Randomized Phase II Study of Fludarabine + Cytosine Arabinoside + Idarubicin ± All-Trans Retinoic Acid ± Granulocyte Colony-Stimulating Factor in Poor Prognosis Newly Diagnosed Acute Myeloid Leukemia and Myelodysplastic Syndrome

By Elihu H. Estey, Peter F. Thall, Sherry Pierce, Jorge Cortes, Miloslav Beran, Hagop Kantarjian, Michael J. Keating, Michael Andreeff, and Emil Freireich

Preclinical data suggest that retinoids, eg, all-trans retinoic acid (ATRA), lower concentrations of antiapoptotic proteins such as bcl-2, possibly thereby improving the outcome of acute myeloid leukemia (AML) chemotherapy. Granulocyte colony-stimulating factor (G-CSF) has been considered to be potentially synergistic with ATRA in this regard. Accordingly, we randomized 215 patients with newly diagnosed AML (153 patients) or high-risk myelodysplastic syndrome (MDS) (refractory anemia with excess blasts [RAEB] or RAEB-t, 62 patients) to receive fludarabine + ara-C + idarubicin (FAI) alone, FAI + ATRA, FAI + G-CSF, or FAI + ATRA + G-CSF. Eligibility required one of the following: age over 71 years, a history of abnormal blood counts before M.D. Anderson (MDA) presentation, secondary AML/MDS, failure to respond to one prior course of chemotherapy given outside MDA, or abnormal renal or hepatic function. For the two treatment arms containing ATRA, ATRA was given 2 days (day-2) before beginning and continued for 3 days after completion of FAI. For the two treatment arms including G-CSF, G-CSF began on day-1 and continued until neutrophil recovery. Patients with white blood cell (WBC) counts >50,000/µL began ATRA on day 1 and G-CSF on day 2. Events (death, failure to achieve complete remission [CR], or relapse from CR) have occurred in 77% of the 215 patients. Reflecting the poor prognosis of the patients entered, the CR rate was only 51%, median event-free survival (EFS) time once in CR was 36 weeks, and median survival time was 28 weeks. A Cox regression analysis indicated that, after accounting for patient prognostic variables, none of the three adjuvant treatment combinations (FAI + ATRA, FAI + G, FAI + ATRA + G) affected survival, EFS, or EFS once in CR compared with FAI. Similarly, there were no significant effects of either ATRA ignoring G-CSF, or of G-CSF ignoring ATRA. As previously found, a diagnosis of RAEB or RAEB-t rather than AML was insignificant. There were no indications that the effect of ATRA differed according to cytogenetic group, diagnosis (AML or MDS), or treatment schedule. Logistic regression analysis indicated that, after accounting for prognosis, addition of G-CSF ± ATRA to FAI improved CR rate versus either FAI or FAI + ATRA, but G-CSF had no effect on the other outcomes. We conclude that addition of ATRA ± G-CSF to FAI had no effect on CR rate, survival, EFS, or EFS in CR in poor prognosis, newly diagnosed AML or high-risk MDS.

© 1999 by The American Society of Hematology.

R ECENT PUBLICATIONS have reported associations between high levels of the protein bcl-2 and unfavorable outcome of therapy of acute myeloid leukemia (AML).1 2 These findings have marked bcl-2 and related proteins, which, like bcl-2, decrease the rate of cell death (apoptosis) as potential targets of anti-AML treatment.3 4 In particular, some have speculated that decreases in the concentrations of such antiapoptotic proteins could increase the effectiveness of chemotherapy. In the late 1980s and early 1990s, data from McCulloch’s laboratory in Toronto indicated that addition of retinoids, eg, all-trans retinoic acid (ATRA), to cultures of AML cells containing ara-C or daunorubicin increased killing of those clonogenic cells thought responsible for perpetuation of the disease.5 6 Subsequent data from the same lab suggested that this effect was mediated by downregulation of bcl-2,7 and that granulocyte colony-stimulating factor (G-CSF) and retinoids might be synergistic in this regard.8 9

With such in vitro data in mind and knowing that ATRA added to chemotherapy decreases recurrence rate in acute promyelocytic leukemia (APL),10 11 admittedly a different disease, we conducted a randomized study designed to assess whether addition of ATRA and/or G-CSF to chemotherapy might improve complete remission (CR) or survival rates in AML and high-risk myelodysplastic syndromes (MDS). We have found such syndromes (refractory anemia with excess blasts [RAEB] or RAEB-t) identical to AML with regard to these outcomes.12 As described below, we limited eligibility to newly diagnosed patients having a particularly poor prognosis. As of July 1998, failures have occurred in over three quarters of the 215 patients treated. This has led us to report our current results.

MATERIALS AND METHODS

Patients with AML (but not APL), RAEB-t, or RAEB were eligible provided any of the following criteria were present: (1) age greater than 71 years, (2) an antecedent hematologic disorder (AHD) defined as a history of an abnormal blood count (hemoglobin <12 g/dL, or neutrophils <1,500/µL, or white blood cells (WBC) >10,000/µL or <4,000/µL, or platelet count <150,000/µL) documented to be present for at least 1 month before M.D. Anderson presentation, (3) AML or MDS arising after chemotherapy for another cancer (secondary AML/ MDS), (4) failure to respond to one course of induction therapy containing ara-C + anthracycline and delivered at another hospital, (5) serum bilirubin greater than 2.9 mg/mL, or (6) serum creatinine >1.5
mg/mL. We conducted the study between September 28, 1995 and November 12, 1997. During this time, we saw 355 patients with newly diagnosed AML (excluding APL), RAEB-t, or RAEB of whom 258 (73%) qualified for the study. A total of 215 of the 258 (83%) were entered. All but four of the 215 were eligible, and we include all 215 in this report. The principal reason that 43 eligible patients did not go on this study was the existence of a competing protocol for topotecan + ara-C in patients with RAEB or RAEB-t. A total of 32 of the 43 patients went on this study. The remaining 11 patients were given standard regimens by patient choice. There were no significant differences in survival or event-free survival (EFS), CR rate, or event-free survival once in CR between the 215 treated on the study and the 43 who were eligible, but not treated on the study (P > .05 for all tests).

We randomly assigned the 215 patients to receive (1) chemotherapy (fludarabine + ara-C + idarubicin = FAI), (2) FAI + G-CSF, (3) FAI + ATRA, or (4) FAI + ATRA + G-CSF. Details of the treatments follow. Patients were stratified at enrollment according to cytogenetics and WBC count and randomized using a dynamic allocation scheme (see Statistical Methods) to avoid imbalances between treatment groups with regard to these variables. At MDA, cytogenetic results are available within 3 working days. Patients whom attending physicians felt required treatment before cytogenetic results were in hand were included in this report. The success rate of 36%. The dynamic allocation scheme of Pocock and Simon13 was used to achieve balance between the treatment groups with regard to cytogenetics (group 1, inv(16) or t(8;21); group 2, -5, -7, +8, or 11q-; group 3, other) and WBC count (group 1, WBC ≥ 10,000/µL; group 2, 10,000/µL < WBC < 50,000/µL; group 3, WBC ≥ 50,000/µL). The balanced allocation of patients by WBC count before randomization was designed to address the possibility that differences between the ATRA arm and the ATRA + G-CSF arm might result only from a difference in ATRA schedule, which differed according to WBC count as noted above, between the arms.

As in the past, we treated patients above age 50 years in laminar air flow rooms, a protected environment (PE) whenever such rooms were available. Patients routinely received trimethoprim/sulfamethoxazole and fluconazole by mouth to prevent infection. It should be noted that fluconazole has been reported to increase plasma concentrations of ATRA.13 However, our patients did not routinely receive other medications reported to affect these concentrations.13 Patients who had persistent disease (>20% blasts in a marrow that was at least 20% cellular in AML or RAEB-t, >5% blasts in a similarly cellular marrow in RAEB) 14 and 21 days after the start of chemotherapy without improvement between these dates received a second course identical to the first. The same criteria, in two consecutive marrow samples, were used for starting a second course in patients whose marrow had decreased blasts or was less than 20% cellular on days 14 or 21, but in whom disease reappeared. Day 14 and 21 marrows included a biopsy, as well as an aspirate; thereafter only weekly aspirates were performed until CR status was established. CR was defined as a marrow sample showing <5% blasts, a platelet count > 100,000/µL, and a neutrophil count > 1,000/µL. Patients not in CR after two courses of treatment were removed from the study and offered other therapies. Once in CR, patients alternated courses of ara-C 100 mg/m² daily × 5 days by continuous infusion with courses of fludarabine 30 mg/m² daily days 1 to 2, ara-C 1 g/m² daily days 1 to 2, and idarubicin 8 mg/m² day 3. Patients assigned to receive G-CSF and/or ATRA during induction received the same treatment during and for the first 3 days after the completion of each postremission course. The doses were those used during induction. Therapy continued until 6 months had elapsed from CR date, by which time most patients had received four to five courses of postremission therapy. Relapse was defined as a marrow with >5% blasts unrelated to recovery of blood counts from the preceding course of chemotherapy. At relapse, patients received salvage therapies as previously described. Although the intent was to transplant only at relapse, two patients (both with AML) received an allogeneic transplant in first CR. These two constituted 2% of the 110 patients achieving CR.

We did not censor these patients at time of transplant because we could not be sure that the physician’s decision to transplant and the chances of relapse or death in CR were independent of each other. Barring such independence, censoring is invalid.

Statistical Methods

T trial design. The planned study sample size of 212 patients, equivalently 53 patients per arm, was determined to detect a .20 difference in the probability of success (alive and in CR at 6 months from start of treatment) between the baseline treatment group (neither ATRA nor G-CSF and each of the three other treatment groups based on two-sided tests having individual power .80 and overall type I error rate .05). This computation was based on a logistic regression model including patient covariates fit to historical data from similar studies to estimate relevant quantities, which in particular yielded a baseline success rate of 36%. The dynamic allocation scheme of Pocock and Simon13 was used to achieve balance between the treatment groups with regard to cytogenetics (group 1, inv(16) or t(8;21); group 2, -5, -7, +8, or 11q-; group 3, other) and WBC count (group 1, WBC ≥ 10,000/µL; group 2, 10,000/µL < WBC < 50,000/µL; group 3, WBC ≥ 50,000/µL). The balanced allocation of patients by WBC count before randomization was designed to address the possibility that differences between the ATRA arm and the ATRA + G-CSF arm might result only from a difference in ATRA schedule, which differed according to WBC count as noted above, between the arms.

Data analysis. Associations between patient characteristics (covariates) were assessed graphically for pairs of numerical variables by examining scatterplots, by Wilcoxon-Mann-Whitney and Kruskal- Wallis test statistics17 for categorical and continuous variables, and by the Fisher exact test18 and its generalizations17 for pairs of categorical variables. We examined the following covariates as assessed before treatment: age, hemoglobin, WBC count, platelet count, performance status (Zubrod 0-2 vs 3-4), treatment in the PE, treatment as an emergency, receipt of one prior course of induction therapy given outside MDA, cytogenetics [normal karyotype, including patients with insufficient metaphases for analysis (see below) v inv(16) or t(8;21) v −5, 5q−, -7 or 7q− (−5i/7i) v other abnormalities], and presence of an AHD. Logistic regression was used to assess the ability of treatment arm or the patients’ characteristics to predict the probability of CR, with goodness-of-fit assessed by residual and partial residual scatter plots and likelihood ratio (LR) statistics. Unadjusted survival and EFS analyses were performed using Kaplan-Meier plots.15 Events were defined as death or relapse, and, in a separate analysis, as death, relapse, or failure to achieve CR due to resistant disease (scored at the time the patient was removed from study). Unadjusted comparisons of survival and EFS between patient subgroups were made using the logrank test.19 The Cox proportional hazards model20 and its generalizations21 were used to assess the ability of treatment groups and patient characteristics to predict survival and EFS, with goodness-of-fit assessed by the Grambsch-Therneau test,22 Schoenfeld residual plots, martingale residual plots,21 and LR statistics. All scatterplots were smoothed using the loess method of Cleveland,23 with variables transformed as appropriate based on these plots. Multivariate logistic and Cox models were obtained by performing a backward elimination with P value cutoff .05, then
allowing treatment-covariate interactions or any variable previously deleted to reenter the final model if its P value was < .05. All computations were performed on a DEC Alpha 2100 5/250 system computer (Digital Electronics Corporation, Nashua, NH) in StatXact (Cytel Software Corporation, Cambridge, MA) and Splus,24 using both standard Splus functions and the survival analysis package of Therneau.25

RESULTS

In terms of the eligibility criteria described above, 32% of the 215 patients were over age 71 years, with a median of 65 years. Two thirds had an AHD, with a median duration of 7 months. A total of 26% had secondary AML/MDS, and 11% had failed to respond to one prior induction course given outside MDA. A total of 12% had a creatinine above 1.5 mg/mL, and 3% had a bilirubin greater than 2.9 mg/mL. Cytogenetic abnormalities were defined using standard ISCN guidelines. Reflecting the demographic features noted above, only 2% of the patients had inv(16) or (8;21). In contrast, 31% had monosomy, or loss of the long arm of, chromosomes 5 and/or 7 (“-5/-7”), with the number of such patients (66) close to the number with a normal karyotype (73). Note that for purposes of subsequent analyses, we combined patients with a normal karyotype with patients with insufficient metaphases for analysis. There were 5, 1, 1, and 3 patients with insufficient metaphase in the FAI, FAI + G, FAI + ATRA, and FAI + ATRA + G arms, respectively, constituting 9%, 2%, 2%, and 6% of the patients in these arms. We have combined the normal karyotype and insufficient metaphase groups because our previous experience suggests that the prognosis of the latter group most resembles that of patients with a normal karyotype.36 Only 12% of the patients had Zubrod performance status 3 or 4, but 16% were treated as responders to one prior induction course given outside MDA. A total of 65% of the patients were admitted to the PE. Table 1 compares the distribution of these and other pretreatment characteristics in the four different groups, showing that there were no significant imbalances for any of these covariates, with the exception that patients given FAI + ATRA were less likely to have secondary AML/MDS.

In view of the patients’ pretreatment characteristics, overall results were poor. The CR rate was 51% (110 of 215) with a median actuarial EFS duration of 36 weeks (95% confidence interval [CI], 27 to 47 weeks) from CR date. Events have occurred in 166 of the 215 patients (77%), with 105 patients failing to achieve CR, 55 having disease recurrence, and six dying in CR. Recalling that the study aimed for a 6-month EFS rate of .56, the probability of EFS at 6 months is currently .39 (95% CI, .33 to .47). Figure 1 illustrates that the actuarial median survival time is currently only 28 weeks (95% CI, 21 to 37 weeks) and that approximately 85% of the patients are projected to be dead within 2 years. Median follow-up time in 75 of the 215 patients who are currently alive is 35 weeks (range, up to 108 weeks).

Within this context, there was some initial evidence for a beneficial effect of G-CSF and, particularly, ATRA. The CR rates in the four arms were: FAI, 21 of 53, 40%; FAI + G-CSF (“FAI + G”), 29 of 53, 55%; FAI + ATRA, 28 of 55, 51%; and FAI + ATRA + G, 32 of 54, 59%, with the P value equal to .087 for the 2×2 comparison of treatment with and without G-CSF. More interestingly, log rank tests comparing survival with FAI to survival with each of the other three treatment arms gave the following P values: FAI versus FAI + G P = .15, FAI versus FAI + ATRA P = .023, and FAI versus FAI + ATRA + G P = .055, thus suggesting a beneficial effect of ATRA on survival without any added benefit from G-CSF. Figure 2 depicts Kaplan-Meier plots for the four treatment groups. It should be noted that only one of 36 patients who survived, but failed to achieve CR on one of the four study arms, achieved a subsequent CR. Hence, the survival data are largely a function of response to initial treatment. Intertreatment comparisons of EFS from start of therapy gave analogous results: FAI versus FAI + G P = .32, versus FAI + ATRA P = .053, versus FAI + G + ATRA P = .095. In the best arm, FAI + ATRA, the current probability of EFS at 6 months is .52 (95% CI, .40 to .67). There was no significant difference between FAI and any of the other three arms with respect to EFS once in CR, with P values .95, .43, and .48 for tests of FAI versus FAI + G, FAI + ATRA, and FAI + ATRA + G, respectively. Given this and given that there appeared to be no effect of ATRA on CR rate, any survival advantage due to ATRA appeared to result from an effect in patients not achieving CR.

Because, in general, there were no statistically significant differences between the four treatment arms in the distribution of the prognostic factors illustrated in Table 1, it appeared plausible that the apparent beneficial effect of ATRA and/or G-CSF might remain after accounting for these factors in Cox and logistic regressions. Table 2 summarizes the multivariate Cox model for survival. After accounting for the covariates indicated in the Table 2, there was no evidence that any of the three adjuvant treatment combinations FAI + G, FAI + ATRA, or FAI + ATRA + G prolonged survival compared with FAI. Additional tests evaluating the effects of G-CSF ignoring ATRA (ie, G-CSF ± ATRA), and of ATRA ignoring G-CSF (ATRA ± G-CSF) were also insignificant, with all P values ≥ .28. Several

| Table 1. Pretreatment Characteristics of Patients on the Four Treatment Arms |
|---------------------------------|-------|-------|-------|-------|-------|
| No. of Patients | FAI (%) | FAI + G (%) | FAI + ATRA (%) | FAI + ATRA + G (%) | P Value |
| Median age in years | | | | | |
| Platelet count (<10^11/µL) | 36 | 36 | 49 | 58 | .28 |
| Placed in PE | 36 (60) | 33 (70) | 33 (60) | 33 (60) | .64 |
| Zbudro PS 3, 4 | 6 (12) | 7 (13) | 4 (7) | 8 (15) | .63 |
| Emergency | 9 (17) | 8 (15) | 9 (16) | 8 (15) | .99 |
| Inv(16)/8;21 | 2 (4) | 2 (4) | 0 | 1 (2) | .43 |
| --5/-7* | 17 (32) | 17 (32) | 15 (27) | 17 (31) | .94 |
| Other abnormal karyotype | 14 (26) | 17 (32) | 18 (33) | 13 (24) | .71 |
| Normal karyotype | 20 (38) | 17 (32) | 22 (40) | 23 (40) | .43 |
| RAEB or RAEB-t | 15 (28) | 16 (30) | 16 (29) | 15 (28) | 1.0 |
| AHD | 38 (72) | 37 (70) | 39 (71) | 31 (57) | |
| Secondary AML, RAEB-t or RAEB | 12 (23) | 17 (32) | 7 (13) | 18 (33) | .04 |
| Failed 1 previous course given outside MDA | 10 (19) | 3 (6) | 4 (7) | 7 (13) | .12 |

*Includes patients with --5, 5q-, --7, and/or 7q-.
points should be noted about the model summarized in Table 2. In particular, patients regarded as emergencies did worse than other patients even after accounting for performance status, which did, and white cell count, which did not, enter the model in Table 2. In contrast, survival was unaffected by whether patients had failed to respond to a previous course of outside treatment (P = .85) or, as previously, whether they had RAEB or RAEB-t rather than AML (P = .54). Although it appeared that whether patients had failed one prior course of outside treatment or, rather, were previously untreated was prognostically insignificant, we repeated the analysis excluding the previously treated patients. Again there was no suggestion that ATRA ± G-CSF or G-CSF ± ATRA affected survival. The presence of secondary AML or MDS, which was least frequent in patients given FAI + ATRA (Table 1) had no affect on survival (P = .37). Finally, there was no evidence for an interaction between ATRA ± G-CSF and cytogenetics (P = .52 for −5/−7, P = .33 for other poor prognosis cytogenetics), diagnosis of RAEB or RAEB-t rather than AML (P = .49) or WBC (P = .31). That is, there was nothing to suggest that the effect of ATRA was different in patients with −5/−7, other cytogenetic abnormalities, a high WBC count, or MDS compared with other patients. The failure to find an interaction between ATRA and WBC count (considered either continuously or dichotomized above and below 10,000/µL) indicates that ATRA schedule had no effect on survival. The multivariate model for EFS from start of treatment was similar to the survival model shown in Table 2 regardless of (1) whether events were considered as death or relapse or, rather, as death, relapse, or failure to achieve CR due to resistant disease, and (2) whether the previously treated patients were excluded. In particular, there was no effect of ATRA +/- G-CSF, or G-CSF +/- ATRA, and there was no interaction between G-CSF and ATRA.

The beneficial effect of G-CSF +/- ATRA on CR rate suggested by the preliminary analysis noted two paragraphs.
above was confirmed by a logistic regression analysis (Table 3). The P value testing the effect of G-CSF \( \pm \) ATRA versus no G-CSF was .018. ATRA +/− G-CSF was not predictive of CR \( (P = .264) \). Again, conclusions were unaffected by exclusion of previously treated patients. Time required from start of treatment to reach a neutrophil count \( > 1,000/\mu\text{L} \) was significantly less in the two arms with G-CSF \( (P < .001, \text{medians of} 24 + 29 \text{ days, considering patients requiring one course to achieve CR, such patients constituting 94% of the CRs}) \). ATRA had no effect on this outcome, and neither G-CSF nor ATRA affected time to reach a platelet count \( > 100,000/\mu\text{L} \). Perhaps the most quantifi-
able measure of a regimen’s toxicity is the early death rate associated with the regimen. There were nine deaths in the first 2 weeks of therapy among 109 patients receiving FAI + ATRA or FAI + ATRA + G versus a rate of 3 of 106 among patients receiving FAI or FAI + G. The P value for this comparison is .084 and must be interpreted in light of data indicating that the difference in death rate between the arms with and without ATRA was less considering deaths occurring in week 1 (5 of 109 v 3 of 106), weeks 1 to 3 (16 of 109 v 10 of 106), or weeks 1 to 4 (22 of 109 v 18 of 106). These data suggest that there was no difference in the early death rate between the arms with and without ATRA. The early death rate was essentially equivalent on FAI + ATRA and FAI + ATRA + G and was similarly equivalent on FAI and FAI + G. The rate of major infection (pneumonia or documented fungal infection) was similar on all four arms, with these infections occurring in 38% of all patients. Other toxicity was infrequent and again similar on all arms. ATRA syndrome was not observed. Nor did the addition of ATRA result in cutaneous or hepatic toxicities or in leukocytosis. This may reflect the short duration of ATRA exposure, and in the case of leukocytosis, the administration of chemotherapy 2 days after starting ATRA in patients presenting with WBC counts \( < 10,000/\mu\text{L} \), and simultaneously with chemotherapy in patients presenting with higher WBC counts and during postre-
therapy mission. Finally, Cox regression confirmed that all four treatments were statistically equivalent with respect to EFS once in CR. \( P \) values for treatments, each compared with FAI, were as follows: ATRA +/− G-CSF, .18; G-CSF +/− ATRA, .40; ATRA without G-CSF, .59; G-CSF without ATRA, .99; and ATRA + G-CSF, .37. Again, results were substantively the same regardless of whether the previously treated patients were excluded.

### DISCUSSION

The principal conclusion of this study is that the addition of ATRA +/− G-CSF to FAI did not affect survival (Table 2), EFS either from start of treatment or from CR date, or CR rate (Table 3) in poor prognosis newly diagnosed AML, RAEB, or RAEB-t. Although addition of G-CSF +/− ATRA to the same chemotherapy improved the CR rate compared with FAI + ATRA without G-CSF \( (P = .018 \text{ on multivariate analysis}) \), it had no effect on the other measures of outcome.

It is a truism that the inferences drawn from a study depend on the methods used to analyze the data. This is certainly the case here. In particular, before accounting for other relevant covariates, evidence suggested that ATRA +/− G-CSF improved survival \( (P < .05, \text{and Fig 2}) \). Yet, despite the balanced randomization and the seemingly even distribution of covariates in the four treatment arms (Table 1, all \( P \) values > .1 with the exception of secondary AML/MDS, which was not prognosti-
cally significant), the conclusion of the study was otherwise. The reason for this is that data summaries such as Table 1 are only one-dimensional representations of a multidimensional phenomenon. The point is that, unless a randomized study is very large, perfect covariate balance between the treatment arms is rarely achieved and cannot be assumed, even if \( P \) values are > .1 for the test of the hypothesis that there is balance with respect to covariates examined individually. Thus, our study illustrates the need to account for non-treatment-related covari-
ates via multivariate analysis in randomized studies.

Our results with G-CSF are reminiscent of those of Dombret et al., who also found in a group of elderly patients with a high prevalence of unfavorable cytogenetic abnormalities, that G-CSF improved CR rate without affecting early mortality. Although this might suggest that G-CSF increases sensitivity to chemother-
apy, this suggestion is difficult to reconcile with the failure of G-CSF to prolong disease-free survival in either study. One explanation for this apparent paradox is that G-CSF increases neutrophil numbers in the marrow, thereby lowering blast

### Table 2. Survival Model

<table>
<thead>
<tr>
<th>Factor</th>
<th>Effect on Survival</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Worse with increasing age</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Platelets</td>
<td>Survival improves as pretreatment platelet count increases from less than 5,000 to approximately 140,000, then worsens with further increase in platelets</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Treated in PE</td>
<td>Favorable</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>PS 3, 4</td>
<td>Unfavorable</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Treated as emergency</td>
<td></td>
<td>.030</td>
</tr>
<tr>
<td>−5/−7†</td>
<td>Unfavorable</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Other poor cytogentic‡</td>
<td></td>
<td>.050</td>
</tr>
</tbody>
</table>

*Versus a normal karyotype or insufficient material for cytogenetic analysis.
†−5/−7 refers to patients with clonal abnormalities involving −5, −7, del 5q, or del 7q.
‡Other poor cytogenetics includes all cytogenetic abnormalities associated with AML/MDS, excluding inv(16), t(8;21), and −5/−7.

### Table 3. CR Model

<table>
<thead>
<tr>
<th>Factor*</th>
<th>Effect on Survival</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Unfavorable</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Platelets</td>
<td>CR rate improves as pretreatment platelet count increases from less than 5,000 to approximately 210,000, then worsens with further increase in platelets</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>PE</td>
<td>Favorable</td>
<td>.002</td>
</tr>
<tr>
<td>PS 3, 4</td>
<td>Unfavorable</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>−5/−7</td>
<td>Unfavorable</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Other poor cytogentic‡</td>
<td></td>
<td>.056</td>
</tr>
<tr>
<td>G-CSF + ATRA</td>
<td>Favorable</td>
<td>.018</td>
</tr>
</tbody>
</table>

*See Table 2 for details.
†Nonsignificant factors include RAEB or RAEB−t (rather than AML) \( (P = .68) \), one previous course of outside therapy \( (P = .58) \), ATRA + G-CSF \( (P = .35) \).
percent, but not the absolute number of blasts. Thus, some of the
remissions seen in the G-CSF cohort are cosmetic only.
Furthermore, there are studies in elderly patients in which
G-CSF was given only after completion of chemotherapy28-29
(as in the Dombret et al study) or in which G-CSF was
given before and during chemotherapy30 (as in the current
study) that found no effect on CR rate, although in two of
these28-29 there was some reduction in days in hospital or
antibiotic use. Our view is that, given the failure to show
beneficial effects on disease-free survival or survival, a suffi-
cient number of trials devoted to G-CSF administration during
induction therapy have been completed, and that patient
resources might be more profitably invested in clinical trials
examining other issues.

For the reasons given at the beginning of this report, the most
interesting part of this trial to us was the use of ATRA in
non-APL AML and RAEB/RAEB-T. Because we are unaware of
any other trials combining similarly myelosuppressive chem-
otherapy with ATRA in these diseases, we are particularly eager
to avoid a falsely negative conclusion about ATRA. As dis-
cussed in Materials and Methods, we believe that we treated
enough patients to detect a medically beneficial effect had one
existed. When we presented results of this study at the 1997
American Society of Hematology (ASH) meeting,31 multivari-
ate analysis indicated that ATRA +/- G-CSF improved sur-
vice and disease-free survival from start of treatment. Between
then and the writing of this report, another 24 events have
occurred, representing a 17% increase. It could be argued the
occurrence of still more events might alter our conclusions.
However, events have occurred in 77% of patients, and there
are, perhaps unfortunately, no conventions as to when to report
results. Lastly, the population we treated was dominated by poor
prognostic features and thus not completely representative of
AML, the chemotherapy we gave was not “standard”, and
other schedules of ATRA could have been tested, eg, one in
which ATRA began only after completion of chemotherapy.
Thus, we emphasize that our results should not be generalized
to AML patients having better prognostic features than those of
the patients in our trial. Given the results of our study, it seems
advisable that subsequent studies of ATRA be designed with
careful consideration of these issues.

REFERENCES
Archimbaud E, Maguad JP, Guyotat D: High expression of bcl-2 protein
in acute myeloid leukemia cells is associated with poor response to
chemotherapy. Blood 81:3091, 1993
2. Maung ZT, MacLean FR, Reid MM, Pearson ADI, Proctor SJ,
Hamilton PJ, Hall AG: The relationship between bcl-2 expression and
response to chemotherapy in acute leukemia. Br J Haematol 88:105,
1994
3. Minn AJ, Rudin CM, Boise LH, Thompson CB: Expression of
bcl-x1 can confer a multidrug resistance phenotype. Blood 86:1903,
1995
4. Banker DE, Groudine M, Norwood T, Appelbaum FR: Measure-
ment of spontaneous and therapeutic agent-induced apoptosis with
bcl-2 protein expression in acute myeloid leukemia. Blood 89:243,
1997
5. Lishner M, Curtis JE, Minkin S, McCulloch EA: Interaction
between retinoic acid and cytosine arabinoside affecting the blast cells
of acute myeloblastic leukemia. Leukemia 3:784, 1989
6. Yang GS, Minden MD, McCulloch EA: Regulation by retinoic
acid and hydrocortisone of the anthracycline sensitivity of blast cells
of acute myeloblastic leukemia. Leukemia 8:2065, 1994
7. Hu Z-B, Minden MD, McCulloch EA: Direct evidence for the
participation of bcl-2 in the regulation by retinoic acid of the ara-C
sensitivity of leukemic stem cells. Leukemia 9:1667, 1995
8. Hu Z-B, Minden MD, McCulloch EA: Regulation of the synthesis
of bcl-2 in acute myeloblastic leukemia: Significance for response
to chemotherapy. Leukemia 10:410, 1996
H, Lowenberg B, Fey M, Archimbaud E, Degos L: Treatment of
newly-diagnosed acute promyelocytic leukemia by all trans retinoic
acid combined with chemotherapy: The European experience. Leuk
Lymphoma 16:431, 1995
11. Tallman MS, Andersen JW, Schiffer CA, Appelbaum FR, Feus-
ner JH, Ogden A, Shepherd L, Wilman C, Bloomfield CD, Rowe J,
Wiernik PH: All-trans-retinoic acid in acute promyelocytic leukemia. N
12. Estey E, Thall P, Beran M, Kantarjian H, Pierce S, Keating M:
Effect of diagnosis (refractory anemia with excess blasts, refractory
anemia with excess blasts in transformation, or acute myeloid leukemia)
14. Pocock SJ, Simon J: Sequential treatment assignment with
balancing for prognostic factors in the controlled clinical trial. Biomet-
ricks 31:103, 1975
15. Snedecor GW, Cochran WG: Statistical Methods, (ed 7). Ames,
IA, Iowa State University Press, 1980
16. Fisher RA: The condition under which P2 measures the discrep-
ancy between observation and hypothesis. J R Stat Soc 87:442, 1924
17. Mehta CR: The exact analysis of contingency tables in medical
18. Kaplan EL, Meier P: Nonparametric estiamtor from incomplete
19. Mantel N: Evaluation of survival data and two new rank order
statistics arising in its consideration. Cancer Chemother Rep 60:163,
1966
20. Cox DR: Regression models and life tables (with discussion). J R
22. Grambsch PM, Therneau TM: Proportional hazards tests and
diagnostics based on weighted residuals. Biometrika 81:515, 1994
23. Cleveland WS: Robust locally-weighted regression and smooth-
Pacific Grove, CA, Wadsworth, 1988
25. Therneau TM: A package for survival analysis in S. Rochester,
MN, Mayo Clinic Foundation, 1994
26. Ravandi-Kashani F, Hayes K, Lovsde D, Pierce S, Estey E:
Prognostic significance of an unsuccessful cytogenetic analysis in
newly-diagnosed AML. Blood 86:776a, 1995 (suppl 1, abstr)
27. Dombret H, Chastang C, Fenaux p, Reiffers J, Bordessoule D,
Bouabdallah R, Mandelli F, Ferrant A, Auzanneau G, Tilly H, Yver A,
Degos L: A controlled study of recombinant human granulocyte
colony-stimulating factor in elderly patients after treatment for acute
Noens L, Szer J, Ganser A, O'Brien C, Matcham J, Barge A: A
randomized, double-blind, placebo controlled, phase III study of
filgrastim in remission induction and consolidation therapy for adults


Randomized Phase II Study of Fludarabine + Cytosine Arabinoside + Idarubicin ± All-Trans Retinoic Acid ± Granulocyte Colony-Stimulating Factor in Poor Prognosis Newly Diagnosed Acute Myeloid Leukemia and Myelodysplastic Syndrome

Elihu H. Estey, Peter F. Thall, Sherry Pierce, Jorge Cortes, Miloslav Beran, Hagop Kantarjian, Michael J. Keating, Michael Andreeff and Emil Freireich