We assessed by multiparametric flow cytometry the levels of minimal residual disease (MRD) in 100 adult patients with acute myelogenous leukemia (AML) achieving complete remission (CR) after intensive chemotherapy. The aim of the study was to determine the optimal threshold, in terms of residual leukemic cells, and the time point of choice, that is, post-induction (post-Ind) or post-consolidation (post-Cons), able to better predict outcome. By applying the maximally selected log-rank statistics, the threshold discriminating MRD from MRD cases was set at $3.5 \times 10^{-4}$ residual leukemic cells, a level that allowed the identification of distinct subgroups of patients, both at post-Ind and post-Cons time points. Post-Cons MRD patients had a superior outcome in terms of relapse rate, overall survival (OS) and relapse-free survival (RFS) ($P<0.001$, for all comparisons), regardless of the MRD status after induction. In particular, patients entering MRD negativity only after consolidation showed the same outcome as those achieving early negativity after induction. Multivariate analysis, including karyotype, age, MDR1 phenotype, post-Ind and post-Cons MRD levels, indicated that the post-Cons MRD status independently affected the outcome ($P<0.001$, for all comparisons). In conclusion: (1) the threshold of $3.5 \times 10^{-4}$ is valid in discriminating risk categories in adult AML and (2) post-Cons MRD assessment is critical to predict disease outcome.

Leukemia advance online publication, 13 July 2006;
doi:10.1038/sj.leu.2404313

Keywords: AML; MRD; multiparametric flow cytometry; leukemia-associated phenotype

Patients and methods

One hundred and thirty-five consecutive adult patients diagnosed with de novo AML (excluding M3) at the Department of Hematology, S Eugenio Hospital, University Tor Vergata, Rome were enrolled in the EORTC/GIMEMA protocols AML-10, AML-12 and AML-13. A "leukemia-associated" immunophenotype was detected by multiparametric flow cytometry in 120 (89%) cases at presentation. Of these, 100 (83%) achieved morphologic CR after induction and were therefore followed up for MRD monitoring. The present analysis includes 56 patients whose results have been already reported in a previous publication. Approval for this study was obtained from the Institutional Review Board. Informed consent was provided according to the Declaration of Helsinki Principles. The EORTC/GIMEMA AML-10 randomized trial included patients aged 18–60 years (Vignetti et al., Blood 2003; 102 (Suppl 1): 611a; abstract). Induction treatment combined cytarabine (100 mg/m$^2$ on days 1–10), etoposide (50 mg/m$^2$ on days 1–5), and on days 1, 3 and 5, either daunorubicin (50 mg/m$^2$), mitoxantrone (12 mg/m$^2$) or idarubicin (10 mg/m$^2$) according to randomization. As consolidation, patients received cytarabine (500 mg/m$^2$/q12 h on days 1–6) and the same anthracycline as in induction. Patients with a human leukocyte antigen (HLA)-compatible sibling were allografted, whereas the others were randomly attractive owing to its high sensitivity (one target cells per $10^{3}$–$10^{6}$) in targeting specific gene aberrations and particularly fusion genes; however, this approach is limited to the fraction of AMLs bearing specific genetic lesion detectable by PCR (approximately 30–40% of cases). Multiparametric flow cytometry may allow a sensitivity of one leukemic cell per $10^{4}$ normal bone marrow (BM) cells and represents a suitable test for 75–85% of the patients. In previous experiences of our group, multiparametric flow cytometry was successfully used to quantify MRD in AML expressing specific leukemia-associated phenotypes. Our results suggested that a post-consolidation (post-Cons) time point was more informative to predict treatment outcome of AML patients than an earlier checkpoint set at the end of induction therapy, although these findings are still a matter of debate.

In the present study, by analyzing a large group of newly diagnosed AML patients in which MRD determinations were available both after induction and consolidation, we were able to validate the value of $3.5 \times 10^{-4}$ residual leukemic cells as the most appropriate threshold level to discriminate risk categories in adult AML, and to demonstrate definitely the higher prognostic significance of a delayed MRD assessment at the end of consolidation to predict disease outcome.

Correspondence: Dr A Venditti, Cattedra di Ematologia, Università di Roma 'Tor Vergata' Ematologia, Azienda Ospedaliera Universitaria Policlinico Tor Vergata, Viale Oxford, Rome 81 – 00133, Italy.
E-mail: adriano.venditti@uniroma2.it
Presented in part at the 46th ASH meeting, 4–7 December 2004, San Diego, CA, USA.
Received 4 November 2005; revised 29 May 2006; accepted 8 June 2006
assigned to peripheral or BM autologous stem cell transplant. In the AML-12 EORTC/GIMEMA trial, patients received the daunorubicin arm of AML-10 as a standard remission induction and cytarabine (500mg/m^2/q12 h on days 1–6) plus daunorubicin (50mg/m^2 on days 4–6) as consolidation. Patients with an HLA-compatible sibling were allografted, whereas the others underwent peripheral blood autologous stem cell transplantation, followed by no further therapy or subcutaneous interleukin-2 maintenance, according to a second randomization. Patients older than 60 years of age were entered onto the EORTC/GIMEMA AML-13 randomized trial. In this protocol, patients received mitoxantrone (7mg/m^2 on days 1, 3 and 5), cytarabine (100mg/m^2 on days 1–7) and etoposide (100mg/m^2 on days 1–3), as an induction therapy. Upon achievement of CR, patients were randomly assigned to receive either an intravenous or an oral consolidation program (two cycles). Intravenous consolidation consisted of idarubicin (8mg/m^2 on days 1, 3 and 5), cytarabine (100mg/m^2 on days 1–5) and etoposide (100mg/m^2 on days 1–3). Oral consolidation consisted of idarubicin (20mg/m^2 on days 1, 3 and 5), etoposide (50mg/m^2 twice a day on days 1–3) and subcutaneous cytarabine (50mg/m^2 twice a day on days 1–5).

**Immunophenotypic studies and MRD detection**

At diagnosis, immunophenotypic, karyotypic and genetic studies were performed with standard techniques, as detailed elsewhere. Once the immunophenotype of the leukemic cells was determined, cases with 'leukemia-associated' phenotypes were selected and re-analyzed by staining with the relevant combinations of antibodies in three/four-color flow cytometry. A given combination of markers was regarded as relevant if expressed in >50% of blasts. This step served to define a 'leukemia immunophenotypic fingerprint', which in turn was used to track possible residual leukemic cells during follow-up at specific time points. At least two antibody combinations were selected in each single case to minimize pitfalls owing to 'phenotypic switches' that have been described to be occasionally associated with relapses. Asynchronous and cross-lineage antigen expression represented the most frequent observed 'leukemia-associated' phenotypes, with the lack of antigen accounting for the remaining cases. Table 1 reports a list of the observed leukemia-associated phenotypes. We also studied 20 BM samples belonging to normal healthy donors and patients affected by lymphoma without BM involvement at steady state and in regenerating phase: the frequency of the observed 'leukemia-associated' phenotypes in these conditions was <0.01%. In the attempt to prevent artifactual results owing to post-chemotherapy hematopoietic regeneration, BM samples were collected after complete reconstitution following each treatment step (induction, consolidations I and II and allogeneic/autologous graft) and 1 week before the delivery of the following cycle. No differences were observed in terms of 'leukemia-associated' phenotypes frequency between samples collected upon reconstitution and those at steady state. The Cell-Quest (Becton Dickinson, Mountain View, CA, USA) software was used for the flow cytometric data acquisition applying 'live gates' on the forward light/orthogonal light scatter (blast region) and fluorescence plots. Samples were then analyzed by using the PAINT-A-GATE software program (Becton Dickinson), as described previously. MRD studies during remission were performed on erythrocyte-lysed whole BM samples using the same antibody combination defining at diagnosis the specific 'leukemia immunologic fingerprint'. During data acquisition, a live-gate including the lymphomonocytic/granuloblastic region and excluding debris and platelet aggregates was used with at least 10^6 total events being acquired in all samples. The acquired events were analyzed with the PAINT-A-GATE software, also applying the MouseTRAX Control option as described elsewhere.

**Statistical analyses**

Overall survival (OS) was calculated from the date of diagnosis to the date of death or last follow-up. Relapse-free survival (RFS) was measured from achievement of CR until relapse. Values of MRD levels, evaluated after induction and consolidation therapies, were tested for possible cutoffs by means of maximally selected log-rank statistics. Relationship of MRD to patient's characteristics and response to treatment was estimated by two-sided χ^2 test (or Fisher's exact test when either group included fewer than 20 cases). The Kaplan–Meier method was used for the estimation of OS and RFS. For comparison of survival and remission duration of two or more groups, the log-rank test was applied. CR and relapse were defined by standardized criteria. To evaluate the simultaneous impact of different variables on relapse rate and duration of OS and RFS, a multivariate analysis was performed using a stepwise regression model. A P-value of 0.05 or less was considered to be significant in all cases.

**Results**

The clinical characteristics of the 100 patients included in the analysis are detailed in Table 2. The levels of MRD were tested to identify the optimal cutoffs, yielding the best separation of AML patients into two groups with different OS and/or RFS probabilities. To do this, we evaluated the trend of standardized log-rank statistics using RFS (Figure 1, upper panels) and OS (Figure 1, lower panels) as dependent variables, and the values of residual leukemic cells, determined both at the post-induction (post-Ind) and post-Cons checkpoints, as independent variable (Figure 1a and b, respectively). The experimental cutoff points,

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Leukemia-associated phenotypes used in the study and their distribution at diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aberrant phenotypes^a^</td>
<td>Frequency^b^</td>
</tr>
<tr>
<td>CD117pos/CD56pos</td>
<td>05/100</td>
</tr>
<tr>
<td>CD33neg/CD15pos/HLA-DRpos</td>
<td>01/100</td>
</tr>
<tr>
<td>CD34pos/CD33pos/CD19pos</td>
<td>05/100</td>
</tr>
<tr>
<td>CD34pos/CD33pos/CD7pos</td>
<td>01/100</td>
</tr>
<tr>
<td>CD34pos/CD33pos/CD15pos/CD19pos</td>
<td>01/100</td>
</tr>
<tr>
<td>CD34pos/CD56pos</td>
<td>02/100</td>
</tr>
<tr>
<td>CD34pos/CD56pos/CD117pos</td>
<td>01/100</td>
</tr>
<tr>
<td>CD34pos/CD56pos/CD117pos</td>
<td>08/100</td>
</tr>
</tbody>
</table>

Abbreviation: HLA, human leukocyte antigen.

^aPos, the presence of a given antigen; neg, the absence of a given antigen; for positive antigens, the intensity of fluorescence (+, ++, ++++) is expressed when it is functional to the identification of an aberrant phenotype.

^bExpressed as no. of positive/no. of cases.

---

Prognostic value of MRD detection in adult AML
F Buccisano et al.
Determination of MRD after induction
After induction, the median level of the residual leukemic cells was $1.0 \times 10^{-3}$ (range 0–$1.0 \times 10^{-5}$). At this time point, 35% of the patients (35/100) were MRD$^+$ and 65% (65/100) were MRD$^–$. Forty-seven of 65 (72%) patients in the MRD$^+$ group underwent relapse at a median time of 7 months (range 1–36), whereas in the MRD$^–$ group 16/35 (46%) had a disease relapse at a median time of 12 months (range 4–41) ($P = 0.010$). The 5-year actuarial probability of RFS and OS (Figure 2a) was 51 and 48%, respectively, for patients in the MRD$^+$ group, compared to 22 and 25%, respectively, for those in the MRD$^–$ group ($P = 3.3 \times 10^{-3}$ and $1.7 \times 10^{-2}$, respectively).

Determination of MRD after consolidation
Seven patients had an early relapse after induction therapy (all were MRD$^+$) and 93 proceeded to receive consolidation; of these, one patient died of septic shock and 92 were evaluable for successive MRD assessment. At post-Cons analysis, the median level of residual leukemic cells in the whole series was $1.1 \times 10^{-3}$ (range 0–$6.7 \times 10^{-5}$). Thirty-six (39%) of the patients were MRD$^+$ and 56 (61%) were MRD$^–$. Nine of 36 (25%) and 47 of 56 (84%) patients in MRD$^+$ and in MRD$^–$ group underwent a relapse at a median time of 15 months (range 5–41) and 10 months (range 2–37), respectively ($P < 0.001$). The 5-year actuarial probability of RFS and OS (Figure 2b) was 71 and 64%, respectively, for patients in the MRD$^+$ group, compared to 13 and 16%, respectively, for those in the MRD$^–$ group ($P = 6.13 \times 10^{-9}$ and $8.77 \times 10^{-6}$, respectively).

Relation between levels of MRD after induction and consolidation checkpoint
Figure 3 shows the kinetics of MRD fluctuations after induction and consolidation therapy in the 92 patients completing both treatment phases. Based on the comparative analysis of MRD levels detected at the two time points, we segregated the series into four groups: (I) Ind$^+$ Cons$^+$ (47 patients), MRD$^+$ both after induction and consolidation; (II) Ind$^+$ Cons$^–$ (10 patients, MRD$^+$ after induction converted into MRD$^–$ after consolidation; (III) Ind$^–$ Cons$^+$ (nine patients), MRD$^–$ after induction converted into MRD$^+$ at the end of consolidation; and (IV) Ind$^–$ Cons$^–$ (26 patients) who were MRD$^–$ at both time points. The analysis of RFS and OS rates showed that the MRD$^–$ status at the end of consolidation was the most significant predictor of outcome, regardless of the levels of MRD after induction. In fact (Figure 4), patients with an Ind$^–$ Cons$^–$ status had the same probability of RFS and OS than those never exceeding the threshold of $3.5 \times 10^{-4}$ residual leukemic cells (Ind$^–$ Cons$^–$). These results suggest a good treatment outcome even in patients with a slow response in terms of MRD clearance.

Impact of transplantation on MRD
Of 68 patients recruited to the AML10/AML12 protocols (Table 2), 11 and 38 underwent allogeneic (SCT) and autologous (AuSCT) stem cell transplantation, respectively. One patient was too early for evaluation, whereas 18 were not transplanted because of early death (one), medical decision (four) and relapse (13). Three additional patients aging more than 60 years were transplanted (three AuSCT) according to the option of AML13 protocol for fit patients. Therefore, 41 patients were submitted to AuSCT and then followed up for MRD determination over the post-transplant period. Before AuSCT, 18/41 (44%) were classified as MRD$^–$ and 23/41 (56%) were MRD$^+$ at the
Among MRD− and MRD+ patients, a significant difference in the 5-year RFS (72 vs 11%, \( P < 0.001 \)), OS (68 vs 27%, \( P = 0.028 \)) and relapse rate (26 vs 89%, \( P < 0.001 \)) was observed. The number of patients selected for allogeneic SCT (11 cases, nine of them belonging to the post-Cons MRD+ group) was too low to draw any significant conclusion regarding the prognostic impact of MRD evaluation in this patient subset.

Relation between MRD and clinicobiologic features at diagnosis
The MRD status after consolidation was not significantly related to gender, white blood cell (WBC) count, and presence of extramedullary involvement or FAB subtype. However, among post-Cons MRD− cases, we observed an association with lower frequency of MDR1 expression (14/38, 37%, \( P < 0.001 \)), poor-risk karyotype (0/33, \( P < 0.001 \)), age > 60 years (5/36, 14%, \( P < 0.001 \)) and a high (>100 x 10^9/l) circulating blast count (8/56, 14%, \( P = 0.013 \)). Relative distributions of clinical variables among MRD− and MRD+ patients are summarized in Table 3.

Prognostic determinants
Finally, all the relevant prognostic variables with a statistical significance in univariate analysis such as age, WBC count, MDR1 phenotype, cytogenetics and MRD status after induction and consolidation were pooled into a multivariate model to determine to what extent they affected independently the outcome of treatment. In this analysis, post-Cons MRD+ status was found to be an independent variable significantly associated with a higher frequency of relapse (\( P < 0.001 \)) and a shorter duration of OS and RFS (\( P < 0.001 \) for both comparisons). Poor-risk karyotype was found to affect independently the duration of OS and RFS (\( P = 0.003 \) and 0.006, respectively) (Table 4). The prognostic impact of post-Cons MRD status remained significant also after adjustment for age < 60 and \( \geq 60 \) years (\( P < 0.001 \)).

Discussion
By sequentially studying MRD with multiparametric flow cytometry during treatment for AML, we found that the post-Cons evaluation is the most important checkpoint to predict disease outcome. Levels of MRD \( \geq 3.5 \times 10^{-4} \) leukemic cells after consolidation therapy were in fact associated with unfavorable prognosis. The present analysis confirmed the threshold of \( 3.5 \times 10^{-4} \) as the most significant value for prognostic purposes by applying the maximally selected log-rank statistics. This test was specifically developed to find out the optimal cutoff for a given biological variable correlating with clinical parameters of interest, and successfully employed by us.
and others for the identification of novel prognosticator cutoffs in other onco-hematological diseases. As shown in the present study, the application of the maximally selected log-rank statistics indicated a threshold value of $3.5 \times 10^{-4}$ residual leukemic cells also after induction therapy. The opportunity to apply, after induction and consolidation, the same reference value as a threshold to define the MRD negativity allowed us to assess the impact of the post-ind therapy on the kinetics of MRD clearance. Moreover, the availability of immunophenotypic data after induction and consolidation for the whole patient population enabled a comprehensive comparative analysis.

In our study, the early achievement of MRD negativity (post-Ind) did not provide an outcome advantage over late (post-Cons) clearance of leukemic cells. In fact, 'slow responders' patients, namely those entering MRD negativity only after consolidation (Ind$^+\text{Cons}^-$) showed the same outcome as those achieving early negativity (Ind$^+\text{Cons}^+$) (Figure 4). The discrepancy between the prognostic value of the post-Cons MRD status and the observation that most of the post-Cons MRD$^+$ patients were already negative after induction raises the question on the impact of post-Ind therapy whose contribution in terms of patient recruitment into the MRD$^+$ group was apparently modest. In fact, roughly 70% (26/36) of patients MRD$^+$ after consolidation were already negative after induction. In this respect, we speculate that an early post-Ind MRD negativity might represent a predisposing factor for favorable prognosis, but the final outcome will rely on the net debulking effect achieved with the whole (induction-consolidation) upfront therapy. This assumption is supported by the statistical observation that the MRD status after induction was significantly associated with relapse rate, OS and RFS in univariate but not in multivariate analysis. Conversely, the MRD status after consolidation retained a statistical significance either in univariate

**Figure 2** RFS and OS of AML patients according to MRD levels after induction and consolidation. Patients were grouped according to MRD cutoff value of $3.5 \times 10^{-4}$ residual leukemic cells, as determined at post-Ind (a) or post-Cons (b), and evaluated for RFS (upper panels) or OS (lower panels).

**Figure 3** Fluctuations of the MRD levels after induction and consolidation cycles. Ninety-two patients were evaluable in terms of comparison between induction and consolidation therapy; 36 (39%) were MRD$^+$ at the end of consolidation; 26 of them were already negative after induction and 10 became MRD$^-$ only after consolidation. Conversely, 56 patients were MRD$^-$ at the end of consolidation; nine of them, MRD$^+$ after induction, progressed into an MRD$^+$ status in spite of administration of the consolidation cycle. The remaining 47 patients were in a MRD$^+$ status throughout the induction and consolidation cycles.

![Graphs](image-url)
Table 3 Correlations between clinical variables and post-Cons MRD status

<table>
<thead>
<tr>
<th>MRD status</th>
<th>MRC cytogenetic adverse-risk</th>
<th>MDR1 phenotype expression</th>
<th>Age ≤ 60 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRD⁺</td>
<td>6/53 (11%)</td>
<td>42/56 (75%)</td>
<td>21/56 (38%)</td>
</tr>
<tr>
<td>MRD⁻</td>
<td>0/33 (0%)</td>
<td>11/35 (31%)</td>
<td>5/36 (14%)</td>
</tr>
<tr>
<td>P-value</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.0176</td>
</tr>
</tbody>
</table>

Abbreviation: MRD, minimal residual disease; post-Cons, post-consolidation.

Table 4 In multivariate analysis, post-Cons MRD status was independently associated with relapse rate, duration of OS and RFS

<table>
<thead>
<tr>
<th>Variable</th>
<th>Relapse rate (P-value)</th>
<th>OS (P-value)</th>
<th>RFS (P-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-Cons MRD</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MRC cytogenetic</td>
<td>NS</td>
<td>0.003</td>
<td>0.006</td>
</tr>
<tr>
<td>risk class</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>MDR1 phenotype</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Post-Ind MRD</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Age ≤ 60 years</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
</tbody>
</table>

Abbreviation: MRD, minimal residual disease; NS, not significant; OS, overall survival; post-Cons, post-consolidation; post-Ind, post-induction; RFS, relapse-free survival.

Figure 4 RFS and OS of AML patients according to the comparison of the MRD status after induction and consolidation cycle. RFS (upper panel) and OS (lower panel) of the 92 patients evaluable after consolidation. Double-headed arrows indicate Ind-Cons+ (nine patients), Ind-Cons- (10 patients), Ind+Cons+ (26 patients) and Ind+Cons- (29 patients) Kaplan–Meier curves with; the P-values obtained by comparing the four Kaplan–Meier curves (log-rank test) are reported at the bottom; double-headed arrows indicate Ind-Cons+ and Ind+Cons+ Kaplan–Meier curves with; the P-values obtained by comparing these latter curves alone are also reported.

or multivariate analysis (Table 4). The predictive value of post-Cons MRD assessment is further supported by two clinical observations: (1) the achievement of a MRD⁺ status after induction does not guarantee persistence of negativity after consolidation; in fact, we observed nine patients who entered a MRD⁺ status after induction but converted to MRD⁻ after consolidation; (2) after the (induction-consolidation) upfront phase, additional intensification by autologous stem cell transplantation may not improve the prognosis of those patients still MRD⁺ at the end of consolidation. In fact, of 41 patients who underwent AuSCT, 23 were MRD⁺ after consolidation and 89% of them had a relapse.

It is generally accepted that, to optimize the risk assignment of AML patients, a combined analysis of pretreatment (MRC profile, cytogenetics, FLT3/NPM mutations) and post-treatment parameters (MRD status, high mobilization of CD34 cells) is warranted. In our series, MDR1 expression correlated with a shorter RFS (P<0.001), whereas karyotype risk classes affected both OS and RFS (P<0.001 and P=0.006, respectively). Furthermore, post-Cons MRD status still maintained its relevance when the analysis was adjusted for age, supporting its independent prognostic role even in biologically defined risk groups.

In conclusion, the evaluation of the MRD status at the end of the (induction-consolidation) upfront treatment of AML allows to identify two patients categories: (1) MRD⁻ patients have a 5-year RFS >70% and are likely to be cured; however, as the relapse rate in this subgroup is 25%, they certainly need to be monitored in order to anticipate overt relapse on the basis of MRD expansion and (2) MRD⁺ patients have the worst prognosis, and may be candidate to additional therapy. According to our previous experience, a standard intensification by means of AuSCT seems not to be effective in this setting, whereas it is conceivable to investigate in these patients the impact of early allogeneic stem cell transplant from related (if available) or even unrelated donor or novel targeted therapies. These agents may be combined to subsequent intensification cycles after upfront therapy with the aim of lowering a higher fraction of patients to be rescued in the MRD⁻ category. Finally, our data support the utility of incorporating sequential MRD assessment in the current protocols for the
treatment of AML in order to deliver tailored therapies on a risk-based assessment.

Acknowledgements

The study was supported in part by Ministero della Salute (Ricerca Finalizzata IRCCS and ‘Allaenza contro il Cancro’), Rome, Italy. We thank Antonella Zucchetto, Clinical and Experimental Hematology Research Unit, CRO Aviano for her support in statistical analyses.

References