Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

MRD Negativity Does Not Homogenize Prognosis in Pediatric AML: A TARGET-AML Study

Jing-wen Wang
Sun Yat-sen University
Yu Li
Sun Yat-sen University
Hong Dong
Sun Yat-sen University
Lu-Hong Xu
xulhong@mail.sysu.edu.cn

Sun Yat-sen University

Research Article

Keywords: Pediatric acute myeloid leukemia, Measurable residual disease, CBF-AML, Genetic subtype, Prognosis, Interaction

Posted Date: November 1st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11264577/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

The 2025 ELN-DAVID consensus emphasizes genetic-subgroup-specific MRD interpretation in AML. Whether genetic subtypes retain prognostic information after early MFC-MRD negativity in pediatric AML remains unclear.

Methods

Among 885 children with AML in the TARGET-AML database, the primary analysis cohort comprised 295 patients achieving MFC-MRD < 0.05% at both end of Course 1 and Course 2 after excluding KMT2A-r (CBF-AML 123, non-CBF/non-KMT2A 172). Multivariable Cox and Fine-Gray models were used for EFS, OS, and CIR; Fine-Gray results are reported as subdistribution hazard ratios (sHR). MRD × genetic subtype interaction was assessed in the extended MRD-evaluable cohort (Course 1: n = 679; Course 2: n = 594), because MRD status has only one level in the double-negative cohort.

Results

CBF-AML was associated with significantly better OS (HR 0.50, 95% CI 0.28–0.89, P = 0.019) but not EFS (HR 0.88, P = 0.546) or CIR (sHR 0.96, 95% CI 0.63–1.46, P = 0.860). Event composition analysis showed similar proportions of patients with relapse (30.1% vs 30.8%), but among patients who relapsed, the proportion of post-relapse death was lower in CBF-AML (37.8% vs 69.8%; Fisher P = 0.005; crude OR 3.74, 95% CI 1.44–10.15, for non-CBF/non-KMT2A vs CBF-AML), suggesting that the OS association may be primarily related to a lower proportion of post-relapse death rather than reduced relapse risk. The OS association remained essentially unchanged after excluding CR1 HSCT recipients (HR 0.50, P = 0.019) and after additional adjustment for FLT3-ITD and NPM1 (HR 0.48, P = 0.017). MRD status significantly modified the association between genetic subtype and EFS at the end of both Course 1 (LRT P = 0.030) and Course 2 (LRT P = 0.0005); MRD positivity was strongly associated with worse EFS in non-CBF/non-KMT2A (Course 1 HR 2.94; Course 2 HR 4.60) and KMT2A-r (Course 1 HR 2.54; Course 2 HR 2.11), whereas no statistically significant association was observed in CBF-AML (Course 1 HR 1.28; Course 2 HR 0.90).

Conclusions

Among children with AML achieving early MFC-MRD negativity, CBF-AML was associated with significantly better OS, but this association was not accompanied by improvements in EFS or CIR, possibly reflecting a lower proportion of post-relapse death rather than reduced relapse risk. MRD status modified the prognostic impact of genetic subtype, consistent with the broader principle of the ELN-

DAVID 2025 framework; however, this study does not validate a specific MRD threshold. These findings suggest that MRD negativity does not define a uniform risk state in pediatric AML within this historical cohort.

Background

Prognostic stratification in pediatric acute myeloid leukemia (AML) traditionally relies on genetic features and early treatment response. The 2025 ELN-DAVID MRD consensus emphasizes that MRD interpretation should be contextual and tailored to genetic subgroups, rather than applying a uniform MRD standard across all AML[1]. This principle reflects the biological heterogeneity within AML: the same MRD threshold may carry different prognostic meaning across genetic contexts. However, in large historical cohorts such as TARGET-AML, CBF-specific qPCR-MRD data are not routinely available, and MFC-MRD remains the widely used method for assessing early treatment response[2]. This raises the question of whether genetic subtypes retain prognostic information under a conventional MFC-MRD negativity definition; this question has not yet been systematically validated in a large pediatric AML cohort restricted to early MFC-MRD double negativity.

CBF-AML was selected as the primary focus of this study for three reasons. First, CBF-AML accounts for approximately 30% of pediatric AML and is the largest favorable-risk subgroup, yet multiple independent studies report a relapse rate of around 30%[3–6], suggesting that "favorable genetics plus MRD negativity" does not equate to absence of relapse risk[7]. Second, CBF-AML exhibits substantial internal molecular heterogeneity: the review by Hu et al. (2025) noted that RUNX1::RUNX1T1 and CBFB::MYH11 differ significantly in co-mutation profiles, with ASXL1/ASXL2 and cohesin mutations predominantly found in RUNX1::RUNX1T1 rather than CBFB::MYH11; Yang et al. (2025) also noted that the two subtypes differ in clinical characteristics and treatment response[8, 9], indicating that the two subtypes should not be simply merged. Third, Liao et al. (2026) reassessed the MFC-MRD threshold in the TARGET database (n = 1,205) and found that 0.05% outperformed the conventional 0.1% [2]; however, this study analyzed the overall cohort and did not validate this conclusion within specific genetic subgroups such as CBF-AML, nor did it assess whether genetic subtypes retain prognostic information after MFC-MRD negativity.

Recent studies suggest that MRD negativity may not fully offset the adverse prognosis of specific high-risk genetic subtypes. For example, Abla et al. observed that among children with MLLT10 fusion AML, long-term outcomes remained significantly inferior to other subtypes even after achieving EOI-2 MFC-MRD negativity[10]; Kato et al. demonstrated in RUNX1::RUNX1T1 pediatric AML that KIT exon 17 mutations retained adverse prognostic impact after MFC-MRD negativity[11], with a similar trend reported by Srinivasan et al. in an independent cohort[12]. van Weelderen et al. demonstrated in 1,130 KMT2A-r AML patients that MRD and fusion partner independently predicted prognosis[13]; Cheng et al. demonstrated in East Asian pediatric AML that multiple high-risk genetic subtypes were independently associated with worse 5-year OS[14]. Together, these findings suggest that early MFC-MRD negativity does not necessarily imply a uniform risk state across genetic subtypes.

However, no study has systematically compared the long-term prognosis of CBF-AML with other genetic subtypes among children with AML who have achieved early MFC-MRD double negativity in a large pediatric cohort. This study aimed to answer the following question: among children with AML achieving early MFC-MRD negativity (< 0.05% at both end of Course 1 and Course 2), do genetic subtypes still retain independent prognostic information? Using CBF-AML as the primary validation target and non-CBF/non-KMT2A as the comparator, we evaluated differences in EFS, OS, and CIR, and explored potential mechanisms through event composition analysis. In addition, we performed a genetic × MRD interaction analysis in the extended MRD-evaluable cohort (Course 1: n = 679; Course 2: n = 594) to assess whether MRD status modifies the prognostic association of genetic subtypes.

Methods

Data Source

The TARGET-AML database (project ID: TARGET-AML) was accessed through the Genomic Data Commons (GDC) portal and the TARGET Data Matrix. Clinical data including MRD, genetics, treatment protocol, and survival outcomes were extracted. MFC-MRD data in TARGET-AML were derived from three clinical trials: CCG-2961, AAML03P1, and AAML0531. This study was conducted under the TARGET data use agreement and required no additional ethical approval.

Study Population

Three analysis cohorts were defined based on the TARGET-AML database. This database does not include patients with acute promyelocytic leukemia (APL, FAB M3). All cohorts excluded patients aged ≥ 18 years, those lacking MFC-MRD data, or those with missing key genetic data. Primary analysis cohort: inclusion criteria were (1) age < 18 years; (2) diagnosis of AML; (3) MFC-MRD < 0.05% at both end of Course 1 and end of Course 2 (MRD double negativity); (4) exclusion of KMT2A-r to avoid potential contamination of the non-CBF comparator by the internal fusion partner heterogeneity within this subgroup, thereby focusing the comparison between CBF-AML and a relatively homogeneous non-CBF/non-KMT2A population. After excluding KMT2A-r from the 365 MRD double-negative patients, the final cohort comprised 295 patients (123 CBF-AML, 172 non-CBF/non-KMT2A). Interaction analysis cohort: inclusion criteria were (1) age < 18 years; (2) diagnosis of AML; (3) available MFC-MRD data (regardless of negativity). This cohort did not exclude KMT2A-r, in order to assess the interaction of MRD status across CBF-AML, KMT2A-r, and non-CBF/non-KMT2A subtypes. Because MRD status has only one level (all negative) in the double-negative cohort, the interaction term cannot be estimated within that subset; therefore, interaction analysis was performed in the extended MRD-evaluable cohort. Course 1 had 679 patients with MFC-MRD data and defined genetic subtype, and Course 2 had 594 patients. Threshold comparison cohort: inclusion criteria were (1) age < 18 years; (2) diagnosis of CBF-AML; (3) available MFC-MRD data (regardless of negativity). This cohort included CBF-AML patients from the primary analysis cohort and additionally included CBF-AML patients who were MRD-positive or had MRD

data at only one timepoint, in order to compare EFS under the 0.05% versus 0.1% thresholds. A total of 200 patients were included (Course 1: 152 negative / 48 positive; Course 2: 154 negative / 46 positive).

Variable Definitions

MRD status: MFC-MRD from the TARGET database was used. MFC-MRD < 0.05% was defined as negative and ≥ 0.05% as positive. Importantly, MRD negativity in this study refers to MFC-MRD negativity and is not equivalent to the RUNX1::RUNX1T1 or CBFB::MYH11 qPCR-MRD negativity recommended by ELN-DAVID 2025 for CBF-AML. ELN-DAVID 2025 recommends ≥ 3-log reduction from baseline for relapse-risk stratification in CBF-AML; however, the TARGET database does not provide quantitative MRD at diagnosis, precluding the use of this standard in the present study. Genetic subtypes: CBF-AML (RUNX1::RUNX1T1 + CBFB::MYH11), KMT2A-r (any KMT2A fusion), non-CBF/non-KMT2A (all other patients who were neither CBF-AML nor KMT2A-r, including rare fusions such as NUP98::NSD1, CBFA2T3::GLIS2, and unclassified cases), KMT2A-HRPG (KMT2A::MLLT10 + KMT2A::AFDN + KMT2A::AFF1). KMT2A-HRPG is a high-risk subgroup of KMT2A-r. Endpoints: EFS (time from diagnosis to relapse, induction failure, second malignancy, or death from any cause, whichever occurred first), OS (time from diagnosis to death from any cause), CIR (cumulative incidence of relapse, with NRM as the competing event). Covariates: age (years), WBC at diagnosis ($\times 10^9$/L), gender, treatment protocol (AAML03P1/AAML0531), HSCT in first CR. In sensitivity analyses, Cox models were additionally adjusted for FLT3-ITD and NPM1 mutation status. Missing data: Baseline characteristics were compared using available data without imputation. The primary and sensitivity analyses used complete-case analysis.

Statistical Analysis

Continuous variables were expressed as median (IQR), and categorical variables as n (%). Baseline comparisons used the Wilcoxon rank-sum test (continuous) and Pearson's Chi-squared test (categorical). Primary analysis: in the 295-patient primary analysis cohort, multivariable Cox proportional hazards models compared EFS and OS between CBF-AML and non-CBF/non-KMT2A, adjusted for age, WBC, gender, treatment protocol, and HSCT in first CR. Fine-Gray competing-risk models compared CIR with NRM as the competing event; Fine-Gray results are reported as subdistribution hazard ratios (sHR). Sensitivity analyses: (1) excluding patients who underwent HSCT in first CR, the Cox models were repeated; (2) using all non-CBF AML (including KMT2A-r) as comparator, the Cox models were repeated; (3) the Cox model was additionally adjusted for FLT3-ITD and NPM1 mutation status. Threshold comparison: among 200 CBF-AML patients with available MRD data, MRD negativity was defined at 0.05% and 0.1% thresholds, and EFS was compared between the two thresholds. Interaction analysis: in the interaction analysis cohort, likelihood ratio tests (LRT) compared Cox models with and without the genetic × MRD interaction term. Simple effects were estimated to calculate HRs for MRD + vs MRD− within each genetic subtype. Event composition analysis in the primary cohort tabulated relapse events, NRM, and post-relapse deaths by group; Fisher's exact test compared post-relapse death proportions between groups. All analyses used R 4.3.0; two-sided $P < 0.05$ was considered significant.

Results

Study Population

The TARGET-AML database contained 885 pediatric AML patients. MRD data were available in 743 patients; 467 had MRD < 0.05% at end of Course 1, 485 at end of Course 2, and 365 (41.2%) at both timepoints. After excluding KMT2A-r, the primary analysis cohort comprised 295 patients (123 CBF-AML, 172 non-CBF/non-KMT2A). In the interaction analysis cohort, 679 patients were evaluable at Course 1 and 594 at Course 2. The threshold comparison cohort comprised 200 CBF-AML patients with available MRD data (Course 1: 152 negative / 48 positive; Course 2: 154 negative / 46 positive). The study flowchart is shown in Fig. 1.

Baseline Characteristics

In the primary analysis cohort, CBF-AML and non-CBF/non-KMT2A were balanced for age, gender, bone marrow blast percentage, and treatment protocol, but differed significantly in WBC (median 34.9 vs 18.5 $\times 10^9$/L, $P = 0.018$), FLT3-ITD mutation rate (4.1% vs 23%, $P < 0.001$), NPM1 mutation rate (0% vs 19%, $P < 0.001$), and CR1 HSCT rate (0% vs 23%, $P < 0.001$). No CBF-AML patient received HSCT in first CR (Table 1). In addition, we explored the prognostic difference between KMT2A-HRPG and NHRPG within the MRD double-negative cohort (Table S1, Figure S1); given the limited sample size, this analysis was reported as exploratory only.

Table 1
Baseline characteristics of the primary analysis cohort (n = 295)

Characteristic	Overall N = 295[1]	non-CBF/non-KMT2A N = 172[1]	CBF-AML N = 123[1]	p-value[2]
Age, years	10.3 (4.5–14.5)	10.4 (2.9–14.9)	10.2 (5.8–13.7)	0.647
Gender				0.898
Female	151 (51)	87 (51)	64 (52)	
Male	144 (49)	85 (49)	59 (48)	
WBC, ×10⁹/L	23.8 (10.2–67.8)	18.5 (8.5–56.0)	34.9 (11.4–78.2)	0.018
Bone marrow blasts, %	65.0 (45.0–80.0)	64.1 (46.0–81.0)	67.5 (45.0–78.0)	0.857
Missing	16	11	5	
Treatment protocol				0.854
AAML03P1	36 (12)	22 (13)	14 (11)	
AAML0531	259 (88)	150 (87)	109 (89)	
FLT3-ITD mutation	44 (15)	39 (23)	5 (4.1)	< 0.001
NPM1 mutation	32 (11)	32 (19)	0 (0)	< 0.001
HSCT in first CR	39 (13)	39 (23)	0 (0)	< 0.001
Fusion gene				< 0.001
RUNX1::RUNX1T1	63 (21)	0 (0)	63 (51)	
CBFB::MYH11	60 (20)	0 (0)	60 (49)	
Other/Unclassified	172 (58)	172 (100)	0 (0)	

[1]Median (Q1–Q3); n (%), [2]Wilcoxon rank sum test; Pearson's Chi-squared test

The primary analysis cohort was restricted to children with AML achieving early MFC-MRD negativity (< 0.05% at both end of Course 1 and Course 2) after excluding KMT2A-r. Continuous variables are median (IQR); categorical variables are n (%). Wilcoxon rank-sum test for continuous; Pearson's Chi-squared test for categorical. WBC, white blood cell count; HSCT, hematopoietic stem cell transplantation; CR, complete remission; MFC-MRD, multiparameter flow cytometry measurable residual disease.

CBF-AML vs non-CBF/non-KMT2A

In multivariable Cox analysis, CBF-AML was associated with significantly better OS compared with non-CBF/non-KMT2A (HR 0.50, 95% CI 0.28–0.89, *P* = 0.019), but EFS did not differ significantly (HR 0.88, 95% CI 0.58–1.34, P = 0.546). Fine-Gray competing-risk analysis showed no significant difference in CIR (sHR 0.96, 95% CI 0.63–1.46, *P* = 0.860). In sensitivity analyses, the OS association remained essentially unchanged after excluding patients who underwent HSCT in CR1 (HR 0.50, 95% CI 0.28–0.89, *P* = 0.019); because no CBF-AML patient received HSCT in CR1, this analysis primarily excluded HSCT recipients in the non-CBF/non-KMT2A group and therefore does not imply that CR1 HSCT is ineffective. When all non-CBF AML (including KMT2A-r) served as the comparator, the association with better OS was stronger (HR 0.44, 95% CI 0.25–0.76, *P* = 0.003), whereas EFS remained non-significant (HR 0.76, 95% CI 0.51–1.12, *P* = 0.165); in a complete-case analysis additionally adjusted for FLT3-ITD and NPM1 mutation status (n = 295; EFS events = 105; OS events = 64), the OS association remained stable (HR 0.48, 95% CI 0.26–0.88, *P* = 0.017), with EFS remaining non-significant (HR 0.83, 95% CI 0.54–1.29, *P* = 0.415) (Table S2). These results are summarized in Table 2, Table 3, Fig. 2, Fig. 3, and Fig. 4. Complete estimates for all covariates in the multivariable Cox model are shown in Figure S2; full results of the sensitivity analysis excluding CR1 HSCT recipients are provided in Table S3.

Table 2
Multivariable Cox regression for EFS and OS in the primary cohort (n = 295)

Endpoint	Variable	HR (95% CI)	P value
EFS	CBF-AML vs non-CBF/non-KMT2A	0.88 (0.58–1.34)	0.546
	Age, per year	1.03 (0.99–1.07)	0.106
	WBC at diagnosis	1.00 (1.00–1.00)	0.340
	Male vs female	0.82 (0.56–1.20)	0.311
	AAML0531 vs AAML03P1	0.83 (0.48–1.44)	0.507
	HSCT in first CR	1.15 (0.65–2.02)	0.640
OS	CBF-AML vs non-CBF/non-KMT2A	0.50 (0.28–0.89)	0.019
	Age, per year	1.04 (1.00–1.09)	0.061
	WBC at diagnosis	1.00 (1.00–1.00)	0.846
	Male vs female	0.64 (0.39–1.06)	0.081
	AAML0531 vs AAML03P1	0.99 (0.46–2.13)	0.987
	HSCT in first CR	1.46 (0.77–2.77)	0.252

HR, hazard ratio; CI, confidence interval; EFS, event-free survival; OS, overall survival. HSCT in first CR was included as a covariate to account for treatment selection.

Table 3
Fine-Gray competing-risk model for CIR in the primary cohort (n = 295)

Variable	sHR (95% CI)	P value
CBF-AML vs non-CBF/non-KMT2A	0.96 (0.63–1.46)	0.860
Age, per year	1.02 (0.98–1.06)	0.240
WBC at diagnosis	1.00 (1.00–1.00)	0.120
Male vs female	0.81 (0.53–1.22)	0.310
AAML0531 vs AAML03P1	0.77 (0.43–1.37)	0.370

CIR, cumulative incidence of relapse; NRM, non-relapse mortality (competing event). sHR, subdistribution hazard ratio. HSCT in first CR was not included as a covariate because it is a post-treatment, time-dependent variable.

Event Composition Analysis

To explore the pattern of better OS without significant differences in EFS or CIR between CBF-AML and non-CBF/non-KMT2A, we analyzed event composition in the primary analysis cohort. The proportion of patients with relapse was similar between groups (37/123 [30.1%] vs 53/172 [30.8%]). However, among patients who relapsed, the proportion of post-relapse death was significantly lower in CBF-AML than in non-CBF/non-KMT2A (14/37 [37.8%] vs 37/53 [69.8%]; Fisher $P = 0.005$; crude OR 3.74, 95% CI 1.44–10.15, for non-CBF/non-KMT2A vs CBF-AML). NRM events were infrequent in both groups (3 in CBF-AML vs 10 in non-CBF/non-KMT2A). This event-composition pattern suggests that the association between CBF-AML and better OS may be primarily related to a lower proportion of post-relapse death rather than to a lower risk of relapse; however, this interpretation should be made cautiously given limited information on post-relapse treatment in TARGET (Table S4).

Threshold Comparison

Among 200 CBF-AML patients with available MRD data, 200 evaluable MRD results were available at both Course 1 and Course 2. MRD classification was based directly on the original records in the TARGET database and was not re-adjudicated from raw MRD values. In this dataset, MRD negativity/positivity classifications defined by the 0.05% and 0.1% thresholds were identical (Course 1: 152 negative / 48 positive; Course 2: 154 negative / 46 positive). One patient had an MRD1 value and one had an MRD2 value between 0.05% and 0.1%; both were recorded as MRD-negative in the original database. Therefore, in this dataset, when the original MRD records were used, the two thresholds did not yield different classifications, and this analysis could not determine whether the 0.05% threshold provides additional prognostic discrimination over 0.1%. With MRD positivity as the reference in unadjusted Cox models, MRD negativity at the end of Course 1 was not significantly associated with EFS (HR 1.10, 95% CI 0.63–1.90, $P = 0.739$), whereas MRD negativity at the end of Course 2 showed a borderline significant association with improved EFS (HR 0.61, 95% CI 0.37–1.01, $P = 0.056$), with imprecise estimates. These results are summarized in Table S5 and Figure S3.

Secondary Analysis

Among MRD double-negative patients, EFS and OS were compared between RUNX1::RUNX1T1 (n = 63, OS events = 9) and CBFB::MYH11 (n = 60, OS events = 8) (Table S6, Figure S4). Using CBFB::MYH11 as the reference in multivariable Cox models adjusted for age, WBC, sex, and treatment protocol, no statistically significant difference was observed between RUNX1::RUNX1T1 and EFS (HR 0.78, 95% CI 0.38–1.62, *P* = 0.504). For OS, the point estimate suggested that RUNX1::RUNX1T1 might have worse OS (HR 2.07, 95% CI 0.62–6.93, *P* = 0.239), but this did not reach statistical significance, and the wide confidence interval indicates an imprecise estimate and limited statistical power. OS events occurred in 9 of 63 patients (14.3%) in RUNX1::RUNX1T1 and 8 of 60 patients (13.3%) in CBFB::MYH11. Therefore, a potential difference in OS between the two subtypes cannot be excluded and warrants validation in larger cohorts.

Interaction Analysis

Because MRD status has only one level in the MRD double-negative cohort, interaction analysis was performed in the extended MRD-evaluable cohort (Course 1: n = 679; Course 2: n = 594). The number of MRD-positive/negative patients and events in each subgroup are shown in Table 4. All models were multivariable Cox models adjusted for age, WBC, sex, treatment protocol, and HSCT in first CR. Likelihood ratio tests showed that MRD status significantly modified the association between genetic subtype and EFS (Course 1: LRT P = 0.030; Course 2: LRT P = 0.0005). In simple effect analyses using MRD negativity as the reference, MRD positivity was significantly associated with worse EFS in non-CBF/non-KMT2A (Course 1: HR 2.94, 95% CI 2.23–3.88; Course 2: HR 4.60, 95% CI 3.34–6.34) and KMT2A-r (Course 1: HR 2.54, 95% CI 1.49–4.33; Course 2: HR 2.11, 95% CI 1.11–4.01); in CBF-AML, no statistically significant association was observed between MRD positivity and EFS (Course 1: HR 1.28, 95% CI 0.71–2.30; Course 2: HR 0.90, 95% CI 0.32–2.50), although the wide confidence intervals cannot exclude a potential association. Results are summarized in Table 4, Fig. 5, and Fig. 6. Full regression estimates for the four interaction models are provided in Table S7 (Course 1) and Table S8 (Course 2).

Table 4
MRD x genetic subtype interaction and simple effects on EFS

MRD timepoint	Genetic subtype	MRD contrast	n (MRD+/MRD−)	EFS events (MRD+/MRD−)	HR (95% CI)	P value
End of Course 1	non-CBF/non-KMT2A	MRD + vs MRD-	155/220	116/94	2.94 (2.23–3.88)	< 0.001
	CBF-AML	MRD + vs MRD-	32/151	14/56	1.28 (0.71–2.30)	0.411
	KMT2A-r	MRD + vs MRD-	25/96	19/51	2.54 (1.49–4.33)	< 0.001
End of Course 2	non-CBF/non-KMT2A	MRD + vs MRD-	76/246	63/108	4.60 (3.34–6.34)	< 0.001
	CBF-AML	MRD + vs MRD-	14/153	4/50	0.90 (0.32–2.50)	0.840
	KMT2A-r	MRD + vs MRD-	17/88	12/47	2.11 (1.11–4.01)	0.022
Course 1 interaction LRT P = 0.030; Course 2 interaction LRT P = 0.0005.						
Interaction analysis was performed in the MRD-evaluable cohort (Course 1: n = 679; Course 2: n = 594). All models were multivariable Cox models adjusted for age, WBC, sex, treatment protocol, and HSCT in first CR. Reference groups: non-CBF/non-KMT2A and MRD-negative.						

Discussion

This study revealed four core findings in children with AML achieving early MFC-MRD negativity (MFC-MRD < 0.05% at both the end of Course 1 and the end of Course 2): (1) CBF-AML was associated with significantly better OS but comparable EFS and CIR compared with non-CBF/non-KMT2A; (2) event composition analysis showed that, among patients who relapsed, the proportion of post-relapse death was lower in CBF-AML (37.8% vs 69.8%; Fisher *P* = 0.005), suggesting that the OS association may be primarily related to a lower proportion of post-relapse death rather than reduced relapse risk; (3) the prognostic impact of MRD status differed substantially across genetic subtypes, with significant modification of the genetic subtype–EFS association at the end of both Course 1 (LRT *P* = 0.030) and Course 2 (LRT *P* = 0.0005); MRD positivity was strongly associated with worse EFS in non-CBF/non-KMT2A and KMT2A-r, whereas no statistically significant association was observed in CBF-AML; and (4) in CBF-AML, the 0.05% and 0.1% thresholds produced identical classifications in this dataset, with only one patient each having MRD between 0.05% and 0.1%, limiting assessment of whether the lower

threshold provides additional prognostic discrimination. These findings suggest that early MFC-MRD negativity does not represent a uniform risk state in this cohort, and that genetic subtype may retain prognostic information for OS after MRD negativity.

The 2025 ELN-DAVID consensus provides a contextual framework for MRD interpretation. It no longer applies a uniform MRD standard across all AML, but rather provides method- and target-specific MRD thresholds and introduces new qualitative MRD response categories (optimal, warning, high risk of treatment failure) to facilitate contextual MRD interpretation[1]. For CBF-AML specifically, ELN-DAVID 2025 recommends qPCR-MRD for RUNX1::RUNX1T1 or CBFB::MYH11 and uses ≥ 3-log reduction from baseline for relapse-risk stratification, rather than a generic MFC-MRD threshold[1]. The consensus also reports that 11%–27% of MRD-negative RUNX1::RUNX1T1 patients may re-emerge with detectable MRD within 24 months, with similar observations for CBFB::MYH11. This is supported by qPCR-based MRD monitoring studies in pediatric CBF-AML, in which MRD re-emergence predicted hematological relapse[15]. In addition, ELN-DAVID specifically recommends DNA-based NGS for FLT3-ITD MRD testing, noting that ultra-high-sensitivity NGS-MRD can identify patients at increased risk[16], further indicating that different genetic subtypes require specific MRD detection methods.

Our findings are consistent with this framework: MRD positivity was strongly associated with worse EFS in non-CBF/non-KMT2A and KMT2A-r, whereas no statistically significant association was observed in CBF-AML. MRD negativity in our study was defined as MFC-MRD < 0.05% in the TARGET database, which is not equivalent to the CBF-specific molecular MRD negativity recommended by ELN-DAVID 2025. Therefore, our study does not directly validate the ELN-DAVID 2025 consensus or any specific MRD threshold. Rather, it provides clinical cohort evidence supporting the broader principle that MRD interpretation should be contextualized by genetic background. In addition, ELN-DAVID 2025 reports that even MRD-negative CBF-AML patients may re-emerge with detectable MRD, which is consistent with our observation that CBF-AML retains relapse risk after MRD negativity (37/123, 30.1%), further supporting the concept that MRD negativity is not the endpoint of disease risk.

Comparison with prior literature can be organized around three themes. First, MRD negativity does not eliminate the adverse impact of high-risk genetics. MLLT10/KMT2A::MLLT10 fusions and high-risk KMT2A-r fusion partners have been associated with poor outcomes despite MFC-MRD negativity[10, 13, 17]. In our cohort, these subgroups were too small for definitive multivariable analysis and were reported as exploratory only. Second, internal heterogeneity within CBF-AML remains debated. KIT exon 17 has been associated with inferior outcomes in some RUNX1::RUNX1T1 cohorts[11, 12], but was not replicated in another pediatric CBF-AML cohort[18], and RUNX1::RUNX1T1 and CBFB::MYH11 differ biologically and clinically[8, 9]. In our MRD double-negative cohort, no statistically significant EFS or OS difference was observed between RUNX1::RUNX1T1 (n = 63) and CBFB::MYH11 (n = 60), but the OS point estimate suggested that RUNX1::RUNX1T1 might have worse OS (HR 2.07, 95% CI 0.62–6.93, *P* = 0.239), with wide confidence intervals that cannot exclude a potential difference; OS events occurred in 9/63 (14.3%) and 8/60 (13.3%) patients, respectively. This direction is consistent with a recent pediatric CBF-AML study[19], although KIT exon 17 enrichment in that cohort may confound the comparison. Larger

samples are needed for confirmation. Third, our study further suggests that even after MFC-MRD double negativity, CBF-AML and non-CBF/non-KMT2A differ in OS, and this difference appears to arise after relapse rather than from a higher relapse rate. The mechanisms remain uncertain.

This study has three clinical implications, although they should be interpreted cautiously. (1) Risk stratification: after early MFC-MRD negativity, CBF-AML should not automatically be regarded as a biologically homogeneous low-risk state; the proportion of patients with relapse in our CBF-AML cohort was still 30.1%, indicating that MRD negativity does not equate to absence of relapse risk. (2) MRD monitoring: MRD interpretation in CBF-AML may require genetic-specific molecular monitoring rather than reliance exclusively on generic MFC-MRD thresholds; future studies should evaluate the complementary value of qPCR-MRD in this population[20]. (3) Relapse management: the lower proportion of post-relapse death observed in CBF-AML may reflect disease biology, greater sensitivity to salvage therapy, differences in post-relapse treatment strategies, or other unmeasured factors; these possibilities cannot be distinguished in the present dataset, and no inference about optimal salvage therapy can be made. Because no CBF-AML patient in our cohort underwent HSCT in CR1, our data cannot directly evaluate the comparative benefit of CR1 transplantation in this subgroup. Nevertheless, the OS association persisted in sensitivity analysis, suggesting that differences in CR1 transplantation alone are unlikely to fully explain the observed association.

This study has several limitations. First, MRD negativity was defined by MFC-MRD < 0.05%, which is not equivalent to the CBF-specific qPCR-MRD negativity or ≥ 3-log reduction recommended by ELN-DAVID 2025. Second, the primary cohort was restricted to early MRD double negativity, and interaction analyses used a different MRD-evaluable cohort; the findings therefore apply only to this selected population. Third, MFC-MRD data were derived from three historical trials with substantial missingness (365/885, 41.2% analyzable); notably, MRD testing accessibility itself may be associated with mortality risk in pediatric AML (adjusted HR 1.7, 95% CI 1.1–2.5)[21], suggesting that MRD missingness may not be purely statistical. Fourth, CR1 HSCT was a post-treatment decision; although excluding HSCT recipients did not change the OS association, residual confounding cannot be excluded, and HSCT was not included in the Fine-Gray model to avoid immortal time bias. Fifth, the threshold comparison relied on original MRD records without re-adjudication, so we could not determine whether 0.05% is superior to 0.1%. Sixth, the post-relapse mortality analysis was conditional on relapse and may be affected by selection and survivor bias, with limited salvage treatment data. Finally, the non-CBF/non-KMT2A comparator remained genetically heterogeneous, and several small subgroups were exploratory only.

Among children with AML achieving early MFC-MRD negativity (< 0.05% at both the end of Course 1 and the end of Course 2) within this historical TARGET-AML cohort, CBF-AML was associated with significantly better OS, but this association was not accompanied by corresponding improvements in EFS or CIR, possibly reflecting a lower proportion of post-relapse death (37.8% vs 69.8%; Fisher P = 0.005) rather than reduced relapse risk. The prognostic impact of MRD status differed substantially across genetic subtypes, with significant modification of the genetic subtype–EFS association at the end of both Course 1 and Course 2; MRD positivity was strongly associated with worse EFS in non-CBF/non-

KMT2A and KMT2A-r, whereas no statistically significant association was observed in CBF-AML. These findings are consistent with the broader principle emphasized by ELN-DAVID 2025 that MRD interpretation should be contextualized by genetic background, but they do not validate a specific MRD threshold. In this cohort, early MFC-MRD negativity does not represent a uniform risk state, and genetic subtype may retain prognostic information for OS after MRD negativity.

Abbreviations

AML
acute myeloid leukemia
APL
acute promyelocytic leukemia
CBF-AML
core-binding factor acute myeloid leukemia
CI
confidence interval
CIR
cumulative incidence of relapse
CR
complete remission
CR1
first complete remission
EFS
event-free survival
ELN
European LeukemiaNet
ELN-DAVID
European LeukemiaNet-DAVID MRD Working Party
EOI
end of induction
FAB
French-American-British classification
FLT3-ITD
FMS-like tyrosine kinase 3 internal tandem duplication
GDC
Genomic Data Commons
HR
hazard ratio
HSCT
hematopoietic stem cell transplantation

IQR
interquartile range
KMT2A
lysine methyltransferase 2A
KMT2A-r
KMT2A-rearranged
KMT2A-HRPG
KMT2A high-risk partner genes
LRT
likelihood ratio test
MFC-MRD
multiparameter flow cytometry measurable residual disease
MRD
measurable residual disease
NGS
next-generation sequencing
NHRPG
non-high-risk partner genes
NPM1
nucleophosmin 1
NRM
non-relapse mortality
OR
odds ratio
OS
overall survival
qPCR
quantitative polymerase chain reaction
sHR
subdistribution hazard ratio
WBC
white blood cell count

Declarations

Acknowledgement

Thanks are due to TARGET researchers as well as all study participants for public data.

Availability of data and materials

The datasets generated and/or analyzed during the current study are available from the TARGET website: https://target-data.nci.nih.gov/Public/AML/clinical/.

Author Contributions

L.H.X and J.W.W participated in project design, analysis, interpretation, and manuscript drafting. Y.L and H.D obtained, assembled data, analyzed and interpreted the data. All authors contributed to the article and approved the submitted version.

Author information

Jing-wen Wang, Yu-Li and Hong Dong contributed equally to this work.

Conflict of Interest

The authors declare that there is no conflict of interest.

Funding

No funding.

Ethics declarations

Ethic approval and consent for publication

The studies involving human participants were reviewed and approved by TARGET Publications Committee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Clinical trial information

COG studies AAML0531, AAML03P1, and CCG2961 were registered at www.clinicaltrials.gov as #NCT00372593, #NCT00070174, and #NCT00002798, respectively.

References

1. Cloos J, Valk PJM, Thiede C, Dohner K, Roboz GJ, Wood BL, Walter RB, Wang S, Wierzbowska A, Wei AH, et al. 2025 update on MRD in acute myeloid leukemia: a consensus document from the ELN-DAVID MRD Working Party. Blood. 2026;147(11):1147–67.
2. Liao XY, Zheng H, Fang JP, Zhou DH, Qiu KY. Optimal minimal residual disease threshold in pediatric acute myeloid leukemia: A retrospective cohort study based on the TARGET database. PLoS Med. 2026;23(5):e1005088.
3. Tomizawa D, Tawa A, Watanabe T, Saito AM, Kudo K, Taga T, Iwamoto S, Shimada A, Terui K, Moritake H, et al. Excess treatment reduction including anthracyclines results in higher incidence of

relapse in core binding factor acute myeloid leukemia in children. Leukemia. 2013;27(12):2413–6.

4. Chen X, Dou H, Wang X, Huang Y, Lu L, Bin J, Su Y, Zou L, Yu J, Bao L. KIT mutations correlate with adverse survival in children with core-binding factor acute myeloid leukemia. Leuk Lymphoma. 2018;59(4):829–36.
5. Yamato G, Shiba N, Yoshida K, Shiraishi Y, Hara Y, Ohki K, Okubo J, Okuno H, Chiba K, Tanaka H, et al. ASXL2 mutations are frequently found in pediatric AML patients with t(8;21)/ RUNX1-RUNX1T1 and associated with a better prognosis. Genes Chromosomes Cancer. 2017;56(5):382–93.
6. Conneely SE, Rau RE. The genomics of acute myeloid leukemia in children. Cancer Metastasis Rev. 2020;39(1):189–209.
7. Karol SE, Coustan-Smith E, Cao X, Shurtleff SA, Raimondi SC, Choi JK, Ribeiro RC, Dahl GV, Bowman WP, Taub JW, et al. Prognostic factors in children with acute myeloid leukaemia and excellent response to remission induction therapy. Br J Haematol. 2015;168(1):94–101.
8. Yang M, Wang W, Zhang G, Qiu S, Liu B, Mi Y, Wang Y, Wang J, Wei H. Clinical characteristics and therapeutic determinants of RUNX1::RUNX1T1 differ from those of CBFB::MYH11 acute myeloid leukemia. Haematologica. 2025;110(10):2281–92.
9. Hu Z, Tang X, Chen F, Li T, Liu Y, Zhou G, Liu S, Wang Y, Liu S, Mai H, et al. Molecular genetics profiling of core-binding factor acute myeloid leukemia in pediatrics. Ther Adv Hematol. 2025;16:20406207251330064.
10. Abla O, Ries RE, Triche T Jr., Gerbing RB, Hirsch B, Raimondi S, Cooper T, Farrar JE, Buteyn N, Harmon LM, et al. Structural variants involving MLLT10 fusion are associated with adverse outcomes in pediatric acute myeloid leukemia. Blood Adv. 2024;8(8):2005–17.
11. Kato S, Tsujimoto SI, Matsubayashi J, Iwamoto S, Hiramatsu H, Okuno Y, Kamitori T, Ohki K, Deguchi T, Kiyokawa N, et al. Adverse prognostic impact of KIT exon 17 mutations despite negative flow cytometric measurable residual disease in pediatric acute myeloid leukemia with RUNX1::RUNX1T1. Haematologica. 2025;110(1):251–6.
12. Srinivasan S, Dhamne C, Patkar N, Chatterjee G, Moulik NR, Chichra A, Pallath A, Tembhare P, Shetty D, Subramanian PG, et al. KIT exon 17 mutations are predictive of inferior outcome in pediatric acute myeloid leukemia with RUNX1::RUNX1T1. Pediatr Blood Cancer. 2024;71(2):e30791.
13. van Weelderen RE, Klein K, Harrison CJ, Jiang Y, Abrahamsson J, Arad-Cohen N, Bart-Delabesse E, Buldini B, De Moerloose B, Dworzak MN, et al. Measurable Residual Disease and Fusion Partner Independently Predict Survival and Relapse Risk in Childhood KMT2A-Rearranged Acute Myeloid Leukemia: A Study by the International Berlin-Frankfurt-Munster Study Group. J Clin Oncol. 2023;41(16):2963–74.
14. Cheng Z, Yu SL, Yu CH, Chang TC, Chang YH, Jou ST, Lu MY, Lin KH, Chang HH, Chou SW, et al. Genetic landscape of pediatric acute myeloid leukemia in Taiwan. Sci Rep. 2026;16(1):4024.
15. Juul-Dam KL, Ommen HB, Nyvold CG, Walter C, Vålerhaugen H, Kairisto V, Abrahamsson J, Alm SJ, Jahnukainen K, Lausen B, et al. Measurable residual disease assessment by qPCR in peripheral

blood is an informative tool for disease surveillance in childhood acute myeloid leukaemia. Br J Haematol. 2020;190(2):198–208.

16. Hourigan CS, Beugelink L, Othman J, Gui G, Ivey A, Dillon R, Thiede C, Levis MJ, Dillon LW, Wei AH, et al. ELN-DAVID recommendations for NGS-based FLT3-ITD MRD testing for patients with acute myeloid leukemia. Blood. 2026;148(7):910–3.
17. Steidel E, Orhan E, Rasche M, Pigazzi M, Tregnago C, Hoffmeister LM, Walter C, Dworzak M, Muhlegger N, von Neuhoff N, et al. Prognostic Value of Molecular Genetic Measurable Residual Disease (MRD) Monitoring in Pediatric Acute Myeloid Leukemia Expressing KMT2A::MLLT10. Eur J Haematol. 2025;115(5):493–504.
18. Zheng Y, Yu L, Le S, Li N, Feng X, Li C, Zheng M, Mai H, Jiang H, He X et al. KIT and FLT3-ITD mutations do not predict outcomes in pediatric core-binding factor acute myeloid leukemia: findings from the C-HUANAN-AML-15 multicenter cohort study. Ann Hematol 2026, 105(9).
19. Hou B, Yang F, Zhang R, Qin M, Li J, Wang Y, Chen H, Li J, Qi P, Jia C, et al. Prospective Evaluation of KIT Mutations and Long-Term Outcomes in Pediatric Core Binding Factor Acute Myeloid Leukemia: A Single Institutional Study in China. Pediatr Hematol Oncol. 2026;43(2):76–88.
20. Maurer-Granofszky M, Kohrer S, Fischer S, Schumich A, Nebral K, Larghero P, Meyer C, Mecklenbrauker A, Muhlegger N, Marschalek R, et al. Genomic breakpoint-specific monitoring of measurable residual disease in pediatric non-standard-risk acute myeloid leukemia. Haematologica. 2024;109(3):740–50.
21. Rujkijyanont P, Ukritchon S, Winaichatsak A, Kuwatjanakul P, Chotsampancharoen T, Techavichit P, Chainansamit SO, Sathitsamitphong L, Kanchanakamhaeng K, Anurathapan U, et al. Hyperleukocytosis and Access to Minimal Residual Disease Testing Impact Outcomes in Children With Newly Diagnosed Acute Myeloid Leukemia in Thailand. Pediatr Blood Cancer. 2026;73(3):e70004.

Figures

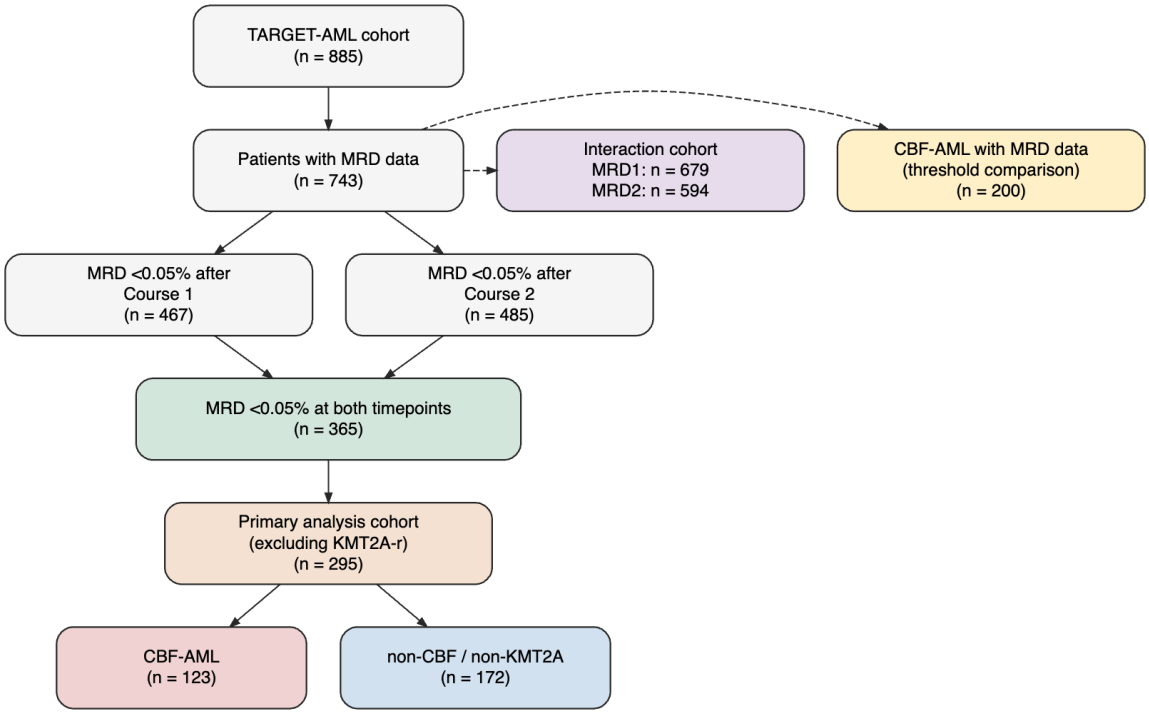


Figure 1

Study flowchart. Study flowchart. Among 885 pediatric AML patients in the TARGET-AML database, 743 had MRD data; 467 had MRD <0.05% at the end of Course 1, 485 at the end of Course 2, and 365 at both timepoints. After excluding KMT2A-r, the primary analysis cohort comprised 295 patients (CBF-AML 123, non-CBF/non-KMT2A 172). Interaction analysis cohort: Course 1 n = 679; Course 2 n = 594. Threshold comparison cohort: 200 CBF-AML patients with available MRD data.

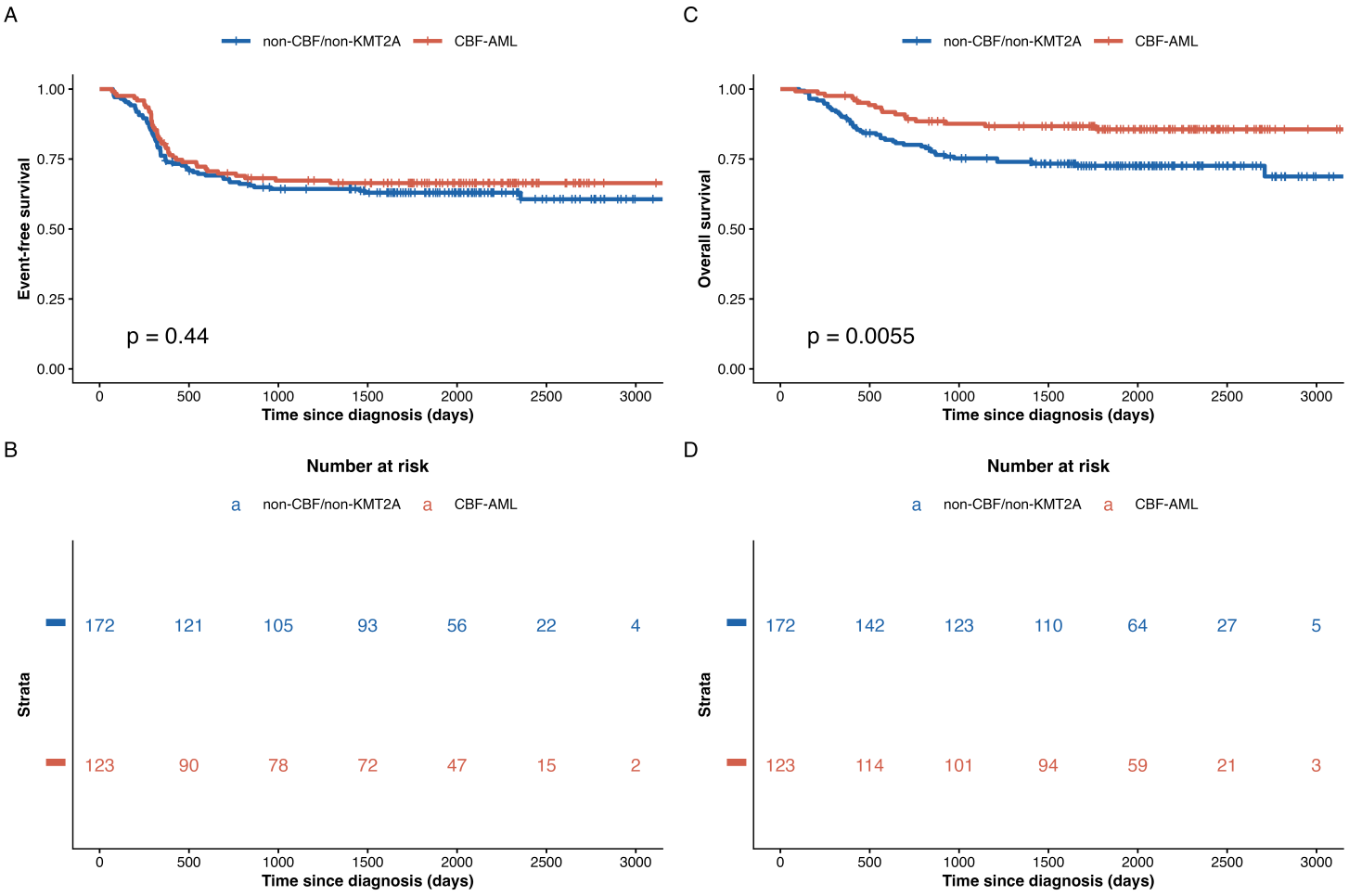


Figure 2

Kaplan-Meier curves for EFS and OS in the primary analysis cohort. Kaplan-Meier curves for EFS and OS comparing CBF-AML with non-CBF/non-KMT2A in the primary analysis cohort. (A) EFS; (B) number at risk for EFS; (C) OS; (D) number at risk for OS. P values are from the log-rank test. CBF-AML was associated with significantly better OS (P = 0.0055) but not EFS (P = 0.44).

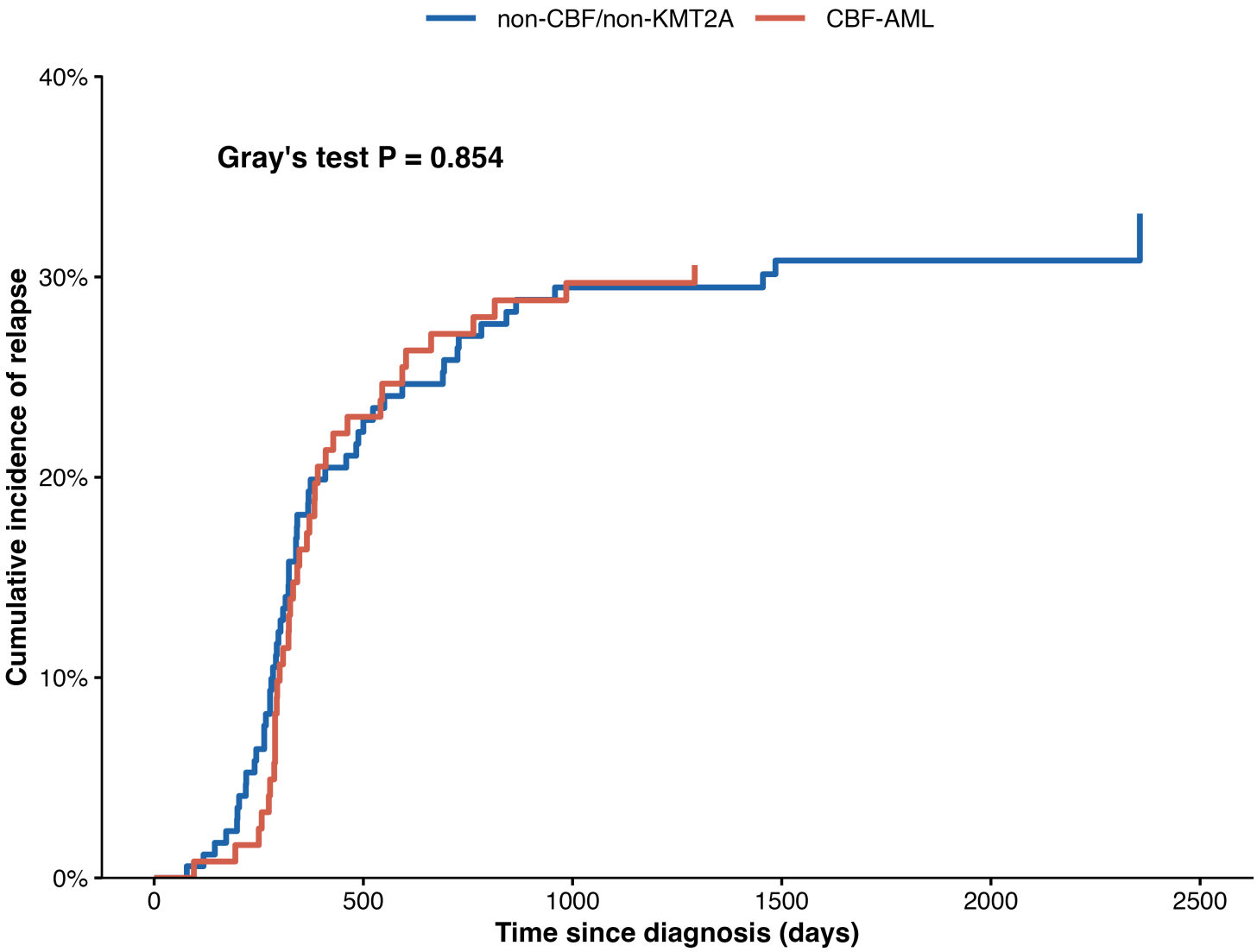


Figure 3

Cumulative incidence of relapse in the primary analysis cohort. Cumulative incidence of relapse (CIR) in the primary analysis cohort comparing CBF-AML with non-CBF/non-KMT2A, with non-relapse mortality (NRM) as the competing event. P value is from Gray's test (P = 0.854). No significant difference in CIR was observed between groups.

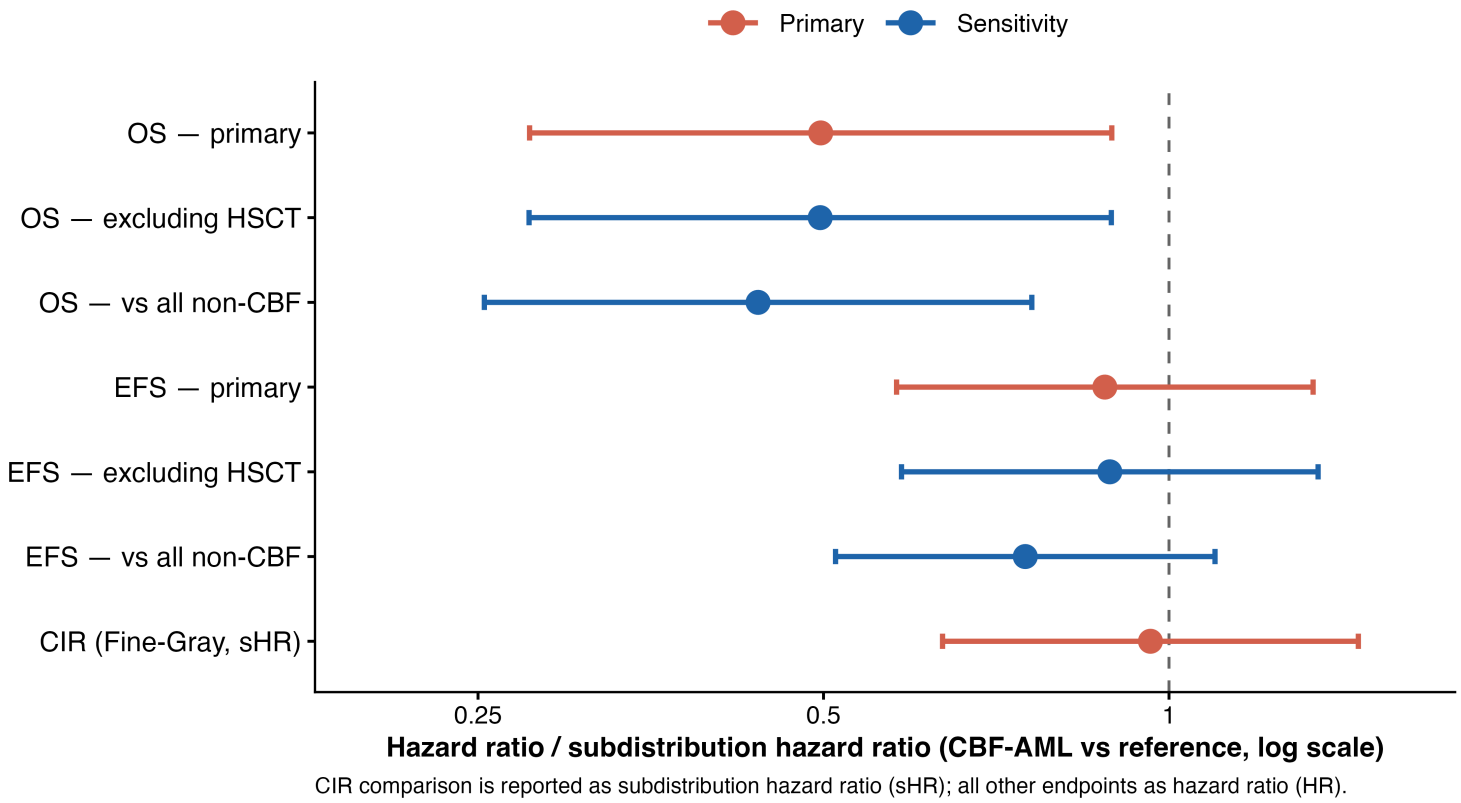


Figure 4

Multivariable and sensitivity analyses for EFS, OS, and CIR. Forest plot of EFS, OS, and CIR for CBF-AML versus the comparator in primary and sensitivity analyses. CIR is reported as subdistribution hazard ratio (sHR); all other endpoints are reported as hazard ratio (HR). Primary: primary analysis; Sensitivity: sensitivity analyses (excluding CR1 HSCT recipients; using all non-CBF AML as comparator). Error bars indicate 95% CI.

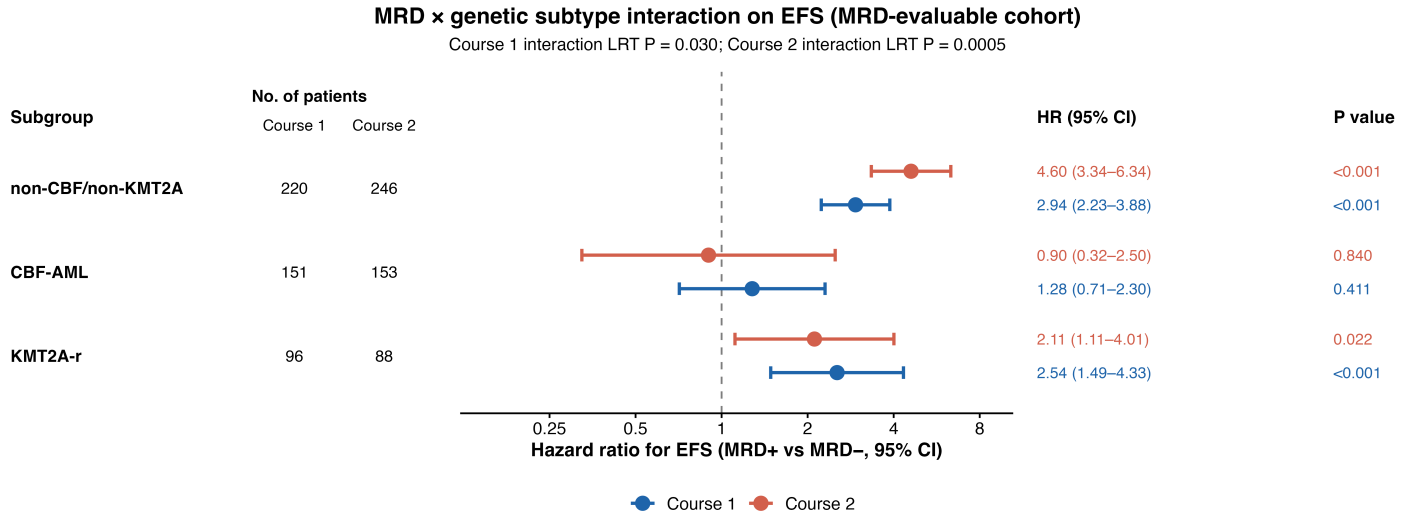


Figure 5

MRD × genetic subtype interaction on EFS. Forest plot of the MRD × genetic subtype interaction on EFS in the MRD-evaluable cohort. Hazard ratios (HR) and 95% CI are shown for MRD positivity versus MRD negativity (reference: MRD-negative) within each genetic subtype. Course 1 interaction LRT P = 0.030; Course 2 interaction LRT P = 0.0005. Numbers of patients per subgroup are shown on the left; HRs and P values on the right.

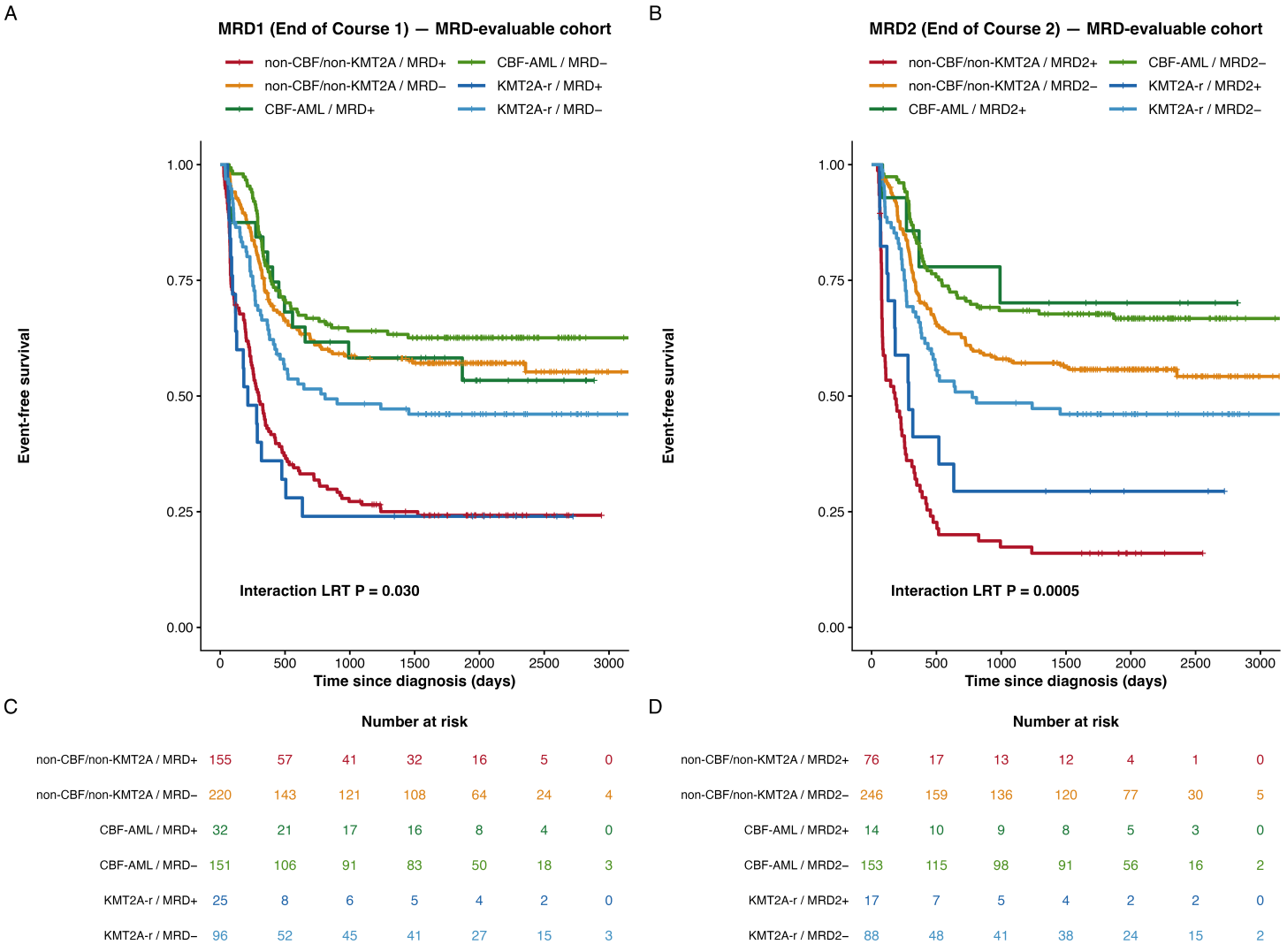


Figure 6

Kaplan-Meier curves for EFS by MRD status and genetic subtype. Kaplan-Meier curves for EFS stratified by genetic subtype and MRD status in the MRD-evaluable cohort. (A) Course 1 (MRD1); (B) Course 2 (MRD2); (C) number at risk for MRD1; (D) number at risk for MRD2. Interaction LRT P values are annotated. MRD+, MRD positivity; MRD−, MRD negativity.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigS1.png
- FigS2.png
- FigS3.png
- FigS4.png
- FigS5.png
- FigS6.png
- TableS1.docx
- TableS10.docx
- TableS2.docx
- TableS3.docx
- TableS4.docx
- TableS5.docx
- TableS6.docx
- TableS7.docx
- TableS8.docx
- TableS9.docx