ORIGINAL PAPER

BJHaem
BRITISH JOURNAL OF HAEMATOLOGY

Development and validation of a prognostic scoring model to risk stratify childhood acute myeloid leukaemia

Jun Li | **Lipeng Liu** | **Ranran Zhang** | **Yang Wan** | **Xiaowen Gong** | **Li Zhang** | **Wenyu Yang** | **Xiaojuan Chen** | **Yao Zou** | **Yumei Chen** | **Ye Guo** | **Min Ruan** | **Xiaofan Zhu**

State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Haihe Laboratory of Cell Ecosystem, Institute of Hematology and Blood Diseases Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Tianjin, China

Correspondence
Min Ruan and Xiaofan Zhu, State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Haihe Laboratory of Cell Ecosystem, Institute of Hematology and Blood Diseases Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Tianjin, China.
Email: ruanmin@ihcams.ac.cn and xfzhu@ihcams.ac.cn

Summary

To create a personal prognostic model and modify the risk stratification of paediatric acute myeloid leukaemia, we downloaded the clinical data of 597 patients from the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) database as a training set and included 189 patients from our centre as a validation set. In the training set, age at diagnosis, -7/del(7q) or -5/del(5q), core binding factor fusion genes, FMS-like tyrosine kinase 3-internal tandem duplication (*FLT3*-ITD)/ nucleophosmin 1 (*NPM1*) status, Wilms tumour 1 (*WT1*) mutation, biallelic CCAAT enhancer binding protein alpha (*CEBPA*) mutation were strongly correlated with overall survival and included to construct the model. The prognostic model demonstrated excellent discriminative ability with the Harrell's concordance index of 0.68, 3- and 5-year area under the receiver operating characteristic curve of 0.71 and 0.72 respectively. The model was validated in the validation set and outperformed existing prognostic systems. Additionally, patients were stratified into three risk groups (low, intermediate and high risk) with significantly distinct prognosis, and the model successfully identified candidates for haematopoietic stem cell transplantation. The newly developed prognostic model showed robust ability and utility in survival prediction and risk stratification, which could be helpful in modifying treatment selection in clinical routine.

KEYWORDS
acute leukaemia, myeloid leukaemia, paediatric haematology, prognostic factors

INTRODUCTION

Acute myeloid leukaemia (AML) is a clinically and molecularly heterogenous disease that accounts for 20%–30% of all leukaemias in children. Despite the relatively good prognosis of childhood AML compared to adults, only 60%–70% patients achieve long-term survival.[1] Relapsed disease poses a great challenge to the treatment of AML, and 30% of children with AML experience relapse and have a poor prognosis. Although new therapeutic agents have been emerging in

recent years, only minor improvements have been achieved in the clinical outcomes of childhood AML.

Combined chemotherapy is the foundation in the treatment of AML. However, chemotherapy alone is still unsatisfactory in some patients, and these patients are often stratified into the high-risk group and are candidates for allogeneic haematopoietic stem cell transplantation (HSCT).[2] In addition, the intensity of chemotherapy should be tailored personally to maximise efficacy and minimise toxicity. Altogether, the treatment for AML requires precise and personal risk stratification. Many efforts have been made in the refinement of risk stratification for adult AML. The European LeukemiaNet (ELN) genetic risk stratification has been widely used in clinical practice and exhibits robust predictive value in not only long-term survival but also survival after HSCT and early mortality.[3–5] Also, many prognostic models for both young adult and elderly AML have been developed and achieved outstanding performances.[6,7] By comparison, few publications have proposed risk stratification for childhood AML, and the discrepancies between children and adults limits the application of the aforementioned models in the paediatric population.[8,9]

Moreover, the implications of some clinical and mutational signatures remain to be determined, and how they mutually influence survival is often hard to quantify. For instance, the presence of *RUNX1-RUNX1T1* fusion gene is considered to be a favourable factor for AML. But it can be complicated by the co-existence with FMS-like tyrosine kinase 3-internal tandem duplication (*FLT3*-ITD), an aberration generally viewed as adverse.[10] Hence, we constructed a prognostic model using some simple-to-obtain clinical features and proposed a risk stratification specifically for paediatric AML. Moreover, we made comparisons to the existing prognostic systems and explored its capability in guiding therapeutic decision-making.

METHODS

Participants and sources of data

The clinical information was downloaded from the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) project database for childhood AML patients (http://ocg.cancer.gov/). A total of 597 patients with AML enrolled between 2006 and 2010 in The Children's Oncology Group Trial COG-AAML0531 (NCT01407757) were included in this study as a training set. Between August 2009 and December 2015, another 189 children with non-M3 AML from the Institute of Haematology, Chinese Academy of Medical Sciences (IHCAMS) were included in this study as the validation set. Patients with missing data or who died before or during initial induction therapy were excluded. The study design and methods complied with the Declaration of Helsinki and was approved by the Ethics Committee and Institutional Review Board of Institute of Haematology

and Blood Diseases Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College. Informed consent was obtained from all patients.

Definitions and outcomes

Complex karyotype was defined as three or more chromosomal abnormalities, excluding the recurrent abnormalities designated in the 2016 World Health Organization (WHO) classification. Complete remission (CR) was defined as bone marrow (BM) blasts <5% by morphology. Relapse was defined as BM blasts ≥5%, or presence of any extramedullary diseases after CR was achieved. Induction failure was defined as BM blasts ≥5% after induction II. Overall survival (OS) was the primary end-point of interest, which was measured from the date of diagnosis to death of any cause or the last date of follow-up. Event-free survival (EFS) was defined as the interval between diagnosis and induction failure, relapse or death, whichever came the first.

Treatment regimen

The treatment regimen of patients enrolled in COG-AAML0531 was detailed in a previous study.[11] Patients treated in IHCAMS received CAMS-2009 regimen. Briefly, introduction therapy consisted of etoposide 150 mg/m^2 intravenously (IV) on days 1–5, idarubicin 8 mg/m^2 IV on days 6–8 and cytarabine (Ara-C) 200 mg/m^2 IV on days 6–12. Re-induction therapy was administered if CR was not achieved. Patients achieving CR were to receive five courses of consolidation therapy. Course 1 and 4 includes idarubicin 10 mg/m^2 IV on day 1 and Ara-C 3 g/m^2 IV on days 1–3. Course 2 and 5 includes etoposide 150 mg/m^2 IV on days 1–3, mitoxantrone 5 mg/m^2 IV on days 4–6 and Ara-C 200 mg/m^2 IV on days 4–8. Course 3 includes etoposide 100 mg/m^2 IV on days 1–5 and Ara-C 2 g/m^2 IV on days 1–5. Patients with *FLT3*-ITD, BM blasts ≥15% after first induction or ≥5% by morphology or ≥0.1% by flow cytometry after re-induction therapy, or relapsed disease are candidates for HSCT.[12]

Predictor

The potential predictors of OS were analysed to develop the prognostic model. Parameters at diagnosis included in this study were as following: age at diagnosis, gender, white blood cell count and BM blasts at diagnosis, -7/del(7q) or -5/del(5q), complex karyotype, gene fusion or mutations including core binding factor (CBF), Nucleoporin 98 (*NUP98*) rearrangements, Lysine methyltransferase 2A (*KMT2A*) rearrangements, *FLT3*-ITD, FMS-like tyrosine kinase 3-tyrosine kinase domain (*FLT3*-TKD), Wilms Tumour 1 (*WT1*), biallelic CCAAT enhancer binding protein alpha (*CEBPA*) and nucleophosmin 1 (*NPM1*), as well as response to induction therapy.

Statistical analysis

Uni- and multivariable Cox analysis was applied to assess the impact of parameters on OS. We included parameters with a $p < 0.05$ in multivariable Cox analysis into model development. The risk score for each subject was calculated based on the coefficient of each parameter in multivariate analysis. Patients were stratified into low-, intermediate- and high-risk groups based on the 25% and 75% quartiles of the risk score of the training set. Model performance was evaluated by discrimination and calibration. Discrimination was assessed by Harrell's concordance index (C-index) and the area under the curve of a time-dependent receiver operating characteristic curve (AUROC). Bias-corrected calibration curves for OS using 1000 bootstrap re-sampling were performed to evaluate the consistency of observed and estimated survival probability.

Continuous variables are presented as median and interquartile range (IQR) and compared using the Mann–Whitney U-test. Categorical variables were analysed using Fisher's exact test. Probability of OS and EFS were estimated by the Kaplan–Meier method and compared using the log-rank test. A two-tailed $p < 0.05$ was considered statistically significant. All statistical analyses were performed using the R software 4.0.2 (The CRAN project, www.r-project.org).

RESULTS

Population characteristics

The flow diagram of this study is shown in Figure 1. The clinical characteristics of both cohorts are compared in Table 1. The median (range) follow-up durations were 66.0 (2.2–103.8) and 20.4 (range 0.7–95.5) months in training and validation set respectively. The 1-, 3- and 5-year OS rates were 86.0% (95% confidence interval [CI] 83.2%–88.8%), 70.6% (95% CI 69.9%–74.4%), 66.5% (95% CI 62.8%–70.6%) in training set. In validation set, the 1- and 3-year OS were 76.9% (95% CI 71.0%–83.3%) and 63.7% (95% CI 56.7%–71.5%), and 5-year OS remained the same as 3-year OS because no deaths occurred after the 3-year follow-up.

Model development and performance in training set

Uni- and multivariable Cox analysis was applied to identify the potential risk factors affecting OS as shown in Table 2. Parameters of age at diagnosis, -7/del(7q) or -5/del(5q), CBF gene fusion, *FLT3*-ITD/*NPM1* status, and biallelic *CEBPA* mutation, *WT1* mutation with a $p < 0.05$ in multivariable analysis were selected as predictors to construct the model (Table S1). The predictive performance of the proposed CAMS-pAML model was evaluated by C-index and AUROC, which demonstrated outstanding discrimination with a C-index of 0.68 (95% CI 0.64–0.71), and 3- and 5-year AUROC 0.71 (95% CI 0.66–0.75) and 0.72 (95% CI 0.67–0.76) respectively (Figure 2A). The calibration curve exhibited high consistency for the 3- and 5-year survival probabilities between the model predicted outcomes and actual observations (Figure 2B,C). In addition, a web-based dynamic calculator (https://ihcams.shinyapps.io/CAMS-pAML/) was developed for survival prediction.

The risk score was calculated for each patient using the formula from the multivariable Cox regression, which was expressed as follow:

$$
\begin{aligned}
\text{Risk score} = &\left[0.380 \times \text{age at diagnosis} \geq 2\right] \\
&+ \left[0.862 \times -7/\text{del}(7q) \text{ or } -5/\text{del}(5q)\right] \\
&- \left(1.452 \times \text{CBF fusion gene}\right) \\
&+ \left(0.054 \times FLT3 - \text{ITD} + /NPM1-\right) \\
&- \left(1.252 \times FLT3 - \text{ITD} - /NPM1+\right) \\
&- \left(0.512 \times FLT3 - \text{ITD} + /NPM1+\right) \\
&- \left(1.166 \times CEBPA \text{ biallelic mutation}\right) \\
&+ \left(0.553 \times WT1 \text{ mutation}\right).
\end{aligned}
$$

By the 25% and 75% quartiles of the risk score, the training set was stratified into three risk groups (low-, intermediate- and high-risk groups). Patients with a risk score <-1.018 were considered as low-risk, whereas those with a risk score >0.380 were considered as high-risk. Patients with intermediate risk have a risk score between -1.018 and 0.380. This resulted in 148, 351, and 98 cases in the low-, intermediate- and high-risk groups, and distinct prognosis of each group. The 5-year OS of low-, intermediate- and high-risk groups were 86.0% (95%
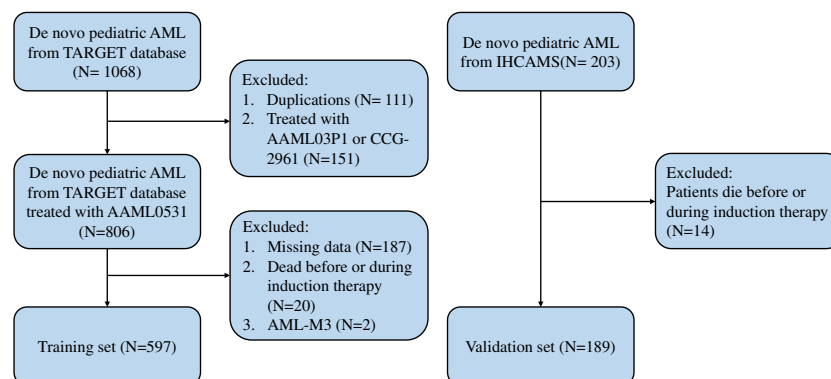

FIGURE 1 Flow diagram describing how the training and validation set were obtained. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Clinical characteristics of all patients

Variable	Level	Overall N = 786	COG-AAML0531 N = 597	CAMS-2009 N = 189	p
Gender, n (%)	Female	366 (46.6)	285 (47.7)	81 (42.9)	0.277
	Male	420 (53.4)	312 (52.3)	108 (57.1)	
Age at diagnosis, years, median (IQR)		9.0 (3.0–13.0)	10.0 (3.0–14.0)	8.0 (5.0–11.0)	0.019
WBC count at diagnosis, ×10^9/L, median (IQR)		23.6 (9.3–73.9)	27.0 (10.8–78.3)	16.5 (8.1–52.3)	0.001
BM blast, %, median (IQR)		68.9 (45.0–86.0)	70.0 (45.0–87.0)	63.6 (44.6–82.5)	0.243
FAB category, n (%)	M0	24 (3.1)	20 (3.4)	4 (2.1)	<0.001
	M1	78 (9.9)	77 (12.9)	1 (0.5)	
	M2	253 (32.2)	152 (25.5)	101 (53.4)	
	M4	165 (21.0)	143 (24.0)	22 (11.6)	
	M5	182 (23.2)	131 (21.9)	51 (27.0)	
	M6	16 (2.0)	13 (2.2)	3 (1.6)	
	M7	40 (5.1)	33 (5.5)	7 (3.7)	
	NOS	28 (3.6)	28 (4.7)	0 (0.0)	
-7/del(7q) or -5/del(5q), n (%)	No	745 (94.8)	561 (94.0)	184 (97.4)	0.089
	Yes	41 (5.2)	36 (6.0)	5 (2.6)	
Complex karyotype, n (%)	No	717 (91.2)	538 (90.1)	179 (94.7)	0.055
	Yes	69 (8.8)	59 (9.9)	10 (5.3)	
CBF fusion gene, n (%)	No CBF fusion gene	524 (66.7)	425 (71.2)	99 (52.4)	<0.001
	RUNX1-RUNX1T1	170 (21.6)	96 (16.1)	74 (39.2)	
	CBFβ-MYH11	92 (11.7)	76 (12.7)	16 (8.5)	
NUP98 fusion gene, n (%)	No	750 (95.4)	562 (94.1)	188 (99.5)	0.001
	Yes	36 (4.6)	35 (5.9)	1 (0.5)	
KMT2A rearrangement, n (%)	No	648 (82.4)	475 (79.6)	173 (91.5)	<0.001
	Yes	138 (17.6)	122 (20.4)	16 (8.5)	
FLT3-ITD/*NPM1* status, n (%)	*FLT3*-ITD-/*NPM1*⁻	637 (81.0)	467 (78.2)	170 (89.9)	0.002
	FLT3-ITD+/*NPM1*⁻	90 (11.5)	77 (12.9)	13 (6.9)	
	FLT3-ITD-/*NPM1*⁺	37 (4.7)	32 (5.4)	5 (2.6)	
	FLT3-ITD+/*NPM1*⁺	22 (2.8)	21 (3.5)	1 (0.5)	
FLT3-TKD, n (%)	Wild-type	742 (94.4)	555 (93.0)	187 (98.9)	0.001
	Mutated	44 (5.6)	42 (7.0)	2 (1.1)	
Biallelic *CEBPA* mutation, n (%)	Wild-type or monoallelic	742 (94.4)	567 (95.0)	178 (94.2)	0.708
	Mutated	44 (5.6)	30 (5.0)	11 (5.8)	
WT1 mutation, n (%)	Wild-type	735 (93.5)	551 (92.3)	184 (97.4)	0.011
	Mutated	51 (6.5)	46 (7.7)	5 (2.6)	
Response to induction therapy, n (%)	CR	608 (77.4)	444 (74.4)	164 (86.8)	<0.001
	Not in CR	178 (22.6)	153 (25.6)	25 (13.2)	

Abbreviations: CAMS, Chinese Academy of Medical Sciences; CBF, core binding factor; *CEBPA*, CCAAT enhancer binding protein alpha; *FLT3*-ITD, FMS-like tyrosine kinase 3-internal tandem duplication; *FLT3*-TKD, FMS-like tyrosine kinase 3-tyrosine kinase domain; IQR, interquartile range; *KMT2A*, lysine methyltransferase 2A; *NPM1*, nucleophosmin 1; *NUP98*, nucleoporin 98; WBC, white blood cell; *WT1*, Wilms tumour 1.

CI 80.4%–91.9%), 66.5% (95% CI 61.6%–71.8%), 36.6% (95% CI 27.8%–48.1%) respectively ($p<0.001$, Figure 2D). The 5-year EFS of low-, intermediate- and high-risk groups were 70.0% (95% CI 62.9%–77.8%), 49.4% (95% CI 44.4%–55.0%), 21.9% (95% CI 15.0%–32.1%) respectively ($p<0.001$, Figure 2E).

Model external validation in validation set

To access the compatibility of the CAMS-pAML model, another independent cohort was included as validation set (n = 189). There were 89, 80 and 20 cases identified in

TABLE 2 Uni- and multivariable Cox analyses of variates impacting overall survival of patients in the training set

Variable	Groups	Univariable analysis		Multivariable analysis	
		HR (95% CI)	p	HR (95% CI)	p
Gender	Female vs. male	1.181 (0.890–1.567)	0.249	1.201 (0.900–1.603)	0.213
Age at diagnosis (years)	<2 vs. ≥2	1.066 (0.733–1.551)	0.737	1.581 (1.039–2.404)	0.032
WBC count at diagnosis (×10⁹/L)	<100 vs. ≥100	1.215 (0.866–1.706)	0.260	1.162 (0.807–1.675)	0.419
BM blast (%)	<70 vs. ≥70	1.211 (0.914–1.606)	0.182	1.103 (0.807–1.508)	0.538
-7/del(7q) or -5/del(5q)	Yes vs. no	2.230 (1.403–3.543)	0.001	2.006 (1.211–3.325)	0.007
Complex karyotype	Yes vs. no	1.641 (1.093–2.464)	0.017	1.274 (0.817–1.987)	0.286
CBF fusion gene	Yes vs. no	0.290 (0.189–0.445)	<0.001	0.244 (0.152–0.393)	<0.001
NUP98 fusion gene	Yes vs. no	1.407 (0.817–2.425)	0.218	0.866 (0.484–1.549)	0.627
KMT2A rearrangement	Yes vs. no	1.372 (0.992–1.879)	0.056	0.970 (0.657–1.434)	0.880
FLT3-ITD/*NPM1* status	*FLT3*-ITD+/*NPM1*- vs. *FLT3*-ITD-/*NPM1*-	1.563 (1.077–2.270)	0.019	0.982 (0.626–1.542)	0.938
	FLT3-ITD-/*NPM1*+ vs. *FLT3*-ITD-/*NPM1*-	0.475 (0.195–1.158)	0.102	0.328 (0.130–0.830)	0.019
	FLT3-ITD+/*NPM1*+ vs. *FLT3*-ITD-/*NPM1*-	0.858 (0.379–1.940)	0.712	0.654 (0.278–1.537)	0.330
FLT3-TKD	Wild-type vs. mutated	0.684 (0.362–1.292)	0.242	0.629 (0.324–1.218)	0.169
Biallelic *CEBPA* mutation	Wild-type or monoallelic vs. mutated	0.454 (0.187–1.105)	0.082	0.298 (0.118–0.751)	0.010
WT1 mutation	Wild-type vs. mutated	2.095 (1.365–3.215)	0.001	1.811 (1.111–2.953)	0.017
Response to induction therapy	CR vs. not in CR	1.669 (1.236–2.252)	0.001	1.234 (0.896–1.700)	0.198

Abbreviations: CBF, core binding factor; *CEBPA*, CCAAT enhancer binding protein alpha; CI, confidence interval; CR, complete remission; *FLT3*-ITD, FMS-like tyrosine kinase 3-internal tandem duplication; *FLT3*-TKD, FMS-like tyrosine kinase 3-tyrosine kinase domain; HR, hazard ratio; *KMT2A*, lysine methyltransferase 2A; *NPM1*, nucleophosmin 1; *NUP98*, nucleoporin 98; WBC, white blood cell; *WT1*, Wilms tumour 1.

the low-, intermediate- and high-risk groups in the validation set. The 5-year OS for low-, intermediate- and high-risk groups of 76.6% (95% CI 67.6%–86.8%), 55.4% (95% CI 44.9%–68.4%), 42.5% (95% CI 25.2%–71.7%) respectively ($p < 0.001$, Figure 3A), and 5-year EFS of 68.0% (95% CI 57.8%–79.9%), 47.5% (95% CI 37.4%–60.4%), 43.3% (95% CI 25.9%–72.4%) respectively ($p = 0.002$, Figure 3B).

Clinical significance of the model

By combining the patients in training and validation sets, we explored the capability of the CAMS-pAML model in identifying HSCT candidates. After removing patients with unavailable information of HSCT, there were 3.0% (seven of 235), 18.2% (73/402) and 35.2% (32/91) patients who received HSCT in first CR in the low-, intermediate- and high-risk groups respectively. It showed that HSCT significantly improved both the OS and EFS of patients in the high-risk group, with 5-year OS of 62.3% (95% CI 46.7%–83.1%) and 29.6% (95% CI 19.6%–44.6%) ($p < 0.001$), and 5-year EFS of 54.9% (95% CI 39.8%–75.7%) and 16.4% (95% CI 9.0%–29.9%) ($p < 0.001$, Figure 4A,B) in patients with and without HSCT respectively. In the intermediate-risk group, the 5-year OS was comparable between patients who underwent HSCT and those who did not (69.6%

[95% CI 59.4%–81.5%] vs. 65.4% [95% CI 60.2%–71.0%], $p = 0.303$, Figure 4C). However, the 5-year EFS was higher in patients who underwent HSCT in the intermediate-risk group (67.8% [95% CI 57.7%–79.7%] vs. 48.6% [95% CI 43.4%–54.5%], $p = 0.004$, Figure 4D). As for the low-risk group, HSCT did not improve the prognosis with patients sharing similar 5-year OS (HSCT versus non-HSCT, 85.7% [95% CI 63.3%–100.0%] vs. 82.7% [95% CI 77.7%–88.1%], $p = 0.870$, Figure 4C) and EFS (HSCT vs. non-HSCT, 85.7% [95% CI 63.3%–100.0%] vs. 69.5% [95% CI 63.6%–76.0%], $p = 0.399$, Figure 4E,F).

Comparison to the existing prognostic systems

The ELN guidelines proposed the widest accepted risk stratifications for paediatric and adult AML in 2012 and 2017 respectively.[5,13] Therefore, we evaluated the performance of the CAMS-pAML model by comparing to them. The CAMS-pAML model outperformed both the 2017 (C-index: 0.68 [95% CI 0.64–0.71] vs. 0.64 [95% CI 0.60–0.68], $p = 0.015$) and the 2012 (C-index: 0.68 [95% CI 0.64–0.71] vs. 0.64 [95% CI 0.61–0.68], $p = 0.015$) ELN prognostic systems with a higher C-index in the training set. Survival analyses revealed the CAMS-pAML model did better in distinguishing intermediate- and high-risk patients (Figure 5A–D). We
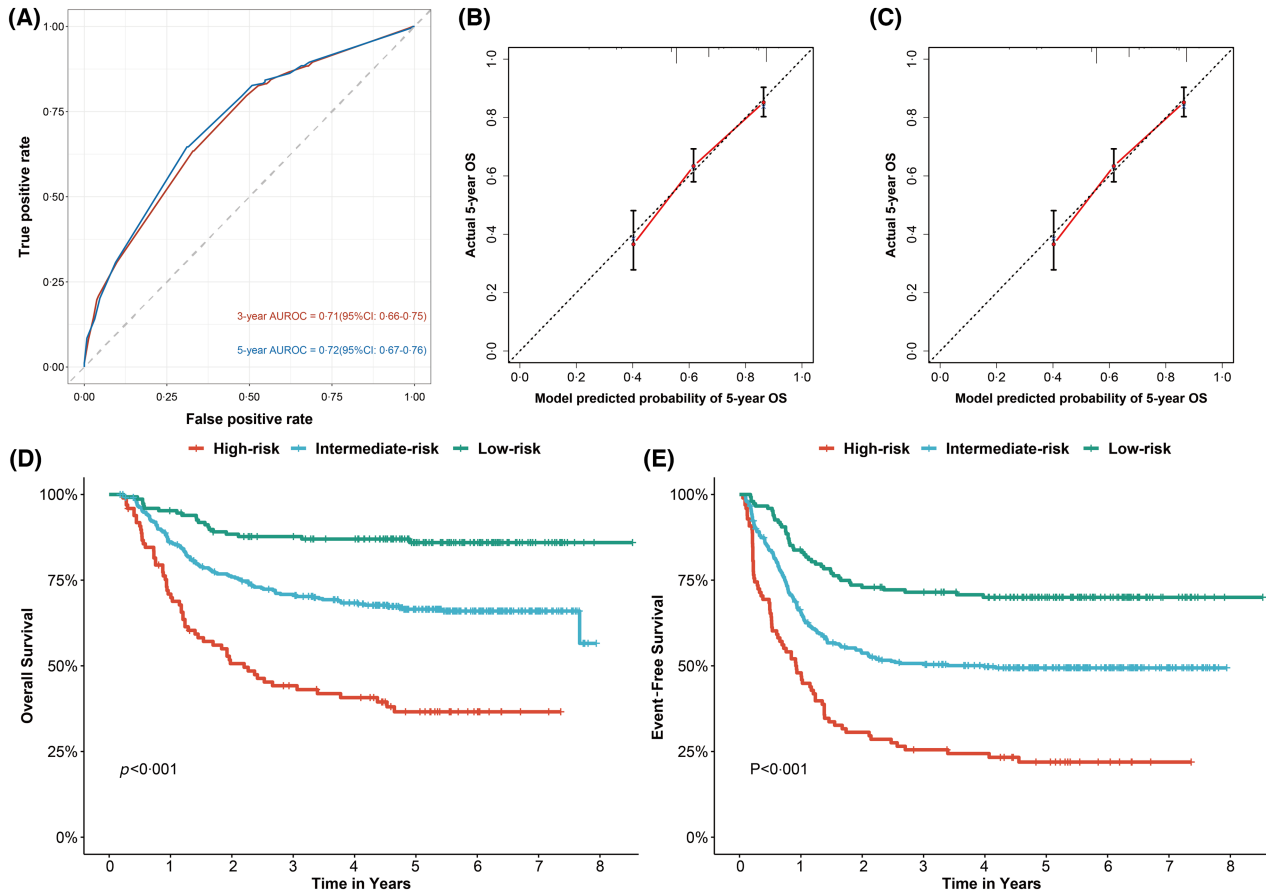
FIGURE 2 Model performance in training set. (A) The discriminative ability of the prognostic model was evaluated by area under the time-dependent receiver operating characteristic curve (AUROC). The 3- and 5-year AUROC was 0.71 and 0.72 respectively. (B and C) The calibration curve of training set showed good agreement of model predicted survival and actual survival. (D and E) Kaplan–Meier curves for overall survival (OS) and event-free survival (EFS) of patients in the training set in different risk groups. [Colour figure can be viewed at wileyonlinelibrary.com]
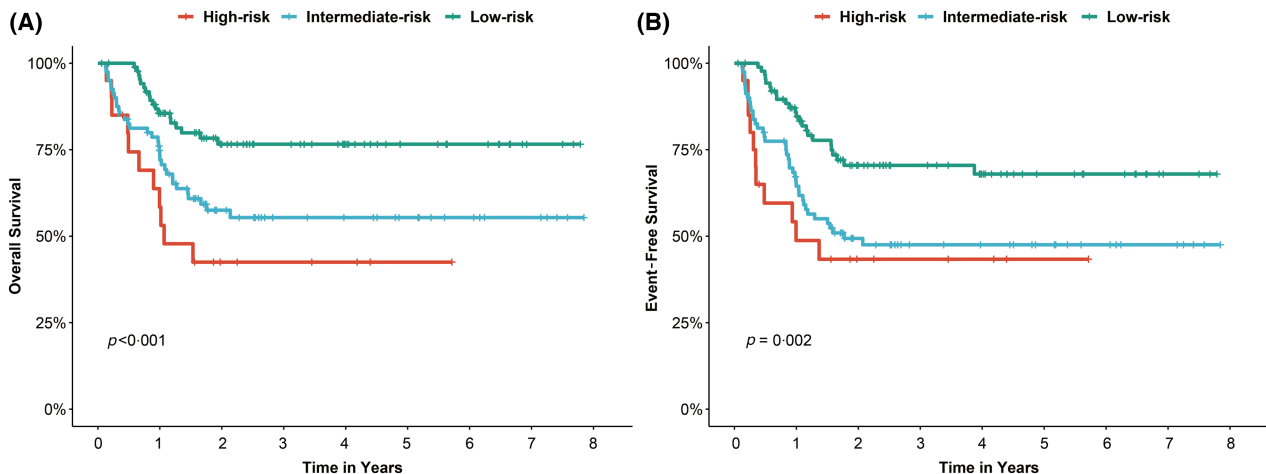

FIGURE 3 Model performance in validation set. (A and B) Kaplan–Meier curves for overall survival and event-free survival of patients in validation set in different risk groups. [Colour figure can be viewed at wileyonlinelibrary.com]

further investigated whether HSCT improved the survival of patients in each of the risk groups stratified by ELN prognostic systems. It turned out that HSCT did not improve the OS of patients in either high-, intermediate-, or low-risk groups identified by the 2012 and 2017 ELN prognostic systems (Figure S1).

DISCUSSION

Here, we developed a model with robust predictive value in long-term survival of childhood AML in both the training and validation sets, which outperformed the existing
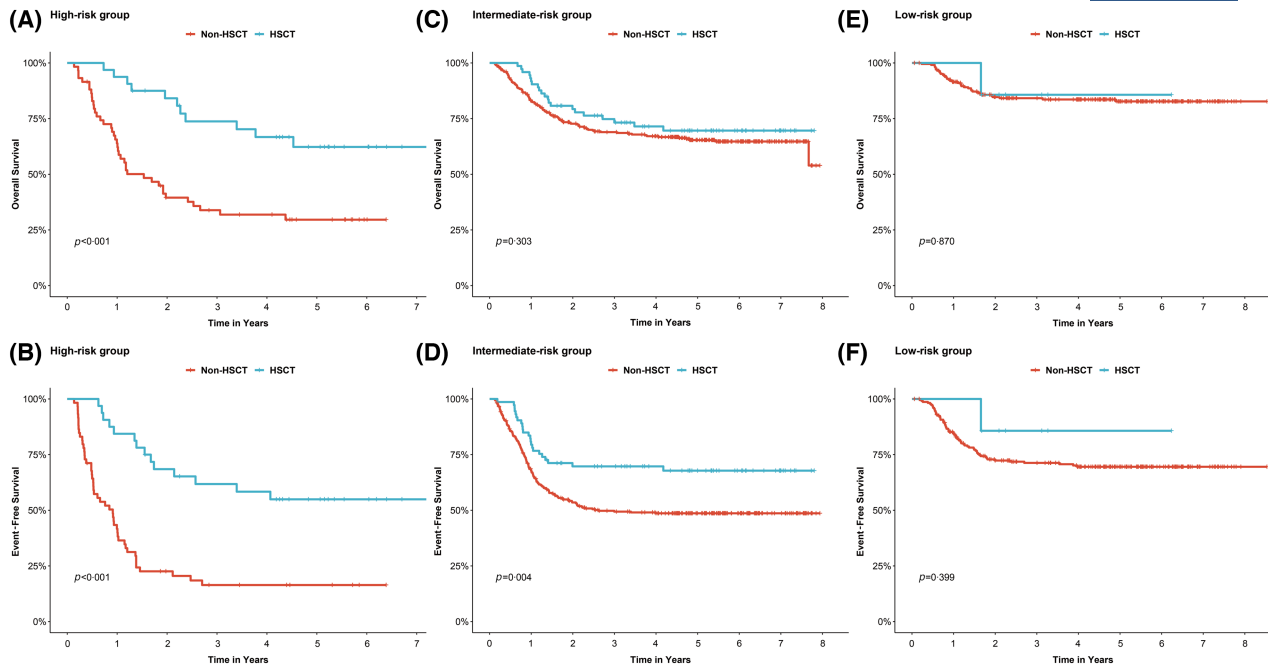
FIGURE 4 Kaplan–Meier curves for overall survival and event-free survival of patients with and without haematopoietic stem cell transplantation (HSCT). (A, B) High-risk group; (C, D) intermediate-risk group; (E, F) low-risk group. [Colour figure can be viewed at wileyonlinelibrary.com]
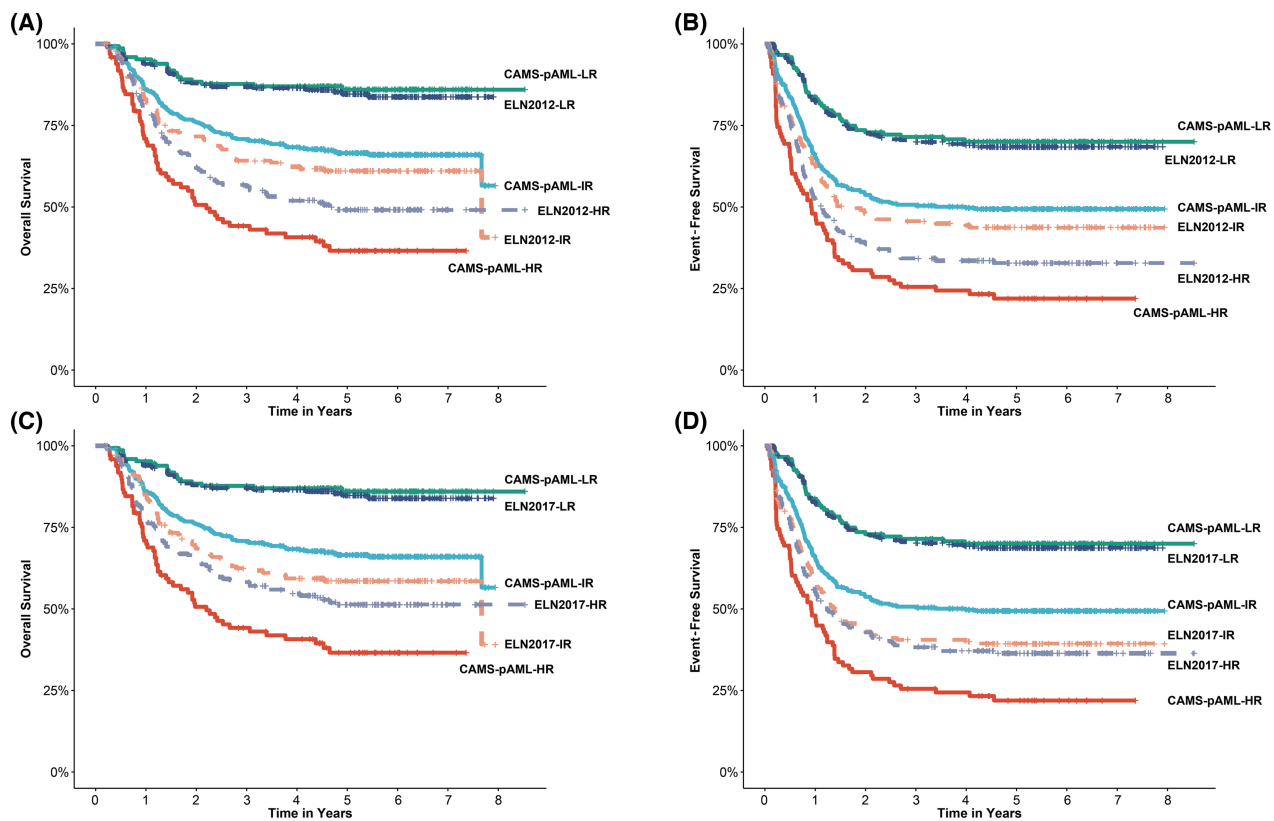

FIGURE 5 Kaplan–Meier curves for overall survival and event-free survival of the risk groups stratified by the CAMS-pAML model (solid lines) and 2012 or 2017 ELN prognostic systems (dashed lines). The CAMS-pAML model distinguished high- and intermediate-risk groups better than the 2012 (A and B) and 2017 (C and D) ELN prognostic systems. CAMS, Chinese Academy of Medical Sciences; ELN, European LeukemiaNet, HR, high-risk; IR, intermediate-risk; LR, low-risk, pAML, paediatric acute myeloid leukaemia. [Colour figure can be viewed at wileyonlinelibrary.com]

prognostic systems. Moreover, the refined risk stratification based on the proposed CAMS-pAML model is able to identify candidates for HSCT, which can be of great help in clinical routine. To date, very few studies have shed light on developing a prognostic model for childhood AML. Elsayed et al.[14] and Wang et al.[15] built models with exceptional predictive ability on the basis of RNA expression. However, gene expression data is not easy to obtain and can be obscure to clinicians. To the best of our knowledge, the CAMS-pAML model includes the most comprehensive clinical predictors, which are also simple to obtain.

Although with wide acceptance globally, both the 2012 and 2017 ELN prognostic systems failed to distinguish the high- and intermediate-risk groups in the training set of our study. This result may be attributed to the unquantified nature of the ELN risk stratifications. When multiple prognostic factors co-exist, it is hard to evaluate the interactions between these factors. Meanwhile, we found that the prognosis of the high-risk group identified by the ELN prognostic systems was not improved by HSCT, indicating our model outperformed the ELN systems in identifying HSCT candidates as well.

Notably, the clinical characteristics varied to a great extent between the training and validation sets, especially on the cytogenetic and mutational profiles. Similar differences have been demonstrated by other large cohort studies of Chinese population as well. Chen et al.[16] conducted a comprehensive descriptive study on fusion genes of Chinese paediatric AML and demonstrated higher occurrence of *RUNX1-RUNX1T1* but lower occurrence of *CBFβ-MYH11* and *KMT2A* gene abnormalities compared to Western reports. More recently, other research found that *FLT3*, *NPM1* and *WT1* mutations are seen less in Chinese paediatric AML.[17] These discrepancies might reflect different backgrounds across ethnicities but verified the compatibility of our model on the other hand.

In our study, cytogenetic and molecular aberrations remained the most powerful predictors for childhood AML prognosis. Previous publications reported similar prognostic significance with -7/del(7q) and -5/del(5q) carrying the poorest prognosis.[18,19] Hasle et al.[20] reported HSCT could not prolong the survival of patients with monosomy 7 or del(7). More effective therapies are warranted for these patients. Although with lower incidence than adult, *FLT3*-ITD/ *NPM1* status and *CEBPA* biallelic mutation also showed significant impact on the survival of our cohorts. *FLT3*-ITD was regarded as an adverse prognostic factor in both National Comprehensive Cancer Network (NCCN) and ELN risk stratifications, and many paediatric cohorts validated its influence as well.[21,22] Despite the development of *FLT3*-targeted therapy, such as midostaurin, sorafenib and gilteritinib, the clinical use of these agents in children is scarce due to toxicity and limited evidence.[23,24] HSCT is thus recommended as consolidation therapy for patients with *FLT3*-ITD. Little research has focussed on *WT1* mutation. Hollink et al.[25] reported that *WT1* mutation could be detected in ~12% of paediatric AML of normal cytogenetics

and is associated with *FLT3*-ITD. Survival analyses revealed its poor prognostic significance with a lower survival rate and higher relapse rate. Our research identified *WT1* mutation as an independent risk factor as well. More studies are needed to further clarify its role in the clinical course of childhood AML.

Core binding factor-AML is a special subtype of leukaemia, characterised by the presence of *RUNX1-RUNX1T1* or *CBFβ-MYH11*. CBF-AML accounts for 25%–30% of paediatric AML and confers a favourable outcome. However, ~40% of patients experience relapse, suggesting other factors influence the prognosis.[26] Unlike in adult CBF-AML where *KIT* mutation is considered to carry poor prognosis, this phenomenon has not been observed in paediatric cohorts. The rate and degree of minimal residual disease (MRD) reduction was believed to be the strongest prognostic factor, which suggested risk stratification of CBF-AML should adjust in line with the response to treatment.[27,28]

Consolidation therapies mainly comprise chemotherapy and HSCT, and currently selection for consolidation therapy is determined based on risk stratification. Consistent with previous studies, patients in the high-risk group identified by our model could benefit from HSCT, which should be conducted as soon as CR is achieved.[29] For patients with intermediate risk, previous studies have shown higher OS and EFS in patients who underwent HSCT, thus, they are recommended to receive HSCT in first CR if they have a matched family donor.[30] Whereas HSCT only improved the EFS but not OS of the intermediate-risk group in our study, indicating the CAMS-pAML model may have better discriminative ability in HSCT candidate selection. Therefore, whether intermediate-risk patients identified by our model receive HSCT may depend on personal situation, especially the response to treatment and surveillance of MRD. In 2021, the ELN updated the consensus on MRD assessment of AML, emphasising its importance in treatment selection and disease surveillance.[31] More specifically, in patients with AML carrying mutated *NPM1*, persistent MRD positivity at a certain level may trigger the decision to consider HSCT. This strategy could help identify patients at high risk in the previous low- or intermediate-risk groups and prevent overt relapse. But more research is needed in MRD assessment before implementing such dynamic risk stratification in AML. Meanwhile, the prognosis of intermediate-risk patients could be improved by adding targeted therapies to chemotherapy in the future.

With the development and clinical application of next-generation sequencing (NGS), more predictors will be discovered to make more precise risk stratification.[32] For example, Sasaki et al.[33] found that the variant allele frequency of *ASXL1*, *DNMT3A*, *JAK2*, *TET2*, *TP53*, and *NPM1* mutations significantly influenced OS of AML. Meanwhile, great progress is being made in targeted therapies. Venetoclax has been widely used due to its significant activity in adult AML and ongoing trails also show a promising future for targeting the *BCL-2* pathway in paediatric patients.[34] More emerging agents targeting *TP53*, *IDH1/2* mutations, CD47 pathway

and many others are expected to bring the treatment of AML to a new era.[35]

In conclusion, we propose a robust prognostic model and risk stratification specifically for paediatric AML, which we believe will contribute to modifying treatment selection and help improve prognosis. With new techniques like NGS more prognostic factors are expected to refine the model.

AUTHOR CONTRIBUTIONS

Jun Li, Lipeng Liu, Min Ruan and Xiaofan Zhu designed the study. Jun Li, Lipeng Liu, Min Ruan, Ranran Zhang and Yang Wan collected the data. Jun Li and Lipeng Liu analysed the data and drafted the manuscript. Min Ruan and Xiaofan Zhu made critical revisions to the manuscript for important intellectual content. Xiaowen Gong, Li Zhang, Wenyu Yang, Xiaojuan Chen, Yao Zou, Yumei Chen and Ye Guo provided guidance on statistical analysis and presentation of results. All authors approved the final manuscript and agreed to submit for publication.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

No conflict of interest declared.

DATA AVAILABILITY STATEMENT

All raw data of this study are available from the corresponding author by request.

ORCID

Jun Li https://orcid.org/0000-0001-8361-9279
Yang Wan https://orcid.org/0000-0002-5914-3327
Min Ruan https://orcid.org/0000-0001-7253-0324

REFERENCES

1. Aung MMK, Mills ML, Bittencourt-Silvestre J, Keeshan K. Insights into the molecular profiles of adult and paediatric acute myeloid leukaemia. Mol Oncol. 2021;15(9):2253–72.
2. Egan G, Chopra Y, Mourad S, Chiang KY, Hitzler J. Treatment of acute myeloid leukemia in children: a practical perspective. Pediatr Blood Cancer. 2021;68(7):e28979.
3. Keren-Froim N, Heering G, Sharvit G, Zlotnik M, Nagler A, Shimoni A, et al. ELN 2017 classification significantly impacts the risk of early death in acute myeloid leukemia patients receiving intensive induction chemotherapy. Ann Hematol. 2022;101(2):309–16.
4. Grimm J, Jentzsch M, Bill M, Goldmann K, Schulz J, Niederwieser D, et al. Prognostic impact of the ELN2017 risk classification in patients with AML receiving allogeneic transplantation. Blood Adv. 2020;4(16):3864–74.
5. Dohner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Buchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;129(4):424–47.
6. Ma TT, Lin XJ, Cheng WY, Xue Q, Wang SY, Liu FJ, et al. Development and validation of a prognostic model for adult patients with acute myeloid leukaemia. EBioMedicine. 2020;62:103126.
7. Liu CJ, Hong YC, Kuan AS, Yeh CM, Tsai CK, Liu YC, et al. The risk of early mortality in elderly patients with newly diagnosed acute myeloid leukemia. Cancer Med. 2020;9(4):1572–80.
8. Short NJ, Rytting ME, Cortes JE. Acute myeloid leukaemia. Lancet. 2018;392(10147):593–606.
9. Juhl-Christensen C, Ommen HB, Aggerholm A, Lausen B, Kjeldsen E, Hasle H, et al. Genetic and epigenetic similarities and differences between childhood and adult AML. Pediatr Blood Cancer. 2012;58(4):525–31.
10. Duployez N, Marceau-Renaut A, Boissel N, Petit A, Bucci M, Geffroy S, et al. Comprehensive mutational profiling of core binding factor acute myeloid leukemia. Blood. 2016;127(20):2451–9.
11. Gamis AS, Alonzo TA, Meshinchi S, Sung L, Gerbing RB, Raimondi SC, et al. Gemtuzumab Ozogamicin in children and adolescents with De novo acute myeloid leukemia improves event-free survival by reducing relapse risk: results from the randomized phase III Children's oncology group trial AAML0531. J Clin Oncol. 2014;32(27):3021–32.
12. Liu LP, Zhang AL, Ruan M, Chang LX, Liu F, Chen X, et al. Prognostic stratification of molecularly and clinically distinct subgroup in children with acute monocytic leukemia. Cancer Med. 2020;9(11):3647–55.
13. Creutzig U, van den Heuvel-Eibrink MM, Gibson B, Dworzak MN, Adachi S, de Bont E, et al. Diagnosis and management of acute myeloid leukemia in children and adolescents: recommendations from an international expert panel. Blood. 2012;120(16):3187–205.
14. Elsayed AH, Rafiee R, Cao X, Raimondi S, Downing JR, Ribeiro R, et al. A six-gene leukemic stem cell score identifies high risk pediatric acute myeloid leukemia. Leukemia. 2020;34(3):735–45.
15. Wang Y, Hu F, Li JY, Nie RC, Chen SL, Cai YY, et al. Systematic construction and validation of a metabolic risk model for prognostic prediction in acute myelogenous leukemia. Front Oncol. 2020;10:540.
16. Chen X, Wang F, Zhang Y, Wang M, Tian W, Teng W, et al. Panoramic view of common fusion genes in a large cohort of Chinese *de novo* acute myeloid leukemia patients. Leukemia & Lymphoma. 2019;60(4):1071–8.
17. Liu T, Rao J, Hu W, Cui B, Cai J, Liu Y, et al. Distinct genomic landscape of Chinese pediatric acute myeloid leukemia impacts clinical risk classification. Nat Commun. 2022;13(1):1640.
18. Klein K, Beverloo HB, Zimmermann M, Raimondi SC, Neuhoff C, Haas V, et al. Prognostic significance of chromosomal abnormalities at relapse in children with relapsed acute myeloid leukemia: a retrospective cohort study of the relapsed AML 2001/01 study. Pediatr Blood Cancer. 2022;69(1):e29341.
19. Johnston DL, Alonzo TA, Gerbing RB, Hirsch B, Heerema NA, Ravindranath Y, et al. Outcome of pediatric patients with acute myeloid leukemia (AML) and −5/5q− abnormalities from five pediatric AML treatment protocols: a report from the Children's oncology group. Pediatr Blood Cancer. 2013;60(12):2073–8.
20. Hasle H, Alonzo TA, Auvrignon A, Behar C, Chang M, Creutzig U, et al. Monosomy 7 and deletion 7q in children and adolescents with acute myeloid leukemia: an international retrospective study. Blood. 2007;109(11):4641–7.
21. Shimada A, Iijima-Yamashita Y, Tawa A, Tomizawa D, Yamada M, Norio S, et al. Risk-stratified therapy for children with FLT3-ITD-positive acute myeloid leukemia: results from the JPLSG AML-05 study. Int J Hematol. 2018;107(5):586–95.
22. Manara E, Basso G, Zampini M, Buldini B, Tregnago C, Rondelli R, et al. Characterization of children with FLT3-ITD acute myeloid leukemia: a report from the AIEOP AML-2002 study group. Leukemia. 2017;31(1):18–25.
23. Sexauer AN, Tasian SK. Targeting FLT3 signaling in childhood acute myeloid leukemia. Front Pediatr. 2017;5:248.
24. McCall D, Roth M, Mahadeo KM, Toepfer L, Nunez C, Short NJ, et al. Gilteritinib combination therapies in pediatric patients with FLT3-mutated acute myeloid leukemia. Blood Adv. 2021;5(23):5215–9.
25. Hollink IHIM, Van Den Heuvel-Eibrink MM, Zimmermann M, Balgobind BV, Arentsen-Peters STCJM, Alders M, et al. Clinical relevance of Wilms tumor 1 gene mutations in childhood acute myeloid leukemia. Blood. 2009;113(23):5951–60.

26. Solh M, Yohe S, Weisdorf D, Ustun C. Core-binding factor acute myeloid leukemia: heterogeneity, monitoring, and therapy. Am J Hematol. 2014;89(12):1121–31.

27. Jourdan E, Boissel N, Chevret S, Delabesse E, Renneville A, Cornillet P, et al. Prospective evaluation of gene mutations and minimal residual disease in patients with core binding factor acute myeloid leukemia. Blood. 2013;121(12):2213–23.

28. Pollard JA, Alonzo TA, Gerbing RB, Ho PA, Zeng R, Ravindranath Y, et al. Prevalence and prognostic significance of KIT mutations in pediatric patients with core binding factor AML enrolled on serial pediatric cooperative trials for de novo AML. Blood. 2010;115(12):2372–9.

29. Klusmann JH, Reinhardt D, Zimmermann M, Kremens B, Vormoor J, Dworzak M, et al. The role of matched sibling donor allogeneic stem cell transplantation in pediatric high-risk acute myeloid leukemia: results from the AML-BFM 98 study. Haematologica. 2012;97(1):21–9.

30. Horan JT, Alonzo TA, Lyman GH, Gerbing RB, Lange BJ, Ravindranath Y, et al. Impact of disease risk on efficacy of matched related bone marrow transplantation for pediatric acute myeloid leukemia: the Children's oncology group. J Clin Oncol. 2008;26(35):5797–801.

31. Heuser M, Freeman SD, Ossenkoppele GJ, Buccisano F, Hourigan CS, Ngai LL, et al. 2021 update on MRD in acute myeloid leukemia: a consensus document from the European LeukemiaNet MRD working party. Blood. 2021;138(26):2753–67.

32. El Achi H, Kanagal-Shamanna R. Biomarkers in acute myeloid leukemia: leveraging next generation sequencing data for optimal therapeutic strategies. Front Oncol. 2021;11:748250.

33. Sasaki K, Kanagal-Shamanna R, Montalban-Bravo G, Assi R, Jabbour E, Ravandi F, et al. Impact of the variant allele frequency of ASXL1, DNMT3A, JAK2, TET2, TP53, and NPM1 on the outcomes of patients with newly diagnosed acute myeloid leukemia. Cancer. 2020;126(4):765–74.

34. Karol SE, Alexander TB, Budhraja A, Pounds SB, Canavera K, Wang L, et al. Venetoclax in combination with cytarabine with or without idarubicin in children with relapsed or refractory acute myeloid leukaemia: a phase 1, dose-escalation study. Lancet Oncol. 2020;21(4):551–60.

35. Liu H. Emerging agents and regimens for AML. J Hematol Oncol. 2021;14(1):49.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.