scientific reports

OPEN

Epidemiology and prognostic nomogram for chronic eosinophilic leukemia: a population-based study using the SEER database

Jinlin Wang[1], Meitong Lin[2] & Fan Wang[3]✉

Chronic Eosinophilic Leukemia (CEL), a rare and intricate hematological disorder characterized by uncontrolled eosinophilic proliferation, presents clinical challenges owing to its infrequency. This study aimed to investigate epidemiology and develop a prognostic nomogram for CEL patients. Utilizing the Surveillance, Epidemiology and End Results database, CEL cases diagnosed between 2001 and 2020 were analyzed for incidence rates, clinical profiles, and survival outcomes. Patients were randomly divided into training and validation cohorts (7:3 ratio). LASSO regression analysis and Cox regression analysis were performed to screen the prognostic factors for overall survival. A nomogram was then constructed and validated to predict the 3- and 5-year overall survival probability of CEL patients by incorporating these factors. The incidence rate of CEL was very low, with an average of 0.033 per 100,000 person-years from 2001 to 2020. The incidence rate significantly increased with age and was higher in males than females. The mean age at diagnosis was 57 years. Prognostic analysis identified advanced age, specific marital statuses, and secondary CEL as independent and adverse predictors of overall survival. To facilitate personalized prognostication, a nomogram was developed incorporating these factors, demonstrating good calibration and discrimination. Risk stratification using the nomogram effectively differentiated patients into low- and high-risk groups. This study enhances our understanding of CEL, offering novel insights into its epidemiology, demographics, and prognostic determinants, while providing a possible prognostication tool for clinical use. However, further research is warranted to elucidate molecular mechanisms and optimize therapeutic strategies for CEL.

Abbreviations

CEL	Chronic eosinophilic leukemia
Chemo	Chemotherapy
CI	Confidence interval
COD	Cause of death
HR	Hazard ration
OS	Overall survival
SEER	Surveillance, epidemiology, and end results

Chronic eosinophilic leukemia (CEL; previously also known as chronic eosinophilic leukemia, not otherwise specified), is an infrequent and challenging subtype of *BCR::ABL* negative myeloproliferative neoplasm (MPN) that primarily targets eosinophils—a crucial component of the body's immune response and allergy regulation[1–3]. Characterized by specific genetic attributes, CEL is set apart by the absence of PDGFRA, PDGFRB, or FGFR1 rearrangements and *PCM1::JAK2*, *ETV6::JAK2*, or *BCR::JAK2* fusion genes[2,4]. This distinctive neoplasm is marked by the persistent proliferation of eosinophilic precursors, resulting in pronounced eosinophilia ($> 1.5 \times 10^9$/L) in

[1]Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, China. [2]Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, China. [3]Department of Hematology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1095 Jiefang Avenue, Wuhan 430030, Hubei, China. ✉email: fanwaang@hust.edu.cn

both peripheral blood and bone marrow[2]. Differentiating CEL from other eosinophilic disorders, often driven by specific genetic alterations or secondary factors, is crucial[5]. While the precise etiology of CEL remains elusive, its diagnosis hinges on the exclusion of alternative causes of eosinophilia. According to the 5th edition of the WHO classification of hematolymphoid tumors, three pivotal factors distinguish CEL from hypereosinophilic syndromes (HES): (a) sustained hypereosinophilia for 4 weeks (a notable change from the 6-month duration requirement in the 4th edition); (b) both the presence of clonality among eosinophils, as indicated by cytogenetic or molecular abnormalities; and (c) bone marrow morphology, such as megakaryocytic or erythroid dysplasia (not required in the 4th edition). Notably, increased blasts ($\geq 2\%$ in peripheral blood or 5–19% in bone marrow) are no longer considered an alternative to demonstrate clonality by the 5th edition of the WHO classification of hematolymphoid tumors[3,5–7]. Additionally, rigorous exclusion of non-myeloid malignancies and myeloid neoplasms with eosinophilia, such as acute myeloid leukemia (AML) with inv(16), chronic myeloid leukemia, myeloproliferative neoplasms, and myelodysplastic syndrome, is imperative[8].

The clinical presentation of CEL is highly variable and contingent on the degree and duration of eosinophilia, the extent of organ involvement, and the risk of progression to AML[9]. Patients may experience a spectrum of symptoms, including fatigue, weakness, weight loss, fever, night sweats, splenomegaly, hepatomegaly, lymphadenopathy, and so on[9,10]. These complications can be life-threatening and require swift recognition and management[10]. Moreover, some patients with CEL may progress to AML, a more aggressive hematological malignancy with a less favorable prognosis[9]. As a result, patients with CEL necessitate regular follow-up and monitoring to assess their disease status and response to therapy.

Managing CEL is a complex endeavor, typically necessitating a multidisciplinary approach[10]. Therapeutic strategies aim to reduce eosinophilia and prevent or reverse organ damage[2]. However, no standardized therapy exists for CEL, and treatment options are often limited in their efficacy[2,10]. The choice of therapy is influenced by the presence of specific genetic abnormalities that may predict a patient's response to certain drugs[9]. For instance, patients with *FIP1L1::PDGFRA* rearrangements or other PDGFRA mutations may benefit from imatinib, a tyrosine kinase inhibitor that targets aberrant signaling pathways[11,12]. Nevertheless, most CEL patients lack these mutations and may exhibit resistance or intolerance to imatinib[12]. For such cases, alternative treatment options encompass corticosteroids, hydroxyurea, interferon-alpha, mepolizumab and chemotherapy[2,13–15]. However, these therapies have variable efficacy, significant toxicity profiles, and may not alter the natural progression of the disease, underscoring the critical need for improved therapeutic strategies[9]. Furthermore, most CEL patients are ineligible for allogeneic stem cell transplantation, a possible curative option, due to their advanced age and significant comorbidities[15]. Participation in clinical trials may offer hope for novel and more effective therapeutic options for this challenging condition.

Data associated with the prognosis of CEL is limited. Prognosis in CEL may dependent on factors such as the degree and duration of eosinophilia, the extent and reversibility of organ damage, the presence of cytogenetic or molecular abnormalities, and the response to therapy[16]. Nevertheless, some patients experience a more aggressive clinical course marked by life-threatening complications or progression to AML[9]. According to a study of a small cohort with 10 patients, the survival durations in CEL vary, ranging from 2.2 months to 10 years (median survival was 22.2 months)[9]. Adverse prognostic markers encompass abnormal karyotypes, increased blast counts, thrombocytopenia, bone marrow fibrosis, atypical megakaryocytes, resistance, or intolerance to imatinib, and transformation to AML[9,16].

Owing to the intricacies of diagnosis and the rarity of CEL, comprehensive investigations among patients meeting the 2016 World Health Organization (WHO) criteria for CEL have been notably limited. The purpose of this study is to investigate the incidence and identify the factors affecting the survival of CEL patients based on a population-based study using the national Cancer Institute's Surveillance, Epidemiology and End Results (SEER) database. Moreover, we endeavor to construct a nomogram to predict the prognosis of patients with CEL.

Materials and methods
Study population and data acquisition
Data for this study were sourced from the Surveillance, Epidemiology, and End Results (SEER) Program (https://seer.cancer.gov/), maintained by the National Cancer Institute (NCI). The data were collected using SEER*Stat software version 8.4.1 (https://seer.cancer.gov/seerstat/, accessed on August 1, 2023). The SEER 17 database [Incidence-SEER Research Data, 17 registries, Nov 2022 Sub (2000–2020)] was utilized to extract data of disease incidence by using the "rate session". Patients diagnosed with CEL between 2001 and 2020 were selected from the "Incidence-SEER Research Plus Data, 17 Registries, Nov 2022 Sub (2000–2020)" database using the "case listing session". Only cases with known age (censored at age 89 years) and malignant behavior were included. The inclusion criteria were as follows: the International Classification of Diseases for Oncology (ICD-O-3) histologic code (9964/3). The exclusion criteria were as follows: (1) the diagnosis confirmation was "unknown"; (2) the patient's survival time was 0 or unknown; (3) age below 20 years old. Ultimately, 487 patients with CEL were included in the final cohort. Ethical approval was not required as SEER data are publicly available and anonymized, precluding patient reidentification. A flowchart depicting the selection process is presented in Fig. 2.

Definition of variables
The analysis encompassed an array of variables, including age, sex, race, marital status, year of diagnosis, primary site, vital status, survival months, cause of death (COD) to site recode, cause-specific death classification, cause of death to site, sequence number, first malignant primary indicator, total number of in situ/malignant tumors for the patient, type of reporting source, diagnostic confirmation, surgery of the primary site, median household income inflation adj. to 2021, rural–urban continuum code, chemotherapy recode, and radiation recode. Age was dichotomized into < 60 years and 60+ years, referencing the age at diagnosis of CEL. Race categories included

African American, White, and Other (comprising "Asian/Pacific Islander," "American Indian/Alaska Native," and "Unknown"). Marital status was classified as married, single, or other (encompassing "divorced," "separated," "widowed," "unmarried or domestic partner," and "unknown"). COD information was derived from the "COD to site recode" field. In the SEER database, CEL-related death was defined as "dead (attributed to this cancer diagnosis)," while CEL-unrelated death was defined as death "dead (attributable to causes other than this cancer diagnosis)." Diagnosis years were categorized into "2001–2005," "2006–2010," "2011–2015," and "2016–2020." To assess the results based on the annual median household income at the county levels, patients were categorized into groups based on their annual household income: < $50,000, $50,000-$75,000, and $75,000+. The classification of residence type was based on the Rural–Urban Continuum Codes. These codes separated metropolitan counties according to the population size of their metropolitan areas, and nonmetropolitan counties by their level of urbanization and proximity to a metropolitan area. Overall survival (OS) time was calculated from the date of diagnosis to death or last follow-up.

Statistical analysis

All statistical analyses were conducted using the R program language (version 4.2.1; R Foundation for Statistical Computing, Vienna, Austria). Patients were randomly allocated to training and validation groups (7:3 ratio), with baseline characteristics compared using Student's t-test and chi-square test for continuous and categorical data, respectively. Prognostic risk factors for CEL were identified through least absolute shrinkage and selection operator (LASSO) regression, followed by univariate and multivariate Cox proportional regression analyses to determine independent prognostic factors. Hazard ratios (HR) and 95% confidence intervals (95% CIs) were reported. The Cox proportional hazards regression model was assessed for proportionality assumptions, with no violations detected. Variables with $P < 0.05$ in the multivariate model were considered significant. Nomograms predicting 3- and 5-year overall survival (OS) probabilities were constructed based on independent prognostic factors. Time-dependent receiver operating characteristic (ROC) curves and corresponding area under the curve (AUC) values assessed discrimination. Calibration curves and decision curve analysis (DCA) nomograms were generated to evaluate performance. Patients from both cohorts were classified into high- and low-risk groups using median nomogram points, with Kaplan–Meier analysis employed for OS estimation and log-rank tests for group comparisons. A two-sided P value < 0.05 indicated statistical significance.

Results

Incidence of CEL

The analysis of data from the SEER database showed that the age-adjusted incidence rate (AIR) of CEL from 2001 to 2020 [age adjusted to the 2000 US Standard Population (19 age groups—Census P25-1130)] was 0.033 per 100,000 person-years (95% CI 0.031–0.036). The annual AIR of CEL was presented in Fig. 1A. Notably, the peak AIR was documented in 2008 with 0.054 per 100,000 person-years (95% CI 0.039–0.073), in comparison with the AIR of 2001–2010 (0.040/100,000 person-years), the incidence rate ratio (IRR) for the 2011–2020 group (AIR 0.028/100,000 person years) was 0.68 (95% CI 0.57–0.81, $P < 0.0001$). Moreover, it is notable that the AIR of CEL increased with age (Fig. 1B). Compared with the AIR of patients < 60 years old (0.024/100,000 person-years), the IRR for the 60+ age group (AIR 0.087/100,000 person years) was 3.65 (95% CI 3.07–4.34, $P < 0.0001$), indicating statistical significance. Investigation into gender differences revealed that the AIR of male (0.042/100,000 person-years) was significantly higher than that of female (0.025/100,000 person-years, IRR 1.66, 95% CI 1.39–1.98, $P < 0.0001$, Fig. 1C, D).

Baseline characteristics of CEL patients

As depicted in Fig. 2, 487 patients were finally identified as CEL in the SEER 17 registry, Nov 2022 Sub (2000–2020) from January 2001 to December 2020. The primary site was bone marrow ($n = 487$, 100%). For all patients in the study cohort, 61.2% were male, 1.6 folds that of female ($n = 189$, 38.8%; Table 1). The average age at diagnosis was 57.0 ± 17.0 years (Range: 20–89), with 53.4% under the age of 60 (< 60) and 46.6% aged 60 or older (60+). Most CEL patients were white (73.9%), African American and Other (including Asian/Pacific Islander and American Indian/Alaska Native) occupied 15.6% and 10.5%, respectively. At diagnosis, 57.3% of patients were married, 24.6% had other marital statuses (such as divorced, separated, widowed, unmarried or domestic partner), and 18.1% were single who had never been married. For all CEL cases in this study, 89.7% were primary CEL and 10.3% were secondary CEL that were secondary to other primary malignancies. For the average household income per year, 11.9% were < $50, 000, 45.4% were $50,000-$75, 000, and 42.7% were above $75, 000. About 41.3% were treated with chemotherapy. At the time of last follow-up, 284 (58.3%) patients were alive; 42 (8.6%) deaths were attributable to CEL, 45 (9.2%) patients died of heart diseases, and an additional 116 (23.8%) patients died due to other causes such as diabetes mellitus, cerebrovascular diseases, septicemia, and so on. For all those variables, there were no statistical difference between the training cohort and validation cohort ($P > 0.05$). The comparison of epidemiologic characteristics was summarized in Table 1.

LASSO regression and independent prognostic factors selection

A total of 9 clinical parameters were included in the training set. According to the results of LASSO Cox regression analysis, age, sex, marital status at diagnosis, household income and sequence were identified for OS risk factors by using the minimum standard value as the criterion (Fig. 3). The Cox regression model was further used to screen the prognostic factors. All the five variables passed the preliminary proportional hazards assumption test: age ($P = 0.366$), sex ($P = 0.355$), marital status ($P = 0.535$), household income ($P = 0.208$) and sequence ($P = 0.454$). Univariate Cox regression analysis revealed that age, marital status at diagnosis, and sequence were significantly associated with OS (Table 2). In the multivariate Cox analysis of OS, age, marital status at diagnosis,
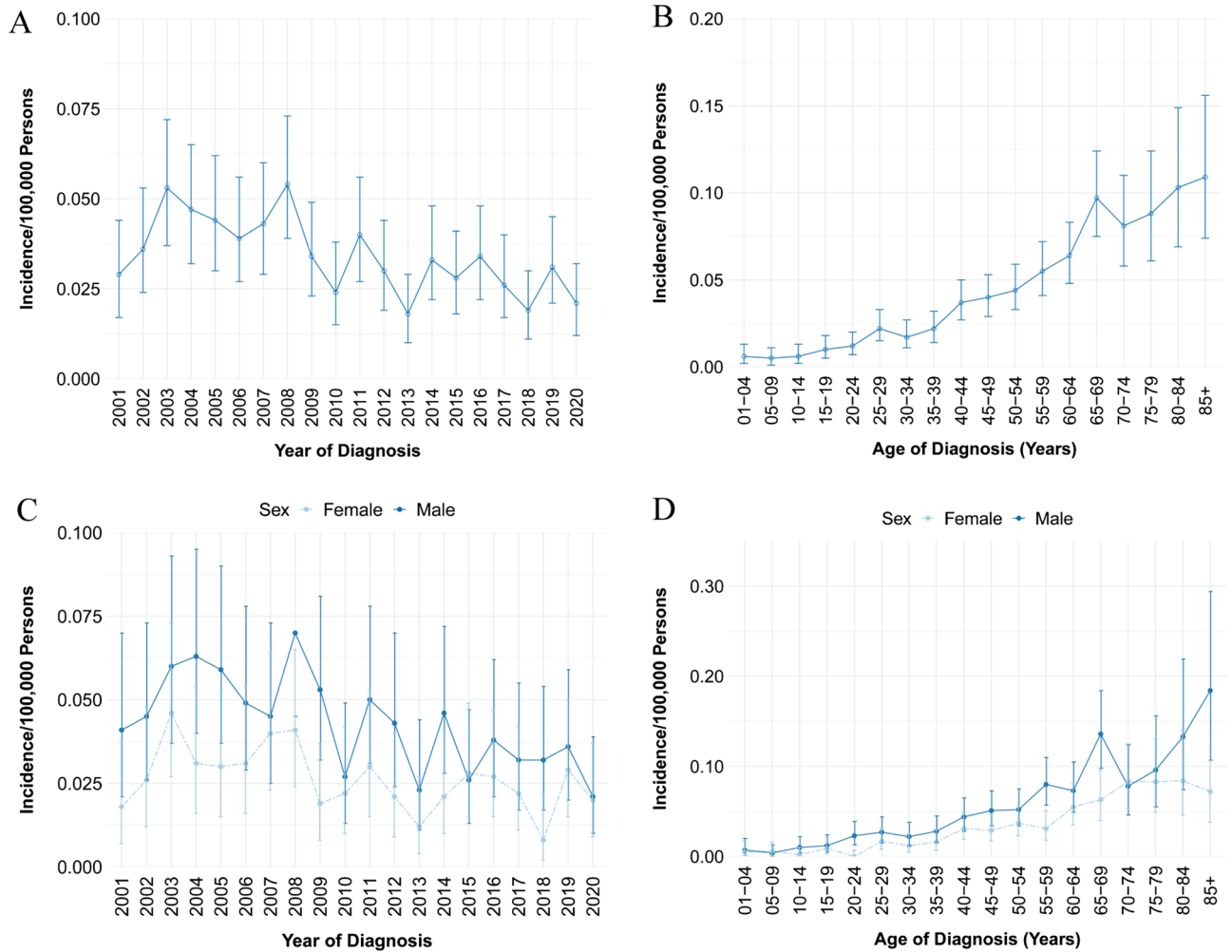
Figure 1. Age-adjusted incidence of CEL from 2001 to 2020 in SEER database. (**A**) Annual age-adjusted incidence of CEL; (**B**) Age-adjusted incidence of CEL based on age of diagnosis; (**C**) Annual age-adjusted incidence of CEL in male and female populations, respectively; (**D**) Age-adjusted incidence of CEL based on age of diagnosis in male and female patients, respectively. CEL, chronic eosinophilic leukemia.
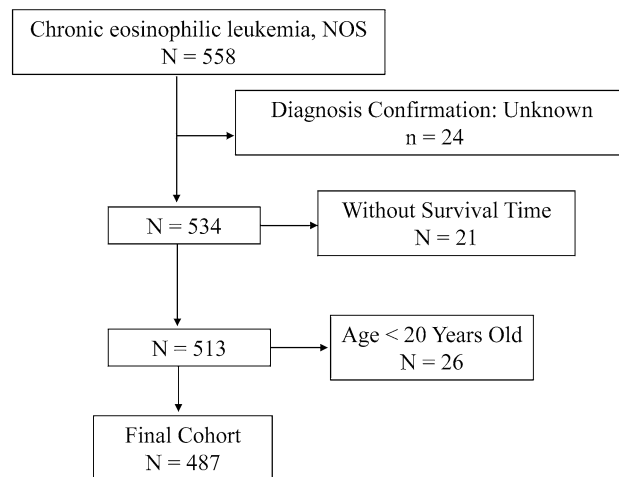

Figure 2. Flow chart of study cohort selection using the SEER database. A flow diagram of selection of patients with CEL in this study. CEL, chronic eosinophilic leukemia; SEER, Surveillance, Epidemiology, and End Results.

Characteristics	Total (N = 487)	Training set (N = 339)	Validation set (N = 148)	P-value
Sex				
Male	298 (61.2%)	205 (60.5%)	93 (62.8%)	0.695
Female	189 (38.8%)	134 (39.5%)	55 (37.2%)	
Age				
< 60	260 (53.4%)	179 (52.8%)	81 (54.7%)	0.769
60 +	227 (46.6%)	160 (47.2%)	67 (45.3%)	
Race				
White	360 (73.9%)	240 (70.8%)	120 (81.1%)	0.053
African American	76 (15.6%)	58 (17.1%)	18 (12.2%)	
Other[1]	51 (10.5%)	41 (12.1%)	10 (6.8%)	
Diagnosis year				
2001–2005	136 (27.9%)	93 (27.4%)	43 (29.1%)	0.762
2006–2010	134 (27.5%)	98 (28.9%)	36 (24.3%)	
2011–2015	113 (23.2%)	78 (23.0%)	35 (23.6%)	
2016–2020	104 (21.4%)	70 (20.6%)	34 (23.0%)	
Sequence				
Primary CEL[2]	437 (89.7%)	305 (90.0%)	132 (89.2%)	0.921
Secondary CEL[3]	50 (10.3%)	34 (10.0%)	16 (10.8%)	
Marital status				
Married	279 (57.3%)	202 (59.6%)	77 (52.0%)	0.251
Single[4]	88 (18.1%)	56 (16.5%)	32 (21.6%)	
Other[5]	120 (24.6%)	81 (23.9%)	39 (26.4%)	
Household income				
< $50,000	58 (11.9%)	39 (11.5%)	19 (12.8%)	0.475
$50,000–$75,000	221 (45.4%)	160 (47.2%)	61 (41.2%)	
$75,000+	208 (42.7%)	140 (41.3%)	68 (45.9%)	
Residence type				
Nonmetropolitan	79 (16.2%)	54 (15.9%)	25 (16.9%)	0.599
Metropolitan < 250,000	41 (8.4%)	29 (8.6%)	12 (8.1%)	
Metropolitan 250,000—1 million	86 (17.7%)	55 (16.2%)	31 (20.9%)	
Metropolitan > 1 million	281 (57.7%)	201 (59.3%)	80 (54.1%)	
Chemo				
Yes	201 (41.3%)	139 (41.0%)	62 (41.9%)	0.934
No/unknown	286 (58.7%)	200 (59.0%)	86 (58.1%)	
COD				
Alive	284 (58.3%)	201 (59.3%)	83 (56.1%)	0.691
CEL[6]	42 (8.6%)	31 (9.1%)	11 (7.4%)	
Heart diseases	45 (9.2%)	31 (9.1%)	14 (9.5%)	
Other cause[7]	116 (23.8%)	76 (22.4%)	40 (27.0%)	

Table 1. Comparison of baseline characteristics of patients with CEL between the training set and validation set from SEER database. *CEL* chronic eosinophilic leukemia, *chemo* chemotherapy, *COD* cause of death. [1]Races including Asian/Pacific Islander, American Indian/Alaska Native and "unknown". [2]Sequence number of "primary only" and "1st of 2 or more primaries", indicating CEL was the primary malignancy. [3]Sequence number of "2nd of 2 or more primaries", "3rd of 3 or more primaries" and "4th of 4 or more primaries", indicating CEL was secondary to other primary malignancies. [4]Marital status at diagnosis was single (never married). [5]Marital statuses of divorced, widowed, separated, unmarried or domestic partner and unknown at diagnosis. [6]Death attributable to CEL. [7]Dead of other causes such as diabetes mellitus, cerebrovascular diseases, septicemia, and so on.

and sequence were independently and significantly associated with OS (Table 2). Older age (HR 3.74, 95% CI 2.51–5.60, $P < 0.001$, Table 2), marital status of single (HR 2.44; 95% CI 1.49–4.00, $P < 0.001$, Table 2), marital status of other (HR 2.08; 95% CI 1.41–3.10, $P < 0.001$, Table 2), secondary CEL (HR 1.98; 95% CI 1.23–3.20, $P = 0.005$, Table 2) were significantly associated with worse overall survival. The detailed data was demonstrated in Table 2.
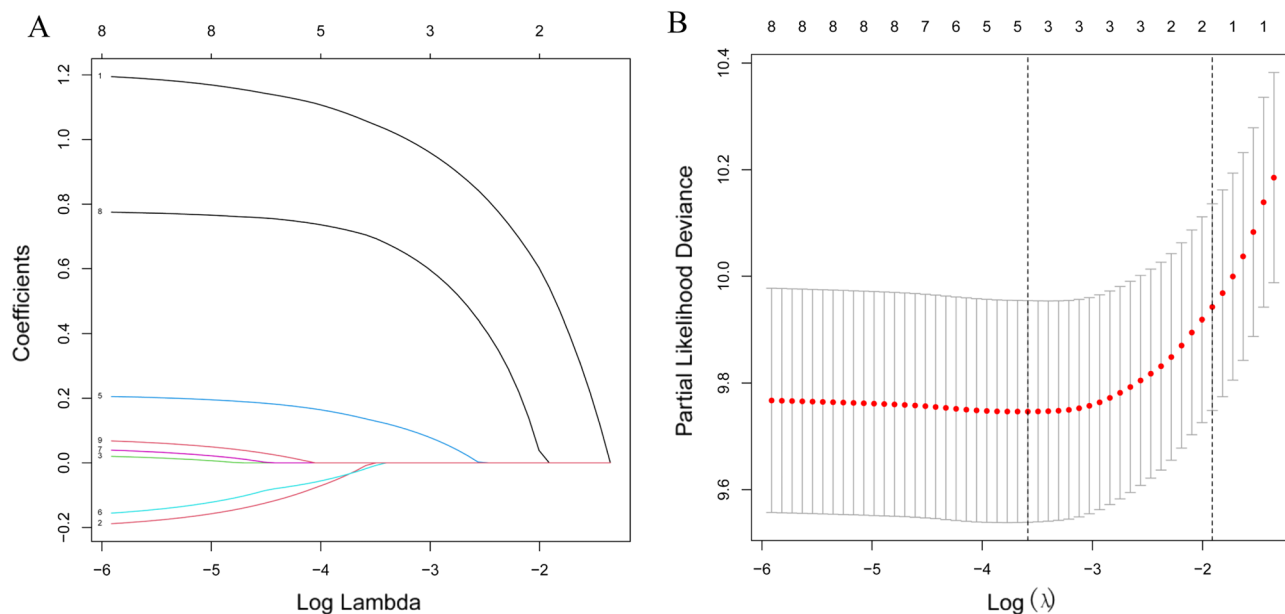
Figure 3. LASSO regression model was used to select characteristic impact factors. (**A**) LASSO coefficients of 7 features; (**B**) Selection of tuning parameter (λ) for LASSO model.

Parameters	Univariable			Multivariable		
	HR	95% CI	*P*-value	HR	95% CI	*P*-value
Age						
< 60	*ref*			*ref*		
60+	3.71	[2.74, 5.02]	< 0.001	3.76	[2.71, 5.20]	< 0.001
Sex						
Male	*ref*			*ref*		
Female	0.98	[0.74, 1.30]	0.868	0.84	[0.62, 1.10]	0.244
Marital status						
Married	*ref*			*ref*		
Single[1]	1.22	[0.83, 1.79]	0.308	2.07	[1.38, 3.10]	< 0.001
Other[2]	1.59	[1.16, 2.19]	0.004	1.49	[1.08, 2.10]	0.016
Household income						
< $50,000	*ref*			*ref*		
$50,000-$75,000	0.84	[0.55, 1.29]	0.423	0.75	[0.49, 1.20]	0.194
$75,000+	0.73	[0.47, 1.13]	0.153	0.71	[0.46, 1.10]	0.135
Sequence						
Primary[3]	ref			ref		
Secondary[4]	3.29	[2.28, 4.74]	< 0.001	2.33	[1.59, 3.40]	< 0.001

Table 2. Univariable and multivariable cox regression analysis for overall survival of patients with CEL. *CEL* chronic eosinophilic leukemia, *Chemo* chemotherapy, *CI* confidence interval, *HR* hazard ration, *OS* overall survival. [1]Marital status of single (never married) at diagnosis. [2]Marital statuses of divorced, widowed, separated, unmarried or domestic partner and unknown at diagnosis. [3]Consisted of "one primary only" and "1st of 2 or more primaries". [4]Consisted of "2nd of 2 or more primaries", "3rd of 3 or more primaries" and "4th of 4 or more primaries", indicating CEL was secondary to other primary malignancies.

Construction of prognostic nomogram

By incorporating the three independent prognostic factors including age, marital status at diagnosis, and sequence, a nomogram was constructed to predict the 3- and 5-year OS probability of CEL patients (Fig. 4). The total points were calculated by integrating scores related to age, marital status, sequence and projected to the bottom scale to predict the OS probability at 3 and 5 years.
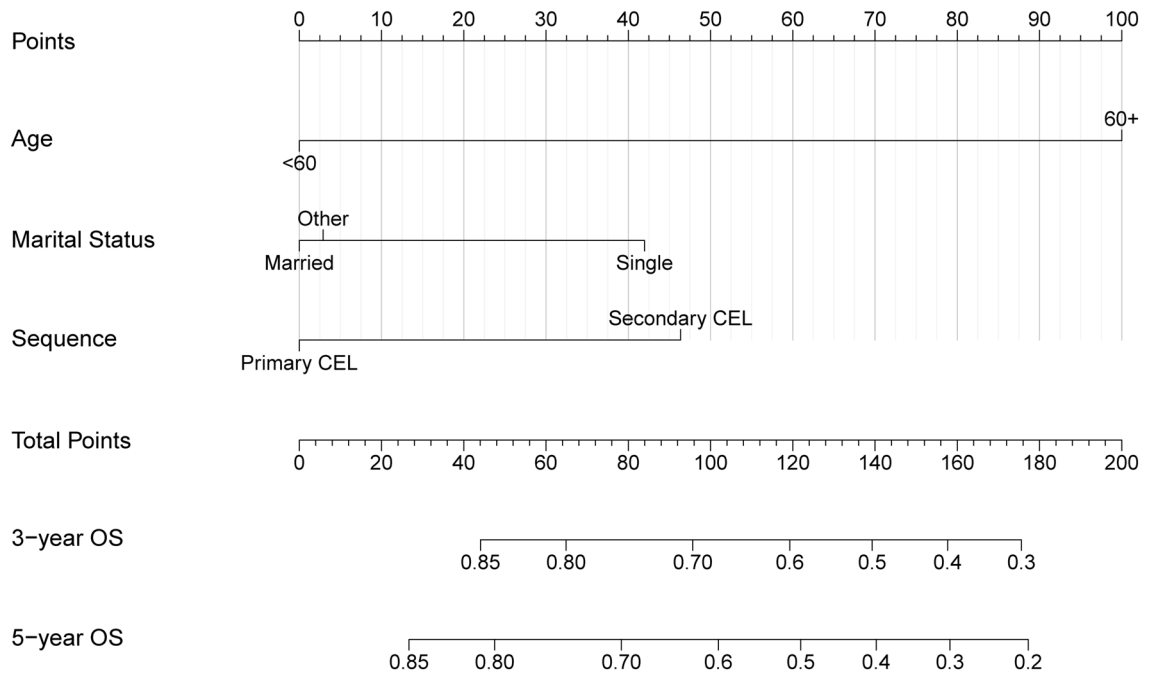
Figure 4. Construction of the prognostic nomogram of CEL patients based on 3 risk factors. The total points were calculated by integrating scores related to age, marital status, sequence and projected to the bottom scale to predict the overall survival probability at 3 and 5 years. CEL, chronic eosinophilic leukemia.

Evaluation and validation of the nomogram

The calibration curve of the nomogram for the training cohort revealed a close match between the predicted and observed OS probability at the 3- and 5-year intervals (Fig. 5A). Additionally, validation cohort calibration plots at 3- and 5 years also showed good agreement between prediction and actual observation (Fig. 5B). Time-dependent ROC analyses showed the accuracy of the nomogram models in predicting 3- and 5-year OS probability in the training set, with AUC values of 0.702 and 0.736, respectively (Fig. 6A), and the 3-year and 5-year AUC of the validation set was of 0.731 and 0.754, respectively (Fig. 6B). The DCA was employed to evaluate the
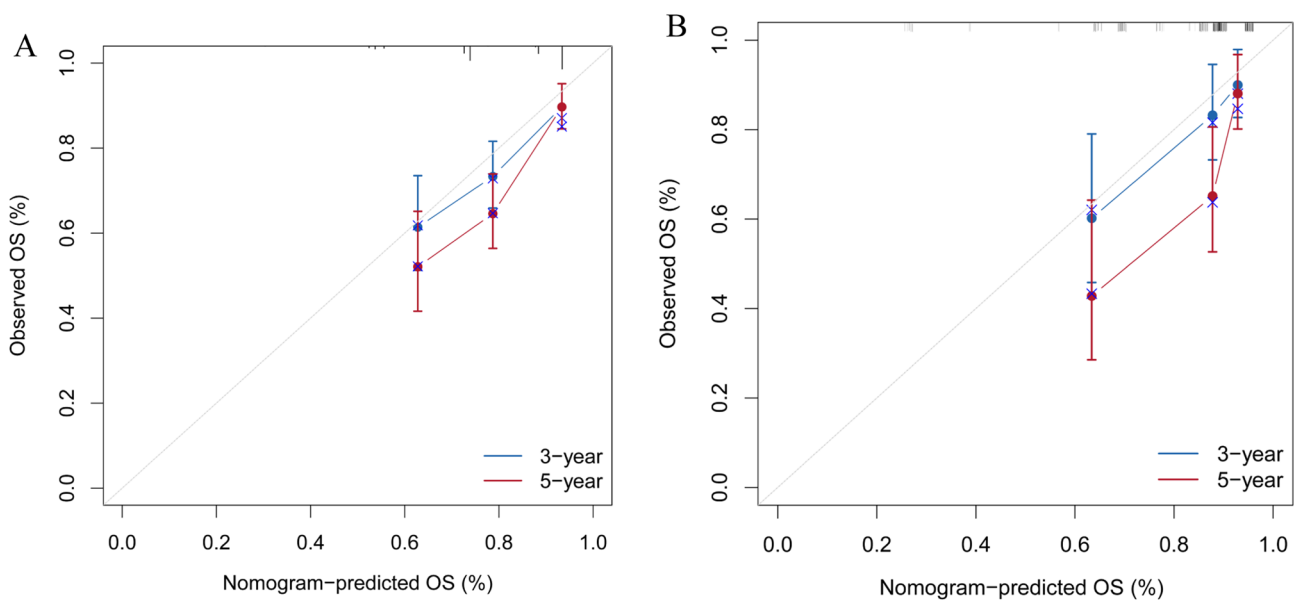
Figure 5. Evaluation of the nomogram of by calibration plot. (**A**) The calibration curve of the training set for the observed overall survival (OS) probability and predicted OS at 3-year and 5-year. (**B**) The calibration curve of the validation set at 3-year and 5-year.
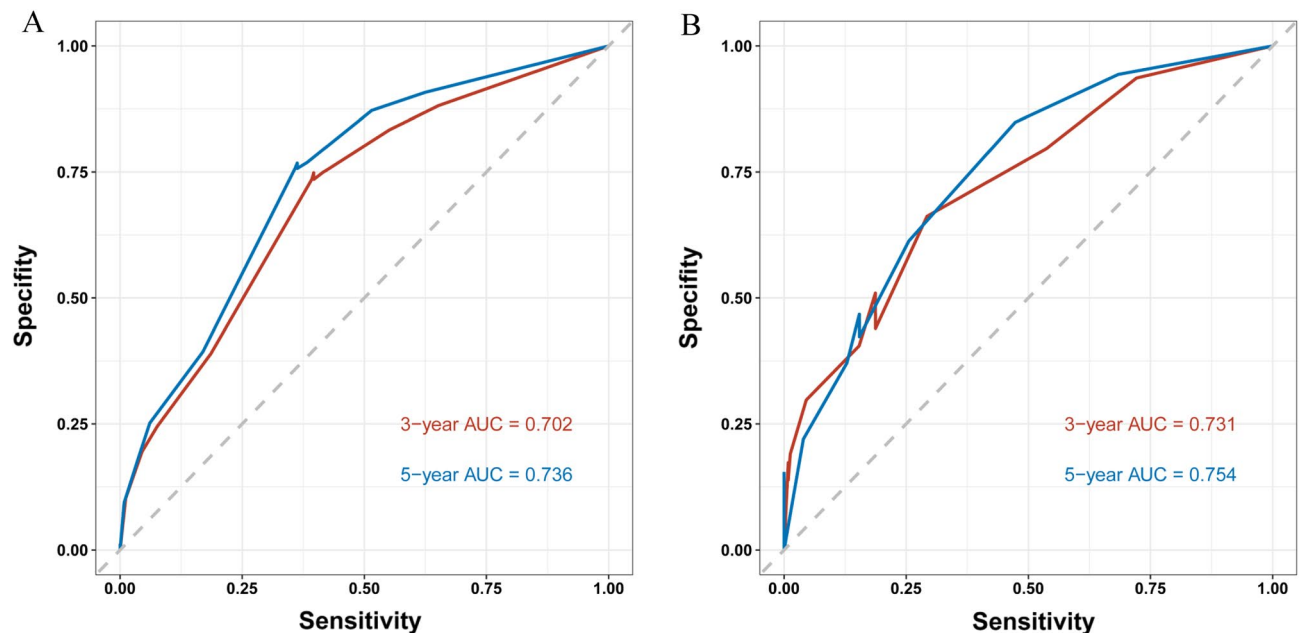
Figure 6. Evaluation of the nomogram of by receiver operating characteristic (ROC) plot. (**A**) Time-dependent ROC curve analyses of the nomograms for the 3 years and 5 years in the training set. (**B**) Time-dependent ROC curve analyses of the validation set.

clinical net benefit of the predictive model. The results showed that the nomogram model has a good net benefit in predicting the 3- and 5-year OS probability both in the training set (Fig. 7A, B) and validation set (Fig. 7C, D).

Survival analysis between the stratified risk groups

The score of each variable was generated from the nomogram and the cumulative scores were calculated for all the patients. The entire cohort was stratified into low- and high- risk subgroups according to the median risk score. Kaplan–Meier analysis of OS revealed significant differences between the low- and high -risk groups for both training set (*P* < 0.0001, Fig. 8A) and validation set (*P* < 0.0001, Fig. 8B), which underscores the exceptional capacity of the nomogram for effective risk stratification.

Discussion

CEL represents a rare and intricate hematological disorder characterized by uncontrolled eosinophilic proliferation[10]. Given its rarity, there is limited published literature on CEL, mostly comprising case reports or small case series, the incidence and clinical characteristics have not been comprehensively studied yet[9,16]. The present study identified 487 CEL patients from 2001 to 2020 using the SEER database, representing the largest cohort describing the incidence and clinical characteristics of CEL patients to date. We also developed and validated a nomogram to predict the 3- and 5-year overall survival probability of CEL patients based on the screened prognostic factors.

The epidemiology of CEL remains incompletely characterized due to its rarity and clinical heterogeneity[1,2]. Available studies suggest an estimated incidence rate of approximately 0.036 per 100,000 individuals for all hypereosinophilic syndromes (HES), including CEL[17]. In this study, analysis of CEL from the SEER database between 2001 and 2020 revealed a low age-adjusted incidence rate (AIR) of 0.033 per 100,000 person-years, and the incidence rate significantly decreased in 2011–2020 compared to 2001–2010. The decreased incidence rate of CEL may be partially attributed to modifications in its diagnostic criteria. In the 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia, a specific group of patients with eosinophilia and gene rearrangements involving PDGFRA, PDGFRB, or FGFR1 were excluded from the diagnosis of CEL and classified as a separate entity, namely "myeloid and lymphoid neoplasms with eosinophilia and abnormalities of PDGFRA, PDGFRB, or FGFR1 (MLNE)", which have distinct clinical and molecular features and respond well to tyrosine kinase inhibitors[18]. Therefore, this reclassification may have reduced the number of cases that were previously diagnosed as CEL based on the older criteria. And the peak incidence in 2008 observed in this study, followed by a decline, coincides with the implementation of the revised WHO classification, suggesting that the initial higher rates may have included cases that would later be reclassified under the new criteria, which supports the hypothesis that the redefinition of CEL has had a significant impact on its reported incidence. Furthermore, the advancement in diagnostic technologies and understanding of molecular genetics over the past two decades has allowed for more accurate diagnosis of myeloproliferative neoplasms (MPNs). This progress may have resulted in more cases being classified into other specific subtypes of MPNs rather than being broadly categorized as CEL.
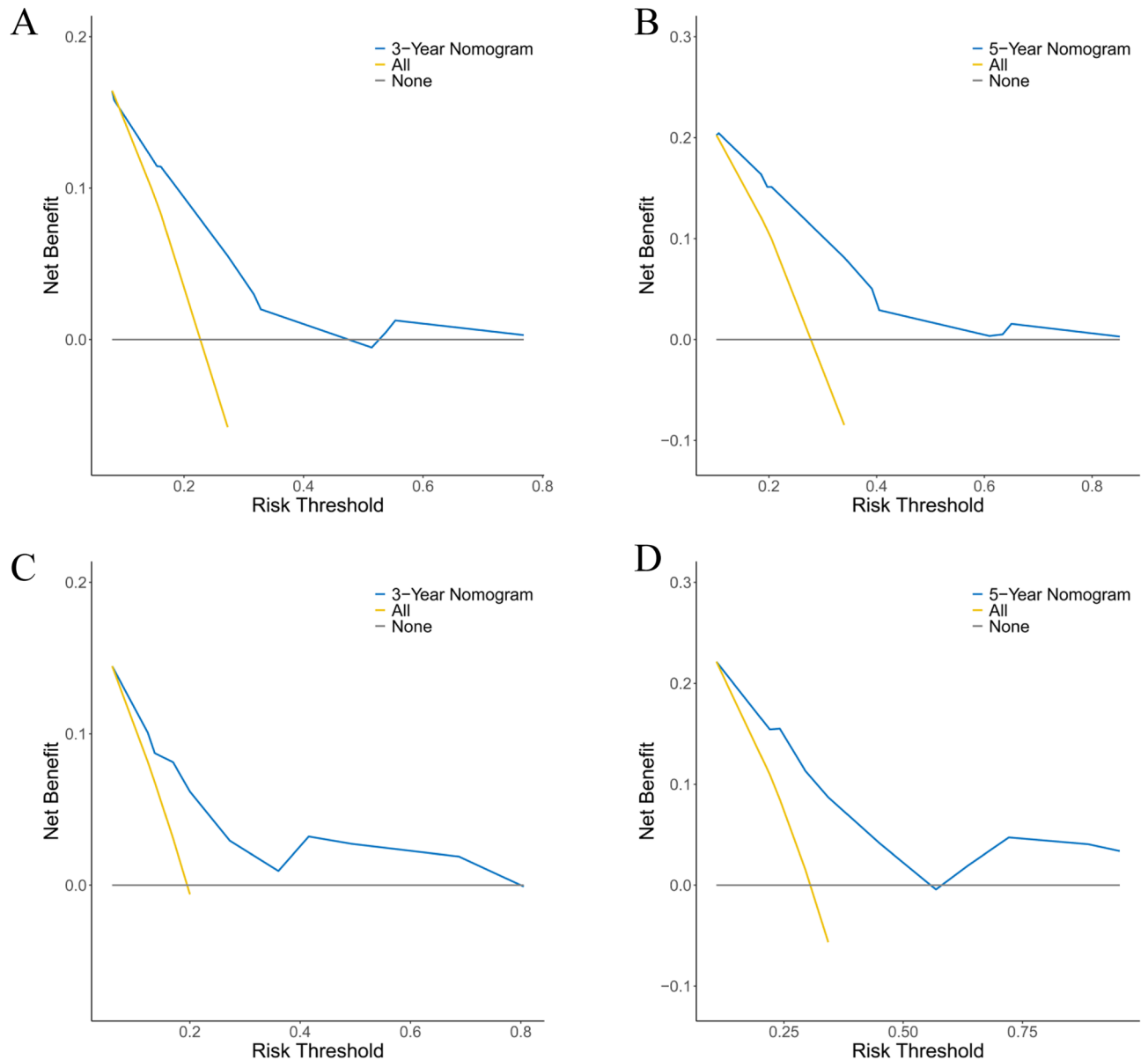
Figure 7. Evaluation of the nomogram of by decision curve analyses. (**A, B**) The decision curve analyses of the nomogram for the 3 years (**A**) and 5 years (**B**) in the training set. (**C, D**) The decision curve analysis of the nomogram for the 3 years (**C**) and 5 years (**D**) in the validation set.

Moreover, it was once reported that CEL exhibits a male predominance, and the median age at diagnosis was 62 years[9]. In this study, the average age of the patients was 57.0 ± 17.0 years, and the IRR of male-to-female was 1.66 (95% CI 1.39–1.98), which is in agreement with the previous research[9,19]. Furthermore, the current study demonstrated that the incidence rate increased with age, with the IRR of the 60+ age group to the < 60 age group being 3.65 (95% CI, 3.07–4.34), which has not been documented in the literature so far. The age distribution of CEL may reflect the accumulation of genetic and epigenetic alterations that lead to clonal expansion of eosinophils over time[20]. The gender difference of CEL may be influenced by hormonal factor or genetic factors that affect the susceptibility or exposure to eosinophilic stimuli[21,22].

A previous study showed that the prognosis of CEL is poor in a cohort of 10 patients and the median survival was 22.2 months, with 5 patients developing acute transformation after median of 20 months from diagnosis[9]. However, it is extremely difficult to draw any conclusion from this study due to the small sample size. In the current study, to identify the prognostic factors for CEL, LASSO Cox regression analysis and multivariate Cox regression analysis were performed on a set of clinical variables. The study unveiled that age, marital status at diagnosis, and sequence were independently associated with overall survival. Older age emerged as a significant adverse prognostic factor, with older individuals facing substantially elevated mortality risks (HR 3.74, 95% CI 2.51–5.60), which is in line with previous studies on MPNs[23]. Marital status was also a significant predictor of survival, with the marital statuses of single and other (divorced, separated, widowed, unmarried or domestic
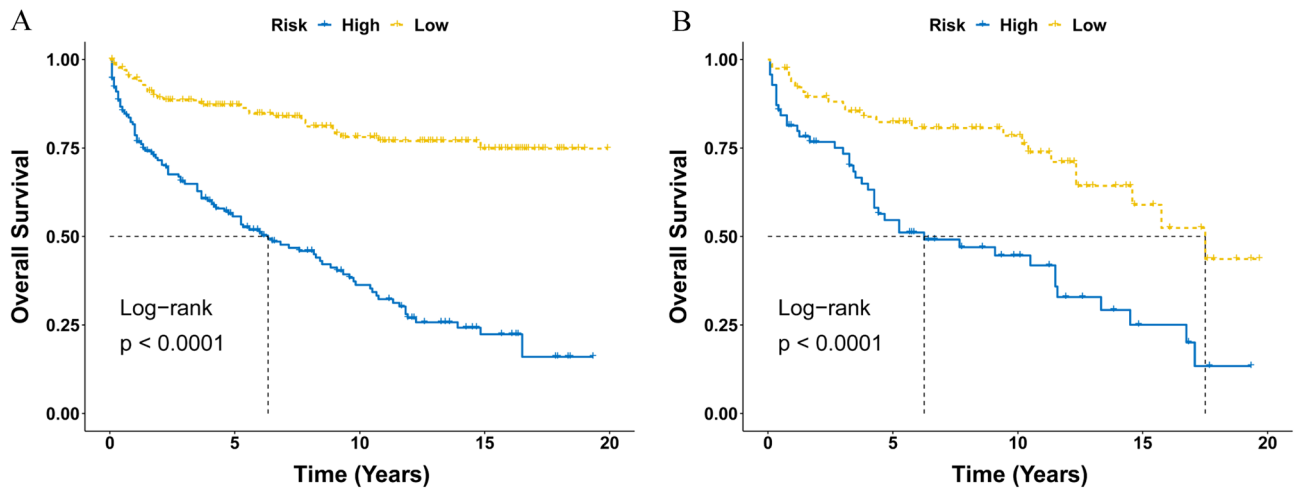
Figure 8. Kaplan–Meier curve of overall survival of CEL patients stratified by the risk stratification system in the training set (**A**) and validation set (**B**). CEL, chronic eosinophilic leukemia.

partner) having a worse outcome than married patients. This may reflect the impact of social support and psychological factors on cancer survival[24–26]. Sequence was another important prognostic factor, with secondary CEL associated with poorer prognosis. This may be due to the presence of other malignancies or comorbidities that affect the treatment response and tolerance[27]. In addition, chemotherapy showed no effect on the OS of CEL patients in this study, which is in accordance with some study showed that CEL patients are usually unresponsive to conventional chemotherapy[9].

Nomograms have been developed and proven to surpass the conventional staging systems in terms of prognostic accuracy for certain types of cancers[24]. Consequently, the integration of nomograms into clinical practice as reliable tools for predicting cancer prognosis has become increasingly prevalent[28,29]. In this study, a prognostic nomogram was constructed to predict the 3- and 5-year overall survival probability of CEL patients based on those three independent prognostic factors: age, marital status at diagnosis, and sequence. The nomogram demonstrated commendable calibration and discriminative performance in both the training and validation cohorts, indicating satisfactory accuracy and consistency of the nomogram. This study is the first to develop and validate a clinical prognostic model for CEL patients. To the best of our knowledge, this is also the largest study ever conducted.

Effective risk stratification is integral to tailoring treatment strategies and optimizing patient outcomes[30]. Utilizing the nomogram, CEL patients were stratified into low- and high-risk groups based on individual risk scores. Kaplan–Meier analysis of OS revealed substantial distinctions between these risk cohorts, underscoring the nomogram's efficacy in risk stratification. This empowers clinicians to identify patients who may benefit from more aggressive therapeutic interventions or intensified surveillance, ultimately contributing to improved patient care and outcomes.

However, there are some limitations with this study. Firstly, the SEER registry did not document other potential prognostic factors that may have a significant impact on CEL patient outcomes, such as genetic mutation, performance status, LDH level, immunophenotypic features, family history and alcohol/smoking consumption history. Secondly, detailed information about therapy was not recorded in the SEER database, making it impossible to analyze the effect of different treatment regimens. Thirdly, this is a retrospective study, which means that there may be unavoidable potential biases such as selection bias. Finally, while the nomograms of CEL were constructed and verified using the same database, they were not further validated using another independent dataset. Thus, although this study provided important insights on CEL due to the rarity and lack of large-scale multicenter prospective study of this disease, the results should still be interpreted with caution.

Conclusions

In conclusion, this study provides novel insights into the epidemiology and prognosis of CEL in the US population using the SEER database. CEL is a very rare disease with a variable clinical presentation and outcome. Age, marital status at diagnosis, and sequence were identified as independent prognostic factors for overall survival, culminating in the development of a prognostic nomogram to predict the 3- and 5-year overall survival probability of CEL patients. This nomogram may help clinicians provide personalized treatment and clinical decisions for CEL patients. To our knowledge, this study represents the largest population-based cohort investigating the epidemiology and survival outcome of CEL patients. However, more clinical research is needed to validate our findings.

Data availability

The data analyzed in this study are from the SEER database (https://seer.cancer.gov/) that are available to the public.

References

1. Arber, D. A. *et al.* The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* **127**, 2391–2405. https://doi.org/10.1182/blood-2016-03-643544 (2016).
2. Shomali, W. & Gotlib, J. World Health Organization-defined eosinophilic disorders: 2019 update on diagnosis, risk stratification, and management. *American Journal of Hematology* **94**, 1149–1167. https://doi.org/10.1002/ajh.25617 (2019).
3. Khoury, J. D. *et al.* The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and histiocytic/dendritic neoplasms. *Leukemia* **36**, 1703–1719, https://doi.org/10.1038/s41375-022-01613-1 (2022).
4. Pardanani, A. *et al.* Long-term follow-up of FIP1L1-PDGFRA-mutated patients with eosinophilia: Survival and clinical outcome. *Leukemia* **26**, 2439–2441. https://doi.org/10.1038/leu.2012.162 (2012).
5. Ogbogu, P. U. *et al.* Hypereosinophilic syndrome: A multicenter, retrospective analysis of clinical characteristics and response to therapy. *J. Allergy Clin. Immunol.* **124**, 1319-1325.e1313. https://doi.org/10.1016/j.jaci.2009.09.022 (2009).
6. Wang, S. A. *et al.* Targeted next-generation sequencing identifies a subset of idiopathic hypereosinophilic syndrome with features similar to chronic eosinophilic leukemia, not otherwise specified. *Mod. Pathol.* **29**, 854–864. https://doi.org/10.1038/modpathol.2016.75 (2016).
7. Campo, E. *et al.* *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues, Revised 4th ed.* (International Agency for Research on Cancer, 2017).
8. Sa, A. W. *et al.* Bone marrow morphology is a strong discriminator between chronic eosinophilic leukemia, not otherwise specified and reactive idiopathic hypereosinophilic syndrome. *Haematologica* **102**, 1352–1360. https://doi.org/10.3324/haematol.2017.165340 (2017).
9. Helbig, G., Soja, A., Bartkowska-Chrobok, A. & Kyrcz-Krzemień, S. Chronic eosinophilic leukemia-not otherwise specified has a poor prognosis with unresponsiveness to conventional treatment and high risk of acute transformation. *Am. J. Hematol.* **87**, 643–645. https://doi.org/10.1002/ajh.23193 (2012).
10. Gotlib, J. World Health Organization-defined eosinophilic disorders: 2017 update on diagnosis, risk stratification, and management. *Am. J. Hematol.* **92**, 1243–1259. https://doi.org/10.1002/ajh.24880 (2017).
11. Helbig, G. *et al.* Durable remission after treatment with very low doses of imatinib for FIP1L1-PDGFRα-positive chronic eosinophilic leukaemia. *Cancer Chemother. Pharmacol.* **67**, 967–969. https://doi.org/10.1007/s00280-011-1582-3 (2011).
12. Metzgeroth, G. *et al.* Safety and efficacy of imatinib in chronic eosinophilic leukaemia and hypereosinophilic syndrome—a phase-II study. *Br. J. Haematol.* **143**, 707–715. https://doi.org/10.1111/j.1365-2141.2008.07294.x (2008).
13. Yamada, O. *et al.* Clinical and cytogenetic remission induced by interferon-α in a patient with chronic eosinophilic leukemia associated with a unique t(3;9;5) translocation. *Am. J. Hematol.* **58**, 137–141. https://doi.org/10.1002/(SICI)1096-8652(199806)58:2%3c137::AID-AJH9%3e3.0.CO;2-T (1998).
14. Basara, N. *et al.* Chronic eosinophilic leukemia: Successful treatment with an unrelated bone marrow transplantation. *Leukemia Lymphoma* **32**, 189–193. https://doi.org/10.3109/10428199809059261 (1998).
15. Dunphy, C. H. Chronic eosinophilic leukemia, not otherwise specified (CEL, NOS). *Curr. Cancer Ther. Rev.* **8**, 30–34, https://doi.org/10.2174/157339412799462512 (2012).
16. Morsia, E., Reichard, K., Pardanani, A., Tefferi, A. & Gangat, N. WHO defined chronic eosinophilic leukemia, not otherwise specified (CEL, NOS): A contemporary series from the Mayo Clinic. *Am. J. Hematol.* **95**, E172–E174. https://doi.org/10.1002/ajh.25811 (2020).
17. Crane, M. M., Chang, C. M., Kobayashi, M. G. & Weller, P. F. Incidence of myeloproliferative hypereosinophilic syndrome in the United States and an estimate of all hypereosinophilic syndrome incidence. *J. Allergy Clin. Immunol.* **126**, 179–181. https://doi.org/10.1016/j.jaci.2010.03.035 (2010).
18. Vardiman, J. W. *et al.* The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: Rationale and important changes. *Blood* **114**, 937–951. https://doi.org/10.1182/blood-2009-03-209262 (2009).
19. Requena, G. *et al.* Clinical profile and treatment in hypereosinophilic syndrome variants: A pragmatic review. *J. Allergy Clin. Immunol.: Pract.* **10**, 2125–2134, https://doi.org/10.1016/j.jaip.2022.03.034 (2022).
20. Aldoss, I., Forman, S. J. & Pullarkat, V. Acute lymphoblastic leukemia in the older adult. *Journal of Oncology Practice* **15**, 67–75. https://doi.org/10.1200/JOP.18.00271 (2019).
21. Garner, W. B. *et al.* Predicting future cancer incidence by age and gender. *J. Clin. Oncol.* **37**, 1559–1559. https://doi.org/10.1200/JCO.2019.37.15_suppl.1559 (2019).
22. Ornos, E. D. *et al.* Molecular basis of sex differences in cancer: Perspective from Asia. *iScience* **26**, 107101, https://doi.org/10.1016/j.isci.2023.107101 (2023).
23. Hultcrantz, M. *et al.* Patterns of survival among patients with myeloproliferative neoplasms diagnosed in Sweden From 1973 to 2008: A population-based study. *J. Clin. Oncol.* **30**, 2995–3001. https://doi.org/10.1200/JCO.2012.42.1925 (2012).
24. Aizer, A. A. *et al.* Marital status and survival in patients with cancer. *J. Clin. Oncol.* **31**, 3869–3876. https://doi.org/10.1200/jco.2013.49.6489 (2013).
25. Ding, Z., Yu, D., Li, H. & Ding, Y. Effects of marital status on overall and cancer-specific survival in laryngeal cancer patients: a population-based study. *Sci. Rep.* **11**, 723. https://doi.org/10.1038/s41598-020-80698-z (2021).
26. Kravdal, Ø. The impact of marital status on cancer survival. *Soc. Sci. Med.* **52**, 357–368. https://doi.org/10.1016/S0277-9536(00)00139-8 (2001).
27. Keegan, T. H. M., Bleyer, A., Rosenberg, A. S., Li, Q. & Goldfarb, M. Second primary malignant neoplasms and survival in adolescent and young adult cancer survivors. *JAMA Oncol.* **3**, 1554–1557. https://doi.org/10.1001/jamaoncol.2017.0465 (2017).
28. Iasonos, A., Schrag, D., Raj, G. V. & Panageas, K. S. How to build and interpret a nomogram for cancer prognosis. *J. Clin. Oncol.* **26**, 1364–1370. https://doi.org/10.1200/JCO.2007.12.9791 (2008).
29. Balachandran, V. P., Gonen, M., Smith, J. J. & DeMatteo, R. P. Nomograms in oncology: More than meets the eye. *Lancet Oncol.* **16**, e173–e180. https://doi.org/10.1016/S1470-2045(14)71116-7 (2015).
30. Watson, E. K. *et al.* Personalised cancer follow-up: risk stratification, needs assessment or both?. *Br. J. Cancer* **106**, 1–5. https://doi.org/10.1038/bjc.2011.535 (2012).

Acknowledgements

Author contributions

Conceptualization, F.W..; methodology, F.W.; software, F.W.; validation, F.W. & JL.W. & MT.L.; formal analysis, F.W. & JL.W.; investigation, F.W.; resources, F.W.; data curation, F.W. & JL.W. & MT.L.; writing-original draft preparation, F.W.; writing-review and editing, F.W. & JL.W. & MT.L.; visualization, F.W.; supervision, F.W.; project administration, F.W.; funding acquisition, F.W. All authors have read and agreed to the published version of the manuscript.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to F.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.