scientific reports

Check for updates

OPEN

# Prediction model establishment of prognosis factors for acute myeloid leukemia based on the SEER database

Gangping Li[1], Di Zhang[2] & Yuewen Fu[1]✉

**Acute myeloid leukemia (AML) with t (9;11) (p22; q23) presents as a varied hematological malignancy. The t (9;11) (p22; q23) translocation is the most common among 11q23/KMT2A rearrangements in AML. This research aimed to develop a nomogram for precise prediction of overall survival (OS) and cancer-specific survival (CSS) in AML with the t (9;11) (p22; q23) translocation. We utilized the Surveillance, Epidemiology, and End Results (SEER) database to identify patients diagnosed with t (9;11) (p22; q23) AML from 2000 to 2021. Prognostic factors for this AML subtype were determined using least absolute shrinkage and selection operator (LASSO) regression, which guided the creation of prognostic nomograms. To evaluate the model's discrimination, accuracy, and effectiveness, we employed the concordance index (C-index), calibration charts, receiver operating characteristic curves (ROC), area under the curve (AUC), and decision-curve analysis (DCA). The research was meticulously planned, executed, and documented in full adherence to the TRIPOD guidelines. The nomogram was developed using key variables including age, race, first primary tumor, and chemotherapy. The concordance indices (C-indices) were 0.704 for OS and for 0.686 for CSS. Patients were classified into high-risk and low-risk groups based on nomogram scores, with significant differences in OS and CSS between these groups ($P < 0.001$). This study developed innovative nomograms that combine clinical and treatment factors to predict 1-, 3-, and 5-year survival rates for patients with t (9;11) (p22; q23) AML.**

**Keywords** Prognosis, Nomogram, T (911) (p22q23) acute myeloid leukemia, SEER, Analysis

Acute myeloid leukemia (AML) with t(9;11)(p22;q23) accounts for 9–12% of pediatric cases and 2% of adult cases, typically presenting with monocytic features[1–5]. Translocations involving chromosome 11q23 are commonly observed in pediatric AML and are linked to a poor prognosis. These translocations usually involve the MLL gene, with over 50 identified partners. The most common is the t(9;11)(p22;q23) translocation, which fuses KMT2A with the MLLT3 gene (formerly AF9)[1]. This translocation defines a distinct disease category in the 2016 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia[6]. In patients under 60 years with t (9;11) (p22; q23) /KMT2A-MLLT3, classified as intermediate risk by the 2017 European Leukemia Net guidelines, outcomes are generally better compared to intermediate-risk patients without this translocation[5]. However, older patients with t(9;11) and those with other 11q23/KMT2A rearrangements experience worse outcome[1].

Prior research has highlighted diverse clinical characteristics and prognostic factors, including age, sex, race, marital status, and chemotherapy utilization, which impact the prognosis of individuals with t (9;11) (p22; q23) AML[7,8]. Given the limitations of relying on single prognostic factors, a predictive model is needed to forecast the occurrence of the outcome.

The predictive models improve accuracy and efficiency in forecasting patient outcomes, crucial in clinical decision-making. The nomogram, a key tool for visualizing Cox proportional hazards models, helps interpret variable risks and identify key predictors. Widely used in oncology and medical science, nomograms serve as valuable diagnostic and prognostic tools[9]. Utilizing the Surveillance, Epidemiology, and End Results (SEER) database, which covers about 30% of the US (United States of America) population[10], we included individuals diagnosed with t(9;11) (p22;q23) AML between 2000 and 2021 in our study. AML can be classified into de

[1]Department of Hematology, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, Zhengzhou, China. [2]Department of Medical Records Management Department, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, Zhengzhou, China. ✉email: zlyyfuyuewen1480@zzu.edu.cn

novo AML and secondary AML based on the presence or absence of a prior medical history[11–13]. Our research aims to explore the clinical and prognostic characteristics of t (9;11) (p22; q23) AML, and develop prognostic nomograms specific to t (9;11) (p22; q23) AML.

## Materials and methods

### Patient selection and data collection

Our study utilized data extracted from the SEER database (SEER*Stat v8.4.3). Patient selection and clinical data collection for t (9;11) (p22; q23) AML were conducted within the SEER research dataset spanning 17 registries from 2000 to 2021. As shown in Fig. 1, inclusion criteria for t (9;11) (p22; q23) AML patients required histology code 9897/3 as per ICD-O-3, along with confirmed favorable cytology or histology. Exclusion criteria included ambiguous or unknown diagnoses and unknown survival times. Key clinical features extracted included age, gender, race, year of diagnosis, marital status at diagnosis, median household income, primary malignancy status, treatment, survival status, survival time, and cause of death. Given the SEER database's anonymized nature and open accessibility, ethical approval from an ethics committee was deemed unnecessary.

### Development, validation, and assessment of nomograms

The development and validation of nomograms for predicting both overall survival (OS) and cancer-specific survival (CSS) in t (9;11) (p22; q23) AML involved several key steps. Initially, least absolute shrinkage and selection operator (LASSO) regression was utilized to select potential prognostic factors, with nine variables (age, sex, race, marital status, median household income, first primary tumor, chemotherapy, radiation, and year of diagnosis) included in the analysis. The selected potential prognostic factors were then integrated into the model, and a nomogram was constructed to visually represent it. Discrimination and calibration of the model were evaluated using receiver operating characteristic (ROC) analysis, concordance index (C-index), and calibration curves. Additionally, 1000 bootstrap samples were used to correct the ROC, C-index, and calibration curves. Decision curve analysis (DCA) assessed clinical utility by assigning cumulative points to predict survival rates, where higher scores indicated lower survival probabilities. Both in the training set and the validation set, patients were categorized into high- and low-risk groups based on the median risk score derived from the predictive model. Those with a risk score equal to or greater than the median were assigned to the high-risk group, while patients with a score below the median were placed in the low-risk group. Differences in OS and CSS between these categories were statistically assessed using log-rank tests and Kaplan–Meier survival analysis.

### Statistical examination

We used the Mann–Whitney U-test to compare continuous variables across different groups, while categorical variables were assessed using either the chi-squared test or Fisher's exact test. OS and CSS were analyzed using the Kaplan–Meier method and compared using the log-rank test. All analyses were conducted using the R Statistical Software (Version 4.2.2) and the Free Statistics Analysis Platform (Version 1.9, Beijing, China; http://www.clinicalscientists.cn/freestatistics). The SEER database data were divided into a training cohort and a validation cohort at a ratio of 7:3 via the "caret" package. LASSO regression was performed using R statistical packages, including "glmnet", "rms", "survival", "timeROC", "foreign", "Hmisc", and "ggDCA". The timeROC package was employed to generate receiver operating characteristic (ROC) curves for both the training and validation cohorts, assessing the model's discriminatory power. A higher area under the curve (AUC) indicates better predictive accuracy. Calibration curves were created using bootstrapping with 1000 repetitions of resampling. To evaluate the clinical

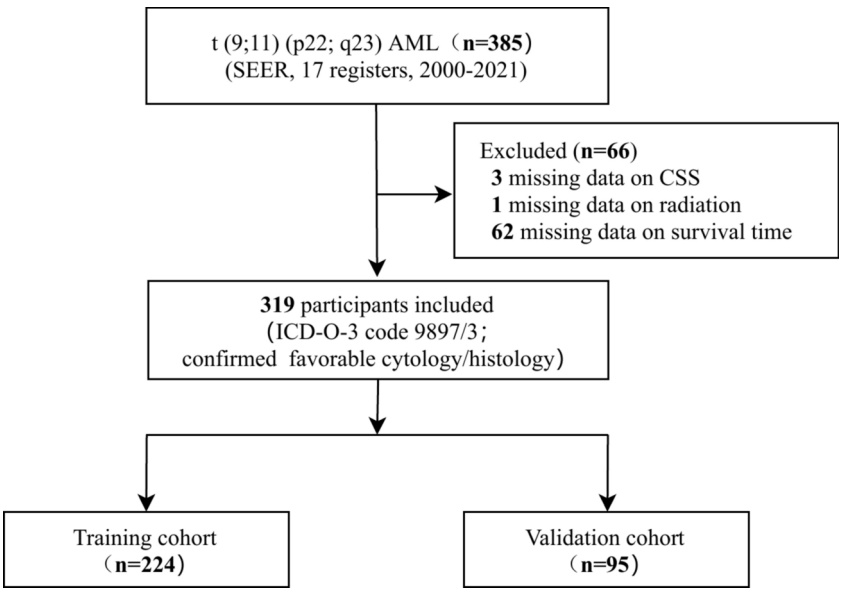


**Fig. 1**. The flow chart of the study.

utility of the model, decision curve analysis was conducted using the ggDCA packages. The **glmnet** function was applied for LASSO regression, where **family = "Gaussian"** indicates a linear regression model, and **alpha = 1** specifies the use of pure LASSO (no ridge regularization). To optimize the regularization parameter **lambda**, the **cv.glmnet** function was employed through cross-validation. The **plot** function was used to visualize the LASSO path and cross-validation performance. Finally, the **coef** function allowed extraction of the model coefficients at a given **lambda** value. A two-tailed P-value less than 0.05 was considered statistically significant.

## Results

### Clinical features of patients

Between 2000 and 2021, the SEER database identified 319 individuals diagnosed with t (9;11) (p22; q23) AML. Table 1 succinctly outlines their demographic and clinical characteristics. The median age of these patients was 50.0 years, ranging from 0 to 85 years. A vast majority were white race (81.5%), female (53.0%) and with most diagnoses occurring between 2010 and 2021 (58%). Moreover, a substantial number (77.1%) of these individuals were initially diagnosed with t (9;11) (p22; q23) AML as their primary tumor. Most patients (85.6%) opted for chemotherapy, while a minority (11.6%) chose radiotherapy for treatment.

| Characteristic | Total (n = 319) | Training cohort (n = 224) | Validation cohort (n = 95) | *P* Value |
|---|---|---|---|---|
| Age at diagnosis (year), Median (IQR) | 50.0 (27.0, 68.0) | 49.5 (25.8, 67.2) | 51.0 (28.5, 68.5) | 0.75 |
| Sex, n (%) | | | | 0.252 |
| Male | 150 (47.0) | 110 (49.1) | 40 (42.1) | |
| Female | 169 (53.0) | 114 (50.9) | 55 (57.9) | |
| Race, n (%) | | | | 0.705 |
| White | 260 (81.5) | 180 (80.4) | 80 (84.2) | |
| Black | 25 (7.8) | 19 (8.5) | 6 (6.3) | |
| Other | 34 (10.7) | 25 (11.2) | 9 (9.5) | |
| Marital status, n (%) | | | | 0.621 |
| Married | 137 (42.9) | 96 (42.9) | 41 (43.2) | |
| Unmarried | 111 (34.8) | 81 (36.2) | 30 (31.6) | |
| Other | 71 (22.3) | 47 (21) | 24 (25.3) | |
| Median household income, (A thousand dollars) | | | | 0.061 |
| < 50 | 14 (4.4) | 7 (3.1) | 7 (7.4) | |
| 50–75 | 127 (39.8) | 84 (37.5) | 43 (45.3) | |
| > 75 | 178 (55.8) | 133 (59.4) | 45 (47.4) | |
| 1st primary tumor, n (%) | | | | 0.167 |
| No | 73 (22.9) | 56 (25) | 17 (17.9) | |
| Yes | 246 (77.1) | 168 (75) | 78 (82.1) | |
| Radiation, n (%) | | | | 0.708 |
| No/Unknown | 282 (88.4) | 199 (88.8) | 83 (87.4) | |
| Yes | 37 (11.6) | 25 (11.2) | 12 (12.6) | |
| Chemotherapy, n (%) | | | | 0.65 |
| No/Unknown | 46 (14.4) | 31 (13.8) | 15 (15.8) | |
| Yes | 273 (85.6) | 193 (86.2) | 80 (84.2) | |
| Status, n (%) | | | | 0.401 |
| Alive | 91 (28.5) | 67 (29.9) | 24 (25.3) | |
| Dead | 228 (71.5) | 157 (70.1) | 71 (74.7) | |
| Cancer-specific survival, n (%) | | | | 0.712 |
| Alive or dead of other cause | 126 (39.5) | 87 (38.8) | 39 (41.1) | |
| Dead (attributable to this cancer) | 193 (60.5) | 137 (61.2) | 56 (58.9) | |
| Survival time (months), Median (IQR) | 14.0 (6.0, 43.5) | 14.0 (6.0, 40.5) | 13.0 (6.0, 68.0) | 0.701 |
| Age, n (%) | | | | 0.847 |
| ≤ 19 | 62 (19.4) | 45 (20.1) | 17 (17.9) | |
| 20–39 | 58 (18.2) | 42 (18.8) | 16 (16.8) | |
| 40–59 | 84 (26.3) | 56 (25) | 28 (29.5) | |
| 60–85 | 115 (36.1) | 81 (36.2) | 34 (35.8) | |
| Year of diagnosis, n (%) | | | | 0.31 |
| 2000–2010 | 134 (42.0) | 90 (40.2) | 44 (46.3) | |
| 2011–2021 | 185 (58.0) | 134 (59.8) | 51 (53.7) | |

**Table 1**. Baseline characteristics of patients diagnosed with t (9;11) (p22; q23) AML from the SEER database.

## Survival analysis

Among the 319 patients with t (9;11) (p22; q23) AML in the research, 71.5% (228/319) experienced mortality, with 84.6% (193/228) specifically succumbing to AML. Employing the log-rank test and Kaplan–Meier survival analysis, we quantified and depicted OS and CSS, stratifying individuals according to age. As illustrated in Fig. 2, t (9;11) (p22; q23) AML patients exhibited a median OS time of 14.0 months, along with a 3-year OS rate of 32.3%. Simultaneously, the median CSS time was 39.7 months, accompanied by a 37% 3-year CSS rate. When categorized by age groups (≤19, 20–39, 40–59, 60–85 years), both OS and CSS show a declining trend with increasing age, with all p-values being less than 0.0001, as illustrated in Fig. 2.

## The selection of prognostic factors and the development of nomograms

Patients with t (9;11) (p22; q23) AML were randomly assigned to training (n = 224) and validation cohorts (n = 95). There were no statistically significant differences observed between the training and validation cohorts across variables, as shown in Table 1. LASSO regression analysis was performed to identify the potential prognostic factors, with nine variables (age, sex, race, marital status, median household income, first primary tumor, chemotherapy, radiation, and year of diagnosis) included in the analysis. To determine the optimal λ (lambda), we analyzed the curves of the regression coefficients and partial likelihood deviation against log(λ). Using λ.1se as the optimal value, we selected four independent prognostic factors related to OS and CSS: age, race,

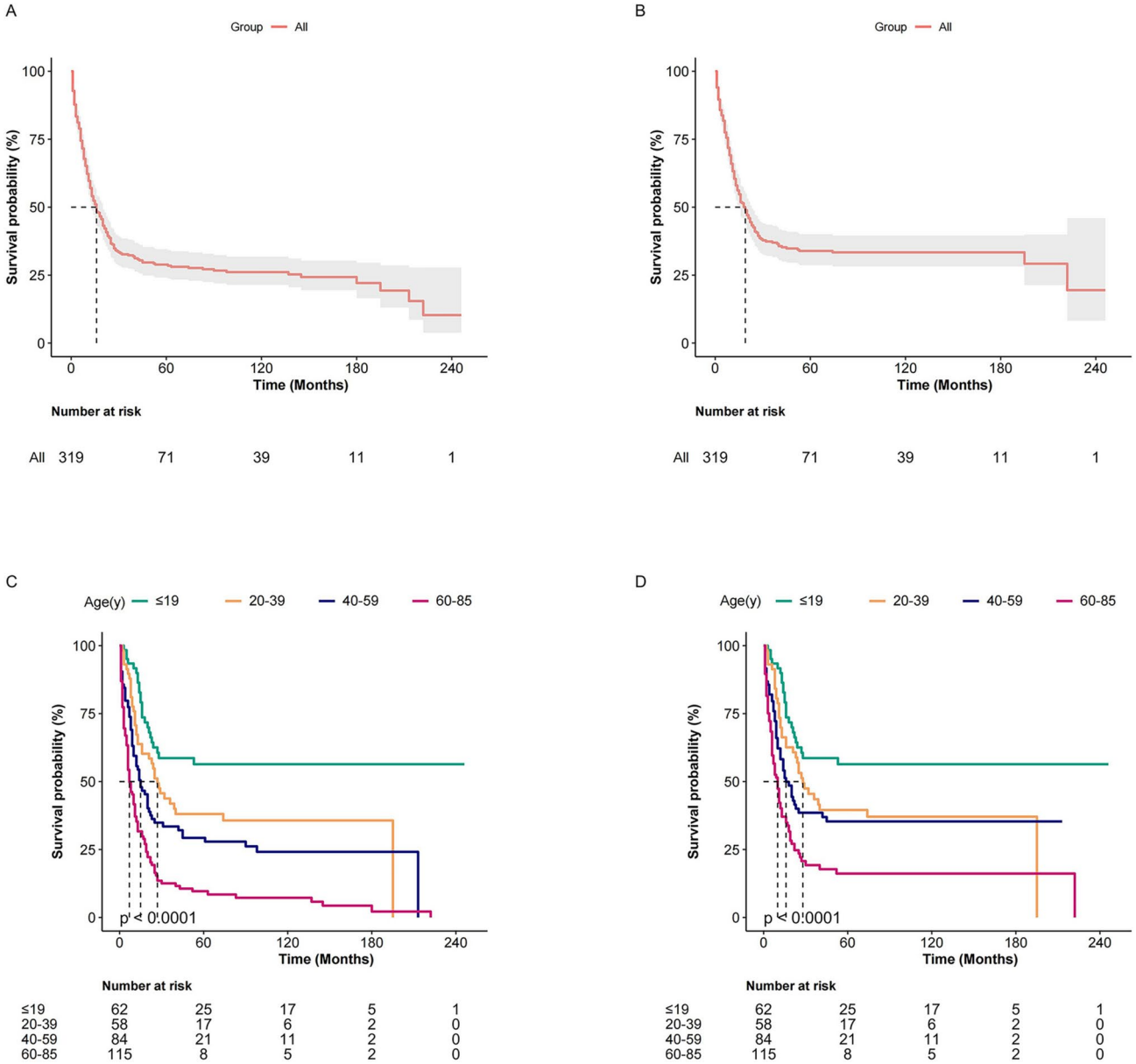


**Fig. 2**. Kaplan–Meier analysis of overall survival (OS) and cancer-specific survival (CSS) in t (9;11) (p22; q23) AML. Kaplan–Meier survival curves of OS for all patients (**A**), CSS for all patients (**B**). Kaplan–Meier survival curves of OS for patients stratified by age (**C**), CSS for patients stratified by age (**D**).

first primary tumor, and chemotherapy (Fig. 3).The LASSO regression analysis results led to the identification of age group (≤ 19, 20–39, 40–59, 60–85 years), first primary tumor (yes, no), race (white, black, other), and chemotherapy (yes, no/unknown) as vital prognostic elements for constructing OS and CSS nomograms (as presented in Fig. 3). In the training group, the C-index values for the OS and CSS nomograms were 0.704 (95% CI 0.663–0.746) and 0.686 (95% CI 0.639–0.733), respectively. In the validation cohort, these C-index values were 0.645 (95% CI 0.573–0.717) for the OS nomogram and 0.633 (95% CI 0.556–0.710) for the CSS nomogram (Fig. 4).

## Nomogram validation and evaluation

We assessed the discrimination and calibration abilities of the OS and CSS nomograms by plotting calibration and ROC curves (Fig. 5). Analysis of the ROC curves revealed AUC values of 0.762, 0.748, and 0.738 for 1-, 3-, and 5-year OS in the training group, and 0.747, 0.721, and 0.716 for 1-, 3-, and 5-year CSS. In the validation cohort, corresponding AUC values were 0.671, 0.679, and 0.726 for OS, and 0.670, 0.676, and 0.711 for CSS across the same time intervals. Moreover, calibration curves demonstrated excellent concordance between observed and nomogram-predicted survival rates at 1, 3, and 5 years in both validation and training cohorts. The clinical utility of the nomograms was further evaluated using DCA, which indicated significant net benefits for both OS and CSS nomograms. These findings underscore the robust clinical applicability and effective predictive accuracy in estimating survival outcomes over different time intervals, as depicted in Fig. 6.

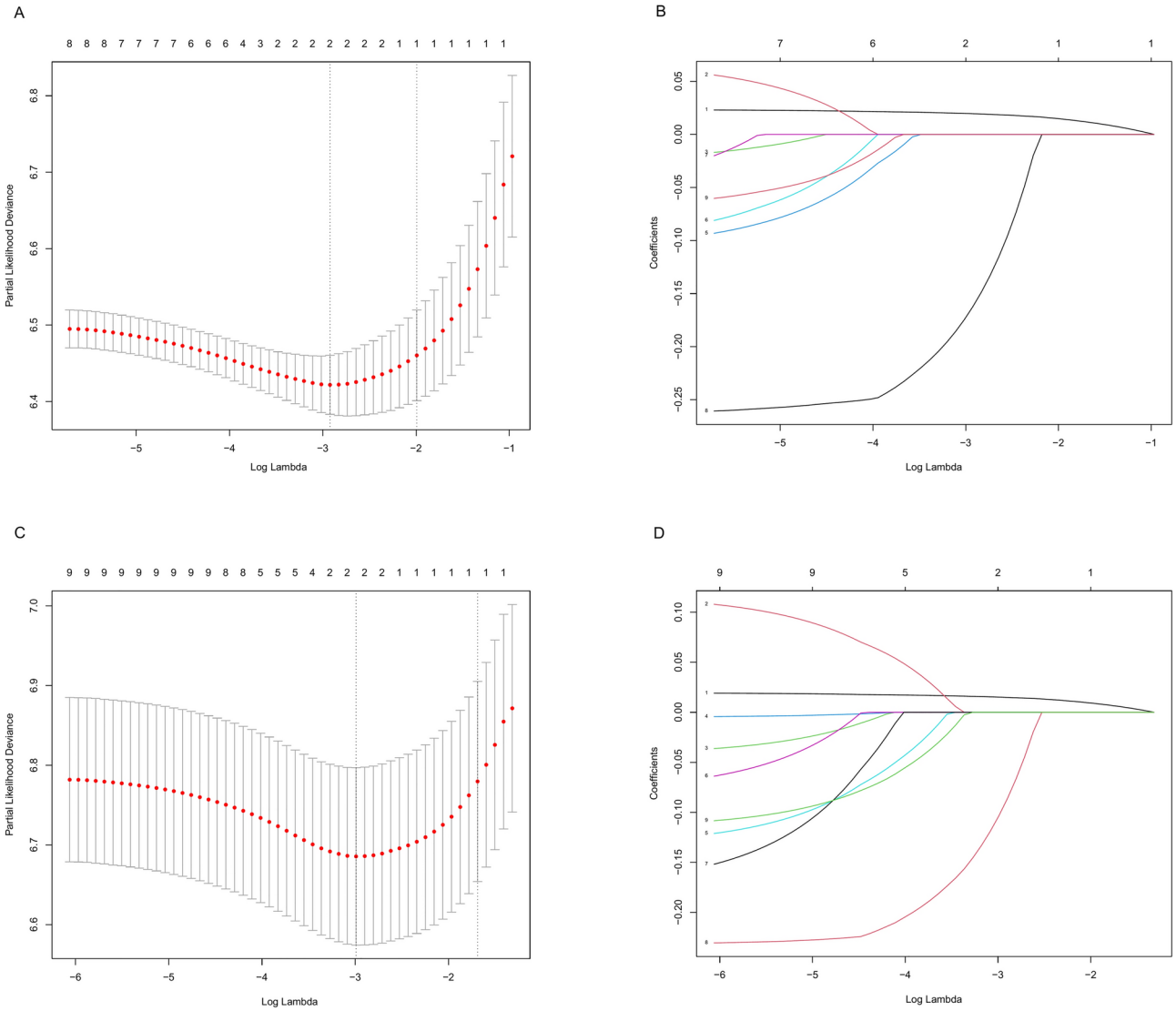


**Fig. 3**. LASSO regression plot for OS and CSS. **A** Plot of partial likelihood deviance for OS; **B** plot of LASSO coefficient profiles for OS; **C** Plot of partial likelihood deviance for CSS; **D** plot of LASSO coefficient profiles for CSS. Each curve illustrates the LASSO coefficient profile of a feature across the log (lambda) sequence.

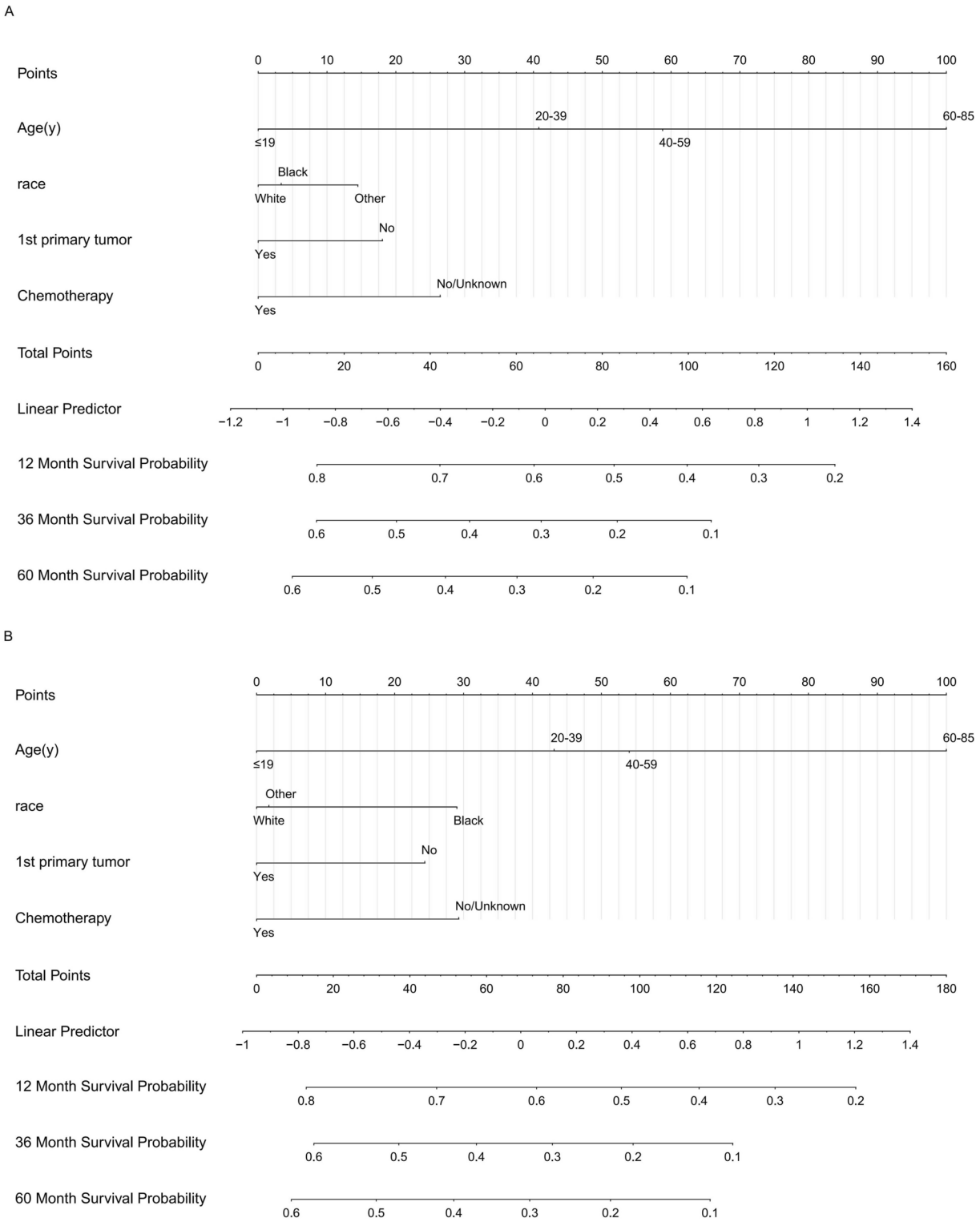


**Fig. 4**. Nomograms for predicting the 1-, 3- and 5-year OS rates (**A**) and CSS rates (**B**) for t (9;11) (p22; q23) AML patients.

## Risk stratification based on the nomogram points

To aid in prognostic stratification, patients with t (9;11) (p22; q23) AML were categorized into two risk groups based on nomogram points: low-risk (lower points) and high-risk (higher points). In the training set, patients were categorized into high- and low-risk groups based on the median risk score derived from the predictive model. Those with a risk score equal to or greater than the median were assigned to the high-risk group, while patients with a score below the median were placed in the low-risk group. In the validation set, the same cutoff

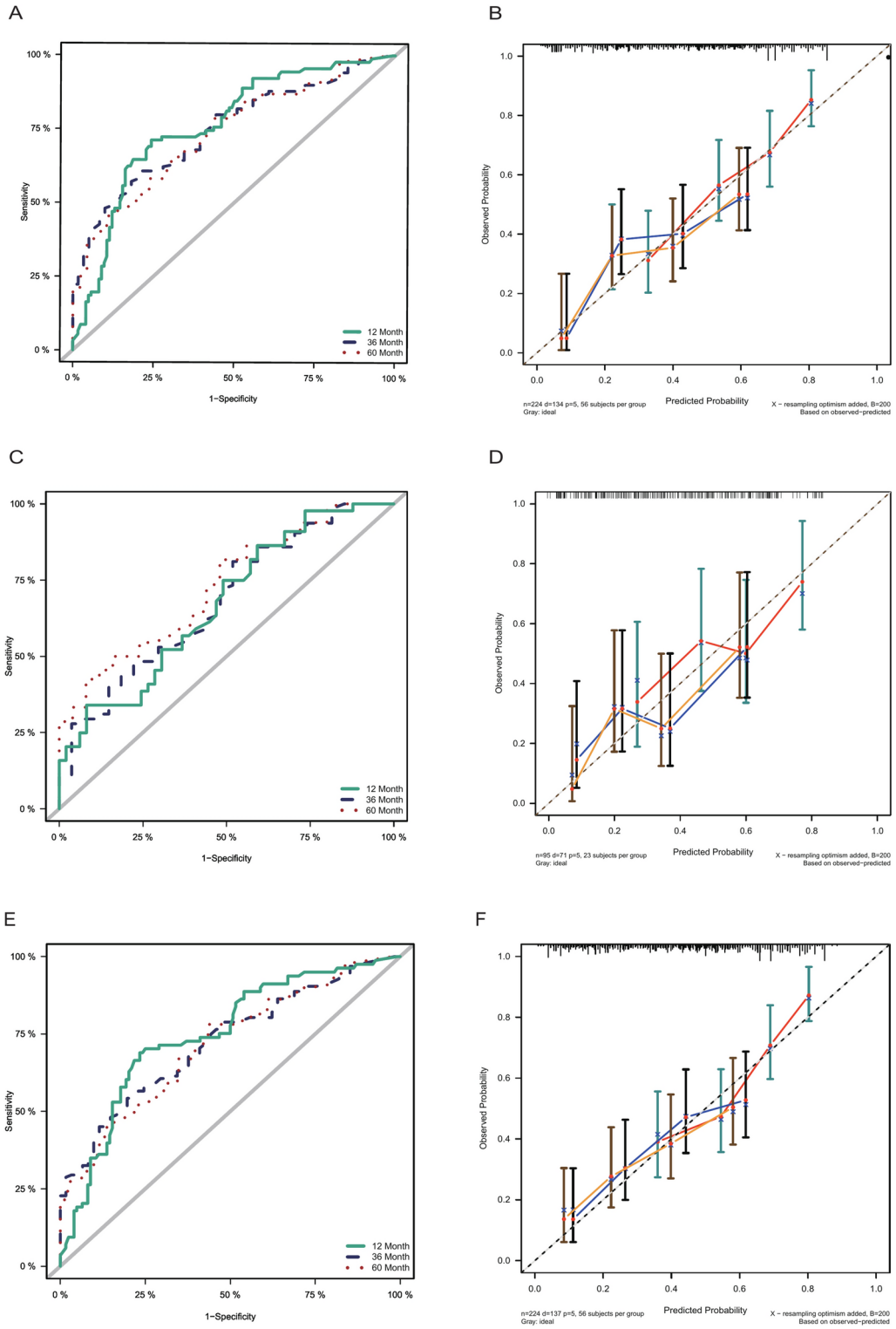


**Fig. 5**. Receiver operating characteristic (ROC) and calibration curves of the nomogram for OS and CSS. In the training cohort, the nomogram achieved AUC values of 0.762, 0.748, and 0.738 for predicting 1-, 3-, and 5-year OS (**A**), and 0.747, 0.721, and 0.716 for predicting 1-, 3-, and 5-year CSS (**E**). In the validation cohort, corresponding AUC values were 0.671, 0.679, and 0.726 for OS (**C**), and 0.670, 0.676, and 0.711 for CSS (**G**) across the same time intervals. Calibration curves for predicting 1-, 3-, and 5-year OS and CSS using the nomogram were generated for both the training cohort (**B**, **F**) and the validation cohort (**D**, **H**). The x-axis represents the predicted survival rate by the model, while the y-axis shows the actual survival rate. The ideal alignment is along the 45-degree line, where predicted and actual survival rates match perfectly. In the plots, red, blue, and orange lines correspond to the model's predictions and actual outcomes for 1-year, 3-year, and 5-year OS, respectively.

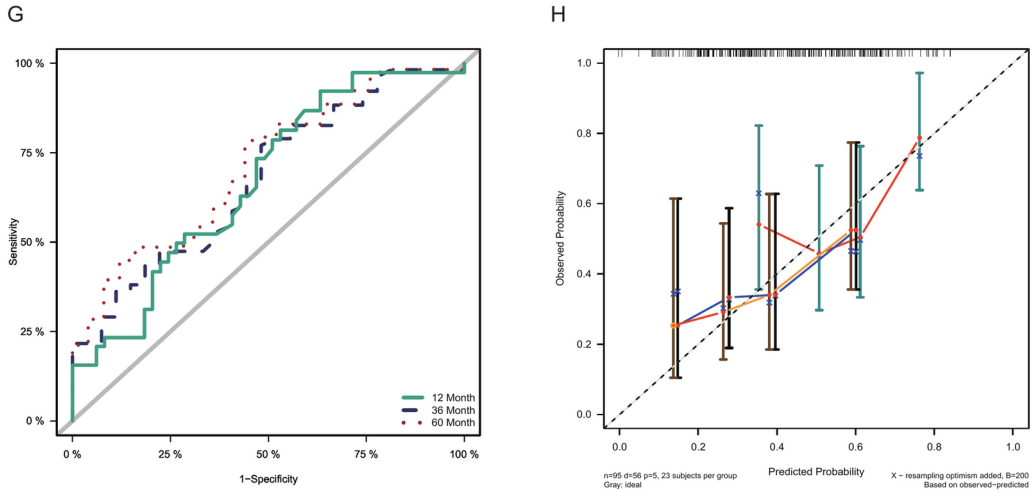


**Figure 5.** (continued)

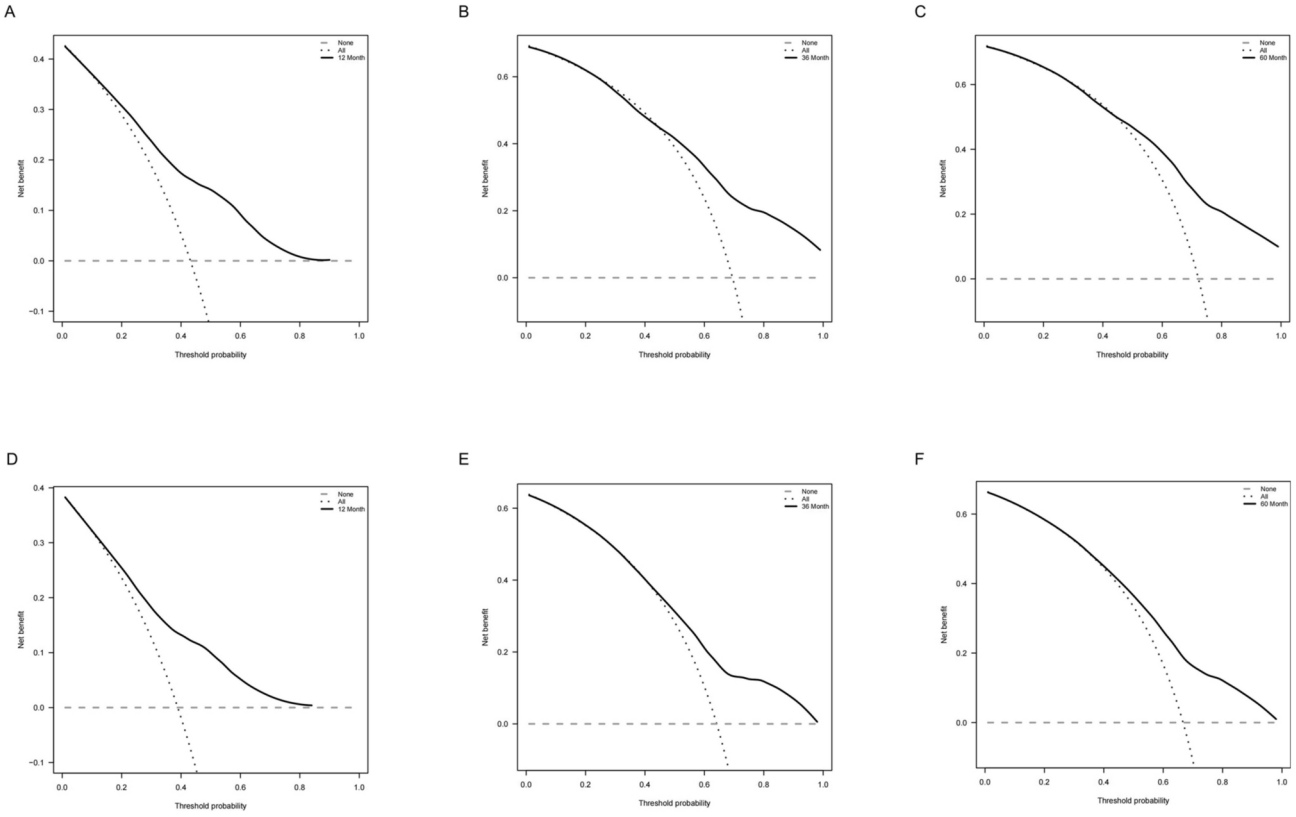


**Fig. 6**. Decision Curve Analysis (DCA) results evaluating a nomogram's predictive accuracy for 1-, 3-, and 5-year OS rates (**A**–**C**) and CSS rates (D-F) in t (9;11) (p22; q23) AML patients. The x-axis shows threshold probabilities, while the y-axis indicates the net benefit. The horizontal line on the x-axis shows no deaths predicted, while the diagonal dashed line indicates all patients predicted to have died. All = assuming all patients survive. None = assuming no patients survive.

point (the median score from the training set) was applied to divide patients into high- and low-risk groups. In the training group, optimal cut-off values for OS (risk score: 0.05) and CSS (risk score: −0.04) nomogram scores were determined. Similarly, in the testing group, cut-off values for OS (risk score:0.03) and CSS (risk score:0.07) nomogram scores were identified (Figs. S1-4). Kaplan–Meier survival analysis clearly stratified patients, revealing a distinct division between those with a more favorable prognosis in the low-risk group and those with a less favorable prognosis in the high-risk group. Significant differences in both OS and CSS were evident, as

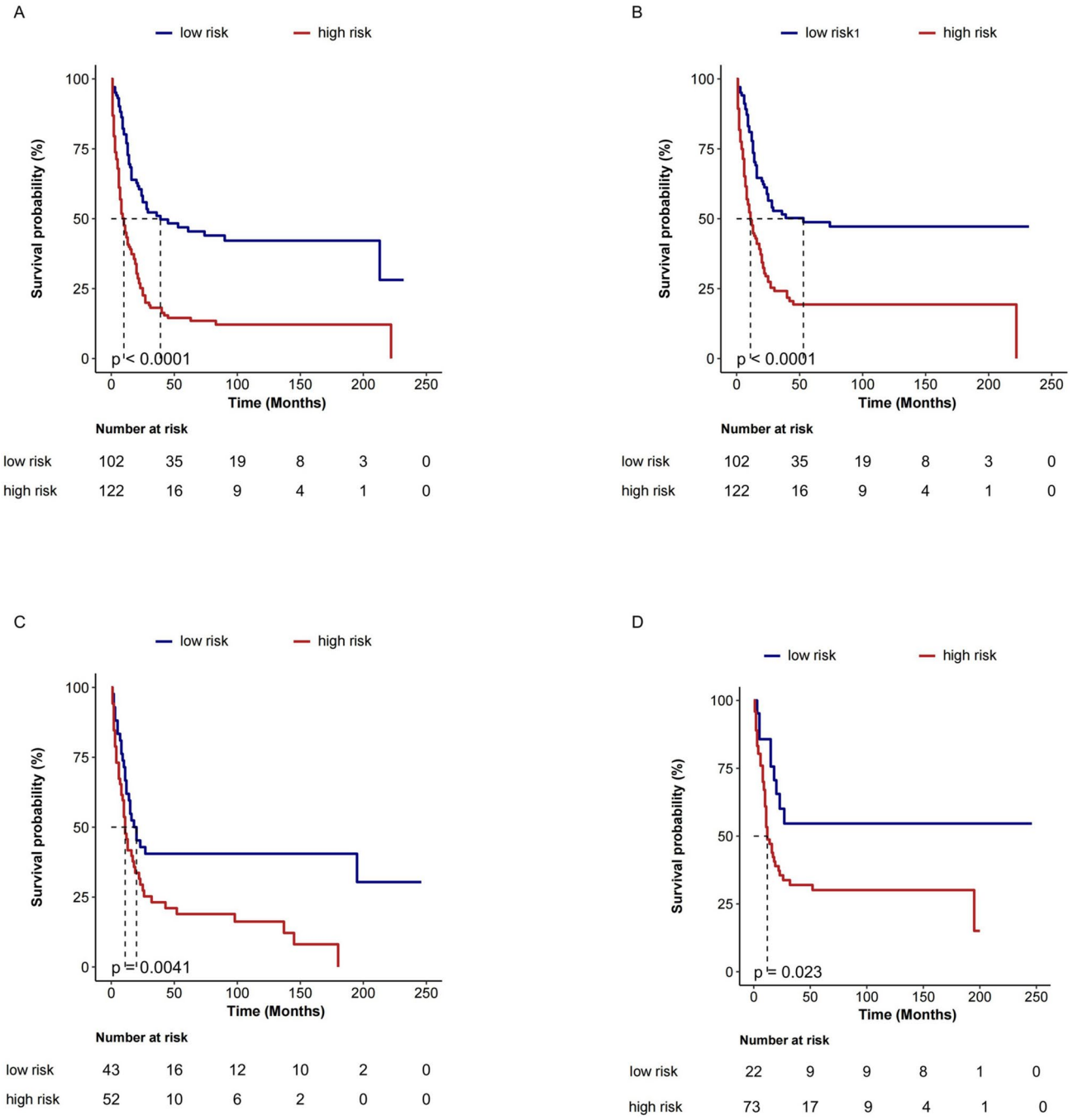


**Fig. 7**. Kaplan–Meier analysis of OS and CSS in t (9;11) (p22; q23) AML patients for risk stratification. Kaplan–Meier survival curves of OS and CSS in t (9;11) (p22; q23) AML patients stratified by the total points in the training cohort (**A**, **B**) and validation cohort (**C**, **D**).

depicted in Fig. 7(A–C and B–D). In summary, Kaplan–Meier analysis effectively classified t (9;11) (p22; q23) AML patients into low- and high-risk groups, highlighting significant disparities in both OS and CSS outcomes.

## Discussion

Nomograms are well-regarded for their accuracy and ease of use in prognostic estimation across various cancers[9,14]. In this study, a detailed analysis was performed on 319 patients with t (9;11) (p22; q23) AML using data from the SEER database. LASSO regression identified age group, race, first primary tumor, and chemotherapy use as independent prognostic factors for OS and CSS in these patients. This research represents the first attempt to develop and validate a nomogram for predicting OS and CSS in t (9;11) (p22; q23) AML patients based on SEER data.

Age has consistently been identified as a significant predictor of tumor prognosis in multiple studies[14–17]. The t(9;11)(p22;q23) translocation results in the MLL-AF9 fusion protein, present in up to 25% of de novo AML cases in children[2]. In our study, age group emerged as the strongest predictor for t (9;11) (p22; q23) AML prognosis. Compared to patients aged 19 or younger, those in the 20–39, 40–59, and over 60 age brackets showed progressively poorer OS and CSS outcomes, with the over-60 group experiencing the worst survival rates. This trend indicates a significant decline in survival with increasing age at diagnosis, consistent with previous research[17].

For the last three decades, intensive acute myeloid leukaemia (AML) therapy includes induction with the 7 + 3 regimen (cytarabine and an anthracycline), followed by consolidation to reduce residual disease. Consolidation options include high-dose cytarabine, allogeneic stem cell transplantation, and maintenance therapy, with the European Leukemia Net recommending specific regimens based on age and risk[13,18]. Chemotherapy remains crucial for managing many AML patients, offering the best chance for long-term remission, particularly when combined with transplantation[19–24]. Some studies suggest that the t (9;11) (p22; q23) subgroup has a better prognosis, possibly due to increased sensitivity to various drugs, as shown in this subgroup[25–28]. In our study, the absence of chemotherapy was found to be a risk factor for survival, consistent with previous results; whereas radiotherapy was not identified as an independent prognostic factor.

AML can be classified as de novo (dn-AML) or secondary (s-AML)[11–13]Secondary AML(s-AML) can develop from an antecedent hematologic disorder, such as myelodysplastic syndrome (MDS) or myeloproliferative neoplasm (MPN), or result from prior cytotoxic chemotherapy or radiation therapy (therapy-related AML, t-AML)[13]. Secondary and therapy-related AML present distinct clinical and biological features compared to dn-AML. Secondary AML(s-AML) is linked to a notably worse prognosis, with consistently lower response rates, event-free survival, and OS compared to dn-AML. Due to frequent poor-risk cytogenetics and high-risk molecular features like TP53 mutations, these leukemic clones are often inherently resistant to chemotherapy[29,30]. Additionally, patients are often older with complex medical histories, which restricts treatment options[31]. In our study, according to the results of the prognostic nomogram models, a history of non-primary tumors was identified as a risk factor for survival, consistent with previous studies.

Consistent with prior studies, our observations align with the indication that race serves as predictors of outcomes in AML patients[32,33]. However, compared to age group and chemotherapy, the weightage assigned to race in the nomogram was lower, indicating a relatively weaker predictive ability for these factors.

By incorporating four prognostic factors (age, race, first primary tumor, and chemotherapy), our study developed nomograms to estimate 1-, 3-, and 5-year OS and CSS for patients with t (9;11) (p22; q23) AML. To our knowledge, no specialized risk assessment model for t (9;11) (p22; q23) AML currently exists. Our nomograms demonstrated enhanced discrimination and calibration, and their potential clinical utility was confirmed in both validation and training cohorts. Additionally, categorizing patients into two distinct risk groups based on nomogram points allowed for precise identification of varying survival outcomes. Our nomograms offer clinicians a straightforward and effective tool for prognostic prediction in t (9;11) (p22; q23) AML.

This study has several limitations. First, the SEER database includes data from selected U.S. cancer registries, which may not fully represent the entire population. Additionally, socio-economic factors like access to care may introduce biases, and SEER does not capture all relevant variables, leading to potential residual confounding. Being retrospective, the study is prone to inherent biases, and missing clinicopathological and biological data may limit the inclusion of new prognostic factors. Furthermore, the SEER database lacks detailed information on relapses, time to relapse, and chemotherapy types, which may affect the analysis. These limitations should be considered when interpreting the findings. The findings of this study are based on a single cohort from the SEER database, which represents only 20–30% of the U.S. population, limiting generalizability. Further validation in independent cohorts, especially from diverse settings and countries with universal healthcare, is needed to confirm the external validity of these results.

## Conclusion

In summary, we have pioneered the development and validation of novel nomograms for t (9;11) (p22; q23) AML patients, being the first to use the SEER database for this purpose. These nomograms offer clinicians a dependable and accurate tool for survival prediction in t (9;11) (p22; q23) AML. Future prospective studies are needed to investigate new prognostic models that include emerging markers.

## Data availability

Publicly available datasets were analyzed in this study. These data are available at www.seer.cancer.gov./.

Received: 4 October 2024; Accepted: 1 January 2025
Published online: 07 January 2025

## References

1. Bill, M. et al. Mutational landscape and clinical outcome of patients with de novo acute myeloid leukemia and rearrangements involving 11q23/KMT2A. *Proc. Natl. Acad. Sci. USA* **117**, 26340–26346. https://doi.org/10.1073/pnas.2014732117 (2020).
2. Heuts BMH, Arza-Apalategi S, Alkema SG, et al (2023) Inducible MLL-AF9 Expression Drives an AML Program during Human Pluripotent Stem Cell-Derived Hematopoietic Differentiation
3. Takahashi, T. et al. Extramedullary relapse of AML with t(9;11)(p22;q23) associated with clonal evolution from trisomy 8 into tetrasomy 8. *Intern. Med.* **49**(5), 447–451. https://doi.org/10.2169/internalmedicine.49.2904 (2010).
4. Tirado, C. A. et al. A t(9;11)(p22;q23) within the context of a complex karyotype is associated with a poor prognosis in a 19-year-old patient with AML. *J. Assoc. Genet. Technol.* **47**, 78–83 (2021).
5. Baer, M. R. et al. Acute myeloid leukemia with 11q23 translocations: Myelomonocytic immunophenotype by multiparameter flow cytometry. *Leukemia* **12**, 317–325. https://doi.org/10.1038/sj.leu.2400933 (1998).

6. 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).
7. Hossain, M. J., Xie, L. & Caywood, E. H. Prognostic factors of childhood and adolescent acute myeloid leukemia (AML) survival: Evidence from four decades of US population data. *Cancer Epidemiol.* **39**, 720–726. https://doi.org/10.1016/j.canep.2015.06.009 (2015).
8. Harrison, C. J. et al. A multicenter evaluation of comprehensive analysis of MLL translocations and fusion gene partners in acute leukemia using the MLL FusionChip device. *Cancer Genet. Cytogenet.* **173**, 17–22. https://doi.org/10.1016/j.cancergencyto.2006.09.006 (2007).
9. Bianco, F. J. Nomograms and medicine. *Eur. Urol.* **50**, 884–886. https://doi.org/10.1016/j.eururo.2006.07.043 (2006).
10. Clegg, L. X. et al. Impact of socioeconomic status on cancer incidence and stage at diagnosis: Selected findings from the surveillance, epidemiology, and end results: National Longitudinal Mortality Study. *Cancer Causes Control* **20**, 417–435. https://doi.org/10.1007/s10552-008-9256-0 (2009).
11. Shallis, R. M., Weiss, J. J., Deziel, N. C. & Gore, S. D. Challenging the concept of *de novo* acute myeloid leukemia: Environmental and occupational leukemogens hiding in our midst. *Blood Rev.* **47**, 100760. https://doi.org/10.1016/j.blre.2020.100760 (2021).
12. Goel, H. et al. Molecular and genomic landscapes in secondary & therapy related acute myeloid leukemia. *Am. J. Blood Res.* **11**, 472–549 (2021).
13. DiNardo, C. D., Erba, H. P., Freeman, S. D. & Wei, A. H. Acute myeloid leukaemia. *Lancet* **401**, 2073–2086. https://doi.org/10.1016/S0140-6736(23)00108-3 (2023).
14. Zhang, S.-L. et al. Incidence, prognostic factors and survival outcome in patients with primary hepatic lymphoma. *Front. Oncol.* **10**, 750. https://doi.org/10.3389/fonc.2020.00750 (2020).
15. Zuo, Z. et al. Survival nomogram for stage IB non-small-cell lung cancer patients, based on the SEER database and an external validation cohort. *Ann. Surg. Oncol.* **28**, 3941–3950. https://doi.org/10.1245/s10434-020-09362-0 (2021).
16. Xing, Z. et al. Incidence, clinical characteristics, and prognostic nomograms for patients with myeloid sarcoma: A SEER-based study. *Front. Oncol.* **12**, 989366. https://doi.org/10.3389/fonc.2022.989366 (2022).
17. Juliusson, G. et al. Age and acute myeloid leukemia: Real world data on decision to treat and outcomes from the Swedish Acute Leukemia Registry. *Blood* **113**, 4179–4187. https://doi.org/10.1182/blood-2008-07-172007 (2009).
18. Forsberg, M. & Konopleva, M. AML treatment: Conventional chemotherapy and emerging novel agents. *Trends Pharmacol. Sci.* **45**, 430–448. https://doi.org/10.1016/j.tips.2024.03.005 (2024).
19. Tang, K., Schuh, A. C. & Yee, K. W. 3+7 combined chemotherapy for acute myeloid leukemia: Is it time to say goodbye?. *Curr. Oncol. Rep.* **23**, 120. https://doi.org/10.1007/s11912-021-01108-9 (2021).
20. Lin, T. L. & Pagano, L. The important role of intensive induction chemotherapy in the treatment of acute myeloid leukemia. *Expert Rev. Hematol.* **14**, 303–314. https://doi.org/10.1080/17474086.2021.1886920 (2021).
21. Stone, R. M. Consolidation chemotherapy for adults with AML in first remission: Is there a best choice?. *J. Clin. Oncol.* **31**, 2067–2069. https://doi.org/10.1200/JCO.2013.48.6886 (2013).
22. Saygin, C. & Carraway, H. E. Emerging therapies for acute myeloid leukemia. *J. Hematol. Oncol.* **10**, 93. https://doi.org/10.1186/s13045-017-0463-6 (2017).
23. Yang, X. & Wang, J. Precision therapy for acute myeloid leukemia. *J. Hematol. Oncol.* **11**, 3. https://doi.org/10.1186/s13045-017-0543-7 (2018).
24. Juliusson, G. et al. Acute myeloid leukemia in the real world: Why population-based registries are needed. *Blood* **119**, 3890–3899. https://doi.org/10.1182/blood-2011-12-379008 (2012).
25. Palle, J. et al. Cellular drug sensitivity in MLL-rearranged childhood acute leukaemia is correlated to partner genes and cell lineage. *Br. J. Haematol.* **129**, 189–198. https://doi.org/10.1111/j.1365-2141.2005.05433.x (2005).
26. Zwaan, C. M. et al. Cellular drug resistance in childhood acute myeloid leukemia is related to chromosomal abnormalities. *Blood* **100**, 3352–3360. https://doi.org/10.1182/blood.V100.9.3352 (2002).
27. Rubnitz, J. E. et al. Favorable impact of the t(9;11) in childhood acute myeloid leukemia. *J. Clin. Oncol.* **20**, 2302–2309. https://doi.org/10.1200/JCO.2002.08.023 (2002).
28. Lie, S. O. et al. Treatment stratification based on initial in vivo response in acute myeloid leukaemia in children without Down's syndrome: Results of NOPHO-AML trials. *Br. J. Haematol.* **122**, 217–225. https://doi.org/10.1046/j.1365-2141.2003.04418.x (2003).
29. Cheung, E. et al. The leukemia strikes back: A review of pathogenesis and treatment of secondary AML. *Ann. Hematol.* **98**, 541–559. https://doi.org/10.1007/s00277-019-03606-0 (2019).
30. Secondary AML Emerging After Therapy with Hypomethylating Agents: Outcomes, Prognostic Factors, and Treatment Options - PubMed. Accessed 13 Dec 2024. https://pubmed.ncbi.nlm.nih.gov/33609248/.
31. Matthews, A. & Pratz, K. W. Optimizing outcomes in secondary AML. *Hematol. Am. Soc. Hematol. Educ. Progr.* **2022**, 23–29. https://doi.org/10.1182/hematology.2022000324 (2022).
32. Lamba, J. K. et al. Pharmacogenomics, race, and treatment outcome in pediatric acute myeloid leukemia. *JAMA Netw. Open* **7**, e2411726. https://doi.org/10.1001/jamanetworkopen.2024.11726 (2024).
33. Eisfeld, A.-K. Disparities in acute myeloid leukemia treatments and outcomes. *Curr. Opin. Hematol.* **31**, 58–63. https://doi.org/10.1097/MOH.0000000000000797 (2024).


## Acknowledgements

The authors would like to thank the Surveillance, Epidemiology, and End Results (SEER) database for the support provided. We thank Jie Liu, PhD (Department of Vascular and Endovascular Surgery, Chinese PLA General Hospital);Haibo Li, PhD (Capital Institute of Pediatrics in China); Dongjie He(Department of Radiation Oncology, Tangdu Hospital, the Second Affiliated Hospital of Air Force Military Medical University, China) and Jinbao Ma (Department of Drug-resistance tuberculosis, Xi'an Chest Hospital, Xi'an, Shaanxi Province, China)for their invaluable assistance with statistical analysis, manuscript review, and commentary.

## Author contributions

GL: data analysis and writing; DZ: data collection; YF review and final approval. All authors read and approved the final manuscript.

## Funding

Henan Province Science and Technology Research and Development Plan (Nos. LHGJ20220188, SBGJ202102069), Henan Province Young and Middle-aged Talent Plan Project (No. YXKC2022046).

## Declarations

### Ethics approval and consent to participate
The data analyzed and used in this study was obtained from Surveillance, Epidemiology, and End Results (SEER) database in accordance with the SEER data use agreement. Therefore, this study did not require approval of ethical board.

### Competing interests
The authors declare no competing interests.

### Consent for publication
Not applicable.

### Additional information
**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-85310-w.

**Correspondence** and requests for materials should be addressed to Y.F.

**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.