CHAPTER 7

Leukemic stem cell quantification in newly diagnosed chronic phase chronic myeloid leukemia (CML-CP) predicts response to nilotinib: results from the ENEST1st stem cell substudy

Submitted

Noortje Thielen[1], Johan Richter[2], Matthias Baldauf[3], Gisela Barbany[4], Thoas Fioretos[5], Frank Giles[6], Bjørn-Tore Gjertsen[7], Andreas Hochhaus[8], Gerrit Jan Schuurhuis[1], Sieghart Sopper[9], Leif Stenke[10], Sarah Thunberg[10], Zlatko Trajanoski[3], Dominik Wolf[11], Gert Ossenkoppele[1], Kimmo Porkka[12], Jeroen Janssen[1], Satu Mustjoki[12]

[1]Department of Hematology, VU University Medical Center, Amsterdam, the Netherlands; [2]Department of Hematology and Vascular Disorders, Skåne University Hospital, Lund, Sweden; [3]Divison of Bioinformatics, Biocenter, Innsbruck Medical University, Austria; [4]Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden; [5]Department of Clinical Genetics, Lund University, Lund, Sweden; [6] NMDTI, Robert H. Lurie Comprehensive Cancer Center of Northwestern University, Chicago, Illinois, USA; [7]Department of Hematology, University of Bergen, Bergen, Norway; [8]Hematology/Oncology, Universitätsklinikum Jena, Jena, Germany; [9]Department of Hematology and Oncology, Innsbruck Medical University and Tyrolean Cancer Research Center, Innsbruck, Austria; [10]Department of Medicine, Karolinska University Hospital, Stockholm, Sweden; [11] Medical Clinic 3, Oncology, Hematology and Rheumatology, University Hospital Bonn (UKB), Bonn, Germany; [12]Hematology Research Unit Helsinki, University of Helsinki and Comprehensive Cancer Center, Helsinki University Hospital, Helsinki, Finland

ABSTRACT

Leukemic stem cells (LSCs) are considered to be an important source of resistance to tyrosine kinase inhibitor therapy in CML. We addressed their predictive value for response to nilotinib therapy both at diagnosis and after 3 months of therapy. The LSC burden was assessed by multiparameter flow cytometry (MPFC; n=44) and sorting of $CD34^{+}CD38^{-}$ cells plus FISH (n=39). By MPFC, residual normal stem cells were detected in 41% of patients at diagnosis, while the mean percentage of Ph+ stem cells by FISH was 85%. LSC burden at diagnosis correlated with blood and bone marrow blast count, hemoglobin, spleen size, Sokal score, molecular response at 3 (MPFC and FISH), 6 (MPFC), 9 (FISH) and 15 months (FISH) ($p < 0.05$) and occurrence of hematological toxicity. During nilotinib therapy, the proportion of LSCs rapidly decreased as measured by both methods. At 3 months, residual LSCs were only detected in 12% of patients by MPFC analysis, while 67% of patients had Ph+ stem cells left when analyzed by FISH (mean percentage of LSCs 9%). In conclusion, the composition of the hematopoietic stem cell compartment as analyzed by two independent methodologies reflects clinical parameters and correlates with outcome in first-line nilotinib treated CML patients.

INTRODUCTION

The introduction of the tyrosine kinase inhibitor (TKI) imatinib and, more recently, of the second generation TKIs nilotinib and dasatinib marked a revolution in the treatment of chronic myeloid leukemia (CML) patients. Most patients attain excellent cytogenetic and molecular responses to these drugs and remain free of progression to accelerated phase and blast crisis.[1-3] However, in the majority of patients the cure of the disease will probably not be achieved by TKIs alone due to the persistence of leukemic stem cells (LSCs). According to *in-vitro* studies LSCs are inherently insensitive to currently available TKIs.[4-10] Patients with high Sokal or Euro scores, based on hematological parameters, spleen size and age, and those with slow reduction of *BCR-ABL1* transcripts in response to starting treatment are at increased risk of treatment failure.[11-13] Moreover, our study groups have previously demonstrated that the composition of the $CD34^{+}CD38^{-}$ stem cell compartment at diagnosis, as assessed by multiparameter flowcytometry (MPFC) or stem and progenitor cell sorting followed by fluorescence in-situ hybridization (FISH), not only predicts later response, but also hematological toxicity during TKI treatment.[14,15] Reliable and early response prediction is of great value as it enables personalized tailoring of therapy before progression to advanced phases in patients who are deemed to respond poorly. In the current study, we set out to assess whether analysis of the stem cell compartment at diagnosis and stem cell dynamics early after treatment initiation, such as the LSC reduction or, conversely, LSC persistence, has additional value in predicting outcome in nilotinib-treated chronic phase CML (CP-CML) patients. This translational project was performed as a substudy to the ENEST1st clinical study in which newly diagnosed CML-CP were treated with nilotinib 300 mg BID. The project is a collaborative effort of the Nordic CML Study Group, the Central European Leukemia Study Group (CELSG) and selected Dutch HOVON centers.

PATIENTS AND MATERIALS

Patients

Patients were eligible for the stem cell study when they were treated in the multicenter and multinational phase IIIb clinical trial Evaluating Nilotinib Efficacy and Safety as First-Line Treatment (ENEST1st study). Patients could be included in this study when they had newly diagnosed chronic phase CML, were 18 years of age or older and had WHO performance status ≤ 2. Treatment consisted of nilotinib 300 mg BID. An adequate bone marrow (BM) aspirate and/or peripheral blood (PB) sample was required at diagnosis, after 1 month and after 3 months of therapy.

The study was performed in accordance with the declaration of Helsinki. All patients provided written informed consent. The study was approved by the local institutional review boards of all participating centers and is registered at ClinicalTrials.gov (identifier NCT01061177).

Materials

For this study, MPFC and sorting plus FISH were performed in parallel and independently in selected and qualified core laboratories (MPFC: Stem Cell Laboratory of the VU University Medical Center, Amsterdam, the Netherlands; sorting plus FISH: Helsinki University Central Hospital, Helsinki, Finland, Department of Clinical Genetics, Lund University, Lund, Sweden and Karolinska Institutet, Stockholm, Sweden).

The assessment of the $CD34^{+}CD38^{-}$ compartment was performed by MPFC and sorting plus FISH at diagnosis, by MPFC after 1 month of treatment and by both MPFC and sorting plus FISH after 3 months of treatment.

Mononuclear cells were isolated from BM and/or PB by density gradient centrifugation using Ficoll-paque (1.077 g/ml; Amersham Biosciences, Freiburg, Germany), followed by a red cell lysis step with lysis buffer (155 mM NH_4Cl, 10 mM $KHCO_3$, 0.1 mM Na_2EDTA, pH 7.4) for 10 min at 4 ºC.

Multiparameter flow cytometry

MPFC was used as previously described for separation of normal and leukemic stem cells.[14] Briefly, freshly isolated mononuclear cells were incubated with monoclonal antibodies for 15 min at room temperature, washed once in PBS containing 0.1% human serum albumin (HSA) and analyzed by flow cytometry. The fluorescent-labeled antibodies used were anti-CD34 phycoerythrin-Cy7 (PE-Cy7), anti-CD45 fluorescein isothiocyanate (FITC), anti-CD38 allophycocyanin (APC), anti-CD7-phycoerythrin (PE), anti-CD11b-PE, anti-CD56-PE, anti-CD90-PE and Via-Probe (7-amino-actinomycin D, 7AAD) (all from BDBiosciences, Erembodegem, Belgium). IgG1-PE was used as a control for PE labeled anti-CD7, -CD56 and -CD90, IgG2a-PE was used for PE-labeled anti-CD11b. PBS was used as a control for anti-CD34-PE-Cy7, CD38-APC and -CD45-FITC. 7-AAD was included to gate out apoptotic/dead cells before stem cell assessment. Data were acquired using a 3-laser FACSCanto or FACSFortessa flow cytometer (BD Biosciences, Erembodegem, Belgium) and analyzed using FACSDiva software (BD Biosciences).

Flow cytometry gating strategy

The gating strategy we used in this study was as previously described.[14] In short, the lowest 1% of CD38 expressing cells in the $CD34^{+}$ compartment was designated as the stem cell compartment. In a CD34CD45 plot, these cells can display two different patterns. In the first, there are two separate populations visible, designated as $CD34^{+}CD45^{high}$ with relatively high forward and side scatter light properties (FSC and SSC respectively) and $CD34^{+}CD45^{low}$ with lower FSC/SSC properties, respectively. As we previously showed, the latter consists of Ph negative cells, always negative for aberrant markers like CD7, CD11b or CD56 and low in CD90 expression, while the cells in the former population are almost all Philadelphia-positive (Ph+) and may display CD7, CD11b or CD56 positivity and high

CD90 expression. The second pattern involves patients with only one single $CD34^+CD45^+$ population. These cells exhibit relatively high FSC/SSC signals and express one or more of the aberrant lineage markers together with high CD90 expression, and were demonstrated to be Ph+. For reasons of clarity the cells in the Ph+ cell population are referred to as LSCs and the Ph-negative cells as nHSCs, although nu functional assessment was performed.

For a reliable distinction between malignant and benign stem cells, at least 10 clustered cells per population were required. To be able to achieve this we aimed to analyze at least 1 million mononuclear cells (MNCs).

Sorting plus FISH

The protocol for sorting plus FISH has been described previously.[15] In short, the mononuclear cell fraction was separated by Ficoll centrifugation (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). Next, $CD34^+$ cells were enriched with paramagnetic beads (Miltenyi Biotech, BergischGladbach, Germany), stained with CD34 and CD38 antibodies (BD Biosciences, San Jose, CA, USA) after which the primitive cell fractions were sorted by a flow cytometer (FACSAria, BD Biosciences) into two populations; the first representing 5% with the lowest expression of CD38 ($CD34^+CD38^-$) and the second representing 80% with highest expression of CD38 ($CD34^+CD38^+$). Like the definition used for MPFC, Ph+ $CD34^+CD38^-$ cells are further designated as LSCs and Ph+ $CD34^+CD38^+$ cells as leukemic progenitor cells (LPCs).

After sorting, cytospin slides were made from both primitive cell fractions and additionally from unseparated whole BM samples from which red cells were lysed with FACS lysing buffer (BD Biosciences). The proportion of Ph+ cells was determined by FISH using dual fusion dual color *BCR-ABL1* probes (Vysis, Abbot, Downers Crove, IL, USA).

By sorting plus FISH 1.000 cells were analyzed in each fraction, in contrast to conventional FISH which most often uses only 200 unselected cells. Therefore, the sensitivity increases considerably up to a detection limit of one Ph+ stem cell out of 100.000 to 1.000.000 mononuclear cells. Although in this study, neither for FISH nor for MPFC functional stem cell assays were performed, in the following, for simplicity, we refer to these populations as stem- and progenitor cells, respectively.

Before the trial start, quality control rounds were performed in the different stem cell laboratories to assess inter-laboratory consistency. The same FISH sample slides were circulated among the laboratories and FISH slides were quantitatively evaluated. In addition, at two different time points, a BM aspirate sample was taken from a newly diagnosed CML patient and an aliquot of the sample was sent to each participating stem laboratory to assess reproducibility for the sorting and FISH procedure. The results were compared between various centers and a > 90% concordance was achieved among the different laboratories.

STATISTICAL ANALYSIS

The primary objective of this study addressed the question whether the composition of the $CD34^+CD38^-$ compartment at diagnosis and during follow up, determined by either MPFC or sorting and FISH, has a predictive value for reaching MR^4 at 18 months.
For analysis of MPFC results, samples were divided in two classes: either nHSCs detectable (in any quantity) or only LSCs detectable at diagnosis. Conversely, at follow up, a binominal distinction was made between samples with only nHSCs and samples with detectable residual LSCs. Differences between those groups regarding a set of parameters were investigated using the Exact Wilcoxon Mann-Whitney Rank Sum test.
For sorting plus FISH analysis, outcome was given on a continuous scale, i.e. the percentage of LSCs of all stem cells. Correlation coefficients were calculated using Kendall's tau-b statistic. Sokal and Euro-scores were calculated according to Sokal et al. and Hasford et al.[12,13] Molecular response was centrally analyzed in the EUTOS reference laboratory. Major molecular response (MR^3) was defined as a *BCR–ABL1* transcript level of ≤ 0.1%, MR^4 as a measurable *BCR-ABL1* transcript level of ≤ 0.01% or an undetectable transcript level with at least 10.000 *ABL1* copies and $MR^{4.5}$ as a measurable *BCR-ABL1* transcript level of ≤ 0.0032% or an undetectable transcript level with at least 32.000 *ABL1* copies, all according to the international scale. For calculation of mean *BCR-ABL1* levels, undetectable transcript levels with at least 10.000 *ABL1* copies were counted as zero.
Statistical analyses were performed with the R software under consideration of a) Fisher's exact test for pairs of categorical variables, b) Exact Wilcoxon-Mann-Whitney Rank Sum test for dichotomous independent/continuous dependent variables, and c) Kendall's rank correlation tau-b test for pairs of continuous random variables. A logistic regression model was used for predicting response for percentage of Ph+ stem cells by FISH. For all analyses, p-values <0.05 were considered to indicate a statistically significant difference. For all analyses, p-values <0.05 were considered to indicate statistical significance outcome.

RESULTS

The demographic and clinical parameters of all patients are shown in Table 1. MPFC of the $CD34^+CD38^-$ cell compartment (BM or PB) was available for 48 patients at diagnosis. Four patient samples were not evaluable due to either insufficient cell numbers in the stem cell compartment or indistinct flow patterns. Of the remaining 44 patients, residual nHSCs were detectable in 18 patients (41%), in either BM and/or PB. In 5 of these patients, the $CD34^+CD38^-$ compartment consisted only of nHSCs, thus without the presence of LSCs. In 3 of these, FISH was performed as well, and all had low LSC burden (1.2%, 14.4% and 29.5%). In the samples of the other 26 patients (59%) nHSCs could not be identified. In 28 out of 30 patients who had both BM and PB samples evaluable, the stem cell patterns (either residual nHSCS or no residual nHSCS) were identical between both compartments.

Only two patients had inconsistent patterns in BM and PB: in one patient residual nHSCs were detected in PB, while BM revealed only LSCs, whereas in the second patient the opposite was observed.

To assess whether the total LSC burden as measured by MPFC differed between patients with or without residual nHSCs, we measured the percentage of LSCs from the total MNCs (the number of designated LSCs divided by the total number of MNCs): LSC burden was higher in patients without residual nHSCs compared to those with residual nHSCs (mean 0.018% vs 0.003% respectively; median 0.011% vs 0.001% respectively, range 0.0007-0.11% and 0.0002-0.02%, $p < 0.001$).

For sorting plus FISH, 46 diagnostic patient samples were available, of which 39 were assessable (Table 1). The average percentage of Ph+ stem cells within the $CD34^{+}CD38^{-}$ population in the diagnostic samples was lower than in the progenitor and total BM fraction (85% vs 96% vs 96%, respectively, $p = 0.023$ and $p = 0.12$) (Figure 1A). Figure 1B presents the concordance between the two methods. The Kendall's tau-b rank correlation coefficient between the two techniques based on the percentages of LSCs as detected by MPFC versus the percentage of LSCs by sorting plus FISH was 0.297 ($p = 0.027$).

Table 1. Demographic and clinical data at diagnosis.

	All	MPFC	sorting + FISH
n	50	48	40*
Sex (M:F)	33:17	31:17	25:15
Age (yr)	59 (20-82)	58 (20-75)	58 (27-82)
WBC (10^9/l)	61 (3.7-307)	62 (3.7-307)	76 (9.7-307)
Platelet count (109/l)	377 (98-1485)	388 (98-1485)	402 (98-1485)
Spleen size (cm)	0 (0-30)	0 (0-30)	0 (0-30)
Sokal score	0.9 (0,46-2,28)	0.9 (0.46-2.28)	0.9 (0.66-2.28)
Euro score	857 (83-1775)	857 (83-1775)	857 (83-1775)

For MPFC, 48 patients were assessable, either at diagnosis (n = 44), at 1 (n = 34) and/or at 3 months (n = 34). *For sorting plus FISH, 40 patients were assessable; either at diagnosis (n = 39) and/or at 3 months (n = 15). Of one patient only a 3-month sample was available.

Data for white blood counts (WBC), platelet counts, spleen sizes, Sokal and Euro scores are shown as medians with ranges in parentheses.

Abbreviations: n, number; M, male; F, female; yr, years

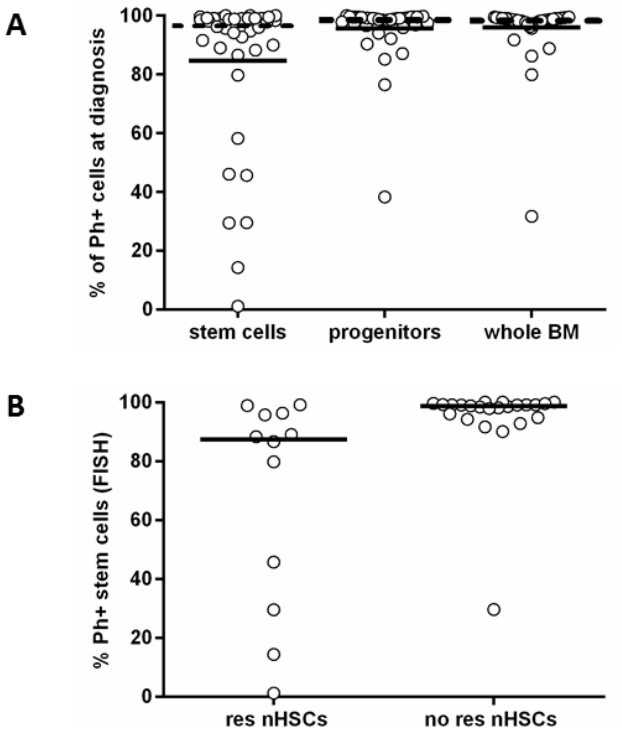


Figure 1. The proportion of Ph+ cells at diagnosis in the stem and progenitor cell compartments and whole BM.

(A) The proportion of Ph+ cells was analyzed by sorting plus FISH in stem (CD34+CD38-), progenitor (CD34+CD38+) and whole BM compartments. The mean and median percentages of Ph+ cells were 85% and 98% in the stem cell fraction, 96% and 99% in the progenitor fraction and 96% and 99% in the whole BM, respectively. Black lines represent mean and dashed lines median values.

(B) Concordance between results of MPFC (residual nHSCs; blue, versus no residual nHSCs; red) and sorting plus FISH (percentage of Ph+ stem cells) at diagnosis (n = 34). Black lines represent median values. Abbreviations: Res nHSCs, residual nHSCs present; no res nHSCs, no residual nHSCs present.

LSC burden by MPFC and sorting plus FISH at diagnosis correlates with clinical and laboratory characteristics

Patients with residual nHSCs by MPFC and patients with low LSC burden by sorting plus FISH had lower blood and marrow blast counts ($p < 0.001$ and $p = 0.044$ for MPFC; $p < 0.001$ and $p = 0.003$ for sorting plus FISH respectively), smaller spleen size ($p < 0.001$ for MPFC; $p = 0.006$ for sorting plus FISH), lower platelet counts ($p = 0.027$ for MPFC; NS for sorting plus FISH), lower Sokal scores ($p < 0.001$ for MPFC; $p = 0.035$ for sorting plus FISH) (Figure 2), lower WBC counts ($p = 0.006$ for sorting plus FISH; NS for MPFC) and higher hemoglobin levels ($p = 0.010$ for MPFC; $p = 0.001$ for sorting plus FISH).

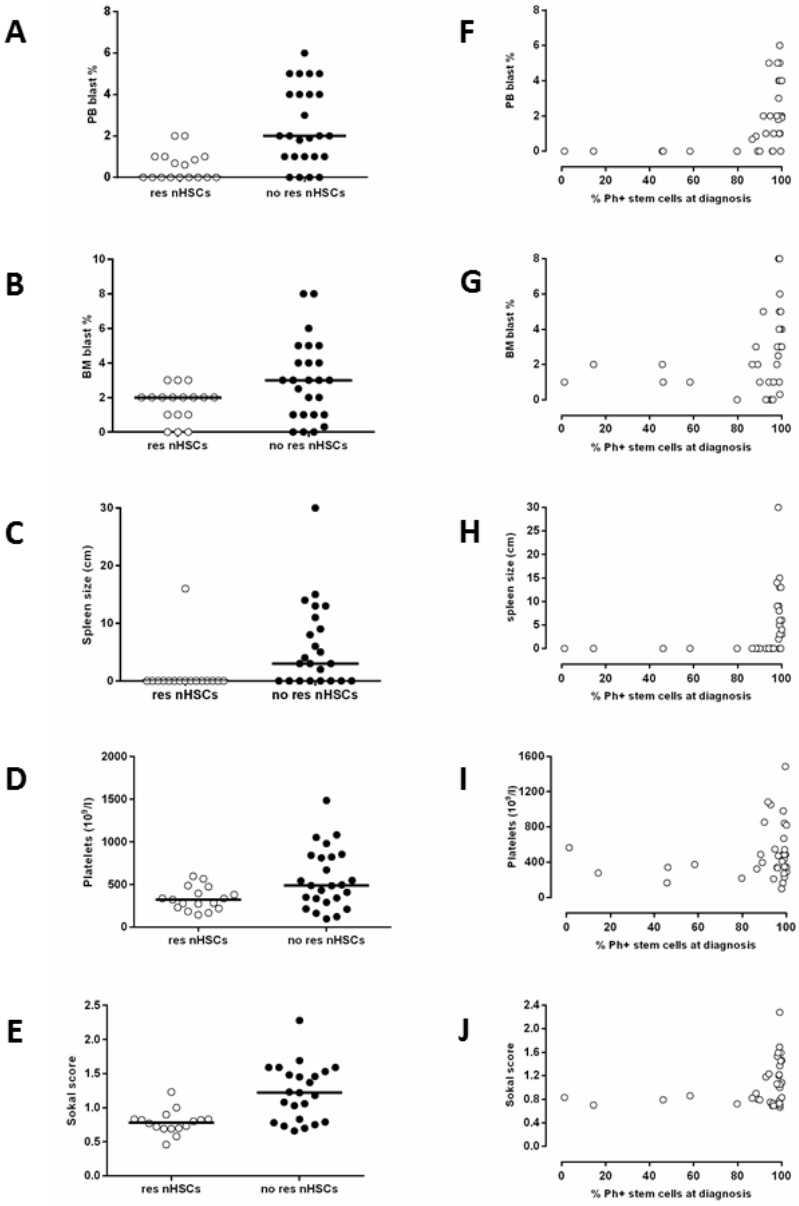


Figure 2. Correlation between LSC burden and clinical parameters.

At the diagnosis, LSCs were measured by MPFC (A-E) and sorting plus FISH (F-J), and results were correlated with biological disease variables. Panels show PB blast percentages (Fig A and F: $p < 0.001$ and $p < 0.001$), BM blast percentages (Fig B and G: $p = 0.044$ and $p = 0.003$), spleen sizes (Fig C and Fig H: $p < 0.001$ and $p = 0.006$), platelet counts (Fig D and Fig I: $p = 0.027$ and $p = 0.33$), and Sokal scores (Fig E and Fig J: $p < 0.001$ and $p = 0.035$). Black lines represent median values. Abbreviations: Res nHSCs, residual nHSCs present; no res nHSCs, no residual nHSCs present.

LSC burden at diagnosis correlates with molecular response during early and late treatment

Patients with residual nHSCs and thereby lower LSC levels at diagnosis as detected by MPFC attained deeper molecular responses at 3 and 6 months after starting nilotinib treatment ($p = 0.020$ and $p = 0.024$ respectively) (Figure 3A) and attained MR^3 earlier than patients with only LSCs, although the difference did not reach statistical significance. Although all evaluable patients in both MPFC groups achieved a *BCR-ABL1* level < 10% at 3 months, only 50% of patients (10 of 20) without nHSCs at diagnosis reached a *BCR-ABL1* level of ≤ 1% at 3 months, as compared to 100% of the 16 patients with residual nHSCs at diagnosis achieving *BCR-ABL1* levels below 1%. In addition, all patients of the latter group reached a *BCR-ABL1* level ≤ 1% at 6 months, whereas only 78% of patients (18 of 23) with only LSCs at diagnosis did. Twenty-five percent of patients with only LSCs and 14% of patients with residual nHSCs at diagnosis did not attain MR^3 at 12 months. After 18 months, 22% of patients with only LSCs and 19% of patients with residual nHSCs did not attain MR^3. There was no significant difference between the groups in respect of achieving MR^4 at 18 months (primary endpoint), although a slightly higher proportion of patients with residual nHSCs at diagnosis accomplished this endpoint (50% versus 39%, $p = 0.53$).

LSC burden at diagnosis as a predictive factor for later molecular response could also be validated with the sorting plus FISH method. Patients with lower LSC burden achieved superior molecular responses at 3 months ($p = 0.023$), 9 months ($p = 0.007$) and 15 months ($p = 0.034$) (Figure 3B), and although not significant, a tendency towards a better response was seen at 6, 12, 18 and 21 months (p value between 0.05 and 0.08) for patients with lower LSC burden. FISH analysis also predicted milestones; all patients not achieving MR^3 at 12 or 18 months had a high LSC burden at diagnosis (> 80% Ph+ cells) (Figure 3B).

Based on the logistic regression analysis, the proportion of LSCs did not significantly impact the likelihood of achieving MR^4 at 18 months ($p = 0.136$). However, the odds for a patient with a hypothetical LSC value of zero being in MR^4 at 18 months was as high as 26.6, and for every one unit increase (+1 %) in the LSC number, the odds decreased by 1.04. Therefore, the odds for a patient showing MR^4 at 18 months decrease with increasing LSC number.

Although in most studies with large patient cohorts Sokal score predicts response depth, this was not observed in our small study population (Supplemental Figure 1). In fact, patients with intermediate Sokal scores obtained the best responses. As shown in Figure 3 and in Supplemental Figure 1, in this rather small group of patients, stem cell compartment analysis by MPFC or FISH thus seems to have a better predictive value for response than the conventional risk scores.

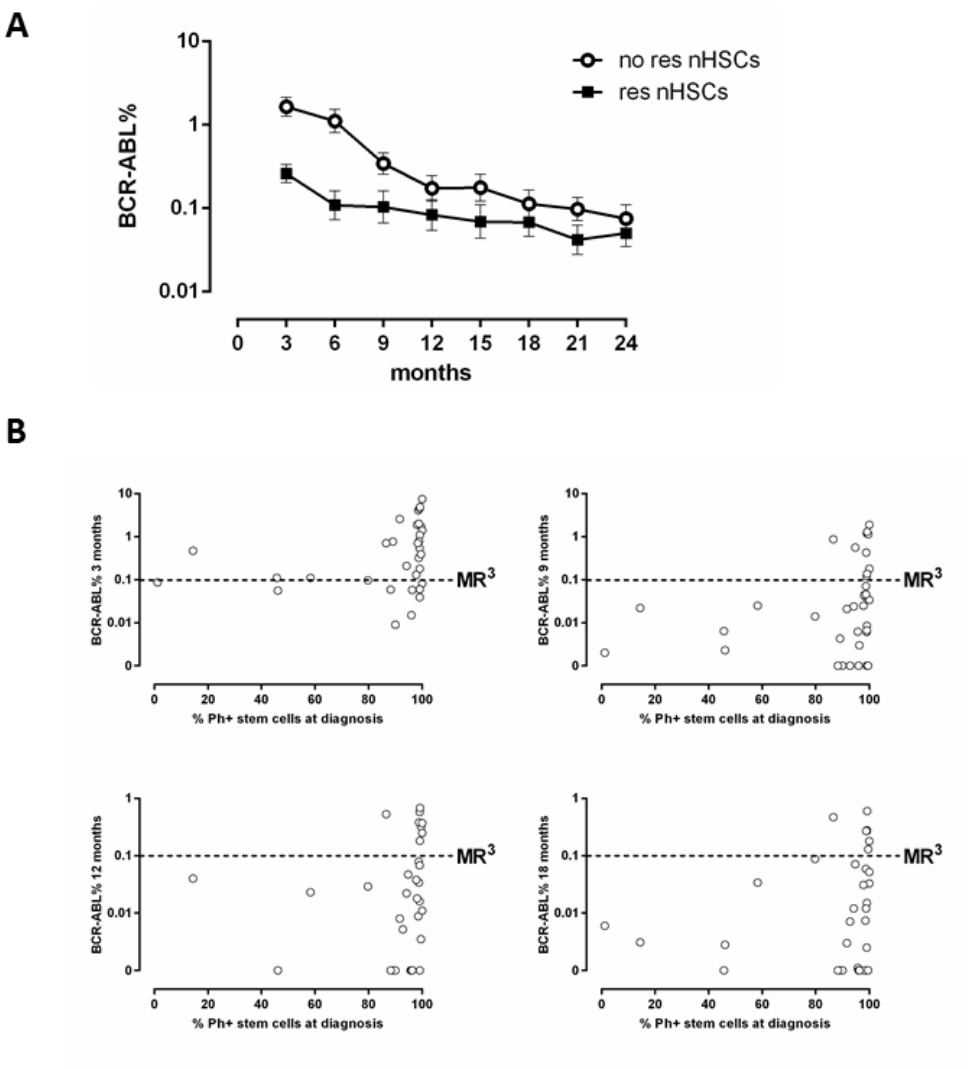


Figure 3. Correlation of LSCs with molecular response.

BCR-ABL1 transcript levels were measured every 3 months after the start of nilotinib treatment. All patients were followed for at least 24 months, but from 3 patients response data were not available. **(A)** Mean *BCR-ABL1* levels during treatment of patients with (blue line) and without (red line) residual nHSCs at diagnosis measured by MPFC. Error bars represent standard errors of the mean. P-values for the difference between groups for 3 and 6 months response were 0.020 and 0.024, respectively. Abbreviations: Res nHSCs, residual nHSCs present; no res nHSCs, no residual nHSCs present.

(B) Correlation between LSC burden at diagnosis measured by sorting plus FISH and molecular response at 3 (p = 0.023), 9 (p = 0.007), 12 (p = 0.066) and 18 months (p = 0.082).

Hematological adverse events are more frequent in patients with high LSC burden

During the 24 month observation period, hematological adverse events (cytopenias) were more frequent in patients without detectable nHSCs at diagnosis (Figure 4A). Interestingly, in six out of seven patients who experienced any hematological toxicity (grade 2-4) no residual nHSCs could be detected by MPFC at diagnosis. Notably, this observation was concordant with the data generated by sorting plus FISH, as all patients experiencing any hematological toxicities had a high LSC burden (> 89% Ph+ cells) at diagnosis (Figure 4B).

LSCs rapidly disappear during nilotinib treatment

In-vitro studies have suggested that TKIs do not effectively target LSCs or that their survival is BCR-ABL1 independent.[4-10] We here show by using two independent methods, that TKI treatment *in-vivo* rapidly reduces LSC burden in the BM. At 1 and 3 months after starting the treatment 45 and 43 patient samples, respectively, were available for MPFC. From both of these time points samples were evaluable by MPFC in 34 patients. While at diagnosis LSCs were detectable by MPFC in 89% of patients (39 of 44), the level dropped to 18% (6 of 34 evaluable patients) at 1 month and 12% (4 of 34 evaluable patients) at 3 months. Notably, in all 4 patients with persisting LSCs at month 3, the stem cell compartment at diagnosis consisted of LSCs only. One of these patients never reached MR^3, two eventually attained a good response (MR^3 at 9 and 12 months, respectively) and in one patient the molecular response was unknown.

For sorting plus FISH, at 3 months 15 patient samples were evaluable out of 23 samples received. Residual LSCs were detected in the majority of patients (67%), however, the levels were relatively low (mean 10%, median 0.3%), and only three patients (20%) had more than 5% Ph+ $CD34^+CD38^-$ cells (Supplemental Figure 2). Two of three patients with >5% of LSCs by FISH had also detectable LSCs by MPFC method and one patient was not evaluable. One of these three patients lost MR^3 at 18 months, one obtained a good response and one patient was lost to follow-up. Similarly in the progenitor and overall bone marrow compartments the percentage of Ph+ cells was low at 3 months time-point (Supplemental Figure 2).

Due to the rapid LSC elimination kinetics during nilotinib therapy and the small patient numbers with persisting LSC at the early time-points after treatment initiation, we were not able to properly analyze the association between molecular response and persisting LSCs at month 1 and 3.

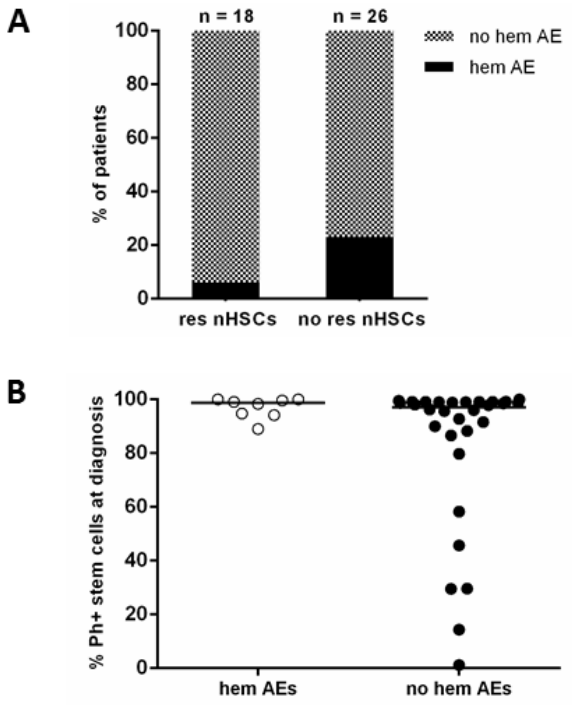


Figure 4. Correlation of LSCs with the occurrence of hematological toxicity during nilotinib therapy. Grade 2-4 hematological toxicity during nilotinib therapy was defined according to standard clinical trial procedures and patients were divided in 2 groups based on the occurrence of any hematological toxicity.

(A) The proportion of patients with hematological toxicity according to stem cell compartment composition as measured by MPFC. Abbreviations: Res nHSCs, residual nHSCs present; no res nHSCs, no residual nHSCs present; AE, adverse events.

(B) The proportion of LSCs as measured by sorting plus FISH in patients with and without hematological toxicity.

DISCUSSION

LSCs in CML are quiescent, self-renewing cells that derive from nHSCs after acquiring the chromosomal translocation (9;22). As a consequence of the translocation, the LSCs are characterized by deregulated cell-cycle activity, reduced apoptosis and adhesion, evasion of innate immunity, increased self-renewal and longevity.[16,17] As LSCs have been suggested to persist during TKI treatment, they may act as a reservoir for resistant clones and eventually be responsible for progression to advanced phases. In this study we were able to demonstrate that LSC identification and quantification, by two independent methods (MPFC and stem cell sorting plus FISH) reflects the biology of the disease in newly diagnosed CML-CP patients and allows the prediction of important response milestones as well as hematological toxicities during therapy with the 2nd generation TKI nilotinib.

The proportion of LSCs markedly differed between individual CML patients at the diagnosis although all patients were in chronic phase. The LSC burden correlated well with PB and BM blast percentage, hemoglobin, spleen size and Sokal risk score. This result is in agreement with our previous studies done both with the MPFC and sorting plus FISH.[14,15] Interestingly, in the current study we were not able to detect LSCs at diagnosis by MPFC in five patients. This could be due to limited sensitivity of the MPFC assay in case LSC numbers are low, and indeed, with the sorting plus FISH analysis we were able to determine that these patients had the lowest percentages of Ph+ cells. However, it could also be due to the fact that in the MPFC assay we only analyzed the lowest 1% of $CD34^{+}CD38^{-}$ cells whereas in the FISH method the stem cell compartment was defined as the lowest 5%.

Although the majority of patients achieve extremely good responses with the 2nd generation TKIs, in our patient cohort, there was a clear association between the composition of the stem cell compartment at diagnosis and later response. The presence of residual nHSCs detected by MPFC or < 20% Ph+ cells detected by sorting plus FISH predicted for a 86% and 100% MR^3 rate at 12 months, while 25% and 27% of patients who did not meet these conditions by MPFC or sorting plus FISH failed to achieve this endpoint. Notably, stem cell analysis by MPFC or FISH seems to outcompete the conventional risk scores, such as the Sokal score, in the prediction of molecular response, but this needs to be confirmed in a larger patient cohort.

Achievement of an early molecular response at 3 and 6 months has recently been included in the updated ELN criteria.[18] It has been demonstrated by several studies, that the *BCR-ABL1* level at these time points is a good predictor for later major molecular response, overall survival and progression free survival.[11,19-24] Although all patients in this study reached the ELN defined molecular goal of a *BCR-ABL1* ratio <10% at 3 months (which shows the high efficacy of nilotinib), *BCR-ABL1* levels at that time point were significantly lower for patients with residual nHSCs at diagnosis than in those without detectable nHSCs. Whether this fact has significance for long-term response depth and can predict for successful therapy discontinuation is an interesting subject, which needs longer follow-up.

In vitro data have shown that TKI therapy is not able to eliminate LSCs, which is at least partly due to *BCR-ABL1* independent survival of the quiescent cells.[4,6,9,10,25-27] In concordance with the *in-vitro* data, it has been shown that in almost all patients in sustained complete molecular response after discontinuation of imatinib, the *BCR-ABL1* rearrangement DNA can still be detected in genomic DNA.[28] In our study, in most patients upon nilotinib treatment, the majority of LSCs were rapidly eliminated, and only in a minority of patients relative high levels of LSCs persisted 3 months after treatment initiation. No other previous study has evaluated the disappearance of LSCs during nilotinib treatment, but in a study by Delfina et al the persistence of $CD34^{+}$ progenitor cells was examined during nilotinib treatment.[29] Ph+ cells were only detected in 1 out of 20 patients who had

been treated with nilotinib a longer time (median 22 months), and surprisingly in none of 5 first-line patients who had only been on nilotinib treatment for 3 months.[29] In contrast, in our study cohort, residual Ph+ stem cells were detected in 67% of patients at 3 months. The limited number of $CD34^+$ cells analyzed in the former study (median 100 cells) may at least in part explain the differences between these two studies as the level of Ph+ cells found in our study was very low (median 0.3% based on 1000 cells analyzed). This low level is in accordance with our previous study analyzing first-line imatinib and dasatinib treated patients[15] where we also found that TKI therapy rapidly eliminated the majority of LSCs (median 0.2% at 3 months). Similarly to our previous studies[14,15], significantly less therapy related hematological toxicity was observed in patients who had residual nHSCs or low LSC percentage at the diagnosis. We postulate that this may mirror a larger residual nHSC pool allowing maintenance of normal hematopoiesis even though the leukemic progenitors and LSCs are efficiently eliminated during the initial phase of nilotinib treatment. Overall, none of the previous studies evaluated two independent stem cell quantification methods. We demonstrate a good correlation between the results of MPFC and sorting plus FISH, even though slightly different definitions of the stem cell compartments were used. MPFC has the additional benefit of enabling sorting of viable leukemic and normal stem cells into separate fractions. This permits functional evaluation as well as expression profiling and comparison of LSCs versus nHSCs that may lead to the identification of novel targets on LSCs and ultimately facilitate the development of LSC specific targeted therapy concepts. On the other hand, the cell sorting plus FISH method appears to be more sensitive than MPFC when it comes to the detection of residual disease during therapy. However, the continued development of both methods and other methods utilizing new markers of CML LSCs, could allow them to become valuable tools for assessing patients with CML at diagnosis and during the course of therapy.

In conclusion, we here demonstrate the ability of MPFC and/or sorting plus FISH to predict achievement of molecular response to nilotinib therapy in first-line treated CML patients already at the diagnosis by discrimination of malignant from normal stem cells. This may allow personalized patient management and tailoring of the therapy in the future to select the most potent TKI with respect to LSC elimination, which may be a prerequisite for long-term disease control and also for treatment-free remission upon TKI discontinuation. However, the use of our methods to guide therapy selection should be confirmed in a prospective clinical trial.

REFERENCE LIST

(1) Deininger M, O'Brien S, Guilhot F, goldman J, Hochhaus A, Hughes T, et al. international randomized study of interferon vs STI571 (IRIS) 8-year follow up: sustained survival and low risk for progression or events in patients with newly diagnosed chronic myeloid leukemia in chronic phase (CML-CP) treated with imatinib. Blood 2009;114:Abstract 1126.

(2) Rosti G, Palandri F, Castagnetti F, Breccia M, Levato L, Gugliotta G, et al. Nilotinib for the frontline treatment of Ph(+) chronic myeloid leukemia. Blood 2009;114(24):4933-8.

(3) Cortes JE, Jones D, O'Brien S, Jabbour E, Ravandi F, Koller C, et al. Results of dasatinib therapy in patients with early chronic-phase chronic myeloid leukemia. J Clin Oncol 2010; 28(3):398-404.

(4) Graham SM, Jorgensen HG, Allan E, Pearson C, Alcorn MJ, Richmond L, et al. Primitive, quiescent, Philadelphia-positive stem cells from patients with chronic myeloid leukemia are insensitive to STI571 in vitro. Blood 2002;99(1):319-25.

(5) Copland M, Hamilton A, Elrick LJ, Baird JW, Allan EK, Jordanides N, et al. Dasatinib (BMS-354825) targets an earlier progenitor population than imatinib in primary CML but does not eliminate the quiescent fraction. Blood 2006;107(11):4532-9.

(6) Jorgensen HG, Allan EK, Jordanides NE, Mountford JC, Holyoake TL. Nilotinib exerts equipotent antiproliferative effects to imatinib and does not induce apoptosis in CD34+ CML cells. Blood 2007;109(9):4016-9.

(7) Konig H, Holtz M, Modi H, Manley P, Holyoake TL, Forman SJ, et al. Enhanced BCR-ABL kinase inhibition does not result in increased inhibition of downstream signaling pathways or increased growth suppression in CML progenitors. Leukemia 2008;22(4):748-55.

(8) Konig H, Holyoake TL, Bhatia R. Effective and selective inhibition of chronic myeloid leukemia primitive hematopoietic progenitors by the dual Src/Abl kinase inhibitor SKI-606. Blood 2008;111(4):2329-38.

(9) Corbin AS, Agarwal A, Loriaux M, Cortes J, Deininger MW, Druker BJ. Human chronic myeloid leukemia stem cells are insensitive to imatinib despite inhibition of BCR-ABL activity. J Clin Invest 2011;121(1):396-409.

(10) Chu S, McDonald T, Lin A, Chakraborty S, Huang Q, Snyder DS, et al. Persistence of leukemia stem cells in chronic myelogenous leukemia patients in prolonged remission with imatinib treatment. Blood 2011;118(20):5565-72.

(11) Marin D, Ibrahim AR, Lucas C, Gerrard G, Wang L, Szydlo RM, et al. Assessment of BCR-ABL1 transcript levels at 3 months is the only requirement for predicting outcome for patients with chronic myeloid leukemia treated with tyrosine kinase inhibitors. J Clin Oncol 2012;30(3):232-8.

(12) Sokal JE, Cox EB, Baccarani M, Tura S, Gomez GA, Robertson JE, et al. Prognostic discrimination in "good-risk" chronic granulocytic leukemia. Blood 1984;63(4):789-99.

(13) Hasford J, Pfirrmann M, Hehlmann R, Allan NC, Baccarani M, Kluin-Nelemans JC, et al. A new prognostic score for survival of patients with chronic myeloid leukemia treated with interferon alfa. Writing Committee for the Collaborative CML Prognostic Factors Project Group. J Natl Cancer Inst 1998;90(11):850-8.

(14) Janssen JJ, Deenik W, Smolders KG, van Kuijk BJ, Pouwels W, Kelder A, et al. Residual normal stem cells can be detected in newly diagnosed chronic myeloid leukemia patients by a new flow cytometric approach and predict for optimal response to imatinib. Leukemia 2012;26(5):977-84.

(15) Mustjoki S, Richter J, Barbany G, Ehrencrona H, Fioretos T, Gedde-Dahl T, et al. Impact of malignant stem cell burden on therapy outcome in newly diagnosed chronic myeloid leukemia patients. Leukemia 2013;27(7):1520-6.

(16) Rice KN, Jamieson CH. Molecular pathways to CML stem cells. Int J Hematol 2010;91(5):748-52.

(17) Nicholson E, Holyoake T. The chronic myeloid leukemia stem cell. Clin Lymphoma Myeloma 2009;9 Suppl 4:S376-S381.

(18) Baccarani M, Deininger MW, Rosti G, Hochhaus A, Soverini S, Apperley JF, et al. European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. Blood 2013;122(6):872-84.

(19) Hughes TP, Hochhaus A, Branford S, Muller MC, Kaeda JS, Foroni L, et al. Long-term prognostic significance of early molecular response to imatinib in newly diagnosed chronic myeloid leukemia: an analysis from the International Randomized Study of Interferon and STI571 (IRIS). Blood 2010;116(19):3758-65.

(20) Hanfstein B, Muller MC, Hehlmann R, Erben P, Lauseker M, Fabarius A, et al. Early molecular and cytogenetic response is predictive for long-term progression-free and overall survival in chronic myeloid leukemia (CML). Leukemia 2012;26(9):2096-102.

(21) Marin D, Hedgley C, Clark RE, Apperley J, Foroni L, Milojkovic D, et al. Predictive value of early molecular response in patients with chronic myeloid leukemia treated with first-line dasatinib. Blood 2012;120(2):291-4.

(22) Neelakantan P, Gerrard G, Lucas C, Milojkovic D, May P, Wang L, et al. Combining BCR-ABL1 transcript levels at 3 and 6 months in chronic myeloid leukemia: implications for early intervention strategies. Blood 2013;121(14):2739-42.

(23) Hochhaus A, Hughes TP, Saglio G, Guilhot F, Al-Ali HK, Rosti G, et al. Outcome of patients with chronic myeloid leukemia in chronic phase (CML-CP) based on early molecular response and factors associated with early response: 4-year follow-up data from Enestnd (Evaluating Nilotinib Efficacy and Safety in Clinical Trials Newly Diagnosed Patients). Blood 2012;120:abstract 167.

(24) Jain P, Kantarjian HM, Nazha A, Jabbour E, Quintas-Cardama A, Benjamini O, et al. Early molecular and cytogenetic responses predicts for significantly longer event free survival (EFS) and overall survival (OS) in patients (pts) with newly diagnosed chronic myeloid leukemia (CML) in chronic phase (CP) - an analysis of 4 tyrosine kinase inhibitor (TKI) modalities (standard dose imatinib, high dose imatinib, dasatinib and nilotinib. Blood 2012;120:abstract 70.

(25) Jamieson CH. Chronic myeloid leukemia stem cells. Hematology Am Soc Hematol Educ Program 2008;436-2.

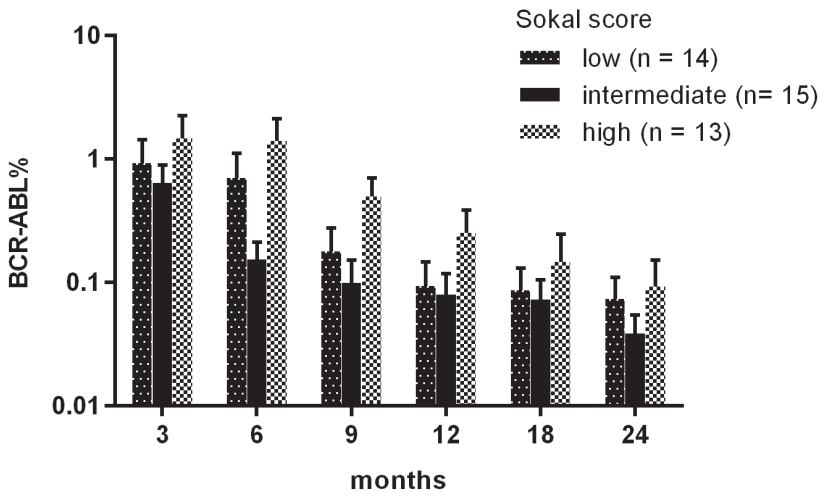
(26) Kavalerchik E, Goff D, Jamieson CH. Chronic myeloid leukemia stem cells. J Clin Oncol 2008;26(17):2911-5.

(27) Hamilton A, Helgason GV, Schemionek M, Zhang B, Myssina S, Allan EK, et al. Chronic myeloid leukemia stem cells are not dependent on Bcr-Abl kinase activity for their survival. Blood 2012;119(6):1501-10.

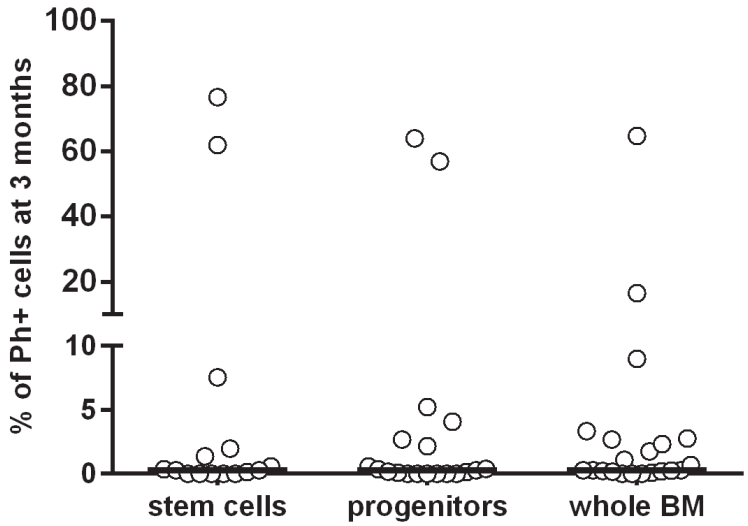
(28) Ross DM, Branford S, Seymour JF, Schwarer AP, Arthur C, Bartley PA, et al. Patients with chronic myeloid leukemia who maintain a complete molecular response after stopping imatinib treatment have evidence of persistent leukemia by DNA PCR. Leukemia 2010;24(10):1719-24.

(29) Defina M, Ippoliti M, Gozzetti A, Abruzzese E, Castagnetti F, Crupi R, et al. Evaluation of residual CD34(+) Ph(+) progenitor cells in chronic myeloid leukemia patients who have complete cytogenetic response during first-line nilotinib therapy. Cancer 2012;118(21):5265-9.

SUPPLEMENTAL MATERIAL



Supplemental Figure 1. Responses at 3, 6, 9, 12, 18 and 24 months according to Sokal risk group. Bars represent mean values and black lines represent standard errors of the mean. The differences between the groups did not reach statistical significance.



Supplemental Figure 2. The proportion of Ph+ cells at diagnosis in the stem and progenitor cell compartment and whole BM as measured by sorting plus FISH. The mean and median percentages of Ph+ cells were 10% and 0.3% in the LSC fractions, 7.2% and 0.3% in the LPC fractions and 5.1% and 0.3% in the whole BM fractions. Black lines represent mean values.