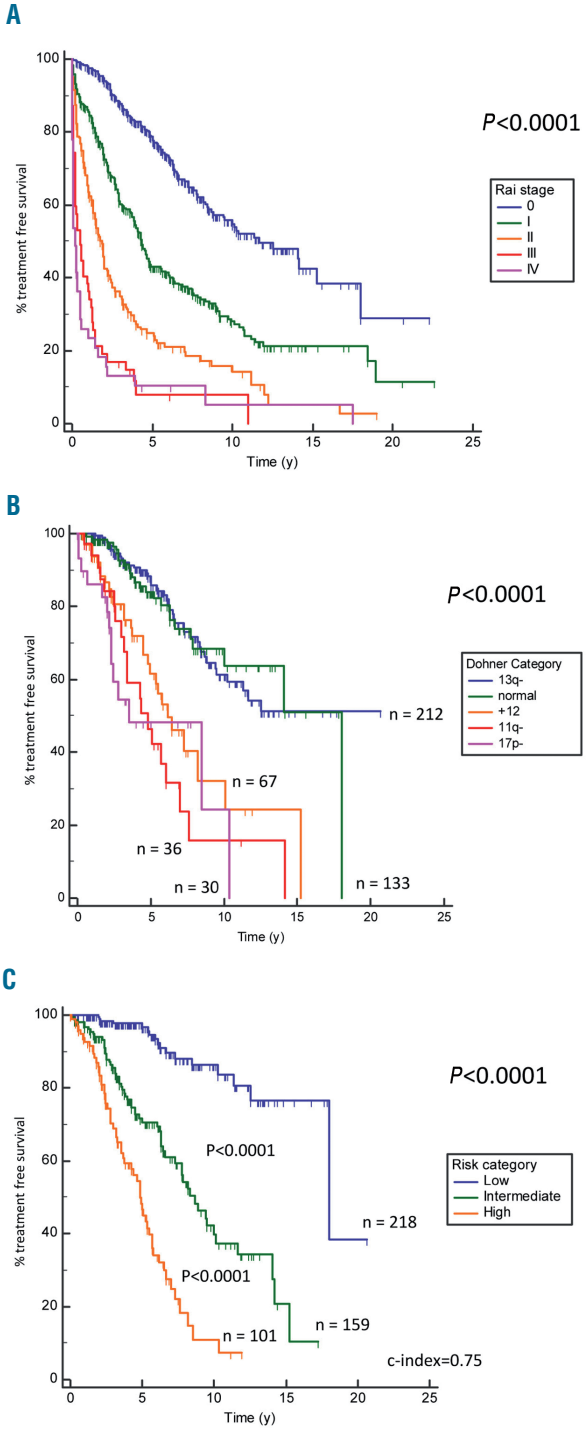

Figure 2. Determination of the Rai 0 training cohort and laboratory-based prognostic calculator (CRO) scoring. Kaplan-Meier curves demonstrating % treatment-free survival (TFS) for (A) a consecutive series of 1,201 cases of chronic lymphocytic leukemia (CLL) referred to our center between 2006 and 2017 from which our training cohort was derived; stratified by Rai stage. The training cohort of 478 Rai 0 cases organized according to (B) Dohner's hierarchical model and (C) CRO score. n: number; y: years. c-index: concordance index.

To construct a scoring system (hereafter designated as the CRO score) using these five independent predictors, a point value of 1 or 2 was assigned to variables according to their respective normalized hazard ratios as follows: i) 2 points to del17p, and UM IGHV; ii) 1 point to del11q, tris12, and WBC count >32K (Table 1). Then, three risk groups, based on point cut offs of 0 (low risk, 218 cases), 1-2 (intermediate risk, 159 cases), and 3-5 (high risk, 101 cases) were established by recursive partitioning analysis (*Online Supplementary Figure S1*). The median TFS (95%CI) was 216 m (216-216 m), 104 m (93-140 m) and 58 m (44-68 m) (P<0.0001) for the low-, intermediate-, and high-risk groups, respectively, with a C-index of 0.75 (Figure 2C). A comparison with the CLL-IPI (possible in 418 of 478 cases with available B2M data) indicated a C-index of 0.76 for the CRO score compared to 0.69 when patient risk groups were split according to the CLL-IPI (P<0.0001) (*Online Supplementary Figure S4A*).

External validation of the CRO score

The CRO score was then validated in four independent cohorts of Rai 0 CLL; baseline patient characteristics are reported in *Online Supplementary Table S1*. These cohorts demonstrated similar TFS survival curves when compared to each other and the training series (*Online Supplementary Figure S5*). Results were as follows (see also Figure 3):

i) the 144 cases of the Gemelli cohort (Figure 3A) had a median (95%CI) follow up of 91 m (83-103 m). Overall, median (95% CI) TFS was 86 m (80-94 m) (*Online Supplementary Figure S5*), while median (95%CI) TFS for the low- (96 cases), intermediate- (36 cases) and high-risk (12 cases) groups was 239 m (range, 239-239 m), 98 m (92-132 m) and 85 m (60-109 m), respectively (P=0.002 between low- and intermediate-risk groups, P=0.09 between intermediate- and high-risk groups). In this cohort, the C-indices were 0.64 and 0.61 for the CRO score and the CLL-IPI, respectively (P<0.0001) (*Online Supplementary Figure S4B*);

ii) the Cardiff cohort (395 cases) had a median (95%CI) follow up of 94 m (83-104 m). Median (95%CI) TFS was 74 m (67-81 m) overall (*Online Supplementary Figure S5*), and not reached (NR), 111 m (97-146 m) and 70 m (29-114 m) for the low- (206 cases), intermediate- (136 cases), and high-risk (53 cases) groups, respectively (P<0.001 between low- and intermediate-risk groups, P=0.009 between intermediate- and high-risk groups) (Figure 3B); C-index was 0.63 for both the CRO score and the CLL-IPI (P=not significant, ns) (*Online Supplementary Figure S4C*);

iii) the Mayo cohort (540 cases) had a median (95%CI) follow up of 77 m (68-88 m). Median (95%CI) TFS was 127 m (96 m-NR) overall (*Online Supplementary Figure S5*), and NR, 76 m (range, 64 m-NR) and 36 m (range, 31-59 m) for the low- (278 cases), intermediate- (168 cases) and high-risk (94 cases) groups, respectively (P<0.0001) (Figure 3C); C-indices were 0.72 and 0.68 for the CRO score and the CLL-IPI, respectively (P<0.0001) (*Online Supplementary Figure S4D*);

iv) the multicenter O-CLL cohort (322 cases) had a median (95%CI) follow up of 89 m (85-95 m), while median (95%CI) TFS was NR overall (*Online Supplementary Figure S5*), and NR, 96 m (83-110 m) and 48 m (39-67 m) for the low- (189 cases), intermediate- (84 cases) and high-risk (49 cases) groups, respectively (P<0.001 between low- and intermediate-risk groups, P=0.003 between intermediate- and high-risk groups; C-index 0.74) (Figure 3D). In the 239 cases with available B2M data, the C-indices were 0.71 and 0.70 for the CRO score and the CLL-IPI, respectively (P<0.001) (*Online Supplementary Figure S4E*).

The composite TFS curve included 1,401 Rai 0 cases by combining cases from the four validation cohorts. Median (95%CI) TFS was 175 m (143-201 m) overall and NR, 106 m (96-134 m) and 45 m (39-60 m) for the low-, intermediate- and high-risk groups, respectively (P<0.0001; C-index 0.69) (Figure 3E). In the 1,318 cases with available B2M data, the C-indices were 0.68 for the CRO score and 0.66 for the CLL-IPI (P<0.0001) (*Online Supplementary Figure S6*). In this context, AIC was 4,881 for the CRO score *versus* 4,912 for the CLL-IPI, in keeping with a relative better quality of the former as a TFS predictor.

Sub-analyses of the CRO score high-risk group

By combining the training and four validation cohorts, 309 of 1,879 cases (16%) were identified as having relatively higher risk of early progression and treatment according to the CRO score (scores 3, 4 or 5). In this subset, the vast majority of cases had a CRO score of 3 (223 cases, 72%) or 4 (72 cases, 23%); however, a small minority of patients (14 cases, 4.5%) had a CRO score of 5, i.e. presented with a WBC count >32K in the context of disease bearing both del17p and an UM IGHV gene status (*Online Supplementary Table S2*). Although the median TFS of cases with CRO scores 3 and 4 was similar, a significantly shortened median TFS was demonstrated in

Table 1. Univariable and reduced multivariable analysis of six factors used to generate our risk calculator.

Factor	Univariable analysis			Reduced multivariable analysis			
	P	HR	95% CI	*P*	HR	95% CI	Weight
WBC>32K cells/L	<0.0001	2.96	2.10-4.16	<0.0001	2.39	1.69 - 3.38	1
FISH category							
del17p	<0.0001	4.38	2.46 - 7.80	0.0002	3.03	1.69 - 5.44	2
del11q	<0.0001	4.02	2.45 - 6.19	0.0049	2.13	1.26 - 3.62	1
tri12	<0.0001	2.85	1.84 - 4.42	0.025	1.7	1.07 - 2.71	1
UM *IGHV*	<0.0001	4.08	2.86 - 5.80	<0.0001	2.91	1.97 - 4.29	2
CD49d$^+$	0.001	1.78	1.27 - 2.51				
Age>65 years	0.0536	1.4	0.99 - 1.98				
Male	0.9232	0.98	0.70 - 1.38				

Fluorescence *in situ* hybridization (FISH) categories were as reported by Dohner *et al.*[7] Weights were determined using the proportion of normalized hazard ratios rounded to the nearest whole integer. HR: hazard ratio; CI: confidence interval; WBC: white blood cell; UM: unmutated.

patients classified as high-risk with a CRO score of 5 (27.6 m, 95%CI: 14.4-28.8; *P*=0.01) (*Online Supplementary Table S2* and *Online Supplementary Figure S7*).

Application of the CRO score in Rai I patients

To assess the generalizability of the CRO score in patients beyond Rai 0 disease, we applied our scoring system to a consecutive series of Italian multicenter patients with Rai I CLL (375 cases) referred to our center for immunocytogenic analyses between 2006 and 2017. Our prognostic calculator demonstrated excellent predictive performance in this cohort (c-index 0.67) with a median (95%CI) TFS of 37 m (47-57 m) (*Online Supplementary Figure S8*).

CRO score variables and TP53 and *NOTCH1* mutations

Data of *TP53* and *NOTCH1* mutations were available in 304 of 478 cases from the training cohort. Therefore, a

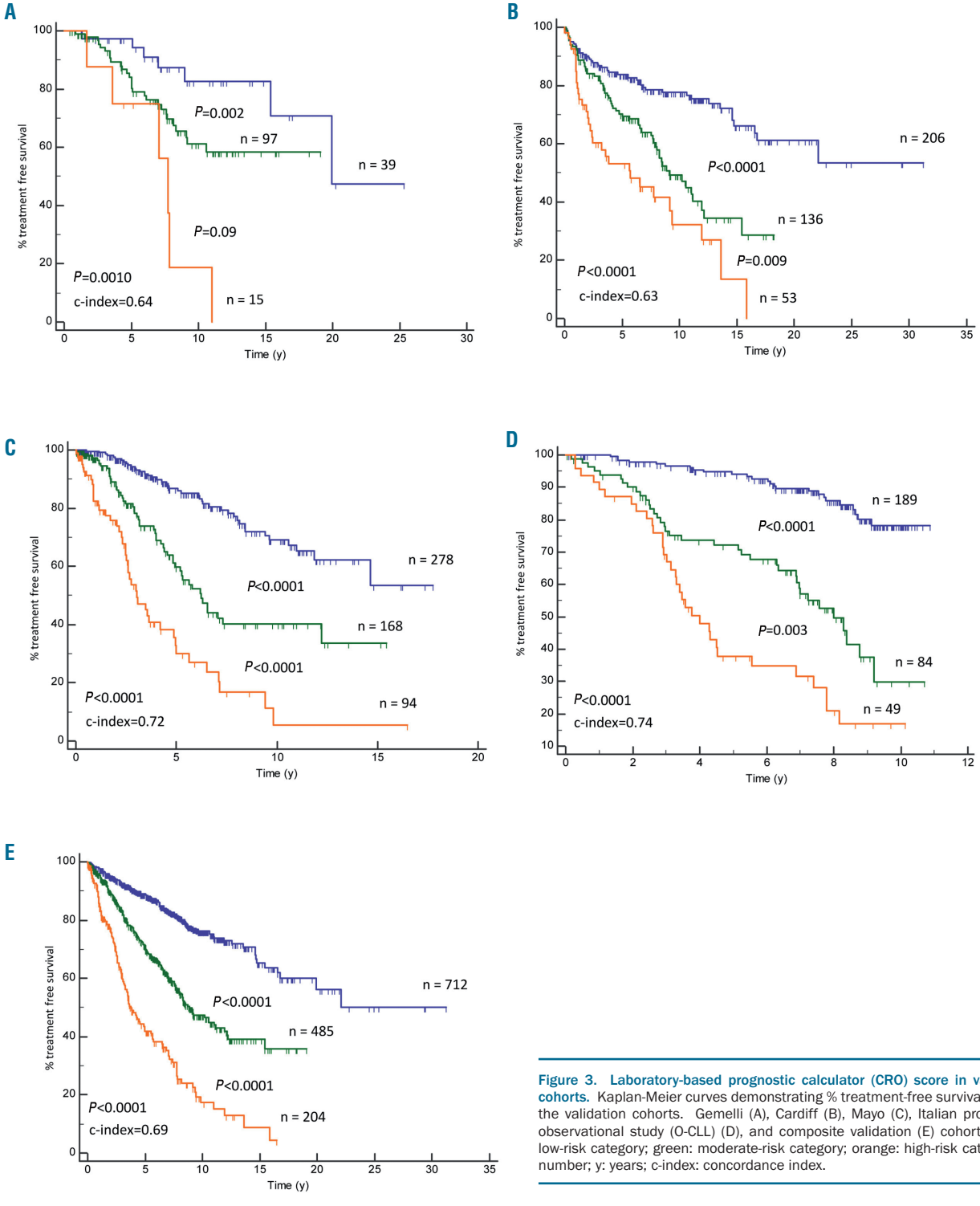

Figure 3. Laboratory-based prognostic calculator (CRO) score in validation cohorts. Kaplan-Meier curves demonstrating % treatment-free survival (TFS) in the validation cohorts. Gemelli (A), Cardiff (B), Mayo (C), Italian prospective observational study (O-CLL) (D), and composite validation (E) cohorts. Blue: low-risk category; green: moderate-risk category; orange: high-risk category. n: number; y: years; c-index: concordance index.

multivariable analysis was performed in these cases by including the same variables (WBC count, del11q, tris12, IGHV gene status, CD49d, age and gender), adding *NOTCH1* and *TP53* mutations and re-classifying del17p and/or *TP53* mutated cases as TP53 disrupted.[5,29-31] As shown in *Online Supplementary Table S3*, the CRO score variables WBC count, del11q, tris12, and IGHV gene status maintained the ability to independently predict short TFS along with TP53 disruption, in a model that included *TP53* and *NOTCH1* gene mutations.

Discussion

The clinical staging systems for CLL, described by Rai and Binet approximately 40 years ago, are still used in clinical practice today to inform prognosis and guide treatment decisions.[1, 2] However, their predictive powers are limited.[3] For example, Pflug *et al.* reported C-indices of 0.56 and 0.58 for Rai and Binet systems, respectively, when applied to a cohort of 1,948 patients.[13] Novel model scoring systems developed in recent years have significantly improved the accuracy of prognostication by incorporating new biomarkers and hold the potential for the development of more individualized treatment strategies,[11-14, 26] especially in the age of an increasing sophistication of novel agents alone or in combination. In this regard, the CLL-IPI was developed as an integrative tool to evaluate overall survival for all clinical stages of disease, and although it demonstrated consistency in subgroup analyses circumscribed to early-stage disease,[15,32] it was not developed specifically to predict TFS.

To our knowledge, this study represents one of the largest attempts to integrate novel biomarkers with traditional clinical factors, with the specific aim of predicting TFS in the setting of Rai 0 CLL. Given the multiplicity of new biomarkers, our goals were: i) to determine which ones individually influence TFS; and ii) to develop a scoring system to stratify risk in patients traditionally thought to harbor indolent disease.

Our training cohort was selected from a consecutive series of 1,201 CLL cases referred to a single center for immunocytogenetic analyses between 2006 and 2017. With respect to TFS, this cohort was stratified into independent risk groups using Rai staging; however, a satisfactory further sub-stratification of Rai 0 cases was not demonstrated using the canonical Dohner's hierarchical classification alone.[7] This observation provided the stimulus to investigate the potential prognostic significance of additional biomarkers in early-stage disease. One of our main strategies was to integrate known prognostic markers that are commonly used in clinical practice today to increase the accessibility and cost-effectiveness of the risk tool.

Our results demonstrate that the CRO score is a powerful tool for guiding treatment prediction in patients with Rai 0 CLL. Notably, a subset analysis of the so-called high-risk category according to the CRO score (i.e. scores 3-5), revealed that a very small subset of cases (14 cases) in a composite cohort of 1,879 training and validation cases, characterized by high WBC counts in the setting of del17p and UM IGHV gene status (i.e. CRO score 5), progressed within two years, significantly more rapidly than the other so-called high-risk cases with CRO scores of 3 or 4. Conversely, in low-risk patients, the CRO score predicted TFS at 10 years of approximately 85%, arguing for its expanded utility in allowing clinicians to confidently provide reassurance of disease quiescence to such patients. Furthermore, in comparison to the CLL-IPI,[11] our model demonstrated superior performance in the training cohort and in 3 out of the 4 validation cohorts, lending credence to its role in the current compendium of comprehensive risk tools in the setting of Rai 0 CLL.

We observed significant heterogeneity in patient characteristics among the five cohorts included in our study (*Online Supplementary Table S1*). For example, 64% of patients in the Gemelli cohort were aged ≥65 years compared to 30% in the O-CLL cohort, and only one patient in the O-CLL had a B2M >3.5 mg, compared to 23% of the patients in the Cardiff cohort. We attribute these differences to the heterogeneity of clinical settings from which each cohort was derived, as has been observed in previous studies comparing 'real world' *versus* observational study patients, single *versus* multicenter registries, and cases from community *versus* tertiary/referral centers.[33-35] In contrast to the results in these studies, which show inconsistent performance of several prognostic indexes across dissimilar cohorts, our scoring system retained powerful predictive capacity throughout, showcasing its generalizability and strength as a clinically useful decision-making tool.

In the training, composite and two out of the four validation cohorts, the proposed prognostic score approached or exceeded C-index values of 0.7, a threshold necessary to confer utility at the individual patient level.[36] In this regard, however, a more precise evaluation of the individual predictive potential may require the application of complex statistical methods, as recently proposed.[37,38]

This study raises questions regarding the appropriate timing of immunocytogenetic analysis in early-stage disease, which today is often postponed until the time of disease progression and first treatment. We appreciate the cost-effectiveness of a 'watch-and-wait' approach, particularly since studies investigating the early use of chlorambucil and fludarabine monotherapy as well as FCR regimens (fludarabine/cyclophosphamide + rituximab) have failed to demonstrate improved outcomes in CLL patients.[18,39,40] However, the role of the novel inhibitors in this setting remains to be elucidated,[41] and the results of this study support the notion that early testing can aid risk-adapted treatment strategies and early intervention, particularly in the modern chemo-free era. In this regard, the CLL12 trial (a phase III clinical study currently underway in Germany) is evaluating the efficacy and safety of ibrutinib compared to a 'watch-and-wait' approach in Binet A CLL using a similar comprehensive scoring system to identify high-risk patients.[36] Another randomized phase II study currently underway at the Mayo Clinic is comparing the efficacy of the BTK inhibitor acalabrutinib alone and in combination with the anti-CD20 obinutuzumab in treating patients with early-stage CLL who are classified as high- or very high-risk according to the CLL-IPI (Sameer Parikh *et al.*, 2018, NCT03516617). Further clinical studies are needed to aid identification of progressive cases of early-stage disease who may benefit from risk-adapted treatment approaches. An important caveat to the approach of up-front testing is that cytogenetic and *TP53* mutational analysis must be repeated at the time of disease progression and/or treatment particularly in previously so-called "TP53 non-disrupted" cases to identify

those that have undergone clonal evolution which could affect treatment decisions.

We recognize that our study has several limitations. For example, we did not include *TP53* gene mutation, an important adverse prognostic factor that, together with del17p, recapitulates the so-called "TP53 disrupted" cases.[31] While the established cut off of 10% for del17p has little biological substantiation, its selection helps mitigate false positive rates. In the era of next generation sequencing, however, *TP53* mutational analysis is admittedly preferred. Despite the exclusion of *TP53* mutational analysis in this study, we were able to achieve superior prognostic power with respect to the CLL-IPI, in keeping with the notion that the majority of del17p cases also bear *TP53* mutations in the undeleted allele,[29,30] and that the clinical impact of subclonal *TP53* mutations,[31] especially if detected alone in early-stage disease, is still not completely understood. Furthermore, subgroup analysis of 304 cases from the training cohort demonstrated the preservation of CRO score variables even in the presence of *TP53* and *NOTCH1* gene mutations (*Online Supplementary Table S3*). We also excluded from our analysis other gene mutations usually associated with disease progression, namely *BIRC3* and *SF3B1*[42-45] although, notably, these mutations have mainly failed to operate as independent predictors when tested in large cohorts.[5,46] Similarly, we did not include in our panel of biomarkers the evaluation of serum thymidine kinase levels, that, according to some studies, is a test with independent clinical relevance as a predictor of overall survival.[13] This assay, however, is of limited application in CLL and is currently not routinely employed in many US or European clinical laboratories.

We have not overlooked the uniqueness of including WBC count as a prognostic biomarker in this study. More commonly, B-cell lymphocyte count is employed in the diagnosis and response to therapy in CLL.[18] However, we demonstrate here that, commensurate with previously published studies,[47] WBC count may deserve consideration as a clinically useful surrogate marker of disease burden particularly in the setting of untreated disease where, alongside del17p and unmutated IGHV gene status, it appears to demonstrate prognostic significance. This observation is consistent with previous data demonstrating WBC count, along with IGHV, as independent predictors of TFS in Binet A CLL.[48]

Finally, CD49d, a well-proven independent prognosticator in CLL,[19,49] including in cases of early-stage disease,[50] failed, perhaps surprisingly, to emerge in multivariable analysis as an independent factor in our Rai 0 training cohort. We hypothesize that the dropout of CD49d in multivariate analysis was due to the close relationship between CD49d and tris12,[51] the latter maintained in the final multivariable model; this idea is supported by bivariate analysis of these two variables wherein CD49d lost significance (*data not shown*).

In conclusion, we present here a novel laboratory-based scoring system for Rai 0 CLL to aid clinical decision making in cases of early-stage disease. These findings advocate a role for immunocytogenetic analysis in Rai 0 CLL at the time of diagnosis to aid prognosis, particularly in today's chemo-free era where early intervention is acquiring momentum in the investigative setting. Further investigation is needed to definitively validate its utility in risk-adapted treatment.

Acknowledgements

The study was supported by the Associazione Italiana Ricerca Cancro (AIRC), Investigator Grants IG-21687; Progetto Ricerca Finalizzata PE 2016-02362756, Ministero della Salute, Rome, Italy; Associazione Italiana contro le Leucemie, Linfomi e Mielomi (AIL), Venezia Section, Pramaggiore Group, Italy; Linfo-check - Bando ricerca - contributo art. 15, comma 2, lett b) LR 17/2014; "5x1000 Intramural Program", Centro di Riferimento Oncologico, Aviano, Italy. SAP is a Scholar in the Mayo Clinic Paul Calabresi Program in Translational Research (K12 CA090628).

References

1. Binet JL, Leporrier M, Dighiero G, et al. A clinical staging system for chronic lymphocytic leukemia: prognostic significance. Cancer. 1977;40(2):855-864.
2. Rai KR, Sawitsky A, Cronkite EP, et al. Clinical staging of chronic lymphocytic leukemia. Blood. 1975;46(2):219-234.
3. Letestu R, Levy V, Eclache V, et al. Prognosis of Binet stage A chronic lymphocytic leukemia patients: the strength of routine parameters. Blood. 2010;116(22):4588-4590.
4. Gribben JG, O'Brien S. Update on therapy of chronic lymphocytic leukemia. J Clin Oncol. 2011;29(5):544-550.
5. Dal Bo M, Bulian P, Bomben R, et al. CD49d prevails over the novel recurrent mutations as independent prognosticator of overall survival in chronic lymphocytic leukemia. Leukemia. 2016 10;30(10):2011-2018.
6. 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.
7. Dohner H, Stilgenbauer S, Benner A, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med. 2000;343(26):1910-1916.
8. Hallek M, Wanders L, Ostwald M, et al. Serum beta(2)-microglobulin and serum thymidine kinase are independent predictors of progression-free survival in chronic lymphocytic leukemia and immunocytoma. Leuk Lymphoma. 1996;22(5-6):439-447.
9. Hamblin TJ, Davis Z, Gardiner A, et al. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood. 1999;94(6):1848-1854.
10. Zenz T, Eichhorst B, Busch R, et al. TP53 mutation and survival in chronic lymphocytic leukemia. J Clin Oncol. 2010;28(29):4473-4479.
11. 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.
12. Bulian P, Rossi D, Forconi F, et al. IGHV gene mutational status and 17p deletion are independent molecular predictors in a comprehensive clinical-biological prognostic model for overall survival prediction in chronic lymphocytic leukemia. J Transl Med. 2012;10:18.
13. 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.
14. 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.
15. Gentile M, Shanafelt TD, Rossi D, et al. Validation of the CLL-IPI and comparison with the MDACC prognostic index in newly diagnosed patients. Blood. 2016;128(16):2093-2095.
16. Molica S, Shanafelt TD, Giannarelli D, et al. The chronic lymphocytic leukemia international prognostic index predicts time to first treatment in early CLL: Independent validation in a prospective cohort of early stage patients. Am J Hematol. 2016;91(11):1090-1095.
17. Matutes E, Owusu-Ankomah K, Morilla R,

et al. The immunological profile of B-cell disorders and proposal of a scoring system for the diagnosis of CLL. Leukemia. 1994;8(10):1640-1645.
18. 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.
19. Bulian P, Shanafelt TD, Fegan C, et al. CD49d is the strongest flow cytometry-based predictor of overall survival in chronic lymphocytic leukemia. J Clin Oncol. 2014;32(9):897-904.
20. Else M, Wade R, Oscier D, et al. The long-term outcome of patients in the LRF CLL4 trial: the effect of salvage treatment and biological markers in those surviving 10 years. Br J Haematol. 2016;172(2):228-237.
21. Oscier D, Wade R, Davis Z, et al. Prognostic factors identified three risk groups in the LRF CLL4 trial, independent of treatment allocation. Haematologica. 2010;95(10):1705-1712.
22. Pospisilova S, Gonzalez D, Malcikova J, et al. ERIC recommendations on TP53 mutation analysis in chronic lymphocytic leukemia. Leukemia. 2012;26(7):1458-1461.
23. D'Agaro T, Bittolo T, Bravin V, et al. NOTCH1 mutational status in chronic lymphocytic leukaemia: clinical relevance of subclonal mutations and mutation types. Br J Haematol. 2018;182(4):597-602.
24. Pozzo F, Bittolo T, Arruga F, et al. NOTCH1 mutations associate with low CD20 level in chronic lymphocytic leukemia: evidence for a NOTCH1 mutation-driven epigenetic dysregulation. Leukemia. 2015;30(1):182-189.
25. Rossi D, Spina V, Bomben R, et al. Association between molecular lesions and specific B-cell receptor subsets in chronic lymphocytic leukemia. Blood. 2013;121(24):4902-4905.
26. 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.
27. Malcikova J, Tausch E, Rossi D, et al. ERIC recommendations for TP53 mutation analysis in chronic lymphocytic leukemia–update on methodological approaches and results interpretation. Leukemia. 2018;32(5):1070-1080.
28. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Ass. 1999;94(446):496-509.
29. Gonzalez D, Martinez P, Wade R, et al. Mutational status of the TP53 gene as a predictor of response and survival in patients with chronic lymphocytic leukemia: results from the LRF CLL4 trial. J Clin Oncol. 2011;29(16):2223-2229.
30. Rossi D, Cerri M, Deambrogi C, et al. The prognostic value of TP53 mutations in chronic lymphocytic leukemia is independent of Del17p13: implications for overall survival and chemorefractoriness. Clin Cancer Res. 2009;15(3):995-1004.
31. Rossi D, Khiabanian H, Spina V, et al. Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. Blood. 2014;123(14):2139-2147.
32. Molica S, Giannarelli D, Levato L, et al. A prognostic algorithm including a modified version of MD Anderson Cancer Center (MDACC) score predicts time to first treatment of patients with clinical monoclonal lymphocytosis (cMBL)/Rai stage 0 chronic lymphocytic leukemia (CLL). Int J Hematol. 2014;100(3):290-295.
33. Bulian P, Tarnani M, Rossi D, et al. Multicentre validation of a prognostic index for overall survival in chronic lymphocytic leukaemia. Hematol Oncol. 2011;29(2):91-99.
34. Shanafelt TD, Jenkins G, Call TG, et al. Validation of a new prognostic index for patients with chronic lymphocytic leukemia. Cancer. 2009;115(2):363-372.
35. Shanafelt TD, Kay NE, Rabe KG, et al. Hematologist/oncologist disease-specific expertise and survival: lessons from chronic lymphocytic leukemia (CLL)/small lymphocytic lymphoma (SLL). Cancer. 2012;118(7):1827-1837.
36. Langerbeins P, Bahlo J, Rhein C, et al. The CLL12 trial protocol: a placebo-controlled double-blind Phase III study of ibrutinib in the treatment of early-stage chronic lymphocytic leukemia patients with risk of early disease progression. Future Oncol. 2015;11(13):1895-1903.
37. Gerstung M, Papaemmanuil E, Martincorena I, et al. Precision oncology for acute myeloid leukemia using a knowledge bank approach. Nat Genet. 2017;49(3):332-340.
38. Grinfeld J, Nangalia J, Baxter EJ, et al. Classification and Personalized Prognosis in Myeloproliferative Neoplasms. N Engl J Med. 2018;379(15):1416-1430.
39. Dighiero G, Maloum K, Desablens B, et al. Chlorambucil in indolent chronic lymphocytic leukemia. French Cooperative Group on Chronic Lymphocytic Leukemia. N Engl J Med. 1998;338(21):1506-1514.
40. 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.
41. Hallek M. On the architecture of translational research designed to control chronic lymphocytic leukemia. Hematology Am Soc Hematol Educ Program. 2018;2018(1):1-8.
42. Dal Bo M, Del Principe MI, Pozzo F, et al. NOTCH1 mutations identify a chronic lymphocytic leukemia patient subset with worse prognosis in the setting of a rituximab-based induction and consolidation treatment. Ann Hematol. 2014;93(10):1765-1774.
43. Fabbri G, Rasi S, Rossi D, et al. Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. J Exp Med. 2011;208(7):1389-1401.
44. Rossi D, Bruscaggin A, Spina V, et al. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. Blood. 2011;118(26):6904-6908.
45. Rossi D, Fangazio M, Rasi S, et al. Disruption of BIRC3 associates with fludarabine chemorefractoriness in TP53 wild-type chronic lymphocytic leukemia. Blood. 2012;119(12):2854-2862.
46. Baliakas P, Hadzidimitriou A, Sutton LA, et al. Recurrent mutations refine prognosis in chronic lymphocytic leukemia. Leukemia. 2015;29(2):329-336.
47. Rossi D, Gaidano G. Lymphocytosis and ibrutinib treatment of CLL. Blood. 2014;123(12):1772.
48. Del Giudice I, Mauro FR, De Propris MS, et al. White blood cell count at diagnosis and immunoglobulin variable region gene mutations are independent predictors of treatment-free survival in young patients with stage A chronic lymphocytic leukemia. Haematologica. 2011;96(4):626.
49. Shanafelt TD, Geyer SM, Bone ND, et al. CD49d expression is an independent predictor of overall survival in patients with chronic lymphocytic leukaemia: a prognostic parameter with therapeutic potential. Br J Haematol. 2008;140(5):537-546.
50. Rossi D, Zucchetto A, Rossi FM, et al. CD49d expression is an independent risk factor of progressive disease in early stage chronic lymphocytic leukemia. Haematologica. 2008;93(10):1575-1579.
51. Zucchetto A, Caldana C, Benedetti D, et al. CD49d is overexpressed by trisomy 12 chronic lymphocytic leukemia cells: evidence for a methylation-dependent regulation mechanism. Blood. 2013;122(19):3317-3321.