Quantitative polymerase chain reaction-based chimerism in bone marrow or peripheral blood to predict acute myeloid leukemia relapse in high-risk patients: results from the KIM-PB prospective study

Allogeneic hematopoietic stem cell transplantation (HSCT) represents the best curative option for many patients with acute myeloid leukemia (AML). Nevertheless, relapses are extremely frequent, especially in patients with a high or very high disease risk index.[1,2] Although salvage treatments are expectedly more effective when given before overt recurrence, it is difficult to identify this clinically relevant time-window in such patients, because of the rapid growth kinetics of the disease leading to early relapse. Moreover, the genetic heterogeneity and clonal plasticity of AML hamper the identification of reliable and stable genetic markers to monitor minimal residual disease.[3]

After myeloablative allogeneic HSCT, reappearance of host-specific hematopoietic chimerism has been strongly associated with relapse,[4] and thus represents a practical surrogate marker of minimal residual disease. The development of quantitative polymerase chain reaction (PCR)-based assays has increased the sensitivity of chimerism monitoring dramatically[5] and a number of studies have shown its clinical utility in relapse prediction.[6-9] However, most of these studies were performed retrospectively and in highly heterogeneous cohorts of patients, and it has not yet been addressed whether the increased sensitivity of this approach might allow disease reappearance to be monitored using peripheral blood (PB) samples and in very high-risk AML patients.

To answer these questions we designed an exploratory, prospective, non-interventional, single-center study to compare chimerism monitoring in PB or bone marrow (BM) in patients undergoing myeloablative allogeneic HSCT for high-risk AML (the "KIM-PB" study, approved by the San Raffaele Ethics Committee on September 1st, 2014). The primary endpoint of the study was prediction of relapse. To monitor post-transplant fluctuations in chimerism not related to the disease, we included a control group of patients with Hodgkin or non-Hodgkin lymphomas without BM disease involvement. In the initial design, the study group comprised 30 AML patients with a high or very high disease risk index, and the control group was formed of 15 patients. Conditioning regimens and graft-versus-host disease prophylaxis were similar in both groups, and mainly based on treosulfan plus fludarabine and sirolimus plus mycophenolate. Between September 2014 and March 2016, 29 patients were enrolled into the study group (one patient was left out of the study because his donor refused to participate in the study), but only 20 were evaluable for the study endpoint, since four were excluded because of early non-relapse-related deaths, two because of disease persistence at first hematologic evaluation, two because of randomization to a non-myeloablative conditioning regimen, and one because of graft rejection. We enrolled eight patients into the control group, excluding one because of early non-relapse-related death and, upon a planned *ad interim* analysis, considered the seven evaluable patients sufficient to perform the relevant comparisons. Patient and transplant characteristics are summarized in Table 1. Chimerism was monitored using commercially available quantitative PCR-based assays (KMRtype and KMRtrack, GenDx, Utrecht, the Netherlands) (see *Online Supplementary Methods*). BM evaluations were performed on days 30, 60, 90, 180, 270

Table 1. Patient and transplant characteristics.

	Study Group (n=20) Number (%)	Control Group (n=7) Number (%)
Median age, years (range)	54 (29-72)	30 (19-36)
Sex		
Male	13 (65.0)	2 (28.6)
Female	7 (35.0)	5 (71.4)
Disease		
Acute myeloid leukemia	20 (100)	–
Hodgkin lymphoma	–	3 (42.9)
Non-Hodgkin lymphoma	–	4 (57.1)
Disease status at HSCT		
Complete remission	9 (45.0)	4 (57.1)
Active disease	11 (55.0)	3 (42.8)
Disease Risk Index		
Low	0	1 (14.2)
Intermediate	0	3 (42.9)
High	17 (85.0)	3 (42.9)
Very high	3 (15.0)	0
Donor type		
Matched related	1 (5.0)	1 (14.3)
Matched unrelated	5 (25.0)	3 (42.9)
Haploidentical	14 (70.0)	3 (42.9)
Median follow-up after HSCT, days (range)	351 (56-375)	372 (26-385)
Median time to engraftment, days (range)	20 (13-31)	18 (3-25)
Relapse	7 (35.0)	0
Median time to relapse, days (range)	72 (60-91)	
Transplant-related mortality	2 (10.0)	2 (28.6)
Median time to TRM, days (range)	84 (56-112)	70 (26-114)
End-of-study survivors	11 (55.0)	5 (71.4)

AML: acute myeloid leukemia; HL: Hodgkin leukemia; NHL: non-Hodgkin leukemia; HSCT: hematopoietic stem cell transplantation; TRM: transplant-related mortality.

and 360 after allogeneic HSCT, according to the practice of our center. PB evaluations were performed on days 3, 7 and 15, then twice a month for the first 4 months and monthly from month 5 to 12. The 1-year timespan for study follow-up was selected because approximately 80% of relapses in patients with AML a high or very high disease risk index occur during the first 12 months after allogeneic HSCT.[2] The sampling schedule is summarized in Figure 1A.

Of the 20 patients in the study group, seven relapsed (median time to relapse: 73 days; range, 61-93), 11 were alive and in complete remission at the end of follow-up and two died in remission before the end of the follow-up (at day 58 and 112 after HSCT) (Figure 1B). In the control group, five of the seven evaluable patients were alive at the end of follow-up and two died before that (at days 25 and 116), with no patient developing BM involvement. Altogether, we collected and analyzed 409 samples (PB, n=310; BM, n=99). When we compared the values of host-specific chimerism obtained from paired PB and BM samples (n=97) we detected a significantly higher host

signal in BM compared to PB, especially at late time-points (Figure 1C) and a moderate correlation between the two measurements ($R^2=0.7$; $P<0.0001$) (Figure 1D). Among several models tested, the best results in terms of relapse prediction were achieved for both PB and BM considering positive only those samples with increasing mixed chimerism of more than 19.4% the value of the previous determination. This threshold accounts for the 95% confidence interval (95% CI) of experimental error of the quantitative PCR technique (see *Online*
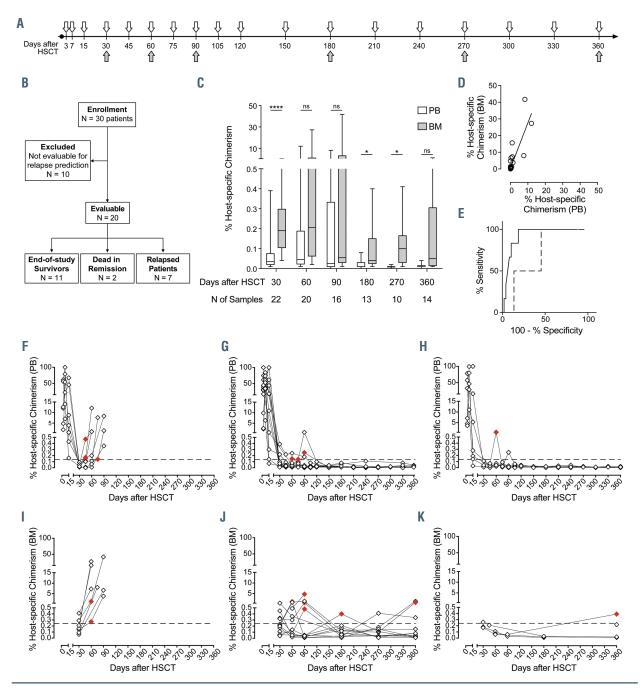

Figure 1. Outline and results of the KIM-PB prospective study. (A) Study sampling schedule. White arrows indicate peripheral blood (PB) sampling, gray arrows BM sampling. (B) Diagram summarizing the enrollment and outcome of patients in the study group. (C) Percentage of host-specific chimerism detected in 97 paired PB (white box-and-whisker plots) and BM (gray box-and-whisker plots) samples collected during the post-transplantation follow-up. Boxes display median and interquartile range, whiskers minimum and maximum values. Paired *t* tests were used for all comparisons (ns=not significant; *$P<0.05$; ****$P<0.0001$). (D) Correlation analysis between the host-specific chimerism values detected in PB (x axis) and BM (y axis) in the same 97 pairs of samples displayed in panel C. Results of a two-sided Pearson correlation analysis are shown, with a linear regression line and 95% confidence interval (95% CI). (E) Receiver operating characteristic (ROC) curves, showing sensitivity (y axis) and 1-specificity in percentages (x axis) obtained by taking into account 19 PB samples (full line) and 14 BM samples (dashed line) showing increasing mixed chimerism as compared to the previous determination. (F-K) Kinetics of host-specific chimerism during the post-transplantation follow-up measured in the PB (panels F-H) and BM (panels I-K) of patients from the study group who experienced relapse (leftmost panels), of patients from the study group who remained in remission throughout the observation period (center panels) and of patients from the control group (rightmost panels). Red diamonds indicate samples with host chimerism increasing by more than 19.4% of the value of the previous determination, and exceeding the threshold values of 0.13% for PB and 0.24% for BM (dashed lines). These are defined as "true positives" if they were observed in study group patients who relapsed, or "false positives" if they were observed in study group patients who did not relapse or in control group patients.

Supplementary Methods). Using this model, we generated receiver operating characteristic (ROC) curves comparing PB and BM in relapse prediction (Figure 1E), although it should be pointed out that a ROC curve is a suboptimal tool in this setting, due to the use of multiple longitudinal determinations for each patient. The area under the curve was 0.9201 (95% CI: 0.8438-0.9964) for PB and 0.7045 (95% CI: 0.4292-0.9799) for BM, with superior performance for PB at all possible thresholds. Based on the results of the ROC analysis and on the maximal reproducible sensitivity of the method, we further refined our model considering positive only values above 0.13% for PB and 0.24% for BM.

Based on these criteria, we documented that PB chimerism predicted four of seven relapses (57%) (Figure 1F) and gave false positive results in three of 13 patients from the study group who did not relapse (23%) (Figure 1G) and in one of seven patients from the control group (14%) (Figure 1H), all in the first 4 months after allogeneic HSCT. BM chimerism analysis predicted relapse in two of seven patients (29%) (Figure 1I) and provided false positive results in six of 13 non-relapsing patients from the study group (46%) (Figure 1J) and in one of the seven patients in the control group (14%) (Figure 1K). Of note, BM false positive results spanned over the entire follow-up, including some at very late time-points. Overall, PB chimerism showed a sensitivity of 57.14% and a specificity of 76.92% in relapse prediction, while BM analyses showed a sensitivity of 28.57% and a specificity of 45.45%, without statistically significant differences. Finally, we compared the median time from detection of increasing mixed chimerism to relapse for the two sites of sampling: the median time was 17 days for PB (range, 8-44) and 33 days for BM (range, 30-59), indicating that in most cases relapse was anticipated by a single episode of over-threshold increasing mixed chimerism. Two patients from the study cohort relapsed after the first year of follow-up (on days 603 and 661). In line with the short temporal window of prediction of chimerism by quantitative PCR, none of these patients had scored positive during the study period.

This study is the first, to our knowledge, to prospectively address the clinical utility of quantitative PCR-based chimerism monitoring in AML patients with a high or very high disease risk index undergoing allogeneic HSCT. These patients represent a significant clinical challenge since, due to the aggressiveness and chemoresistance of their disease, even if they achieve remission, this is often short-lived. This has prompted the search for approaches to anticipate the detection of leukemia recurrence and to allocate pre-emptive therapies rationally.

Although a number of recent studies have documented encouraging results obtained by quantitative PCR chimerism monitoring,[6,7,9] semiquantitative analysis of short tandem repeat polymorphisms[8,10] remains the gold standard, probably because there are approved guidelines.[11] The higher sensitivity of quantitative PCR could be perfect for PB monitoring, allowing a tighter follow-up of high-risk patients. Indeed, our results suggest clear, although not statistically significant, superiority of PB monitoring over BM. This was not only due to the more frequent sampling, allowing increasing mixed chimerism to be captured before overt relapse, but also to the higher specificity of positive signals, since BM analysis detected significant "background noise" at all time-points, possibly explained by aspiration of host BM stromal cells. Despite being superior, results obtained from PB monitoring are far from perfect, and evidently inferior to those obtained in similar studies in which specific mutations are tracked.[12] In our study cohort, only four of 20 (20%) patients carried nucleophosmin-1 mutations and for all of them, relapse was predicted by both mutation monitoring and PB quantitative PCR chimerism.

One of the main issues evidenced by our study is that the median time to relapse in high-risk patients is exceptionally short. This makes the value of early determinations crucial, especially in order to understand whether early positive chimerism is the sign of still active clearance of the patient's hematopoiesis or the first sign of disease reappearance. Tighter monitoring and the use of techniques with higher accuracy than quantitative PCR, such as droplet digital PCR,[13-15] may be key to address this crucial hurdle. It should be noted in fact that in several true positive cases from our study, multiple serial determinations confirmed the initial positivity, whereas in false positive cases there was often a decrease of chimerism at the netx time-point. The short experimental turn-around time of quantitative PCR and its relatively affordable cost could easily allow weekly PB determinations in very high-risk patients or in case of a first warning.

The very narrow time window of prediction also raises the relevant issue of whether and how the information acquired could be translated into a therapeutic benefit: whereas immune interventions are known to take 2 to 4 weeks to show any effect, new drugs, such as FLT3 inhibitors, have extremely rapid activity. This might allow physicians to implement them in a pre-emptive fashion, based on an increase in host chimerism, sparing the toxicities and economic burden that accompany prophylactic administration[16] and possibly also the cytopenias that are commonly observed upon use in overt relapse.[17]

In conclusion, even considering its limited sensitivity and specificity in relapse prediction, the low invasiveness, rapid analytical turnaround and nearly universal applicability of PB quantitative PCR-based monitoring might make it a useful complement to the still too limited armamentarium that is available for monitoring high-risk AML. We do, however, have to note that the cohort analyzed in our study is rather small, due to its single-center nature and to the intrinsic hurdles of prospective trials; therefore, further and larger multicenter studies should be undertaken to confirm these results, and to address the crucial issue of standardization of this method across different laboratories.

Valentina Gambacorta,[1] Riccardo Parolini,[1] Elisabetta Xue,[1,2] Raffaella Greco,[2] Evelien E. Bouwmans,[3] Cristina Toffalori,[1] Fabio Giglio,[2] Andrea Assanelli,[2] Maria Teresa Lupo Stanghellini,[2] Alessandro Ambrosi,[4] Benedetta Mazzi,[5] Wietse Mulder,[3] Consuelo Corti,[2] Jacopo Peccatori,[2] Fabio Ciceri[2,4#] and Luca Vago[1,2#]

[1]Unit of Immunogenetics, Leukemia Genomics and Immunobiology, IRCCS San Raffaele Scientific Institute, Milano, Italy; [2]Hematology and Bone Marrow Transplantation Unit, IRCCS San Raffaele Scientific Institute, Milano, Italy; [3]GenDx, Utrecht, the Netherlands; [4]Vita-Salute San Raffaele University, Milano, Italy and [5]HLA and Chimerism Laboratory, IRCCS San Raffaele Scientific Institute, Milano, Italy

[#]FC and LV contributed equally as co-senior authors.

Correspondence:

LUCA VAGO - vago.luca@hsr.it

References

1. Horowitz M, Schreiber H, Elder A, et al. Epidemiology and biology of relapse after stem cell transplantation. Bone Marrow Transplant. 2018;53(11):1379-1389.
2. Armand P, Gibson CJ, Cutler C, et al. A disease risk index for patients undergoing allogeneic stem cell transplantation. Blood. 2012;120(4):905-913.
3. Schuurhuis GJ, Heuser M, Freeman S, et al. Minimal/measurable residual disease in AML: a consensus document from the European LeukemiaNet MRD Working Party. Blood. 2018;131(12):1275-1291.
4. Bader P, Kreyenberg H, Hoelle W, et al. Increasing mixed chimerism defines a high-risk group of childhood acute myelogenous leukemia patients after allogeneic stem cell transplantation where pre-emptive immunotherapy may be effective. Bone Marrow Transplant. 2004;33(8):815-821.
5. Alizadeh M, Bernard M, Danic B, et al. Quantitative assessment of hematopoietic chimerism after bone marrow transplantation by real-time quantitative polymerase chain reaction. Blood. 2002; 99(12):4618-4625.
6. Qin X-Y, Li G-X, Qin Y-Z, et al. Quantitative chimerism: an independent acute leukemia prognosis indicator following allogeneic hematopoietic SCT. Bone Marrow Transplant. 2014;49(10):1269-1277.
7. Jacque N, Nguyen S, Golmard J-L, et al. Chimerism analysis in peripheral blood using indel quantitative real-time PCR is a useful tool to predict post-transplant relapse in acute leukemia. Bone Marrow Transplant. 2015;50(2):259-265.
8. Ahci M, Stempelmann K, Buttkereit U, et al. Clinical utility of quantitative PCR for chimerism and engraftment monitoring after allogeneic stem cell transplantation for hematologic malignancies. Biol Blood Marrow Transplant. 2017;23(10):1658-1668.
9. Sellmann L, Rabe K, Bünting I, et al. Diagnostic value of highly-sensitive chimerism analysis after allogeneic stem cell transplantation. Bone Marrow Transplant. 2018;53(11):1457-1465.
10. Bader P, Niethammer D, Willasch A, Kreyenberg H, Klingebiel T. How and when should we monitor chimerism after allogeneic stem cell transplantation? Bone Marrow Transplant. 2005;35(2):107-119.
11. Clark JR, Scott SD, Jack AL, et al. Monitoring of chimerism following allogeneic haematopoietic stem cell transplantation (HSCT): technical recommendations for the use of short tandem repeat (STR) based techniques, on behalf of the United Kingdom National External Quality Assessment Service. Br J Haematol. 2015;168(1):26-37.
12. Xue E, Tresoldi C, Sala E, et al. Longitudinal qPCR monitoring of nucleophosmin 1 mutations after allogeneic hematopoietic stem cell transplantation to predict AML relapse. Bone Marrow Transplant. 2016;51(3):466-469.
13. Valero-Garcia J, González-Espinosa MDC, Barrios M, et al. Earlier relapse detection after allogeneic haematopoietic stem cell transplantation by chimerism assays: digital PCR versus quantitative real-time PCR of insertion/deletion polymorphisms. PLoS One. 2019;14(2):e0212708.
14. Waterhouse M, Pfeifer D, Duque-Afonso J, et al. Droplet digital PCR for the simultaneous analysis of minimal residual disease and hematopoietic chimerism after allogeneic cell transplantation. Clin Chem Lab Med. 2019;57(5):641-647.
15. Stahl T, Böhme MU, Kröger N, Fehse B. Digital PCR to assess hematopoietic chimerism after allogeneic stem cell transplantation. Exp Hematol. 2015;43(6):462-468.
16. Levis MJ, Chen Y-B, Hamadani M, Horowitz MM, Jones RJ. FLT3 inhibitor maintenance after allogeneic transplantation: is a placebo-controlled, randomized trial ethical? J Clin Oncol. 2019;37(19):1604-1607.
17. De Freitas T, Marktel S, Piemontese S, et al. High rate of hematological responses to sorafenib in FLT3-ITD acute myeloid leukemia relapsed after allogeneic hematopoietic stem cell transplantation. Eur J Haematol. 2016;96(6):629-636.