RESEARCH ARTICLE
DOI: 10.4274/tjh.galenos.2026.16046

Predictive Value of the Zhang Scoring System for Imatinib Failure in Chronic Myeloid Leukemia: A Comparative Analysis with Established Prognostic Models

Kronik Miyeloid Lösemide İmatinib Tedavi Başarısızlığını Öngörmede Zhang Skorlama Sisteminin Değeri: Yerleşik Prognostik Modellerle Karşılaştırmalı Analiz

Hakan Oral, Ümit Yavuz Malkan, Elifcan Aladagğ Karakulak, Haluk Demiroğlu, Hakan Göker, İbrahim Celalettin Haznedaroğlu
Hacettepe University Faculty of Medicine, Department of Internal Medicine, Division of Hematology, Ankara, Türkiye

Hakan Oral, M.D., Hacettepe University Faculty of Medicine, Department of Internal Medicine, Division of Hematology, Ankara, Türkiye
hknorl.06@gmail.com

June 26, 2026
October 20, 2026

Abstract
Objective: The aim of this study was to evaluate the predictive performance of the Zhang score in comparison with other chronic myeloid leukemia (CML) risk scores for predicting imatinib treatment failure.
Materials and Methods: A total of 201 patients with CML treated between January 1, 2000, and January 10, 2025, were included [94 (46.7%) female and 107 (53.3%) male; mean age 47 ± 16 years; median follow-up 107.0 months (minimum-maximum) (6.4-359.6)].
Results: Compared with the Sokal, Hasford, EUTOS, and ELTS scores, the Zhang score demonstrated the highest discriminative ability for predicting treatment failure, with the highest 5-year time-dependent AUC (0.775, 95% CI 0.700–0.849). In competing-risk regression analysis, both the intermediate-risk (subdistribution hazard ratio (SHR)=5.36, 95% CI 2.19–13.10; p<0.001) and high-risk (SHR=11.40, 95% CI 4.49–29.00; p<0.001) Zhang groups were significantly associated with treatment failure compared with the low-risk group.
Conclusion: In conclusion, the Zhang score showed better predictive performance than the other four CML scoring systems for first-line imatinib treatment failure.
Keywords: Leukemia, Chronic Myeloid, Tyrosine Kinase Inhibitors, Treatment Failure, Risk Scores

Öz
Amaç: Bu çalışmanın amacı, imatinib tedavi başarısızlığını öngörmede Zhang skorunun performansını diğer kronik miyeloid lösemi (KML) risk skorları ile karşılaştırmalı olarak değerlendirmektir.

uncorrected proof

Gereç ve Yöntemler: Çalışmaya 1 Ocak 2000 ile 10 Ocak 2025 tarihleri arasında tedavi edilen toplam 201 KML hastası dahil edildi [94 (%46,7) kadın ve 107 (%53,3) erkek; ortalama yaş 47±16 yıl; medyan takip süresi 107,0 ay (minimum-maksimum: 6,4–359,6)].
Bulgular: Sokal, Hasford, EUTOS ve ELTS skorları ile karşılaştırıldığında, Zhang skoru tedavi başarısızlığını öngörmede en yüksek ayırt edici performansı gösterdi ve 5 yıllık zaman bağımlı ROC analizinde en yüksek eğri altında kalan alan (AUC) değerine sahipti (0,775; %95 GA: 0,700–0,849). Yarışan risk regresyon analizinde, düşük risk grubuna kıyasla hem orta risk (subdistribüsyon hazard oranı [SHR]=5,36; %95 GA: 2,19–13,10; p<0,001) hem de yüksek risk (SHR=11,40; %95 GA: 4,49–29,00; p<0,001) Zhang grupları tedavi başarısızlığı ile anlamlı olarak ilişkili bulundu.
Sonuç: Sonuç olarak, Zhang skoru birinci basamak imatinib tedavi başarısızlığını öngörmede diğer dört KML skorlama sistemine göre daha üstün bir performans göstermiştir.
Anahtar Kelimeler: Lösemi, Kronik Myeloid, Tirozin Kinaz İnhibitörleri, Tedavi Başarısızlığı, Risk Skorları

Introduction

Chronic myeloid leukemia (CML) is a clonal hematopoietic stem cell disorder resulting from a BCR: : ABL1 fusion caused by a translocation of the Philadelphia chromosome [1]. CML is classified into three phases according to the World Health Organization (WHO) 2016 criteria. Although the 2022 WHO update suggested removing the accelerated phase concept, both classifications continue to be used clinically [2,3].

With the beginning of tyrosine kinase inhibitors (TKI) in clinical practice in CML patients, their life expectancy has become comparable to that of the background population [4]. Following the implementation of TKIs into clinical practice for patients with CML, previous studies reported a 10-year overall survival rate of 82% and a 10-year progression-free survival rate of 80% among patients treated with imatinib [5]. Although TKIs have a high clinical efficacy in CML, treatment resistance can be seen in 24-25% of cases [6]. Despite the substantial improvement in outcomes achieved with TKI treatment, a subset of patients with CML continues to experience an unfavorable prognosis. To identify these high-risk patient groups at diagnosis, several prognostic scoring systems have been developed, including the Sokal score, Euro (Hasford) score, EUTOS (European Treatment and Outcome Study for CML) score, ELTS (EUTOS Long-Term Survival) score, and the Zhang score. These scoring systems have evolved alongside changes in CML treatment strategies, with newer models developed to address the prognostic needs of the TKI era [7-11]. They are widely used in clinical practice to estimate survival at the time of initial diagnosis and to predict treatment failure and disease progression during TKI treatment. The Zhang score was developed to predict treatment failure in patients with CML receiving TKI treatment [11]. However, its comparative performance against established scores remains limited, particularly in patients receiving first-line imatinib. In this context, our study aimed to evaluate the effectiveness of the Zhang score compared with other established prognostic scoring systems in predicting treatment failure and survival outcomes in patients diagnosed with CML at our center, using clinical, laboratory, genetic, and pathological data.

Materials and Methods

Patients

A total of 380 patients with suspected CML who were admitted to Hacettepe University Adult Hospital between January 1, 2000, and January 10, 2025, were screened for eligibility. Patients without a confirmed CML diagnosis or with an uncertain diagnosis (n=30), those aged <18 years (n=20), and those who had received alternative treatments, including hydroxyurea, interferon, or transplantation, before imatinib initiation (n=30) were excluded. Thus, 300 patients with confirmed CML who received imatinib as upfront first-line treatment were identified. Among these patients, those diagnosed with CML at another center (n=32), those with insufficient follow-up (n=42), those with missing data, and those who underwent TKI switching for reasons not consistent with the ELN 2020 and 2025 criteria (n=25) were excluded. Of these 25 patients, 8 had insufficient cytogenetic data at diagnosis. No imputation method was used for missing data. After applying the eligibility criteria, 201 patients were included in the final analysis **(Figure 1)**. Informed consent was waived due to the retrospective design of the study. Demographic characteristics of all patients, comorbidities were recorded. CML diagnosis was made based on clinical, laboratory, genetic, and pathological findings according to the WHO 2022 [3], European LeukemiaNet (ELN) 2020 [12], and 2025 criteria [13].

Treatment failure and laboratory parameters

uncorrected proof

Treatment failure was defined in accordance with the 2020 and 2025 ELN recommendations and included failure to meet the recommended ELN milestones at 3, 6, or 12 months, loss of response (complete hematologic response, complete cytogenetic response, or major molecular response), progression to accelerated or blast phase, TKI resistance associated with ABL1 mutations, or CML-related death. Failure-free survival (FFS) was defined as the time from initiation of imatinib treatment to the occurrence of a failure event or death related to CML, with censoring at the time of imatinib discontinuation, transition to a second- or third-generation TKI due to adverse events, death unrelated to CML, or last follow-up.

Laboratory parameters of hemoglobin, leukocytes, lymphocytes, monocytes, neutrophils, eosinophils, basophils, platelets, blast percentage of blood, spleen size below the costal margin, and high-risk additional cytogenetic anomaly (ACA) were recorded at diagnosis of CML. Conventional cytogenetic analysis of bone marrow samples was routinely performed at the time of CML diagnosis in our institution. High-risk ACA included +8, a second Ph-chromosome, i(17q), +19, +17, +21, −7/7q-, 11q23, or 3q26.2 aberrations, and complex aberrant karyotypes [13].

CML scores

The formulas of CML scores are shown in appendices 1.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics version 22.0 and R version 4.5.3. Categorical variables were presented as frequencies and percentages. Normality of continuous variables was evaluated using either the Kolmogorov–Smirnov test or the Shapiro–Wilk test, as appropriate. Continuous variables following a normal distribution were summarized as mean ± standard deviation, whereas those not conforming to normality were described using median and interquartile range (IQR; 25th–75th percentiles). For comparisons between two independent groups, the independent samples t-test was applied for normally distributed variables, while the Mann–Whitney U test was used for non-normally distributed data. Categorical variables were compared using the chi-square test. Median follow-up duration and FFS estimates were obtained using the Kaplan–Meier method, and differences between survival curves were assessed with the log-rank test. Kaplan–Meier and Cox regression analyses were used to evaluate failure-free survival, whereas competing-risk analyses were performed to estimate the cumulative incidence of treatment failure while accounting for competing events. For the FFS analysis, transition to a second- or third-generation TKI due to adverse events and CML-unrelated death were treated as censoring events, whereas these events were treated as competing events in the competing-risk analysis of treatment failure. Cumulative incidence functions were compared across risk strata using Gray's test. The relationship between prognostic scoring systems and treatment failure risk was further examined through Fine–Gray subdistribution hazard models, with subdistribution hazard ratios (SHRs) and 95% confidence intervals (CIs) reported. Variables with $p < 0.05$ in univariable analyses were included in the corresponding multivariable models. Prognostic scoring systems were analyzed separately and were not included in multivariable models because of substantial overlap between their component variables and the individual clinical, laboratory, and cytogenetic parameters. In addition, the ability of prognostic models to discriminate treatment failure over time was assessed using time-dependent receiver operating characteristic (ROC) analysis with inverse probability of censoring weighting (IPCW) under competing risk settings. Time-dependent area under the curve (AUC) values were estimated at 1, 2, 3, 5, and 10 years. Cox proportional hazards regression analysis was also conducted to explore associations between prognostic groups and failure-free survival, and hazard ratios (HRs) with 95% CIs were presented. Where appropriate, Bonferroni correction was applied to adjust for multiple testing. A two-sided p value <0.05 was considered statistically significant.

This study was approved by the Hacettepe University Clinical Research Ethics Committee (approval date: March 4, 2025; approval number: SBA 25/090).

Results

The final cohort comprised 201 patients with CML who received imatinib as first-line treatment were included in the study. Ninety-four (46.7%) of patients were female, and 107 (53.3%) were male; the mean age at diagnosis (mean ± SD) was 47±16 years. The median follow-up time (minimum-maximum) was 107.0 (6.4-359.6) months. Three patients died from treatment failure during the first year of treatment, while all remaining patients were observed beyond 12 months. The high-risk additional cytogenetic abnormalities seen at the time of diagnosis with CML were 6.96% (n=14), including +8 (n=4), a second Ph-chromosome (n=4), +19 (n=3), and a complex karyotype (n=3). Treatment failure was observed in 30 (14.93%), 51, 59, 66, and 74 patients at 1, 2, 3, 5, and 10 years, respectively. The reasons for treatment failure at 10 years (n=74) included failure to meet ELN milestones at 3 months (n=6), 6 months (n=16), and 12 months (n=15); loss of TKI response after 12 months (n=30); progression to accelerated or blast phase CML (n=4); and TKI resistance associated with ABL1 mutations (n=3). According to the Sokal, Hasford, Zhang, EUTOS, and ELTS scores, 76 (37.8%), 117

Uncorrected proof

(58.2%), 52 (25.9%), 192 (95.5%), and 124 (61.7%) patients were classified as low risk, respectively. The corresponding proportions for intermediate risk were 75 (37.3%), 65 (32.3%), 122 (60.7%), not applicable, and 51 (25.4%), and for high risk were 50 (24.9%), 19 (9.5%), 27 (13.4%), 9 (4.5%), and 26 (12.9%), respectively.
When all patients with CML were stratified according to 1-year treatment failure status, those with 1-year treatment failure had lower hemoglobin levels [10.7 (9.4–12.4) vs. 12.2 (10.9–13.3); p=0.007], higher leukocyte counts [158.3 (44–241.8) vs. 64 (34.4–136); p=0.01], a higher percentage of basophils [1.65 (0.7–2.8) vs. 0.60 (0.18–2.2); p=0.009], a higher percentage of peripheral blood blasts [4% (2–4) vs. 0% (0–3); p<0.001], and greater spleen size below the costal margin [6 (2–8) vs. 2 (0–6); p=0.001]. There were no significant differences in eosinophil percentage and platelet count between the groups **(Table 1)**.
In time-dependent ROC analyses assessing the predictive performance of CML risk scores for treatment failure varied over time, the AUC for 5-year treatment failure was 0.692 for the Sokal score, 0.667 for the Hasford score, 0.775 for the Zhang score, 0.730 for the EUTOS score, and 0.761 for the ELTS score **(Figure 2)**. Harrell's concordance index (C-index) values were 0.659 for Sokal, 0.622 for Hasford, 0.724 for Zhang, 0.683 for EUTOS, and 0.714 for ELTS. The Zhang score demonstrated the highest discrimination for treatment failure among the evaluated scoring systems.
Overall FFS differed significantly across risk categories of the CML scoring systems. The Zhang score demonstrated median (95% CI) FFS durations of 40.3 (9.4-71.2) months in the high-risk group, whereas the median FFS was not reached in the low- and intermediate-risk groups ($p < 0.001$). 5-year cumulative incidences of imatinib failure in the 3 risk subgroups of Zhang score were 6.7% (6.4-7%), 38.9% (38.5-39.3%), 66.2% (64.3-68.1%) (p<0.001). Similarly, the ELTS score showed the median FFS was not reached in low risk, median FFS values of 27.7 (8.1-47.3) months, 43.1 (0-92.6) months, across the intermediate-, and high-risk groups, respectively ($p < 0.001$). 5-year cumulative incidences of imatinib failure in the 3 risk subgroups of ELTS score were 20.1% (19.8-20.4%), 58.8% (57.8-59.8%), 56.9% (54.7-59.1%) (p<0.001). Post-hoc pairwise comparisons of FFS among risk groups were performed using the log-rank test with Bonferroni correction. Although significant differences among the four CML risk scores were observed across the low-, intermediate-, and high-risk groups after Bonferroni correction, a significant difference between the intermediate- and high-risk groups was identified only for the Zhang score (p<0.001). For the EUTOS score, the median FFS was not reach in the low and high risk group (p: 0.143) **(Table 2) (Figure 3)**.
Competing risk regression analysis showed that the Sokal, Hasford, Zhang, and ELTS scoring systems were significantly associated with treatment failure. The Zhang score demonstrated the strongest association with treatment failure, with significant effects observed in both the intermediate-risk group (SHR=5.36, 95% CI 2.19–13.10, p<0.001) and the high-risk group (SHR=11.40, 95% CI 4.49–29.00, p<0.001). Likewise, the ELTS intermediate- and high-risk groups were significantly associated with treatment failure (SHR=3.68, 95% CI 2.22–6.12, p<0.001 and SHR=3.57, 95% CI 1.95–6.52, p<0.001, respectively). In contrast, the EUTOS score was not significantly associated with treatment failure (SHR=2.11, 95% CI 0.66–6.70, p=0.21).
HR analyses for FFS further supported these findings. The Zhang score showed the highest hazard ratios, with HRs of 5.322 ($p < 0.001$) for intermediate- versus low-risk and 12.031 ($p < 0.001$) for high- versus low-risk. For the ELTS score, HRs were 3.757 ($p < 0.001$) for intermediate- versus low-risk and 3.767 ($p < 0.001$) for high- versus low-risk **(Table 3)**. Univariable and multivariable regression analyses of clinical, laboratory, and cytogenetic parameters associated with treatment failure and FFS are presented in **Appendices 2**.

Discussion

Evolution of CML Prognostic Scores

In CML, the Sokal score was developed during the chemotherapy era to estimate survival, while the Hasford score was subsequently developed with the introduction of interferon-based treatment to improve prognostic stratification and survival prediction [7,8]. With the initiation of imatinib in the treatment of CML, the EUTOS score was developed [9]. Following the availability of newer-generation TKIs, the EUTOS score has been used more specifically to evaluate survival outcomes [10]. In addition to their impact on overall survival, these scoring systems have also been employed in various studies to assess treatment response and progression-free survival in patients with CML.
In 2024, Zhang et al. developed a novel scoring system to predict treatment failure in patients with CML, addressing a critical unmet need in this field [11]. The Zhang score shares common parameters with the Sokal, Hasford, and ELTS scores, including age, spleen size, and peripheral blood blast percentage, while it overlaps with the EUTOS score only in terms of spleen size. Distinctive features of the Zhang score compared with other prognostic models include gender, hemoglobin level, and the presence of ACA.

Uncorrected Proof

Previous studies have demonstrated that hemoglobin levels <10 g/dL and ≤9 g/dL at the time of CML diagnosis are associated with adverse prognostic outcomes [14,15]. Moreover, high-risk ACA, a key component of the Zhang score, has been associated with increased mortality and treatment failure [16,17].

Performance of the Zhang Score and Comparison with Established Scores

In this study, we aimed to evaluate the predictive performance of the Zhang score for TKI treatment failure in comparison with other established CML risk scores. In our cohort, treatment failure at 1 year was observed in 30 patients and did not differ significantly according to gender ($p = 0.229$) or the presence of ACA (10% vs. 6.43%; $p = 0.444$). The lack of significant associations for individual components such as gender and ACA does not necessarily contradict the overall predictive performance of the Zhang score, as its discrimination may reflect the combined contribution and weighting of multiple variables rather than the effect of each component considered separately. However, significant differences were observed between patients with and without treatment failure in hemoglobin levels ($p = 0.007$), leukocyte count ($p = 0.01$), basophil percentage ($p = 0.009$), spleen size ($p = 0.001$), and peripheral blood blast percentage ($p < 0.001$), findings that are consistent with the Zhang score. In our study, eosinophil percentage, a component of the Hasford score, and platelet count, which is included in both the Sokal and Hasford scores, were not significantly associated with treatment failure ($p=0.906$ and $p=0.084$, respectively). The superior performance of the Zhang score may partly reflect the inclusion of baseline anemia and high-risk ACAs, which may capture adverse hematological and clonal features of CML not included in the ELTS and EUTOS scores.

In a previously published cohort study evaluating the Zhang score, the highest AUC for treatment failure was reported as 0.68, whereas the ELTS score demonstrated a comparable AUC at year 3 but lower AUC values at years 1 and 5 [18]. In our analysis of 5-year treatment failure, the Zhang score demonstrated the best predictive performance in ROC analysis (AUC=0.775), outperforming the other scoring systems, in line with previously published data. The ELTS score (AUC= 0.761), and the EUTOS score (AUC= 0.730) ranked after the Zhang score in predicting treatment failure. In contrast, the Sokal (AUC=0.692) and Hasford (AUC=0.667) scores demonstrated lower discriminatory performance for treatment failure than the other CML prognostic scores in our cohort. Previous studies have shown that the EUTOS and ELTS scores perform better than the Sokal and Hasford scores in predicting event-free survival, which is consistent with our findings and further supports the relatively limited predictive value of the Sokal and Hasford scores in this setting [19-21]. EUTOS demonstrated good discrimination when analyzed as a continuous score; however, its clinical utility as a categorical risk stratification tool was limited in our cohort due to the very small proportion of patients with high risk (4.5%), which may have limited statistical power and contributed to the lack of statistical significance and wider confidence intervals.

In the FFS analysis, the high-risk Zhang score group exhibited the shortest median FFS duration, the highest 5-year cumulative incidence of treatment failure, and the highest SHR [40.3 (9.4–71.2) months; 66.2% (95% CI: 64.3–68.1%); SHR= 11.4 (95% CI: 4.49–29); $p < 0.001$]. These findings are consistent with those of the original Zhang study and suggest superior predictive performance compared to other CML risk scores. Zhang et al. developed the predictive model for treatment failure, while Moinard et al. subsequently evaluated its performance in a multicenter European-type cohort [11,18]. In contrast, our study is, to our knowledge, the first in the English literature to compare the Zhang score with four established CML risk scoring systems (Sokal, Hasford, ELTS, and EUTOS) in a real-world cohort receiving first-line imatinib. While previous studies have primarily compared the Zhang score with the Sokal and ELTS scores, our study represents the first analysis in the English literature to evaluate the Zhang score alongside four established scoring systems, including the Hasford and EUTOS scores.

Strengths and Limitations

In our cohort, imatinib was used as the first-line TKI owing to reimbursement policies in our country. The limitations of this study include its retrospective design and the lack of patients receiving non-imatinib TKIs as first-line treatment, and the retrospective application of current ELN criteria across the long study period, during which molecular monitoring practices and treatment targets evolved substantially. Therefore, the applicability of our findings to patients receiving second-generation TKIs or other contemporary first-line therapies remains uncertain. Additionally, formal pairwise statistical comparisons of the time-dependent AUCs were not performed; therefore, differences in AUC values should be interpreted descriptively. The strengths of our study include the long follow-up period, use of real-life data, which enabled the evaluation of treatment failure over a 10-year period, and the fact that this is the first study to compare the Zhang score with the four currently used CML risk scoring systems.

Conclusion

In conclusion, the Zhang score demonstrated superior performance compared with the other four CML scores in predicting treatment failure among patients treated with first-line imatinib. These findings suggest that the Zhang score may be a useful tool in clinical practice for predicting imatinib failure in CML. Nevertheless, further prospective studies are warranted to validate these results.

Uncorrected Proof

Author contributions: All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by all authors. The first draft of the manuscript was written by all authors and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Acknowledgements: N/A

Informed Consent: Informed consent was waived due to the study's retrospective design.

References

1. Apperley JF. Chronic myeloid leukaemia. Lancet. 2015;385(9976):1447-59.
2. Haznedaroglu IC, Kuzu I, Ilhan O. WHO 2016 Definition of Chronic Myeloid Leukemia and Tyrosine Kinase Inhibitors. Turk J Haematol. 2020;37(1):42-7.
3. Khoury JD, Solary E, Abla O, Akkari Y, Alaggio R, Apperley JF, Bejar R, Berti E, Busque L, Chan JKC, Chen W, Chen X, Chng WJ, Choi JK, Colmenero I, Coupland SE, Cross NCP, De Jong D, Elghetany MT, Takahashi E, Emile JF, Ferry J, Fogelstrand L, Fontenay M, Germing U, Gujral S, Haferlach T, Harrison C, Hodge JC, Hu S, Jansen JH, Kanagal-Shamanna R, Kantarjian HM, Kratz CP, Li XQ, Lim MS, Loeb K, Loghavi S, Marcogliese A, Meshinchi S, Michaels P, Naresh KN, Natkunam Y, Nejati R, Ott G, Padron E, Patel KP, Patkar N, Picarsic J, Platzbecker U, Roberts I, Schuh A, Sewell W, Siebert R, Tembhare P, Tyner J, Verstovsek S, Wang W, Wood B, Xiao W, Yeung C, Hochhaus A. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. Leukemia. 2022;36(7):1703-19.
4. Bower H, Bjorkholm M, Dickman PW, Hoglund M, Lambert PC, Andersson TM. Life Expectancy of Patients With Chronic Myeloid Leukemia Approaches the Life Expectancy of the General Population. J Clin Oncol. 2016;34(24):2851-7.
5. Hehlmann R, Lauseker M, Saußele S, Pfirrmann M, Krause S, Kolb HJ, Neubauer A, Hossfeld DK, Nerl C, Gratwohl A, Baerlocher GM, Heim D, Brümmendorf TH, Fabarius A, Haferlach C, Schlegelberger B, Müller MC, Jeromin S, Proetel U, Kohlbrenner K, Voskanyan A, Rinaldetti S, Seifarth W, Spieß B, Balleisen L, Goebeler MC, Hänel M, Ho A, Dengler J, Falge C, Kanz L, Kremers S, Burchert A, Kneba M, Stegelmann F, Köhne CA, Lindemann HW, Waller CF, Pfreundschuh M, Spiekermann K, Berdel WE, Müller L, Edinger M, Mayer J, Beelen DW, Bentz M, Link H, Hertenstein B, Fuchs R, Wernli M, Schlegel F, Schlag R, de Wit M, Trümper L, Hebart H, Hahn M, Thomalla J, Scheid C, Schafhausen P, Verbeek W, Eckart MJ, Gassmann W, Pezzutto A, Schenk M, Brossart P, Geer T, Bildat S, Schäfer E, Hochhaus A, Hasford J. Assessment of imatinib as first-line treatment of chronic myeloid leukemia: 10-year survival results of the randomized CML study IV and impact of non-CML determinants. Leukemia. 2017;31(11):2398-406.
6. Zhang XS, Gale RP, Zhang MJ, Huang XJ, Jiang Q. A predictive scoring system for therapy-failure in persons with chronic myeloid leukemia receiving initial imatinib therapy. Leukemia. 2022;36(5):1336-42.
7. Sokal JE, Cox EB, Baccarani M, Tura S, Gomez GA, Robertson JE, Tso CY, Braun TJ, Clarkson BD, Cervantes F, Rozman C, Italian Cooperative CML Study Group. Prognostic discrimination in "good-risk" chronic granulocytic leukemia. Blood. 1984;63(4):789-99.
8. Hasford J, Pfirrmann M, Hehlmann R, Allan NC, Baccarani M, Kluin-Nelemans JC, Alimena G, Steegmann JL, Ansari H. A new prognostic score for survival of patients with chronic myeloid leukemia treated with interferon alfa. Writing Committee for the Collaborative CML Prognostic Factors Project Group. J Natl Cancer Inst. 1998;90(11):850-8.
9. Hasford J, Baccarani M, Hoffmann V, Guilhot J, Saussele S, Rosti G, Guilhot F, Porkka K, Ossenkoppele G, Lindoerfer D, Simonsson B, Pfirrmann M, Hehlmann R. Predicting complete cytogenetic response and subsequent progression-free survival in 2060 patients with CML on imatinib treatment: the EUTOS score. Blood. 2011;118(3):686-92.
10. Pfirrmann M, Clark RE, Prejzner W, Lauseker M, Baccarani M, Saussele S, Guilhot F, Heibl S, Hehlmann R, Faber E, Turkina A, Ossenkoppele G, Höglund M, Zaritskey A, Griskevicius L, Olsson-Strömberg U, Everaus H, Koskenvesa P, Labar B, Sacha T, Zackova D, Cervantes F, Colita A, Zupan I, Bogdanovic A, Castagnetti F, Guilhot J, Hasford J, Hochhaus A, Hoffmann VS. The EUTOS long-term survival (ELTS) score is superior to the Sokal score for predicting survival in chronic myeloid leukemia. Leukemia. 2020;34(8):2138-49.

Uncorrected proof

11. Zhang X, Liu B, Huang J, Zhang Y, Xu N, Gale RP, Li W, Liu X, Zhu H, Pan L, Yang Y, Lin H, Du X, Liang R, Chen C, Wang X, Li G, Liu Z, Zhang Y, Liu Z, Hu J, Liu C, Li F, Yang W, Meng L, Han Y, Lin L, Zhao Z, Tu C, Zheng C, Bai Y, Zhou Z, Chen S, Qiu H, Yang L, Sun X, Sun H, Zhou L, Liu Z, Wang D, Guo J, Pang L, Zeng Q, Suo X, Zhang W, Zheng Y, Huang X, Jiang Q. A predictive model for therapy failure in patients with chronic myeloid leukemia receiving tyrosine kinase inhibitor therapy. Blood. 2024;144(18):1951-61.
12. Hochhaus A, Baccarani M, Silver RT, Schiffer C, Apperley JF, Cervantes F, Clark RE, Cortes JE, Deininger MW, Guilhot F, Hjorth-Hansen H, Hughes TP, Janssen JJWM, Kantarjian HM, Kim DW, Larson RA, Lipton JH, Mahon FX, Mayer J, Nicolini F, Niederwieser D, Pane F, Radich JP, Rea D, Richter J, Rosti G, Rousselot P, Saglio G, Saußele S, Soverini S, Steegmann JL, Turkina A, Zaritskey A, Hehlmann R. European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia. Leukemia. 2020;34(4):966-84.
13. Apperley JF, Milojkovic D, Cross NCP, Hjorth-Hansen H, Hochhaus A, Kantarjian H, Lipton JH, Malhotra H, Niederwieser D, Radich J, Rousselot P, Saussele S, Schiffer CA, Silver R, Soverini S, Stenke L, Turkina A, Casado LF, Castagnetti F, Cervantes F, Clark RE, Cortes J, Deininger M, Hughes TP, Janssen J, Jiang Q, Kim DW, Larson RA, Mahon FX, Mauro M, Mayer J, Nicolini FE, Pane F, Rea D, Richter J, Rosti G, Saglio G, Hehlmann R. 2025 European LeukemiaNet recommendations for the management of chronic myeloid leukemia. Leukemia. 2025;39(8):1797-813.
14. Ko PS, Yu YB, Liu YC, Wu YT, Hung MH, Gau JP, Liu CJ, Hsiao LT, Chen PM, Chiou TJ, Liu CY, Liu JH. Moderate anemia at diagnosis is an independent prognostic marker of the EUTOS, Sokal, and Hasford scores for survival and treatment response in chronic-phase, chronic myeloid leukemia patients with frontline imatinib. Curr Med Res Opin. 2017;33(10):1737-44.
15. Liu Z, Shi Y, Yan Z, He Z, Ding B, Tao S, Li Y, Yu L, Wang C. Impact of anemia on the outcomes of chronic phase chronic myeloid leukemia in TKI era. Hematology. 2020;25(1):181-5.
16. Hehlmann R, Voskanyan A, Lauseker M, Pfirrmann M, Kalmanti L, Rinaldetti S, Kohlbrenner K, Haferlach C, Schlegelberger B, Fabarius A, Seifarth W, Spieß B, Wuchter P, Krause S, Kolb HJ, Neubauer A, Hossfeld DK, Nerl C, Gratwohl A, Baerlocher GM, Burchert A, Brümmendorf TH, Hasford J, Hochhaus A, Saußele S, Baccarani M; SAKK and the German CML Study Group. High-risk additional chromosomal abnormalities at low blast counts herald death by CML. Leukemia. 2020;34(8):2074-86.
17. Ratajczak B, Przybyłowicz-Chalecka A, Czerwińska-Rybak J, Kanduła Z, Ustaszewski A, Gil L, Lewandowski K, Jarmuż-Szymczak M. The presence of additional cytogenetic aberrations in chronic myeloid leukemia cells at the time of diagnosis or their appearance on tyrosine kinase inhibitor therapy predicts the imatinib treatment failure. Leuk Res. 2023;132:107349.
18. Moinard S, Lebecque B, Lachaise T, Johnson-Ansah H, Doublet C, Roth-Guepin G, Rigal-Huguet F, Roy L, Parry A, Meunier M, Penot A, Legros L, Pante V, Escoffre-Barbe M, Rousselot P, Denis G, Monjanel H, Hamroun D, Khlaifia A, Besombes J, Pereira B, Bourgne C, Berger MG. Evaluating a predictive model of tyrosine kinase inhibitor therapy failure in a European-type cohort: a step towards population-specific tools. Leukemia. 2025;39(10):2375-83.
19. Sato E, Iriyama N, Tokuhira M, Takaku T, Ishikawa M, Nakazato T, Sugimoto KJ, Fujita H, Kimura Y, Fujioka I, Asou N, Komatsu N, Kizaki M, Hatta Y, Kawaguchi T. The EUTOS long-term survival score predicts disease-specific mortality and molecular responses among patients with chronic myeloid leukemia in a practice-based cohort. Cancer Med. 2020;9(23):8931-9.
20. Iriyama N, Hatta Y, Kobayashi S, Uchino Y, Miura K, Kurita D, Kodaira H, Inoue M, Takei M. The European Treatment and Outcome Study score is associated with clinical outcomes and treatment response following European LeukemiaNet 2013 recommendations in chronic-phase chronic myeloid leukemia. Int J Hematol. 2014;100(4):379-85.
21. Uz B, Buyukasik Y, Atay H, Kelkitli E, Turgut M, Bektas O, Eliacik E, Isik A, Aksu S, Goker H, Sayinalp N, Ozcebe OI, Haznedaroglu IC. EUTOS CML prognostic scoring system predicts ELN-based 'event-free survival' better than Euro/Hasford and Sokal systems in CML patients receiving front-line imatinib mesylate. Hematology. 2013;18(5):247-52.

uncorrected proof

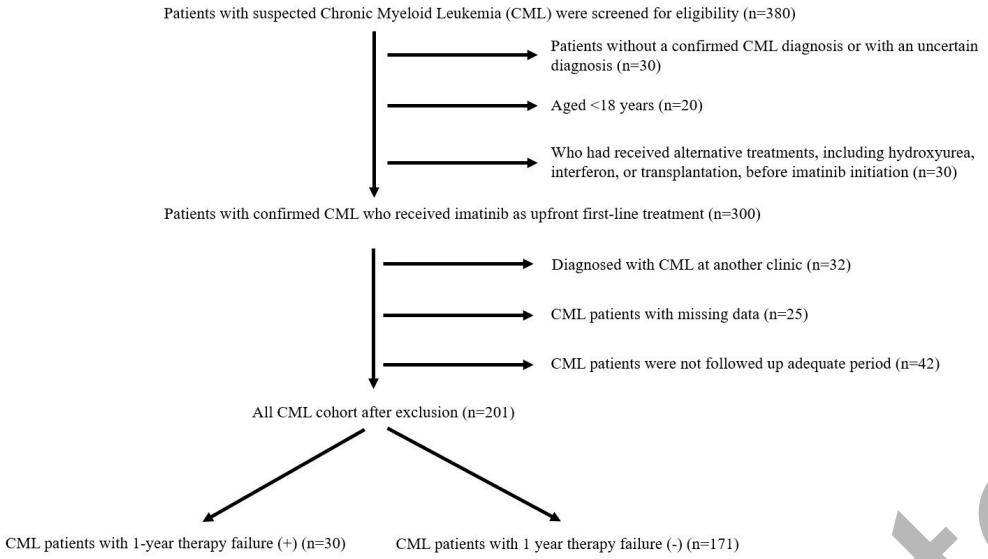


Figure 1. Flow chart of patients included and excluded in the study

uncorrected proof

a

	1-year Theraphy-failure	2-year Theraphy-failure	3-year Theraphy-failure	5-year Theraphy-failure	10-year Theraphy-failure
	AUC (CI 95%)	AUC (CI 95%)	AUC (CI 95%)	AUC (CI 95%)	AUC (CI 95%)
Sokal score	0.676 (0.582–0.770)	0.682 (0.599–0.765)	0.691 (0.610–0.772)	0.692 (0.608–0.776)	0.673 (0.572–0.774)
Hasford score	0.652 (0.542–0.763)	0.635 (0.544–0.726)	0.654 (0.566–0.742)	0.667 (0.579–0.754)	0.654 (0.557–0.751)
Zhang score	0.737 (0.644–0.829)	0.762 (0.688–0.836)	0.770 (0.698–0.842)	0.775 (0.700–0.849)	0.763 (0.676–0.849)
EUTOS score	0.691 (0.588–0.794)	0.705 (0.624–0.787)	0.713 (0.635–0.792)	0.730 (0.652–0.808)	0.766 (0.681–0.850)
ELTS score	0.749 (0.661–0.837)	0.743 (0.667–0.818)	0.752 (0.680–0.825)	0.761 (0.686–0.836)	0.754 (0.667–0.841)

b

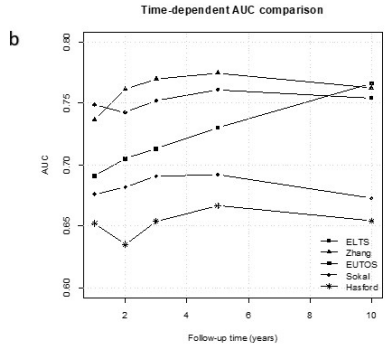


Figure 2. Time-dependent ROC analysis of CML scores, AUC values of CML scores (a), Time dependent AUC comparison (b)

uncorrected proof

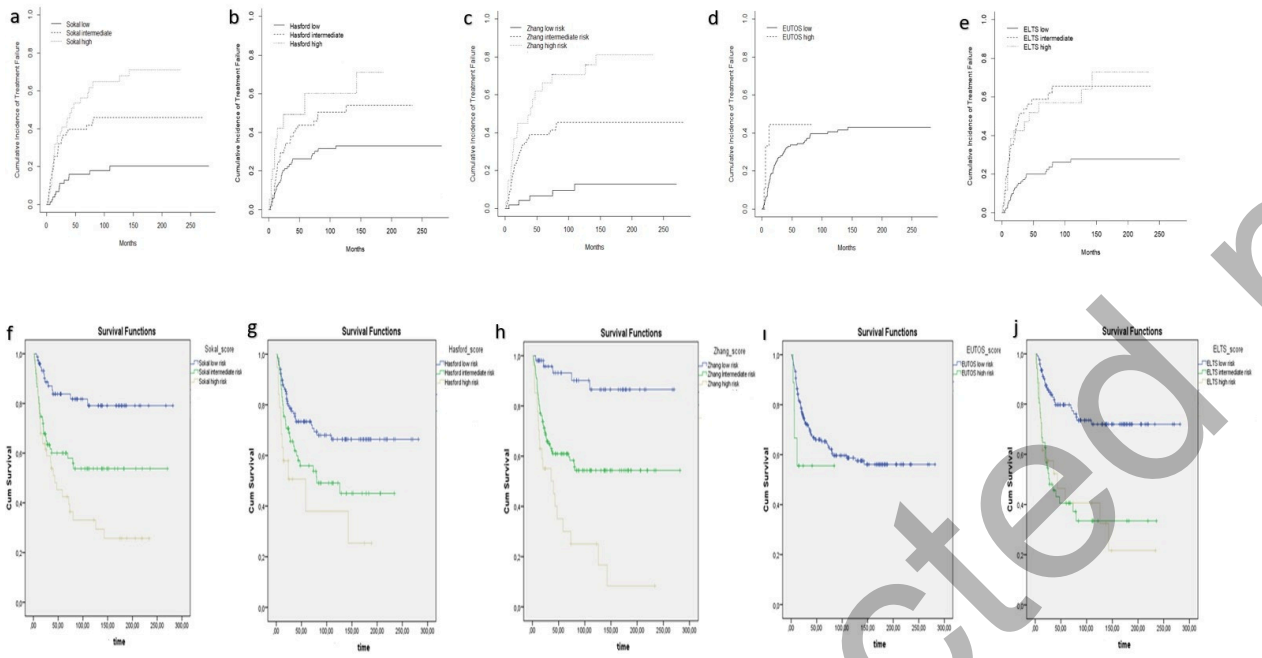


Figure 3. Cumulative incidence of treatment failure of CML scores (a) Sokal score, (b) Hasford score, (c) Zhang score, (d) EUTOS score, (e) ELTS score, Failure free survival of CML scores (f) Sokal score, (g) Hasford score, (h) Zhang score, (ı) EUTOS score, (j) ELTS score

uncorrected proof

Table 1. Demographic Characteristics, Comorbidities and Laboratory parameters of CML Patients With and Without Treatment Failure

Characteristics	Total	1-year Treatment failure (+)	1-year Treatment failure (-)	*p* value
Number of patients	201	30	171	
Age (mean±SD)	47±16	48±16	47±17	0.918
Number of Female	94	11	83	0.229
Number of Male	107	19	88	0.229
Female age (Mean±SD)	49±16	52±19	49±16	0.489
Male age (Mean±SD)	46±17	45±15	46±17	0.802
High risk additional cytogenetic abnormalities	14 (6.96%)	3 (10%)	11 (6.43%)	0.444
Comorbidities				
Diabetes mellitus (%)	30 (14.92%)	3 (10%)	27 (15.79%)	0.412
Coronary artery disease (%)	27 (13.43%)	3 (10%)	24 (14.04%)	0.550
Hypertension (%)	45 (22.38%)	6 (20%)	39 (22.81%)	0.734
Chronic kidney disease (%)	5 (2.48%)	0 (0%)	5 (2.92%)	1.000
Laboratory parameters				
Hemoglobin (g/dL)	12.1 (10.8-13.2)	10.7 (9.4-12.4)	12.2 (10.9-13.3)	**0.007**
Leukocyte (x10³/µL)	67.2 (36.8-154)	158.3 (44-241.8)	64 (34.4-136)	**0.010**
Eozinophil (%)	1.3 (0.5-2.3)	1.4 (0.4-2.6)	1.3 (0.6-2.2)	0.906
Basophil (%)	0.8 (0.2-2.4)	1.65 (0.7-2.8)	0.60 (0.18-2.20)	**0.009**
Platelet (x10³/µL)	402 (259-621)	367 (233-446)	405 (259-646)	0.084
Blood blasts, %, median (IQR)	1 (0-3)	4 (2-4)	0 (0-3)	**<0.001**
Spleen size, cm below costal margin, median (IQR)	3 (0-7)	6 (2-8)	2 (0-6)	**0.001**
High risk additional cytogenetic abnormalities				
Choromosome +8	4 (1.99%)	1 (3.33%)	3 (1.75%)	0.479
Second Ph-chromosome	4 (1.99%)	1 (3.33%)	3 (1.75%)	0.479
Choromosome +19	3 (1.49%)	1 (3.33%)	2 (1.17%)	0.386
Complex karyotype	3 (1.49%)	0 (0%)	3 (1.75%)	1.000

uncorrected proof

Table 2. Failure-Free Survival and Cumulative Incidence of Treatment Failure According to CML Risk Scores

	Median (95% CI) (months)	*p* value	1-year Treatment failure	2-year Treatment failure	3-year Treatment failure	5-year Treatment failure	10-year Treatment failure
Sokal low risk	NR*		4% (3.9-4.1)	11.3% (11-11.6)	12.9% (12.6-13.2)	15.9% (15.5-16.3)	20.4% (19.9-20.9)
Sokal intermediate risk	NR*		20% (19.6-20.4)	32.1% (31.5-32.7)	38.1% (37.5-38.7)	39.7% (39.0-40.4)	45.9% (45.1-46.7)
Sokal high risk	43.1 (17.5-68.7)	**<0.001**	24% (23.3-24.7)	38.5% (37.5-39.5)	45.6% (39.5-44.5)	56.1% (54.9-57.3)	64.9% (63.7-66.1)
Sokal low vs. intermediate risk		**<0.001**					
Sokal low vs. high risk		**<0.001**					
Sokal intermediate vs. high risk		0.059					
Hasford low risk	NR*		10.3% (10.1-10.5)	20.4% (20.1-20.7)	23.3% (23-23.6)	26.3% (26-26.6)	33.1% (32.7-33.6)
Hasford intermediate risk	NR*		16.9% (16.5-17.3)	29.3% (28.7-29.9)	38% (37.2-38.8)	43.7% (42.9-44.5)	50.4% (49.5-51.3)
Hasford high risk	58.4 (0-120.5)	**0.001**	36.8% (34.2-39.4)	49.3% (46.2-52.4)	49.3% (46.2-52.4)	60.2% (55.9-64.5)	60.2% (55.9-64.5)
Hasford low vs. intermediate risk		**0.015**					
Hasford low risk vs. high risk		**0.001**					
Hasford intermediate vs. high risk		0.125					
Zhang low risk	NR*		1.9% (1.8-2)	4.2% (4-4.4)	4.2% (4-4.4)	6.7% (6.4-7)	12.9% (12.3-13.5)
Zhang intermediate risk	NR*		17.2% (17-17.4)	30.5% (30.1-30.9)	36.9% (36.5-37.3)	38.9% (38.5-39.3)	45.4% (44.9-45.9)
Zhang high risk	40.3 (9.4-71.2)	**<0.001**	29.6% (28.0-31.2)	44.9% (43-46.8)	49.2% (47.2-51.2)	66.2% (64.3-68.1)	70.5% (68.6-72.4)
Zhang low vs. intermediate risk		**<0.001**					
Zhang low risk vs. high risk		**<0.001**					
Zhang intermediate vs. high risk		**0.002**					
EUTOS low risk	NR*		13.6% (13.5-13.7)	25.1% (24.9-25.3)	29.8% (29.6-30)	34.3% (34-34.6)	40.7% (40.4-41)
EUTOS high risk	NR*	0.143	44.4% (38.1-50.7)	44.4% (38.1-50.7)	44.4% (38.1-50.7)	44.4% (38.1-50.7)	
ELTS low risk	NR*		6.5% (6.4-6.6)	14.3% (14.1-14.5)	17.2% (17-17.4)	20.1% (19.8-20.4)	27.7% (27.3-28.1)
ELTS intermediate risk	27.7 (8.1-47.3)		27.5% (26.7-28.3)	45.2% (44.2-46.2)	53.8% (52.8-54.8)	58.8% (57.8-59.8)	65.6% (64.4-66.8)
ELTS high risk	43.1 (0-92.6)	**<0.001**	30.8% (29.1-32.5)	42.6% (42.4-44.9)	47% (44.9-49.1)	56.9% (54.7-59.1)	56.9% (54.7-59.1)
ELTS low vs. intermediate risk		**<0.001**					
ELTS low risk vs. high risk		**<0.001**					
ELTS intermediate vs. high risk		0.990					

*NR: Not reach, †EUTOS: European Treatment and Outcome Study, ‡ ELTS: European Treatment and Outcomes Study Long-Term Survival

uncorrected proof

Table 3. Univariable Fine–Gray competing risk regression and Cox proportional hazards analyses of CML scores for treatment failure and failure-free survival

Overall failure-free survival	Subdistribution hazard ratio (CI95%)	p value	Hazard ratio (CI 95%)	p value
Sokal low risk vs intermediate risk	3.07 (1.62-5.85)	**<0.001**	3.068 (1.609-5.848)	**0.001**
Sokal low risk vs high risk	4.84 (2.59-9.03)	**<0.001**	4.933 (2.576-9.448)	**<0.001**
Hasford low risk vs intermediate risk	1.82 (1.12-2.95)	**0.015**	1.816 (1.115-2.956)	**0.016**
Hasford low risk vs high risk	3.04 (1.52-6.09)	**0.002**	3.146 (1.587-6.237)	**0.001**
Zhang low risk vs intermediate risk	5.36 (2.19-13.1)	**<0.001**	5.322 (2.122-13.348)	**<0.001**
Zhang low risk vs high risk	11.4 (4.49-29)	**<0.001**	12.031 (4.499-32.174)	**<0.001**
EUTOS low risk vs high risk	2.11 (0.66-6.70)	0.210	2.093 (0.761-5.759)	0.153
ELTS low risk vs intermediate risk	3.68 (2.22-6.12)	**<0.001**	3.757 (2.255-6.260)	**<0.001**
ELTS low risk vs high risk	3.57 (1.95-6.52)	**<0.001**	3.767 (2.042-6.950)	**<0.001**

uncorrected proof

Appendices

Appendices 1- Formulas of CML scores

Sokal score = exp (0.0116 x [in years] – 43.4) + 0.0345 x (spleen size [cm below costal margin] – 7.51) + 0.1880 x ((platelet count [in 109/L]/ 700)2 – 0.563) + 0.0887 X (blast [% in peripheral blood] – 2.10))
Hasford score= (0.6666 x age [0 when age <50 years; 1, otherwise] + 0.0420 x spleen size [cm below costal margin] + 0.0584 x blasts [% in peripheral blood] + 0.0413 x eosinophils [% in peripheral blood] + 0.2039 x basophils [0 when basophils [% in peripheral blood] <3; 1, otherwise] + 1.0956 x platelet count [0 when platelet count [in 109 /L] <1,500; 1, otherwise]) x 1000
Zhang score = 0.1919 x Sex (male =1, female=0= + 1.6160 x (Age/100) + 0.3105 x $(Hemoglobin/100)^{-2}$ + 0.1087 x Blood blasts + 0.0671 x Spleen size +0.5461 x High_risk ACAs (Yes =1, No=0)
EUTOS score = 7 x basophil [% in peripheral blood] + 4 x spleen size [cm below costal margin]
ELTS score = 0.0025 x (age in completed years/10)3 + 0.0615 x spleen size [cm below costal margin] + 0.1052 x blasts [% in peripheral blood] + 0.4104 x (platelet count/1000)-0.5

uncorrected proof

Appendices 2- Univariable Fine–Gray competing risk regression and Cox proportional hazards analyses of CML scores for treatment failure and failure-free survival

	Unvivariable SHR* (CI 95%)	p value	Multivariable SHR* (CI 95%)	p value	Univariable HR† (CI 95%)	p value	Multivariable HR† (CI 95%)	p value
Age at diagnosis	0.991 (0.977-1.010)	0.220			0.991 (0.977-1.005)	0.228		
Gender	2.200 (1.340-3.600)	**0.002**	1.668 (0.987-2.820)	0.056	2.266 (1.388-3.699)	**0.001**	1.743 (1.040-2.922)	**0.035**
High risk additional cytogenetic abnormalities	2.270 (1.150-4.480)	**0.018**	2.535 (1.310-4.910)	**0.006**	2.386 (1.184-4.808)	**0.015**	2.749 (1.343-5.628)	**0.006**
Hemoglobin (g/dL)	0.836 (0.757-0.924)	**<0.001**	0.985 (0.861-1.130)	0.830	0.830 (0.746-0.923)	**0.001**	0.973 (0.851-1.112)	0.973
Leukocyte (x10³/µL)	1.000 (1.000-1.010)	**<0.001**	1.002 (1.000-1.000)	0.066	1.004 (1.002-1.006)	**<0.001**	1.002 (1.000-1.005)	0.103
Eozinophil (%)	1.030 (0.908-1.160)	0.690			1.022 (0.900-1.1160)	0.740		
Basophil (%)	1.020 (0.951-1.090)	0.630			1.018 (0.949-1.092)	0.626		
Platelet (x10³/µL)	1.020 (0.951-1.090)	0.630			0.999 (0.999-1.000)	0.090		
Blood blasts (%)	1.040 (1.010-1.080)	**0.008**	1.019 (0.982-1.060)	0.310	1.044 (1.019-1.070)	**<0.001**	1.022 (0.987-1.057)	0.219
Spleen size, cm below costal margin	1.120 (1.080-1.170)	**<0.001**	1.068 (1.011-1.130)	**0.019**	1.122 (1.077-1.168)	**<0.001**	1.061 (1.002-1.124)	**0.044**

*SHR: Subdistribution hazard ratio †HR: Hazard Ratio

uncorrected proof