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.

# Delayed Early MRD Clearance and Outcomes in Pediatric Acute Myeloid Leukemia: A Retrospective Cohort Study

**jingwen wang**
Sun Yat-sen Memorial Hospital
**Yu li**
Sun Yat-sen Memorial Hospital
**Hong Dong**
Sun Yat-sen Memorial Hospital
**Lu-Hong Xu**
xulhong@mail.sysu.edu.cn

Sun Yat-sen Memorial Hospital

**Research Article**

**Keywords:** acute myeloid leukemia, pediatric, measurable residual disease, MRD kinetics, competing risk, relapse

**Posted Date:** November 3rd, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11357128/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: Measurable residual disease (MRD) is an established prognostic marker in pediatric acute myeloid leukemia (AML), but whether the trajectory of MRD clearance provides prognostic information beyond MRD status at a later treatment time point remains uncertain.

Methods: We conducted a retrospective cohort study of 418 children with newly diagnosed AML in the TARGET-AML database who had paired quantitative multiparameter flow cytometry-based MRD measurements at the end of induction 1 (EOI1) and EOI2. Using a 0.05% threshold, patients were classified as Neg/Neg, Pos/Neg, or Pos/Pos according to EOI1/EOI2 MRD status; 7 patients with a Neg/Pos pattern were included in the Pos/Pos group because of the small subgroup size. The primary outcome was 5-year cumulative incidence of relapse (CIR), with non-relapse mortality as a competing event.

Results: Compared with Neg/Neg, Pos/Neg was associated with higher CIR after adjustment for age, ELN 2022 risk, white blood cell count, and treatment protocol (subdistribution hazard ratio, 1.66; 95% CI, 1.14–2.42; P=.008). Pos/Neg was also associated with inferior overall survival (hazard ratio, 2.03; P=.003) and event-free survival (hazard ratio, 1.89; P<.001). Among patients who were MRD-negative at EOI2, Pos/Neg remained associated with higher CIR (subdistribution hazard ratio, 1.70; 95% CI, 1.16–2.47; P=.006). Continuous ΔMRD did not improve time-dependent discrimination beyond categorical MRD kinetics.

Conclusions: Early MRD trajectory was associated with relapse risk beyond EOI2 MRD status alone. Categorical MRD clearance pattern, but not continuous ΔMRD, may refine risk stratification in pediatric AML. Prospective studies are warranted to evaluate longitudinal MRD assessment for risk stratification in pediatric AML.


# Key Points

- Pos/Neg MRD kinetics were associated with higher relapse risk and inferior OS and EFS despite subsequent EOI2 MRD negativity.
- Early MRD trajectory provided prognostic information beyond EOI2 MRD status, whereas continuous ΔMRD did not improve model discrimination.

# Introduction

Pediatric acute myeloid leukemia (AML), remains a major therapeutic challenge, as relapse continues to be a leading cause of treatment failure despite improvements in risk stratification and treatment(1) Identifying patients at high risk of relapse during therapy is therefore important for refining risk-adapted treatment strategies.

Measurable residual disease (MRD) provides an objective measure of treatment response and has been consistently associated with outcomes in pediatric AML(2). Multiparametric flow cytometry (MFC) is a widely used approach for MRD assessment in pediatric AML(3), and recent reviews have emphasized the importance of standardized MRD interpretation in pediatric acute leukemia (4). In the Children's Oncology Group AAML0531 study, early MFC-MRD was strongly associated with survival and relapse outcomes, supporting its role in response assessment during induction therapy(5). However, the prognostic interpretation of MFC-MRD depends on both the threshold used to define positivity and the timing of assessment(2, 3).

MFC-MRD is an independent prognostic factor in pediatric AML (6–9), but its optimal threshold remains uncertain. A TARGET-AML study (10) showed that a 0.05% threshold provided better prognostic discrimination than 0.1% after induction. These findings provided the rationale for using 0.05% as the prespecified threshold in the present analysis. However, patients with the same MRD status at a later treatment time point may have followed different response trajectories. Whether the trajectory of early MRD clearance provides additional prognostic information beyond the MRD status at the later assessment remains unclear.

Recent evidence in adult AML suggests that the timing of MRD clearance may be clinically relevant. Jen et al. reported that patients who were MRD-positive after induction but became MRD-negative after subsequent chemotherapy had inferior relapse-free survival compared with patients who were MRD-negative at both assessments(11). Whether this observation also applies to pediatric AML has not been established.

In this study, we investigated the prognostic value of early MRD clearance patterns in children with AML who had paired MFC-MRD measurements at the end of induction 1 (EOI1) and end of induction 2 (EOI2).Using the 0.05% threshold evaluated in the recent TARGET-based analysis(10), we compared patients who were MRD-negative at both assessments (Neg/Neg) with those who converted from MRD-positive to MRD-negative (Pos/Neg) and those with persistent MRD positivity(Pos/Pos). We further evaluated whether the continuous change in MRD burden between EOI1 and EOI2 provided additional prognostic information beyond categorical MRD status.

# Methods

Study Design and Population

This was a multicenter, retrospective cohort study using the TARGET-AML public database (NCI TARGET project; COG trials AAML03P1 and AAML0531). Inclusion criteria were: (1) newly diagnosed pediatric AML (age < 18 years), non-M3; (2) morphological complete remission (CR; bone marrow blasts < 5%) after induction; (3) paired quantitative MFC-MRD results at EOI1 and EOI2; and (4) treatment on AAML03P1 or AAML0531. Exclusion criteria were: (1) missing key baseline covariates; and (2) BCR-ABL1 positivity. Patients who did not achieve CR, including those with induction failure, or who lacked paired EOI1/EOI2 MRD data were excluded because they did not meet the inclusion criteria. AAML03P1 and

AAML0531 were selected because they used uniform MFC-MRD assessment and contemporary treatment strategies; patients treated on CCG-2961 were excluded because MRD data were largely unavailable and treatment differed substantially. A total of 418 patients were included.

Risk stratification

ELN 2022 risk(12, 13) was assigned retrospectively based on available cytogenetic and molecular data. In the TARGET-AML database, the original risk categories were Low, Standard, and High, which correspond to ELN 2022 Low, Intermediate, and Adverse, respectively.

MRD Assessment and Grouping

MFC-MRD was assessed by multiparameter flow cytometry. Using the 0.05% threshold, patients were classified as: Neg/Neg (both EOI1 and EOI2 < 0.05%), Pos/Neg (EOI1 ≥ 0.05%, EOI2 < 0.05%), and Pos/Pos (both ≥ 0.05%). The Neg/Pos group (n = 7) was merged into Pos/Pos due to small sample size.

MRD kinetic groups were defined a priori according to MFC-MRD status at EOI1 and EOI2 as Neg/Neg, Pos/Neg, and Pos/Pos, representing persistent MRD negativity, early MRD clearance, and persistent MRD positivity, respectively. These study-specific categories were used to characterize early MRD clearance trajectories and were not intended to directly reproduce the qualitative MRD response categories proposed by the ELN-DAVID MRD Working Party.

Calculation of ΔMRD

$\Delta MRD = \log2(MRD_EOI1 + 10^{-6}) - \log2(MRD_EOI2 + 10^{-6})$. Undetectable values were assigned $10^{-6}$. A larger ΔMRD indicates faster clearance.

Endpoint Definitions

The primary endpoint was 5-year cumulative incidence of relapse (CIR), defined as time from diagnosis to first hematological relapse, with non-relapse mortality (NRM) as the competing event. Secondary endpoints included event-free survival (EFS), and overall survival (OS).

# Statistical Analysis

CIR was estimated using the Aalen-Johansen estimator and compared with Gray's test. Univariable Fine-Gray competing risk models were used to assess each variable. Three prespecified multivariable Fine-Gray models (14) were constructed, all adjusted for age, ELN 2022 risk, log10(WBC + 1), and treatment protocol: Model 1: MRD kinetics group (Neg/Neg, Pos/Neg, Pos/Pos) plus covariates. Model 2: static EOI2 MRD status (≥ 0.05% vs < 0.05%) plus covariates. Model 3: static EOI2 MRD status plus ΔMRD plus covariates. A Model 4 was further constructed to assess whether the association of ΔMRD was independent of EOI1 MRD burden by additionally adjusting for log10(EOI1 MRD + $10^{-6}$).

Model comparison between Model 2 and Model 3 used the 5-year time-dependent AUC estimated with the timeROC package, categorical and continuous NRI(15), pseudo-AIC, and a likelihood ratio test based on pseudo-likelihood. OS and EFS were analyzed with multivariable Cox proportional hazards models with robust standard errors. Interaction tests were performed using cause-specific Cox models with robust standard errors

Because the TARGET-AML public database does not provide the exact date of EOI2 assessment, a formal EOI2 landmark analysis could not be performed. We therefore performed fixed-time landmark sensitivity analyses at 3, 6, and 9 months from diagnosis. At each landmark, patients who had experienced relapse or death before the landmark were excluded, and CIR was recalculated from the landmark time. Sensitivity analyses included excluding early relapses within 6 months, alternative ΔMRD definitions, alternative thresholds, cause-specific Cox models, excluding the Neg/Pos group, and multiple imputation. All tests were two-sided with $p < .05$ considered significant. R 4.3 + was used.

# Results

Baseline Characteristics

A total of 418 patients were included. Median age was 9.6 years (IQR 2.7–13.9), and 207 (49.5%) were male. ELN 2022 risk: Low 186 (44.5%), Intermediate 184 (44.0%), Adverse 48 (11.5%). MRD kinetics groups: Neg/Neg 307 (73.4%), Pos/Neg 56 (13.4%), Pos/Pos 55 (13.2%). The distribution of EOI1 and EOI2 MRD values is shown in Fig. 1. During follow-up, 165 relapses (39.5%) and 21 non-relapse deaths (5.0%) occurred. Significant differences were observed among groups in WBC, ELN 2022 risk, FLT3-ITD, inv(16), and SCT rates (Table 1).

Table 1
Baseline clinical and genetic characteristics of 418 pediatric AML patients stratified by MRD kinetics (Neg/Neg, Pos/Neg, Pos/Pos).

| Characteristic | Overall | Neg/Neg | Pos/Neg | Pos/Pos | p-value |
|---|---|---|---|---|---|
| | N = 418 | N = 307 | N = 56 | N = 55 | |
| **Age, years** | **9.6 (2.7–13.9)** | **9.4 (2.9–13.8)** | **8.0 (2.0-12.5)** | **11.2 (3.5–14.7)** | **.255** |
| **Gender** | | | | | **.995** |
| Male | 207 (49.5%) | 152 (49.5%) | 28 (50.0%) | 27 (49.1%) | |
| Female | 211 (50.5%) | 155 (50.5%) | 28 (50.0%) | 28 (50.9%) | |
| **WBC, x10^9/L** | **27.3 (11.1–81.4)** | **23.9 (10.8–75.1)** | **43.1 (19.1–138.0)** | **27.6 (11.6-104.6)** | **.022** |
| **Bone marrow blasts, %** | **70.0 (49.0–85.0)** | **69.5 (48.8–82.8)** | **73.0 (49.5–90.0)** | **71.1 (49.0-87.2)** | **.844** |
| **ELN 2022 risk** | | | | | **< .001** |
| Low | 186 (44.5%) | 151 (49.2%) | 24 (42.9%) | 11 (20.0%) | |
| Intermediate | 184 (44.0%) | 130 (42.3%) | 24 (42.9%) | 30 (54.5%) | |
| Adverse | 48 (11.5%) | 26 (8.5%) | 8 (14.3%) | 14 (25.5%) | |
| **Treatment protocol** | | | | | **.657** |
| AAML03P1 | 50 (12.0%) | 37 (12.1%) | 5 (8.9%) | 8 (14.5%) | |
| AAML0531 | 368 (88.0%) | 270 (87.9%) | 51 (91.1%) | 47 (85.5%) | |
| **FLT3-ITD mutation** | | | | | **.002** |
| No | 354 (84.9%) | 268 (87.6%) | 48 (85.7%) | 38 (69.1%) | |
| Yes | 63 (15.1%) | 38 (12.4%) | 8 (14.3%) | 17 (30.9%) | |
| **NPM1 mutation** | | | | | **.819** |
| No | 377 (90.8%) | 276 (90.8%) | 50 (89.3%) | 51 (92.7%) | |
| Yes | 38 (9.2%) | 28 (9.2%) | 6 (10.7%) | 4 (7.3%) | |
| **t(8;21)** | | | | | **.192** |

Median (IQR) for continuous variables; n (%) for categorical variables. ELN 2022 risk was mapped from the original TARGET categories (Low/Standard/High). Kruskal-Wallis rank sum test, Pearson's Chi-squared test, or Fisher's exact test as appropriate. Missing data: bone marrow blasts (n = 23); FLT3-ITD (n = 1); NPM1 (n = 3); t(8;21) (n = 4); inv(16) (n = 4); KMT2A (n = 4). WBC, white blood cell count; SCT, stem cell transplantation; CR, complete remission; MRD, measurable residual disease.

| Characteristic | Overall | Neg/Neg | Pos/Neg | Pos/Pos | p-value |
|---|---|---|---|---|---|
| | N = 418 | N = 307 | N = 56 | N = 55 | |
| No | 342 (82.6%) | 245 (80.6%) | 49 (89.1%) | 48 (87.3%) | |
| Yes | 72 (17.4%) | 59 (19.4%) | 6 (10.9%) | 7 (12.7%) | |
| **inv**(16) | | | | | **.035** |
| No | 350 (84.5%) | 249 (81.9%) | 49 (89.1%) | 52 (94.5%) | |
| Yes | 64 (15.5%) | 55 (18.1%) | 6 (10.9%) | 3 (5.5%) | |
| **KMT2A rearrangement** | | | | | **.346** |
| No | 350 (84.5%) | 255 (83.9%) | 50 (90.9%) | 45 (81.8%) | |
| Yes | 64 (15.5%) | 49 (16.1%) | 5 (9.1%) | 10 (18.2%) | |
| **SCT in first CR** | | | | | **.001** |
| No | 350 (83.7%) | 269 (87.6%) | 39 (69.6%) | 42 (76.4%) | |
| Yes | 68 (16.3%) | 38 (12.4%) | 17 (30.4%) | 13 (23.6%) | |

Median (IQR) for continuous variables; n (%) for categorical variables. ELN 2022 risk was mapped from the original TARGET categories (Low/Standard/High). Kruskal-Wallis rank sum test, Pearson's Chi-squared test, or Fisher's exact test as appropriate. Missing data: bone marrow blasts (n = 23); FLT3-ITD (n = 1); NPM1 (n = 3); t(8;21) (n = 4); inv(16) (n = 4); KMT2A (n = 4). WBC, white blood cell count; SCT, stem cell transplantation; CR, complete remission; MRD, measurable residual disease.

Cumulative Incidence of Relapse and Non-Relapse Mortality

The 5-year CIR was 34.5% for Neg/Neg, 53.7% for Pos/Neg, and 49.8% for Pos/Pos (Gray's test p=.004; Fig. 2, Supplementary Table S1). The corresponding 5-year NRM was 2.9%, 9.1%, and 12.8%, respectively (Gray's test p=.003). Pos/Neg had the highest CIR, whereas Pos/Pos had the highest NRM, partly explaining its numerically lower CIR compared with Pos/Neg.

Univariable and Multivariable Fine-Gray Analyses

In univariable analysis, both Pos/Neg (SHR 1.78, 95% CI 1.21–2.60, p = .003) and Pos/Pos (SHR 1.66, 95% CI 1.08–2.57, p = .022) were significantly associated with CIR (Table 2). In multivariable analysis adjusted for age, ELN 2022, logWBC, and protocol, Pos/Neg remained independently associated (SHR 1.66, 95% CI 1.14–2.42, p = .008), whereas Pos/Pos was not (SHR 1.41, 95% CI 0.89–2.22, p = .145) (Table 3, Model 1). In Model 2, static EOI2 MRD positivity (≥ 0.05% vs < 0.05%) was not significantly associated with CIR (SHR 1.28, 95% CI 0.82–1.99, p = .284). In Model 3, ΔMRD was independently associated (SHR 1.03, 95% CI 1.00–1.06, p = .021), while static EOI2 MRD positivity remained non-significant (SHR 1.40, 95% CI 0.88–2.21, p = .157) (Table 3, Model 3).

Table 2
Univariable Fine-Gray competing risk analysis for cumulative incidence of relapse.

| Variable | Comparison | SHR (95% CI) | p-value |
|---|---|---|---|
| **MRD kinetics** | **Pos/Neg vs Neg/Neg** | **1.78 (1.21–2.60)** | **.003** |
| **MRD kinetics** | **Pos/Pos vs Neg/Neg** | **1.66 (1.08–2.57)** | **.022** |
| **ΔMRD (per log2 unit)** | **per log2 unit** | **1.03 (1.00-1.06)** | **.032** |
| **Age** | **per year** | **0.99 (0.96–1.02)** | **.535** |
| **Gender** | **Female vs Male** | **0.96 (0.71–1.31)** | **.807** |
| **log10(WBC + 1)** | **per unit** | **1.66 (1.26–2.20)** | **< 0.001** |
| **Bone marrow blasts (%)** | **per %** | **1.01 (1.00-1.02)** | **.011** |
| **ELN 2022 risk** | **Intermediate vs Low** | **2.06 (1.47–2.89)** | **< 0.001** |
| **ELN 2022 risk** | **Adverse vs Low** | **2.08 (1.28–3.40)** | **.003** |
| **Treatment protocol** | **AAML0531 vs AAML03P1** | **0.97 (0.62–1.50)** | **.879** |
| **FLT3-ITD** | **Yes vs No** | **1.41 (0.94–2.10)** | **.093** |
| **NPM1 mutation** | **Yes vs No** | **0.74 (0.41–1.33)** | **.316** |
| **t(8;21)** | **Yes vs No** | **0.40 (0.24–0.67)** | **< 0.001** |
| **inv**(16) | **Yes vs No** | **0.96 (0.63–1.45)** | **.843** |
| **KMT2A rearrangement** | **Yes vs No** | **1.35 (0.90–2.03)** | **.150** |
| **SCT in first CR** | **Yes vs No** | **0.94 (0.62–1.42)** | **.773** |

SHR, subdistribution hazard ratio; CI, confidence interval; MRD, measurable residual disease; WBC, white blood cell count; SCT, stem cell transplantation; CR, complete remission; ELN, European LeukemiaNet. All estimates were derived from univariable Fine-Gray competing risk models with non-relapse mortality as the competing event. Reference group: Neg/Neg for MRD kinetics; Low for ELN 2022 risk; AAML03P1 for treatment protocol; No for binary variables.

Table 3
Multivariable Fine-Gray competing risk models for cumulative incidence of relapse.

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Pos/Neg vs Neg/Neg | 1.66 (1.14–2.42) | --- | --- |
| | p .008 | | |
| Pos/Pos vs Neg/Neg | 1.41 (0.89–2.22) | --- | --- |
| | p .145 | | |
| EOI2 MRD ( > = 0.05% vs < 0.05%) | --- | 1.28 (0.82–1.99) | 1.40 (0.88–2.21) |
| | | p .284 | p .157 |
| ΔMRD (per log2 unit) | --- | --- | 1.03 (1.00-1.06) |
| | | | p .021 |
| Age (per year) | 1.02 (0.99–1.05) | 1.01 (0.98–1.04) | 1.01 (0.98–1.04) |
| | p .299 | p .349 | p .365 |
| ELN 2022 Intermediate vs Low | 2.20 (1.54–3.15) | 2.18 (1.52–3.12) | 2.18 (1.52–3.12) |
| | p < .001 | p < .001 | p < .001 |
| ELN 2022 Adverse vs Low | 1.66 (0.98–2.79) | 1.73 (1.03–2.90) | 1.65 (0.98–2.79) |
| | p .058 | p .039 | p .061 |
| log10(WBC + 1) | 1.64 (1.24–2.17) | 1.68 (1.27–2.23) | 1.66 (1.26–2.20) |
| | p .001 | p < .001 | p < .001 |
| AAML0531 vs AAML03P1 | 0.97 (0.61–1.55) | 1.00 (0.63–1.60) | 0.98 (0.62–1.56) |
| | p .913 | p .989 | p .944 |

Model 1: MRD kinetics group + covariates. Model 2: static EOI2 MRD ( > = 0.05% vs < 0.05%) + covariates. Model 3: static EOI2 MRD + ΔMRD + covariates. Covariates: age, ELN 2022 risk, log10(WBC + 1), and treatment protocol. A Model 4 further adjusted for EOI1 MRD burden is shown in Supplementary Table S2. SHR, subdistribution hazard ratio; CI, confidence interval; MRD, measurable residual disease.

Model Comparison

The 5-year AUC was 69.5% for the static model (Model 2) and 69.2% for the dynamic model (Model 3), with ΔAUC − 0.3% (95% CI − 2.2to 1.6, p=.704) (Table 4). The likelihood ratio test showed that adding ΔMRD improved model fit (LR = 4.33, df = 1, p=.037), and pseudo-AIC was slightly lower (1873.0vs 1875.3). Categorical NRI was1.6%, and continuous NRI was + 29.0%. Overall, ΔMRD improved fit but not discrimination or categorical reclassification.

Table 4
Model comparison between static and dynamic Fine-Gray models for cumulative incidence of relapse.

| Metric | Static model (m2) | Dynamic model (m3) | Comparison |
|---|---|---|---|
| 5-year AUC | 69.5% | 69.2% | −0.3% (− 2.2 to 1.6), p=.704 |
| Pseudo log-likelihood | -931.64 | -929.48 | --- |
| Pseudo-AIC | 1875.3 | 1873.0 | ΔAIC = 2.3 |
| Likelihood ratio test | --- | --- | LR = 4.33, df = 1, p = .037 |
| Categorical NRI | --- | --- | 1.6% |
| Continuous NRI | --- | --- | 29.0% |

AUC, area under the time-dependent receiver operating characteristic curve, estimated using timeROC at 5 years. NRI, net reclassification improvement; AIC, Akaike information criterion; LR, likelihood ratio. Model 2: static EOI2 MRD ( > = 0.05% vs < 0.05%) + covariates. Model 3: static EOI2 MRD + ΔMRD + covariates. The continuous NRI is exploratory; 95% CI and p-value were not calculated.

Sensitivity Analysis of ΔMRD Adjusted for EOI1 MRD Burden

In a sensitivity model further adjusted for EOI1 MRD burden (Model 4), the association of ΔMRD was attenuated and no longer statistically significant (SHR 0.959, 95% CI 0.857–1.073; p=.463), whereas EOI1 MRD itself was not independently associated with CIR (SHR 1.287, 95% CI 0.887–1.868; p=.184) (Supplementary Table S2). These findings suggest that continuous ΔMRD may partly reflect the initial MRD level rather than an independent clearance kinetic marker.

Subgroup Analysis

In the EOI2 MRD < 0.05% subgroup (n = 363), Pos/Neg was independently associated with CIR (SHR 1.70, 95% CI 1.16–2.47, p=.006; 5-year CIR 53.7% vs 34.5%; Fig. 3A, Supplementary Table S1). Stratified by ELN 2022, the association was observed in the intermediate-risk subgroup (SHR 2.28, 95% CI 1.37–3.79, p=.001), and in the AAML0531 protocol subgroup (SHR 1.65, 95% CI 1.10–2.46, p=.015), but not in the low-risk subgroup (SHR 1.25, p=.520) or in AAML03P1 (SHR 1.30, p=.659) (Fig. 3B, Table 5). Formal interaction tests for MRD kinetics × ELN 2022 (p = .238) and MRD kinetics × treatment protocol (p = .283) were not significant (Supplementary Table S3), so subgroup estimates should be interpreted as descriptive rather than as evidence of differential effects.

Table 5
Subgroup analysis of Pos/Neg versus Neg/Neg for cumulative incidence of relapse.

| Subgroup | n | SHR (95% CI) | p-value |
|---|---|---|---|
| Overall (EOI2 < 0.05%) | 363 | 1.70 (1.16–2.47) | .006 |
| ELN 2022 Low | 175 | 1.25 (0.63–2.47) | .520 |
| ELN 2022 Intermediate | 154 | 2.28 (1.37–3.79) | .001 |
| AAML03P1 | 42 | 1.30 (0.41–4.12) | .659 |
| AAML0531 | 321 | 1.65 (1.10–2.46) | .015 |

SHR, subdistribution hazard ratio; CI, confidence interval; MRD, measurable residual disease; ELN, European LeukemiaNet; EOI, end of induction. The ELN 2022 Adverse subgroup was not fitted because only 8 patients had Pos/Neg status. Formal interaction tests were not significant (MRD x ELN 2022, p = .238; MRD x Protocol, p = .283), so subgroup estimates should be interpreted as descriptive rather than as evidence of differential effects.

Secondary Endpoints

In multivariable Cox models, Pos/Neg was independently associated with inferior OS (HR 2.03, 95% CI 1.28–3.22, p=.003) and EFS (HR 1.89, 95% CI 1.32–2.71, p = < .001). Pos/Pos was also associated with inferior OS (HR 1.79, 95% CI 1.10–2.91, p=.020) and EFS (HR 1.75, 95% CI 1.16–2.64, p=.007) (Table 6, Fig. 4).

Table 6
Multivariable Cox proportional hazards models for overall survival and event-free survival.

| Endpoint | Comparison | HR (95% CI) | p-value |
|---|---|---|---|
| OS | Pos/Neg vs Neg/Neg | 2.03 (1.28–3.22) | .003 |
| OS | Pos/Pos vs Neg/Neg | 1.79 (1.10–2.91) | .020 |
| EFS | Pos/Neg vs Neg/Neg | 1.89 (1.32–2.71) | < .001 |
| EFS | Pos/Pos vs Neg/Neg | 1.75 (1.16–2.64) | .007 |

OS, overall survival; EFS, event-free survival; HR, hazard ratio; CI, confidence interval; MRD, measurable residual disease. HRs were derived from multivariable Cox proportional hazards models with robust standard errors, adjusted for age, ELN 2022 risk, log10(WBC + 1), and treatment protocol. Reference group: Neg/Neg.

Exploratory assessment of the ELN 2025 MRD relapse definition

Because the ELN-DAVID 2025 update introduced a qualitative MRD relapse definition, we explored its applicability within the Pos/Pos group. Only 7 patients met the ≥ 1-log10 rise criterion and 9 met the ≥ 0.5-log10 criterion, precluding formal statistical comparison. The estimated 5-year CIR was numerically similar between the ELN-MRD relapse and stable positive groups (42.9% vs 50.8% for the ≥ 1-log10 criterion; 44.4% vs 50.9% for the ≥ 0.5-log10 criterion) (Supplementary Table S4, Supplementary Figure

S1). These results should be regarded as exploratory rather than as a validation of the ELN 2025 criteria in children.

Sensitivity Analyses

Sensitivity analyses confirmed the robustness of the main findings (Table 7). Excluding early relapses (n = 399), 0.1% threshold, cause-specific Cox, excluding Neg/Pos (n = 411), and MICE were all consistent with the primary analysis. The only inconsistent result was the absolute difference replacing ΔMRD (SHR 0.99, 95% CI 0.96–1.01, p=.330), a known phenomenon when applying absolute differences to log-normally distributed data.

Table 7
Sensitivity analyses of the association between MRD kinetics / ΔMRD and cumulative incidence of relapse.

| Sensitivity analysis | Key result | Consistent with main analysis |
|---|---|---|
| Excluding early relapses within 6 months (n = 399) | ΔMRD SHR 1.04 (1.01–1.07); P = .019 | Yes |
| Absolute difference replacing ΔMRD | ΔMRD SHR 0.99 (0.96–1.01); P = .330 | No |
| Alternative 0.1% threshold | Pos/Neg SHR 1.70 (1.16–2.50); P = .007 | Yes |
| Cause-specific Cox model | ΔMRD HR 1.03 (1.01–1.06); P = .014 | Yes |
| Excluding Neg/Pos group (n = 411) | Pos/Neg SHR 1.66 (1.14–2.42); P = .008 | Yes |
| Multiple imputation (MICE) | ΔMRD pooled coefficient 0.036 vs 0.034 | Yes |
| Fixed-time landmark at 3 months | Pos/Neg SHR 1.62 (1.12–2.37); P = .011 | Yes |
| Fixed-time landmark at 6 months | Pos/Neg SHR 1.86 (1.25–2.75); P = .002 | Yes |
| Fixed-time landmark at 9 months | Pos/Neg SHR 2.17 (1.40–3.37); P < .001 | Yes |

SHR, subdistribution hazard ratio; HR, hazard ratio; CI, confidence interval; MRD, measurable residual disease; MICE, multiple imputation by chained equations. Fixed-time landmark analyses at 3, 6, and 9 months were performed because the TARGET-AML public database does not provide the exact date of EOI2 assessment. Landmark analyses were restricted to the primary analysis cohort (n = 418); the number at risk decreased from 415 at 3 months to 352 at 9 months, and the EOI2 < 0.05% subgroup decreased from 360 to 315.

In fixed-time landmark analyses at 3, 6, and 9 months from diagnosis, the association between Pos/Neg and higher CIR remained consistent (all p < .05), with SHRs of 1.62,1.86, and 2.17, respectively

(Supplementary Table S5). In the EOI2 < 0.05% subgroup, the corresponding SHRs were1.66, 1.89, and 2.15. These results reduce concern that the findings were driven by differential survival to the second MRD assessment.

# Discussion

In this retrospective cohort study of children with AML, we found that early MRD clearance patterns were associated with subsequent clinical outcomes. Patients who were MRD-positive at EOI1 but became MRD-negative at EOI2 (Pos/Neg) had a higher cumulative incidence of relapse than patients who were MRD-negative at both assessments (Neg/Neg). Importantly, this association remained when the analysis was restricted to patients who were MRD-negative at EOI2. The Pos/Neg pattern was also associated with inferior overall survival and event-free survival. These findings suggest that the MRD result at an earlier treatment time point may retain prognostic information even after MRD becomes undetectable at a later assessment. Consistent with the recent ELN-DAVID framework emphasizing contextual and longitudinal interpretation of MRD response(16), our findings suggest that the early trajectory from EOI1 to EOI2 may provide prognostic information beyond MRD status at EOI2, although our study-specific trajectory groups should not be considered direct equivalents of the ELN-DAVID qualitative response categories. Notably, the 5-year CIR was numerically higher in the Pos/Neg group (53.7%) than in the Pos/Pos group (49.8%), despite the latter representing persistent MRD positivity. This counterintuitive finding is explained by the higher NRM in the Pos/Pos group (12.8% vs 9.1%): NRM, as a competing event, compressed the observed CIR.

Our findings are consistent with previous studies showing that MRD measured during induction is prognostic in pediatric AML(2). In the COG AAML0531 study, early flow-cytometric MRD was associated with overall survival, relapse, and disease-free survival(5). The AML02 study further demonstrated the feasibility of incorporating sequential MRD measurements into risk-directed therapy for childhood AML(6).A prospective multicenter study by Steinbach et al. further validated that standardized MRD testing has a strong prognostic impact independent of established risk factors in childhood AML(17).Our study differs from these previous analyses by focusing on the sequence of MRD results rather than the prognostic value of MRD at a single treatment landmark. The finding that Pos/Neg patients had worse outcomes than Neg/Neg patients, even when both groups were MRD-negative at EOI2, suggests that the timing of achieving MRD negativity may provide information that is not captured by the later MRD result alone.

This observation is also consistent with the recent study by Jen et al. in adults with AML, in which patients who achieved MRD negativity only at a later treatment assessment had shorter relapse-free survival than those who were MRD-negative at both assessments(11). Our results extend this question to pediatric AML and suggest that early MRD clearance patterns may also be relevant in children. However, the present study does not establish that the same biological mechanisms underlie delayed MRD clearance in children and adults, and the prognostic findings should not be considered a direct validation of the adult data.

Our study used the 0.05% threshold previously identified as optimal in a TARGET-based analysis (10), but addressed a different question: whether the sequence of MRD results between EOI1 and EOI2 provides prognostic information beyond the later MRD status. Together with prior work on MRD clearance kinetics (11), these findings suggest two complementary aspects of MFC-MRD assessment in pediatric AML: the threshold used to define positivity and the trajectory of MRD response over time.

We also examined MRD clearance as a continuous variable. Although ΔMRD was statistically associated with cumulative incidence of relapse after adjustment for EOI2 MRD status, adding ΔMRD did not improve time-dependent AUC or categorical net reclassification. Thus, the statistical association observed for ΔMRD did not translate into clear improvement in prognostic discrimination in this cohort. These results suggest that a simple categorical description of early MRD trajectories may be more clinically interpretable than a continuous change measure. Because ΔMRD is mathematically derived from MRD measurements at two time points, its interpretation should also be made cautiously, particularly with respect to the influence of the initial MRD level. In a sensitivity model further adjusted for EOI1 MRD burden, the association of ΔMRD was attenuated, suggesting that continuous ΔMRD may partly reflect the initial MRD level rather than an independent clearance kinetic marker. Sensitivity analyses excluding early relapses and using alternative thresholds produced consistent results, although replacing ΔMRD with an absolute difference measure did not replicate the association, consistent with the known limitations of applying absolute differences to log-normally distributed data.

The association between the Pos/Neg pattern and relapse appeared particularly evident in patients with ELN 2022 intermediate-risk disease(12). This finding is of potential clinical interest because MRD trajectory may provide additional information in patients whose baseline disease characteristics do not place them clearly into the highest or lowest risk categories. However, formal interaction tests were not significant, and this subgroup analysis was not the primary endpoint of the study and should be interpreted cautiously because of the smaller sample size. Further studies are needed to determine whether MRD clearance patterns can improve risk stratification within specific genetic risk groups.

We also explored the applicability of the 2025 ELN MRD framework in this pediatric cohort(16). The updated ELN-DAVID recommendations provide a more detailed framework for interpreting MRD response in AML, but their applicability to pediatric AML requires evaluation in pediatric populations(16). In our cohort, only 7 patients fulfilled the prespecified MRD relapse criterion based on a 1-log increase, which was insufficient for a reliable formal analysis. Therefore, our results should be regarded as exploratory rather than as a validation of the updated ELN MRD criteria in children.

A key concern in studies using EOI2-defined exposures is immortal-time bias, because patients must survive to EOI2 to be classified. In our fixed-time landmark analyses at 3, 6, and 9 months, the association between Pos/Neg and higher CIR remained consistent, with SHRs increasing over time. This suggests that the association is unlikely to be explained by differential survival to the second MRD assessment. However, because the TARGET-AML public database does not provide the exact date of

EOI2 assessment, a formal EOI2 landmark analysis could not be performed, and residual selection bias cannot be fully excluded.

This study has several limitations. First, its retrospective design limits causal interpretation, and the analysis included only patients with paired EOI1 and EOI2 MRD measurements. Second, the Neg/Pos group was small (n = 7) and was therefore combined with the Pos/Pos group in the primary analysis; although sensitivity analysis excluding these patients produced consistent results, the small sample size prevented a separate assessment of this trajectory. Third, the patients were treated in historical COG trials, and treatment strategies and supportive care have evolved over time. Fourth, the small number of patients meeting the ELN 2025 MRD relapse criterion prevented formal evaluation of that framework. Fifth, the TARGET-AML public database does not provide the exact date of EOI2 assessment, so a formal EOI2 landmark analysis could not be performed; fixed-time landmark analyses yielded consistent results, but residual selection bias cannot be fully excluded. Finally, although ΔMRD was statistically associated with relapse, it did not improve discrimination or reclassification, and its association was attenuated after adjustment for EOI1 MRD burden, limiting its current clinical utility.

In conclusion, children with AML who converted from MRD-positive at EOI1 to MRD-negative at EOI2 had a higher risk of relapse than those who were MRD-negative at both assessments. This association persisted among patients who were MRD-negative at EOI2, suggesting that early MRD trajectory may provide prognostic information beyond later MRD status alone. Further prospective studies should determine whether longitudinal MRD assessment can improve risk stratification and inform future risk-adapted treatment strategies in pediatric AML.

Data Sharing Statement

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/Harmonized

# Abbreviations

AIC Akaike information criterion

AML Acute myeloid leukemia

AUC Area under the receiver operating characteristic curve

BCR-ABL1 Breakpoint cluster region-Abelson murine leukemia viral oncogene homolog 1

CI Confidence interval

CIR Cumulative incidence of relapse

COG Children's Oncology Group

| | |
|---|---|
| CR | Complete remission |
| EFS | Event-free survival |
| ELN | European LeukemiaNet |
| EOI1 | End of induction 1 |
| EOI2 | End of induction 2 |
| FAB | French-American-British |
| FLT3-ITD | FMS-like tyrosine kinase 3 internal tandem duplication |
| HR | Hazard ratio |
| IQR | Interquartile range |
| KMT2A | Lysine methyltransferase 2A (formerly MLL) |
| LR | Likelihood ratio |
| MFC | Multiparameter flow cytometry |
| MICE | Multiple imputation by chained equations |
| MRD | Measurable residual disease |
| NPM1 | Nucleophosmin 1 |
| NRM | Non-relapse mortality |
| NRI | Net reclassification improvement |
| OS | Overall survival |
| SCT | Stem cell transplantation |
| SHR | Subdistribution hazard ratio |
| TARGET | Therapeutically Applicable Research to Generate Effective Treatments |
| WBC | White blood cell count |

# Declarations

Data Sharing Statement

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/Harmonized

Acknowledgement

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

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.

Funding:

No funding.

Author information

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

Competing interest

The authors declare that there is no conflict of interest.

Ethics declarations

Ethic approval and consent for publication

This study is a secondary analysis of de-identified, publicly available data from the TARGET-AML database. Because the data are publicly available and de-identified, this analysis did not require approval from an institutional review board or ethics committee. Access to the TARGET-AML data was approved by the TARGET Publications Committee.

Clinical trial information

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

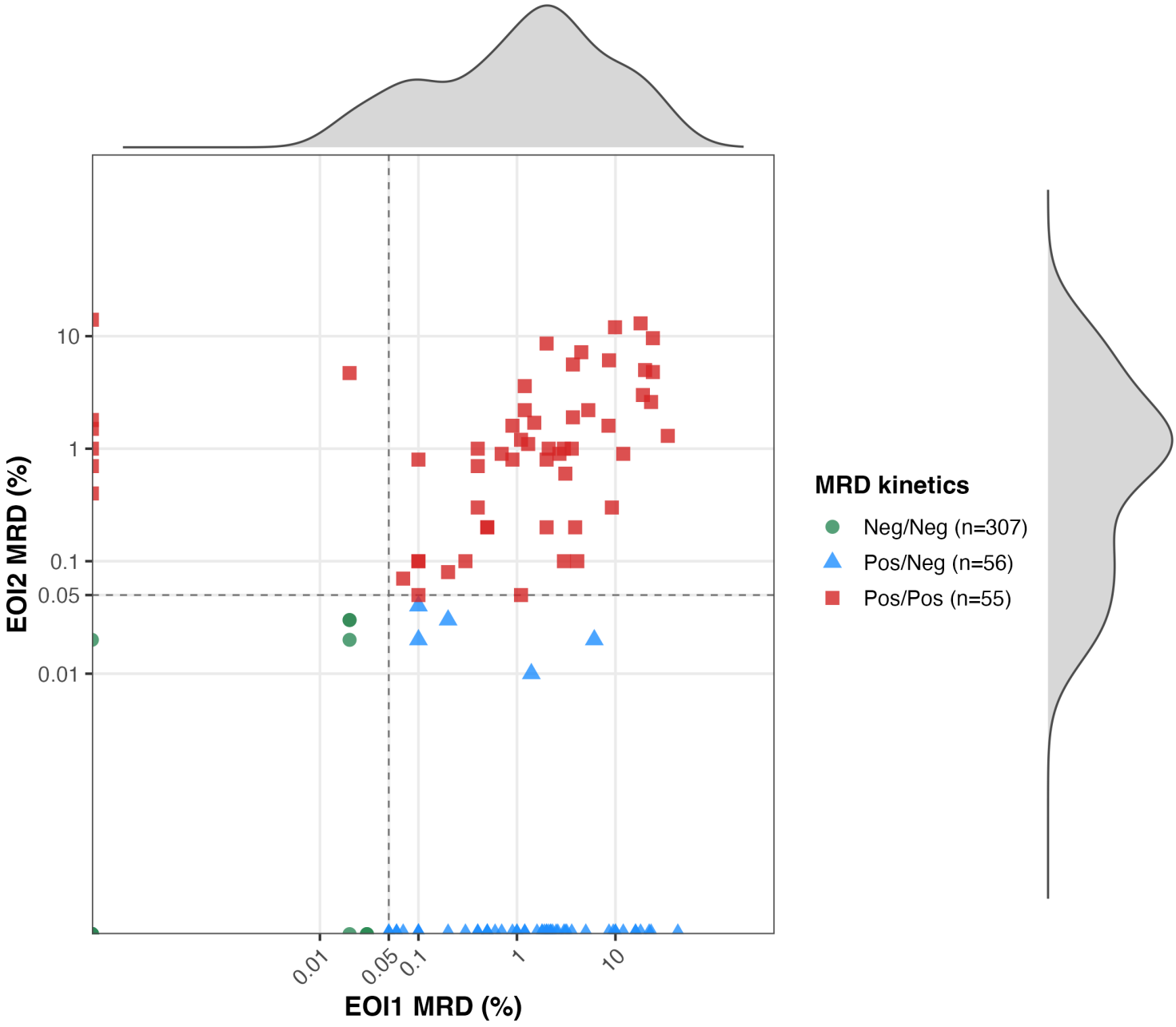
# References

1. Zwaan CM, Tasian SK, Aplenc R, Brodersen LE, Buldini B, De Moerloose B et al (2026) Diagnosis and management of AML in pediatric patients: consensus recommendations from an international expert panel. Blood 147(14):1532–1561

2. Segerink WH, de Haas V, Kaspers GJL (2021) Measurable residual disease in pediatric acute myeloid leukemia: a systematic review. Expert Rev Anticancer Ther 21(4):451–459
3. Langebrake C, Creutzig U, Dworzak M, Hrusak O, Mejstrikova E, Griesinger F et al (2006) Residual disease monitoring in childhood acute myeloid leukemia by multiparameter flow cytometry: the MRD-AML-BFM Study Group. J Clin Oncol 24(22):3686–3692
4. Eidenschink Brodersen L, Kovach AE, Zuna J (2026) Measurable Residual Disease in Pediatric Acute Leukemia: Application and Interpretation of Current and Emerging Methodologies. Am Soc Clin Oncol Educ Book 46(3):e516962
5. Brodersen LE, Gerbing RB, Pardo ML, Alonzo TA, Paine D, Fritschle W et al (2020) Morphologic remission status is limited compared to ΔN flow cytometry: a Children's Oncology Group AAML0531 report. Blood Adv 4(20):5050–5061
6. Rubnitz JE, Inaba H, Dahl G, Ribeiro RC, Bowman WP, Taub J et al (2010) Minimal residual disease-directed therapy for childhood acute myeloid leukaemia: results of the AML02 multicentre trial. Lancet Oncol 11(6):543–552
7. Buldini B, Rizzati F, Masetti R, Fagioli F, Menna G, Micalizzi C et al (2017) Prognostic significance of flow-cytometry evaluation of minimal residual disease in children with acute myeloid leukaemia treated according to the AIEOP-AML 2002/01 study protocol. Br J Haematol 177(1):116–126
8. Tierens A, Bjørklund E, Siitonen S, Marquart HV, Wulff-Juergensen G, Pelliniemi TT et al (2016) Residual disease detected by flow cytometry is an independent predictor of survival in childhood acute myeloid leukaemia; results of the NOPHO-AML 2004 study. Br J Haematol 174(4):600–609
9. van der Velden VH, van der Sluijs-Geling A, Gibson BE, te Marvelde JG, Hoogeveen PG, Hop WC et al (2010) Clinical significance of flowcytometric minimal residual disease detection in pediatric acute myeloid leukemia patients treated according to the DCOG ANLL97/MRC AML12 protocol. Leukemia 24(9):1599–1606
10. Liao XY, Zheng H, Fang JP, Zhou DH, Qiu KY (2026) Optimal minimal residual disease threshold in pediatric acute myeloid leukemia: A retrospective cohort study based on the TARGET database. PLoS Med 23(5):e1005088
11. Jen WY, Sasaki K, Ravandi F, Kadia TM, Wang SA, Wang W et al (2025) Impact of measurable residual disease clearance kinetics in patients with AML undergoing intensive chemotherapy. Blood Adv 9(4):783–792
12. Döhner H, Wei AH, Appelbaum FR, Craddock C, DiNardo CD, Dombret H et al (2022) Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. Blood 140(12):1345–1377
13. Rausch C, Rothenberg-Thurley M, Dufour A, Schneider S, Gittinger H, Sauerland C et al (2023) Validation and refinement of the 2022 European LeukemiaNet genetic risk stratification of acute myeloid leukemia. Leukemia 37(6):1234–1244
14. Nolan EK, Chen HY (2020) A comparison of the Cox model to the Fine-Gray model for survival analyses of re-fracture rates. Arch Osteoporos 15(1):86

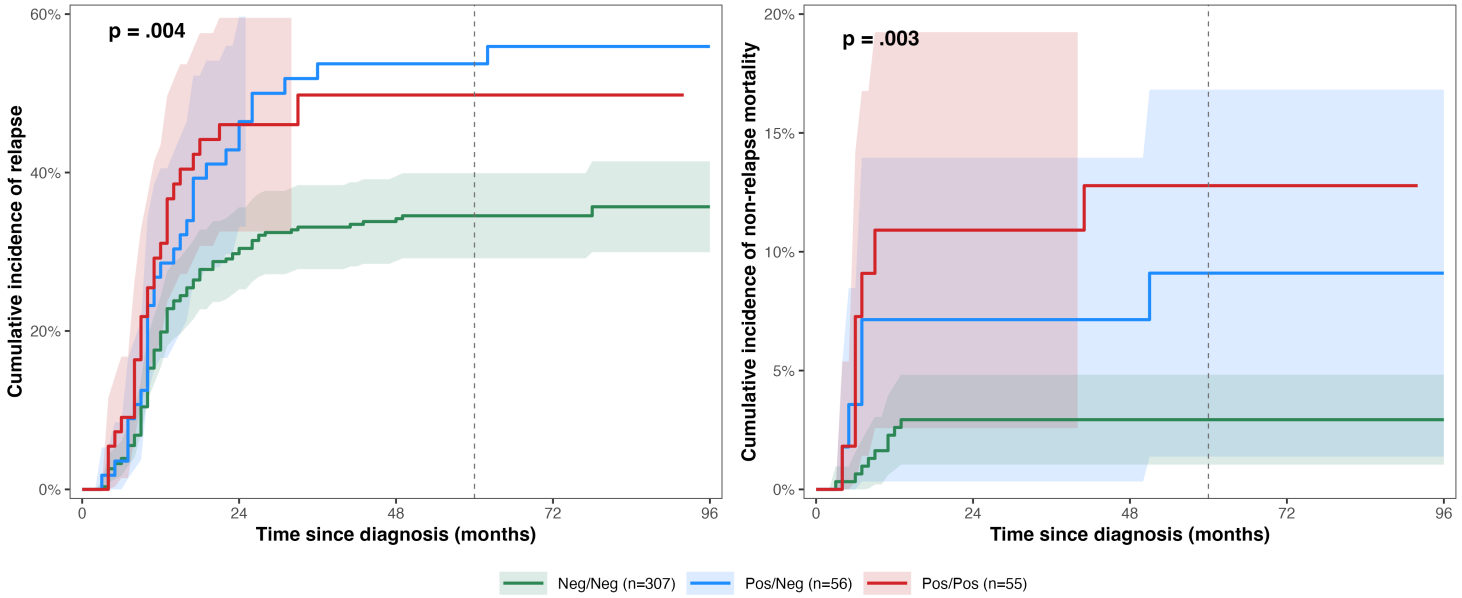
15. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB Jr., Vasan RS (2008) Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 27(2):157–172 discussion 207–12
16. Cloos J, Valk PJM, Thiede C, Döhner K, Roboz GJ, Wood BL et al (2026) 2025 update on MRD in acute myeloid leukemia: a consensus document from the ELN-DAVID MRD Working Party. Blood 147(11):1147–1167
17. Steinbach D, Bader P, Willasch A, Bartholomae S, Debatin KM, Zimmermann M et al (2015) Prospective validation of a new method of monitoring minimal residual disease in childhood acute myelogenous leukemia. Clin Cancer Res 21(6):1353–1359

## Figures



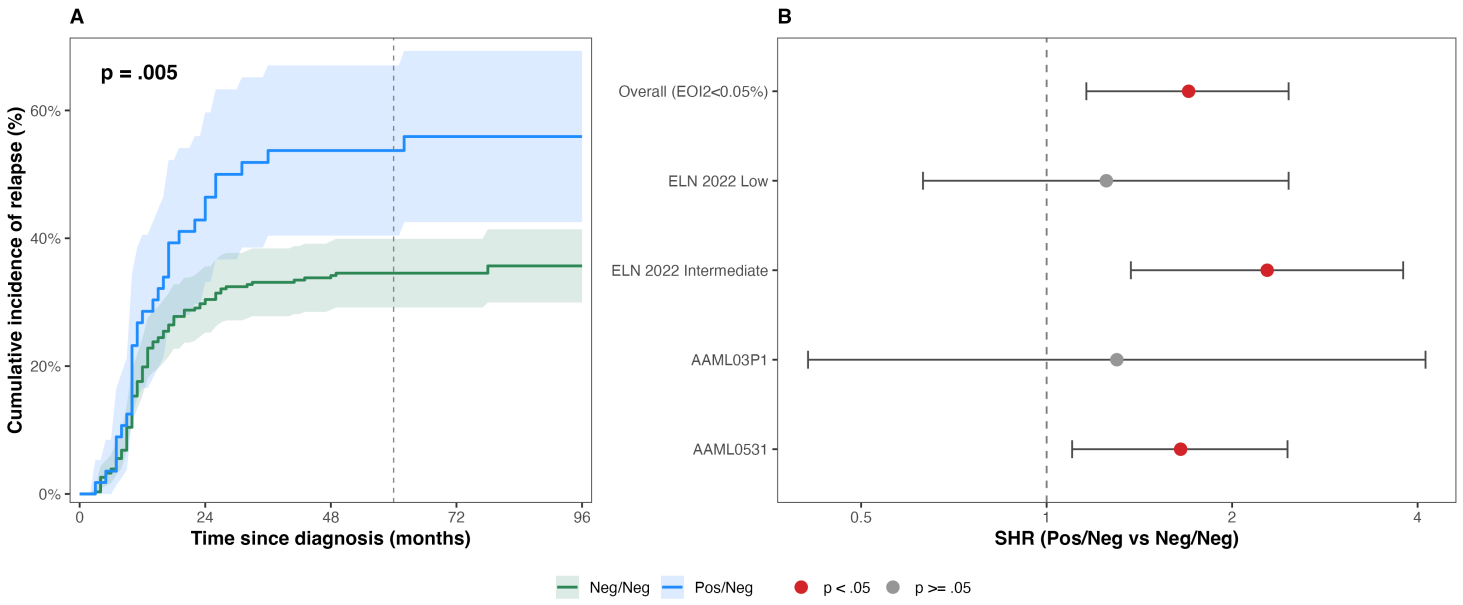
## Figure 1

**Distribution of MFC-MRD at end of induction 1 (EOI1) and end of induction 2 (EOI2) in pediatric AML (n = 418).** Dashed lines indicate the 0.05% MRD threshold. Points are colored and shaped by MRD kinetics group: Neg/Neg (green circles, n = 307), Pos/Neg (blue triangles, n = 56), and Pos/Pos (red squares, n = 55). Marginal density plots summarize the distribution of EOI1 (top) and EOI2 (right) MRD values. MRD, measurable residual disease; EOI1, end of induction 1; EOI2, end of induction 2; MFC, multiparameter flow cytometry.




## Figure 2

**Cumulative incidence of relapse (CIR, left) and non-relapse mortality (NRM, right) by MRD kinetics group in pediatric AML.** Curves were estimated using the Aalen-Johansen estimator in a competing risk framework. Patients were classified according to the 0.05% MRD threshold as Neg/Neg (green, n = 307), Pos/Neg (blue, n = 56), or Pos/Pos (red, n = 55). Shaded areas represent 95% confidence intervals. Dashed vertical line indicates 5 years after diagnosis. P-values were calculated using Gray's test (CIR: p = .004; NRM: p = .003). Note that Y-axis scales differ between panels; CIR and NRM should be interpreted separately. MRD, measurable residual disease; CIR, cumulative incidence of relapse; NRM, non-relapse mortality.




## Figure 3

**(A) Cumulative incidence of relapse (CIR) in patients achieving EOI2 MRD <0.05%, stratified by EOI1 MRD status (Neg/Neg vs Pos/Neg).** Curves were estimated using the Aalen-Johansen estimator with non-relapse mortality as a competing event. Shaded areas represent 95% confidence intervals. Dashed vertical line indicates 5 years after diagnosis. P-value was calculated using Gray's test (p = .005). (B) Forest plot of subdistribution hazard ratios (SHR) for Pos/Neg versus Neg/Neg across subgroups. Point estimates and 95% confidence intervals are shown, with numeric values provided on the right. Red points indicate p < .05; grey points indicate p ≥ .05. Dashed line indicates SHR = 1. MRD, measurable residual disease; CIR, cumulative incidence of relapse; SHR, subdistribution hazard ratio; CI, confidence interval; ELN, European LeukemiaNet; EOI, end of induction.

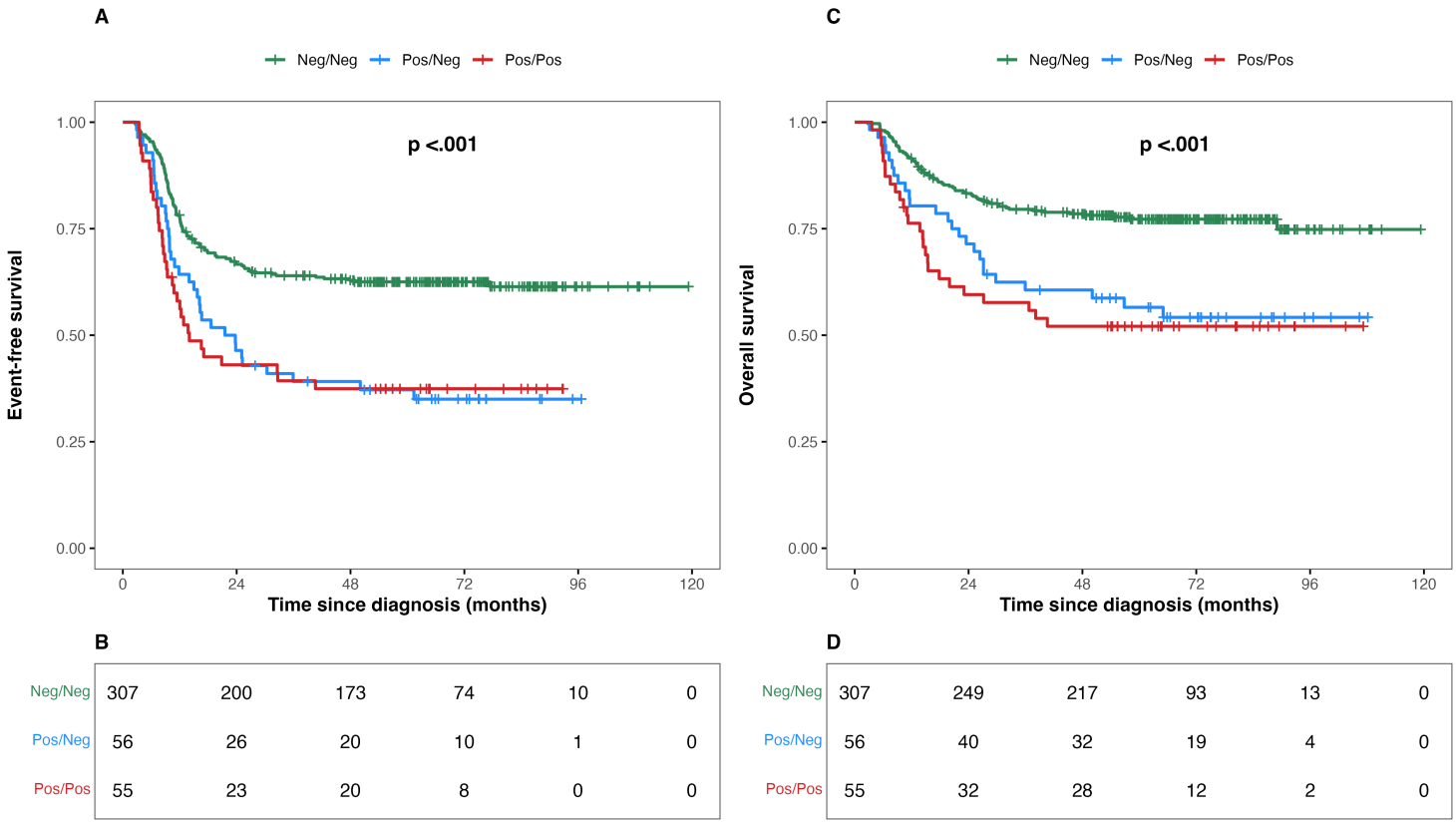



**Figure 4**

**Kaplan-Meier estimates of event-free survival (EFS, A) and overall survival (OS, C) by MRD kinetics group in pediatric AML.** Patients were classified according to the 0.05% MRD threshold as Neg/Neg (green, n = 307), Pos/Neg (blue, n = 56), or Pos/Pos (red, n = 55). P-values were calculated using the log-rank test (both p < .001). Number at risk is shown below each survival panel (B for EFS; D for OS). MRD, measurable residual disease; EFS, event-free survival; OS, overall survival.

# Supplementary Files

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

- Table5.docx
- Table6.docx
- Table2.docx
- Table3.docx
- Table4.docx
- Table1.docx
- Table7.docx
- SupplementaryFigureS1.tiff

- SupplementaryTableS3.docx
- SupplementaryTableS2.docx
- SupplementaryTableS1.docx
- SupplementaryTableS5.docx
- SupplementaryTableS4.docx