Outcome after induction chemotherapy for older patients with acute myeloid leukemia is not improved with mitoxantrone and etoposide compared to cytarabine and daunorubicin: a Southwest Oncology Group study

Jeanne E. Anderson, Kenneth J. Kopecky, Cheryl L. Willman, David Head, Margaret R. O'Donnell, Frederick W. Luthardt, Thomas H. Norwood, I-Ming Chen, Stanley P. Balcerzak, David B. Johnson, and Frederick R. Appelbaum

Complete remission and long-term survival rates are low for older adults treated for acute myeloid leukemia (AML). Because of favorable phase 2 data using mitoxantrone and etoposide, we conducted a phase 3 study (SWOG-9333) in which patients over 55 years of age with previously untreated AML were randomized to receive mitoxantrone (10 mg/m^2 per day × 5) and etoposide (100 mg/m^2 per day × 5) [ME], or cytarabine (200 mg/m^2 per day × 7) and daunorubicin (45 mg/m^2 per day × 3) [AD] as induction therapy. The randomization was stratified by age, onset of leukemia, and multidrug resistance phenotype. Over a 4-year period, 328 eligible patients from 66 institutions were enrolled. The complete remission rate was 34% (95% confidence interval [CI] 26%-41%) for patients in the ME and 43% (CI 35%-51%) for patients in the AD treatment arm (one-tailed P value .96). The rates of resistant disease were 43% (CI 35%-51%) and 34% (CI 27%-42%), respectively, for the 2 treatment arms (one-tailed P value .95). The estimated overall survival at 2 years was 11% (CI 6%-15%) and 19% (CI 12%-25%) for patients randomized to ME and to AD induction therapy, respectively (one-tailed P value .99). After accounting for the independent prognostic factors associated with survival (karyotype, performance status, age, white blood cell count), exploratory analysis suggested there was a worse survival for patients who received ME compared with AD induction therapy (2-tailed P value .0066). We conclude that the results of our study do not demonstrate any benefit to the use of ME induction chemotherapy instead of AD in older patients with AML. (Blood. 2002;100: 3869-3876)

© 2002 by The American Society of Hematology

Introduction

The incidence of acute myeloid leukemia (AML) increases rapidly with increasing age resulting in a median age at diagnosis of approximately 65 years.[1] Unfortunately, the majority of advances in therapy for AML have been restricted to younger patients. Intensive induction and consolidation therapy with agents such as idarubicin[2,3] and high-dose cytarabine,[4-7] and with hematopoietic stem cell transplantation[8] have predominantly benefited younger rather than older patients with AML. Numerous factors may contribute to the inferior outcome with therapy in older patients. Older patients more often have an antecedent myelodysplastic syndrome (MDS) and poor-risk karyotype. Their disease more often exhibits a multidrug resistance (MDR) phenotype resulting in greater intrinsic resistance of leukemic blasts to many therapeutic agents.[9-14] Furthermore, older patients often have more comorbid conditions than younger patients and tolerate prolonged pancytopenia and intensive therapy less well.[1,11,15] Therefore, clinical trials specifically designed to target the therapeutic challenges of older AML patients are needed.

The use of conventional anthracycline and cytarabine regimens in older patients (usually defined as over 55-65 years of age) results in complete remission (CR) rates of approximately 25% to 55% and long-term survival rates of approximately 10%.[1,13,14,16,17] For example, in an earlier Southwest Oncology Group (SWOG) study of 211 patients over 55 years of age (SWOG-9031) who received cytarabine 200 mg/m^2 per day for 7 days and daunorubicin 45 mg/m^2 per day for 3 days, there was a 45% CR rate, 8-month median overall survival (OS), and 9-month median relapse-free survival (RFS).[16] Mitoxantrone is an anthraquinone that has replaced anthracyclines in some cytarabine-containing induction regimens. Early studies suggested that mitoxantrone might be better tolerated than anthracyclines.[18,19] Etoposide, an epipodophyllotoxin with antileukemic activity, has shown efficacy when incorporated into induction, consolidation, and salvage chemotherapy regimens, and when incorporated into myeloablative regimens for hematopoietic stem cell transplantation in patients with AML.[20,21] A regimen combining mitoxantrone and etoposide was first reported in 61 patients with relapsed or refractory AML, with a CR rate of 43%.[22] In a subsequent phase 2 study, 21 patients with previously untreated MDS-related AML were reported to have

a 57% CR rate.[23] Finally, a phase 2 study with this regimen as first-line therapy in 67 older patients with AML reported a 55% CR rate, 9.2-month median OS, and 8.4-month median RFS.[24] Because of these encouraging results in high-risk patients, we conducted a randomized, phase 3 study, SWOG-9333, comparing standard cytarabine and daunorubicin (AD) induction with mitoxantrone and etoposide (ME) in patients over 55 years of age with previously untreated AML. The primary objectives were to determine whether induction therapy with ME resulted in improved remission rates, OS, or RFS compared with conventional therapy with AD, and to compare the toxicities of the 2 regimens. We further wanted to know if patients induced with ME and treated with AD consolidation would have improved outcome compared with patients exposed only to AD.

Patients, materials, and methods

Eligibility

Patients 56 years of age or older with a morphologically confirmed diagnosis of previously untreated AML, except for acute promyelocytic leukemia, were eligible.[25,26] Patients with blast crisis of chronic myeloid leukemia were excluded. Prior treatment for hyperleukocytosis with agents such as hydroxyurea, or for suspected or proven central nervous system involvement with intrathecal chemotherapy was allowed. Patients with secondary AML, defined as AML arising after a morphologic diagnosis of MDS[27] or AML arising after prior chemotherapy or radiotherapy, were not excluded. Prior treatment for MDS with cytarabine (doses ≤ 100 mg/m^2) administered more than 30 days before registration was allowed. Patients were required to have performance status 0-3 by SWOG criteria. Intact solid organ function was required, as defined by (1) serum bilirubin ≤ 2 times the institutional upper limits of normal (IULN); (2) serum creatinine less than or equal to 1.3 times IULN and/or creatinine clearance more than or equal to 50 cc/min; (3) ejection fraction more than or equal to 50% as measured by either multigated cardiac blood pool (MUGA) scan or echocardiography; and (4) absence of unstable cardiac arrhythmia or unstable angina. Informed consent in accordance with local institutional guidelines and federal regulations was required of all patients.

Eligibility for the study also required submission of a bone marrow aspirate sample collected within 14 days before randomization to the SWOG Myeloid Repository at the University of New Mexico, for measurement of the expression of the MDR glycoprotein MDR-1.[9,10] MDR-1 expression by leukemic blasts was measured using the MDR-1–specific antibody MRK16, as previously described,[9,10] for use as a stratification factor.

Study design

Randomization, stratification, and induction therapy. Patients were randomly assigned to one of 2 remission induction treatment arms: (ME) mitoxantrone 10 mg/m^2 intravenously over 30 minutes daily on days 1 through 5 and etoposide 100 mg/m^2 intravenously over 30 minutes daily on days 1 through 5; or (AD) cytarabine 200 mg/m^2 daily as 24-hour continuous intravenous infusion on days 1 through 7, and daunorubicin 45 mg/m^2 intravenous push daily on days 1 through 3. The doses of drugs in the AD regimen were chosen to be the same as those used in the preceding SWOG study in older patients with AML, SWOG-9031.[16] A bone marrow aspirate and biopsy was obtained on day 10 and, if the day-10 marrow blast percentage was more than or equal to 5%, again on day 14. Once a marrow blast percentage less than 5% was documented, each patient received treatment with yeast-derived granulocyte macrophage–colony-stimulating factor (GM-CSF) at 250 mg/m^2 intravenously once daily until the absolute neutrophil count (ANC) reached 1500/μL.[28] If the day-14 marrow showed more than or equal to 5% blasts, then a second cycle of induction chemotherapy was administered upon the earliest of recovery of marrow cellularity to more than 20% or of white blood cells (WBCs) to more than 2500/μL, or day 21. The dose of drugs administered for cycle 2 was the same as cycle 1, except as follows: mitoxantrone and daunorubicin were

reduced by 50% for serum bilirubin between 2 to 4 times the IULNs or discontinued for higher serum bilirubin; etoposide was reduced by 50% for creatinine between 25 cc/min and 50 cc/min or by 75% for creatinine clearance less than 25 cc/min.

The randomization was stratified by age (56-64 years vs 65 years and greater), onset of leukemia (secondary vs de novo), and MDR-1 expression (negative vs positive vs indeterminate).

Supportive care. All patients received antimicrobial prophylaxis as follows: (1) ciprofloxicin 500 mg orally or 200 mg intravenously twice daily until administration of broad spectrum antibiotics for febrile neutropenia; (2) fluconazole 200 mg orally or intravenously daily until administration of amphotericin B, and (3) among patients who were seropositive for herpes simplex virus, acyclovir 800 mg orally or 250 mg/m^2 intravenously twice daily. If patients developed febrile neutropenia, ciprofloxicin was discontinued and parenteral broad-spectrum antipseudomonal antibiotic coverage was instituted. Antibiotics were continued until the ANC was over 500/μL, and clinical signs of infection resolved. Guidelines for transfusion support included prophylactic platelet transfusion at a threshold of 10 000/μL in asymptomatic patients and red cell transfusion at a threshold of hemoglobin of 8 gm/dL. Early institution of hyperalimentation was recommended for patients with poor oral intake.

Postremission therapy. Patients who achieved a CR after 1 or 2 cycles of induction chemotherapy were registered to receive 2 cycles of postremission therapy, consisting of daunorubicin 30 mg/m^2 intravenous push daily on days 1 and 2 and cytarabine 200 mg/m^2 daily as 24-hour continuous intravenous infusion on days 1 through 5. Performance status 0-2 by SWOG criteria and serum bilirubin less than or equal to twice the IULN were required for postremission therapy, as was a MUGA scan demonstrating ejection fraction more than or equal to 45% for patients with cardiac toxicity during or after induction therapy. The use of GM-CSF with postremission therapy was left to the discretion of the treating physician.

Definition of outcomes

Response to remission induction therapy was evaluated according to SWOG criteria as modified from National Cancer Institute–sponsored workshop guidelines.[25] Patients who failed to achieve CR after induction therapy were classified according to type of failure: resistant disease (RD), death during aplasia, or indeterminate.[25] OS was measured from the date of randomization until death from any cause, with observation censored at the date of last contact for patients last known to be alive. RFS was measured from the date of CR until relapse or death from any cause, with observation censored at the date of last contact for patients last known to be alive without report of relapse. Times to neutrophil recovery were measured from the start of protocol induction therapy until the first day that the ANC recovered to more than or equal to 500/μL for 2 consecutive measurements on different days. Observation was censored for patients who died while neutropenic or whose ANC data were inadequate to determine the date of recovery. Time to platelet recovery to more than or equal to 30 000/μL (for at least 7 consecutive days) was defined analogously. Duration of hospitalization was measured from the start of protocol induction therapy until the first discharge from the hospital or death, whichever occurred first. Toxicities were defined and graded according to the National Cancer Institute (NCI) Common Toxicity Criteria, version 2.0.

Laboratory and cytogenetic data

MDR-1 expression by the leukemic blasts was measured using MRK16 (Kamiya Biomedical, Seattle, WA) in 3-color flow cytometric assays where blasts were costained with CD34 (hematopoietic stem/progenitor cell antigen) and CD33 (panmyeloid antigen) as previously described.[9,12] This costaining allows accurate analysis of blast populations for MDR-1 proteins, CD33 and CD34 expression. MRK16 staining was detected in a 2-step approach using first the biotin-labeled goat anti–mouse IgG2a antibody (Southern Biotechnology Associates, Birmingham, AL) to detect MRK16, and then streptavidin-RED670 (Gibco-BRL, Rockville, MD) to detect the biotin label. The biotin-avidin detection system enhances the dim staining of MDR-1 expression on leukemic cells. Appropriately matched isotype controls were used in all assays.[9,12] To assess functional drug efflux of the leukemia blasts, 2 fluorescent dyes, Di(OC)$_2$ and Rhodamine 123

(Rh123), were measured in a single-color flow cytometric assay as described in previous reports.[9,12]

Cytogenetic studies on bone marrow or unstimulated peripheral blood samples were performed in SWOG-approved cytogenetic laboratories and reviewed by at least 3 members of the SWOG Cytogenetic Committee as previously described.[10,29] Cytogenetic abnormalities were grouped according to published criteria adopted by SWOG as favorable, intermediate, unfavorable, or unknown.[10,29]

Statistical methods

Comparisons of treatment outcomes between arms were based on intent-to-treat analyses of all eligible patients, that is, patients who received none or only part of their assigned induction therapy were included in their original treatment arm. CR, RD, and induction toxicity rates were analyzed using the Fisher exact test and logistic regression analysis.[30] Distributions of OS and RFS were estimated by the method of Kaplan and Meier, and analyzed according to treatment assignment and patient and disease characteristics using proportional hazards regression models.[31,32]

The primary null hypotheses of this study were that the CR rates produced by AD and ME induction regimens are equal, that is, that the logistic regression coefficient representing the treatment difference is 0, and that the hazard ratios representing the OS and RFS differences are 1.0. The study design called for randomization of 400 eligible patients. If the true CR rate is 50% with AD, this would provide statistical power of 90% (one-tailed test at 5% critical level) to detect an increased CR rate of 65% with ME, corresponding to an alternative hypothesis that the logistic regression coefficient is 0.62. Assuming a median OS of 6 months for the AD arm, 400 patients accrued over an expected 4 years, with one additional year's follow-up, would also provide 98% power under the alternative hypothesis that the mortality hazard ratio (ME/AD) is 0.67. Since the principal objective of the study was to assess whether treatment outcomes are superior with ME, the statistical significance of treatment differences in outcomes was measured by one-sided P values for superior outcome in the ME arm. All other test results are reported using 2-sided P values. Confidence intervals (CIs) were calculated at the 95% confidence level. Results are based on data available February 28, 2001.

Early termination of the study

A Data and Safety Monitoring Committee (DSMC), of which the investigators were not members, monitored this study throughout its accrual and follow-up phases. There were 2 interim analyses performed as required by the study protocol. These analyses allowed the DSMC to assess, on the basis of predefined significance levels, whether the study should be terminated earlier than planned because there was sufficient evidence either to reject the null hypotheses, suggesting that ME is superior to AD, or to reject the alternative hypotheses specified above, suggesting that ME is not sufficiently superior to AD to warrant its clinical use. In the second interim analysis, based on data available September 14, 1998, the alternative hypotheses were rejected at sufficiently high levels of significance for both CR ($P = .0001$) and OS ($P < .0001$). Based on these results the DSMC directed the early closure of the study to accrual of new patients.

Results

Patient accrual and characteristics

A total of 334 patients from 66 institutions were enrolled in this study between January 1995 and December 1998. There were 6 patients who were ineligible because of MDS diagnosis (n = 4), acute promyelocytic leukemia diagnosis (n = 1), or presence of a concomitant malignancy (n = 1) at time of registration. Of the 328 eligible patients, 167 were randomized to the ME arm and 161 to the AD arm. Patient and disease characteristics were generally similar in the 2 treatment arms (Tables 1-2). About half of the patients expressed MDR-1 or cyclosporine A (CsA)–inhibited efflux at moderate or bright levels, and the quantitative level of

Table 1. Characteristics of 328 eligible patients, by induction treatment arm

	AD (n = 161)		ME (n = 167)	
	n	%	n	%
Sex				
Female	78	48	66	40
Male	83	52	101	60
Race/ethnicity				
White (non-Hispanic)	141	88	149	89
Black (non-Hispanic)	11	7	12	7
Hispanic	5	3	5	3
Asian or Pacific Islander	3	2	0	
Other/Unknown	1	1	1	1
Age, y				
56-64	57	35	57	34
65+	104	65	110	66
AML onset				
Secondary	36	22	38	23
De novo	125	78	129	77
MDR stratum				
MDR+	98	61	105	63
MDR−	63	39	62	37
Performance status				
0-1	125	78	127	76
2-3	36	22	40	24
WBC count				
Below 10 000/μL	77	48	81	49
10 000/μL or above	84	52	86	51
FAB subtype (local)				
M1	32	20	41	25
M2	59	37	48	29
M4	31	19	34	20
M5	15	9	18	11
M6	5	3	5	3
M7	3	2	0	
M0	12	7	16	10
Other/unknown	4	2	5	3
CD34				
Negative	51	32	42	25
Low	6	4	5	3
High	104	65	120	72
MDR-1 expression*				
Negative	64	40	65	39
Dim	19	12	16	10
Moderate	30	19	25	15
Bright	48	30	61	37
CsA-inhibited Di(OC)₂ efflux†				
Negative	69	44	66	40
Dim	5	3	8	5
Moderate	24	15	22	13
Bright	59	38	69	42
CsA-inhibited Rh123 efflux†				
Negative	66	46	63	42
Dim	4	3	5	3
Moderate	26	18	20	13
Bright	47	33	61	41

AD indicates cytarabine/daunorubicin induction; and ME, mitoxantrone/etoposide induction.

*MDR-1 categories based on Kolmogorov-Smirnov D value: Negative D < 0.15, Dim 0.15 ≤ D < 0.20, Moderate 0.20 ≤ D < 0.30, Bright D ≥ 0.30.

†Efflux categories based on Kolmogorov-Smirnov D value: Negative D < 0.20, Dim 0.20 ≤ D < 0.25, Moderate 0.25 ≤ D < 0.40, Bright D ≥ 0.40.

MDR-1 expression was strongly correlated with CsA-inhibited efflux (Spearman rank order correlation coefficients 0.75 and 0.77 for Di(OC)₂ and RH123 efflux, respectively). Pretreatment cytogenetic results classified by SWOG criteria[10,29] are shown in Table 3. Cytogenetic results were not available for 56 (17%) eligible patients, either because specimens were not submitted

Table 2. Overall characteristics of 328 eligible patients, by induction treatment arm

	AD (n = 161)		ME (n = 167)	
	Median	Range	Median	Range
Age, y	68	56-84	67	56-86
Marrow blasts, %	60	6-99	61	5-99
WBC count, × 1000/μL	11.6	0.8-273	10.5	0.6-241
Peripheral blasts, %	22	0-99	27	0-99
Peripheral blast count, × 1000/μL	3.1	0-237	2.5	0-207
Hemoglobin level, g/dL	9.1	4.3-14.4	9.1	5.9-14.9
Platelet count, × 1000/μL	51.5	2-1200	55.0	3-690

to approved laboratories (n = 12), or evaluable studies were not obtained (n = 44). Among the 272 patients with cytogenetic data, the proportion with normal karyotypes was just below 40% in both treatment arms; however, a slightly higher proportion of the ME patients had unfavorable karyotypes (46%) compared with the AD patients (39%). Due to the small number of patients with favorable cytogenetics (n = 13) or low CD34 expression (n = 11), they were combined with the intermediate cytogenetic risk group and CD34$^-$ patients, respectively, for most analyses.

Response to induction therapy

As shown in Table 4, 56 (34%) of the 167 ME patients achieved CR (CI 26%-41%), which was not superior to the CR rate of 69/161 or 43% (CI 35%-51%) observed for the AD patients (one-tailed P = .96 based on logistic regression analysis with adjustment for stratification by age, MDR-1 status, and AML onset). Only 44 (26%) of the ME patients achieved CR after a single course of induction therapy, compared with 56 (35%) of the AD patients. Among the patients who did not achieve CR after a single course, nearly equal proportions received the planned second course of identical induction therapy: 56/123 (46%) on the ME arm and 47/105 (45%) on the AD arm. However, the response rate to the second induction course was also somewhat lower with ME (12/56 or 21%) compared with AD (13/47 or 28%). Responses of 20 patients could not be determined; however, this incompleteness of response data could not explain the inferior result with ME. Even if all 8 ME patients and none of the 12 AD patients with undetermined responses had achieved CR, the CR rate would still be lower with ME (38%) than AD (43%).

Of 167 ME patients, 72 (43%) had resistant disease (CI 35%-51%) following one (n = 34) or 2 (n = 38) courses. In contrast, 55 (34%) of the 161 AD patients had resistant disease (CI 27%-42%) after one (n = 27) or 2 (n = 28) courses (one-tailed

Table 3. Cytogenetic characteristics of 328 eligible patients, by induction treatment arm

	AD (n = 161)		ME (n = 167)	
	n	%	n	%
Evaluable				
Yes	133	83	139	83
No	28	17	28	17
Normal cytogenetics				
Yes	52	39	51	37
No	81	61	88	63
Cytogenetic risk group				
Favorable	8	6	5	4
Intermediate	71	53	67	48
Unfavorable	52	39	64	46
Unknown	2	2	3	2

AD indicates cytarabine/daunorubicin induction; and ME, mitoxantrone/etoposide induction.

Table 4. Response to induction therapy of 328 eligible patients, by treatment arm

	AD (n = 161)		ME (n = 167)	
	n	%	n	%
Complete response	69	43	56	34
Resistant disease	55	34	72	43
Died during aplasia	4	2	8	5
Died within 7 days after treatment	12	7	12	7
Died without marrow assessment	9	6	11	7
Not adequately assessed*	12	7	8	5

AD indicates cytarabine/daunorubicin induction; and ME, mitoxantrone/etoposide induction.

*Responses of these 20 patients could not be adequately assessed using the NCI response criteria (see Cheson et al[25]). Of these patients, 3 (2 AD and 1 ME) died within 40 days of randomization; 14 (8 AD and 6 ME) died between 130 days and 27 months after randomization; and 3 were alive at 22, 41, and 42 months, respectively, after randomization.

P = .95 in stratified analysis). As in the comparison of CR rates, the inadequate response assessments of 20 patients could not explain the higher RD rate of ME patients. Even if all 12 inadequately assessed AD patients had RD, and none of the 8 ME patients did, the RD rate would not be lower with ME (43%) than with AD (42%).

The results above led us to consider whether the CR and RD rates might be significantly worse with ME than with AD. Since these questions were not part of the study protocol, but were prompted by examination of the results, the 2-tailed P values for the treatment comparisons above (P = .089 for CR, P = .11 for RD) are appropriate measures of statistical significance and do not indicate that the differences are statistically significant.

Responses to induction chemotherapy, according to selected clinical and disease characteristics, are shown in Table 5. In univariate analyses without adjustment for stratification, the CR rate decreased significantly with increasing age (P = .0015) and was significantly lower for patients with secondary AML compared with de novo AML (P = .011). The 34% CR rate among the 116 patients with unfavorable cytogenetics did not differ significantly from the 44% rate of the 151 patients with favorable or intermediate cytogenetics (P = .13), although it was noted that 10 (77%) of the 13 with favorable cytogenetics achieved CR. The CR rate decreased slightly but not significantly with increasing MDR-1 expression (P = .53) or CsA-inhibited efflux (P = .092 and .19 for Di(OC)$_2$ and Rh123, respectively). In multiple logistic regression analyses, after accounting for the significant effects of age (P = .0026) and onset of leukemia (P = .018), none of the other factors in Tables 1-3 had statistically significant prognostic associations with CR rate. Furthermore, after accounting for the effects of age and onset of leukemia, there was no evidence of benefit with ME induction therapy (one-tailed P = .97).

In univariate analyses (Table 5), the incidence of RD increased with increasing expression of MDR phenotype (P = .0038 for MRK16 expression, and P = .0012 and .0063 for cyclosporine-inhibited functional efflux of DiOC2 and Rh123, respectively). RD was also somewhat more frequent in patients with secondary AML (P = .048), performance status 2-3 (P = .043), or unfavorable karyotype (P = .048). In multivariate logistic regression analyses, after adjusting for the significant effect of CsA-inhibited Di(OC)$_2$ efflux, the effects of AML onset (P = .078), performance status (P = .047), and unfavorable karyotype (P = .069) were not clearly significant, nor were the effects of any other factors in Tables 1 and 2. MDR-1 expression retained no independent prognostic significance for RD after accounting for the effect of CsA-inhibited Di(OC)$_2$ efflux (P = .42), since the 2 were so strongly correlated.

Table 5. Treatment outcomes of 328 AML patients, by stratification factors, patient and disease characteristics, CD34 and MRK16 expression, CsA-inhibited efflux, and cytogenetics

	Complete response (CR)			Resistant disease (RD)			Overall survival (OS)			Relapse-free survival (RFS)		
	% CR	95 CI, %	P*	% RD	95 CI, %	P*	OS at 2 years, %	95 CI, %	P*	RFS at 2 years, %	95 CI, %	P*
Age			.0015			.17			.0056			.79
56-64 y	47	38-57		34	26-44		18	11-25		16	6-26	
65 y or older	33	27-40		41	34-48		13	8-17		18	9-27	
AML onset			.011			.048			.11			.89
De novo	42	36-48		36	30-42		16	11-20		18	11-26	
Secondary	26	16-37		49	37-61		10	3-17		11	1-33	
Performance status			.10			.043			.0011			.99
0-1	40	34-47		42	36-48		15	11-20		18	10-26	
2-3	30	20-42		29	19-40		12	4-19		12	0-25	
WBC count, × 1000/µL			.48			.19			.027			.40
Below 10.0	41	33-49		40	32-48		17	11-22		20	10-30	
10.0 or above	35	28-43		38	30-45		12	7-17		14	5-23	
CD34 expression			.69			.074			.25			.012
Neg/low	37	27-47		32	23-42		20	13-28		30	15-45	
High	39	32-46		42	35-49		12	7-16		11	5-18	
MRK16 expression†			.53			.0038			.97			.25
Negative	40	31-49		29	21-37		14	8-20		17	6-27	
Dim	37	21-55		40	24-58		7	0-15		15	0-35	
Moderate	38	25-52		44	30-58		27	15-39		29	9-48	
Bright	37	28-46		48	38-57		11	5-17		13	2-23	
CsA-inhibited Di(OC)₂ efflux†			.092			.0012			.19			.98
Neg/dim	42	34-50		28	21-36		16	10-22		17	7-26	
Moderate	37	23-52		50	35-65		13	3-23		6	0-29	
Bright	34	26-43		48	40-57		14	8-20		23	10-35	
CsA-inhibited Rh123 efflux†			.19			.0063			.23			.62
Neg/dim	42	34-51		30	22-38		16	10-22		18	8-28	
Moderate	39	25-55		46	31-61		20	8-31		22	3-41	
Bright	35	26-45		48	38-58		12	6-18		16	4-27	
Cytogenetic group			.13			.048			<.0001			.042
Fav/int	44	36-52		34	26-42		22	15-29		22	12-32	
Unfav	34	26-44		46	42-55		7	2-11		13	2-23	

*Two-tailed P value from simple logistic (CR, RD) or proportional hazards (OS, RFS) regression analyses. Age, WBC count, MRK16, and efflux are treated as quantitative variables, all others as dichotomous variables.

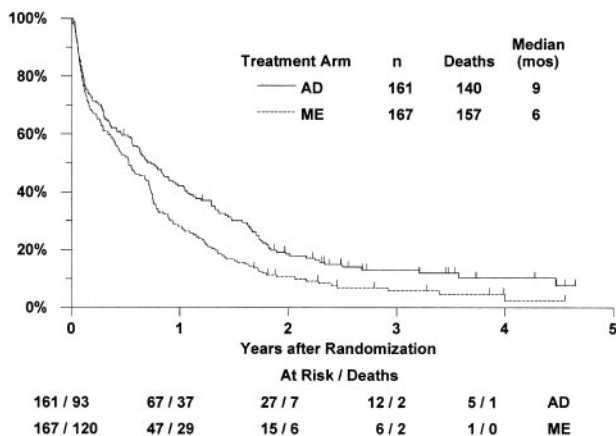†Categories as defined in Table 1.

There was no evidence of lower RD rates in the ME arm ($P = .91$) after adjusting for the effect of CsA-inhibited Di(OC)₂ efflux.

Overall survival by treatment arm

Of the 328 eligible patients, 297 (157 ME and 140 AD) have died. The remaining 31 were reported alive between 6 and 56 months (median 32 months) after randomization (Figure 1). For patients randomized to receive ME induction, the median survival was 6 months (CI 5-9 months) and the estimated probability of 2-year

Figure 1. Estimated distributions of overall survival, by induction treatment arm. Tickmarks indicate censored observations for patients last known to be alive.

survival was 11% (CI 6%-15%). For patients randomized to receive AD induction, the median survival was 9 months (CI 7-11 months) and the estimated 2-year survival was 19% (CI 12%-25%). Based on proportional hazards regression analysis adjusted for the stratification by age group, MDR-1 status, and onset of leukemia, survival was not significantly better with ME induction (one-tailed $P = .99$). The estimated hazard ratio ("relative risk") of death in the ME arm, compared with the AD arm, was 1.32 (CI 1.04-1.69). There was some evidence that survival was significantly worse with ME induction therapy compared with the AD therapy (2-tailed $P = .022$).

Results of the univariate proportional hazards regression analyses without adjustment for stratification are shown in Table 5. Survival decreased significantly with increasing age ($P = .0056$) or WBC count ($P = .027$), and was significantly poorer for patients with performance status 2-3 compared with 0-1 ($P = .0011$) and for patients with unfavorable compared with favorable/intermediate cytogenetics ($P < .0001$). The significant effect of performance status was due to the very high early mortality among patients with performance status 2-3, whose estimated probability of surviving 6 months was only 33% (CI 22%-43%) compared with 63% (CI 57%-69%) for those with performance status 0-1. In multiple proportional hazards regression analyses of survival, 4 factors were found to have statistically significant independent prognostic effects: survival was significantly poorer in patients with unfavorable cytogenetics (2-tailed $P = .0001$, compared with favorable/intermediate), or with performance status 2-3 ($P = .029$ compared

with performance status 0-1), and decreased significantly with increasing age ($P = .024$) or WBC count ($P = .053$). After accounting for these factors, there was still no evidence of benefit with ME induction (one-tailed $P = 1.00$), and exploratory analysis suggested the possibility of worse survival with ME (2-tailed $P = .0066$). In this multivariate analysis, the estimated relative risk of death in the ME arm, compared with the AD arm, was 1.44 (CI 1.11-1.87), slightly larger than the estimate of 1.32 from the stratified analysis described above. Thus, the poorer survival of the ME patients could not be attributed to imbalances in the significant prognostic factors. After accounting for the effects of the 4 significant factors listed above, none of the other factors available had a significant independent prognostic effect, including AML onset ($P = .34$), MRK16 expression ($P = .66$), or CsA-inhibited Di(OC)$_2$ or Rh123 efflux ($P = .51$ and .70, respectively).

Relapse-free survival by treatment arm

Of the 125 patients who achieved complete remission (56 with ME induction, 69 with AD induction), 100 have relapsed (47 ME, 53 AD) and 8 have died without report of relapse (4 in each arm; Figure 2). For patients randomized to receive ME induction who achieved CR, the median RFS was 7 months (CI 6-9 months) and the estimated probability of 2-year RFS was 16% (CI 6%-26%). For patients in the AD arm who achieved CR, the median RFS was 9 months (CI 7-12 months) and the 2-year RFS estimate was 18% (CI 9%-27%). Based on proportional hazards regression analysis adjusted for the stratification by age group, MDR-1 status, and onset of leukemia, RFS was not significantly better with ME induction therapy (one-tailed $P = .83$). The estimated hazard ratio ("relative risk") of relapse or death in the ME arm, compared with the AD arm, was 1.22 (CI 0.80-1.87). In exploratory analysis, RFS was not significantly worse with ME (2-tailed $P = .35$).

In univariate analyses without adjustment for the stratification (Table 5), patients with unfavorable cytogenetics had a somewhat shorter RFS survival compared with those with favorable or intermediate cytogenetics ($P = .042$). In addition, high CD34 expression on leukemic blasts was associated with shorter RFS ($P = .012$). In unstratified multivariate analyses that adjusted for the effect of CD34 expression, none of the other variables were significantly associated with RFS, including unfavorable cytogenetics ($P = .079$). In addition, there was no evidence of benefit from ME induction ($P = .71$ after adjusting for the effect of CD34 expression).
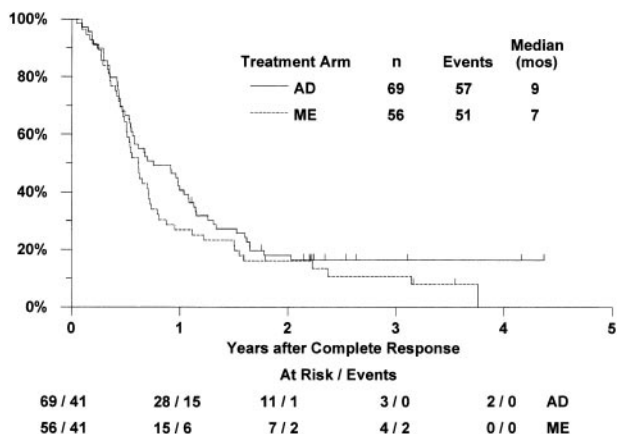

Figure 2. Estimated distributions of relapse-free survival, by induction treatment arm. Tickmarks indicate censored observations for patients last known to be alive without report of relapse.

Toxicity and hospitalization by treatment arm

There were 2 patients in the AD arm who were not evaluable for induction toxicity due to refusal of therapy (n = 1) and death within one day of starting treatment (n = 1). Among the 326 remaining patients, the incidence of fatal toxicity was slightly higher in the ME arm: 38 of 167 (23%, CI 17%-30%), compared with 28 of 159 (18%, CI 12%-24%) in the AD arm (one-tailed $P = .90$ by Fisher exact test). Most of these deaths were due to or involved infection (34 ME, 23 AD), with the remaining due to solid organ failure or hemorrhage. An additional 27 patients in each arm suffered a wide variety of grade 4 nonhematologic toxicities. Comparison of the incidence of grade 3 and higher toxicities between the 2 treatment arms identified only stomatitis as significantly different: 14% of patients in the ME arm, compared with 4% of patients in the AD arm (2-tailed $P = .0016$). There were no statistically significant differences in other gastrointestinal toxicities between the 2 treatment arms.

The estimated median times to neutrophil recovery were 33 days (CI 28-38 days) for the ME arm and 30 days (CI 26-32 days) for the AD arm. The estimated median times to platelet recovery were 33 days (CI 29-38 days) for the ME arm and 34 days (CI 28-38 days) for the AD arm. Time to hospital discharge was also determined and did not differ significantly between the 2 treatment arms with estimated medians of 30 days (CI 27-32 days) for the ME arm and 28 days (CI 27-30 days) for the AD arm.

Postremission therapy

Of the 125 patients who achieved complete remission, 18 (8 ME, 10 AD) were not registered for postremission therapy on study and one registered patient did not receive treatment due to investigator preference. The reasons for not registering were variable and included low platelet count (n = 3) and low ejection fraction (n = 3). There were no reports of the use of alternative consolidation, such as higher-dose cytarabine or transplantation. Among the 106 patients who received protocol postremission therapy, 99 received all planned consolidation therapy, that is, 2 cycles or until relapse. The other 7 (5 in the ME arm, 2 in the AD arm) did not receive the second cycle due to toxicity or other medical reasons (2 patients), refusal or other personal circumstances (4 patients), or investigator error (1 patient). There were 2 consolidation-related deaths, both due to infections in patients who received AD induction. There were 11 additional patients (5 ME, 6 AD) who had grade 4 nonhematologic consolidation toxicities, most commonly infections (5 patients), hypokalemia (2 patients), edema (3 patients), and cardiovascular toxicity (1 patient).

Discussion

AML is a devastating diagnosis in older patients because most patients are either not candidates for intensive therapy or because such therapy results in low complete remission rate, high induction toxicity, and short relapse-free and overall survivals. We conducted a phase 3, randomized clinical trial to assess whether the use of an induction regimen of mitoxantrone and etoposide was associated with improved outcome compared with a conventional induction regimen of cytarabine and daunorubicin. Unfortunately, the study was closed early after a scheduled interim analysis demonstrated that such an improved outcome was unlikely to be the case. Specifically, the hypothesized improvements in CR rate and OS that would warrant use of ME rather than AD induction were convincingly rejected. Furthermore, additional analyses exploring

the possibility of inferior outcomes with ME induction demonstrated a markedly lower survival rate ($P = .0066$) for patients in the ME arm, after correction for other significant factors (age, cytogenetics, performance status, and WBC count). While this evidence for possibly inferior outcomes with ME induction should be considered suggestive rather than definitive, since the study was not prospectively designed to address this possibility, the results certainly do not warrant routine use of ME induction chemotherapy over conventional AD induction.

The rationale for comparing ME to conventional induction therapy was based on the results reported in phase 2 studies of this regimen in patients over 60 years of age,[24] patients with MDS-related AML,[23] and patients with relapsed or refractory AML.[22] These 3 pilot studies each consisted of high-risk patients and the complete remission rates ranged from 43% to 57%.[22-24] The largest of these studies consisted of 67 patients with a median of 68 years of age (range, 60-80 years) who received the same ME regimen for initial treatment of AML followed by consolidation with cytarabine 500 mg/m^2 twice daily for 12 doses.[24] Many of the patient characteristics of this single-institution study appear to be relatively similar to our study, including 39% with secondary AML (as defined in our study), 64% with CD34$^+$ immunophenotype, and 48% with abnormal karyotype.[24] However, in contrast, there appears to be a greater occurrence of poor performance status in our study (24% with performance status 2-3) compared with their study (median of 80% Karnofsky performance status, with 95% confidence limits of 70%-90%). The CR rate of 34% (CI 26%-41%) for ME patients in our phase 3 study was markedly lower than the 55% (CI 43%-67%) reported in this phase 2 study. The results of these 2 studies should be compared with due caution considering the possible impacts of differences in performance status or numbers of participating institutions (ie, multicenter vs single center), and perhaps in definition and implementation of response criteria. Comparison of the phase 2 and phase 3 results clearly demonstrates the need to perform randomized, controlled trials before an investigational regimen is adopted in routine practice.

In choosing to study the ME regimen, we had hoped to find an improved CR rate and survival due to both improved antileukemic efficacy and reduced toxicity. However, the CR rate and overall survival for patients in the ME arm appeared to be lower and the incidence of RD higher than for patients in the AD arm. Because induction regimens containing cytarabine have shown no marked difference when combined with either daunorubicin or mitoxantrone,[18,33] our results suggest that the use of etoposide instead of cytarabine may have been the reason for the potentially worse outcome in the experimental treatment arm. While there was a somewhat higher proportion with unfavorable karyotype in the ME arm, this imbalance does not explain the inferior survival with ME

because even after correction for karyotype distribution in the multiple regression analysis, the inferior survival in the ME arm persists. In this study, we prospectively monitored toxicity, duration of neutropenia and thrombocytopenia after induction therapy, and duration of first hospitalization. There was no reduction in the incidence of fatal induction toxicity in the ME arm, and times to hematologic recovery and duration of hospitalization were similar in the 2 treatment arms. There was significantly more stomatitis with the ME regimen. In our study, patients on both arms were eligible to receive a second cycle of induction therapy if more than or equal to 5% blasts remained 14 or more days after the first cycle. This strategy appears to have benefited a significant minority of patients since about 25% of patients who received a second cycle did achieve a CR. However, whether or not this timing of 2 cycles of induction therapy is optimal needs to be addressed in future studies.

Once CR was achieved, RFS was similarly short for patients in both treatment arms. We had hypothesized that patients who achieved a CR with ME and consolidated with AD might have improved RFS by virtue of being exposed to 4 different active agents rather than just 2. Unfortunately, we found no evidence to support this hypothesis. The consolidation regimen we chose was somewhat less intensive than that used by some other groups.[24,28] We chose to use the same consolidation therapy in this study as used in the earlier SWOG study of patients over 55 years of age (SWOG-9031) in order to assess more accurately any impact of the change in induction strategy. Further, at the time this study was initiated, there were no definitive randomized trials demonstrating that more intensive consolidation was of benefit for older patients. Subsequent studies suggesting a benefit of more intensive cytarabine-containing regimens for patients with favorable cytogenetics, at least those less than 60 years of age,[4] should be considered in choosing therapy for selected patients 55 to 60 years of age. Clearly, improvements in the survival of older patients with AML will depend on improvements in both the induction and postinduction therapy administered.

In our study, the randomization was stratified according to MDR-1 expression status, secondary versus de novo onset of leukemia, and age, because the earlier SWOG study of patients over 55 years of age (SWOG-9031) had shown that these 3 factors had independent prognostic effects on outcomes (MDR-1 status and onset of leukemia were associated with CR rate, and age with survival).[10,16] Unfavorable cytogenetics was also a significant and independent prognostic factor (associated with CR rate, and overall and relapse-free survival) in SWOG-9031, but this factor could not be used for stratification because cytogenetic results were not uniformly available at the time of randomization. In addition, increasing WBC count was associated with shorter survival in SWOG-9031. Although the eligibility criteria and the participating institutions were essentially the same in these 2 SWOG studies,

Table 6. Results of the multivariate analyses of 328 patients in the current study compared with 211 patients in the prior SWOG study for similar patients

	Current study (SWOG-9333)	Prior study (SWOG-9031)*
Lower complete remission rate	Secondary AML ($P = .010$)	Secondary AML ($P = .0035$)
	Increasing age ($P = .0025$)	Unfavorable karyotype ($P = .0031$)
		MDR-1 + phenotype ($P = .0041$)
Higher resistant disease rate	CsA-inhibited Di(OC)$_2$ efflux ($P = .0012$)	MDR-1 + phenotype ($P = .0007$)
		Unfavorable karyotype ($P = .017$)
Shorter overall survival	Unfavorable karyotype ($P = .0001$)	Unfavorable karyotype ($P < .0001$)
	Increasing age ($P = .030$)	Increasing age ($P = .014$)
	Increasing WBC count ($P = .051$)	Increasing WBC count ($P = .029$)
	Poorer performance status ($P = .026$)	
Shorter relapse-free survival	High CD34 expression ($P = .012$)	Unfavorable karyotype ($P = .028$)

*See Leith et al[10] and Godwin et al[16] for details.

there were some differences in the results of the multivariate analyses of treatment outcomes (Table 6). In the current study, the factors most predictive of poor response were increasing age (lower CR rate and OS), secondary onset of AML (lower CR rate), unfavorable cytogenetics (lower OS), CsA-inhibited Di(OC)$_2$ efflux (higher RD rate), expression of CD34 (lower RFS), and performance status 2-3 and increasing WBC count (lower OS). One potential reason for the difference in outcome of these 2 sequential SWOG studies may be that the prognostic effect of performance status was not analyzed in the first study. A second reason may be that the proportion of patients with unfavorable cytogenetics was higher in the second study (43%) compared with the first (32%), predominantly due to a significant increase in the frequency of complex karyotypes (20% in the second study versus 3% in the first study, 2-tailed $P < .0001$). Except for the MDR-1 phenotype results, the predictive factors determined in our 2 SWOG studies have also been found to be significant in other large studies of older patients with AML.[13,34,35]

In conclusion, the outcome for patients over 55 years of age newly diagnosed with AML is poor, with few patients alive 3 years after diagnosis. The results of our study do not demonstrate any rationale for administering mitoxantrone and etoposide in place of cytarabine and daunorubicin as induction chemotherapy. In fact, there is a possibility that this regimen results in a shorter survival than conventional therapy. The most important factors that determine outcome for older patients with AML who are candidates for intensive induction therapy are age, leukemia onset, karyotype, and performance status. Further study of novel approaches is warranted.

References

1. Yin JA, Johnson PRE. Clinical annotation. Acute myeloid leukaemia in the elderly: biology and treatment. Br J Haematol. 1993;83:1-6.

2. Wiernik PH, Banks PLC, Case DC Jr, et al. Cytarabine plus idarubicin or daunorubicin as induction and consolidation therapy for previously untreated adult patients with acute myeloid leukemia. Blood. 1992;79:313-319.

3. Berman E, Heller G, Santorsa J, et al. Results of a randomized trial comparing idarubicin and cytosine arabinoside with daunorubicin and cytosine arabinoside in adult patients with newly diagnosed acute myelogenous leukemia. Blood. 1991;77:1666-1674.

4. Mayer RJ, David RB, Schiffer CA, et al. Intensive postremission chemotherapy in adults with acute myeloid leukemia. N Engl J Med. 1994;331:896-903.

5. 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.

6. 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.

7. Cassileth PA, Lynch E, Hines JD, et al. Varying intensity of postremission therapy in acute myeloid leukemia. Blood. 1992;79:1924-1930.

8. Cassileth PA, Harrington DP, Appelbaum FR, et al. Chemotherapy compared with autologous or allogeneic bone marrow transplantation in the management of acute myeloid leukemia in first remission. N Engl J Med. 1998;339:1649-1656.

9. Leith CP, Kopecky KJ, Chen I-M, et al. Frequency and clinical significance of the expression of the multidrug resistance proteins MDR-1/P-glycoprotein, MRP1, and LRP in acute myeloid leukemia. a Southwest Oncology Group study. Blood. 1999;94:1086-1099.

10. Leith CP, Kopecky KJ, Godwin J, et al. Acute myeloid leukemia in the elderly: assessment of multidrug resistance (MDR-1) and cytogenetics distinguishes biologic subgroups with remarkable distinct responses to standard chemotherapy. a Southwest Oncology Group study. Blood. 1997;89:3323-3329.

11. Baudard M, Marie JP, Cadiou M, Viguie F, Zittoun R. Acute myelogenous leukaemia in the elderly: retrospective study of 235 consecutive patients. Br J Haematol. 1994;86:82-91.

12. Leith CP, Kopecky KJ, Willman CL, et al. Acute myeloid leukemia in the elderly: high frequency of trilineage dyspoiesis suggests that elderly AML is biologically similar to AML arising secondary to myelodysplasia [abstract]. Blood. 1996;88:559a.

13. Latagliata R, Petti MC, Mandelli F. Acute myeloid leukemia in the elderly: 'per aspera ad astra'? Leukemia Research. 1999;23:603-613.

14. Hiddemann W, Kern W, Schoch C, et al. Management of acute myeloid leukemia in elderly patients. J Clin Onc. 1999;17:3569-3576.

15. Ryan DH, Head D, Shiaer SM, Hynes HE, Gumbart CH. Analysis of treatment failure in acute nonlymphocytic leukemia patients over fifty years of age: a Southwest Oncology Group study. Am J Clin Oncol. 1992;15:69-75.

16. Godwin JE, Kopecky KJ, Head DR, et al. A double-blind placebo-controlled trial of granulocyte colony-stimulating factor in elderly patients with previously untreated acute myeloid leukemia: a Southwest Oncology Group study. Blood. 1998;91:3607-3615.

17. Bennett JM, Young ML, Andersen JW, et al. Long-term survival in acute myeloid leukemia: the Eastern Cooperative Oncology Group experience. Cancer Supplement. 1997;80:2205-2209.

18. Arlin Z, Case DC, Moore J, et al. Randomized multicenter trial of cytosine arabinoside with mitoxantrone or daunorubicin in previously untreated adult patients with acute nonlymphocytic leukemia (ANLL). Leukemia. 1990;4:177-183.

19. Yin JAL, Johnson PRE, Daview JM, Flanagan NG, Gorst DW, Lewis MJ. Mitozantrone and cytosine arabinoside as first-line therapy in elderly patients with acute myeloid leukaemia. Br J Haematol. 1991;79:415-420.

20. Ho AD, Brado B, Haas R, Hunstein W. Etoposide in acute leukemia: past experience and future perspectives. Cancer. 1991;67:281-284.

21. Bishop JF, Lowenthal RM, Joshua D, et al. Etoposide in acute nonlymphocytic leukemia. Blood. 1990;75:27-32.

22. Ho AD, Lipp T, Ehninger G, et al. Combination of mitoxantrone and etoposide in refractory acute myelogenous leukemia—an active and well-tolerated regimen. J Clin Onc. 1988;6:213-217.

23. Knauf WU, Berdel WE, Ho AD, Kreuser E-D, Thiel E. Combination of mitoxantrone and etoposide in the treatment of myelodysplastic syndromes transformed into acute myeloid leukemia. Leukemia and Lymphoma. 1994;12:421-425.

24. Bow EJ, Sutherland JA, Kilpatrick MG, et al. Therapy of untreated acute myeloid leukemia in the elderly: remission-induction using a non-cytarabine-containing regimen of mitoxantrone plus etoposide. J Clin Onc. 1996;14:1345-1352.

25. Cheson BD, Cassileth PA, Head DR, et al. Report of the National Cancer Institute-sponsored workshop on definitions of diagnosis and response in acute myeloid leukemia. J Clin Onc. 1990;8:813-819.

26. Head DR, Lizardo C, Savage RA, et al. Institutional performance in application of the FAB classification of acute leukemia. Cancer. 1985;55:1979-1986.

27. Bennett JM, Catovsky D, Daniel MT, et al. Proposals for the classification of the myelodysplastic syndromes. Br J Haematol. 1982;51:189-199.

28. Rowe JM, Andersen JW, Mazza JJ, et al. A randomized placebo-controlled phase III study of granulocyte-macrophage colony-stimulating factor in adult patients (>55 to 70 years of age) with acute myelogenous leukemia: a study of the Eastern Cooperative Oncology Group (E1490). Blood. 1995;86:457-462.

29. Slovak ML, Kopecky KJ, Cassileth PA, Harrington KS, et al. Karyotypic analysis predicts outcome of preremission and postremission therapy in adult acute myeloid leukemia: a Southwest Oncology Group/Eastern Cooperative Oncology Group study. Blood. 2000;96:4075-4083.

30. Cox DR. Analysis of Binary Data. London, United Kingdom; Chapman & Hall, 1970.

31. Cox DR. Regression models and life tables. J R Stat Soc. 1972;34:187-202.

32. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958;53:457-481.

33. Rowe JM, Nauberg D, Friedenberg W, Bennett JM, et al. A phase III study of daunorubicin vs idarubicin vs mitoxantrone for older adult patients (> 55 years) with acute myelogenous leukemia (AML): a study of the Eastern Cooperative Oncology Group (E3993) [abstract]. Blood. 1998;92:313a.

34. Löwenberg B, Suciu S, Archimbaud E, et al. Mitoxantrone versus daunorubicin in induction-consolidation chemotherapy—the value of low-dose cytarabine for maintenance of remission, and an assessment of prognostic factors in acute myeloid leukemia in the elderly: final report of the Leukemia Cooperative Group of the European Organization for the Research and Treatment of Cancer and the Dutch-Belgian Hemato-Oncology Cooperative Hovon Group Randomized Phase III Study AML-9. J Clin Onc. 1998;16:872-881.

35. Johnson PRE, Hunt LP, Yin JAL. Prognostic factors in elderly patients with acute myeloid leukaemia: development of a model to predict survival. Br J Haematol. 1993;85:300-306.