MRD in AML: does it already guide therapy decision-making?

Gert Ossenkoppele and Gerrit Jan Schuurhuis

Department of Hematology, VU University Medical Center Amsterdam, Amsterdam, The Netherlands

Prognostic factors determined at diagnosis are predictive for outcome whereas achievement of morphological complete remission (CR) is still an important end point during treatment. Residual disease after therapy may reflect the sum of all diagnosis and postdiagnosis resistance mechanisms/factors; its measurement could hypothetically be very instrumental for guiding treatment. The possibility of defining residual disease (minimal residual disease [MRD]) far below the level of 5% blast cells is changing the landscape of risk classification. In this manuscript, the various methods, all different in sensitivity, specificity, and phase of development, to assess MRD are discussed. Currently, the 2 methods mostly used are flow cytometry–based immune MRD (multiparameter flow cytometry [MPFC]) and molecular MRD assessed by real-time quantitative polymerase chain reaction. Both have advantages and disadvantages that are summarized in detail. Many studies in children as well as adults already demonstrated that MRD detection by MPFC or molecular MRD provides strong prognostic information in acute myeloid leukemia (AML) after both induction and consolidation. These studies are summarized in this review. The general conclusion of this review is that a better definition of disease burden than morphological CR is now emerging. MRD assessed by flow or molecular techniques should become standard in every clinical trial in AML. Harmonization of antibody panels, introduction of single-cell tube systems (for determination of residual leukemic stem cells), and standardized analytical programs will pave the way for individual risk assessment and become a surrogate end point for survival in studies investigating new drugs, hopefully resulting in faster drug approval in AML.

Learning Objectives

- To have knowledge of the various methods available to assess MRD and their appropriate application
- To know the clinical utility of MRD

Acute myeloid leukemia (AML) is a heterogeneous disease characterized by a multitude of molecular abnormalities. Better understanding of this complex mutational landscape has not resulted in dramatic changes in treatment in the last decades.[1] At diagnosis, several factors have prognostic impact for outcome, although achievement of morphological complete remission (CR) is still an important end point. There is no cure without CR. The current definition of remission is based on those established decades ago. The morphologic definition of CR, that is, bone marrow (BM) should contain <5% blast cells (and neutrophil count $>1.0 \times 10^9$/L and platelets $>100 \times 10^9$/L) is practical, cheap, and clinically relevant: below the 5% cutoff level, it defines a group of patients performing relatively well compared with those above the cutoff level.

Currently, the term remission now also includes CRp (CR with platelets $<100 \times 10^9$/L) and CRi (all criteria for CR except for neutrophil and/or platelet recovery). It has been shown that CRi and CRp are associated with shorter survival, underlining the poorer quality of these types of remission.

Independent of the definition of CR, the majority of CR patients relapse within a few years after diagnosis. A multitude of factors at

diagnosis, including clinical parameters and cytogenetic as well as molecular factors and biological properties of the leukemic cells, have important prognostic impact for outcome in the whole patient population. Such risk factors were shown to correlate with quality of remission reflected by frequencies of residual disease.[2-5]

The outcome of AML treatment is highly variable and not individually predictable. It thus seems that prognosticators at diagnosis will not be able to reach the ultimate goal of individualized risk assessment. One important issue is that the prognostic impact of these factors in the present risk groups does not take into account the contribution of several factors such as cellular resistance mechanisms at diagnosis, and moreover, postdiagnosis factors, which include dosage, compliance, pharmacological resistance, and probably other unknown features. The only postdiagnosis prognostic factor that is now included in risk stratification is the CR status after the first course of induction remission therapy. Patients achieving late CR have a more dismal outcome. The earlier-mentioned factors, which corroborate proper risk classification, are only partly covered by inclusion of CR status. To illustrate the complexity of risk stratification, on the one hand relapses occur in very-good-risk patients, but, on the other hand, cures are possible in very-poor-prognosis AML. Because residual disease after different cycles of therapy may reflect the sum of all diagnosis and postdiagnosis resistance mechanisms/factors, its measurement could hypothetically be very instrumental for guiding treatment.

The possibility of defining residual disease far below the level of 5% blast cells is changing the landscape of risk classification. This

so-called minimal residual disease (MRD; also referred to as measurable residual disease) approach at present establishes the presence of leukemia cells down to levels of 1:1000 to 1:10[6] white blood cells (WBCs), compared with 1:20 for morphology.[6]

Different platforms are available for assessing residual disease in AML including traditional light microscopy, fluorescence in situ hybridization, cytogenetics, multiparameter flow cytometry (MPFC), including leukemia stem cells (LSCs), real-time quantitative polymerase chain reaction (RT-qPCR), and next-generation sequencing (NGS), all different in sensitivity, specificity, and phase of development.[6]

Morphological assessment of blast percentage is highly unreliable because of limited sensitivity and interobserver variability. Striking discrepancies between standard morphology and flow cytometry–determined CR status have been clearly shown: part of BM classified as CR by morphology appeared to have in excess of 5% blasts by MPFC, whereas part of morphologically assessed non-CR cases, probably due to BM regeneration, in fact showed no or very few leukemic blasts by MPFC.[7] An urgent need for a more sensitive, specific, and objective assessment of MRD than classical morphology is obvious in order to develop a more individualized approach of treatment decisions.

Methods for MRD assessment
MRD detection by PCR
RT-qPCR allows MRD detection in cases with chimeric fusion genes generated by balanced chromosomal rearrangements. Common targets are PML-RARA/t(15;17) RUNX1-RUNX1T1/t(8;21) CBFB-MYH11/inv(16)/t(16;16) DEK-CAN (NUP214)/t(6;9), t(11q23)/MLL fusions, t(5;11)/NUP98-NSD1 patients. Other genetic alterations that can be used for MRD detection include: (1) insertions/duplications (eg, NPM1, FLT3-ITD, MLL-PTD); (2) point mutations (CEBPA, IDH1/2, KIT, RAS, RUNX1, TP53, JAK2); (3) gene overexpression (WT1, EVI1, ERG).[8,9]

Altogether, in about 50% of patients, PCR assessment of MRD is, in principle, possible.

Apart from t(15;17), RUNX1-RUNX1T1, and CBFB-MYH11, currently, NPM1 is the best-validated molecular marker for MRD assessment.[10] Overexpression of WT1 has been proven to be informative to predict outcome in selected cases.[11] This assay will not be generally applicable because of limited sensitivity, lack of specificity, and better alternatives.

Real-time qPCR (RT-qPCR) is an advantage over RT-PCR for MRD detection because it is more reliable and can be readily standardized. Moreover, quantitative assays with the capacity to measure the absolute level of leukemic transcripts allow for the assessment of whether these levels are rising or falling, which is extremely important to inform therapy. This methodology has been standardized for several molecular markers for clinical implementation in the Europe Against Cancer (EAC) program.[12] It is most likely that appropriate cutoff levels will differ between the various genomic aberrations.

A major drawback of RT-qPCR is that it can only be applied in around 50% of AML cases. MPFC, on the other hand, has been shown by many authors to enable MRD assessment in the vast majority of the AML patient population.

MRD detection by MPFC
There are 2 main approaches for detecting MPFC-MRD: (1) using leukemia-associated immune phenotypes (LAIPs)[13] and (2) identifying "different from normal" patterns.[14,15]

The use of LAIP. The basic principle is to identify, at diagnosis, immunophenotypically aberrant populations (LAIPs) that differ from the majority of normal hematopoietic cells, and to use these to trace residual leukemia after treatment. It is important to stress that these LAIPs consist of normally occurring marker/marker combinations, that are present in low or very-low frequencies in normal and regenerating BM. Different types of aberrancies can be distinguished; the major are cross-lineage expression of antigens, antigen overexpression, lack of antigen expression, and asynchronous expression of antigens. LAIPs generally consist of the pan leukocyte marker CD45, a primitive marker (CD34, CD117, CD133), a myeloid antigen (CD33, CD13), together defining normally occurring populations, and 1 or more markers/marker combinations defined as aberrantly expressed, and which may include the above-mentioned markers too. Although low, the background levels of LAIPs in normal and regenerating BM prevent specific detection of aberrancies with sensitivities higher than 1:10 000. To cover all possible aberrancies, this requires, for each new AML patient, the use of an extensive panel of monoclonal antibodies. The number of different LAIPs in AML may amount to about 100 as reported in a recent study.[16]

Identifying "different from normal" patterns. Essentially, aberrant patterns of differentiation are recognized and translated into quantifiable aberrant cell populations (in fact LAIPs). It uses a standard fixed antibody panel to recognize leukemic cells based on their difference with normal hematopoietic cells at all stages of disease/treatment. The advantage is that it does not a priori restrict MRD assessment based on LAIPs defined at diagnosis only, thereby taking into account possible immunophenotype "shifts".

Newer technologies
Digital PCR
Digital PCR (dPCR) is a refinement of conventional PCR methods that can be used to directly quantify and clonally amplify nucleic acids. The key difference between dPCR and traditional PCR lies in the method of quantifying nucleic acids. PCR carries out 1 reaction per single sample. dPCR also carries out a single reaction within a sample, however, the sample is separated into a large number of partitions and the reaction is carried out in each partition individually. This separation allows a more reliable collection and sensitive measurement of nucleic acid amounts and does not need a standard curve. It has been shown to be a reliable tool for MRD assessment in lymphoid malignancies with greater applicability and reduced labor intensiveness than RT-qPCR.[17] Also in NPM1[+] AML, dPCR has been shown to be applicable for a large variety of NPM1 mutation subtypes without the need for plasmid standards.[18] It is also reliable for identifying patients at risk during follow-up.[18]

Next-generation sequencing
In a French study assessment of MRD in the NPM1-mutated genetic subgroup, IDH1/2 and DNMT3a mutations were quantified by NGS.[19] IDH1/2 mutations were reliable markers for prediction of relapse in 100% of cases whereas DNMT3a mutations were not correlated to relapse after long follow-up. Persistence of this mutation in 40% of the patients on the one hand fits in the concept of clonal hematopoiesis, but, on the other hand, makes DNMT3a less suitable for specific MRD detection.

Table 1. Eight-color LSC detection tube

Tube	FITC	PE	PerCP-CY5.5	Pe-Cy7	APC	APC-H7	BV421	HV500c
1	CD45RA	CLL-1 TIM-3 CD7 CD11b CD22 CD56	CD123	CD33	CD38	CD44	CD34	CD45

APC, allophycocyanin; FITC, fluorescein isothiocyanate; PE, phycoerythrin; PerCP, peridinin chlorophyll.

RUNX 1 is a frequently recurring mutation in AML. Because it lacks a mutational hotspot it requires patient-specific, RT-qPCR which, due to labor intensiveness and costs, is not attractive for MRD application. In a recent series, this mutation was detected by deep sequencing in 25.9% of 814 patients.[20] The prognostic impact of residual levels of this mutation, based on separation according to median mutation burden, resulted in significant differences in outcome.

Heuser and colleagues used NGS to detect NPM1 and FLT3-ITD mutations at follow-up in remission. Parallel assessment of MRD by NGS and RT-qPCR in NPM1-mutated patients was concordant in 95% of analyzed samples.[21] Many other mutations (like, eg, TET2, ASXL1) detected by NGS can probably be used for MRD assessment, however, clinical data are currently lacking. A possible risk is that these mutations can be present in a substantial number of cases in the premalignant lesion implying that one has to be assured that MRD positivity represents recurrence of the disease. NGS sensitivity reflects a promising tool to assess MRD, although it is currently applicable for only a restricted number of mutations. Although promising, various hurdles have to be taken, implicating that introduction for MRD assessment in clinical practice will not be realized in very short time. NGS will probably refine our MRD measurements further, however, standardization is currently lacking and the sensitivity level is about 1% (dependent on depth of sequencing) which cannot compete with other MRD measurements techniques discussed in the previous sections.

False-positivity and false-negativity in MRD
Careful examination of the studies published thus far reveals that, in most cases, a single cutoff level that defines "MRD+" and "MRD−" patient groups results in a small proportion of patients who, while classified as MRD+, remain in CR. This may, at least in part, be related to persistence of preleukemic clones, for example, for a molecular aberration like t(8;21). Clinically, false-positivity may lead to overtreatment, which is especially important as to the decision to treat or not with allogeneic transplantation. Increasing the threshold level that defines positivity usually results in better identification of patients at risk, but at the same time increases the number of MRD− patients with poor prognosis. However, in most studies, with cutoff levels in the order of 0.1%, already 20% to 40% of the patients classify as MRD false-negative. Further increase of these percentages would unacceptably decreases the value of MRD-based risk stratification. From this, one might argue that MRD-based risk stratification may ultimately benefit from the introduction of 2 instead of 1 cutoff level, thereby defining 3 patient groups in which different prognostic impact could then be combined with the other well-known prognostic parameters. Given that, it is clearly of importance to better identify the possible underlying causes of MRD negativity. There are several: (1) sensitivity of the assay, (2) phenotypic shifts, (3) differences in time of onset of outgrowth to relapse, (4) other biologically important factors

such as frequencies of residual LSCs, and (5) quality of the aspirate. These causes are explored in further detail in the following sections.

Sensitivity of the assay
As referred to previously, assay sensitivity is restricted by its relatively low specificity.

Phenotypic shifts
Initially present aberrancies used for MRD detection can change or disappear while the leukemia-initiating clone/cell population remains (reviewed in Zeijlemaker et al[22]).

Differences in time of onset of outgrowth to relapse
Such differences are not adequately covered by assessing MRD at 1 particular time point after therapy.

Other biologically important factors such as frequencies of residual LSCs
LSCs are thought to be chemotherapy resistant and thereby to be at the basis of outgrowth of MRD cells to overt relapse. Different cellular compartments may contain LSCs. Based on relative therapy resistance, ability to grow out in different mouse models, and correlation with leukemic engraftment,[23-25] as well as prognostic strength at diagnosis,[26] $CD34^+CD38^-$ LSCs qualify best for assessment under MRD conditions, at least in $CD34^+$ AML cases. Identification of LSCs is challenging because of the very low frequency, down to 1 in 5×106 WBCs. $CD34^+CD38^-$ LSCs often aberrantly express a plethora of cell surface markers in highly heterogeneous patterns with, similar to MRD, immunophenotype changes between diagnosis and relapse. For these reasons, similar to MRD, extensive antibody panels are required. Thanks to the high specificity of LSC detection,[27] it was possible to replace a multitube 8-color antibody panel by a single 1 tube 8-color LSC assay (Table 1). This enables $CD34^+$ $CD38^-$ LSC detection in a broadly applicable, less expensive, and more efficient manner than current detection strategies. A further advantage is that a more objective analysis is possible with much less requirement of extensive experience of normal and leukemic BM differentiation patterns as is necessary for MPFC MRD. It remains to be established whether both MRD and LSC assays are necessary for optimal prognostification.

Quality of the aspirate
MRD levels in peripheral blood (PB) are lower than in BM so contamination with blood, when not sending in the first tap for examination but continuing to aspirate from the same site, will result in hemodilution and cause underrepresentation of the number of leukemic cells. This has been clearly shown in childhood acute lymphoblastic leukemia (ALL). For an optimal MRD assessment, it is required that aspirates are taken from separate puncture sites. Unequal infiltration of the BM space, BM fibrosis, and adhesive properties of blasts can all cause false-negative

MRD results. Also, a heterogeneous response on therapy in the BM compartment may cause false MRD negativity. Positron emission tomography, using the proliferation marker 3′-deoxy-3′-[^{18}F]fluorol-thymidine, for early assessment of treatment response in patients with AML, showed a marked heterogeneity in response especially in patients with refractory disease.[28] An aspirate for assessment of response is typically a single point measurement. Consequently, this might sample a region negative for residual disease despite the presence of residual disease in other parts of BM. In 1987, Martens and Hagenbeek showed in a rat model that distribution of leukemic cells after chemotherapy was heterogenous, which contrasted to the homogeneous distribution at diagnosis.[29] Theoretically, the number of MRD cells in peripheral blood may better represent the overall leukemic burden at different BM sites.

PB MRD vs BM MRD

PB to assess MRD would thus be an attractive alternative reliable source for MRD detection, also for specific purposes allowing a frequent sampling, facilitating sequential MRD monitoring. Sensitivity for MPFC is lower,[30,31] whereas specificity may be higher as compared with the BM MRD assay.[30] The latter is likely caused by a much lower frequency in PB compared with BM of "background" normal progenitor populations. Studies in ALL have already proven that assessment of BM MRD can be replaced by PB MRD.

In addition, molecular MRD (in patients with NPM1 mutant, core-binding factor [CBF], or with overexpressed WT1) on PB paralleled BM assessment. For MPFC, Maurillo et al found promising results of PB MRD in a series of samples from 50 patients.[31] Recently, in a larger study (378 paired samples in 114 patients), MPFC-based MRD significantly correlated between PB and BM ($r = 0.67$, $P < .001$).[30] In 78 patients, it was shown that the cumulative incidence of relapse (CIR) 1 year after induction therapy was 29% for PB MRD$^-$ patients and 89% for MRD$^+$ patients ($P < .001$). Although less sensitive, the high specificity of PB MRD may thus outweigh this disadvantage, offering a promise for a prominent role in future clinical treatment decisions.

In conclusion, it is clear that various methods to assess MRD are available. Some are more standardized, others are less, although a few are still in development phase. Concerning the molecular and flow cytometric MRD, a broad combined US and European initiative to generate technical as well as clinical recommendations for MRD has been undertaken under the umbrella of European LeukemiaNet (ELN) and will be finalized soon.

Clinical aspects

Although the concept of MRD negativity as an indicator for the quality of treatment response is the same in AML and other hematological diseases such as chronic myeloid leukemia, multiple myeloma, and ALL, application of MRD assessment in AML has lagged behind.

Standard therapy produces CRs even in patients whose cytogenetic/molecular features suggest a poor outcome. However, the quality of the remission measured by MRD affects its duration. Incorporation of on-treatment data is becoming increasingly important. It is now generally accepted that levels of MRD assessed at particular time points during treatment offer an independent biomarker to predict outcome. Most of the available data have been derived from clinical studies with intensive AML treatment. The predictive value of MRD assessment in less-intensive treatment schedules is not studied very extensively. At least the timing of assessment will be different. From a cost-effectiveness perspective, we would recommend performing the assessment after 2 cycles of intensive chemotherapy to predict outcome.

Selected data from clinical studies using molecular MRD

Acute promyelocytic leukemia

MRD assessment has been introduced as a component of the standard response criteria in acute promyelocytic leukemia (APL). Grimwade et al showed in a large Medical Research Council (MRC) study with 406 APL patients that MRD assessed using the standardized EAC PML-RARA RT-qPCR proved to be more powerful independent predictor for relapse as compared with WBCs that usual dictates treatment.[32] Given the high survival rates in low-risk APL, sequential monitoring is most useful in high-risk APL and relapsed APL. BM is the recommended source for monitoring because PB has a 1.5 log lower sensitivity.

CBF leukemias

RUNX1-RUNX1. MRD assessment by RT-qPCR in AML with the RUNX1-RUNX1 transcript has been shown to be important. Molecular remission or a significant reduction in RUNX1-RUNX1 transcripts is strongly associated with a favorable outcome as shown in several studies. In a study of 96 t(8;21) AML patients, high WBC count, the presence of KIT mutations and of FLT3-ITD and FLT3-TKD, as well as a <3-log MRD reduction in univariate analysis showed a higher hazard rate for relapse.[33] However, in multivariate analysis, MRD reduction remained the only prognostic factor. At 3 years, the CIR was 22% vs 54% in patients who achieved ≥3-log MRD reduction vs the others. These data were confirmed by Yin et al who, in a series of 163 patients taken from the MRC AML-15 trial, showed that >3 log reduction in RUNX1-RUNX1 transcripts in BM after 1 cycle of induction therapy was the strongest prognostic variable for relapse in multivariate analysis: relapse rate was only 4%, compared with >30% for those who did not reach this threshold.[34] The fact that this did not translate into a better overall survival (OS) can be explained by excellent rescue, mostly by allogeneic stem cell transplantation (alloSCT), after relapse. In all studies, rising MRD levels, assessed by serial monitoring, predicted relapse.

CBFB-MYH11. Similar to AML, t(8;21) assessment of CBFB-MYH112 studies has been shown to be very valuable for predicting outcome. In 98 patients with either inv(16)(p13.1;q22) or t(16;16) (p13.1;q22), Jourdan et al showed that MRD level after the second consolidation cycle of chemotherapy was the only significant prognostic parameter in multivariate analysis for relapse-free survival (RFS).[33] Similarly, for inv16 AML as well, MRD proved to be a very good predictor of relapse: Yin et al identified cutoff MRD thresholds in BM (>50 copies) and PB (>10 copies) with a 100% relapse rate.[34] In this study, it was also shown that there was a significant correlation between BM and PB levels in both groups of CBF AMLs during follow-up. Based on their data, serial monitoring should be performed at 3-month intervals during follow-up, creating a 3- to 4-month window for preemptive treatment.

Importantly, in t(8;21) patients, MRD negativity is not a condition sine qua non in contrast to inv(16) patients.

NPM1-mutated AML

Mutations of the NPM1 gene are among the most common molecular aberrancies in AML, present in roughly 30% of all AML and approaching 60% in normal karyotype AML. The prognosis is dependent on the presence or absence of additional mutations (FLT3-ITD and DNMT3A): patients with FLT3-ITD or mutated DNMT3A have a poorer outcome as compared with NPM1 single

Table 2. Clinical studies on the prognostic value of MRD in adult and childhood AML by MPFC

Ref.	Multi-center	Study pop.	% LAIP	No. of pts	MRD measurement after	Cutoff MRD level			Univariate analysis significant for	Multivariate analysis significant for	Study details
						I	C	Post-Tx			
44	N	A	46	53	I, C	0.05%	0.2%		RFS, OS	RFS	
45	Y	A	70	56	I, C	0.045%	0.035%		I; -C: RFS, OS	I; -C: RFS, OS	
46		A	75	126	I	<0.01%, 0.01%-0.1% 0.1%-1% >1%			RFS, OS	RFS	MRD >1%: 3-y RR: 85% MRD 0.1-1.0%: 3-y RR: 45% MRD 0.01%-0.1%: 3-y RR: 14% MRD <0.01%: 3-y RR: 0%
47	N	Ch	85	46	I₁, I₂	0.1%			I₁: RR, OS I₂: RR, OS	I₁: OS I₂: OS	
48		A	79	31	Pre Tx				RR after Tx		
15	Y	Ch	?	252	I₁	0.5%	0.035%		RFS, OS	RFS, OS	3-y OS 69% (MRD⁻) vs 41% (MRD⁺)
49	Y	A	100	106	Day 16	Log diff. 2.11			CR, EFS, RFS, OS	EFS, RFS	
50	N	A	100	62	I, C	Log diff. 2.11	Log diff. 2.53		I: RFS C: RFS, OS	I: RFS C: RFS	
51	N	A	100	72	I₁, I₂, C, PBSCT	I₁: 1% I₂: 0.14%	0.11%		I₁, I₂, C, PBSCT; RFS, OS	I₁, I₂, C, PBSCT; RFS, OS	
52*		A	89	100	I, C	0.035%	0.035%		I and C: RR, RFS, OS	I: -C: RR, RFS, OS	5-y RFS 72% (MRD⁻) vs 11% (MRD⁺)
53	N	A	94	45	I, C	0.05%-0.1%	0.05%-0.1%		I: RFS C: RFS		5-y RFS 85% (MRD⁻) vs 42% (MRD⁺)
54	Y	A, Ch	?	150	Day 15, I₁, I₂, C	0.1%-2%	0.1%-1.3%		Day 15, I; RFS		MRD similar EFS as traditional risk factors
55	N	A	?	41	Day after Tx			0.1%	RFS, OS	RFS, OS	
56	Y	A	?	142	I, C	0.035%	0.035%		I and C; RFS, OS	I and C; RFS, OS	5-y RR 60% (MRD⁺) vs 16% (MRD⁻)
57	N	A	94	54	I, C	0.15%	0.15%		I: RFS, OS C: RFS, OS	I: RFS, OS C: -	
58	Y	Ch	?	94	I₁, I₂, C, end of Tx	<0.1% 0.1%-0.5% >0.5%			I₁: RFS, OS	I₁: RFS, OS	3-y RFS 64% (MRD⁺) vs 14% (MRD⁻)
14	Y	Ch	100	188	I₁, I₂, end of Tx	>0%,	0-1%		I₁: OS, RFS I₂: RFS, OS	I₁: OS, RFS I₂: RFS, RR	RR at 3 y 60% vs 29%
7	Y	Ch	?	203	I₁, I₂, end of Tx	<0.1%, 0.1%-1% >1%			I₁: EFS, RFS I₂: EFS, RFS	I₁: EFS, RFS I₂: EFS, RFS	Morphological assessment has limited value in comparison with flow cytometry.
16	Y	A	89	517	I₁, I₂,	0.1%	0.1%		I₁: RFS, OS I₂: RFS, OS	I₁: RFS, OS I₂: RFS, OS	Cutoff points between 0.05 and 0.8 are all significant.
38	Y	A	93	427	I₁, I₂,	0.1%	0.1%		I₁: RFS, OS I₂: RFS, OS	I₁: RFS, OS I₂: RFS, OS	3-y OS 38% (MRD⁺) vs 18% (MRD⁻) after cycle 2

?, not known; A, adult; C, consolidation; Ch, children; DFS, disease-free survival; diff., difference; I, induction treatment; I₁, induction cycle 1; I₂, induction cycle 2C; N, no; PBSCT, peripheral blood stem cell transplantation; Pop., population; pts, patients; Ref., reference; RR, relapse risk; Tx, transplantation; Y, yes.

*Includes 56 patients previously reported by Venditti et al.[45]

Table 2. (continued)

Ref.	Multi-center	Study pop.	% LAIP	No. of pts	MRD measurement after	Cutoff MRD level			Univariate analysis significant for	Multivariate analysis significant for	Study details
						I	C	Post-Tx			
39	N	A, Ch	100	253	Pre Tx	0.1%					MRD predictive in CR1 and CR2
59	Y	A	?	210	I, C	0.035%	0.035%		I, C: DFS, OS	I, C: DFS, OS	MRD negativity gives 5-y DFS: 57% vs 13% in elderly AML
60	N	A	100	359	Pre Tx	0.1%				OS, PFS, RFS	3-y RR 67% (MRD⁺) vs 22% (MRD⁻)
61	Y	A	100	306	At the time of morphologic CR	<0.01%, 0.01%-0.1% >0.1%			RFS	RFS	Multivariate analysis revealed MRD, age, and cytogenetics as independent variables. Cytogenetics and MRD are complementary in a scoring system.
62	Y	Ch	78	101	Day 15, pre C	0.1%	0.1%		Day 15: EFS, OS Pre C: EFS, OS	Day 15: EFS, OS Pre C: EFS, OS	EFS at 5 y 65% (MRD⁻) vs 22% (MRD⁺)

?, not known; A, adult; C, consolidation; Ch, children; DFS, disease-free survival; diff., difference; I, induction treatment; I_1, induction cycle 1; I_2, induction cycle 2C; N, no; PBSCT, peripheral blood stem cell transplantation; Pop., population; pts, patients; Ref., reference; RR, relapse risk; Tx, transplantation; Y, yes.
*Includes 56 patients previously reported by Venditti et al.[45]

mutated as well as non FLT3ITD cases. The NPM1 mutations are heterogeneous (>50 reported) but provide an ideal leukemia-specific target for MRD detection by RT-qPCR because 3 types of mutations (A, B, D) account for 90% of NPM1-mutated cases. This contrasts to RUNX1 mutations as discussed previously. Krönke et al showed in 245 patients that, after double induction, RT-qPCR negativity for NPM1 had a CIR of 6.5% at 4 years vs 53% for NPM1⁺ patients.[10] This was confirmed in retrospective studies.[9,35] Also, rising levels of NPM1 after chemotherapy or alloSCT were predictive for relapse.

Recently, Ivey et al showed the complexity of NPM1-mutated AML. By molecular profiling >150 subgroups could be distinguished, of which, in univariate analysis, only the presence of FLT3-ITD and DNMT3A mutations were associated with outcome.[36] In multivariate analysis, only MRD status (measured by RT-qPCR of NPM1) remained as a significant prognostic factor. The median sensitivity of the assay was 1×10^{-5}. Risk of relapse at 3 years was 30% in case of an absence of NPM1 transcripts vs 82% in those with detectable transcripts. On sequential monitoring, relapse could be predicted by rising levels of NPM1 transcripts.

Other targets
Other targets include (1) Wilms tumor 1 (WT1) gene expression, (2) MLL-MLLT3, and (3) FLT3-ITD, explored in further detail in the following sections.

WT1 gene expression. WT1 is not leukemic-specific, limiting the capacity to distinguish low-level MRD from expression in normal PB and BM cells. However, in an ELN study, it was shown to be possible to assess at least a 2-log reduction in transcript level in about 45% of the cases which provided independent prognostic information. Because of lack of specificity and limited sensitivity, and because more suitable tools are available for MRD detection, WT1 is not a likely candidate for routine MRD assessment.[11]

MLL-MLLT3. MLL-MLLT3 is very rare in adults (2% of AMLs) in contrast to childhood AML where frequencies of over 10% have been reported. A sensitive RT-qPCR for different fusion products has been developed and found useful for MRD assessment to predict outcome. MRD negativity by this assay showed a very low relapse rate of 11% and an OS of 70% at 48 months while all MRD⁺ patients relapsed and died.[37]

FLT3-ITD. FLT3-ITD mutations are present in 25% to 30% of AML patients and are correlated with poor outcome. Because of technical limitations (necessity for clone-specific primer/probes for patient-specific QR-PCR) and instability of the marker, the clinical applicability for MRD assessment is currently limited.

In conclusion, at present for routine clinical application of molecular MRD assessment in adult AML, it is only the CBF and NPM1⁺ AML cases that are good candidates.

Selected data from clinical studies using MPFC
Many studies in children as well as adults already demonstrated that MRD detection by MPFC provides strong prognostic information in AML after both induction and consolidation therapy (summarized in Table 2). Importantly, one has to realize that most of these studies were retrospective and in a single institute setting, thus introducing potential bias. Marker panels and instrumentations varied. By identifying cutoff values in the order of 10^{-3} to 10^{-4}, it has been possible in most of
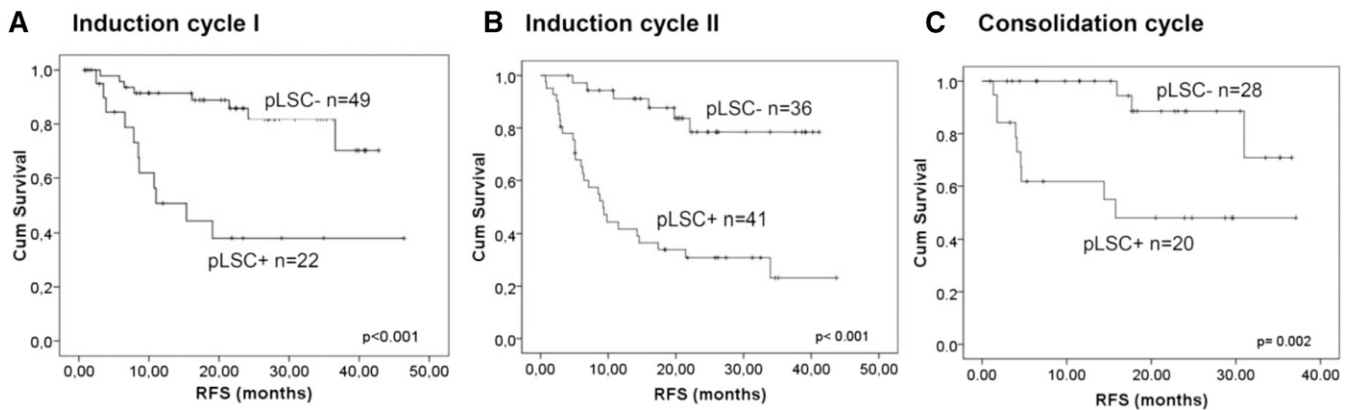
Figure 1. Prognostic value of frequencies of CD34+CD38− putative LSC (pLSC) compartment at follow-up. The Kaplan-Meier analyses for RFS for the CD34+CD38− pLSC compartment at follow-up for 3 consecutive therapy cycles are shown. The optimal cutoff levels were chosen to define high pLSC frequency (pLSC+) and low pLSC frequency (pLSC−) after first induction cycle (0.0003%, which is 3 pLSCs in 1 000 000 WBCs) and after second induction cycle and consolidation therapy (0.0001%, 1 pLSC in 1 000 000 WBCs). After the first induction cycle (A, 71 patients), second induction cycle (B, 77 patients), and after consolidation therapy (C, 48 patients), pLSC+ patients showed significantly poorer outcome as compared with pLSC− patients.

these studies to identify 2 patient groups with either relatively poor or relatively good prognosis. Determining the optimal cutoff is mostly done by empirically setting a threshold below or above which patient outcome is statistically significantly different. In part of these cases, the thresholds found were validated in a prospective setting. Various statistical methods have been used, such as maximally selected log-rank statistics or receiver operating characteristic analysis. A couple of studies have now also been performed in a multicenter and, in 1 instance, also a multinational setting. Two of the largest studies are summarized here.

The HOVON group

The HOVON group established the value of immunophenotypic-determined MRD in patients younger than 61 years.[16] MRD was evaluated in BM in 389 patients available for analysis. After all courses of therapy, low MRD values distinguished patients with relatively favorable outcome from those with high relapse rate and adverse RFS and OS. In the whole patient group and in the subgroup with intermediate-risk cytogenetics, MRD was an independent prognostic factor. RFS at 4 years was 23% in MRD+ patients vs 52% in MRD− patients and the rate of relapse at 4 years was 72% vs 42%, respectively. Multivariate analysis after cycle 2, when decisions about consolidation treatment have to be made, confirmed that high MRD values ($>0.1\%$ of WBCs) were associated with a higher risk of relapse after adjustment for consolidation treatment time-dependent covariate risk score and early or later CR.

The UK MRC group

The UK MRC group assessed MRD in 427 patients older than 60 years. MRD negativity conferred significantly better 3-year survival from CR (42% v 26% in MRD+ patients after cycle I, and 38% v 18%, respectively, after cycle II).[38] Reduced relapse and higher risk of early relapse in MRD+ patients (median time to relapse, 8.5 vs 17.1 months, respectively) was shown. In multivariate analysis, MRD status was an independent prognostic factor, identifying a subgroup of intermediate-risk patients with poor outcome.

Collectively, these studies showed that low levels of MRD were associated with improved survival and lower risk of relapse. Many studies included a multivariate analysis showing that MRD was an independent predictor of RFS, OS, event-free survival (EFS), or combinations, superior to other well-defined prognostic factors such as AML type, age, WBC count at diagnosis, and classification of cytogenetic risk.

Pretransplant MRD

Evidence is accumulating that presence of MRD pretransplant is a powerful predictor of outcome after alloSCT in AML. Most of the available data comes from the Walter group at Fred Hutchinson Cancer Research Center (FHCRC; Seattle, WA). They showed that in 253 patients transplanted after myeloablative therapy in CR1, the 3-year OS was 73% for pretransplant MRD− patients vs 32% for pretransplant MRD+ patients and 73% vs 44%, respectively, for patients in CR2. Relapse rates were 21% for MRD− patients vs 58% for MRD+ CR1 patients and 19% vs 68% for CR2 patients. In a recent update, Araki et al[60] showed that in 359 adults, the 3-year relapse rates were 67% in MRD+ as compared with 22% in MRD− patients, resulting in an OS of 26% vs 73%, respectively.[39] These authors also showed that MRD status has the same predictive value in the nonmyeloablative transplant setting. Although this conclusion was based on retrospective data, this study does not support the preferential use of myeloablative conditioning in case of pretransplant MRD positivity. The quality of the response preceding transplantation seems to be the most important predictor of transplant outcome. The same group presented data that MRD positivity before myeloablative cord blood transplantation in patients has no impact on relapse rate and survival, indicating that the greater graft-versus-leukemia effect of this procedure can overcome the persistence of residual disease present before transplantation.

The recent finding that conversion of MRD positivity pretransplant to MRD negativity after myeloablative conditioning does not improve relapse rate or OS is disappointing.

Obviously, additive peritransplant strategies, like posttransplant epigenetic or immunotherapeutic approaches, are necessary to improve outcome.

LSCs as predictive marker for outcome

Treatment failure in AML is most probably caused by the presence at diagnosis of leukemia-initiating cells, also referred to as LSCs, and their persistence after therapy. We developed flow cytometric
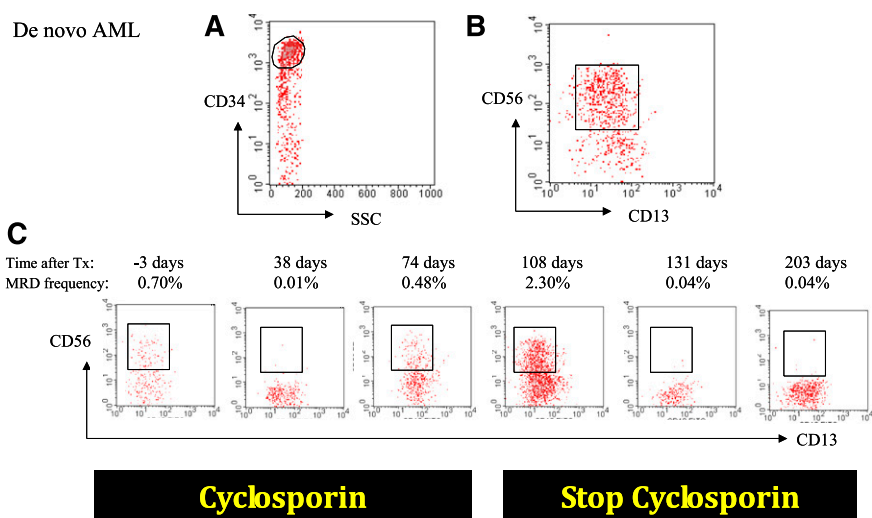
Figure 2. Sequential MRD monitoring in an allogeneically transplanted patient. De novo AML patient is characterized by a CD34$^+$ population (A) with an aberrant expression of CD56 (B). Three days before transplantation, the percentage of MRD is high (0.7%) which is known to be highly predictive for a fast relapse; 38 days after allogeneic transplantation, the percentage of MRD has dropped down to 0.01%, but unfortunately rises to 2.3% at day 108 after transplantation. Stopping immune suppression at that time point led to a quick decrease of MRD in the next months (0.04% at days 131 and 203), probably due to the graft-versus-leukemia effect resulting from discontinuation of immune suppression.

methods using LSC-associated markers and newly defined aberrancies to identify LSCs and to distinguish these from normal hematopoietic stem cells (HSCs).

At diagnosis, the frequency of the thus-defined neoplastic part of the CD34$^+$CD38$^-$ putative stem cell compartment had a strong prognostic impact.[26] After different courses of therapy too, higher percentages of neoplastic CD34$^+$CD38$^-$ cells in CR strongly correlated with (shorter) patient survival (Figure 1). Discrimination between putative LSCs and HSCs in this study thus allowed for demonstration of the clinical importance of putative CD34$^+$CD38$^-$ LSCs in AML. Moreover, combining neoplastic CD34$^+$CD38$^-$ frequencies with frequencies of MRD cells, which reflect the total neoplastic burden, revealed 4 patient groups with different survival. As discussed before, a single-cell tube has been constructed that is now prospectively validated in a large HOVON study.[30]

Current status of clinical application of MRD
MRD negativity is becoming accepted as a potential surrogate for clinical benefit in some hematological malignancies like chronic lymphocytic leukemia, ALL, and APL. For AML, this is still doubtful because of the heterogeneity of the disease and the complexity of the MRD assessment.

Two nonrandomized studies point clearly in the direction that treatment intensification guided by MRD measurement improves the outcome; they are explored in further detail in the following sections.

Risk-directed study
Rubnitz et al applied risk-directed therapy in childhood AML.[40] Risk was determined by risk profile at diagnosis and level of MRD after the first cycle of chemotherapy. MRD was measured by flow cytometric determination of LAIP. Levels of MRD were used to intensify treatment (addition of Mylotarg) and timing of the second induction cycle. The outcome was superior to other comparable trials performed and reported in childhood AML, suggesting that this success was due to this risk-stratification strategy based on MRD findings.

Risk-adapted study
Zhu et al performed a risk-adapted nonrandomized study in CBF AML.[41] Patients were assigned to alloSCT after the second consolidation treatment if MRD$^+$. In case of MRD negativity, the treatment consisted of chemotherapy/autologous SCT. Some patients decided not to follow the assigned treatment based on the MRD status. It was shown that the patients receiving treatment other than that assigned by the risk status did worse than those who completed their assigned treatment. Although the authors showed that both groups were balanced, a bias cannot be excluded in this retrospective study design.

It is clear that randomized studies are lacking, showing unambiguously the clinical benefit that can be achieved with risk-adapted MRD-based treatment. At least the UK MRC AML18 trial is trying to answer this question by randomizing for MRD assessment.

However, although perhaps premature, many trial groups and individual centers start to guide treatment based on MRD status. The decision to proceed with an alloSCT can be adapted by taking the MRD status into account. Also posttransplantation monitoring of MRD could guide the prescription of immune-suppressive drugs (see Figure 2).

Whether MRD can be used as a surrogate end point for survival is still debatable. However, once established, it would be very helpful, not only to redefine risk groups, but also in the evaluation of new drugs by offering the possibility of faster drug approval or, in contrast, stopping development of drugs and treatment strategies that are ineffective and suboptimal. Currently, 2 studies may serve as examples of how MRD can be used as a surrogate for survival end points; they are explored in the following sections.

Daunorubicin study
Prebet et al showed that, in CBF-AML, a higher dosage of daunorubicin (90 mg/m^2, 3 days) has a better clinical outcome than the standard dosage (60 mg/m^2, 3 days), which was associated with a significantly lower level of MRD.[42]

Gemtuzumab ozogamicin study

In the ALFA-0701 trial, the addition of Mylotarg (gemtuzumab ozogamicin [GO]) to standard-induction therapy did not result in a higher CR rate but showed an improvement in 2-year OS (53% vs 42%) and EFS (41% vs 17%). Interestingly, MRD negativity (assessed by NPM1 RT-qPCR) was significantly more frequent in the GO arm as compared with the control arm: 39% vs 7% after induction and 91% vs 61% at the end of treatment.[43]

Conclusion

A definition of disease burden better than morphologic CR is now emerging. MRD assessed by flow or molecular techniques should become standard in every clinical trial in AML. Harmonization of antibody panels, introduction of single-cell tube systems (LSCs) and standardized analytical programs will pave the way for MRD-based individual risk assessment and for MRD to become a surrogate end point for survival in studies investigating new drugs, thereby hopefully resulting in faster drug approval in AML.

Correspondence

Gert Ossenkoppele, VU University Medical Center Amsterdam, Amsterdam, The Netherlands; e-mail: g.ossenkoppele@vumc.nl.

References

1. Döhner H, Weisdorf DJ, Bloomfield CD. Acute myeloid leukemia. *N Engl J Med*. 2015;373(12):1136-1152.
2. van der Pol MA, Pater JM, Feller N, et al. Functional characterization of minimal residual disease for P-glycoprotein and multidrug resistance protein activity in acute myeloid leukemia. *Leukemia*. 2001;15(10): 1554-1563.
3. van Stijn A, Kok A, van Stalborch MA, et al. Minimal residual disease cells in AML patients have an apoptosis-sensitive protein profile. *Leukemia*. 2004;18(4):875-877.
4. Hess CJ, Feller N, Denkers F, et al. Correlation of minimal residual disease cell frequency with molecular genotype in patients with acute myeloid leukemia. *Haematologica*. 2009;94(1):46-53.
5. van den Ancker W, van Luijn MM, Chamuleau ME, et al. High class II-associated invariant chain peptide expression on residual leukemic cells is associated with increased relapse risk in acute myeloid leukemia. *Leuk Res*. 2014;38(6):691-693.
6. Hourigan CS, Karp JE. Minimal residual disease in acute myeloid leukaemia. *Nat Rev Clin Oncol*. 2013;10(8):460-471.
7. Inaba H, Coustan-Smith E, Cao X, et al. Comparative analysis of different approaches to measure treatment response in acute myeloid leukemia. *J Clin Oncol*. 2012;30(29):3625-3632.
8. Grimwade D, Freeman SD. Defining minimal residual disease in acute myeloid leukemia: which platforms are ready for "prime time"? *Blood*. 2014;124(23):3345-3355.
9. Kayser S, Walter RB, Stock W, Schlenk RF. Minimal residual disease in acute myeloid leukemia–current status and future perspectives. *Curr Hematol Malig Rep*. 2015;10(2):132-144.
10. Krönke J, Schlenk RF, Jensen KO, et al. Monitoring of minimal residual disease in NPM1-mutated acute myeloid leukemia: a study from the German-Austrian acute myeloid leukemia study group. *J Clin Oncol*. 2011;29(19):2709-2716.
11. Cilloni D, Renneville A, Hermitte F, et al. Real-time quantitative polymerase chain reaction detection of minimal residual disease by standardized WT1 assay to enhance risk stratification in acute myeloid leukemia: a European LeukemiaNet study. *J Clin Oncol*. 2009;27(31): 5195-5201.
12. Gabert J, Beillard E, van der Velden VH, et al. Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia - a Europe Against Cancer program. *Leukemia*. 2003;17(12):2318-2357.
13. Ossenkoppele GJ, van de Loosdrecht AA, Schuurhuis GJ. Review of the relevance of aberrant antigen expression by flow cytometry in myeloid neoplasms. *Br J Haematol*. 2011;153(4):421-436.
14. Loken MR, Alonzo TA, Pardo L, et al. Residual disease detected by multidimensional flow cytometry signifies high relapse risk in patients with de novo acute myeloid leukemia: a report from Children's Oncology Group. *Blood*. 2012;120(8):1581-1588.
15. Sievers EL, Lange BJ, Alonzo TA, et al. Immunophenotypic evidence of leukemia after induction therapy predicts relapse: results from a prospective Children's Cancer Group study of 252 patients with acute myeloid leukemia. *Blood*. 2003;101(9):3398-3406.
16. Terwijn M, van Putten WL, Kelder A, et al. High prognostic impact of flow cytometric minimal residual disease detection in acute myeloid leukemia: data from the HOVON/SAKK AML 42A study. *J Clin Oncol*. 2013;31(31):3889-3897.
17. Drandi D, Kubiczkova-Besse L, Ferrero S, et al. Minimal residual disease detection by droplet digital PCR in multiple myeloma, mantle cell lymphoma, and follicular lymphoma: a comparison with real-time PCR. *J Mol Diagn*. 2015;17(6):652-660.
18. Bacher U, Dicker F, Haferlach C, et al. Quantification of rare NPM1 mutation subtypes by digital PCR. *Br J Haematol*. 2014;167(5):710-714.
19. Debarri H, Lebon D, Roumier C, et al. IDH1/2 but not DNMT3A mutations are suitable targets for minimal residual disease monitoring in acute myeloid leukemia patients: a study by the Acute Leukemia French Association. *Oncotarget*. 2015;6(39):42345-42353.
20. Kohlmann A, Nadarajah N, Alpermann T, et al. Monitoring of residual disease by next-generation deep-sequencing of RUNX1 mutations can identify acute myeloid leukemia patients with resistant disease. *Leukemia*. 2014;28(1):129-137.
21. Thol F, Kölking B, Damm F, et al. Next-generation sequencing for minimal residual disease monitoring in acute myeloid leukemia patients with FLT3-ITD or NPM1 mutations. *Genes Chromosomes Cancer*. 2012;51(7):689-695.
22. Zeijlemaker W, Gratama JW, Schuurhuis GJ. Tumor heterogeneity makes AML a "moving target" for detection of residual disease. *Cytometry B Clin Cytom*. 2014;86(1):3-14.
23. Costello RT, Mallet F, Gaugler B, et al. Human acute myeloid leukemia CD34+/CD38- progenitor cells have decreased sensitivity to chemotherapy and Fas-induced apoptosis, reduced immunogenicity, and impaired dendritic cell transformation capacities. *Cancer Res*. 2000;60(16): 4403-4411.
24. Pearce DJ, Taussig D, Zibara K, et al. AML engraftment in the NOD/SCID assay reflects the outcome of AML: implications for our understanding of the heterogeneity of AML. *Blood*. 2006;107(3):1166-1173.
25. van Rhenen A, Feller N, Kelder A, et al. High stem cell frequency in acute myeloid leukemia at diagnosis predicts high minimal residual disease and poor survival. *Clin Cancer Res*. 2005;11(18):6520-6527.
26. Terwijn M, Zeijlemaker W, Kelder A, et al. Leukemic stem cell frequency: a strong biomarker for clinical outcome in acute myeloid leukemia. *PLoS One*. 2014;9(9):e107587.
27. Zeijlemaker W, Kelder A, Oussoren-Brockhoff YJ, et al. A simple one-tube assay for immunophenotypical quantification of leukemic stem cells in acute myeloid leukemia. *Leukemia*. 2016;30(2):439-446.
28. Vanderhoek M, Juckett MB, Perlman SB, Nickles RJ, Jeraj R. Early assessment of treatment response in patients with AML using [(18)F]FLT PET imaging. *Leuk Res*. 2011;35(3):310-316.
29. Martens AC, Hagenbeek A. Detection of minimal disease in acute leukemia using flow cytometry: studies in a rat model for human acute leukemia. *Cytometry*. 1985;6(4):342-347.
30. Zeijlemaker W, Kelder A, Oussoren-Brockhoff YJ, et al. Peripheral blood minimal residual disease may replace bone marrow minimal residual disease as an immunophenotypic biomarker for impending relapse in acute myeloid leukemia. *Leukemia*. 2016;30(3):708-715.
31. Maurillo L, Buccisano F, Spagnoli A, et al. Monitoring of minimal residual disease in adult acute myeloid leukemia using peripheral blood as an alternative source to bone marrow. *Haematologica*. 2007;92(5): 605-611.

American Society of Hematology

32. Grimwade D, Jovanovic JV, Hills RK, et al. Prospective minimal residual disease monitoring to predict relapse of acute promyelocytic leukemia and to direct pre-emptive arsenic trioxide therapy. *J Clin Oncol.* 2009; 27(22):3650-3658.

33. Jourdan E, Boissel N, Chevret S, et al. Prospective evaluation of gene mutations and minimal residual disease in patients with core binding factor acute myeloid leukemia. *Blood.* 2013;121(12):2213-2223.

34. Yin JA, O'Brien MA, Hills RK, Daly SB, Wheatley K, Burnett AK. Minimal residual disease monitoring by quantitative RT-PCR in core binding factor AML allows risk stratification and predicts relapse: results of the United Kingdom MRC AML-15 trial. *Blood.* 2012;120(14): 2826-2835.

35. Hubmann M, Köhnke T, Hoster E, et al. Molecular response assessment by quantitative real-time polymerase chain reaction after induction therapy in NPM1-mutated patients identifies those at high risk of relapse. *Haematologica.* 2014;99(8):1317-1325.

36. Ivey A, Hills RK, Simpson MA, et al; UK National Cancer Research Institute AML Working Group. Assessment of minimal residual disease in standard-risk AML. *N Engl J Med.* 2016;374(5):422-433.

37. Scholl C, Schlenk RF, Eiwen K, Döhner H, Fröhling S, Döhner K; AML Study Group. The prognostic value of MLL-AF9 detection in patients with t(9;11)(p22;q23)-positive acute myeloid leukemia. *Haematologica.* 2005;90(12):1626-1634.

38. Freeman SD, Virgo P, Couzens S, et al. Prognostic relevance of treatment response measured by flow cytometric residual disease detection in older patients with acute myeloid leukemia. *J Clin Oncol.* 2013;31(32): 4123-4131.

39. Walter RB, Buckley SA, Pagel JM, et al. Significance of minimal residual disease before myeloablative allogeneic hematopoietic cell transplantation for AML in first and second complete remission. *Blood.* 2013;122(10):1813-1821.

40. Rubnitz JE, Inaba H, Dahl G, et al. Minimal residual disease-directed therapy for childhood acute myeloid leukaemia: results of the AML02 multicentre trial. *Lancet Oncol.* 2010;11(6):543-552.

41. Zhu HH, Zhang XH, Qin YZ, et al. MRD-directed risk stratification treatment may improve outcomes of t(8;21) AML in the first complete remission: results from the AML05 multicenter trial. *Blood.* 2013; 121(20):4056-4062.

42. Prebet T, Bertoli S, Delaunay J, et al. Anthracycline dose intensification improves molecular response and outcome of patients treated for core binding factor acute myeloid leukemia. *Haematologica.* 2014;99(10): e185-e187.

43. Lambert J, Lambert J, Nibourel O, et al. MRD assessed by WT1 and NPM1 transcript levels identifies distinct outcomes in AML patients and is influenced by gemtuzumab ozogamicin. *Oncotarget.* 2014;5(15): 6280-6288.

44. San Miguel JF, Martínez A, Macedo A, et al. Immunophenotyping investigation of minimal residual disease is a useful approach for predicting relapse in acute myeloid leukemia patients. *Blood.* 1997;90(6): 2465-2470.

45. Venditti A, Buccisano F, Del Poeta G, et al. Level of minimal residual disease after consolidation therapy predicts outcome in acute myeloid leukemia. *Blood.* 2000;96(12):3948-3952.

46. San Miguel JF, Vidriales MB, López-Berges C, et al. Early immunophenotypical evaluation of minimal residual disease in acute myeloid leukemia identifies different patient risk groups and may contribute to postinduction treatment stratification. *Blood.* 2001;98(6):1746-1751.

47. Coustan-Smith E, Ribeiro RC, Rubnitz JE, et al. Clinical significance of residual disease during treatment in childhood acute myeloid leukaemia. *Br J Haematol.* 2003;123(2):243-252.

48. Venditti A, Maurillo L, Buccisano, et al. Pretransplant minimal residual disease level predicts clinical outcome in patients with acute myeloid leukemia receiving high-dose chemotherapy and autologous stem cell transplantation. *Leukemia.* 2003;17(11);2178-2182.

49. Kern W, Voskova D, Schoch C, Schnittger S, Hiddemann W, Haferlach T. Prognostic impact of early response to induction therapy as assessed by multiparameter flow cytometry in acute myeloid leukemia. *Haematologica.* 2004;89(5):528-540.

50. Kern W, Voskova D, Schoch C, Hiddemann W, Schnittger S, Haferlach T. Determination of relapse risk based on assessment of minimal residual disease during complete remission by multiparameter flow cytometry in unselected patients with acute myeloid leukemia. *Blood.* 2004;104(10): 3078-3085.

51. Feller N, van der Pol MA, van Stijn A, et al. MRD parameters using immunophenotypic detection methods are highly reliable in predicting survival in acute myeloid leukaemia. *Leukemia.* 2004;18(8):1380-1390.

52. Buccisano F, Maurillo L, Gattei V, et al. The kinetics of reduction of minimal residual disease impacts on duration of response and survival of patients with acute myeloid leukemia. *Leukemia.* 2006;20(10): 1783-1789.

53. Laane E, Derolf AR, Björklund E, et al. The effect of allogeneic stem cell transplantation on outcome in younger acute myeloid leukemia patients with minimal residual disease detected by flow cytometry at the end of post-remission chemotherapy. *Haematologica.* 2006;91(6):833-836.

54. MRD-AML-BFM Study Group, Langebrake C, Creutzig U, Dworzak M, et al. Residual disease monitoring in childhood acute myeloid leukemia by multiparameter flow cytometry: the MRD-AML-BFM Study Group. *J Clin Oncol.* 2006;24(22):3686-3692.

55. Díez-Campalo M, Pérez-Simón JA, Pérez J, et al. Minimal residual disease monitoring after allogeneic transplantation may help to individualize post-transplant therapeutic strategies in acute myeloid malignancies. *Am J Hematol.* 2009;84(3):149-152.

56. Maurillo L, Buccisano F, Del Principe MI, et al. Toward optimization of postremission therapy for residual disease-positive patients with acute myeloid leukemia. *J Clin Oncol.* 2008;26(30):4944-4951.

57. Al-Mawali A, Gillis D, Hissaria P, et al. Incidence, sensitivity, and specificity of leukemia-associated phenotypes in acute myeloid leukemia using specific five-color multiparameter flow cytometry. *Am J Clin Pathol.* 2008;129:934-945.

58. van der Velden VH, van der Sluijs-Geling A, Gibson BE, et al. Clinical significance of flowcytometric minimal residual disease detection in pediatric acute myeloid leukemia patients treated according to the DCOG ANLL97/MRC AML12 protocol. *Leukemia.* 2010;24(9):1599-1606.

59. Buccisano F, Maurillo L, Piciocchi A, et al. Minimal residual disease negativity in elderly patients with acute myeloid leukemia may indicate different postremission strategies than in younger patients. *Ann Hematol.* 2015;94(8):1319-1326.

60. Araki D, Wood BL, Othus M, et al. Allogeneic hematopoietic cell transplantation for acute myeloid leukemia: time to move toward a minimal residual disease-based definition of complete remission? *J Clin Oncol.* 2016;34(4):329-336.

61. Vidriales MB, Pérez-López E, Pegenaute C, et al. Minimal residual disease evaluation by flow cytometry is a complementary tool to cytogenetics for treatment decisions in acute myeloid leukaemia. *Leuk Res.* 2016;40:1-9.

62. Tiersens A, Bjørklund E, Siitonen S, et al. Residual disease detected by flow cytometry is an independent predictor of survival in childhood acute myeloid leukaemia; results of the NOPHO-AML 2004 study. *Br J Haematol.* 2016;174(4):600-609. doi:10.1111/bjh.14093.