

Published in final edited form as:

Lancet Oncol. 2010 June ; 11(6): 543–552. doi:10.1016/S1470-2045(10)70090-5.

Minimal Residual Disease-Directed Therapy for Childhood Acute Myeloid Leukemia: Results of the AML02 Multicenter Trial

Jeffrey E. Rubnitz, MD, PhD¹, Hiroto Inaba, MD, PhD¹, Gary Dahl, MD², Raul C. Ribeiro, MD¹, W. Paul Bowman, MD³, Jeffrey Taub, MD⁴, Stanley Pounds, PhD¹, Bassem I Razzouk, MD¹, Norman J. Lacayo, MD², Xueyuan Cao, PhD¹, Soheil Meshinchi, MD, PhD⁵, Barbara Degar, MD⁶, Gladstone Airewele, MD⁷, Susana C. Raimondi, PhD¹, Mihaela Onciu, MD¹, Elaine Coustan-Smith, MS¹, James R. Downing, MD¹, Wing Leung, MD, PhD¹, Ching-Hon Pui, MD¹, and Dario Campana, MD, PhD¹

¹St. Jude Children's Research Hospital and the University of Tennessee Health Science Center, College of Medicine, Memphis, TN

²Lucile Packard Children's Hospital and Stanford Cancer Center, Palo Alto, CA

³Cook Children's Medical Center, Fort Worth, TX

⁴Children's Hospital of Michigan, Detroit, MI

⁵Fred Hutchinson Cancer Research Center, Seattle, WA

⁶Dana-Farber Cancer Institute, Boston, MA

⁷Texas Children's Cancer Center, Houston, TX

Summary

Background—We sought to improve outcome of childhood acute myeloid leukemia (AML) by applying risk-directed therapy based on the genetic abnormalities of the leukemic cells and measurements of minimal residual disease (MRD) as determined by flow cytometry during treatment.

Methods—From October 13, 2002 to June 19, 2008, 232 patients with de novo AML (n=206), therapy- or myelodysplasia-related AML (n=12), or mixed-lineage leukemia (n=14) were enrolled at eight centers. Block, nonblinded randomization, stratified by cytogenetic or morphologic subtype, assigned patients to high-dose (18 g/m², n=113) or low-dose (2 g/m², n=117) cytarabine (A), given together with daunorubicin (D) and etoposide (E) (Induction I); achievement of MRD negative status was the primary endpoint. Induction II consisted of ADE with or without

© 2010 Elsevier Ltd. All rights reserved

Correspondence: Jeffrey E. Rubnitz, MD, PhD Department of Oncology St. Jude Children's Research Hospital 262 Danny Thomas Place Memphis, TN 38105-2794 Phone: 901-595-2388, FAX: 901-521-9005 jeffrey.rubnitz@stjude.org.

Current affiliation for BIR: Children's Center for Cancer and Blood Diseases, St. Vincent Children's Hospital, Indianapolis, IN

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Contributors JER was the principal investigator of the clinical trial, analyzed data, and wrote the manuscript; HI, GD, RCR, WPB, JT, BIR, NJL, SM, BD, GA, and WL participated in the design of the trial, took care of patients on the trial, and reviewed the manuscript; SP and XC performed the statistical design and analysis and reviewed the manuscript; SCR reviewed all cytogenetic data and reviewed the manuscript; MO performed central pathology review of all cases and reviewed the manuscript; EC-S and DC participated in the design of the trial, performed all MRD analyses, analyzed data, and reviewed the manuscript; JRD and C-HP participated in the design of the trial, analyzed data, and reviewed the manuscript.

Conflicts of Interest The authors declare no conflicts of interest.

gemtuzumab ozogamicin (GO); consolidation therapy included three additional courses of chemotherapy or hematopoietic stem cell transplantation (HSCT). Levels of MRD were used to allocate GO and determine the timing of Induction II; both MRD and genetic abnormalities at diagnosis were used to determine final risk classification. Low-risk patients (n=68) received 5 courses of chemotherapy, whereas high-risk patients (n=79), as well as standard-risk patients (n=69) with matched sibling donors, were eligible for HSCT (performed in 48 high and 8 standard-risk patients). All randomized patients (n=230) were analyzed for the primary endpoint. The other analyses were limited to the 216 patients with AML, excluding mixed-lineage leukemia. This trial, closed to accrual, is registered with ClinicalTrials.gov, number NCT00136084.

Findings—The complete remission rates were 80% (173 of the 216) after Induction I and 94% (203 of 216) after Induction II. Induction failures included two toxic deaths and 10 cases of resistant leukemia. The introduction of high-dose cytarabine did not significantly lower the rate of MRD positivity after Induction I therapy (34% vs. 42%, $P=0.17$). The cumulative incidences of grade 3 or greater infection were $79.3\% \pm 4.0\%$ and $75.5\% \pm 4.2\%$ for patients treated on the high-dose or low-dose arms. The 3-year estimates (\pm SE) of event-free and overall survival were $63.0\% \pm 4.1\%$ and $71.1\% \pm 3.8\%$, respectively. Achievement of MRD $< 0.1\%$ after Induction II identified a large group of patients (80%) with a cumulative incidence of relapse of only $17\% \pm 3\%$. Post-Induction I MRD $\geq 1\%$ was the only independent adverse prognostic factor that was statistically significant ($P < 0.05$) for both event-free (HR, 2.41; CI 1.36–4.26; $P=0.003$) and overall survival (HR, 2.11; CI 1.09–4.11; $P=0.028$).

Interpretation—Our findings suggest that the use of targeted chemotherapy and HSCT, in the context of a comprehensive risk-stratification strategy based on genetic features and MRD findings, can improve the outcome of childhood AML.

Introduction

With improvements in risk-directed therapy and supportive care, event-free survival (EFS) rates for children with acute lymphoblastic leukemia now approach 90%.¹ By contrast, EFS rates for children with acute myeloid leukemia (AML) range from 49% to 62% in recent trials.^{2–4} Improving clinical results in AML will require not only the development of new drugs and better supportive care, but also a more precise application of risk-directed therapy.

It is well known that genetic abnormalities of leukemic blasts are associated with clinical outcome in patients with AML.⁵ Methods for detecting minimal residual disease (MRD) allow much more precise determinations of early reduction in leukemic burden than were possible in the past; the results of these tests are powerful and independent predictors of relapse in adults and children with AML.^{6–10} We therefore designed a multicenter study, AML02 (ClinicalTrials.gov identifier: NCT00136084), that relied on presenting genetic features and sequential evaluation of MRD to determine the final risk assignment and treatment of children with AML. In addition, because earlier studies suggested that higher doses of cytarabine during induction therapy may be associated with lower relapse rates,^{11–13} we tested in a randomized fashion whether high-dose cytarabine (18 g/m^2) would produce better results than lower doses of this agent (2 g/m^2). Gemtuzumab ozogamicin was administered to patients with poor early response; those with high-risk features were eligible for hematopoietic stem cell transplantation (HSCT).

Patients and Methods

Patients

From October 13, 2002 to June 19, 2008, 232 children with de novo AML (n=206), therapy- or MDS-related AML (n=12), or mixed-lineage leukemia (n=14) were enrolled in the AML02 trial at eight centers. Mixed-lineage leukemia was defined as described in the WHO

2008 classification.^{14,15} Their ages at diagnosis ranged from 2 days to 21.4 years (median, 9.1 years). Patients with acute promyelocytic leukemia or Down syndrome were excluded. The protocol was approved by the institutional review boards and written informed consent was obtained from all patients or their guardians or parents.

Cytogenetic analysis was performed as previously described.¹⁶ The *AML1-ETO*, *MLL-AF9*, and *CBFβ-MHY11* fusion transcripts were detected by reverse transcriptase polymerase chain reaction (RT-PCR),¹⁷ *MLL* gene rearrangements by fluorescent in situ hybridization, and *FLT3* internal tandem duplications (ITD) and *FLT3* point mutations by PCR.

Risk Classification and Measurements of Treatment Response

At diagnosis, patients were provisionally classified as having low-risk AML if their leukemic cells had *t(8;21)/AML1-ETO*, *inv(16)/CBFβ-MHY11*, or *t(9;11)/MLL-AF9*. High-risk cases included those with *-7*, *FLT3-ITD*, *t(6;9)*, megakaryoblastic leukemia, treatment-related AML, or AML arising from myelodysplastic syndrome. All other patients were provisionally classified as having standard-risk AML.

Responses to each course of therapy, as assessed by morphologic and flow cytometric studies of MRD, determined the final risk classification. Complete remission (CR) was defined as trilineage hematopoietic recovery with less than 5% blasts in the marrow. For MRD studies, leukemia-associated immunophenotypes were identified in diagnostic bone marrow specimens and marker combinations that allowed detection of 1 leukemia cell per 1000 mononuclear bone marrow cells were applied to subsequent samples (Supplementary Methods; Table S1).⁷ Results are reported as percentage of nucleated cells with the leukemia-associated immunophenotype. MRD positivity was defined as 1 or more leukemic cell per 1000 mononuclear bone marrow cells (i.e., $\geq 0.1\%$).

Therapy

Patients were randomized to receive daunorubicin (50 mg/m² on days 2, 4, and 6) and etoposide (100 mg/m² on days 2–6) plus either high-dose cytarabine (3 g/m² every 12 hours on day 1, 3, and 5) or low-dose cytarabine (100 mg/m² every 12 hours on days 1–10) during Induction I (Fig 1). Bone marrow was examined on day 22 to evaluate treatment response. Patients with $\geq 1\%$ leukemic blasts by flow cytometry proceeded to Induction II immediately; whereas those with less than 1% blasts began Induction II at the time of count recovery (absolute neutrophil count $\geq 300/\mu\text{l}$ and platelet count $\geq 30,000/\mu\text{l}$). If greater than 7 days passed between day 22 and count recovery, a second bone marrow examination was performed to determine the final response to Induction I. In the nine cases with discrepancies between day 22 and final Induction I MRD values, the final Induction I MRD value was used for clinical and statistical purposes. During Induction II, patients received low-dose cytarabine, daunorubicin, and etoposide (ADE) alone or in combination with 3 mg/m² gemtuzumab ozogamicin (GO). Initially, we limited the eligibility to receive ADE plus GO to patients with $>25\%$ bone marrow blasts by flow cytometry after Induction I because there had been no reports on the use of this combination in children or adults with AML; patients with MRD $\geq 0.1\%$ after Induction II were given GO (6 mg/m²) as Induction III. Because ADE plus GO was well tolerated, we amended the protocol in February 2005 extending the eligibility for t ADE plus GO to patients with MRD levels $\geq 1\%$ after Induction I; GO was no longer used as Induction III.

Consolidation therapy was based on the initial risk classification and response to therapy. Low-risk patients received three courses of cytarabine-based chemotherapy (Fig. 1). Patients with poor responses to therapy ($>25\%$ blasts after Induction I or persistent MRD after 3 courses of therapy) were considered high risk and were eligible for HSCT. All other patients

were classified as standard risk and were eligible to receive HSCT only if they had matched sibling donors.

At the beginning of the trial, patients without central nervous system (CNS) disease received five doses of age-adjusted intrathecal cytarabine, given at the beginning of each course of therapy. Patients with CNS leukemia at diagnosis received weekly intrathecal therapy until the cerebrospinal fluid was clear of leukemia cells (minimum four doses) and then four additional doses. In July 2003, the protocol was amended to replace intrathecal cytarabine with triple intrathecal therapy consisting of methotrexate, hydrocortisone, and cytarabine.

Supportive care guidelines initially included the use of prophylactic voriconazole for all patients. As recently described,¹⁸ guidelines were later modified to include the use of prophylactic vancomycin and ciprofloxacin after each course of chemotherapy (beginning when the absolute neutrophil count was < 1000/ μ l and continuing until it was > 100/ml and rising).

Statistical Design and Analysis Methods

The primary study aim was to compare the rates of MRD negativity on day 22 of Induction I between patients randomized to receive high-dose versus low-dose cytarabine during remission induction, based on the O'Brien-Fleming group sequential method for comparing two binomial distributions.¹⁹ East software (Cytel Corp., Cambridge, MA) was used to develop the design and compute its statistical properties. The design specified enrollment of 186 patients who were evaluable for MRD, and included four interim analyses and one final analysis. Each interim analysis and the final analysis were performed using a Monte-Carlo approximation of the Cochran-Mantel-Haenzel test (stratified by initial risk assignment) based on 10,000 permutations. The study provided an overall power of 80% at an overall 5% level for a two-sided test to detect a difference of 20% between the two arms, assuming one arm had an MRD-negative rate of 50%.

Zelen's block-randomization method with block-size of 6 was used to assign subjects in a stratified manner to high-dose or low-dose cytarabine.²⁰ For patients with detectable cytogenetic abnormalities, there were five strata: inv(16), t(8;21), t(9;11), MDS-related AML or secondary AML or monosomy 7 or M7, and others. For patients without such data, the five strata were M1/M2 with Auer rods, M4Eo, M5, MDS-related AML or secondary AML or M7, and others. The St. Jude Pharmacy used a program developed and maintained by the St. Jude Biostatistics Department to perform the randomization. The program kept the treatment assignments concealed until the assignment was needed for a newly enrolled patient. The results of the treatment assignment were provided to the treating physician, the participants, and the data analysts; no masking was performed. Additionally, the randomization was not stratified by center.

The exact chi-square test was used to compare categorical features and outcomes, while the exact Kruskal-Wallis test was used to compare quantitative features and outcomes. EFS was defined as the time from study enrollment to induction failure, relapse, death, or study withdrawal for any reason, with event-free patients censored on the date of the last follow-up. Overall survival (OS) was defined as the time from study enrollment to death, with living patients censored on the date of the last follow-up. The Kaplan-Meier method²¹ was used to estimate EFS and OS rates, and the exact log-rank test was used to compare survival curves. Survival estimates are reported with standard errors determined by the method of Peto and Pike.²² Survival plots include the number of patients at risk.²³

Cox proportional hazards regression models²⁴ were used to explore the associations of Induction I MRD findings, treatment arm assignment, and presenting features with outcome.

Potential variables included in the AIC-based model selection were randomization arm, age, WBC count, Induction I MRD, *FLT3*-ITD, t(8;21)/inv(16), t(9;11), 11q23 translocations other than t(9;11), M7 with t(1;22), and M7 without t(1;22). The model was selected using overall survival data; a model with the selected variables was fit to EFS data. The exploratory multivariable regression model was fit using data from subjects with data available for all variables. Gray's methods²⁵ were used to estimate and compare the cumulative incidence of important events (relapse, death in remission, induction failure) while adjusting for competing risks. Cumulative incidence estimates are reported with standard errors. Monte Carlo approximations based on 10,000 permutations were used to compute the P-value for the chi-square and log-rank tests. All analyses were performed with SAS software (SAS Institute, Cary, NC), Windows version 9.1.

Role of the funding source

The funding source of this study had no role in study design; collection, analysis, or interpretation of the data; or in the writing of the report. JER, SP, and XC had access to the complete trial database. JER had the final responsibility to submit for publication.

Results

Early Response According to Cytarabine Randomization

Among the 232 eligible patients, 230 (216 with AML and 14 with mixed lineage leukemia) were randomized to receive high-dose cytarabine (n=113) or low-dose cytarabine (n=117) (Fig. 2). Two patients were not randomized because of physician choice or parent refusal. Presenting features of the randomized patients are shown in Table 1.

On day 22 of remission induction therapy, the rate of MRD-positivity was lower (though not significantly) in patients treated in the high-dose cytarabine arm compared with the low-dose arm (34% vs. 42%, P=0.17). The result was similar when the analysis was limited to patients with de novo AML (33% vs. 40%, P=0.22). MRD positivity rates were not significantly different between the two treatment arms when the analysis was performed within risk categories defined by presenting features: low risk (12% in the high-dose arm vs. 14% in the low-dose arm, P=1.0), standard risk (33% vs. 40%, P=0.62), and high risk (53% vs. 68%, P=0.31).

Overall, there were no differences in the incidence of any grade 3 or greater toxicity during Induction I between patients treated on the high-dose (98 of 113, 88%) or low-dose arm (101 of 117, 86%), nor was there a difference in the cumulative incidence of bacterial infections between the two arms. However, patients treated on the high-dose arm had a higher cumulative incidence of grade 2 or greater fungal infection compared with those on the low-dose arm (23.6% ± 4.2% vs. 13.6% ± 3.3% at 6 months, P=0.058, Table 2).

Overall Treatment Outcome

Complete remission was achieved in 173 (80%) of the 216 patients with AML (excluding patients with mixed-lineage leukemia) after Induction I and 203 (94%) after Induction II. Two patients (0.9%) died during induction (one from sepsis and one from intracranial hemorrhage), one withdrew from the study, and 10 (4.6%) had resistant leukemia. The 10 patients with refractory leukemia included 5 with *FLT3*-ITD and one each with 5q-/7q-, t(16;21), t(6;7), t(11;19)/*MLL-ENL*, and MDS. Five of these 10 patients subsequently achieved remission and are alive after HSCT.

The 3-year EFS and OS estimates were 63.0% ± 4.1% and 71.1% ± 3.8%, with median failure times of 191 days and 292 days for EFS and OS, respectively (Fig. 3). Among

patients who achieved complete remission, the 3-year cumulative incidences of relapse and of death unrelated to relapse were $20.8\% \pm 2.9\%$ and $8.7\% \pm 2.0\%$, respectively (Fig. 3). The median time to relapse was 265 days. Bone marrow was the major site of relapse, accounting for 37 of the 43 relapses. Of the first 32 patients enrolled, whose CNS-directed therapy consisted of intrathecal cytarabine alone, three had CNS relapses. The protocol was therefore amended to substitute intrathecal cytarabine with triple intrathecal therapy, and only one of the subsequent 184 patients had a CNS relapse. This amendment decreased the 3-year cumulative incidence of CNS relapse from $9.4\% \pm 5.3\%$ to $0.6\% \pm 0.6\%$ ($P=0.0007$).

Outcome differed significantly among patients with low-risk ($n=68$), standard-risk ($n=69$) and high-risk ($n=79$) leukemia: 3-year EFS rates were $82.1\% \pm 6.1\%$ vs. $62.9\% \pm 7.0\%$ vs. $46.7\% \pm 6.8\%$ ($P<0.0001$, Fig. 4A), and 3-year OS rates were $89.1\% \pm 4.8\%$ vs. $72.5\% \pm 6.4\%$ vs. $54.8\% \pm 6.8\%$ ($P<0.0001$, Fig. 4B). Accounting for time to transplant, there was no difference in overall survival between high-risk patients undergoing HSCT ($n=48$) or receiving chemotherapy only ($n=31$): $57.5\% \pm 8.6\%$ vs. $50.5\% \pm 10.7\%$ ($P=0.85$). When the analysis was limited to high-risk patients with MRD $>1\%$ after Induction I, there was a trend towards better overall survival among those who underwent HSCT: $43.5\% \pm 12.4\%$ vs. $23.1\% \pm 10.1\%$ ($P=0.14$).

Patients randomized to receive high-dose or low-dose cytarabine during Induction I fared essentially the same in comparisons of EFS ($60.2\% \pm 5.9\%$ vs. $65.7\% \pm 5.7\%$, $P=0.41$), OS ($68.8\% \pm 5.5\%$ vs. $73.4\% \pm 5.3\%$, $P=0.41$), cumulative incidence of relapse ($17.5\% \pm 3.8\%$ vs. $21.5\% \pm 4.0\%$, $P=0.50$), and cumulative incidence of death unrelated to relapse at 3 years ($11.9 \pm 3.3\%$ vs. $5.5 \pm 2.2\%$, $P=0.13$) (Table 2). When the analyses were performed within each risk category, there were also no significant differences in EFS or OS between patients treated with high-dose and those with low-dose cytarabine (Table 2).

Clinical Significance of Flow-cytometry-based MRD Findings

MRD after Induction I was studied by flow cytometry in 202 (93.5%) of the 216 AML patients and was $\geq 0.1\%$ in 74 (36.6%). Of the 14 patients without Induction I MRD studies, one died before receiving therapy, 11 had leukemic cells at diagnosis that lacked distinctive immunophenotypes for an MRD assay with at least 0.1% sensitivity, one had no cells available at diagnosis to define a suitable immunophenotype, and one had a bone marrow sample with inadequate cellularity post-Induction I. The odds of having detectable MRD after Induction I were significantly lower ($P<0.0001$) for patients with core binding factor (CBF) leukemia [$t(8;21)/AML1-ETO$ or $inv(16)/CBF\beta-MHY11$] (6 of 56 patients) and significantly greater ($P<0.0001$) for those with *FLT3*-ITD-positive leukemia (21 of 27 patients).

The presence of MRD after Induction I was significantly associated with an adverse outcome: 3-year cumulative incidences of relapse or induction failure were $38.6\% \pm 5.8\%$ for MRD-positive patients ($n=74$) and $16.9\% \pm 3.4\%$ for MRD-negative patients ($n=128$) ($P<0.0001$, Fig. 5A). The corresponding 3-year EFS rates were $43.1\% \pm 6.9\%$ and $73.6\% \pm 5.0\%$ for the MRD-positive and MRD-negative patients ($P<0.0001$). MRD after Induction I was not a significant predictor of relapse or induction failure among patients with low-risk ($P=0.55$) or standard-risk AML ($P=0.76$). By contrast, among patients with high-risk AML, those with positive MRD after Induction I had a significantly higher 3-yr cumulative incidences of relapse or induction failure ($45.1\% \pm 7.3\%$) than those with negative MRD ($20.5\% \pm 8.5\%$, $P=0.01$).

Among the 74 patients with detectable MRD after Induction I, 50 had high levels ($\geq 1\%$) and 24 had lower levels (0.1% to $<1\%$). The cumulative incidence of relapse or induction failure was significantly higher among those with high levels of MRD ($49.2\% \pm 7.4\%$) compared to

patients with low levels of MRD ($16.7\% \pm 7.8\%$, $P < 0.0001$, Fig. 5B). The outcome for patients with low levels of MRD after Induction I was identical to that of patients with negative MRD.

MRD was studied after Induction II in 193 (89.4%) of the 216 patients, and was $\geq 0.1\%$ in 38 (19.7%). Death before therapy ($n=1$), lack of a suitable immunophenotype ($n=11$), lack of diagnostic sample ($n=1$), inadequate cellularity ($n=1$) or no follow up sample ($n=4$) accounted for 18 of the 23 patients without Induction II MRD testing. The remaining five did not receive Induction II. MRD levels after Induction II were also important predictors of relapse. The 3-year cumulative incidence of relapse or induction failure estimates were $56.3\% \pm 8.4\%$ for the 38 patients with positive MRD and $16.7\% \pm 3.1\%$ for the 155 patients who were MRD-negative ($P < 0.0001$, Fig. 5C). The 3-year EFS estimates for these groups of patients were $35.8\% \pm 8.6\%$ and $71.2\% \pm 4.7\%$, respectively ($P < 0.0001$). In contrast to findings after Induction I, even low levels of MRD after Induction II were associated with a very poor outcome: the 3-year cumulative incidence of relapse or induction failure was $64.7\% \pm 12.4\%$ for those with high levels ($\geq 1\%$, $n=17$) and $48.7\% \pm 11.6\%$ for patients with low levels (0.1% to $< 1\%$, $n=21$, Fig. 5D).

MRD measurements were further used to assess response to ADE plus GO, which was given to patients with poor responses to Induction I. Overall, 27 of the 29 evaluable patients who received ADE plus GO had reductions in MRD levels. Of the patients who received this combination because of MRD $> 25\%$, all eight had reductions in MRD, with four becoming MRD negative. Of the 21 patients who subsequently received these agents because of MRD $\geq 1\%$, 19 had MRD reductions and nine became negative.

Prognostic Factor Analysis

We analyzed associations between treatment outcome and known risk factors in childhood AML (Table 3; Table S2). In addition to the associations between MRD levels and outcome described above, we observed significant differences in 3-year EFS between patients with or without *FLT3*-ITD ($39.3\% \pm 12.5\%$ vs. $67.6\% \pm 4.4\%$, $P=0.002$) and between patients with or without genetic abnormalities involving *CBF* ($85.8\% \pm 6.1\%$ vs. $55.5\% \pm 4.9\%$, $p < 0.001$). Multivariable analysis, restricted to patients with data available for all variables, identified Induction I MRD $\geq 1\%$ and the absence of *CBF* abnormalities as independent adverse prognostic factors for EFS and Induction I MRD $\geq 1\%$, age > 10 years at diagnosis, 11q23 translocations other than t(9;11), and M7 without t(1;22) for OS (Table 4). Of all the included variables in each model, only age showed significant evidence ($p = 0.03$) of violating the proportional hazards assumption for the OS analysis. However, because our study was not powered to detect all prognostic factors, we cannot exclude that other factors might have some prognostic relevance.

Discussion

This multicenter study of risk-adapted therapy based on genetic features and sequential MRD measurements achieved a 3-year EFS rate of 63% and an OS rate of 71%. These results represent substantial gains over the outcomes of recent trials conducted in the United States, including St. Jude AML97 (44% and 50%),²⁶ the Pediatric Oncology Group Study 9421 (36% and 54%),²⁷ and the Children's Cancer Group 2961 (42% and 52%).³ They also compared favorably with results reported by the Medical Research Council (MRC, 48% and 56%),^{28,29} the Nordic Society for Paediatric Haematology and Oncology (NOPHO, 48% and 64%),³⁰ and the Berlin-Frankfurt-Muenster study group (BFM, 49% and 62%),² and are similar to those of the AML99 study recently reported by the Japanese Childhood AML Cooperative Group (62% and 76%).⁴ However, we included older patients (up to 21 years of age, compared to 18 years in the Japanese, BFM, and NOPHO studies, and 15 years in the

MRC study), a higher proportion of black patients, and a higher percentage of patients with megakaryoblastic leukemia (12% vs. 5–8% in the other trials), features that have been associated with an inferior outcome in childhood AML.^{31–33} In addition, AML02 included 10 patients with treatment-related AML, a high-risk diagnosis that was not eligible for the Japanese or European studies, and excluded patients with acute promyelocytic leukemia (included in the BFM and NOPHO studies), or patients with Down syndrome (included in the MRC trial). Finally, the proportion of patients with *CBF* abnormalities (a favorable genetic feature) was exceptionally high in the Japanese study (37%), compared to 24% in our study.

A combination of several protocol components is likely to be responsible for our improved results. The use of MRD to intensify treatment either by intensifying the timing of Induction II or by the addition of GO to Induction II, helped reduce the rate of refractory disease to 4.6%. In this context, careful application of supportive care measures undoubtedly contributed to the increased efficacy of our modified treatment, as the death rate during remission induction therapy was only 0.9%. In addition, the substitution of intrathecal cytarabine with triple intrathecal therapy decreased the 3-year cumulative incidence of CNS relapse from $9.4\% \pm 5.3\%$ to $0.6\% \pm 0.6\%$ ($P=0.0007$). However, this result should be interpreted cautiously, as only 4 CNS relapses occurred in the entire cohort. Because earlier studies suggested that increases in the intensity of induction therapy were associated with improved outcomes,^{11,12,28,34,35} we tested whether remission induction therapy containing high-dose cytarabine (18 g/m^2) would secure better results than induction therapy containing lower doses of this agent (2 g/m^2). Despite this rationale, there were no significant differences in response rates between the two study arms in terms of remission induction, EFS or OS, in contrast to results achieved in some adult studies.^{11,12} The presumed lack of benefit from the higher dose of cytarabine during Induction I might be due to the relatively high cumulative doses of the drug (34.1 to 49.6 g/m^2) during the subsequent four courses of therapy. It is also possible that any potential gains from the higher dose of cytarabine during Induction I were tempered by the shorter exposure time (5 days vs. 10 days in the lower-dose arm).

Morphologic assessment of treatment response is notoriously subjective and lacks sensitivity, providing a rationale for using flow cytometric or molecular methods to track residual leukemia. Flow cytometry is the only method that can be used in virtually all childhood AML patients (94.9% in our study; Table S1). The prognostic value of MRD as assessed by flow cytometry was demonstrated by prior studies.^{7,9,10} In studies conducted by the Children's Oncology Group⁹ and at our institution,⁷ the detection of residual disease by flow cytometry was the most powerful independent predictor of poor outcome. A study performed in an AML-BFM 98 cohort also demonstrated that MRD was a significant prognostic factor, even in analyses restricted to patients without detectable blasts by morphology after the initial course of chemotherapy.¹⁰ Moreover, when the results of flow cytometry on day 15 were included in a multivariate analysis controlling for a risk classification schema based on FAB subtype, cytogenetics and blasts on day 15 by morphology, flow cytometry appeared to be a better predictor of failure-free survival. Only when adding the additional covariate of blast by morphology on day 28, did the schema have a better predictive value than MRD alone. We interpret these findings as a further indication of the utility of MRD for risk assignment, which can be complemented by presenting features, such as genetic abnormalities.

In the present study, we were able to monitor MRD after each course of therapy in more than 90% of patients, and used the results to establish the final risk classification and to determine the timing of subsequent courses of therapy. This also allowed us to directly measure the anti-leukemic effects of ADE plus GO in patients who had responded poorly to

prior therapy. Interestingly, the outcome for patients with low levels of MRD (0.1% to <1%) after Induction I was similar to that of patients with undetectable MRD levels (<0.1%), suggesting that our treatment strategy abrogates the unfavorable prognosis typically associated with a slow clearance of leukemic cells. However, the presence of high levels of MRD ($\geq 1\%$) after Induction I remained a significant adverse predictor, suggesting that novel therapies are needed for this high-risk subset of patients. Although it is possible to detect MRD by flow cytometry at a sensitivity of 1 in 10,000 cells (0.01%) in some patients with AML,⁶ the maximum sensitivity achievable by current methods in most cases is 0.1%, a threshold that we adopted to define MRD positivity in AML02. We suggest that a more sensitive test would have detected MRD in at least some of the 20 Induction I MRD-negative patients who subsequently relapsed. Thus, newer flow cytometric methods based on more than 4-color analysis should improve the sensitivity of MRD testing and may well improve the identification of patients who are likely to have a poor early treatment response.

Our results, together with those recently reported by the Japanese Childhood AML Cooperative Study Group, demonstrate that survival rates of 70% can be achieved in children with AML, and that for some subgroups of patients, the outcome now approaches that of patients with low-risk ALL. For example, patients with low-risk AML defined by genetic features had a 90% survival in AML02, confirming previous reports.^{4,17} Moreover, among the 57 patients with *CBF* abnormalities, three of the five deaths were related to sepsis, underscoring the low relapse rate for this subgroup and the importance of good supportive care. Improved survival rates for other subtypes of AML will require additional refinements in risk assignment and the introduction of novel agents that can circumvent the development of drug resistance. To this end, our current AML08 protocol incorporates newly discovered molecular predictors of outcome, highly sensitive MRD tests, novel agents such as clofarabine, sorafenib, and valproic acid, and natural killer cell therapy.³⁶

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Sheila Shurtleff, PhD, for performing molecular diagnostic analyses, Kathy Jackson for data collection, and Julie Groff for preparing the figures. This work was supported by R01 CA115422 and Cancer Center Support (CORE) P30 CA021765 grants from the National Institutes of Health, and by the American Lebanese Syrian Associated Charities (ALSAC).

Funding. Supported by R01 CA115422 and Cancer Center Support (CORE) P30 CA021765 grants from the National Institutes of Health, and by the American Lebanese Syrian Associated Charities (ALSAC).

References

1. Pui CH, Campana D, Pei D, et al. Treating childhood acute lymphoblastic leukemia without cranial irradiation. *N Engl J Med*. 2009; 360:2730–2741. [PubMed: 19553647]
2. Creutzig U, Zimmermann M, Lehnbecher T, et al. Less toxicity by optimizing chemotherapy, but not by addition of granulocyte colony-stimulating factor in children and adolescents with acute myeloid leukemia: results of AML-BFM 98. *J Clin Oncol*. 2006; 24:4499–4506. [PubMed: 16983120]
3. Lange BJ, Smith FO, Feusner J, et al. Outcomes in CCG-2961, a children's oncology group phase 3 trial for untreated pediatric acute myeloid leukemia: a report from the children's oncology group. *Blood*. 2008; 111:1044–1053. [PubMed: 18000167]
4. Tsukimoto I, Tawa A, Horibe K, et al. Risk-stratified therapy and the intensive use of cytarabine improves the outcome in childhood acute myeloid leukemia: the AML99 trial from the Japanese

- Childhood AML Cooperative Study Group. *J Clin Oncol*. 2009; 27:4007–4013. [PubMed: 19620491]
5. Grimwade D, Hills RK. Independent prognostic factors for AML outcome. *Hematology Am Soc Hematol Educ Program*. 2009:385–395. [PubMed: 20008224]
 6. Campana D. Determination of minimal residual disease in leukaemia patients. *Br J Haematol*. 2003; 121:823–838. [PubMed: 12786792]
 7. Coustan-Smith E, Ribeiro RC, Rubnitz JE, et al. Clinical significance of residual disease during treatment in childhood acute myeloid leukaemia. *Br J Haematol*. 2003; 123:243–252. [PubMed: 14531905]
 8. Maurillo L, Buccisano F, Del Principe MI, et al. Toward optimization of postremission therapy for residual disease-positive patients with acute myeloid leukemia. *J Clin Oncol*. 2008; 26:4944–4951. [PubMed: 18606980]
 9. Sievers EL, Lange BJ, Alonzo TA, et al. Immunophenotypic evidence of leukemia after induction therapy predicts relapse: results from a prospective Children's Cancer Group study of 252 patients with acute myeloid leukemia. *Blood*. 2003; 101:3398–3406. [PubMed: 12506020]
 10. Langebrake C, Creutzig U, Dworzak M, et al. Residual disease monitoring in childhood acute myeloid leukemia by multiparameter flow cytometry: the MRDAML-BFM Study Group. *J Clin Oncol*. 2006; 24:3686–3692. [PubMed: 16877738]
 11. Bishop JF, Matthews JP, Young GA, et al. A randomized study of high-dose cytarabine in induction in acute myeloid leukemia. *Blood*. 1996; 87:1710–1717. [PubMed: 8634416]
 12. Weick JK, Kopecky KJ, Appelbaum FR, et al. A randomized investigation of high-dose versus standard-dose cytosine arabinoside with daunorubicin in patients with previously untreated acute myeloid leukemia: a Southwest Oncology Group study. *Blood*. 1996; 88:2841–2851. [PubMed: 8874180]
 13. Kern W, Estey EH. High-dose cytosine arabinoside in the treatment of acute myeloid leukemia: Review of three randomized trials. *Cancer*. 2006; 107:116–124. [PubMed: 16721819]
 14. Vardiman JW, Thiele J, Arber DA, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*. 2009; 114:937–951. [PubMed: 19357394]
 15. Rubnitz JE, Onciu M, Pounds S, et al. Acute mixed lineage leukemia in children: the experience of St Jude Children's Research Hospital. *Blood*. 2009; 113:5083–5089. [PubMed: 19131545]
 16. Raimondi SC, Zhou Y, Mathew S, et al. Reassessment of the prognostic significance of hypodiploidy in pediatric patients with acute lymphoblastic leukemia. *Cancer*. 2003; 98:2715–2722. [PubMed: 14669294]
 17. Rubnitz JE, Raimondi SC, Tong X, et al. Favorable impact of the t(9;11) in childhood acute myeloid leukemia. *J Clin Oncol*. 2002; 20:2302–2309. [PubMed: 11981001]
 18. Kurt B, Flynn P, Shenep JL, et al. Prophylactic antibiotics reduce morbidity due to septicemia during intensive treatment for pediatric acute myeloid leukemia. *Cancer*. 2008; 113:376–382. [PubMed: 18459178]
 19. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics*. 1979; 35:549–556. [PubMed: 497341]
 20. Zelen M. The randomization and stratification of patients to clinical trials. *J Chronic Dis*. 1974; 27:365–375. [PubMed: 4612056]
 21. Kaplan EL, Meier P. Non-parametric estimation for incomplete observations. *J Am Stat Assoc*. 1958; 53:457–481.
 22. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. *Br J Cancer*. 1977; 35:1–39. [PubMed: 831755]
 23. Pocock SJ, Clayton TC, Altman DG. Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. *Lancet*. 2002; 359:1686–1689. [PubMed: 12020548]
 24. Cox DR. Regression models and life-tables. *J R Stat Soc B*. 1972; 34:187–220.
 25. Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Annals of Statistics*. 1988; 16:1141–1154.

26. Rubnitz JE, Crews KR, Pounds S, et al. Combination of cladribine and cytarabine is effective for childhood acute myeloid leukemia: results of the St Jude AML97 trial. *Leukemia*. 2009; 23:1410–1416. [PubMed: 19242495]
27. Becton D, Dahl GV, Ravindranath Y, et al. Randomized use of cyclosporin A (CsA) to modulate P-glycoprotein in children with AML in remission: Pediatric Oncology Group Study 9421. *Blood*. 2006; 107:1315–1324. [PubMed: 16254147]
28. Stevens RF, Hann IM, Wheatley K, Gray RG. Marked improvements in outcome with chemotherapy alone in paediatric acute myeloid leukemia: results of the United Kingdom Medical Research Council's 10th AML trial. MRC Childhood Leukaemia Working Party. *Br J Haematol*. 1998; 101:130–140. [PubMed: 9576193]
29. Gibson BE, Wheatley K, Hann IM, et al. Treatment strategy and long-term results in paediatric patients treated in consecutive UK AML trials. *Leukemia*. 2005; 19:2130–2138. [PubMed: 16304572]
30. Lie SO, Abrahamsson J, Clausen N, et al. Long-term results in children with AML: NOPHO-AML Study Group--report of three consecutive trials. *Leukemia*. 2005; 19:2090–2100. [PubMed: 16304571]
31. Razzouk BI, Estey E, Pounds S, et al. Impact of age on outcome of pediatric acute myeloid leukemia: a report from 2 institutions. *Cancer*. 2006; 106:2495–2502. [PubMed: 16639734]
32. Aplenc R, Alonzo TA, Gerbing RB, et al. Ethnicity and survival in childhood acute myeloid leukemia: a report from the Children's Oncology Group. *Blood*. 2006; 108:74–80. [PubMed: 16537811]
33. Athale UH, Razzouk BI, Raimondi SC, et al. Biology and outcome of childhood acute megakaryoblastic leukemia: a single institution's experience. *Blood*. 2001; 97:3727–3732. [PubMed: 11389009]
34. Woods WG, Kobrinsky N, Buckley JD, et al. Time-sequential induction therapy improves postremission outcome in acute myeloid leukemia: a report from the Children's Cancer Group. *Blood*. 1996; 87:4979–4989. [PubMed: 8652810]
35. Woods WG, Neudorf S, Gold S, et al. A comparison of allogeneic bone marrow transplantation, autologous bone marrow transplantation, and aggressive chemotherapy in children with acute myeloid leukemia in remission. *Blood*. 2001; 97:56–62. [PubMed: 11133742]
36. Rubnitz JE, Inaba H, Ribeiro RC, et al. NKAML: a pilot study to determine the safety and feasibility of haploidentical natural killer cell transplantation in childhood acute myeloid leukemia. *J Clin Oncol*. 2010; 28:955–959. [PubMed: 20085940]



Figure 1. AML02 treatment schema

After enrollment, patients were randomized to receive daunorubicin (50 mg/m² on days 2, 4, and 6) and etoposide (100 mg/m² on days 2–6) plus either high-dose cytarabine (HDAC, 3 g/m² every 12 hours on days 1, 3, and 5) or low-dose cytarabine (LDAC, 100 mg/m² every 12 hours on days 1–10), constituting the Induction I phase of treatment. Induction II consisted of low-dose cytarabine (100 mg/m² every 12 hours on days 1–8), daunorubicin, and etoposide (ADE) with or without 3 mg/m² gemtuzumab ozogamicin (GO). Standard-risk patients who lacked matched related donors and all low-risk patients received three courses (CI, CII, CIII) of consolidation therapy, whereas standard-risk patients with matched related donors and high-risk patients were eligible to receive hematopoietic stem cell transplantation (HSCT). CI consisted of cytarabine (500 mg/m²/day by continuous infusion for 5 days) and cladribine (9 mg/m² on days 1–5) for patients with t(9;11) and inv(16); cytarabine (3 g/m² every 12 hours on days 1–3) and etoposide (125 mg/m² on days 2–5) for patients with M4 or M5 AML without t(9;11) or inv(16); and cytarabine (3 g/m² every 12 hours on days 1–3) and mitoxantrone (10 mg/m² on days 3–4) for all other patients. CII consisted of cytarabine (3 g/m² every 12 hours on days 1, 2, 8, 9) and L-Asparaginase (6000 Units/m² 3 hours after the 4th and 8th doses of cytarabine). CIII included mitoxantrone (10 mg/m² on days 1–3) and cytarabine (1 g/m² every 12 hours on days 1–3).

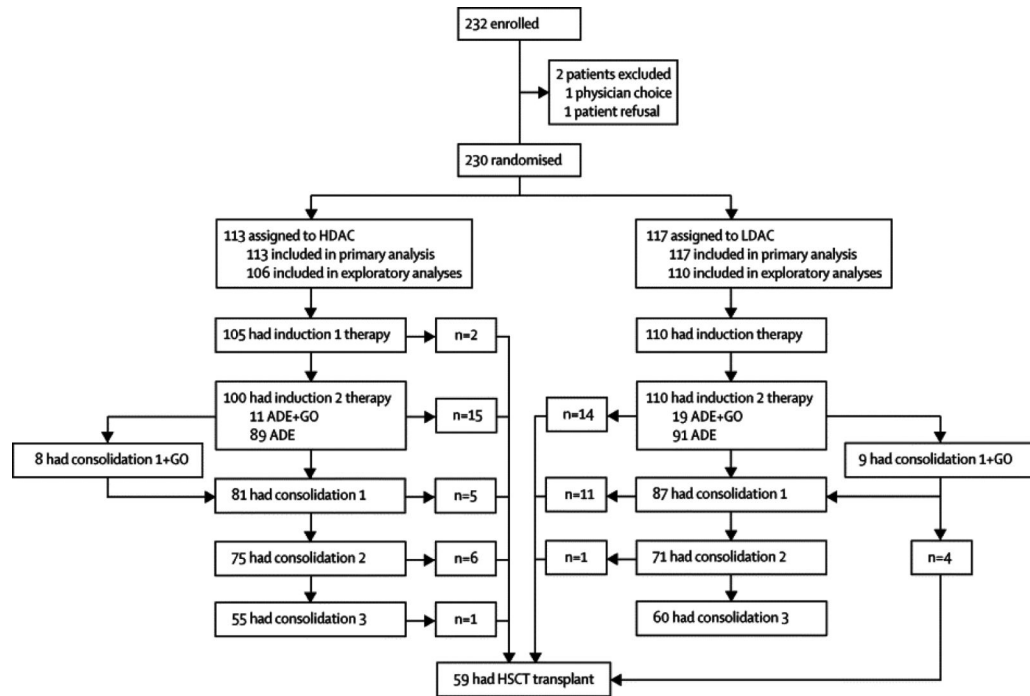


Figure 2.
Patient Flow Chart

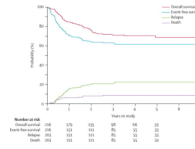


Figure 3. Outcome for patients treated in the AML02 study
 Upper curves show the probabilities of overall survival and event-free survival; lower curves show the cumulative incidences of relapse and of death unrelated to relapse.

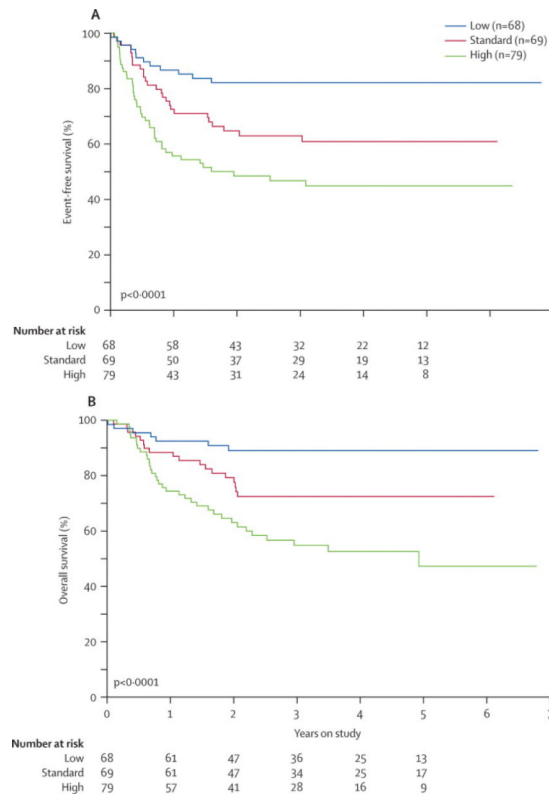


Figure 4. Outcome according to risk group

(A) Probability of event-free survival according to risk group. (B) Probability of overall survival according to risk group.

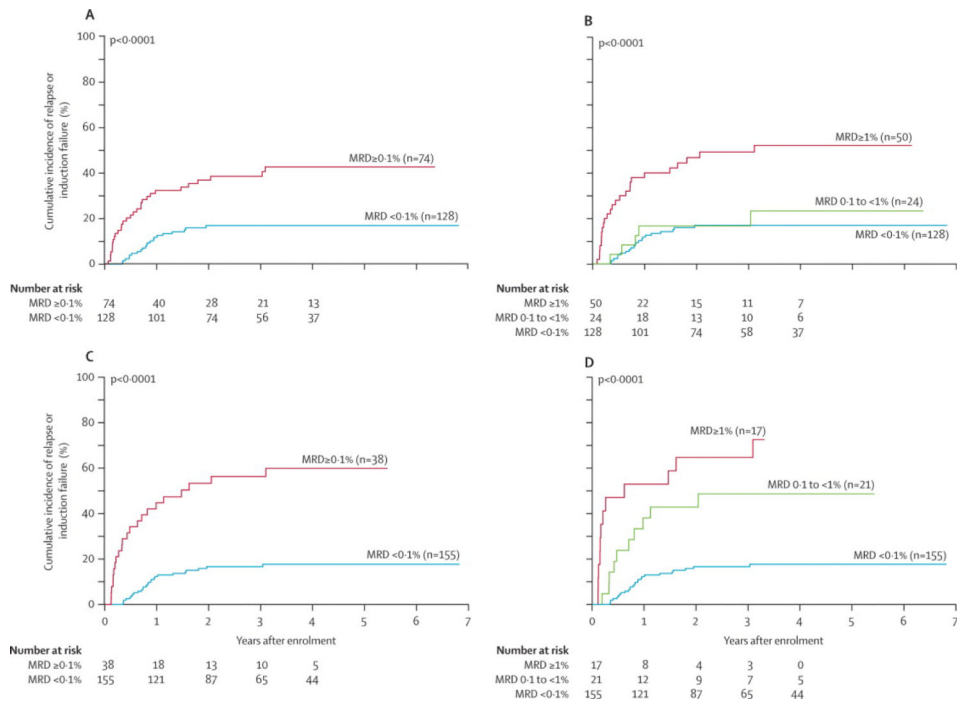


Figure 5. Cumulative incidence (CI) of relapse or induction failure according to minimal residual disease (MRD)

(A) CI for patients who were positive or negative for MRD after Induction I. (B) CI for patients with negative, low, or high MRD after of Induction I. (C) CI for patients who were positive or negative for MRD after Induction II. (D) CI for patients with negative, low, or high MRD after Induction II.

Table 1

Characteristics of Randomized Patients

Characteristic	High-dose cytarabine	Low-dose cytarabine
Gender		
Female	52	49
Male	61	68
Race		
Black	22	21
White	76	84
Other	15	12
FAB		
M0	2	2
M1	13	14
M2	20	17
M4	12	21
M4Eo	10	13
M5	27	27
M6	2	0
M7	13	12
Not available ^a	14	11
Karyotype		
inv(16)	11	15
t(8;21)	17	14
t(9;11)	7	8
Other 11q23	14	13
Normal	18	35
Other	41	31
Not available	5	1
Age (years)		
Median	9.3	9.0
Range	0.1–19.3	0.0–21.4
WBC count		
Median	16.2	34.6
Range	0.8–340	0.3–286
<i>FLT3</i> status		
ITD	13	19
PM	2	6
Wild type	89	85
Not available	9	7

Abbreviations: WBC, white blood cell; ITD, internal tandem duplication; PM, point mutation

^aFAB classification was not performed on 25 cases because bone marrow aspirations and biopsies were not performed (only peripheral blood was examined) or because samples were inadequate for a conclusive morphologic classification.

Table 2Overall Results and Causes of Failure Among Randomized AML Patients^a

	High-dose arm (n=106)	Low-dose arm (n=110)	P value ^b
3-year cumulative incidence of failure			
Induction failure	4.7 ± 2.1	5.5 ± 2.2	0.81
Death unrelated to relapse	11.9 ± 3.3	5.5 ± 2.2	0.13
Any relapse	17.5 ± 3.8	21.5 ± 4.0	0.46
Hematologic relapse	14.7 ± 3.5	18.8 ± 3.8	0.44
CNS relapse	2.8 ± 1.6	0.9 ± 0.9	0.30
6-month cumulative incidence of infection			
Grade 2 or greater bacterial infection	70.8 ± 4.5	70.0 ± 4.4	0.42
Grade 2 or greater invasive fungal infection	23.6 ± 4.2	13.6 ± 3.3	0.06
Grade 3 or greater infection	79.3 ± 4.0	75.5 ± 4.2	0.17
Grade 4 or greater infection	18.9 ± 3.8	10.0 ± 2.9	0.06
3-year event-free survival			
All patients	60.2 ± 5.9	65.7 ± 5.7	0.41
Low risk	79.0 ± 9.3	85.2 ± 7.5	0.36
Standard risk	51.9 ± 10.4	72.4 ± 8.7	0.12
High risk	51.2 ± 8.7	42.0 ± 10.7	0.37
3-year overall survival			
All patients	68.8 ± 5.5	73.4 ± 5.3	0.41
Low risk	84.6 ± 8.3	93.5 ± 5.1	0.13
Standard risk	67.1 ± 9.6	77.0 ± 8.3	0.34
High risk	57.7 ± 8.8	51.3 ± 10.3	0.50

^aExcluding patients with mixed-lineage leukemia^bP values determined by Gray's test for cumulative incidence comparisons and by log-rank test for survival comparisons

Table 3

Analysis of Prognostic Factors Among Patients with AML^a

Factor	N	3-year EFS ± SE (%)	P value	3-year OS ± SE (%)	P value ^b
Gender			0.87		0.69
Female	95	63.2 ± 5.8		71.9 ± 5.4	
Male	121	62.8 ± 5.8		70.4 ± 5.4	
Race			0.07		0.12
Black	41	53.4 ± 10.1		61.0 ± 10.2	
White	151	62.6 ± 5.0		70.7 ± 4.6	
Other	24	82.6 ± 8.6		91.0 ± 6.4	
Age (years)			0.20		0.40
<10	112	66.3 ± 5.3		74.1 ± 4.9	
≥ 10	104	59.3 ± 6.5		67.9 ± 6.0	
WBC count			0.57		0.80
< 50	151	62.4 ± 4.7		71.5 ± 4.4	
≥ 50	65	64.3 ± 8.2		70.1 ± 7.5	
Induction I MRD			<0.001		<0.001
< 0.1%	128	73.6 ± 5.0		80.5 ± 4.5	
≥ 0.1%	74	43.1 ± 6.9		54.1 ± 6.9	
Induction I MRD			<0.001		<0.001
< 0.1%	128	73.6 ± 5.0		80.5 ± 4.5	
0.1 to < 1%	24	65.9 ± 11.6		74.3 ± 10.9	
≥ 1%	50	32.1 ± 7.6		44.5 ± 8.0	
FLT3-ITD			0.002		0.004
Present	28	39.3 ± 12.5		47.6 ± 12.2	
Absent	173	67.6 ± 4.4		75.2 ± 4.1	
CBF leukemia			<0.001		<0.001

Factor	N	3-year EFS ± SE (%)	P value	3-year OS ± SE (%)	P value ^b
Present	57	85.8 ± 6.1		90.6 ± 5.1	
Absent	154	55.5 ± 4.9		65.6 ± 4.7	
t(9;11)			0.63		0.50
Present	15	66.7 ± 14.5		80.0 ± 11.9	
Absent	196	63.4 ± 4.3		71.6 ± 4.0	
Other 11q23			0.19		0.08
Present	26	50.0 ± 11.8		55.9 ± 11.7	
Absent	185	65.6 ± 4.4		74.6 ± 4.0	
M7 without t(1;22)			0.04		0.003
Present	18	36.5 ± 11.9		44.0 ± 12.4	
Absent	177	67.0 ± 4.5		75.5 ± 4.1	

Abbreviations: EFS, event-free survival; OS, overall survival; WBC, white blood cell; ITD, internal tandem duplication

^aExcluding patients with mixed-lineage leukemia

^bP values are from a 10,000 permutation Monte Carlo approximation to the exact log-rank test.

Table 4

Multivariable analysis of prognostic factors

Factors	EFS			OS		
	HR	95% CI	P-value	HR	95% CI	P-value
Treatment Arm	0.83	(0.50, 1.39)	0.483	0.83	(0.46, 1.53)	0.554
Induction I MRD ≥ 1%	2.41	(1.36, 4.26)	0.003	2.11	(1.09, 4.11)	0.028
CBF	0.32	(0.13, 0.79)	0.013	0.37	(0.12, 1.12)	0.079
Age at diagnosis ^a	1.65	(0.92, 2.95)	0.095	2.55	(1.23, 5.27)	0.012
Other 11q23	2.05	(0.96, 4.34)	0.062	3.29	(1.34, 8.04)	0.009
M7 without t(1;22)	2.17	(0.86, 5.45)	0.099	5.64	(1.99, 16.0)	0.001
FLT3-YTD	1.64	(0.82, 3.27)	0.163	1.93	(0.84, 4.44)	0.123

Abbreviations: EFS, event-free survival; OS, overall survival; HR, hazard ratio; CI, confidence interval; ITD, internal tandem duplication;

^a Age at diagnosis is dichotomized as less than or great than 10 years.