**International Prognostic Score for**

**Asymptomatic Early stage Chronic Lymphocytic Leukemia**

**Running title**: CLL IPS-E

Adalgisa Condoluci1,2\*,Lodovico Terzi Di Bergamo1\*, Petra Langerbeins3\*, Manuela Hoechstetter4\*, Carmen D. Herling3, Lorenzo De Paoli5, Julio Delgado6, Kari G. Rabe7, Massimo Gentile8, Michael Doubek9,10, Francesca R. Mauro11, Giorgia Chiodin12, Mattias Mattsson13, Jasmin Bahlo3, Giovanna Cutrona14, Jana Kotaskova9,10, Clara Deambrogi5, Karin E. Smedby15, Valeria Spina1, Alessio Bruscaggin1, Wei Wu1,16, Riccardo Moia5,Elena Bianchi2, Bernhard Gerber2, Emanuele Zucca2, Michele Ghielmini2, Franco Cavalli2, Georg Stussi2, Mark A. Hess17, Antonino Neri18, Manlio Ferrarini19, Richard Rosenquist15, Francesco Forconi12,20, Robin Foà11, Sarka Pospisilova9,10, Fortunato Morabito21,22, Stephan Stilgenbauer23, Hartmut Döhner23, Sameer A. Parikh7†, William G. Wierda17†, Emili Montserrat6,24†, Gianluca Gaidano5†, Michael Hallek3†, and Davide Rossi1,2†

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

\*A.C., L.T.D.B., P.L. and M.H. equally contributed

†S.A.P., W.G.W., E.M., G.G., M.H. and D.R. equally contributed

Presented at the 15-International Conference on Malignant Lymphoma

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

**Text word count**: 3523; **Abstract word count**: 248; **Figures**: 4; **References**: 57.

**Category**: Lymphoid neoplasia

**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

**Abstract**

**Purpose.** Most patients with chronic lymphocytic leukemia (CLL) are diagnosed with early stage disease and managed with active surveillance. The individual course of subjects 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 CLL patients with early, asymptomatic disease (IPS-E). **Patients and methods.** Individual patient data from 11 international cohorts of patients with early stage CLL (n=4933) were analyzed to build and validate the prognostic score. **Results.** Three covariates were consistently and independently correlated with TTFT: unmutated IGHV genes, absolute lymphocyte count >15 x109/l, and presence of palpable lymph nodes. The IPS-E was the sum of the covariates (one point each), and separated low-risk (score 0), intermediate-risk (score 1) and high-risk patients (score 2-3) 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-years cumulative risk for need of treatment was 8.4%, 28.4%, and 61.2% among low-risk, intermediate-risk, and high-risk patients, respectively. **Conclusions.** The IPS-E is a simple and robust prognostic model that predicts the likelihood of need for therapy in patients with early stage CLL. The IPS-E can be useful in clinical management and in the design of early intervention clinical trials.

**Keywords**: chronic lymphocytic leukemia, Binet stage A, prognosis, IPS-E

**Introduction**

Chronic lymphocytic leukemia (CLL) is characterized by the relentless accumulation of monoclonal B lymphocytes with a distinct immunophenotype (i.e. 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 its 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.3-7 Indeed, different early intervention strategies based on chemo-chemoimmunotherapy failed to show a survival benefit in clinical trials for asymptomatic CLL patients.8-12 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, while others show a stable or a slowly progressive disease not requiring treatment for decades.13 Because of this, the management of patients with early stage CLL is challenging and shaped by uncertainty. Of note, it is estimated that over 400,000 individuals in Europe and US diagnosed with CLL are followed by active surveillance.1,14 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 proved by early intervention with novel agents in patients who are at risk of impending treatment need.15 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 of 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 number, NCT03436524). We obtained the individual patient datasets from 11 cohorts, including 8 series that were previously used to generate or validate scoring systems in CLL (Table 1 and 2).11,16-21 Inclusion criteria were: i) presentation after 1996 confirmed by flow cytometry;3,4 ii) Binet stage A at diagnosis as defined by blood cell count and physical examination;3,4 and iii) active surveillance as initial management, defined by no treatment requirement within the first 3 months after diagnosis. The REMARK criteria were followed throughout the study.22 Patients provided informed consent in accordance with local IRB requirements and Declaration of Helsinki. The ethics committee approved the study (BASEC 2018-00341). Early stage CLL was defined according to the Binet system in 10 out 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 number, NCT00262782) (N=547) and CLL7 (ClinicalTrials.gov number, NCT00275054) (N=339) trials of the German CLL Study Group,11,20 and the O-CLL1 GISL trial (ClinicalTrials.gov number, NCT00917540) (N=312).17,21 The M.D. Anderson Cancer Center (MDACC) (N=1225),16 the University of Eastern Piedmont (UEP) (N=333),19 the Barcelona (N=355),21 the Brno University Hospital (Brno) (N=269),21 the Southampton (N=226), and the Sapienza University (SU) (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 case-control study.18 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 dataset of the training cohort (UEP). Subsequently, we analyzed the external-validation datasets to confirm the findings from the full analysis dataset. The study endpoint was time to first treatment (TTFT), defined as the time between presentation and start of first treatment for 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 one month from initial presentation, were initially considered as covariates for construction of the prognostic index. These covariates were clinical characteristics (age, sex, >1 palpable lymph nodes with a diameter >1 cm and palpable spleen by investigator’s physical examination),6 laboratory values (hemoglobin level, platelet count, absolute lymphocyte count, and beta-2-microglobulin), cytogenetic abnormalities as assessed by FISH [del(17p), del(11q), trisomy 12, del(13q)], and gene mutations (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 translocation were excluded from the study. Continuous variables were categorized by published thresholds (age, hemoglobin, platelet count, beta-2-microglobulin, IGHV identity),20,23-26 or by identifying the best cut off through recursive partitioning (absolute lymphocyte count) (Figure S1 in the Supplementary Appendix). Blood cell count was assessed by local laboratories. Beta-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 was 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 cut off. 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*, *SF3B1* were assessed in the training cohort by targeted deep next generation sequencing using an allele frequency >10% for defining a positive result.27 For the validation datasets, only variables composing the final model were included in the analyses. In the training and validation cohorts, minimal missing data were documented among covariates utilized for development and validation of the prognostic score. Therefore, no missing data imputation procedures were used. All statistical tests were two-sided. Statistical significance was defined as p value <0.05. The analysis was performed with the Statistical Package for the Social Sciences (SPSS) software v.22.0 (Chicago, IL) and with R statisticalpackage v3.4.1 (http://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.28 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.29,30 The adjusted association between exposure variables and TTFT was estimated by Cox regression.31 In the training cohort, Cox regression included exposure variables showing an univariable association with TTFT with a Bonferroni corrected p value <0.05 to account for multiple testing. Backward elimination using likelihood ratio statistics with selection criterion p<0.05 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.32 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.33 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.20 We stratified patients in different risk groups by recursive partitioning of the score.34 Two major steps were utilized to derive the best decision tree: i) growing an initial tree under the following constraints and stopping rules; a) split criteria of p<0.01 according to the log-rank test; b) >20 patients in a node in order to be considered for splitting; c) >10 patients in a terminal node; and ii) applying a pruning algorithm based on the complexity parameter (cp >0.01). Model discrimination was computed using the Harrel’s c-index.35 The 9 Binet A validation cohorts were meta-analyzed to interpolate the validation series in a single plot. TTFT and corresponding 95% confidence intervals were estimated in each of the Binet A validation cohorts at 6 months intervals. At each timepoint, a random-effect model was used to provide a meta-analytic estimate of TTFT and corresponding 95% confidence intervals 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% confidence intervals, 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-years cumulative risk of treatment need, and number of events per 100 person years, and bootstrapping was used to calculate 95% confidence interval. Heterogeneity was assessed by using the I2 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 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 similar to those reported in first line clinical trials (Table S1 in the Supplementary Appendix), thus mitigating the risk of biases due to 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 of treatment need (i.e. TTFT) with a Bonferroni corrected p value <0.05 were palpable lymph nodes, absolute lymphocyte count (>15 x109/l), mild anemia (hemoglobin comprised between 10-13 g/dl in men, and 10-12 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 (Table S2 in the Supplementary Appendix). By multivariable analysis in the training series, four variables independently associated with TTFT were identified, including unmutated IGHV genes, absolute lymphocyte count >15 x109/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 >15 x109/l, palpable lymph nodes, and trisomy 12 were not penalized and therefore not shrunk to zero. Unmutated IGHV genes, absolute lymphocyte count >15 x109/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 >15 x109/l, and palpable lymph nodes maintained the highest percentage of selection in the final model, also after including in the multivariable analyses del(17p) and elevated beta-2-microglobulin, which are part of the CLL-IPI score (Figure 1D).20

After pruning the model from the less consistent covariates, unmutated IGHV genes, absolute lymphocyte count >15 x109/l, and palpable lymph nodes were used to build the score and received a weight of one due to 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 to establish the final international prognostic score to classify patients with asymptomatic CLL in early stage disease (IPS-E) according to the risk of treatment need. Patients with a score of 0 had the longest TTFT, patients with a score of 2 or 3 had the shortest TTFT, while patients with a score of 1 had an intermediate TTFT (Figure S2 in the Supplementary Appendix).

Patients from the training cohort were segregated into three distinct risk categories according to the IPS-E (Figure 2A): low- (score 0), intermediate- (score 1), and high-risk (score 2-3). Sensitivity analysis accounting for death before treatment as competing risk consistently confirmed that IPS-E discriminates three risk groups with significantly different TTFT (Figure S3A in the Supplementary Appendix). 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 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; SU, N=223, median follow-up 11.3 years, number of treatment events 117) vs referral centers (MDACC, 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; Figure S4A-I in the Supplementary Appendix) and by sensitivity analysis accounting for death before treatment as competing risk (Figure S3B-J in the Supplementary Appendix). By meta-analysis across the cohorts, the IPS-E c-index for TTFT was 0.70 (95% C.I.=0.69-0.72, Figure 3) while the c-index of the sole IGHV mutation status was 0.67 (95% C.I.=0.66-0.68, Figure S5 in the Supplementary Appendix).

Approaching the same meta-analytic assessment across the study cohorts, IPS-E showed a slightly higher discrimination capacity of TTFT compared to CLL-IPI (c-index=0.68, 95% C.I.=0.67-0.70, Figure S6 in the Supplementary Appendix).

The training and validation series were meta-analyzed to provide a more precise estimate of the risk of 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 of treatment need (not accounting for death due to competing risk) was <0.1% after 1 year of surveillance, and 8.4% after 5 years. The risk of 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 of treatment need (not accounting for death due to 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 of treatment need (not accounting for death due to competing risk) was 14.1% after 1 year of surveillance, and 61.2% after 5 years. The risk of treatment was 16.1 events per 100 person years among high-risk patients. IPS-E stratified three risk groups over sequential time periods, which is consistent with the lack of updates of treatment initiation criteria in the guidelines3,4,6 and the steadiness of the threshold for commencement of therapy adopted over time by the physicians (Figure S7 in the Supplementary Appendix).

Two different clinical staging systems are used to stratify CLL patients, namely the Binet staging, mostly used in European countries, and the Rai staging, mostly used in the US. 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 three 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 (Figure S3K in the Supplementary Appendix), thus confirming the generalizability of IPS-E also in patients with CLL staged according to Rai system.

**Discussion**

The majority of patients with CLL are diagnosed in asymptomatic, early phases of the disease and are followed with no therapy.2 The clinical evolution of these patients is heterogeneous and difficult to predict. While 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 (i.e. Binet stage A) has not significantly changed over the last decades.2 The management of these patients constitutes a major challenge, and should be made based on 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.20,21,36-41 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 of such models overall survival is the main endpoint.20,21 The uniqueness of CLL patients diagnosed with asymptomatic early disease and the challenges posed by their management make advisable a specific prognostic model with TTFT as a main endpoint.

An important general caveat is that outcome indicators derived from patients requiring therapy do not necessarily apply to patients not requiring intervention. Among CLL patients managed with active surveillance, the IGHV unmutated status is the biomarker with the strongest impact 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.42-44 Patients having *ATM* abnormalities or *NOTCH1* mutations are highly enriched of unmutated IGHV genes.However, a large proportion of IGHV unmutated patients does not harbor neither *ATM* abnormalities nor *NOTCH1* mutation.45-47 Therefore, IGHV unmutated status not only captures *ATM* and *NOTCH1* abnormalities but 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 CLL patient 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 IPS-E high-risk patient, while it was 2.1% and 28.4%, for intermediate-risk patients, and <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 due to the likelihood of requiring treatment sooner.

IPS-E calculation is simple, being the sum of three variables. Compared to other scores, IPS-E calculation needs the assessment of only one molecular variable, namely the IGHV mutation status, whose testing is broadly available and standardized.48 IGHV status has been recognized by the current guidelines as a predictive biomarker for treatment tailoring.5-7 Our results further support the indication for testing IGHV mutations in CLL patients. IGHV status never changes during the course of disease and might as well be evaluated at timepoint of first diagnosis in order to give the patient and the treating physician an estimate on TTFT. The results of our study also support that, as recommended by guidelines, FISH analysis and molecular testing for *TP53* has no clinical utility when done 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 should not be evaluated in early stage asymptomatic patients lacking treatment indication (in line with the iwCLL criteria and CLL guidelines).3-7 CLL can exhibit diverse growth patterns, including a continued stable state, never requiring therapy after diagnosis.49,50 However, a short lymphocyte doubling time is rarely an indication to start therapy by itself, as shown in previous reports51 and in this study, and therefore was as 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 high-risk early stage CLL patients, an issue that has gained momentum due to the availability of effective small molecules with manageable toxicity.52,53

The development of the IPS-E followed an extensively used approach for the generation of prognostic scores in hematology.20,54-57 However, the risk of biases related to the timing of scheduled evaluations and premature censoring cannot be discarded. On the other hand, although the IPS-E has an outcome discrimination capacity that is generally considered as acceptable58 and similar to that of other CLL prognostication scores now used in clinical practice and for clinical trial design,20 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.

**Financial support**: This study was supported by Swiss Cancer League, ID HSR-4660-11-2018, Bern, Switzerland; Research Advisory Board of the Ente Ospedaliero Cantonale, Bellinzona, Switzerland; European Research Council (ERC) Consolidator Grant CLLCLONE ID: 772051; Grant No. 320030\_169670/1 Swiss National Science Foundation, Berne, Switzerland; Fondazione Fidinam, Lugano, Switzerland; Nelia & Amadeo Barletta Foundation, Lausanne, Switzerland; Fond’Action, Lausanne, Switzerland; Translational Research Program, No. 6594-20, The Leukemia & Lymphoma Society, New York; AIRC 5x1000, No. 21198, Associazione Italiana per la Ricerca sul Cancro Foundation, Milan, Italy; PRIN 2017 No. 2015ZMRFEA\_004 , MIUR, Rome, Italy; Italian Ministry of Health 5x1000 funds 2014 and 2016, Compagnia S. Paolo Turin Italy, project 2017.0526; Swedish Cancer Society, Swedish Research Council, Knut and Alice Wallenberg Foundation, Karolinska Institutet, Karolinska University Hospital, and Radiumhemmets Forskningsfonder, Stockholm; MH CR grants No. (NV19-03-00091) and DRO (FNBr, 65269705), MEYS CR under the project CEITEC 2020 (LQ1601).

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

**Conflict-of-interest disclosure**: D.R. received honoraria from Abbvie, AstraZeneca, Gilead, Janssen, Roche, and research grants from Abbvie, Gilead, Janssen, Cellestia. M. Hallek received research support and honoraria from Roche, Gilead, Mundipharma, Janssen, Celgene, Pharmacyclics, Abbvie. G.G. received honoraria from Abbvie, Janssen, AstraZeneca, 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, Abbvie and a research grant from Gilead. C.D.H. received travel support and research funding from Roche. G.C. received a research grant from Gilead (fellowship program 2017).

**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. 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; 11(13):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: v78-84, (suppl 5).
6. 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.
7. Wierda WG, Byrd JC, Abramson JS, et al. NCCN Guidelines: Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma, Version 4.2020. https://www.nccn.org/professionals/physician\_gls/pdf/cll.pdf. Accessed 15 January 2020.
8. 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.
9. 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.
10. 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; 122:524.
11. 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.
12. CLL Trialists' Collaborative Group.Chemotherapeutic options in chronic lymphocytic leukemia: a meta-analysis of the randomized trials. *JNCI*. 1999; 91(10): 861-868.
13. Hallek M, Shanafelt TD, Eichhorst B. Chronic lymphocytic leukaemia. *Lancet*. 2018; 391 (10129):1524-1537.
14. Sant M, Allemani C, Tereanu C, et al. Incidence of hematologic malignancies in Europe by morphologic subtype: results of the HAEMACARE project. *Blood*. 2010; 116(19):3724-3734.
15. 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 [abstract]. *Haematol Oncol*. 2019;S2:38-40.
16. 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.
17. 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.
18. 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.
19. 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.
20. 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.
21. 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.
22. McShane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor marker prognostic studies (REMARK). *J Natl Cancer Inst*. 2005; 97(16):1180-1184.
23. 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.
24. 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.
25. 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.
26. 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.
27. Rossi D, Condoluci A, Spina V, Gaidano G. Methods for Measuring ctDNA in Lymphomas. *Methods Mol Biol*. 2019; 1881:253-265.
28. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958; 53:457-481.
29. 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; 13:92.
30. Dignam JJ, Kocherginsky MN. Choice and interpretation of statistical tests used when competing risks are present. *J Clin Oncol*. 2008; 26(24):4027-4034.
31. Cox DR. Regression models and life tables. *J R Stat Assoc*. 1972; 34:187-220.
32. 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.
33. Schoenfeld D. Partial residuals for the proportional hazard regression model. *Biometrika*. 1982; 69:239-241.
34. 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.
35. Harrell FE Jr. Regression modeling strategies. New York, NY: Springer-Verlag; 2001.
36. Rai KR, Sawitsky A, Cronkite EP, Chanana AD, Levy RN, Pasternack BS. Clinical staging of chronic lymphocytic leukemia. *Blood*. 1975; 46(2):219-234.
37. 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.
38. 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.
39. 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.
40. 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.
41. Montserrat E, Gale RP. Predicting the outcome of patients with chronic lymphocytic leukemia: progress and uncertainty. *Cancer*. 2019; 125(21):3699-3705.
42. 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):212-4.
43. 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):957-64.
44. 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.
45. 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):2122-30.
46. Montserrat E, Baumann T, Delgado J. Present and future of personalized medicine in CLL. *Best Pract Res Clin Haematol*. 2016; 29(1): 100-110.
47. Baliakas P, Hadzidimitriou A, Sutton LA, et al. Recurrent mutations refine prognosis in chronic lymphocytic leukemia. Leukemia. 2015; 29(2):329-36.
48. Ghia P, Stamatopoulos K, Belessi C, et al. ERIC recommendations on IGHV gene mutational status analysis in chronic lymphocytic leukemia. *Leukemia*. 2007; 21(1):1-3.
49. Gruber M, Bozic I, Leshchiner I, et al. Growth dynamics in naturally progressing chronic lymphocytic leukaemia. *Nature*. 570(7762):474-9, 2019.
50. Montserrat E, Sánchez-Bisonó M, Viñolas N, Rozman C. Lymphocyte doubling time in chronic lymphocytic leukaemia: analysis of its prognostic significance. *Br J Haematol* 1986; 62(3): 567-575.
51. 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.
52. Sharma S, Rai KR. Chronic lymphocytic leukemia (CLL) treatment: so many choices, such great options. *Cancer*. 2019; 125(9):1432-1440.
53. Hallek M, Fürstenau M. How to approach CLL in clinical practice. *Hematol Oncol*. 2019;37:38-42, (suppl 1).
54. Project TIN-HsLPF. A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. *N Engl J Med*. 1993; 329(14):987-994.
55. Solal-Céligny P, Roy P, Colombat P, et al. Follicular lymphoma international prognostic index. *Blood*. 2004; 104(5):1258-1265.
56. 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):4555-62.
57. Hoster E, Dreyling M, Klapper W, et al. A new prognostic index (MIPI) for patients with advanced-stage mantle cell lymphoma. *Blood*. 2008; 111(2):558-565.
58. Hosmer DW, Lemeshow S. Applied logistic regression. John Wiley & Sons, New York, 2nd Edition; 2000.

**Table 1. Characteristics of the training cohort**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **UEP** | | |
|  | **N=333** | | |
| **Variable** | **N** | **%** | **Missing** |
| Age >65 | 208 | 62.4 | - |
| Male gender | 178 | 53.4 | - |
| Palpable lymph nodes | 68 | 20.4 | - |
| Palpable spleen | 20 | 6.0 | - |
| Lymphocytes >15 x109/l | 65 | 19.5 | - |
| Hb <13 g/dl (M)/<12 g/dl (F) | 39 | 11.7 | - |
| Platelets <150 x109/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 | - |

UEP, University of Eastern Piedmont; Hb, hemoglobin; B2M, beta-2-microglobulin; FISH, fluorescence in situ hybridization; Del, deletion; IGHV, immunoglobulin heavy variable gene

**Table 2. Characteristics of the validation cohorts**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | |  | | | | | | | | | | | | | | | | | | | | |
|  | **MDACC** | | |  | **CLL1** | | |  | **Barcelona** | | |  | **CLL7** | | |  | **O-CLL1** | | | |  | **Brno** | | |  | **Southampton** | | |  | **SCAN** | | |  | **SU** | | |  | **Mayo Clinic** | | |
|  | **N=1225** | | |  | **N=547** | | |  | **N=355** | | |  | **N=339** | | |  | **N=312** | | | |  | **N=269** | | |  | **N=226** | | |  | **N=223** | | |  | **N=223** | | |  | **N=881** | | |
| **Variable** | **N** | **%** | **Missing** |  | **N** | **%** | **Missing** |  | **N** | **%** | **Missing** |  | **N** | **%** | **Missing** |  | **N** | **%** | **Missing** | |  | **N** | **%** | **Missing** |  | **N** | **%** | **Missing** |  | **N** | **%** | **Missing** |  | **N** | **%** | **Missing** |  | **N** | **%** | **Missing** |
| Age >65 | 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 gender | 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 x109/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 |  |

MDACC, M.D. Anderson Cancer Center; SU, Sapienza University; B2M, beta-2-microglobulin; FISH, fluorescence in situ hybridization; Del, deletion; IGHV, immunoglobulin heavy variable gene

**Figure legends**

**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 time to first treatment by univariable analysis. Solid boxes indicate the hazard ratio, horizontal lines indicate the 95% Confidence Intervals. 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 0.05 significant level. (**B**) Forest plot of the hazard ratio for the 5 covariates assessed for association with time to first treatment by multivariable analysis. Solid boxes indicate the hazard ratio, horizontal lines indicate the 95% Confidence Intervals. 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 0.05 significant 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.

**Figure 2. IPS-E stratified time to first treatment in early stage CLL patients 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 timepoint. The shadow shows the cubic splines fitted on the meta-analytic estimate of the 95% confidence interval of the cumulative proportion of TTFT at each timepoint. (**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.

**Figure 3. Meta-analysis of the estimate of the IPS-E discrimination capacity of time to first treatment 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% Confidence Intervals. The red rhombus represents the overall estimate by the meta-analysis.

**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-years 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 year 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% Confidence Intervals. The red rhombus represents the overall estimate by the meta-analysis.