


RESEARCH

Open Access



The molecular prognostic scoring system for normal karyotype myelodysplastic syndromes

Wei Wang^{1,2,3,4†}, Yudi Zhang^{1,2,3,4†}, Wenli Yang^{1,2,3,4†}, Xiaozhen Liu^{1,2,3,4}, Lingxu Jiang^{1,2,3,4}, Wei Lang^{1,2,3,4}, Yingwan Luo^{1,2,3,4}, Xinping Zhou^{1,2,3,4}, Lu Wang^{1,2,3,4}, Li Ye^{1,2,3,4}, Ying Xu^{1,2,3,4}, Liya Ma^{1,2,3,4} and Hongyan Tong^{1,2,3,4*} 

Abstract

Background Molecular-clinical prognostic models for Myelodysplastic syndromes (MDS) offer more accurate prognosis predictions, yet existing models often overlook the heterogeneity of mutational profiles against the cytogenetic background. Moreover, how to apply these models in regions where large panel NGS is unaffordable remains a significant challenge to be addressed.

Methods A total of 237 NK MDS patients from our center were used as the training set to screen for key variables and develop a prognostic model with overall survival (OS) as the endpoint. The C-index was used as the main evaluation metric to assess the model's performance. The IWG-PM cohort ($n = 691$) was used as an external independent validation set to evaluate the generalizability of the model.

Results We developed a seven-parameter molecular-clinical prognostic model, the Molecular Prognostic Scoring System for NK MDS (NK-PSS-M), which only incorporates three gene mutations as parameters. The NK-PSS-M can reliably predict OS and leukemia-free survival (LFS). The performance of NK-PSS-M was comparable to that of the Molecular International Prognostic Scoring System (IPSS-M), and it significantly outperformed the Revised International Prognostic Scoring System for MDS (IPSS-R).

Conclusions The NK-PSS-M model improved the risk stratification of non-molecular models and provided a reliable alternative to the IPSS-M. This strategy provides insights into how resource-scarce regions can apply molecular-clinical models.

[†]Wei Wang, Yudi Zhang, and Wenli Yang contributed equally to this work.

*Correspondence:

Hongyan Tong
tonghongyan@zju.edu.cn

¹Department of Hematology, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

²Myelodysplastic Syndromes Diagnosis and Therapy Center, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

³Cancer Center, Zhejiang University, Hangzhou, Zhejiang 310058, China

⁴Zhejiang Provincial Key Lab of Hematopoietic Malignancy, Zhejiang University, Hangzhou, Zhejiang, PR China



Background

Myelodysplastic syndrome (MDS), a heterogeneous clonal hematopoietic disorder, is characterized by cytopenia, lineages dysplasia, and an increased risk of transformation to acute myeloid leukemia [1]. MDS patients exhibit significant heterogeneity in clinical manifestations and prognoses; some experience rapid deterioration within months of diagnosis, whereas others may live with the disease for decades [2]. Therefore, accurate predictive models are essential for the management of MDS which can assess disease progression at the initial stage of diagnosis and assist clinicians in formulating clinical treatment strategies [3, 4].

Unlike the International Prognostic Scoring System (IPSS) and the Revised IPSS (IPSS-R) which was based on cytopenia, BM blast burden, and cytogenetic abnormalities, modern prognostic models include molecular features as key variables [5, 6]. Integrating molecular factors into non-molecular prognostic models like IPSS or IPSS-R can significantly enhance predictive accuracy. For instance, Nazha et al. [7] combined age, *TP53*, *SF3B1*, and *EZH2* with the IPSS-R score, improving the model's predictive performance (concordance index [C-index] = 0.73 vs. 0.69 for the IPSS-R). The molecular IPSS (IPSS-M) integrates clinical features, cytogenetic data, and information of 31 gene mutations reclassifying nearly half of IPSS-R model patients into more appropriate risk categories [8].

However, next-generation sequencing (NGS) is an expensive process, which can pose significant challenges for resource-limited centers [6]. Moreover, fixed and limited panels may not provide sufficient data to meet the requirements of the molecular model and can inevitably lead to missing data. Although the IPSS-M has addressed this issue and introduced targeted remedies, the accuracy of model predictions can be significantly affected by gene deletions, with accuracy dropping below 50% when deletion numbers are high [9]. There is an urgent need for reliable molecular-clinical prognostic models that can be applied in resource-limited situations.

Another issue worth noting is that models developed have often been constructed based on either all MDS patients or subgroups categorized based on the IPSS-R, which may lead to the oversight of cytogenetic background in the identification of genes as model parameters. In our previous investigation focusing on the molecular profiles of 928 cases of normal karyotype (NK) MDS across two cohorts, we unveiled distinct patterns of gene mutation frequency and prognostic outcomes among these patients. For instance, *TP53* gene mutation is commonly regarded as an adverse prognostic indicator and frequently incorporated into various molecular-clinical prognostic models for MDS, while in the context of normal karyotype MDS, it exhibited lower mutation

incidence, reduced variant allele frequency (VAF), diminished occurrence of multi-hit events, and lacked prognostic relevance [10–12]. Accordingly, the prognostic mutation spectrum identified in the broader MDS population may not reflect the specific mutation patterns observed in the subset of patients with normal karyotype [13, 14]. Incorporating molecular features that do not accurately represent the subset of patients with normal karyotype into the MDS prognostic model may lead to decreased prediction precision and could inadvertently result in the misallocation of resources.

Drawing on insights from our comprehensive exploration of the unique molecular characteristics of NK MDS, we constructed a lightweight molecular-clinical model, NK-PSS-M, which is more likely to be applied in resource-limited centers. This model has sufficient capability to serve as a reliable alternative to IPSS-M when its application is limited.

Materials and methods

Patient cohort, genomic information acquisition, and analysis

The study was conducted under the approval of the Ethics Committee of the First Affiliated Hospital of Zhejiang University and in accordance with the Declaration of Helsinki, with informed consent obtained from all participants. The baseline characteristics of our patient cohort, genomic information acquisition, and analysis methodologies were extensively described in our previous publications [14]. Accordingly, we enrolled 237 de novo NK MDS patients based on WHO 2016 criteria [15] who underwent comprehensive pre-treatment examinations including NGS, bone marrow (BM) examination, cytogenetic, and hematological analysis to confirm MDS diagnosis. To ensure diagnostic accuracy, at least 20 metaphase spreads were analyzed using G-banding or R-banding to detect the presence of NK in MDS [16, 17]. For external validation, data from NK primary MDS patients from the International Working Group for the Prognosis of MDS (IWG-PM) cohort of the IPSS-M study, available through the cBioPortal platform, were used (<https://www.cbioportal.org/>). All patients included in this study were treatment-naïve and had primary normal karyotype MDS. The patient selection process was not influenced by clinical characteristics, treatment, or other biases. The parameters for our prognostic model—patient age, peripheral blood cell count, bone marrow blast percentage, and gene mutation data—were chosen for their widespread availability in clinical practice, ensuring the model's ease of implementation. Genomic analysis entailed targeted NGS on individual BM cells with a predefined gene list, identifying pathogenic mutations with a variant allele frequency > 2%. Genes with mutation frequencies < 1% were excluded, culminating in

a focused list of 23 significant genes including *SETBP1*, *U2AF1*, *ASXL1*, *CBL*, *CEBPA*, *CSF3R*, *DNMT3A*, *ETV6*, *EZH2*, *IDH1*, *IDH2*, *KRAS*, *NPM1*, *NRAS*, *PHF6*, *PTPN11*, *RUNX1*, *SF3B1*, *SRSF2*, *TP53*, *WT1*, *ZRSR2*, and *TET2*.

Statistical analysis

Overall survival (OS) represents the duration from the initial diagnosis to either the date of death or the most recent follow-up. Leukemia-free survival (LFS) measures the time from the point of diagnosis to the occurrence of leukemia transformation or death due to any reason. Survival probabilities were evaluated using Kaplan–Meier estimates, and the statistical significance of differences was analyzed with the log-rank test. To normalize the distribution of skewed platelet (PLT) count data, we applied a natural logarithm transformation; no modifications were necessary for the other hematological data. Age was categorized using a 60-year threshold.

In total, 29 variables were included in the analysis, with age and mutation status being binary variables and the remaining variables being continuous variables. A fixed random seed was set to ensure the reproducibility of the analysis. Preliminary variable filtering for factors affecting OS was performed using univariate Cox regression, and variables with a P value < 0.05 were considered as candidates for analysis. To develop a parsimonious model, we used the least absolute shrinkage and selection operator (LASSO) algorithm to filter and select variables; this algorithm can exclude weaker variables by increasing the penalty while retaining stronger variables. The R package “glmnet” [18] was used to perform LASSO regression and select variables associated with OS, with 10-fold cross-validation used to estimate the LASSO penalty weight (λ). The λ value (lambda.min) corresponding to the minimum mean squared error was selected. The resulting parameters were used for stepwise Cox regression to obtain the final model and regression coefficients. Time-dependent receiver operating characteristic (ROC) curves were generated for different prediction times. Harrell’s C -index evaluated model discriminatory ability and the R package “compareC” was used to test for differences among models. Calibration curves were used to evaluate the consistency between predicted and observed probabilities. Decision curve analysis (DCA) evaluated the clinical effectiveness of the models. Categorical covariates were analyzed using Fisher’s test or χ^2 test. We used simplified P values for some results, where “ns” indicates not significant and *, **, ***, and **** indicate P values of < 0.05 , < 0.01 , < 0.001 , and < 0.0001 , respectively. Statistical analyses were conducted, and figures were created, using R software (version 4.2.1; R Development Core Team, Vienna, Austria).

Results

Patient cohorts and genetic profiles

In keeping with our earlier descriptions, the development of our model utilized two cohorts: a derivation set comprising 237 cases from our center and an external validation set of 691 NK MDS patients from the IWG-PM cohort. These cohorts showed typical geographical variations in aspects such as patient age, bone marrow blast percentage, and cytopenias (Table S1) [14]. Broadly, the profiles of the top 10 genes by mutation frequency showed general similarity across both cohorts (Fig. S1A, B). Additionally, a discernible decline in *TP53* mutation frequency was observed, consistent with trends noted in our previous research.

Variable selection and model construction

Univariate Cox regression analysis revealed that, of the Age, $\ln(\text{PLT})$, HB, BM Blasts, and seven genes (*CEBPA*, *EZH2*, *NPM1*, *NRAS*, *RUNX1*, *U2AF1*, and *SRSF2*) that correlated with OS, only $\ln(\text{PLT})$ and hemoglobin (HB) were protective factors; all other clinical features and genes were identified as risk factors (Fig. 1A). The absolute neutrophil count was not significantly associated with OS and LFS, which is consistent with IPSS-M, and therefore was excluded from further analysis. Among the seven genes, the *CEBPA* mutation was significantly associated with OS and had the highest HR (4.97, 95% CI: 2.45–10.1, $P < 0.001$).

The features with $P < 0.05$ were further subjected to LASSO analysis in the univariate Cox regression analysis. While both “lambda.1se” and “lambda.min” were provided as potential λ value, we selected “lambda.min” as our criterion considering that “lambda.1se” would exclude PLT from the model which has been recognized as a crucial clinical parameter in MDS and has been widely incorporated into well-established prognostic scoring systems such as IPSS-R and IPSS-M. Under “lambda.min”, Age, $\ln(\text{PLT})$, HB, BM Blasts, *CEBPA*, *U2AF1*, *RUNX1*, and *NRAS* were identified as significant risk predictors (Fig. S1C–D). The backward stepwise Cox regression analysis retained the seven parameters except *NRAS*. The final model included age, $\ln(\text{PLT})$, HB, BM Blasts, *CEBPA*, *U2AF1*, and *RUNX1* as risk predictors. Based on the regression coefficients, a risk score was calculated using the following formula: Risk score = age (≥ 60) * 0.84 + BM Blasts * 0.1 – HB * 0.15 – $\ln(\text{PLT})$ * 0.27 + *CEBPA* (mutation) * 0.84 + *U2AF1* (mutation) * 0.94 + *RUNX1* (mutation) * 0.55 (Fig. 1B). Using maximally selected rank statistics, a “– 1.1” risk score was identified as the optimal cut-off value. High- and low-risk patients had values higher and lower than the cut-off value, respectively (Fig. 1C). Next, we divided patients into very high- (risk score > 0 , 17%), high- (risk score > -1.1 , 25%), intermediate- (risk score > -1.7 ,

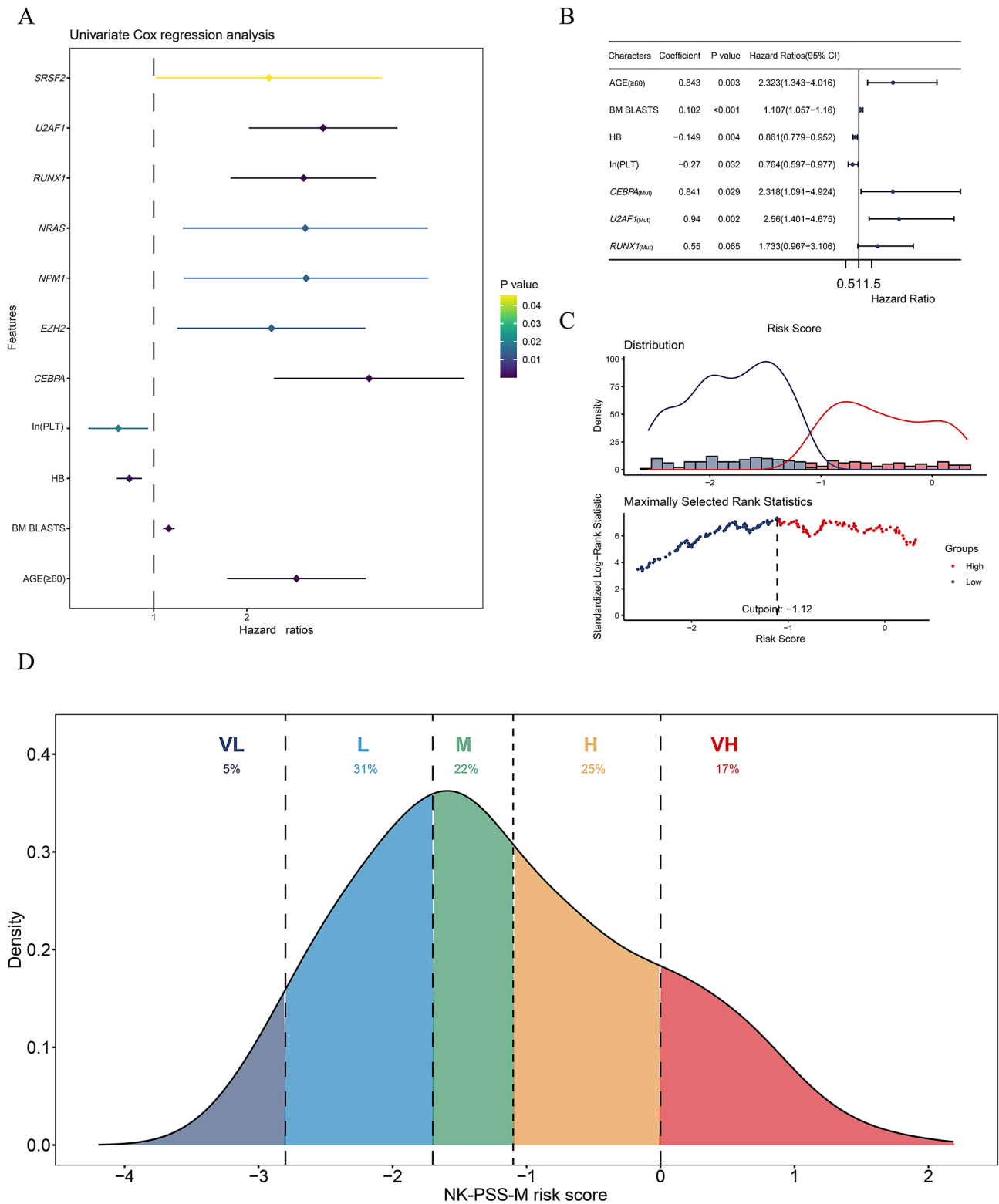


Fig. 1 Development and risk classification of NK-PSS-M. **(A)** Univariate Cox regression forest plot, with colors indicating the significance of the variables. **(B)** Multivariate Cox regression forest plot, with diamonds representing hazard ratios (HRs) and line segments representing confidence intervals. **(C)** The optimal risk score cut-off value is calculated using maximally selected rank statistics to classify patients into high- and low-risk groups. **(D)** Density plot of the NK MDS risk scores from 237 patients in the training set. The x-axis represents the risk score calculated by NK-PSS-M; the vertical dashed line represents the cut-off value. The groups are represented by their abbreviated names, and the numbers below indicate the proportion of each group. VL, very low; L, low; M, intermediate; H, high; VH, very high

22%), low- (risk score > -2.8 , 31%), and very low- (risk score ≤ -2.8 , 5%) risk groups (Fig. 1D). Prognoses varied significantly among these different groups, with the very low- and low-risk groups not reaching median survival. Conversely, the intermediate-, high-, and very high-risk groups had a median survival of 6.46 (95% CI: 6.142–NA), 2.81 (95% CI: 1.8–NA), and 1.21 (95% CI: 0.926–1.92) years, respectively, ($P < 0.0001$; Fig. 2A). The 5-year time-dependent area under the ROC curve values were 0.78, 0.86, 0.86, 0.86, and 0.89, respectively (Fig. 2B). We applied the same parameters to calculate the risk score of all patients in the validation set and assigned patients to the same risk groups based on the thresholds.

According to the survival analysis results, the model maintained excellent discriminatory ability in the validation cohort. The very low-risk group did not reach median survival, whereas the low-, intermediate-, high-, and very high-risk groups had median survival times of 7.18 (95% CI: 5.91–9.03), 4.71 (95% CI: 3.46–7.62), 2.45 (95% CI: 2.02–3.38), and 1.42 (95% CI: 1.1–1.90) years, respectively ($P < 0.0001$). The 5-year time-dependent area under the ROC curve values were 0.7, 0.74, 0.74, 0.73, and 0.74, respectively (Fig. S2A, B).

In the training set, the very low-risk group not reach median LFS, whereas the low-, intermediate-, high-, and very high-risk groups had median LFS times of 8.92 (95% CI: 7.384–NA), 6.46 (95% CI: 6.142–NA), 2.55 (95% CI: 1.233–NA), and 1.06 (95% CI, 0.863–1.63) years, respectively ($P < 0.0001$) (Fig. 3A). In the validation set, the median LFS time was not reached for very low-risk group, whereas the low-, intermediate-, high-, and very high-risk groups had median LFS times of 6.68 (95% CI: 5.814–9.03), 3.33 (95% CI: 2.751–6.55), 1.82 (95% CI: 1.507–2.45), and 1.03 (95% CI: 0.638–1.61) years, respectively ($P < 0.0001$) (Fig. S3A). The 5-year time-dependent area under ROC curve values were > 0.7 for both cohorts (Fig. 3B; Fig. S3B).

Model performance

In addition to the IPSS-M and IPSS-R models, we compared two additional models in this study, the IPSS-R Age-adjusted (referred to as IPSS-R-Age below) and Nazhe et al. (referred to as Nazha-2016 below). Patients with Moderate Low and Moderate High-risk levels in the IPSS-M model were combined as Intermediate to create a five-category variable, as NK-PSS-M, IPSS-R, and IPSS-R-Age are five-category models. The predictive ability of all models for LFS and OS were evaluated (Figs. 2 and 3; Fig. S2–S4). Moreover, we evaluated the ability of four five-category models (except Nazha-2016) to predict clinical outcomes (Table S2 and S3).

For both OS and LFS, NK-PSS-M and IPSS-M improved the poor differentiation of IPSS-R/IPSS-R-Age for higher-risk patients (Very High and High).

NK-PSS-M and IPSS-M had similar time-dependent area under ROC curve values (0.78–0.89 and 0.76–0.85, respectively), which were higher than those of IPSS-R (0.71–0.8) (Fig. 2; Fig. S5A). However, after adjustment for age, the gap between IPSS-R-Age, NK-PSS-M, and IPSS-M was reduced (0.74–0.85; Fig. 2H). The C-index for NK-PSS-M (0.776) was not significantly different from that of IPSS-M (0.75, $P = 0.264$); however, it was significantly higher than the C-index of IPSS-R-Age (0.728, $P = 0.0244$), IPSS-R (0.695, $P < 0.0001$), and Nazha-2016 (0.7, $P < 0.001$) (Fig. 4A).

In the validation set, NK-PSS-M demonstrated high discriminatory ability for the OS and LFS in risk-stratified patients and improved the poor differentiation of IPSS-R and IPSS-R-Age for high-risk patients (Fig. S2, 3). Furthermore, the C-index of NK-PSS-M (0.692) was not significantly different from that of IPSS-M ($P = 0.693$); however, it was significantly higher than the C-index of IPSS-R (0.665, $P = 0.0101$) and Nazha-2016 (0.631, $P < 0.0001$). Although the difference was not statistically significant, the C-index of IPSS-R-Age (0.678) remained lower than that of NK-PSS-M ($P = 0.141$) (Fig. 4A).

The calibration curve demonstrated that the NK-PSS-M maintained a consistent relationship between the predicted and observed OS for NK MDS patients in the training and validation cohorts (Fig. S5B, C). Compared with IPSS-M and IPSS-R, NK-PSS-M showed comparable or better calibration across different time points (1-year, 3-year, and 5-year OS) (Fig. S5D–I). In addition, decision curve analysis was conducted to evaluate the clinical effectiveness of the five models. The findings revealed that NK-PSS-M and IPSS-M had similar benefits at various risk thresholds and outperformed IPSS-R, IPSS-R-Age, and Nazha-2016 (Fig. 4B, C).

Restratification of patients from IPSS-R and IPSS-M to NK-PSS-M

We compared the efficacy of the novel NK-PSS-M risk classification systems with the established IPSS-M and IPSS-R systems. The NK-PSS-M system reclassified a significant proportion of patients in the IPSS-M and IPSS-R risk categories (44% and 50%, respectively; Fig. 5A, B); similar reclassification proportions were observed in the validation cohort (53% and 42% of patients, respectively; Fig. S6A, B). Although there was substantial reclassification of patients, no statistically significant difference was observed in performance between NK-PSS-M and IPSS-M (Fig. 5C, D; Fig. S6C, D). However, the NK-PSS-M demonstrated superiority over the IPSS-R in reclassifying patients with a briefer OS among those classified as intermediate- or low-risk by IPSS-R in the training and validation cohorts (Fig. 5E; Fig. S6E). IPSS-R was ineffective in reclassifying patients based on NK-PSS-M (Fig. 5F; Fig. S6F).

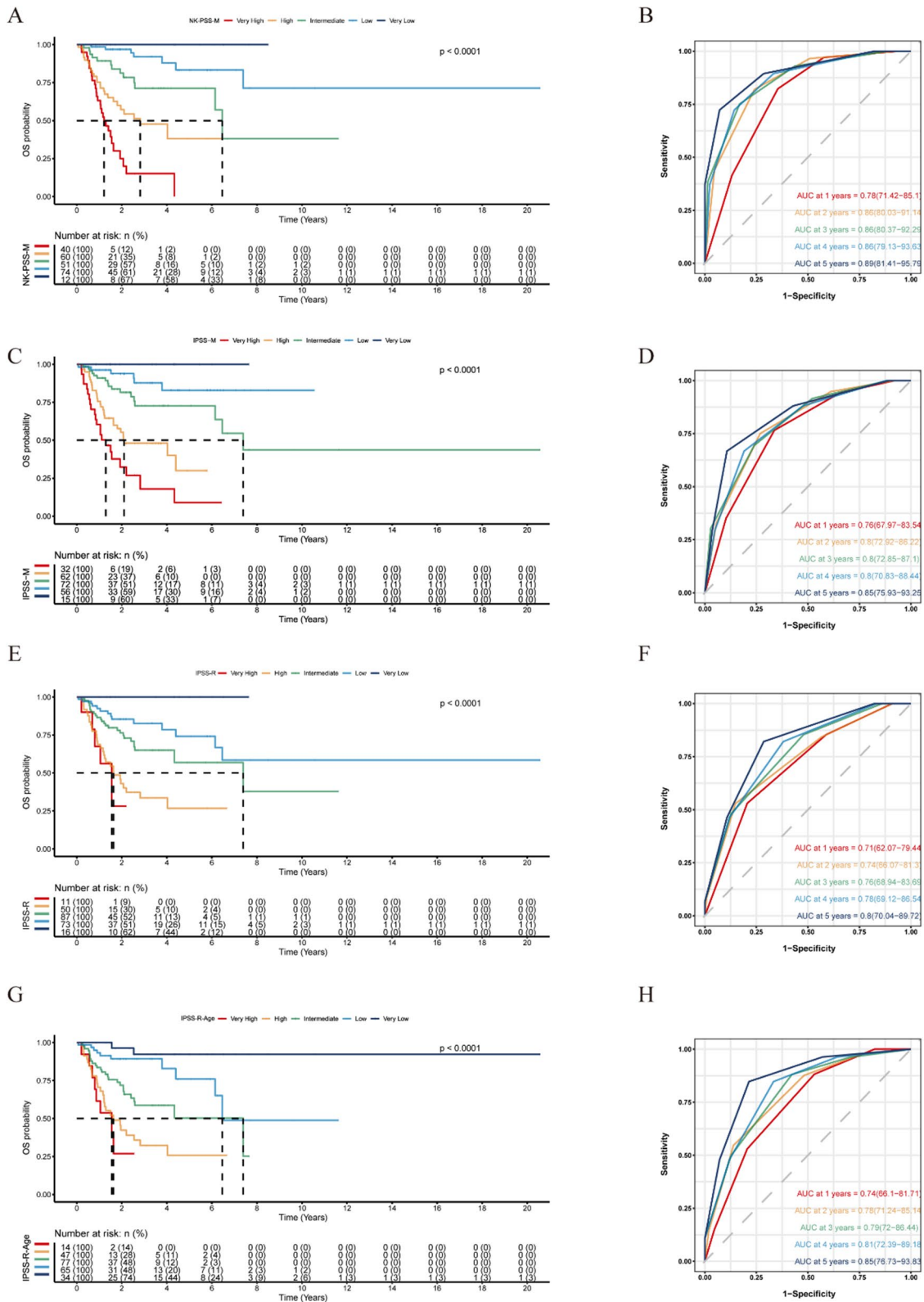


Fig. 2 Overall survival (OS) in the training cohort according to four models. (A–B) Kaplan–Meier probability estimates of OS are presented across NK-PSS-M risk categories and in time-dependent receiver operating characteristic (ROC) curves. (C–D) IPSS-M risk categories. (E–F) IPSS-R risk categories. (G–H) IPSS-R Age-adjusted (IPSS-R-Age) risk categories

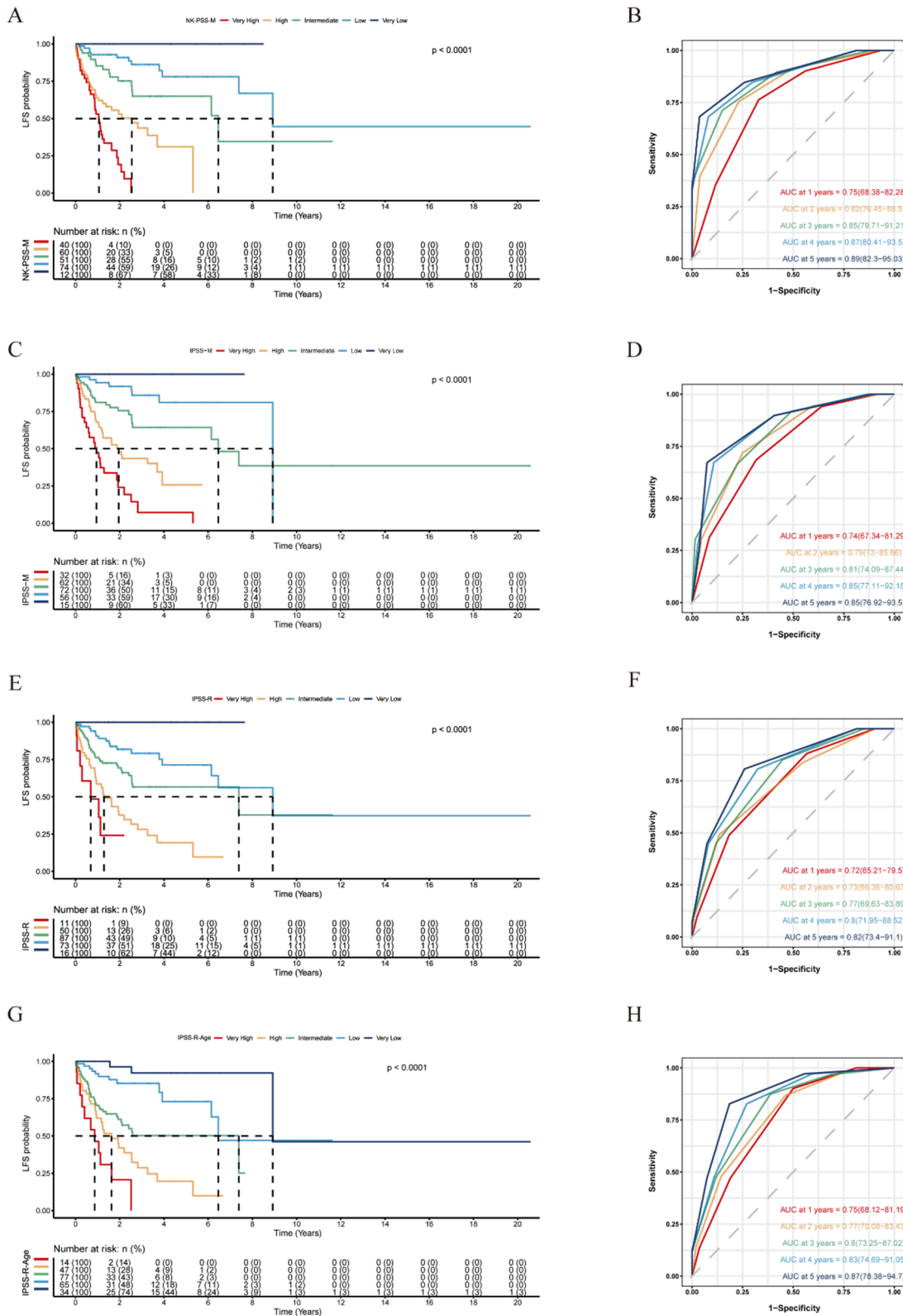


Fig. 3 Leukemia-free survival (LFS) in the training cohort according to four models. (A–B) Kaplan–Meier probability estimates of LFS are presented across NK-PSS-M risk categories and in time-dependent receiver operating characteristic (ROC) curves. (C–D) IPSS-M risk categories. (E–F) IPSS-R risk categories. (G–H) IPSS-R Age-adjusted (IPSS-R-Age) risk categories

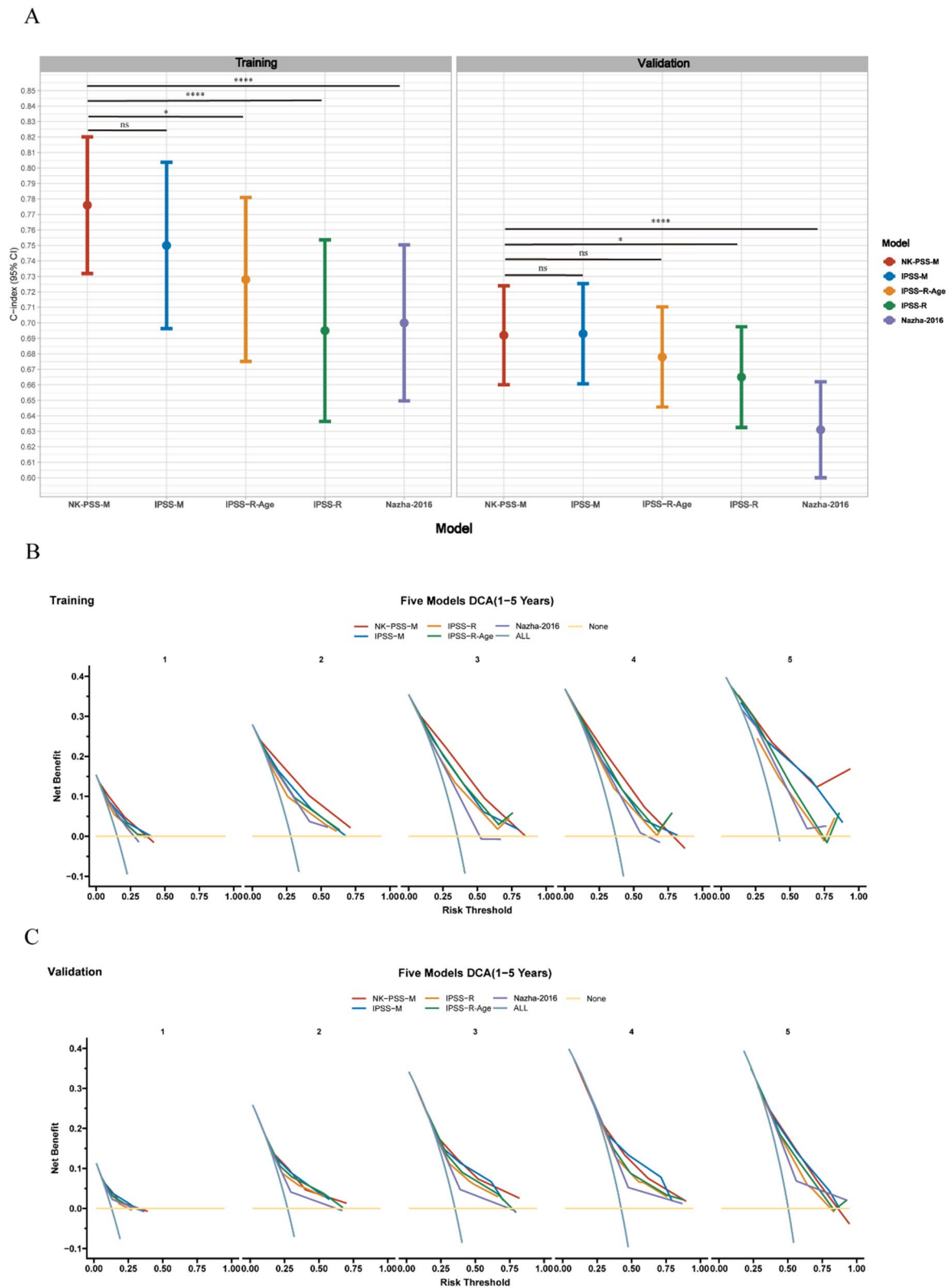


Fig. 4 Comparison of the NK-PSS-M model with IPSS-M, IPSS-R-Age, IPSS-R, and Nazha-2016. **(A)** The concordance index (C-index) of the five prognostic models and the significance of the differences between NK-PSS-M and the other four models were calculated in the training and validation cohorts. **(B–C)** Decision curve analysis of the five prognostic models in the training and validation cohorts, respectively. Each color represents a different model, with red representing NK-PSS-M, blue representing IPSS-M, orange representing IPSS-R, green representing IPSS-R-Age, and purple representing Nazha-2016

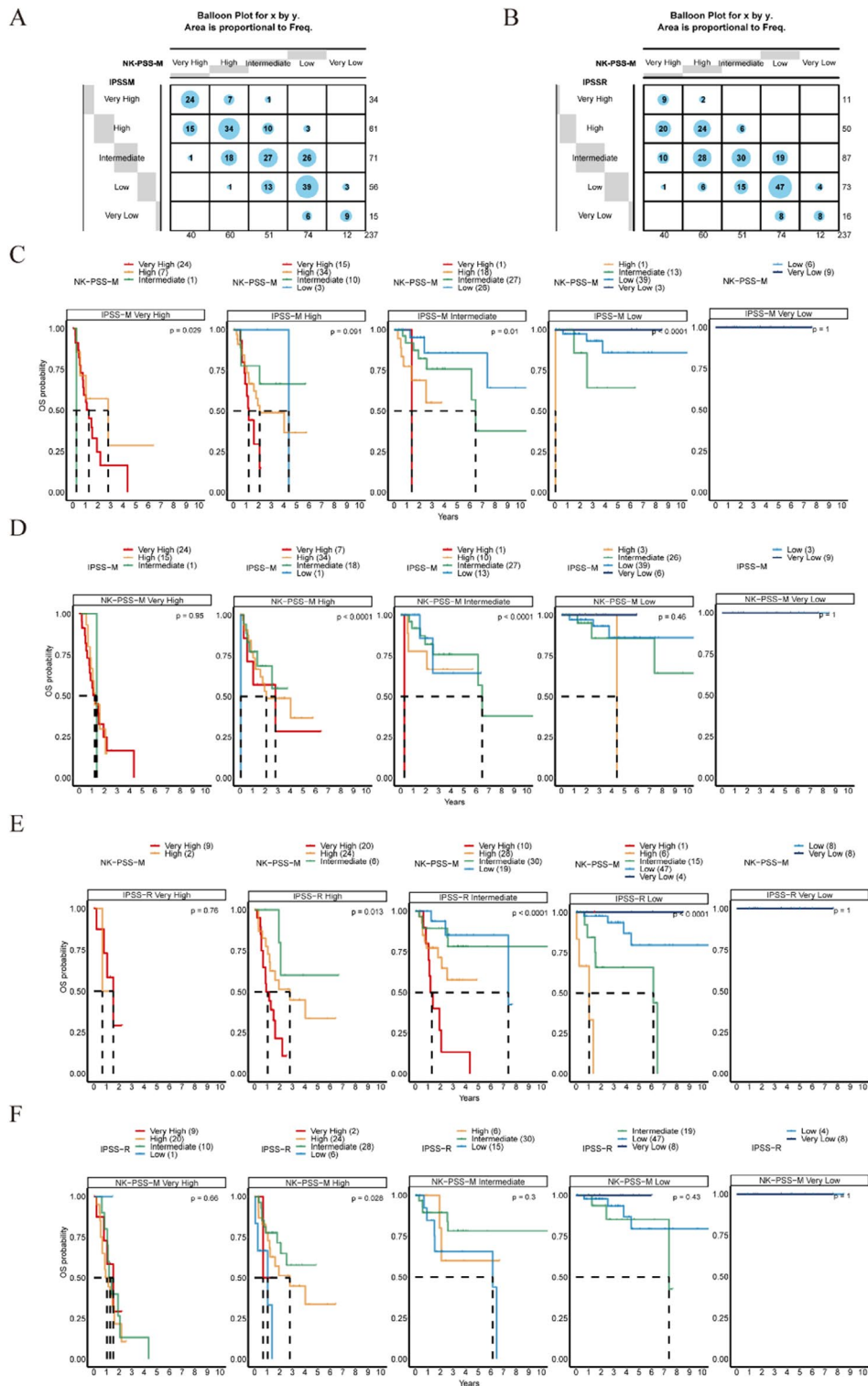


Fig. 5 Restratification of patients from IPSS-M and IPSS-R to NK-PSS-M in the training cohort. **(A)** Balloon plot shows the number of patients reclassified in each of the five IPSS-M (row) and five NK-PSS-M (column) categories. **(B)** The balloon plot shows the number of patients reclassified in each of the five IPSS-R (row) and five NK-PSS-M (column) categories. **(C)** The Kaplan-Meier (KM) curve evaluates overall survival (OS) based on the NK-PSS-M classification within each IPSS-M category. **(D)** The KM curve evaluates OS based on the IPSS-M classification within each NK-PSS-M category. **(E)** The KM curve evaluates the OS based on the NK-PSS-M classification within each IPSS-R category. **(F)** The KM curve evaluates the OS based on the IPSS-R classification within each NK-PSS-M category

Among the patients who were re-stratified by the NK-PSS-M, we observed that those who were upstaged were generally older, while those who were downgraded were generally younger in both cohorts (Fig. 6A; Fig. S8E). Furthermore, patients who were upstaged tended to have more gene mutations identified by the NK-PSS-M model, with the occurrence of ≥ 2 gene mutations being almost exclusively found in the (very) high-risk groups, suggesting that the occurrence and accumulation of gene mutations have a significant impact on the prognosis (Fig. 6B; Fig. S8F).

Our analysis reveals that BM blasts, Hb, PLT, and ANC concentrations indeed play roles in the NK-PSS-M's re-stratification process, particularly for patients categorized as (very) low risk in IPSS-R. However, for those at a relatively higher risk, the NK-PSS-M model distinctly emphasizes the significance of age and gene mutations, offering a more nuanced consideration of these factors (Fig. S7A-D and S8A-D).

We assessed the performance of different prognostic models within specific categories of bone marrow blast cell percentages. Our findings indicated that for normal karyotype MDS patients with low blasts (especially blasts $\leq 2\%$), both the NK-PSS-M and IPSS-M models can effectively distinguish patients with potential higher risk, who are often categorized as lower risk according to the IPSS-R model (Fig. 6C-E; Fig. S7E-G). Further analysis of model discrimination in different blast groups showed that NK-PSS-M exhibited superior discriminative ability in patients with blasts $\leq 2\%$ and maintained comparable performance with IPSS-M in patients with blasts $> 2\%$, while both models demonstrated better discrimination than IPSS-R and Nazha-2016 (Fig. S9A-B).

Performance and validation of the simplified NK-PSS-M model

While the NK-PSS-M model is effective, it requires formulaic calculations to determine each patient's prognostic score. Here, we provide an alternative simplified version of the NK-PSS-M model to achieve comparable predictive efficacy (referred to as the NK-PSS-M easy model). The streamlined NK-PSS-M (easy model) classifies patients into five distinct prognostic categories—Very low, Low, Intermediate, High, and Very high—based on a composite score from previously described seven parameters (Fig. S10A). The simplified model maintains robust predictive performance in both the training and validation cohorts ($P < 0.001$ in all) (Fig. S10B). Moreover, the C-index of the NK-PSS-M easy model surpasses that of IPSS-R and is comparable to IPSS-M, with values for the training cohort at 0.767 for the easy model, 0.750 for IPSS-M, and 0.695 for IPSS-R; and for the validation cohort at 0.675 for the easy model, 0.693 for IPSS-M, and 0.665 for IPSS-R (Fig. 10C).

A practical application strategy for NK-PSS-M in clinical practice

To further validate the potential value of NK-PSS-M in real-world clinical practice, we tested the model's predictive efficacy under certain extreme scenarios. First, we created a panel that included only three genes: *CEBPA*, *U2AF1*, and *RUNX1*. Under this condition, the predictive results of IPSS-M were significantly affected, while NK-PSS-M appeared to handle this situation well (Fig. 7A-C).

Secondly, considering the genetic heterogeneity of MDS, it may lead to misjudgments when the number of genes included in the model is insufficient. Therefore, we evaluated the model's predictive performance in patients with mutations other than the three genes included in NK-PSS-M. The results demonstrated that NK-PSS-M retained excellent predictive performance for these patients, which was confirmed in the validation cohort (Fig. 7D, E).

Based on these findings, we believe that NK-PSS-M is a reliable molecular-clinical prognostic model capable of handling the complex situations encountered in real-world clinical practice. Here, we provide a recommended application strategy for clinical practice (Fig. 7F). For patients who can afford NGS testing, we suggest performing NGS at the initial diagnosis and primarily using IPSS-M for prognostic stratification. For NK MDS patients whose NGS panels cannot cover the genes required by IPSS-M, we recommend considering NK-PSS-M. For patients who cannot afford NGS, we advise using IPSS-R as the basis for diagnosis and treatment in non-normal karyotype MDS patients. For normal karyotype MDS patients, if IPSS-R indicates a higher-risk category (IPSS-R score > 3.5), no additional action is needed. However, for lower-risk patients (IPSS-R score ≤ 3.5) or those with blasts $\leq 2\%$, we strongly recommend performing mutation detection of the three genes (*CEBPA*, *RUNX1*, and *U2AF1*) to apply NK-PSS-M for prognostic assessment.

Discussion

MDS, a highly heterogeneous hematological disorder, demands accurate and reasonable stratification for optimal clinical management. Although the MDS prognostic model has entered the molecular era, clinicians rely primarily on the IPSS or IPSS-R. Multiple studies have demonstrated that integrating molecular information into the model can significantly enhance its predictive accuracy. Molecular prognostic models are in the early stages of development and have limited validity, and additional concerns must also be considered. These include the cost of implementing the model and the suitability of the local NGS panels at medical centers to meet the model requirements [6].

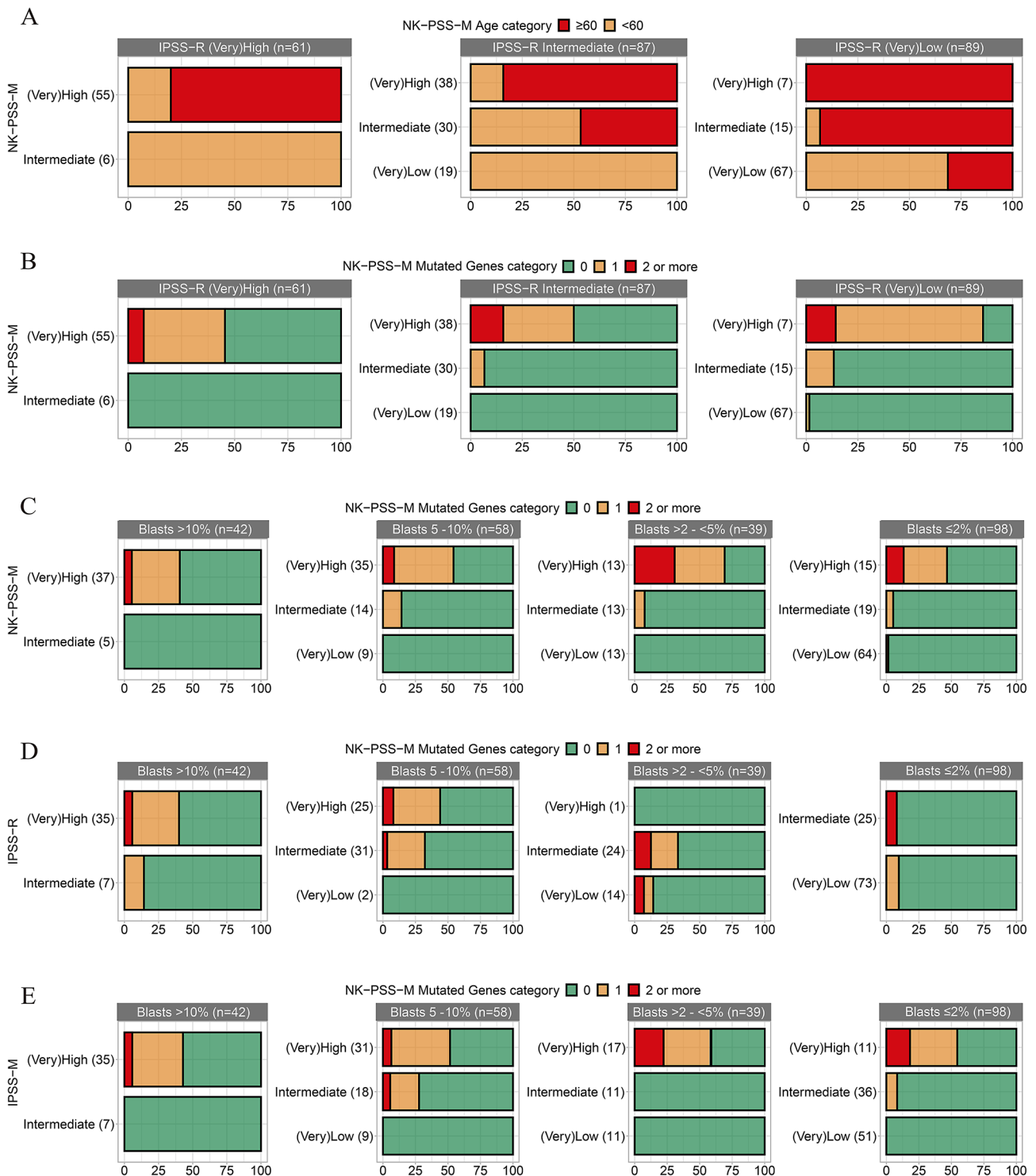


Fig. 6 NK-PSS-M parameters and their impact on risk stratification in MDS. Utilizing Simplified Risk Category Models from NK-PSS-M, IPSS-M, and IPSS-R. (A-B) Relationship between age (A) and the number of mutated NK-PSS-M genes (B) with patient restratification. (C-E) Relationship between different IPSS-R blast categories and the definition of patient risk stratification in relation to the number of mutated NK-PSS-M genes across the three models. The numbers inside parentheses on the Y-axis represent the number of cases, while the numbers on the X-axis represent percentages

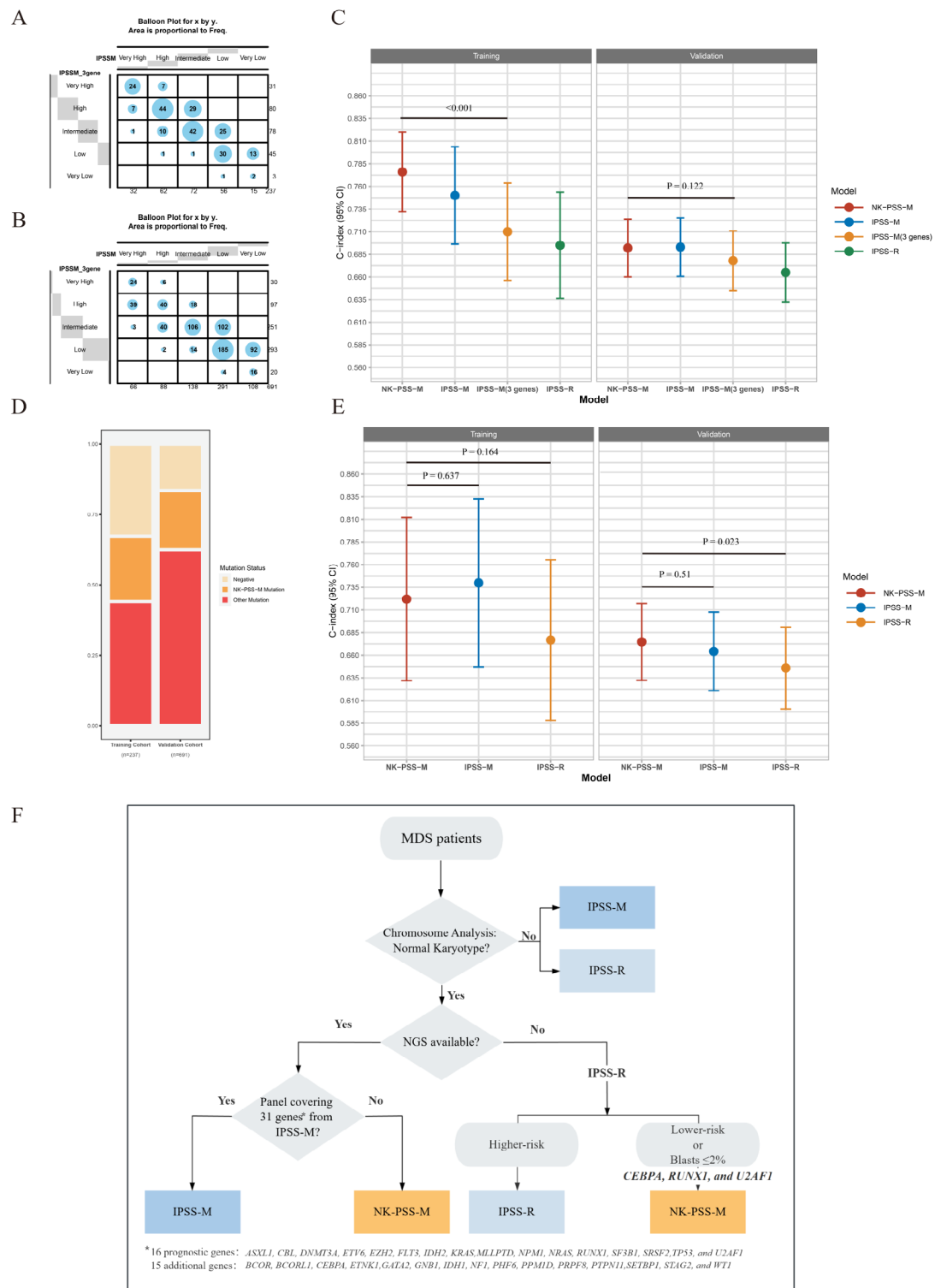


Fig. 7 NK-PSS-M demonstrates robust performance in extreme scenarios and a recommended application strategy. **(A-C)** Comparison of the predictive performance of NK-PSS-M, IPSS-M, and IPSS-R when the gene panel only includes *CEBPA*, *U2AF1*, and *RUNX1*. **(A-B)** The balloon plots show the changes in IPSS-M prediction results in the training and validation cohorts, respectively. **(C)** C-index values of each model in the training and validation cohorts. **(D-E)** Evaluation of the predictive performance of NK-PSS-M, IPSS-M, and IPSS-R in patients with mutations other than the three genes. **(D)** Proportion of patients with different mutation status. **(E)** C-index values of each model in the training and validation cohorts. **(F)** A recommended application strategy for NK-PSS-M in clinical practice

Furthermore, most MDS prognostic models are based on the entire population and incorporate cytogenetic information as a variable. Because the prognostic significance of genes varies in different cytogenetic backgrounds, these models may not be optimal for specific cytogenetic MDS patients. Although our research revealed that the mutation frequency of most NK MDS genes is similar to that of the entire MDS population, variations may be detected in specific genes. Some popular prognostic genes, such as *TP53*, have lost their association with prognosis in NK MDS, revealing heterogeneity in prognostic significance across different cytogenetic backgrounds [14]. Incorporating these ‘unimportant’ molecular factors into the model can compromise its predictive performance, particularly for molecular prognostic models with few parameters. Such as Nazha-2016, the addition of age and molecular factors to the model does not significantly enhance its predictive power for NK MDS compared to IPSS-R.

We developed a parsimonious molecular-clinical prognostic model based on 237 NK MDS patients. The NK-PSS-M includes age, HB, PLT, BM Blasts, and three genes. Clinical characteristics remained significant even after incorporating molecular factors, indicating that a single molecular feature cannot represent MDS specificity. Clinical characteristics represent unique aspects of MDS that differ from molecular features, which is consistent with a previous study that developed a prognostic model based on only gene mutation. Although the pure molecular prognostic model demonstrates certain advantages, its performance does not match that of the molecular-clinical model [19].

The NK-PSS-M model allows for personalized risk assessment based on patient characteristics, with higher scores indicating a greater risk and poorer prognosis. In addition, we developed a five-class classification system based on these scores, which differentiates patients across risk categories and predicts outcomes, such as OS and LFS. Compared to our central cohort, the IWG-PM cohort displayed typical differences between Western and Asian patients, such as lower BM blast burden and milder cytopenias [20–22]. Although the IWG-PM cohort is considered a relatively low-risk group compared to our training cohort according to established models like IPSS-M and IPSS-R, our model demonstrated excellent generalizability and performed equally well in predicting outcomes across both cohorts. Furthermore, the performance of our model was comparable to that of IPSS-M, and it significantly outperformed IPSS-R and other molecular prediction models based on it.

The superior performance of NK-PSS-M over IPSS-R can be attributed to the inclusion of age and gene mutation information. Compared to the Nazha-2016, the NK-PSS-M incorporates a gene list with more

prognostic relevance in NK MDS. Compared to the IPSS-M, the advantage of NK-PSS-M is its parsimonious design, which includes only three common molecular features. This reduces the demand for sequencing resources and panel restrictions, making it more suitable for resource-limited centers.

Furthermore, based on the parameters identified in NK-PSS-M, we developed a version that simplifies the calculation process. The simplified version of the NK-PSS-M model has proven to be an effective tool for the stratification of MDS patients, maintaining robust predictive accuracy with less complexity. This model offers another choice that can enhance clinical usability, facilitating quicker and more efficient patient assessment and decision-making in routine practice.

The robust performance and practical applicability of NK-PSS-M offer several advantages in clinical settings. First, as a model specifically designed for NK MDS patients, it provides more precise risk stratification for this particular subgroup. Notably, NK-PSS-M could identify an additional 10% of high-risk patients among those previously classified as (very) low risk by IPSS-R. Second, NK-PSS-M demonstrates reliable prognostic accuracy across different time points and populations, making it a dependable tool for long-term survival prediction and treatment planning. Third, the model’s parsimonious nature, incorporating only three genes along with readily available clinical parameters, provides a streamlined and cost-effective approach to risk stratification. This feature is particularly valuable in resource-limited settings where extensive genetic testing may not be accessible or affordable. Finally, the model’s excellent performance in both Asian and Western populations suggests its broad applicability across different ethnic backgrounds and healthcare settings. These characteristics position NK-PSS-M as a practical and reliable tool for improving risk-adapted treatment strategies in NK MDS patients, potentially influencing decisions regarding treatment intensity, timing of interventions, and clinical trial eligibility.

Although our study showed promising results, it had some limitations. It was a single-center retrospective study that included 237 NK MDS patients. We only included 23 common genes in the final panel to ensure study uniformity and clinical applicability. In addition, we focused on OS as the outcome during variable selection and model construction, which may have limited the predictive ability for LFS and AML transformation. Our study provides an alternative to overcome the challenges associated with the application of IPSS-M, particularly for NK MDS. In addition, our findings suggest the importance of considering different cytogenetic backgrounds in MDS molecular model development.

In conclusion, we developed and validated a molecular-clinical prognostic model for NK MDS patients, which

can predict OS and LFS. NK-PSS-M is superior to the IPSS/IPSS-R, which currently guides clinical practice, and has stronger universality while maintaining comparable predictive ability to IPSS-M, making it a reliable alternative. However, further validation using larger and more diverse clinical cohorts is necessary to promote its clinical translation.

Abbreviations

MDS	Myelodysplastic syndromes
NK	Normal karyotype
NK-PSS-M	Molecular Prognostic Scoring System for Normal karyotype Myelodysplastic syndromes
IPSS	International Prognostic Scoring System
IPSS-R	Revised International Prognostic Scoring System
IPSS-M	Molecular International Prognosis Scoring System
NGS	Next-generation sequencing
VAF	Variant allele frequency
OS	Overall survival
LFS	Leukemia-free survival
BM	Bone Marrow
ROC	Receiver operating characteristic

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-024-05995-x>.

Supplementary Material 1

Acknowledgements

We sincerely appreciate the selfless sharing of clinical and mutation data from the IPSS-M study by the IWG-PM.

Author contributions

H-Y T, WW, and Y-D Z conceptualized and designed the study. WW conducted the statistical analysis and prepared the figures. WW, Y-D Z, and W-L Y wrote the paper. X-Z L, W-L Y, L-X J, and WL collected patient samples and clinical data. Y-W L, X-P Z, LW, LY, YX, and L-Y M collected the mutational data. H-Y T supervised the study and revised the manuscript. All authors reviewed the manuscript.

Funding

This work was supported by grants from the National Natural Science Foundation of China (82200159, 82270146) and the Key R&D Program of Zhejiang (2024C03164).

Data availability

The clinical dataset used during the current study are available from the corresponding author on reasonable request.

Declarations

Ethical approval

This study was approved by the Ethics Committee of the First Affiliated Hospital, Zhejiang University School of Medicine, and conducted in compliance with the Helsinki Declaration.

Consent for publication

All authors have consented to the publication of this manuscript in this journal.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 19 October 2024 / Accepted: 16 December 2024

Published online: 16 January 2025

References

- Li H, Hu F, Gale RP, et al. Myelodysplastic syndromes. *Nat Rev Dis Primers*. 2022;8:74.
- Cazzola M. Myelodysplastic syndromes. *N Engl J Med*. 2020;383:1358–74.
- Platzbecker U. Treatment of MDS. *Blood*. 2019;133:1096–107.
- Greenberg PL, Stone RM, Al-Kali A, et al. NCCN Guidelines(R) insights: myelodysplastic syndromes, Version 3.2022. *J Natl Compr Canc Netw*. 2022;20:106–17.
- Greenberg P, Cox C, LeBeau MM, et al. International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood*. 1997;89:2079–88.
- Xie Z, Chen EC, Stahl M, et al. Prognostication in myelodysplastic syndromes (neoplasms): molecular risk stratification finally coming of age. *Blood Rev*. 2023;59:101033.
- Nazha A, Narkhede M, Radivoyevitch T, et al. Incorporation of molecular data into the revised International Prognostic Scoring System in treated patients with myelodysplastic syndromes. *Leukemia*. 2016;30:2214–20.
- Bernard E, Tuechler H, Greenberg PL, et al. Molecular International Prognostic Scoring System for Myelodysplastic syndromes. *NEJM Evid*. 2022;1:EVIDoa2200008.
- Sauta E, Robin M, Bersanelli M et al. Real-world validation of Molecular International Prognostic Scoring System for Myelodysplastic syndromes. *J Clin Oncol*:Jco2201784, 2023.
- Montalban-Bravo G, Kanagal-Shamanna R, Benton CB, et al. Genomic context and TP53 allele frequency define clinical outcomes in TP53-mutated myelodysplastic syndromes. *Blood Adv*. 2020;4:482–95.
- Wang W, Routbort MJ, Tang Z, et al. Characterization of TP53 mutations in low-grade myelodysplastic syndromes and myelodysplastic syndromes with a non-complex karyotype. *Eur J Haematol*. 2017;99:536–43.
- Daver NG, Maiti A, Kadia TM, et al. TP53-Mutated myelodysplastic syndrome and Acute Myeloid Leukemia: Biology, current therapy, and future directions. *Cancer Discov*. 2022;12:2516–29.
- Tefferi A, Idossa D, Lasho TL, et al. Mutations and karyotype in myelodysplastic syndromes: TP53 clusters with monosomal karyotype, RUNX1 with trisomy 21, and SF3B1 with inv(3)(q21q26.2) and Del(11q). *Blood Cancer J*. 2017;7:658.
- Wang W, Zhang Y, Yang W, et al. Mutation landscape of normal karyotype myelodysplastic syndromes and their prognostic impact. *Am J Hematol*. 2024;99:E51–4.
- Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127:2391–405.
- Rack KA, van den Berg E, Haferlach C, et al. European recommendations and quality assurance for cytogenomic analysis of haematological neoplasms. *Leukemia*. 2019;33:1851–67.
- Hook EB. Exclusion of chromosomal mosaicism: tables of 90%, 95% and 99% confidence limits and comments on use. *Am J Hum Genet*. 1977;29:94–7.
- Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized Linear models via Coordinate Descent. *J Stat Softw*. 2010;33:1–22.
- Haferlach T, Nagata Y, Grossmann V, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia*. 2014;28:241–7.
- Wu J, Zhang Y, Qin T, et al. IPSS-M has greater survival predictive accuracy compared with IPSS-R in persons ≥ 60 years with myelodysplastic syndromes. *Exp Hematol Oncol*. 2022;11:73.
- Matsuda A, Germing U, Jinnai I, et al. Difference in clinical features between Japanese and German patients with refractory anemia in myelodysplastic syndromes. *Blood*. 2005;106:2633–40.
- Jiang Y, Eveillard JR, Couturier MA et al. Asian Population is more Prone to develop high-risk myelodysplastic syndrome, Concordantly with their propensity to exhibit high-risk cytogenetic aberrations. *Cancers (Basel)* 13, 2021.

Publisher's note

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