

Strikingly different molecular relapse kinetics in *NPM1c*, *PML-RARA*, *RUNX1-RUNX1T1*, and *CBFB-MYH11* acute myeloid leukemias

Hans Beier Ommen,¹ Susanne Schnittger,² Jelena V. Jovanovic,³ Ingrid Beier Ommen,¹ Henrik Hasle,⁴ Mette Østergaard,¹ David Grimwade,³ and Peter Hokland¹

¹Laboratory of Immunohematology, Department of Hematology, Aarhus University Hospital, Aarhus, Denmark; ²Munich Leukemia Laboratory, Munich, Germany; ³Department of Medical and Molecular Genetics, King's College London School of Medicine, London, United Kingdom; and ⁴Department of Pediatrics, Aarhus University Hospital Skejby, Aarhus, Denmark

Early relapse detection in acute myeloid leukemia is possible using standardized real-time quantitative polymerase chain reaction (RQ-PCR) protocols. However, optimal sampling intervals have not been defined and are likely to vary according to the underlying molecular lesion. In 74 patients experiencing hematologic relapse and harboring aberrations amenable to RQ-PCR (mutated *NPM1* [designated *NPM1c*], *PML-RARA*, *RUNX1-RUNX1T1*, and *CBFB-MYH11*), we observed strikingly different relapse kinetics. The me-

dian doubling time of the *CBFB-MYH11* leukemic clone was significantly longer (36 days) than that of clones harboring other markers (*RUNX1-RUNX1T1*, 14 days; *PML-RARA*, 12 days; and *NPM1c*, 11 days; $P < .001$). Furthermore, we used a mathematical model to determine frequency of relapse detection and median time from detection of minimal residual disease to hematologic relapse as a function of sampling interval length. For example, to obtain a relapse detection fraction of 90% and a median time of 60 days, blood sam-

pling every sixth month should be performed for *CBFB-MYH11* leukemias. By contrast, in *NPM1c*⁺/*FLT3-ITD*⁻, *NPM1c*⁺/*FLT3-ITD*⁺, *RUNX1-RUNX1T1*, and *PML-RARA* leukemias, bone marrow sampling is necessary every sixth, fourth, and fourth and second month, respectively. These data carry important implications for the development of optimal RQ-PCR monitoring schedules suitable for evaluation of minimal residual disease-directed therapies in future clinical trials. (Blood. 2010;115:198-205)

Introduction

Relapse remains the event that heralds ultimate treatment failure for most acute myeloid leukemia (AML) patients.¹ Thus, whereas reacquisition of complete remission (CR) is often possible, it always poses a greater therapeutical challenge than the initial cytoreduction, most probably because of selection for therapy-resistant clones.^{1,2}

Consequently, detection of impending relapse remains a major challenge in these patients. Several tools with various sensitivities are available for this task, most notably multicolor flow cytometry (sensitivity as low as 0.01%)³ and real-time quantitative polymerase chain reaction (RQ-PCR; sensitivity as low as 0.0001%).^{4,5} RQ-PCR assays are being applied to fusion transcripts, such as *PML-RARA*,⁶⁻¹¹ *CBFB-MYH11*,^{8,10,12-14} *RUNX1-RUNX1T1*,^{8,10,15-19} and *DEK-CAN*,²⁰⁻²² overexpressed genes, such as *WT1*²³⁻²⁶ and *PRAME*,²⁷ or mutated genes, such as *NPM1*.²⁸⁻³² There is evidence to suggest that assay sensitivity not only varies between minimal residual disease (MRD) markers, but can also differ significantly between patients possessing the same molecular target.^{9,10,32} Furthermore, as follow-up MRD samples are often collected sparsely, even in prospective studies, it has hitherto been difficult to provide firm recommendations for sampling, not only regarding intervals, but equally so with regard to source of sample material (ie, bone marrow [BM] or peripheral blood [PB]).

To address these issues, we recently used the Wilms tumor gene 1 (*WT1*) as a molecular marker to derive a mathematical model enabling the delineation of several quantitative parameters

related to relapse kinetics, such as the power to detect molecular relapse (MR; relapse detection fraction [RDF]) and the median time (t_m) from molecular positivity to hematologic relapse (HR) for different sampling intervals.³³

Given that the sensitivity of the *WT1* assay is inferior to assays detecting recurrent molecular aberrations such as fusion transcripts and mutations, we in this study present data collected from 3 centers handling large numbers of MRD samples. By analyzing relapsing patients with several prerelapse samples available and applying our mathematical model, we have now been able to delineate relapse kinetics in patients harboring 4 different target transcripts, namely *RUNX1-RUNX1T1*, *CBFB-MYH11*, *PML-RARA*, and mutated *NPM1* (designated *NPM1c*).³⁴ Application of the model, which takes into account differences in assay sensitivity and relapse kinetics, will allow development of monitoring schedules tailored according to the molecular lesion, suitable for evaluating the clinical utility of MRD-directed therapies in multicenter clinical trials.

Methods

Patient samples

Patient samples were analyzed at the Laboratory of Immunohematology (IHL), Aarhus University Hospital, from October 1995 to November 2007; the Munich Leukemia Laboratory from July 2005 to March 2008; and the

Submitted April 12, 2009; accepted October 13, 2009. Prepublished online as *Blood* First Edition paper, November 9, 2009; DOI 10.1182/blood-2009-04-212530.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2010 by The American Society of Hematology

Department of Medical and Molecular Genetics/Molecular Oncology Diagnostics Unit, Guy's Hospital from May 2002 to March 2009. All patients were diagnosed and followed for MRD detection using standard published methods.^{8-10,32,35} MRD analyses were conducted with informed patient consent in accordance with the Declaration of Helsinki and subject to ethics committee approval from all participating institutions. The patients with *PML-RARA* acute promyelocytic leukemia were treated with all-*trans* retinoic acid and anthracycline-based chemotherapy, as detailed in Grimwade et al.⁹ Non-acute promyelocytic leukemia (APL) patients were treated with standard combination therapy regimens, as described.^{10,32,33,36}

MRD determination and reporting

MRD values were calculated using the $\Delta\Delta$ cycle threshold (C_t) or the absolute quantification method using plasmid standards (Ipsogen), as described by Beillard et al,³⁷ according to the laboratory and molecular target (for *NPM1c* quantification, see Schnittger et al³²). Samples with low-quality RNA (defined as a threshold cycle number of the control gene β_2 -microglobulin exceeding 25 and/or a threshold cycle number of the control gene Abelson exceeding 30) were excluded.

Assays were run in duplicate (*NPM1c*) or triplicate (other markers), and in accordance with Europe against Cancer criteria, amplification in at least 2 of 3 replicates with C_t values up to 40 (threshold 0.1 [0.05 for *PML-RARA*]) was required to define a result as polymerase chain reaction (PCR) positive for the MRD marker in question.³⁷ For the depiction of relative values on an absolute scale, data were transformed from ΔC_t values using the PCR efficiencies found in routine laboratory assay testing.

Definition of MR, HR, and molecular CR

Definitions of MR, HR, and molecular CR (mCR) followed the recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia,³⁸ with the modifications described in "Low-level positive expression in patients in continuous CR." Thus, patients were considered having a MR at first recurrence of MRD marker PCR positivity, except for those followed using the *NPM1c* or *RUNX1-RUNXIT1* molecular markers in which a MRD level above 5×10^{-5} or 1×10^{-4} , respectively, was required to define MR. Similarly, patients with the 2 latter markers were considered in mCR if PCR expression levels dropped to below these thresholds.

Inclusion and exclusion criteria

To be eligible for relapse kinetics analysis, patients needed to enter CR, have PCR samples taken after the discontinuation of chemotherapy, but before HR, and not receive pre-emptive treatment upon MR. Thus, patients PCR positive after treatment, who were not given further high-dose chemotherapy, could be included in this part of the study ($n = 5$).

To be eligible for relapse modeling based on PCR conversion, patients needed to enter mCR, have samples taken the last year before HR, and not receive pre-emptive treatment upon MR. Thus, patients experiencing a HR with no previous positive samples could be included ($n = 28$). One patient with a central nervous system relapse was excluded from both analyses.

Statistical analyses

To compare relative levels of expression of the molecular target at diagnosis and at relapse, as well as increments in normalized leukemic transcripts before HR, Wilcoxon rank-sum test was used. To compare the performance of PB versus BM testing, the number of paired MRD measurements in which only the BM was positive was compared with the number of MRD measurements in which only the PB was positive using the binomial distribution. These analyses were restricted to paired samples in which the BM and PB yielded RNA of comparable quality, as indicated by the respective levels of control gene expression.

Modeling of relapse patterns

Relapse patterns were modeled as described in Ommen et al.³³ Briefly, prerelapse samples were divided into monthly intervals based on how long

before time of HR they were taken. Patients were considered RQ-PCR negative or positive in all intervals between 2 negative or positive samples, respectively. The fraction of positive samples in each interval was then plotted against time to relapse. For a given positive sampling interval, I , the chance of detecting the relapse, *RDF*, was then given by:

$$RDF = \frac{\int_{-I}^0 F(t) dt}{I},$$

where $F(t)$ describes the distribution of the time span from when the leukemic burden exceeds the sensitivity of a given MRD marker to HR. This function can be approximated based on the distribution of individual patients' conversion to PCR positivity before HR, as described in Ommen et al.³³ The median time from MR to HR, t_m , can be found by solving the integral equation:

$$\int_{t_m}^0 (F(t) - F(t + I)) dt = \int_{t_0}^{t_m} (F(t) - F(t + I)) dt,$$

where t_0 is the intercept with the x -axis.³³ In this manner, RDFs and t_m for sampling intervals 1, 2, 3, 4, 6, 9, and 12 months were found for each molecular marker in both PB and BM. $F(t)$ is approximated using high-grade polynomials for each of the MRD markers. The polynomials used can be found in supplemental Document 1 (available on the *Blood* website; see the Supplemental Materials link at the top of the online article).

Results

MRD marker expression at diagnosis and relapse

A major determinant of the capacity for RQ-PCR to detect residual AML in a background of normal hematopoietic cells is assay sensitivity, which depends on the relative level of expression of the MRD target in leukemic blasts (reviewed in Freeman et al⁵). To compare assay sensitivities, we determined the MRD marker expression in relation to control gene expression in BM at diagnosis from 365 patients (151 *PML-RARA*, 31 *CBFB-MYH11*, 42 *RUNX1-RUNXIT1*, and 141 *NPM1c*) and at HR for all relapsing patients.

As seen from supplemental Figure 1 and supplemental Table 1, median expression of *NPM1c* was higher than that of the fusion transcript MRD markers that exhibited comparable levels of expression. Whereas diagnosis levels were generally higher than relapse levels, median values were within the same order of magnitude, allowing the use of diagnosis expression level as a measurement for MRD assay sensitivity.

Low-level positive expression in patients in continuous CR

Simple qualitative PCR positivity has been reported in several cases in patients who did not later relapse.³⁹⁻⁴¹ To investigate this situation in this large cohort of patients, we analyzed the number of positive samplings in patients who had earlier tested negative (for *PML-RARA* patients, 2 consecutive negative samples were required, as described in Grimwade et al⁹). For all 4 aberrations, a small percentage of positive samples could in fact be observed, the highest fraction in *NPM1c* patients (26% of continuous CR [CCR] samplings), possibly due to the higher level of MRD marker transcripts in these cells or in some cases due to cross-reactivity of the mutant assay with the wild-type allele.¹⁰ However, when a cutoff of 5×10^{-5} , relative to the diagnostic level, was used to define MR, only 1 positive reaction was not followed by a HR (of a total of 104 CCR determinations). Applying a similar cutoff level

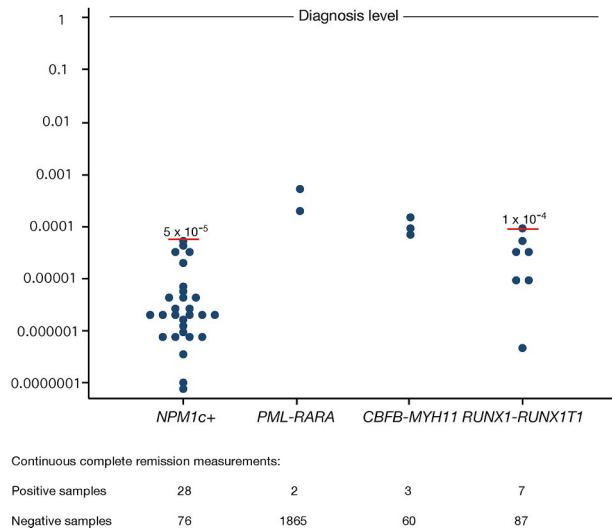


Figure 1. Low-level positive RQ-PCR samplings in patients in CCR. x-axis, MRD target; y-axis, MRD marker expression level relative to diagnostic level, arbitrarily set to 1. Numbers below the figure depict the number of low-level positive samples, and negative samples in patients with available diagnostic levels. Suggested cutoff levels for *RUNX1-RUNX1T1* and *NPM1c* are shown in red bars and given in numbers. Sensitivity of each test can be understood as the lower of the difference between diagnostic level and the sample quality-dependent sensitivity (typically between 1×10^{-5} and 1×10^{-3}) and cutoff level (*NPM1c* and *RUNX1-RUNX1T1*).

for *RUNX1-RUNX1T1* patients at 1×10^{-4} resulted in 1 positive reaction of 87 CCR determinations not followed by a HR. By contrast, positive samplings in CCR using the *CBFB-MYH11* molecular marker occurred in only 3 samples, accounting for 4.8% of the CCR samples. For this aberration, we therefore chose not to use a MR cutoff level, and resorted to the requirement of 2 consecutive positive samplings to define a MR given that *CBFB-MYH11* relapses are generally indolent, allowing for such an approach.⁴² For *PML-RARA* follow-up, positive samples were rare, and we chose not to use a cutoff threshold in the case of this

aberration. For the 2 assays (*NPM1c* and *RUNX1-RUNX1T1*) in which thresholds were introduced to exclude very low-level positivity not followed by a HR, the maximum sensitivity corresponds to the level of the threshold. As the sensitivity of the assays using the other 2 MRD markers is approximately 1×10^{-4} (with a range of 1×10^{-5} to 1×10^{-3} reflecting differences in the relative level of expression of the MRD target in the leukemic blasts between patients, as well as variation in the quality of follow-up samples), we found that the sensitivities of MRD detection, and thereby the ability to detect MR, using the 4 different markers are comparable (Figure 1).

Relapse kinetics

The advantage of using the RQ-PCR methodology in MRD follow-up is its potential for revealing relapse kinetics in the individual patient, provided that enough RQ-PCR-positive samples have been obtained before HR (Table 1).

In Figure 2 the *CBFB-MYH11* patients display a slower rate of rise of leukemic transcripts before HR than all the other aberrations (median BM doubling time 36 days, range 7.4-175 days, versus all other aberrations; $P < .001$). Compared with this, the incremental rise in *NPM1c* transcripts was more rapid (median BM doubling time 11 days, range 2.2-33 days; $P = .002$); however, there was marked heterogeneity with some patients experiencing indolent relapses, whereas others exhibited the shortest doubling times in the dataset. A major reason for this duality was found to be the presence or absence of the *FLT3-ITD* aberration (*NPM1c*⁺/*FLT3-ITD*⁻, median BM doubling time 15 days, range 2.2-33 days; *NPM1c*⁺/*FLT3-ITD*⁺, median BM doubling time 7.4 days, range 3.0-14 days; $P = .031$). *RUNX1-RUNX1T1* relapsing clone reappearance was faster than *CBFB-MYH11* growth (median BM doubling time 14 days, range 12-51 days; $P = .105$), as was *PML-RARA* clone growth (median BM doubling time 12 days, range 4.2-462 days; $P = .093$), with the exception of 1 patient who displayed the slowest relapse growth rate in the entire cohort (BM and PB doubling times of 462 and 54 days, respectively). Both these leukemias progressed with approximately the same speed as the *NPM1*⁺/*FLT3-ITD*⁻

Table 1. Study subjects

MRD marker	<i>NPM1c</i> ⁺	<i>PML-RARA</i>	<i>CBFB-MYH11</i>	<i>RUNX1-RUNX1T1</i>
No. of patients	180	406*	47	56
Patients with HR	54	22*	19	20
Patients in CCR	126	363*	28	36
CCR samples	193	1865	151	170
Relapse kinetics				
Patients included	22	8	9	7
Patients excluded	32	13	10	13
Reasons for exclusion:				
No MR/HR sample	32	9	10	12
Administration of pre-emptive therapy in MR	0	4	0	1
CNS relapse	0	1	0	0
Relapse modeling				
Patients included	31	13	14	11
Samples	70	61	59	46
Patients excluded	23	8	5	9
Reasons for exclusion:				
No mCR	3	1	0	1
No CR samples [†]	20	4	5	7
Administration of pre-emptive therapy in MR	0	3	0	1
CNS relapse	0	1	0	0

MRD indicates minimal residual disease; HR, hematologic relapse; CR, complete remission; CCR, continuous CR; MR, molecular relapse; mCR, molecular CR; and CNS, central nervous system.

*Numbers do not add, as some patients were successfully treated in first MR.

[†]In the 12 months prior to HR, there were no CR samples.

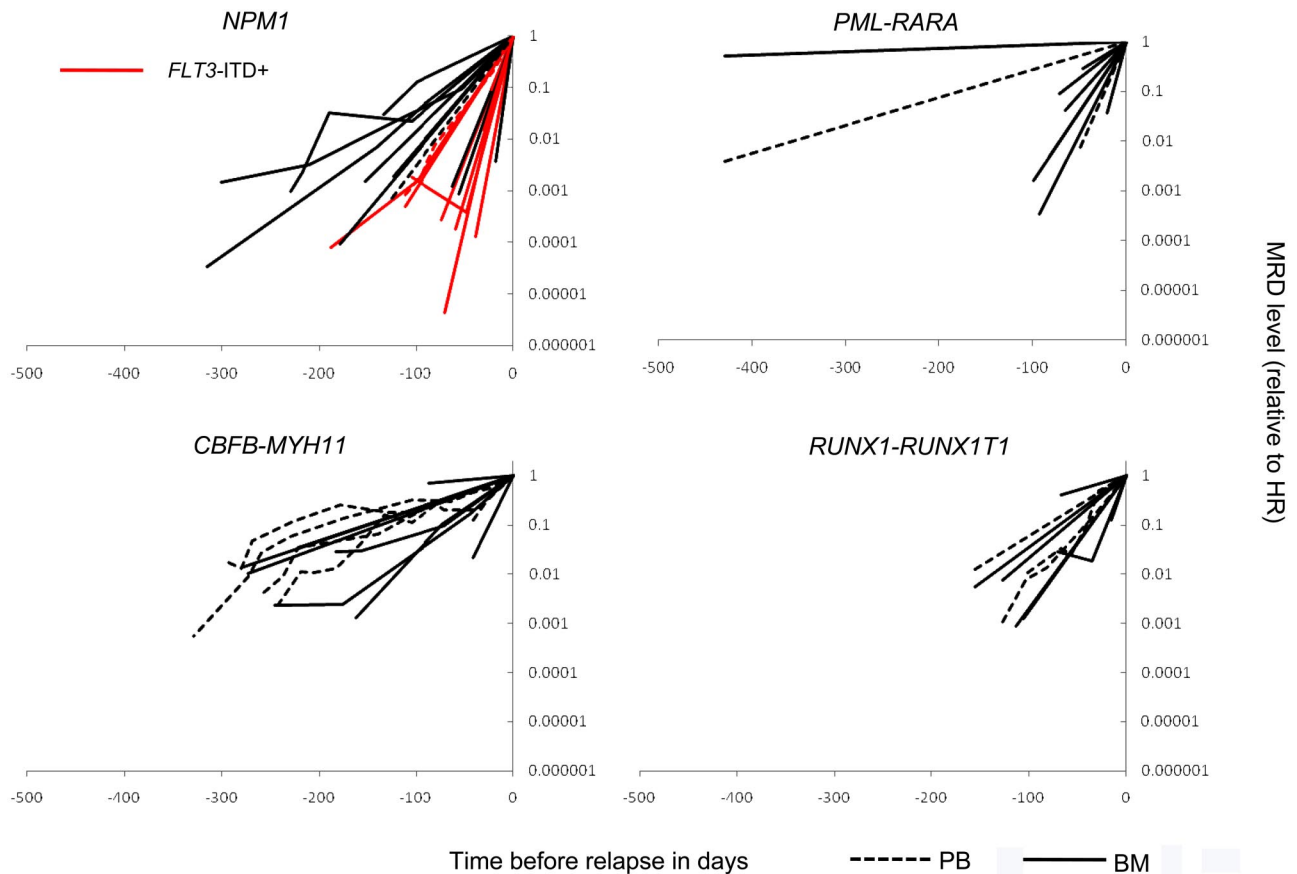


Figure 2. Relapse kinetics according to molecular lesion. Increasing MRD marker levels are shown in patients between time of molecular relapse and diagnosis of HR. Expression normalized to HR level (set to 1). x-axis, Time before HR. y-axis, Normalized MRD levels with PB and BM values depicted separately.

ones ($P = .97$ and $.93$, respectively), but slower than the $NPM1^+/FLT3-ITD^+$ leukemias ($P = .02$ and $.05$, respectively).

Relapse modeling based on conversion to PCR positivity

In some patients subjected to longitudinal RQ-PCR testing, no positive samples before HR will be obtained, either because of very rapid clone growth, poor sampling quality giving rise to false negative results, or simply unfortunate scheduling of MRD assessment relative to the timing of recurrent PCR positivity. We devised a mathematical model, which takes this lack of data into account and allows inclusion of information from such patients.³³ As can be seen in Figure 3, which depicts time before relapse as a function of the fraction of positive samples in each interval for each molecular marker, we were able to confirm and extend the relapse kinetics findings shown in Figure 2. Thus, *CBFB-MYH11* emerged as the MRD marker with the longest lag from MR to HR (50% of the patients tested RQ-PCR positive in BM 8 months before relapse; Figure 3C). By comparison, *RUNX1-RUNX1T1* relapses showed more rapid relapse kinetics with 50% being positive in BM as close as 3 months before relapse (Figure 3D). Moreover, the dichotomy observed for *NPM1c* transcript-based relapse detection was preserved in the model with a 3-month difference in when 50% of *FLT3-ITD*⁻ and *FLT3-ITD*⁺ patients became PCR positive (6.5 months vs 3.5 months before relapse; Figure 3A). For *PML-RARA*-positive relapses, the heterogeneity was even more pronounced, with 1 relapse being detectable 14 months before HR, and 2 being undetectable 71 days and 41 days before HR, respectively (50% detectable 3.5 months before HR; Figure 3B).

A recurring question when RQ-PCR is used for MRD detection has been to what extent PB can substitute for BM sampling. Comparing PB versus BM-based MR detection, it is now apparent that for the majority of *PML-RARA* cases, BM sampling is superior. Thus, in 7 paired samplings, MR was detected in BM only in 5 cases and in PB only in none ($P = .031$). In Figure 3C and D, in contrast suggests that PB and BM sampling are equally useful for *CBFB-MYH11* and *RUNX1-RUNX1T1*, even though the number of paired samplings was too low to draw any firm conclusions.

Taking advantage of the mathematical modeling, we were finally able to calculate the RDFs and $t_{m,s}$ to hematologic relapse as a function of sampling interval using the formulas described in "Modeling of relapse kinetics" (Figure 4). For *NPM1c*-based follow-up, modeling was done for the subgroups containing and lacking *FLT3-ITD* separately. As conclusions regarding *NPM1c*-based PB sampling would be based on very few patients (median number of MRD courses per interval 2.5), no modeling was done for *NPM1c*-based follow-up in this tissue. For comparative purposes and considering that some patients are negative for all the markers described in this study, we included *WT1* data that complement the ones presented in this study, resulting in more than 80% of AML patients having a valid molecular marker for MRD detection.²⁵

Once more, *CBFB-MYH11* leukemia displayed the most tardy relapses with high RDFs and $t_{m,s}$ for sampling intervals as long as 6 months (PB, RDF 90% and $t_{m,s}$ 180 days; BM, RDF 85% and $t_{m,s}$ 150 days). Sampling interval of this length also yielded satisfactory results in the group of patients with *NPM1c* leukemia without

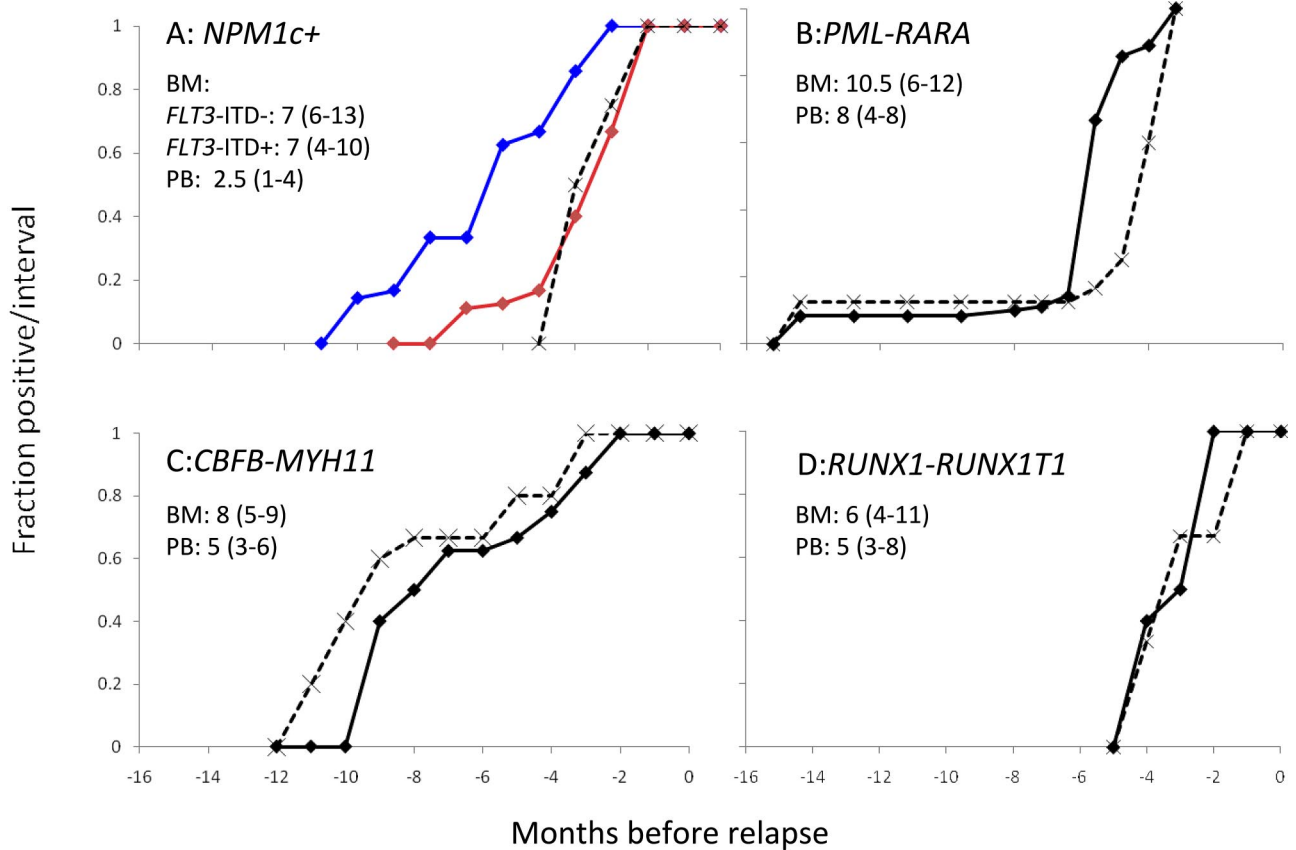


Figure 3. Conversion to PCR positivity. x-axis, Time, HR occurred at $t = 0$. y-axis, Fraction-positive samples in each prerelapse interval. Median number of evaluable MRD profiles in each interval is shown for each molecular marker (panels A-D; range in parentheses). Solid lines, BM; broken lines, PB. A: *NPM1c+* blue, *FLT3-ITD*⁻; red, *FLT3-ITD*⁺. B: *PML-RARA*. C: *CBFB-MYH11*. D: *RUNX1-RUNX1T1*.

FLT3-ITD (6-month intervals, BM: RDF 90%, t_m 120 days). As an intermediate group, for *NPM1c* (with *FLT3-ITD*) and *RUNX1-RUNX1T1* leukemias, application of a 3- to 4-month sampling frequency still yielded satisfactory relapse detection (4-month intervals, *NPM1c*⁺*FLT3-ITD*⁺, BM, RDF 85% and t_m 65 days; *RUNX1-RUNX1T1* PB, RDF 75% and t_m 55 days; BM, RDF 95% and t_m 85 days). Somewhat surprisingly, relapse detection by *WT1*

fell into this category too, at least when BM was used (4-month intervals, *WT1*; BM, RDF 95% and t_m 75 days). Due to the occurrence of 2 patients who were negative in BM close to HR, *PML-RARA* relapses were the most difficult to detect, and 2-month BM sampling will be necessary to obtain the relapse detection efficiencies of the other markers (2-month intervals, *PML-RARA*; RDF 95%, t_m 70 days).

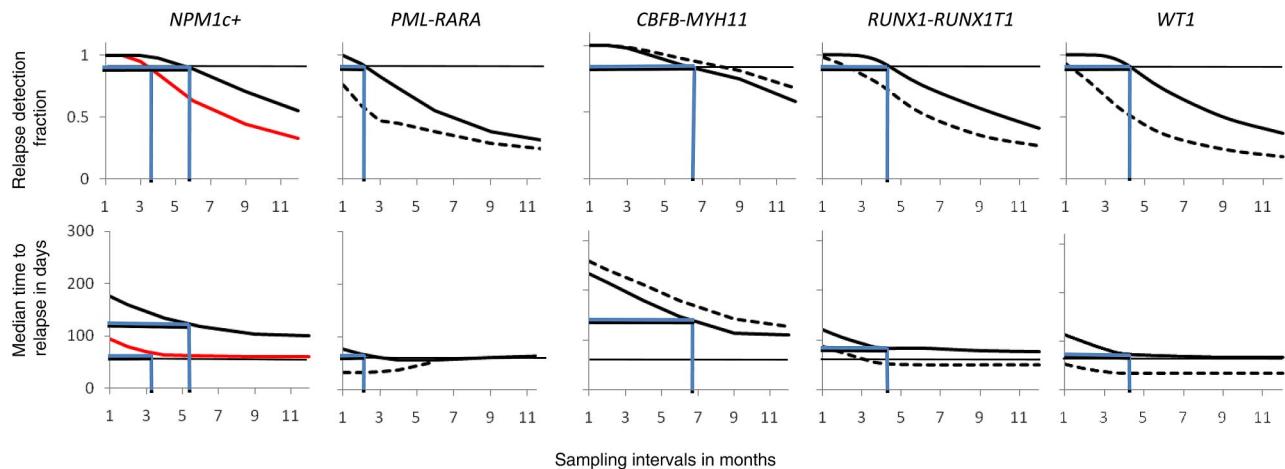


Figure 4. Comparisons between different molecular markers regarding RDF and t_m . x-axis, Sampling time in months. y-axis, RDF (top), median time from MR to HR (t_m ; bottom). Horizontal black lines, top panels: RDF equaling 90%. Bottom panels, t_m equaling 60 days. Blue lines in top panel represent sampling interval necessary to ensure 90% relapse detection. Blue lines in bottom panels show corresponding t_m values. Solid lines, BM; broken lines, PB; *NPM1c*-based MRD follow-up: black line, *FLT3-ITD*⁻ and red line, *FLT3-ITD*⁺. *PML-RARA*-based MRD follow-up: PB sampling intervals above 6 months not shown as model handles situation of long sampling intervals in very heterogeneous populations poorly.

Discussion

Close molecular monitoring using RQ-PCR in AML patients in CR holds the promise of detecting subclinical levels of residual disease in time to institute treatment intervention to prevent overt relapse. Despite an impressive body of data showing that RQ-PCR is excellently suited for early detection of AML relapses,^{6-28,30-32} few investigators have taken clinical action on these findings, although it is by now evident that the risk of a false positive result of a RQ-PCR is minimal, at least when the thresholds to exclude irrelevant low-level amplification in CCR are used, and especially so when molecular conversion is confirmed in a subsequent sample.^{9,33,43,44} To date, benefit from early salvage after conversion has been shown in *PML-RARA* APL,^{9,45,46} and in a preliminary report on the use of donor-lymphocyte infusion upon recurrent *WT1* positivity after allogeneic transplantation,⁴⁷ but has not been otherwise evaluated in non-APL AML patients.

Major reasons for this lack of translation of a powerful and technically standardized molecular method into clinical decision making beyond APL include uncertainty as to (1) the most informative schedules for MRD monitoring, (2) the most appropriate management of confirmed molecular relapse, and (3) whether early treatment intervention is likely to confer any clinical benefit compared with retreatment in overt relapse. To develop optimal MRD monitoring schedules to allow reliable assessment of the clinical utility of MRD monitoring in non-APL patients within large scale clinical trials, we have analyzed data from a very large cohort of patients subject to hematologic relapse (n = 114).

We used 2 different ways of describing the behavior of the leukemic clone before HR. First, we used the quantitative data obtained from patients in whom MR was identified to describe relapsing clone growth before HR. Assuming a constant doubling time, we were able to compare the different leukemia subsets and show that *CBFB-MYH11* leukemia displayed slower leukemic clone growth than AML with mutant *NPM1* or the *RUNX1-RUNX1T1* fusion and, with a single exception, *PML-RARA*⁺ APL. Interestingly, but perhaps not surprisingly, *NPM1c⁺/FLT3-ITD⁺* AML displayed relapses occurring significantly faster than all other molecular subtypes studied.

In the other model presented in this study, we examined pre-HR conversion to PCR positivity. This approach has the advantage of allowing the inclusion of information on patients for whom MR was not necessarily detected before HR. In addition, the assumption of a constant doubling time of the malignant clone is not necessary, although the cost is the lower degree of integration of quantitative values in this model. Thus, the 2 models complement each other, and it is reassuring that the results using these distinct approaches were near identical.

By directly comparing paired samples of PB and BM, we were able to show that *PML-RARA* BM MRD testing is superior to that of PB. Moreover, in the absence of a suitable number of paired samples for the other aberrations, the PCR positivity conversion model suggests that PB and BM are comparable for MRD detection in core binding factor leukemia. This is an encouraging conclusion, as PB sampling is much less troublesome for the patients, even if slightly higher sampling frequencies are necessary. However, further analyses are needed, especially regarding *NPM1c*-based MRD detection in PB, as our data do not support any definitive conclusions in this matter.

One great challenge when optimizing MRD follow-up is to determine the optimal sampling interval. The model presented in

this study allows for the evaluation of suitable sampling intervals for the different analyzed markers.

Establishing guidelines for follow-up sampling should include considerations about their predictive value of MRD assessment, the options for intervention, as well as cost-benefit estimations. Thus, PB sampling every 6 months in patients with *CBFB-MYH11* leukemia from PB will result in a RDF of 90% and a t_m of 180 days. Even with this sampling cadenza, only 10% of relapses will be missed before hematologic relapse. Moreover, if the first sample is obtained 3 months after discontinuation of therapy and each patient is followed for only 3 years (given the literature on AML relapse⁴⁸ and that 15 of 16 *CBFB-MYH11* relapses in the study cohort occurred during the first 3 years after diagnosis), based on the present study, only 14 MRD samples would have to be taken to detect 1 MR. Given such considerations, which can be applied equally well to other RQ-PCR assays, it will be apparent that molecular monitoring might prove to be cost effective with its promise of early, possibly less intensive, intervention.

In conclusion, we in this study present data enabling us to model relapse kinetics in 4 common AML subtypes. We show great difference between these subtypes, and show that the studied markers are generally superior to nonleukemia-specific marker *WT1*. Furthermore, we show that this superiority can be used in the core binding factor leukemias to use PB as the sampling tissue. These data should be useful in cost-benefit calculations regarding MRD monitoring implementation, in clinical decision making in the individual patient, and in statistical power calculations when designing clinical trials.

Acknowledgments

We thank Drs Morten Frydenberg and Steffen Hokland for statistical help. We thank Dr Robert Hills for helpful discussions, and D.G. gratefully acknowledges Prof Alan Burnett and the United Kingdom National Cancer Research Institute AML Working Group for access to data from APL patients treated in the Medical Research Council AML15 trial. P.H. thanks Prof Finn Bo Petersen, Intermountain Health Care Center, Salt Lake City, for helpful discussions.

This study was supported by grants from the Danish Medical Research Council, the Danish Cancer Society, the John and Birthe Meyer Foundation, and the Karen Elise Jensen Foundation (P.H.). D.G. and J.V.J. gratefully acknowledge support from the Leukemia Research Fund (Great Britain) and the European LeukemiaNet. This study was conducted within the framework of the MRD Workpackage (WP12) of the European LeukemiaNet.

Authorship

Contribution: H.B.O. and P.H. designed the study, analyzed the data, and wrote the first draft of the manuscript; S.S. was responsible for the analyses performed at Munich Leukemia Laboratory; J.V.J. and D.G. were responsible for the analyses performed at the Department of Medical and Molecular Genetics; I.B.O. analyzed the data; H.H. was responsible for the clinical care of the pediatric patients followed at the IHL; and M.Ø. was responsible for the analyses performed at IHL. All authors made significant contributions to the final manuscript.

Conflict-of-interest disclosure: S.S. in part owns Munich Leukemia Laboratory. The remaining authors declare no competing financial interests.

Correspondence: Peter Hokland, Aarhus University Hospital, Tage-Hansens Gade 2, Århus, Denmark 8000; e-mail: phokland@ki.au.dk.

References

- Breems DA, Van Putten WL, Huijgens PC, et al. Prognostic index for adult patients with acute myeloid leukemia in first relapse. *J Clin Oncol*. 2005; 23:1969-1978.
- Estey E, Dohner H. Acute myeloid leukemia. *Lancet*. 2006;368:1894-1907.
- Vidriales MB, San-Miguel JF, Orfao A, Coustan-Smith E, Campana D. Minimal residual disease monitoring by flow cytometry. *Best Pract Res Clin Haematol*. 2003;16:599-612.
- Kubista M, Andrade JM, Bengtsson M, et al. The real-time polymerase chain reaction. *Mol Aspects Med*. 2006;27:95-125.
- Freeman SD, Jovanovic JV, Grimwade D. Development of minimal residual disease-directed therapy in acute myeloid leukemia. *Semin Oncol*. 2008;35:388-400.
- Cassinat B, Zassadowski F, Balitrand N, et al. Quantitation of minimal residual disease in acute promyelocytic leukemia patients with t(15;17) translocation using real-time RT-PCR. *Leukemia*. 2000;14:324-328.
- Slack JL, Bi W, Livak KJ, et al. Pre-clinical validation of a novel, highly sensitive assay to detect PML-RAR α mRNA using real-time reverse-transcription polymerase chain reaction. *J Mol Diagn*. 2001;3:141-149.
- 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; 17:2318-2357.
- 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;27:3650-3658.
- Schnittger S, Weissner M, Schoch C, Hiddemann W, Haferlach T, Kern W. New score predicting for prognosis in PML-RARA $^{+}$, AML1-ETO $^{+}$, or CBFMYH11 $^{+}$ acute myeloid leukemia based on quantification of fusion transcripts. *Blood*. 2003; 102:2746-2755.
- Santamaria C, Chillon MC, Fernandez C, et al. Using quantification of the PML-RAR α transcript to stratify the risk of relapse in patients with acute promyelocytic leukemia. *Haematologica*. 2007; 92:315-322.
- Marcucci G, Caligiuri MA, Dohner H, et al. Quantification of CBF β /MYH11 fusion transcript by real time RT-PCR in patients with INV(16) acute myeloid leukemia. *Leukemia*. 2001;15: 1072-1080.
- Krauter J, Hoellge W, Wattjes MP, et al. Detection and quantification of CBF β /MYH11 fusion transcripts in patients with inv(16)-positive acute myeloblastic leukemia by real-time RT-PCR. *Genes Chromosomes Cancer*. 2001;30: 342-348.
- Buonamici S, Ottaviani E, Testoni N, et al. Real-time quantitation of minimal residual disease in inv(16)-positive acute myeloid leukemia may indicate risk for clinical relapse and may identify patients in a curable state. *Blood*. 2002;99: 443-449.
- Marcucci G, Livak KJ, Bi W, Strout MP, Bloomfield CD, Caligiuri MA. Detection of minimal residual disease in patients with AML1/ETO-associated acute myeloid leukemia using a novel quantitative reverse transcription polymerase chain reaction assay. *Leukemia*. 1998;12:1482-1489.
- Krauter J, Wattjes MP, Nagel S, et al. Real-time RT-PCR for the detection and quantification of AML1/MTG8 fusion transcripts in t(8;21)-positive AML patients. *Br J Haematol*. 1999;107:80-85.
- Sugimoto T, Das H, Imoto S, et al. Quantitation of minimal residual disease in t(8;21)-positive acute myelogenous leukemia patients using real-time quantitative RT-PCR. *Am J Hematol*. 2000;64: 101-106.
- Kondo M, Kudo K, Kimura H, et al. Real-time quantitative reverse transcription-polymerase chain reaction for the detection of AML1-MTG8 fusion transcripts in t(8;21)-positive acute myelogenous leukemia. *Leuk Res*. 2000;24:951-956.
- Fujimaki S, Funato T, Harigae H, et al. A quantitative reverse transcriptase polymerase chain reaction method for the detection of leukaemic cells with t(8;21) in peripheral blood. *Eur J Haematol*. 2000;64:252-258.
- Østergaard M, Stentoft J, Hokland P. A real-time quantitative RT-PCR assay for monitoring DEK-CAN fusion transcripts arising from translocation t(6;9) in acute myeloid leukemia. *Leuk Res*. 2004; 28:1213-1215.
- Tobal K, Frost L, Liu Yin JA. Quantification of DEK-CAN fusion transcript by real-time reverse transcription polymerase reaction in patients with t(6;9) acute myeloid leukemia. *Haematologica*. 2004;89:1267-1269.
- Garçon L, Libura M, Delabesse E, et al. DEK-CAN molecular monitoring of myeloid malignancies could aid therapeutic stratification. *Leukemia*. 2005;19:1338-1344.
- Cilloni D, Gottardi E, De Micheli D, et al. Quantitative assessment of WT1 expression by real time quantitative PCR may be a useful tool for monitoring minimal residual disease in acute leukemia patients. *Leukemia*. 2002;16:2115-2121.
- Trka J, Kalinova M, Hrusak O, et al. Real-time quantitative PCR detection of WT1 gene expression in children with AML: prognostic significance, correlation with disease status and residual disease detection by flow cytometry. *Leukemia*. 2002;16:1381-1389.
- Ostergaard M, Olesen LH, Hasle H, Kjeldsen E, Hokland P. WT1 gene expression: an excellent tool for monitoring minimal residual disease in 70% of acute myeloid leukemia patients: results from a single-centre study. *Br J Haematol*. 2004; 125:590-600.
- 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: 5195-5201.
- Steinbach D, Hermann J, Viehmann S, Zintl F, Gruhn B. Clinical implications of PRAME gene expression in childhood acute myeloid leukemia. *Cancer Genet Cytogenet*. 2002;133:118-123.
- Gorello P, Cazzaniga G, Alberti F, et al. Quantitative assessment of minimal residual disease in acute myeloid leukemia carrying nucleophosmin (NPM1) gene mutations. *Leukemia*. 2006;20: 1103-1108.
- Papadaki C, Dufour A, Seibl M, et al. Monitoring minimal residual disease in acute myeloid leukemia with NPM1 mutations by quantitative PCR: clonal evolution is a limiting factor. *Br J Haematol*. 2009;144:517-523.
- Chou WC, Tang JL, Wu SJ, et al. Clinical implications of minimal residual disease monitoring by quantitative polymerase chain reaction in acute myeloid leukemia patients bearing nucleophosmin (NPM1) mutations. *Leukemia*. 2007;21:998-1004.
- Bacher U, Badbaran A, Fehse B, Zabelina T, Zander AR, Kroger N. Quantitative monitoring of NPM1 mutations provides a valid minimal residual disease parameter following allogeneic stem cell transplantation. *Exp Hematol*. 2009;37: 135-142.
- Schnittger S, Kern W, Tschulik C, et al. Minimal residual disease levels assessed by NPM1 mutation specific RQ-PCR provide important prognostic information in AML. *Blood*. 2009;114:2220-2231.
- Ommen HB, Nyvold CG, Braendstrup K, et al. Relapse prediction in acute myeloid leukemia patients in complete remission using WT1 as a molecular marker: development of a mathematical model to predict time from molecular to clinical relapse and define optimal sampling intervals. *Br J Haematol*. 2008;141:782-791.
- Falini B, Mecucci C, Tiacci E, et al. Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. *N Engl J Med*. 2005; 352:254-266.
- Pallisaard N, Hokland P, Riishoj DC, Pedersen B, Jorgensen P. Multiplex reverse transcription-polymerase chain reaction for simultaneous screening of 29 translocations and chromosomal aberrations in acute leukemia. *Blood*. 1998;92: 574-588.
- Olesen LH, Aggerholm A, Andersen BL, et al. Molecular typing of adult acute myeloid leukemia: significance of translocations, tandem duplications, methylation, and selective gene expression profiling. *Br J Haematol*. 2005;131: 457-467.
- Beillard E, Pallisaard N, van der Velden VH, et al. Evaluation of candidate control genes for diagnosis and residual disease detection in leukemic patients using 'real-time' quantitative reverse-transcriptase polymerase chain reaction (RQ-PCR): a Europe against Cancer program. *Leukemia*. 2003;17:2474-2486.
- Cheson BD, Bennett JM, Kopecky KJ, et al. Revised recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia. *J Clin Oncol*. 2003;21: 4642-4649.
- Jurlander J, Caligiuri MA, Ruutu T, et al. Persistence of the AML1/ETO fusion transcript in patients treated with allogeneic bone marrow transplantation for t(8;21) leukemia. *Blood*. 1996;88: 2183-2191.
- Tobal K, Saunders MJ, Grey MR, Yin JA. Persistence of RAR α -PML fusion mRNA detected by reverse transcriptase polymerase chain reaction in patients in long-term remission of acute promyelocytic leukemia. *Br J Haematol*. 1995;90:615-618.
- Miyamoto T, Nagafuji K, Akashi K, et al. Persistence of multipotent progenitors expressing AML1/ETO transcripts in long-term remission patients with t(8;21) acute myelogenous leukemia. *Blood*. 1996;87:4789-4796.
- Stentoft J, Hokland P, Ostergaard M, Hasle H, Nyvold CG. Minimal residual core binding factor AMLs by real time quantitative PCR: initial response to chemotherapy predicts event free survival and close monitoring of peripheral

- blood unravels the kinetics of relapse. *Leuk Res.* 2006;30:389-395.
43. Diverio D, Rossi V, Avvisati G, et al. Early detection of relapse by prospective reverse transcriptase-polymerase chain reaction analysis of the PML/RAR α fusion gene in patients with acute promyelocytic leukemia enrolled in the GIMEMA-AIEOP multicenter "AIDA" trial: GIMEMA-AIEOP Multicenter "AIDA" Trial. *Blood.* 1998;92:784-789.
 44. Raff T, Gokbuget N, Luschen S, et al. Molecular relapse in adult standard-risk ALL patients detected by prospective MRD monitoring during and after maintenance treatment: data from the GMALL 06/99 and 07/03 trials. *Blood.* 2007;109:910-915.
 45. Lo Coco F, Diverio D, Avvisati G, et al. Therapy of molecular relapse in acute promyelocytic leukemia. *Blood.* 1999;94:2225-2229.
 46. Esteve J, Escoda L, Martin G, et al. Outcome of patients with acute promyelocytic leukemia failing to front-line treatment with all-trans retinoic acid and anthracycline-based chemotherapy (PETHEMA protocols LPA96 and LPA99): benefit of an early intervention. *Leukemia.* 2007;21:446-452.
 47. Bacigalupo A, Dominietto A, Di Grazia C, et al. Leukemia relapse after allogeneic transplants in AML: predictive role of WT1 expression and protective effect of DLI. ASH Annual Meeting Abstracts. *Blood.* 2008;112:Abstract 1120.
 48. de Lima M, Strom SS, Keating M, et al. Implications of potential cure in acute myelogenous leukemia: development of subsequent cancer and return to work. *Blood.* 1997;90:4719-4724.