Prognostic and therapeutic implications of minimal residual disease detection in acute myeloid leukemia

By Francesco Buccisano¹, Luca Maurillo¹, Maria Ilaria Del Principe¹, Giovanni Del Poeta¹, Giuseppe Sconocchia², Francesco Lo-Coco¹,⁴, William Arcese¹, Sergio Amadori¹, Adriano Venditti⁵

Running title: MRD in AML.

Keywords: acute myeloid leukemia, minimal residual disease, flow cytometry, gene mutation, polymerase chain reaction.

¹Hematology, Fondazione Policlinico Tor Vergata, Rome, Italy, ²Institute of Translational Pharmacology, Department of Medicine, CNR, Rome, Italy and ³Laboratorio di Neuro-Oncoematologia, Fondazione Santa Lucia, Rome, Italy

Address correspondence to:
Adriano Venditti, MD
Ematologia, Fondazione Policlinico Tor Vergata
Viale Oxford 81 – 00133
Rome, Italy
Telephone: +39 06 20903226
Fax: +39 06 20903221
E-mail: adriano.venditti@uniroma2.it
ABSTRACT
The choice of either induction or post-remission therapy for adults with acute myeloid leukemia is still largely based on the “one size fits all” principle. Moreover, pre-treatment prognostic parameters, especially chromosome and gene abnormalities, may fail in predicting individual patient outcome. Measurement of minimal residual disease (MRD) is nowadays recognized as a potential critical tool to assess the quality of response after chemotherapy and to plan post-remission strategies that are, therefore, driven by the individual risk of relapse. Polymerase chain reaction and multiparametric flow cytometry have become the most popular methods to investigate MRD since they have been established as sensitive and specific enough to allow MRD to be studied serially. In the present review, we will examine the evidence supporting the appropriateness of incorporating MRD detection into the AML risk assessment process. A comprehensive prognostic algorithm, generated by combining pre-treatment cytogenetics/genetics and post-treatment MRD determination, should promote advances in development of personalized therapeutic approaches.
INTRODUCTION

In adult patients with acute myeloid leukemia (AML), intensive chemotherapy achieves complete remission (CR) rates ranging from 50% to 80%. Despite these encouraging results, the majority of responding patients will eventually relapse, with only 30-40% of young and less than 20% of elderly patients being long-term survivors.\textsuperscript{1-5} Advances in biologic characterization are expected to provide a more proper risk stratification of AML, allowing delivery of treatments proportional to the real aggressiveness of disease. In this view, cytogenetic abnormalities represent the most reliable prognosticator in adult AML.\textsuperscript{6-10} Identification of specific gene abnormalities (FLT3, NPM1, CEBPA, DMNT3A, etc) has further improved prognostic allocation of patients with AML, especially within homogeneous karyotypic groups, i.e. intermediate karyotypes or favorable karyotypes, where the possible concomitant mutations of c-kit, occurring in the context of core binding factor (CBF) translocations, confer a negative prognosis.\textsuperscript{11-15} A proper assignment to low- or high-risk category is a critical step in the therapeutic decision making process of patients with AML. In fact, patients belonging to the low-risk category would not benefit from routine use of allogeneic stem cell transplantation (ASCT) in first CR since any advantages in terms of reduced relapse incidence will be counterbalanced by procedure related morbidity and mortality.\textsuperscript{16} On the other hand, missing early identification of high-risk patients might hamper or delay the timely use of intensive treatments such as ASCT.\textsuperscript{17} However, it is also well recognized that cytogenetic/genetic signature cannot always reliably predict outcome in individual patients; as a matter of fact, about 40-50% of those with favorable karyotype will eventually experience a relapse. In this view, measure of minimal residual disease (MRD) promises to be an efficient tool to establish on an individual basis the patient’s leukemia’s susceptibility to treatment, enhancing delivery of risk adapted therapies.\textsuperscript{18-19} In spite of these premises, the systematic applicability of MRD analysis in AML is not yet accomplished. For cases with a genetic signature that account for 60-70% of AML, the molecular approach still suffers from lack of standardized assays and cutoffs. Similarly, although flow cytometry can potentially investigate >85% of AML, its widespread applicability is complicated by lack of standardized procedure. The present review will focus on the technical issues and clinical relevance of MRD determination in non-M3 AML. In particular, we will try to outline how MRD assessment might improve risk stratification.
THE CASE OF OUTCOME PREDICTION

Clonal chromosome alterations are detected in AML in more than 50% of adults and are universally considered the strongest predictor of duration of response and overall survival (OS). Based on cytogenetics, patients are generally stratified into 3 risk groups: favorable, intermediate and unfavorable.\textsuperscript{6-10} Accordingly, patients falling into the category of favorable risk karyotype (F-RK) can expect an approximate 65-70\% likelihood of cure, those who have an intermediate risk karyotype (I-RK) an approximate 40\% chance of long-term disease free survival (DFS) whereas patients belonging to the unfavorable risk karyotype (U-RK) category have a very dismal outcome, with less than 5-10\% becoming long-term survivors.\textsuperscript{6-10} However, it has become clear that cytogenetic risk allocation may help guide therapeutic decision particularly for those who are at the extremes (F-RK or U-RK) whereas, due to the heterogeneity of this category, for patients within I-RK the therapeutic decision may be problematic. In fact, 40-50\% of patients categorized in this group lack any clonal abnormalities on standard cytogenetic analysis. Furthermore, together with cytogenetically normal AML (CN-AML), the I-RK also includes a miscellany of different structural and numerical changes that are too infrequent to be reliably assigned a prognostic significance and are not encompassed in the two other risk groups. In recent years, various molecular markers have been identified allowing to dissect further cytogenetically defined subsets.\textsuperscript{20} This is critical to risk assessment of patients with CN-AML (Table 1). Although the expectations of using genetics to guide therapy were enormous, the delivery to the clinics has made slow progress and the therapeutic strategy for most patients still remains controversial. On one side, risk assessment based upon pre-treatment factors comprises a large group of patients with broad differences in terms of risk of relapse. On the other side, fitting the increasing number of new gene abnormalities into a practical, prospective, decision making algorithm is a very difficult task. In fact, the excess of dissection in homogeneous genetic subsets may lead to a plethora of subgroups failing to provide physicians with an adequate statistical power (Figure 1). A reasonable way to improve outcome prediction in AML might be attained by combining pre-treatment and post-treatment parameters into a common prognostic algorithm. Such a hypothesis found acceptance in a seminal manuscript published by Wheatley et al who generated a prognostic scoring system by integrating pre-treatment cytogenetics and achievement of CR after 1 or 2 cycles of induction therapy.\textsuperscript{21} Patients with U-RK who entered CR after 2 induction courses had the worst outcome. The design of the analysis represented a first attempt to put together into a risk stratification algorithm two different classes of prognostic
parameters: pre-treatment factors and indicators of the quality of response after a chemotherapy-induced morphologic CR. Nowadays, such a working hypothesis can be developed further thanks to the availability of techniques such as polymerase chain reaction (PCR) and multiparametric flow-cytometry (MPFC). These techniques can reliably detect, at high sensitivity, leukemic cells at submicroscopic level, thus offering the opportunity to investigate MRD. Assessment of MRD promises to be a powerful and accurate tool to refine patients’ risk category assignment as initially established on the basis of the sole cytogenetic/genetic findings.
THE QUEST OF RESIDUAL LEUKEMIA

PCR and MPFC have proven sensitive and specific enough to allow MRD to be investigated and represent today the gold standard for MRD monitoring in AML.

MRD detection by PCR

Leukemic fusion genes

The power of PCR relies on cloning of breakpoints of the chromosomal rearrangements and allows their detection in the post-remission phase in at least 30% of the patients, using reverse-trancriptase PCR (RT-PCR) or real-time quantitative PCR (RQ-PCR). Common targets for PCR-based MRD detection are fusion transcripts of CBF-positive AML, e.g. RUNX1-RUNX1T1, CBFB-MYH11 and MLL-gene fusions. RT-PCR in CBF positive AML has a limited clinical applicability as persistent PCR positivity has been observed in long survivors even after ASCT.\(^{24}\) RQ-PCR has proven potentially more valuable owing to its capability to anticipate impending relapse during follow-up monitoring.\(^{25}\) In a recent report, Corbacioglu et al established clinically relevant MRD cutpoints at which persistence of CBFB-MYH11 transcript positivity singled out patients with significantly increased risk of relapse. The authors conclude that monitoring of CBFB-MYH11 transcript levels should be incorporated into future clinical trials to guide therapeutic decisions.\(^{26}\)

Mutations (Table 1)

The discovery of gene mutations in fusion gene-negative AML has potentially increased to 60-70% the percentage of cases suitable for PCR-based MRD monitoring.\(^{27-59}\) The key question about the use of these genes as candidates for MRD monitoring regards their stability over the course of disease. The receptor tyrosine kinase FLT3 is mutated in a relevant proportion of CN-AML. Mutations resulting in the constitutive activation of FLT3 have been identified in two functional domains of the receptor, the juxtamembrane (JM) domain and the split tyrosine kinase domain (TKD).\(^{28,29}\) The JM domain is disrupted by internal tandem duplications (ITD) of various sizes in 28-34% of CN-AML whereas point mutations of the TKD codon 835 or 836 have been reported in 11-14% of CN-AML.\(^{30-32}\) The role of FLT3-ITD in MRD monitoring is controversial: lack of stability of FLT3 mutations in paired samples from diagnosis and relapse raises concerns about the clinical usefulness of this marker.\(^{29,36}\) However, some authors claim that this lack of longitudinal stability might relate to lack of sensitivity. Using high sensitivity RT-PCR, it was observed that in all 25 cases under study FLT3-ITD was detected at diagnosis and relapse. In 4 of them, relapse was predictable based on serially documented reappearance of FLT3-ITD. These results suggest that there is still room for revisiting the role of FLT3-ITD as a...
potential marker for MRD monitoring.\textsuperscript{60} While there are only initial experiences dealing with the use of CEBPA and MLL-PTD in MRD monitoring,\textsuperscript{61,62} increasing knowledge is being accumulated on the role of NPM1 mutations. NPM1 mutations can be identified in 45-60\% of patients with CN-AML, accounting for the most frequent genetic change in this subset. The mutational event modifies specific nucleoli binding and nuclear export signal motifs coded by exon 12 and determines an abnormal cytoplasmic localization of NPM1.\textsuperscript{37-38} Approximately 40\% of patients with NPM1 mutations also carry FLT3-ITD. Some studies indicate that NPM1 mutations are very stable at relapse, and thus that they might have a role in MRD assessment.\textsuperscript{63} Schnittger et al developed a highly sensitive RQ-PCR assay able to prime 17 different mutations of NPM1.\textsuperscript{64} In 252 NPM1-mutated AML, high levels of NPM1\textsuperscript{mut}, quantified at four different time-points, were significantly correlated with outcome at each tested time-point. Multivariate analysis including age, FLT3-ITD status and the level of NPM1, demonstrated that the latter was the most relevant prognostic factor affecting EFS during first-line treatment, also in the subgroup of patients who underwent ASCT. Furthermore, the authors compared the kinetics of NPM1 and FLT3-ITD fluctuations and found that FLT3-ITD levels suffer from an unpredictable instability making it unreliable for MRD purposes.\textsuperscript{29,36} In a further refinement of such an approach, Kronke et al\textsuperscript{65} demonstrated that the level of NPM1\textsuperscript{mut} measured after double induction and consolidation therapy, impacted on OS and cumulative incidence of relapse (CIR) (p<0.001 for all comparisons).

\textbf{Gene overexpression}

Techniques of RQ-PCR intended to quantify genes over-expression theoretically provide the opportunity to study the whole population of AML. Various candidates have been proposed, with WT1 being the most reliable. Retrospective clinical data indicate that WT1 assessment after induction therapy is a predictor of outcome with lower levels being associated with long-term remission.\textsuperscript{66} WT1 over-expression can be exploited as a marker to establish the presence, persistence or reappearance of leukemic hematopoiesis. Cilloni et al\textsuperscript{67} have measured the levels of WT1 after chemotherapy-induced morphologic CR, in peripheral blood (PB) and bone marrow (BM) of patients with AML. In multivariate analysis, they found that a ≥2-log MRD reduction in PB and/or BM was associated with a significantly lower CIR (p=0.004). Although the manuscript is of outstanding relevance due to the provision of a common standardized protocol on the behalf of the LeukemiaNet, only in 46\% and 13\% of PB and BM samples were the levels of WT1 sufficiently over-expressed, as compared to normal samples, to allow a prognostic stratification. Based on
concerns still remain about the confounding role of the physiologic background of WT1 in normal PB and BM.

**MRD detection by flow-cytometry**

MRD monitoring by MSFC relies on the expression of “leukemia-associated immunophenotypes” (LAIPs) defined as the presence of a combination of antigens and/or flow cytometric physical abnormalities that are absent or very infrequent in normal BM. The growing interest surrounding this technique is due to its wide applicability (>85% of AML), quickness, specificity and ability to distinguish viable cells from BM debris and dead cells. Furthermore, the diffusion of devices equipped with multiple lasers allowed implementation of multiple color assays (>4–5 monoclonal antibody combinations) thus favoring increment of sensitivity that, at the present time, can be reasonably placed in between 10^{-4} and 10^{-5}. An additional advantage of multiple color assays consists in a significant attenuation of concerns of phenotypic shifts that can be observed upon recurrence.\(^6\)\(^8\) There is evidence that the use of an expanded 6-9 color polychromatic assay not only increases sensitivity\(^6\)\(^9\) but also results in improved qualitative information on the leukemic clone thanks to the superior definition of its phenotypic composition. In this context, MRD monitoring might not be compromised, especially if serial determinations are performed. The results from the literature demonstrate that detection of MRD by MSFC is technically sound and represents a powerful tool to segregate patients with AML into categories of risk (Table 2). Most studies do not deal with the MRD issue as a simple matter of a positive or negative finding, rather they set a threshold below or above which the outcome can be significantly different. A common approach is to set up empirically the most significant level of MRD by choosing a logarithmic scale (i.e. 10^1, 10^2, 10^3, etc.) or a quartile segregation that correlate with survival estimates and CIR. An alternative approach relies on the evaluation of the prognostic role of MRD in a continuous variable model. In the experience of Kern et al\(^7\)\(^0\),\(^7\)\(^1\) such an analysis demonstrated a strong correlation with outcome when assessed on day 16 from induction and after consolidation. However, when it came stratifying the patients in low or high-risk categories, significant thresholds were identified. At least four manuscripts\(^7\)\(^2\)-\(^7\)\(^5\) tried to address the issue of MRD “prognostically significant levels” applying specific statistical methods that would help selecting thresholds more appropriately. In the paper by Al Mawali,\(^7\)\(^5\) in order to determine the optimal cutoff yielding the best segregation of AML patients into categories of risk, a Receiver Operating Characteristic (ROC) analysis was carried out. The threshold was established at the value of 0.15% residual leukemic cells, therefore, patients with MRD
>0.15% qualified as MRD-positive, whereas those with MRD ≤0.15% as MRD-negative. We evaluated the trend of standardized log-rank statistics using relapse-free survival (RFS) and OS as dependent variables and the values of residual leukemic cells determined after induction and consolidation as independent variables (Maximally Selected log-rank Statistics). Based on the results of this test, we currently utilize the value of 0.035% BM residual leukemic cells to discriminate MRD-negative from MRD-positive cases, both after induction and consolidation. Although the use of a dedicated statistical approach to establish the threshold for MRD negativity/positivity is recommended, the reproducibility of that specific threshold throughout different laboratories still remains an issue. Should we select a particular cutoff which different laboratories can refer to or should any laboratory set up its own? A universal cutoff would be desirable but it might represent a very troublesome mission to accomplish since standardizing thresholds is affected by a number of technical (equipment, fluorochromes, sensitivity of LAIPs, procedures of acquisition and analysis) and clinical factors. Among clinical factors, the different intensity of chemotherapeutic regimens delivered may play a critical role. In fact, the clinically relevant threshold of MRD may depend on the therapy used and might change with changes in therapeutic schedule. Divergence between our experience and that of San Miguel is an example of such a situation. In the study of San Miguel, the induction, consolidation and intensification therapy consisted of the combination of an anthracycline and cytosine arabinoside. In our study, the patients were treated on three drug-based regimens, associating an anthracycline with cytosine arabinoside and etoposide. In addition, cytosine arabinoside administration was prolonged through 10 days instead of conventional seven during the induction phase. Thus, one may assume that in the study of San Miguel, the less intensive therapy delivered was associated with a milder debulking effect that in turn may account for the higher level of MRD at which a significant influence on disease outcome was found. Based on this, we can also assume that each new protocol will most likely define a new threshold.
PB VERSUS BM MRD DETECTION

PB may represent an alternative source of cells for the purpose of MRD studies. This is based on the assumption that the presence of circulating blasts at the time of CR might reflect the persistence of malignant cells in the BM. Studies using RQ-PCR to monitor MRD in CBF-positive AML showed that the transcripts were detectable in PB and BM with a comparable sensitivity.\textsuperscript{79-80} The levels of WT1 transcript measured after consolidation in PB and BM were found to be equally associated with the risk of relapse, and the sensitivity of PB WT1 analysis resulted to be equivalent if not better than that of BM.\textsuperscript{66-67} We have demonstrated the feasibility of MRD detection in PB of 50 adult AML patients using MPFC.\textsuperscript{74} The levels of MRD after induction and consolidation in PB significantly reproduced those observed in BM (r=0.86 p<0.0001 and r=0.82 p<0.0001, respectively). A level of MRD >0.015% in PB after consolidation was associated with a significant likelihood of subsequent relapse and a shorter RFS. Our data suggest that PB may be a complementary source of cells for MRD studies in patients with AML. In addition, combined measurement of MRD in BM and PB might improve the risk stratification process. The clinical impact of MRD contamination of PB apheresis product is a further subject of research since it has been associated with a shorter RFS and OS.\textsuperscript{81} Further studies are warranted to clarify these issues.
WHEN MRD ASSESSMENT IS CLINICALLY RELEVANT?

Timing of PCR-assessment

As soon as experience was accumulated, it became clear that selection of the most appropriate time-point to determine MRD was a key issue in order to have a risk-adapted approach transposed into the clinical reality. Ideally, such a time-point should be the one having the power to provide the most informative prognostic indication, so that the choice of post-remission therapy is driven by the actual risk of relapse. The paradigm of such a situation is represented by acute promyelocytic leukemia (APL), where the persistence of PML-RARA fusion transcript at the end of consolidation therapy or subsequent recurrence of PCR positivity in patients previously in molecular remission was demonstrated to precede overt relapse. CBF-positive cases have been extensively investigated by RQ-PCR for MRD persistence at specific time-points. Perea et al found that MRD positivity was associated with an increased risk of relapse at any time of assessment but only MRD persistence at the end of treatment and during subsequent follow-up significantly anticipated relapse. Similar results have been reported in CBFB-MYH11 AML by Corbacioglu et al who identified in the post-consolidation and early follow-up phase (≤3 months) the critical time-points allowing patients at higher risk of relapse to be recognized. Kronke et al observed that detection of high NPM1mut transcript levels after double induction and consolidation correlated significantly with an increased CIR. In accordance with what we have learned by assessing MRD in APL, experience with CBF-positive and NPM1mut AML appears to indicate that MRD levels as measured at delayed rather early time-points are of superior prognostic relevance. At variance with this assumption, Cilloni et al suggest that risk stratification is indeed improved by early assessment. In fact, they reported an increased risk of relapse when the level of WT1 failed to reduce ≥2-log after induction; the magnitude of WT1 reduction remained independently significant even after adjusting for competitive covariates such as age, white blood cell count and cytogenetics. A final issue pertains to the hypothesis that the optimal sampling interval varies with molecular subgroups. Ommen et al have demonstrated that the kinetics of molecular relapse are remarkably different among NPM1, PML-RARA, RUNX1-RUNX1T1 and CBFB-MYH11 AML. The authors derived a mathematical model to investigate the molecular relapse and the time from molecular relapse to hematological relapse. They found that CBFB-MYH11 AML displayed a slower clone regrowth than AML with other molecular signature. These results will potentially optimize MRD monitoring, allowing the identification of suitable sampling intervals for other molecular signatures.
Timing of MPFC-assessment

With regard to the MPFC approach, the issue of the best time point is even more controversial. The German AML Cooperative group demonstrated that assessment of MRD persistence on day 16 from induction and the log-difference between MRD-positive cells on day 1 and 16 from induction, represented an independent prognostic factor affecting CR, EFS, OS and RFS. However, it is important to note that this analysis is something different from MRD assessment, for it pertains to the concept of “speed of blast clearance”. “Speed of blast clearance” is supposed to reflect the chemosensitivity of the leukemic clone but it does not necessarily relate to the quality of response as determined, later on, upon full hematopoietic reconstitution. To the best of our knowledge, no formal demonstration has been published on the correlation between “fast blast clearance” and achievement of MRD negativity. San Miguel et al demonstrated that a stratification according to levels of MRD <0.01%, 0.01-0.1%, 0.1-1%, and >1% after induction therapy, resulted in significant differences in OS. Al Mawali et al found that a threshold of 0.15% residual leukemic cells discriminated MRD-negative from MRD-positive cases with optimal sensitivity and specificity, allowing impending relapse to be predicted at post-induction and post-consolidation time points. Multivariate analysis showed that the post-induction MRD level affected independently RFS and OS. However, also diverging opinions have been published supporting the hypothesis that delayed time points may be even more informative as compared to earlier ones. We demonstrated that levels of MRD ≥0.035% as measured after consolidation therapy predicted a high frequency of relapse and a short duration of OS and RFS; the prognostic role of MRD positivity after consolidation was confirmed in multivariate analysis. In line with our experience, Kern et al reported that the 75th percentile of the MRD log-difference between the first day of induction and post-consolidation time-point was the sole variable dividing the patients in two groups with significantly different OS. Selecting an early or delayed time point for MRD determination might entail the choice of different therapeutic options: the early time-point option may prove useful to identify as soon as possible high-risk patients for whom a fast allocation to very intensive treatments is required. For these patients, approaches such as dose-dense schedule and/or ASCT could be incorporated into the upfront treatment strategy. On the other hand, opponents of this hypothesis raise concerns about situations of over-treatment for patients showing a slow blast clearance. In our experience, approximately 30% of patients who are MRD-positive after induction, become negative at the end of consolidation and the clinical outcome of these “slow responders” was not significantly
different from that of patients who tested MRD-negative soon after induction. Finally, we believe that special consideration should be given to specific situations where a serial sampling may be recommended. In our experience, sequential MRD monitoring is required in the phase following autologous stem cell transplantation (AuSCT) in order to enhance detection of impending relapse in patients who are MRD-negative after consolidation.
OUTCOME PREDICTION: CYTOGENETICS/GENETICS, MRD OR BOTH?

In AML, cytogenetic and molecular findings at diagnosis are critical determinants of outcome and allow stratification of ~40% of patients into "good risk" (based on the presence of mutated NPM1 without FLT3-ITD or F-RK) or "poor risk" (FLT3-ITD mutations or U-RK) categories. Good risk patients achieve high OS and DFS rates with standard treatments, while poor risk do unsuccessfully without intensified therapy with ASCT. On the other hand, there are no accepted criteria to direct the decision-making process after induction/consolidation for patients included in the I-RK (~60%). For these patients, evaluation of MRD status appears appropriate in order to extrapolate those at high (MRD-positive) or low (MRD-negative) risk of relapse, for whom differentiated treatments may be adopted. Based on this assumption, we have tried to optimize risk assessment of patients with AML by integrating evaluation of pre-treatment cytogenetics/genetics and MRD status at the post-consolidation time-point. Of 143 adult patients, those with F-RK and I-RK who were MRD-negative had 4-yrs RFS of 70% and 63%, and OS of 84% and 67%, respectively. Patients with F-RK and I-RK who were MRD-positive had 4-yrs RFS of 15% and 17%, and OS of 38% and 23%, respectively (p<0.001 for all comparisons, Figure 2, A-B). Likewise, FLT3 wild-type patients achieving a MRD-negative status (Figure 2 C-D), had a better outcome than those who remained MRD-positive after consolidation (4-yrs RFS 54% vs 17% p<0.0001, OS 60% vs 23% p=0.002). Therefore, patients with F-RK, I-RK or FLT3 wild-type had a very different outcome depending on MRD status at the end of consolidation. The Children’s Oncology Group has recently carried out the same analysis, based on which children with standard risk AML and MRD positivity after the first induction were reallocated to the high-risk category, whereas those MRD-negative were added to the favorable-risk cohort. This risk stratification system generated two new categories with a 3-year DFS of 20% and 68%, respectively (p<0.001). The authors conclude that cytogenetics, molecular genotyping, and post-induction MPFC analysis provide a robust means of stratifying pediatric AML into 2 risk groups with significantly different outcomes.
THE LEUKEMIC STEM CELL (LSC) RESIDUAL DISEASE

The role of the LSC in AML recurrence it is now more than a simple hypothesis. Relapses derive from re-expansion of residual leukemic cells escaping the cytotoxic effect of chemotherapy; this leukemic cell population is thought to reside within the stem cell CD34⁺CD38⁻ compartment. Experimental data indicate that LSC are more resistant to chemotherapy than the more mature CD34⁺CD38⁺ progeny and can be distinguished from their normal counterpart owing to the expression of LSC-specific immunophenotype.⁹⁴ Therefore, the concept of LAIP expression is something that applies even to LSC on which aberrant over-expression of CD47,⁹⁵ CD123, CD44 has been described.⁹⁴ Among markers of LCS aberrancy, C-type lectin-like molecule-1 (CLL-1) promises to be one of the most appealing since its expression is lacking on normal stem cells.⁹⁶-⁹⁷ Therefore, CLL-1 has the double role of LSC specific antigen and potential target for future LSC directed-therapy. By using a monoclonal antibody targeting CLL-1, Schuurhuis et al demonstrated that a LSC frequency >1×10⁻³ after induction and >2×10⁻⁴ after the second induction and consolidation predicted a short RFS (p=0.00003 and 0.004, respectively). By combining the residual LSC and the “whole MRD” fractions they came up with 4 different categories whose outcome was the best for the group with a negative residual LSC status.⁹⁸ The persistence of residual LSC may explain why a certain proportion of MRD-negative patients experience disease recurrence: in our experience such a proportion accounts for 20-25% of MRD-negative patients. Therefore, monitoring of LSC-LAIPs represents an additional tool capable to refine the MPFC MRD assessment and the combination of LSC-LAIPs and “whole MRD blast” frequencies might prove useful to guide future therapeutic intervention.
IMPLEMENTING RISK-ADAPTED TRANSPLANTATION POLICY

The choice of post-remission therapy is too often driven by a “genetic randomization” resulting in a given patient being allocated to ASCT once the availability of an HLA-compatible sibling donor is documented. Meta-analysis of large prospective studies indicates that the beneficial effect of ASCT takes place as soon as the risk of relapse exceeds 35-40%; when probabilities of relapse are below those percentages the risk of treatment-related mortality will attenuate the survival advantage of this procedure. Furthermore, the extensive use of ASCT is hampered by the paucity of candidates (25-30%) with a fully matched family donor. For the remaining, even if a timely search of either international registries or cord blood banks is started, the probability of identifying a suitable donor is 46% and 73% at 3 and 6 months, respectively (W. Arcese, personal communication). In the meantime, almost 40% of candidates for transplant die of relapse or are relegated to less effective approaches like chemotherapy or AuSCT. We have demonstrated that an adjusted risk stratification based on pretreatment genetics/cytogenetics and MRD status at the end of consolidation refines the upfront genetic/cytogenetic risk classification. In this view, MRD assessment will greatly enhance selection procedures so that patients are assigned to ASCT not on the basis of donor availability but of the real risk of relapse. Applying this adjusted risk-stratification, we distinguished two categories of patients: 1) Low-risk: F-RK and I-RK that were MRD-negative after consolidation; and 2) High-risk: U-RK, FLT3-ITD mutated cases, F-RK and I-RK that were MRD-positive after consolidation (Figure 3). Following these observations, we started a program for high-risk AML based on a prospective assignment to ASCT that should be delivered in the form of matched sibling donor (MSD), matched unrelated donor (MUD), umbilical cord blood (UCB) or haploidentical related donor (HRD) transplant. We analyzed preliminarily a cohort of 21 high-risk patients who, according to the above-mentioned policy, were assigned to ASCT (8 MSD, 7 MUD/UCB and 6 HRD). For comparative purposes, a matched historical cohort of 36 high-risk patients was analyzed: 12 were given MSD-ASCT and 24, lacking a MSD, received AuSCT. Survival estimates were significantly better for the prospective cohort as compared to the control group (DFS 70% vs 20%, p=0.00047; OS 69% vs 24%, p=0.046) (Figure 4). The prospective cohort also showed a lower relapse rate (20% vs 52%, p=0.003).
SUMMARY AND CONCLUSIONS
The paradigm of treatment for adult AML has largely been based on the “one size fits all” approach: in the short term, this has led to satisfactory rates of CR but in the long term, survival estimates for young and elderly patients are disappointing. In APL and acute lymphoblastic leukemia MRD positivity has consistently been shown to increase probabilities of relapse, therefore attainment of MRD-negative remission represents a “gold standard” leading to prolonged duration of CR. More recently, a trial of prospective MRD-driven therapy has also been initiated in childhood non-M3 AML, demonstrating an improvement of outcome in high-risk patients. All together these observations suggest that, whatever the method, even in non-M3 AML measurement of MRD will potentially contribute refining risk assessment. In this view, a comprehensive risk stratification, generated by integrating the prognostic weight of pre-treatment and post-treatment parameters (MRD), might help to allocate the majority of patients to a more realistic category of risk, thus favoring selection of more appropriate post-remission strategies. In the context of such an integrated approach, pre-treatment cytogenetic/genetic profile will dictate intensity and type of induction therapy, whereas post-treatment MRD status helps modulating intensity of post-remission strategies, allowing a treatment proportional to the individual risk of relapse to be delivered. The final purpose is that patients will be assigned to ASCT not on the basis of donor availability (donor vs no donor approach) but on the basis of their adjusted (cytogenetic/genetic plus MRD) risk of relapse (transplant vs no transplant approach). This will also comply with the primary goal of saving additional lives even without any major advances in chemotherapy or transplant technologies.
REFERENCES


54. Schnittger S, Haferlach C, Ulke M, et al. IDH1 mutations are detected in 6.6% of 1414 AML patients and are associated with intermediate risk karyotype and unfavorable prognosis in adults younger than 60 years and unmutilated NPM1 status. Blood. 2010; 116(25): 5486-96.


Table 1. Gene mutations occurring in cytogenetically normal acute myeloid leukemia.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Incidence</th>
<th>Biological features</th>
<th>Impact on clinical outcome and targeted therapy opportunities</th>
</tr>
</thead>
</table>
| *FLT3* ITD | Found in 28-34% of CN-AML | • Member of the class III receptor tyrosine kinase family  
  • In-frame mutations, mostly in exons 14 and 15 of the juxta-membrane domain;  
  • Level of mutant allele likely of importance  
  • Homozygous *FLT3* mutations as a result of mitotic recombination leading to partial UPD. | • Consistently associated with inferior outcome in terms of OS, RFS, EFS;  
  • Constitutively active FLT3 molecules are targets for specific molecular inhibitors (phase 2-3 trials). |
| *TKD* | Found in 11-14% of CN-AML | • In-frame point mutations of the tyrosine kinase domain (TKD) codon 835 and 836;  
  • Level of mutant allele may be of importance. | • Meta-analysis suggesting a negative prognostic impact;  
  • More recently associated with better OS;  
  • High-level mutants associated with improved OS. |
| *NPM1* | Found in 25-35% of AML (45-62% of CN-AML) | • The mutational event modifies specific nucleoli binding and nuclear export signal motifs coded by the exon 12 and determines an abnormal cytoplasmic localization of NPM1;  
  • Associated with presenting clinical and laboratory features such as female sex, higher BM blast counts and LDH levels, as well as high CD33 but low or absent CD34 levels;  
  • Associated with *FLT3*-ITD and TKD mutations (40% of cases);  
  • NPM1 acts as a co-repressor in retinoic acid (RA)-associated transcriptional regulation. | • *NPM1*mut*/FLT3*-ITDneg genotype associated with a favorable OS, RFS, CR rate;  
  • Patients with *NPM1*mut*/FLT3*-ITDneg genotype may not benefit from MRD allogeneic transplantation in first CR;  
  • Pharmacologic modulation of the RA-signaling pathway has been suggested as a therapeutic option. |
| *CEBPA* | Found in 15-20% of CN-AML | • Transcription factor mediating lineage specification and differentiation of multipotent myeloid progenitors into mature neutrophils;  
  • Concurrent mutations (namely *FLT3*-ITD and *NPM1*) | • Associated with higher CR rate and better RFS and OS;  
  • The presence of CEBPA biallelic (bm) but not single |
were virtually not present in cases with CEBPA-bm as compared to CEBPA-sm.\textsuperscript{44}

<table>
<thead>
<tr>
<th>Allele</th>
<th>PTD found in 5-11% of CN-AML</th>
<th>Usually involves exons 5 through 11 or, less frequently, exons 5 through 12; MLL wild-type allele is silenced, most likely as a result of differential DNA methylation and histone modifications; Rationale for the use of DNA methyltransferase and/or histone deacetylase inhibitors based on in vitro data.</th>
<th>Associated with shorter CR duration, or inferior RFS and EFS;\textsuperscript{60} Allogeneic transplantation may improve outcome.\textsuperscript{11}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MLL</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IDH1/2</strong></td>
<td>Found in 15% of AML</td>
<td>Correlated with accumulation of the cancer associated metabolite 2-hydroxyglutarate; IDH2 gene encodes a mitochondrial protein homologous to IDH1 that also catalyzes isocitrate carboxylation; IDH1 and IDH2 mutations are mutually exclusive.</td>
<td>IDH1 mutations identify patients at higher risk of relapse within the favorable risk-category of NPM1 mutated AML;\textsuperscript{50-53} IDH2 mutations get along with specific pre-treatment features (low WBCc, lack of other mutations and primary resistance to chemotherapy).\textsuperscript{50-53}</td>
</tr>
<tr>
<td><strong>DNMT3A</strong></td>
<td>Found in 22% of AML.</td>
<td>Somatic mutation in DNMT3A, encoding a DNA methyltransferase; Highly enriched in the group of patients with an intermediate-risk cytogenetic profile (56/66, 33.7%); Absent in the subgroup of 79 patients carrying a favorable-risk cytogenetic; A perturbation in the methylation of specific DNA sequences has been postulated.</td>
<td>Independent determinant of dismal prognosis both in the overall study population and, significantly, within high-risk patients (FLT3-ITD, age older than 60 years);\textsuperscript{54} Only ASCT provided a significant benefit.\textsuperscript{54}</td>
</tr>
<tr>
<td><strong>BAALC over-expression</strong></td>
<td>Rare</td>
<td>Expressed in neuroectoderm-derived tissues and hematopoietic precursors; Detected in a subset of karyotypically normal adults under the age of 60 years with \textit{de novo} AML.</td>
<td>Significantly worse CR rates and shorter DFS, EFS and OS than patients with low expression;\textsuperscript{55} SCT in first CR might overcome</td>
</tr>
<tr>
<td><strong>ERG over-expression</strong></td>
<td>Rare</td>
<td>• Identified in CN-AML patients and patients with prognostically unfavorable complex karyotypes with cryptic amplification of chromosome</td>
<td>• Significantly shorter OS and higher CIR.56-57</td>
</tr>
<tr>
<td>------------------------</td>
<td>------</td>
<td>-----------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td><strong>RAS</strong></td>
<td>Found in 9-13% of CN-AML</td>
<td>• Constitutively activated by point mutations either in the guanosine triphosphate (GTP)-binding domain (codon 12/13) or in the GTPase domain (codon 61). • Possible susceptibility blasts to farnesyltransferase inhibitors (FTIs).</td>
<td>• A clear prognostic role could not be elucidated yet.42</td>
</tr>
<tr>
<td><strong>WT1</strong></td>
<td>Found in ~10% of CN-AML</td>
<td>• The majority of mutations in both studies were frameshift mutations of exon 7 while only one was identified in exon 9.</td>
<td>• Initial studies on small patient cohorts suggest association with induction failure.47-49</td>
</tr>
</tbody>
</table>

Abbreviations: BM, bone marrow; LDH, lactate dehydrogenase; CN, cytogenetically normal; ITD, internal tandem duplication; TKD, tyrosine kinase domain; MRD, matched related donor; JM, juxtamembrane domain; UPD, uniparental disomy; OS, overall survival; RFS, relapse-free survival; PTD, partial tandem duplication. NPM1, nucleophosmin (nucleolar phosphoprotein B23, numatrin); FLT3-ITD, internal tandem duplication of the fms-related tyrosine kinase 3 (FLT3) gene; CR, complete remission; EFS, event-free survival; RFS, relapse-free survival; DFS, disease-free survival; OS, overall survival; CEBPA, CCAAT/enhancer binding protein (C/EBP), alpha; CRD, CR duration; MLL-PTD, partial tandem duplication of the myeloid/lymphoid or mixed-lineage leukemia (MLL) gene; IDH1/2, isocitrate dehydrogenase 1 and 2; DNMT3A DNA methyltransferase 3A gene; BAALC, brain and acute leukemia gene, cytoplasmic; ERG, v-erythroblastosis virus E26 oncogene like (avian); CIR, cumulative incidence of relapse.
Table 2. MRD by MPFC studies including more than 50 patients

<table>
<thead>
<tr>
<th>Author</th>
<th>Patients (n°)</th>
<th>Time-point</th>
<th>Threshold post-IND</th>
<th>Threshold post-CONS</th>
<th>Method</th>
<th>Multivariate analysis</th>
<th>Survival parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>San Miguel</td>
<td>76</td>
<td>I,C</td>
<td>0.5%</td>
<td>0.2%</td>
<td>NA</td>
<td>I,C</td>
<td>RFS</td>
</tr>
<tr>
<td>Venditti</td>
<td>56</td>
<td>I,C</td>
<td>0.045%</td>
<td>0.035%</td>
<td>empirical</td>
<td>C</td>
<td>OS,RFS</td>
</tr>
<tr>
<td>San Miguel</td>
<td>126</td>
<td>I</td>
<td>&lt; 0.01%</td>
<td>0.01%-0.1%, 0.1%-1%, &gt; 1%</td>
<td>NA</td>
<td>NA</td>
<td>I</td>
</tr>
<tr>
<td>Kern</td>
<td>62</td>
<td>I,C</td>
<td>Continuous analysis log-difference</td>
<td>Continuous analysis log-difference</td>
<td>75th percentiles of log difference</td>
<td>C</td>
<td>RFS</td>
</tr>
<tr>
<td>Kern</td>
<td>106</td>
<td>Day 16 after start of induction</td>
<td>Continuous analysis log-difference</td>
<td>NA</td>
<td>Median of log difference</td>
<td>Day 16 after start of induction</td>
<td>EFS,RFS</td>
</tr>
<tr>
<td>Buccisano</td>
<td>100</td>
<td>I,C</td>
<td>0.035%</td>
<td>0.035%</td>
<td>Maximally selected log-rank statistic</td>
<td>C</td>
<td>OS,RFS</td>
</tr>
<tr>
<td>Maurillo</td>
<td>142</td>
<td>I,C</td>
<td>0.035%</td>
<td>0.035%</td>
<td>Maximally selected log-rank statistic</td>
<td>C</td>
<td>OS,RFS</td>
</tr>
<tr>
<td>Al Mawali</td>
<td>54</td>
<td>I</td>
<td>0.15%</td>
<td>0.15%</td>
<td>ROC analysis</td>
<td>I</td>
<td>RFS,OS</td>
</tr>
</tbody>
</table>

Abbreviations: post-IND=post-induction; post-CONS=post-consolidation; I=induction; C=consolidation; RFS=relapse free survival; OS=overall survival; CR=complete remission; NA=not available; I=induction; C=consolidation; EFS=event free survival; RFS=relapse free survival; OS=overall survival;
FIGURE LEGENDS

Figure 1. Pie chart depicting the molecular heterogeneity of cytogenetically normal AML. The analysis is based on mutations in the NPM1, CEBPA, MLL, FLT3 [internal tandem duplication (ITD) and tyrosine kinase domain (TKD) mutations at codons D835 and I836], NRAS, and WT1 genes. Data are derived from mutational analysis of 485 younger adult patients with cytogenetically normal AML from the German AML Study Group.90

Figure 2. Subgroup analysis of RFS and CIR of 143 AML patients stratified according to pre-treatment karyotype or FLT3 status and levels of minimal residual disease (MRD) after consolidation. In panels A and B, those with a level of residual leukemic cells < 0.035% are referred to as intermediate karyotype-MRDneg, favorable karyotype-MRDneg or FLT3wt MRDneg while those with levels ≥ 0.035% are categorized as intermediate karyotype-MRDpos or favorable karyotype-MRDpos or FLT3wt MRDpos. Survival outcomes of these subsets and of unfavorable-risk karyotype and FLT3-ITD category are shown (p <0.001 for all comparisons). In panels C and D, FLT3 wild-type patients achieving a MRD negative status show a better outcome than those who remained MRD positive after consolidation (p <0.001).

Figure 3: Risk assessment combining pre-treatment and post-treatment prognosticators. Applying an adjusted risk-stratification, combining genetics/cytogenetics and level of minimal residual disease at the end of consolidation therapy, we distinguished two categories of patients: 1) Low-risk, including favorable and intermediate-risk karyotype that were MRD negative; and 2) High-risk, including unfavorable-risk karyotype, FLT3-ITD cases and favorable/intermediate-risk karyotype that were MRD positive. The first group stands for significantly longer overall survival (73% vs 17%), relapse free survival (58% vs 22%) and cumulative incidence of relapse (17% vs 77%) (p<0.001 for all comparisons).

Figure 4. Comparison of clinical outcome in the “prospective” and “historical” cohort. The prospective cohort included 21 high-risk patients (4 MRDpos favorable-karyotype, 9 MRDpos intermediate-karyotype, 4 unfavorable-karyotype and 4 FLT3-ITD) who underwent allogeneic stem cell transplantation (8 matched sibling donor, 6 haploidentical related and 7 matched unrelated/umbilical cord blood donor). The historical cohort accounted for 36 high-risk patients (6 MRDpos favorable-karyotype, 23 MRDpos...
intermediate-karyotype, 1 unfavorable-karyotype and 6 FLT3-ITD). Allogeneic stem cell transplantation was offered to 12 patients with an available matched sibling donor whereas those lacking this option were given autologous stem cell transplantation (n=24). With a median follow-up of 18 months, survival estimates were significantly superior for the prospective cohort as compared to the historical control (DFS 70% vs 20%, p=0.00047; OS 69% vs 24%, p=0.046).
FIGURE 3

<table>
<thead>
<tr>
<th>Low-Risk</th>
<th>High-Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Favorable-K/MRD-negative</td>
<td>Unfavorable-K FLT3-positive</td>
</tr>
<tr>
<td>Intermediate-K/MRD-negative</td>
<td>Favorable-K/MRD-positive</td>
</tr>
<tr>
<td></td>
<td>Intermediate-K/MRD-positive</td>
</tr>
</tbody>
</table>

Cumulative incidence of death

A

B

C

p=0.0001

p=0.0001

p=0.0001

Low risk

High risk
**FIGURE 4**

<table>
<thead>
<tr>
<th></th>
<th>Prospective Cohort</th>
<th>Retrospective Cohort</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS</td>
<td>69%</td>
<td>24%</td>
<td>0.046</td>
</tr>
<tr>
<td>DFS</td>
<td>70%</td>
<td>20%</td>
<td>0.00047</td>
</tr>
</tbody>
</table>

OS = 70% 69%
DFS = 70% 20%

**TABLE:**

<table>
<thead>
<tr>
<th></th>
<th>Prospective Cohort</th>
<th>Retrospective Cohort</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS</td>
<td>69%</td>
<td>24%</td>
<td>0.046</td>
</tr>
<tr>
<td>DFS</td>
<td>70%</td>
<td>20%</td>
<td>0.00047</td>
</tr>
</tbody>
</table>
Prognostic and therapeutic implications of minimal residual disease detection in acute myeloid leukemia

Francesco Buccisano, Luca Maurillo, Maria Ilaria Del Principe, Giovanni Del Poeta, Giuseppe Sconocchia, Francesco Lo-Coco, William Arcese, Sergio Amadori and Adriano Venditti