The detection and significance of minimal residual disease in acute and chronic leukemia

N.-G. Chung, V. Buxhofer-Ausch & J. P. Radich

Clinical Research Division, Program in Genetics and Genomics, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

Key words
leukemia; minimal residual disease; relapse

Abstract
Minimal residual disease (MRD) can be detected in many patients with leukemia who have achieved complete remission as defined by conventional pathology examination. The detection of MRD, be it by flow cytometry or by polymerase chain reaction assays, has now been found to be associated with subsequent relapses in most leukemia subtypes, either following chemotherapy or following hematopoietic stem cell transplantation. These assays are now increasingly used in clinical trial design to optimize therapy and provide a novel way to assess treatment efficacy.

Introduction
Relapse remains the major obstacle to cure in leukemia. This problem persists despite ever-intensive chemotherapy regimens, new ‘targeted’ molecular therapy, exhaustive study of mechanisms of drug resistance, and attempts to leverage immunotherapy against the leukemia clone. Unfortunately, once frank relapse occurs, the prospects for achieving a subsequent remission are low, and cure, dismal. One strategy to combat relapse is to detect and treat minimal residual disease (MRD) before overt relapse. Over the past decade, the technology to detect miniscule quantities of leukemia has developed and matured, and there are now ample clinical trials testifying to the power of MRD detection to predict relapse.

The problem of defining remission and relapse is shown in Figure 1. At diagnosis, patients may have a leukemia burden of up to $10^{10}$–$10^{12}$ cells. Thus, even after achieving the ‘gold standard’ of a three-log depletion of leukemia following induction therapy, potentially millions of leukemia cells may persist despite the morphological appearance of remission. The goal of MRD detection is to detect, quantify, and hopefully eliminate persistent and relapsing disease before frank hematological relapse.

Methods of MRD detection
The standard approach to evaluate residual leukemia by routine pathologic examination of bone marrow (BM) is limited by the, sometimes, subtle morphologic differences between malignant and normal cells. Fortunately, several techniques can be used to find the ‘needle in the haystack’, and these are summarized in Table 1.

Chromosomal assays
Conventional metaphase cytogenetics can detect approximately 1 leukemia cell in 100 normal cells (denoted here as ‘$10^{-1}$’ sensitivity) if enough metaphases are analyzed (1, 2). Cytogenetics is the method of choice to define cytogenetic remission or to look for new, unanticipated genomic changes during therapy. However, cytogenetics is limited by sampling only those few cells that divide in culture.
Molecular cytogenetics refers to the analysis of genomic alterations using fluorescence in situ hybridization (FISH)-based technology. FISH uses chromosome-specific or locus-specific probes to target tumor-specific genetic aberrations in metaphase or interphase cells (as opposed to conventional metaphase cytogenetics). FISH can rapidly screen 200–10,000 cells for numerical chromosomal aberrations, even in samples insufficient for conventional cytogenetic analysis, and has the ability to identify minor abnormal clones undetected by karyotypic study. Compared with conventional cytogenetics, the FISH sensitivity ranges from $10^{-1}$ to $10^{-3}$ (depending on the number of probes and number of nuclei scored). The sensitivity of FISH is increased by combining it with other special techniques such as using FISH with metaphase preparations, in combination with morphologic analysis, by using it on specific flow-sorted cell populations or by increasing the number of nuclei scored. The sensitivity of FISH is in-creased by combining it with other special techniques such as using FISH with metaphase preparations, in combination with morphologic analysis, by using it on specific flow-sorted cell populations or by increasing the number of metaphases analyzed by hypermetaphase FISH (3–5).

Flow cytometry assays

Cell surface antigen expression can distinguish malignant from normal cells. While truly tumor-specific antigens are rare, malignant cells often express cell surface antigens in subtly different patterns than normal cells (6). By using combinations of multiple antibodies, ‘multiparametric’ flow cytometric assays use aberrant antigen expression patterns to ‘fingerprint’ the malignant clone (Figure 2). If combinations of several antibodies are used to define the aberrant antigen expression, the sensitivity can reach as low as $10^{-2}$–$10^{-3}$ in experienced hands. This technique has been very useful in acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL), where leukemia blasts often have evidence of abnormal antigen expression compared with normal hematopoietic cells. The current state-of-the-art of three or four color flow cytometry is quickly becoming displaced by techniques that allow seven colors, or more. If this increases sensitivity, flow cytometry may supplant polymerase chain reaction (PCR) (see below) for MRD detection in many diseases.

Polymerase chain reaction

The most sensitive approach to detect MRD involves nucleic acid amplification using the PCR (7). In PCR, a specific genetic lesion target is the ‘fingerprint’ of the malignancy for PCR-driven reactions to have the desired sensitivity and specificity (Figure 3). Gene translocations, such as the t(9;22) in chronic myeloid leukemia (CML) and the t(15;17), t(8;21), and inv(16)/t(16;16) found in ALL, are straightforward leukemia-specific markers for the detection of MRD (8–14). In ALL, the most common translocation involves the Ph+ chromosome, although up to 40–50% of ALL (and AML) have some type of translocation that can be detected by PCR assays (15). In ALL, the most used PCR markers involve the leukemia-specific fingerprints caused by the immunoglobulin heavy chain and T-cell receptor rearrangement (16, 17).

The advantages of PCR are the excellent sensitivity and specificity. Disadvantages include the potential of false-positive results from contamination of the amplified product and false negatives due to RNA degradation.

MRD detection in ALL

MRD detection in chemotherapy

In ALL, MRD detection is accomplished by either flow cytometry or PCR assays. Flow cytometry assays rely on the unique and aberrant constellation of surface antigens in the leukemia clone. PCR-based assays rely on the clonal immunoglobulin heavy chain (IgH V-D-J) or T-cell receptor gene rearrangements that can be used to follow MRD in ~70% of ALL cases. This PCR method takes advantage of the unique gene rearrangements that occur in these genes as cells develop into functional B and T cells. While all individual normal B and T cells should have unique gene rearrangements, a clonal outgrowth of malignant cells should all have identical gene rearrangements, and this

Table 1 Methods to detect minimal residual disease

<table>
<thead>
<tr>
<th>Detection method</th>
<th>Target</th>
<th>Sensitivity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathologic examination</td>
<td>Cellular morphology</td>
<td>5</td>
</tr>
<tr>
<td>Cytogenetics</td>
<td>Chromosome structure</td>
<td>1–5</td>
</tr>
<tr>
<td>FISH</td>
<td>Specific genetic marker(s)</td>
<td>0.08–5</td>
</tr>
<tr>
<td>Flow cytometry</td>
<td>Surface antigen expression</td>
<td>0.1–1</td>
</tr>
<tr>
<td>PCR</td>
<td>DNA or RNA sequence</td>
<td>0.0001–0.1</td>
</tr>
</tbody>
</table>

FISH, fluorescence in situ hybridization; PCR, polymerase chain reaction.
fingerprint can later be used to distinguish malignant from normal cells in ‘remission’ samples. While both methods have excellent sensitivity, both are susceptible to clonal ‘shifts’ that occur with disease progression in 10–20% of cases, whereby the relapsing clone appears with a different gene rearrangement or surface antigen pattern than the presenting leukemia (18–21). Recently, combined use of these two methods should enable MRD monitoring in almost all patients and prevent false-negative results due to clonal evolution and phenotypic shift (22, 23).

The clinical impact of MRD strongly depends on the time point at which it is assessed. Detection of MRD can be used to evaluate early treatment response, thereby allows the identification of low-risk and high-risk patients who may profit from therapy reduction and therapy intensification, respectively. Early clearance of leukemia cells is a favorable prognostic indicator in childhood ALL (16, 24, 25). Low or absent MRD in the BM after the completion of induction therapy predicts a favorable outcome, as MRD-negative patients have an overall relapse rate of only 2–10% (Table 2) (16, 26–29). However, high MRD levels at the end of induction treatment were associated with high relapse rates of 70–100% (16, 26–28, 30); these patients at risk for relapse may benefit from considering alternative treatment approaches such as transplantation (16, 31–33). MRD status after induction therapy may be the most significant prognostic factor in pediatric ALL, independent of other clinically relevant risk factors, such as age, blast count, immunophenotype, presence of chromosome aberrations at diagnosis, and response to prednisone (16, 28, 34).

In pediatric ALL cases who relapse after achieving remission, MRD status after re-induction treatment is similarly important (32, 35). Eckert et al. (32) used a quantitative PCR (QPCR) assay to detect MRD in 30 relapsed patients. The event-free survival (EFS) between two groups (MRD divided into <10⁻³ or ≥10⁻³ levels) was significantly different (86% vs 0%, respectively; \( P < 0.0001 \)). Coustan-Smith et al. (35) studied the clinical significance of MRD in ALL after first relapse using flow cytometry for MRD detection. Of the 35 children with morphological remission after induction treatment, 19 (54%) had MRD \( >0.01\% \), and their subsequent relapse rate was 70% compared with 28% for the 16 MRD-negative patients. The value of continuing MRD monitoring during maintenance therapy and off therapy is less clear since frequent monitoring has been difficult to use (25, 28, 36).

MRD studies in adult patients with ALL have lagged behind their pediatric studies. There is increasing evidence that in adult patients with ALL, as in children, an early decrease in MRD levels after induction therapy is an important predictor of prognosis and response to chemotherapy (37–39). Thus, the assessment of the MRD kinetics...
at different follow-up time points defined risk groups based on a specific threshold MRD level in standard-risk adult patients with ALL (40). Patients with a rapid MRD decline to lower than $10^{-4}$ or under the detection limit at day 11 (during first induction) and day 24 (at the end of induction) were at low risk with a 3-year disease-free survival (DFS) and overall survival (OS) of 100%. In contrast, patients with persistent MRD of $10^{-4}$ until week 16 were at high risk with a 3-year DFS of 5.8% and a 3-year OS of 45.1%.

In the adult Ph+ ALL, MRD kinetics can be followed by measuring bcr-abl levels with quantitative real-time PCR (41, 42). Using this approach, 42 patients with complete hematological response following induction therapy were divided into two prognostic groups: good molecular

### Table 2: Comparison of MRD studies in ALL in childhood at the induction time point

<table>
<thead>
<tr>
<th>Study</th>
<th>EORTC [Cave et al. (16)]</th>
<th>iBFM [van Dongen et al. (28)]</th>
<th>St. Jude [Coustan-Smith et al. (26)]</th>
<th>BFM-Austria [Dworzak et al. (27)]</th>
<th>NOPHO [Nyvold et al. (29)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>178</td>
<td>240</td>
<td>195</td>
<td>108</td>
<td>104</td>
</tr>
<tr>
<td>MRD method</td>
<td>lg/TCR-PCR</td>
<td>lg/TCR-PCR &amp; TAL 1 deletion</td>
<td>Flow cytometry</td>
<td>Flow cytometry</td>
<td>lg/TCR-PCR</td>
</tr>
<tr>
<td>MRD study time point</td>
<td>After induction therapy</td>
<td>5 weeks after diagnosis</td>
<td>End of induction</td>
<td>Day 33 after start of induction therapy</td>
<td>Day 29 after start of induction therapy</td>
</tr>
<tr>
<td>MRD status</td>
<td>Number of patients with relapsing / number of patients with given MRD result (relapse rate, %)</td>
<td>Number of patients with relapsing / number of patients with given MRD result (relapse rate, %)</td>
<td>Number of patients with relapsing / number of patients with given MRD result (relapse rate, %)</td>
<td>Number of patients with relapsing / number of patients with given MRD result (relapse rate, %)</td>
<td>Number of patients with relapsing / number of patients with given MRD result (relapse rate, %)</td>
</tr>
<tr>
<td>Negative $\leq 10^{-4}$</td>
<td>7/88 (8)</td>
<td>2/71 (3)</td>
<td>9/123 (10)</td>
<td>4/59 (7)</td>
<td>1/53 (2)</td>
</tr>
<tr>
<td>$&gt; 10^{-2}$</td>
<td>11/15 (73)</td>
<td>20/27 (67)</td>
<td>7/9 (72)</td>
<td>3/3 (100)</td>
<td>5/15 (33)</td>
</tr>
</tbody>
</table>

lg, immunoglobulin; MRD, minimal residual disease; PCR, Ploymerase chain reaction; TCR, T-cell receptor.

EORTC, European Organization for Research and Treatment of Cancer.

iBFM, International Berlin-Frankfurt-Münster (BFM) study group.

NOPHO, Nordic Society for Pediatric Hematology and Oncology.
**Table 3** A comparison of studies examining the relation between MRD levels before HSCT and subsequent relapse

<table>
<thead>
<tr>
<th>Study</th>
<th>Number of Patients</th>
<th>Negative</th>
<th>Low-level positive</th>
<th>High-level positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knechtli et al. (47)</td>
<td>64</td>
<td>8/38</td>
<td>5/9</td>
<td>12/12</td>
</tr>
<tr>
<td>Uzunel et al. (48)</td>
<td>30</td>
<td>0/5</td>
<td>5/10</td>
<td>8/15</td>
</tr>
<tr>
<td>Bader et al. (46)</td>
<td>41</td>
<td>2/14</td>
<td>5/10</td>
<td>11/17</td>
</tr>
<tr>
<td>Krejci et al. (45)</td>
<td>140</td>
<td>12/77</td>
<td>14/27</td>
<td>27/36</td>
</tr>
<tr>
<td>van der Velden et al. (50)</td>
<td>17</td>
<td>2/11</td>
<td>4/6</td>
<td></td>
</tr>
<tr>
<td>Scanchez et al. (51)</td>
<td>24</td>
<td>3/18</td>
<td>4/6</td>
<td></td>
</tr>
<tr>
<td>Sramkova et al. (49)</td>
<td>25</td>
<td>1/17</td>
<td>6/8</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>341</td>
<td>28/180 (15.5%)</td>
<td>29/56 (51.7%)</td>
<td>58/80 (72.5%)</td>
</tr>
</tbody>
</table>

MRD, minimal residual disease.

responders (with >2-log reduction of *bcr–abl* transcript levels after induction and a >3-log reduction after consolidation therapy) and poor molecular responders, who had a higher MRD level at both time points (42). The probability of 2-year OS was 48% for the good molecular responders compared with 0% (*P = 0.0026*) for the poor molecular responders.

Krampera et al. showed that frequent immunophenotypic MRD measurement in the first year of treatment is a useful outcome predictor for adult patients with T-cell-ALL (43). MRD-positive patients before consolidation therapy had a probability of relapse at 2 years of 82% compared with 39% (*P = 0.00078*) in MRD-negative patients. In a larger study on 102 adolescent and adult patients with ALL, MRD at day 35 was the most powerful independent prognostic parameter by multivariate analysis (44). Patients in morphologic complete remission (CR) and with a low MRD level (<0.05%) had a median relapse-free survival (RFS) of 42 months vs 16 months (*P = 0.001*) in case of higher MRD levels. An excellent prognosis with a projected 5-year RFS of 90% was observed when MRD levels at day 14 were very low (<0.03%).

**MRD in ALL after transplantation**

Allogeneic hematopoietic stem cell transplantation (HSCT) is generally reserved for the treatment of high-risk patients in first remission and those who have relapsed after chemotherapy. As the risks of transplant are substantial, it is important to distinguish the patients who will benefit from HSCT from those who will profit from different therapeutic approach or further modification before or after HSCT. MRD may be a method to accomplish this clinically valuable assessment.

MRD detection before and following transplantation both are strongly predictive of outcome. Many reports have shown that MRD burden prior to conditioning is the most powerful single predictive factor for relapse after HSCT [Table 3; (45–51)]. In these studies, the EFS for the MRD-negative groups before HSCT was often more than 70%, while that for MRD-positive group fell to 17–50%. Both PCR and flow cytometry were used to detect MRD in these studies, and various criteria have been used to determine ‘low’ from ‘high’ levels of MRD burden. Taken together, however, the data are fairly consistent and logical, showing a very high risk of relapse (~75%) in cases with high levels of MRD before HSCT, intermediate risk of relapse in cases with a low-level MRD (~50%), and a much smaller risk of relapse in those without MRD (~15%).

MRD detection following transplantation is also predictive of subsequent relapse. This is true for Ph+ ALL and non-Ph+ ALL cases. In the later setting, studies have found the relative risk (RR) of relapse associated with MRD to be roughly fivefold to tenfold, with relapse occurring within a few months of MRD detection (52, 53). The report on Ph+ ALL is slightly more complicated. In Ph+ ALL, PCR must be performed on both the *p210 bcr-abl* variant (found in CML and some ALL) and the *p190 bcr-abl* (found only rarely in CML but common in ALL). All studies of Ph+ ALL have shown excess relapses in MRD-positive cases following transplantation (54–56). However, the data suggest that the risk of relapse associated with *p190 bcr-abl* MRD is stronger than that with *p210 bcr-abl*. Thus, in 90 patients, the RR of relapse with any MRD posttransplant assay (compared with patients with all negative MRD assays) was 4.4 (56). For *p190*-positive MRD, the RR was 8.7, while for *p210 bcr-abl*, the RR was 2.2. Note that this difference in risk with different *bcr-abl* variants was also seen in the MRD assessment before transplantation. The median time from first MRD detection to relapse was 75 days, ample time to intervene with such agents as imatinib, interferon (IFN), or donor leukocytes.

Most MRD studies have used BM samples rather than peripheral blood (PB). In precursor B cell and Ph+ ALL,
the level of MRD is higher in BM than in PB, which makes
BM the preferred source of disease monitoring (57). However, the use of PB to detect MRD has several practical advantages over the use of marrow, especially in children, as PB sampling is simple and relatively painless, allowing more frequent monitoring, which may offset the sensitivity advantage of BM (58, 59).

**MRD detection in AML**

**MRD detection in chemotherapy**

In AML, the use of PCR-based detection of MRD is essentially limited to patients with AML containing unique breakpoint fusion regions, such as t(15;17), t(8;21), and inv(16). However, flow cytometry assays can be used to detect abnormal patterns of lineage and differentiation-specific markers suggestive of leukemia blast in the majority of AML cases.

Acute promyelocytic leukemia (APL) frequently harbors the t(15;17), which causes the PML/RARα gene juxtaposition. Patients with APL who have residual disease after induction or consolidation therapy have a significant risk of relapse (60, 61). Serial quantitative MRD monitoring by PCR until completion of maintenance therapy might allow identifying patients at high risk of relapse in APL (62). Thus, the study of 70 patients with APL identified the MRD level after first consolidation as the most powerful predictor of relapse. Patients with a MRD level of >10⁻³ had a >10-fold higher relapse rate at 5 years compared with a level of <10⁻³ (P = 0.001). In APL, there is evidence that the analysis of PB maybe as informative as that of BM (63).

Disruptions of the core-binding factor (CBF) genes AML1 and CBFB are involved in the t(8;21) and inv (16) ‘good’ risk translocations, found collectively in 15–20% of adult and pediatric cases with AML. The monitoring of the core AML1–CBFB chimeric messenger RNA (mRNA) poses an interesting problem. The qualitative detection (yes or no) of the AML1–MGT8 and CBFB–MYH11 transcript has a limited value in monitoring MRD and predicting relapse since these transcripts persist during long-term remission (64–72). Therefore, quantification of MRD is essential in these AML variants (11, 67). Several groups have attempted to define the critical level of MRD that accurately segregates good and poor outcome. In general, a posttherapy MRD level of <1% has been associated with a relatively low risk of relapse (73–77). MRD investigation at the end of treatment might have the best prognostic value in CBF leukemia (78). Schnittger et al. created a score for risk stratification of AML by combining the transcription ratios at diagnosis and after 3–4 month during treatment (79).

Immunophenotypic detection of MRD by flow cytometry can be used for identifying risk groups after induction chemotherapy and for providing important information to postinduction treatment strategies. In a study of 72 patients, repeat flow cytometry studies enabled the identification of specific cutoff values of MRD associated with significant survival impact after the induction and consolidation therapy (80). A cutoff level of 1% after the first induction, 0.14% after the second induction and of 0.11% after the consolidation resulted in a 6.1-, 3.4- and 7.2-fold higher RR of relapse, respectively. Similar conclusions can be drawn from other studies (81, 82). Cousan-Smith et al. (83) found MRD in 85% of children, with a sensitivity of 10⁻³ or more by using flow cytometric detection of abnormal immunophenotypic combinations. Residual disease was found in 34.1% and 27.8% of patients after first and second induction, respectively, by flow cytometry, with a 2-year OS was approximately 30% compared with >70% in patients without detectable MRD. Sievers et al. (84) found immunophenotypic abnormalities in 41 (16%) of 252 children with AML who achieved morphologic CR with initial induction therapy, and MRD was highly associated with a poor outcome.

The Wilm’s tumor gene (WT1) may be a molecular target for MRD detection, as it is overexpressed in around 80% of cases with AML (85, 86). The ability to use this marker is complicated by the fact that it is also expressed in normal hematopoietic cells, although at a lesser degree; thus, the tricky issue is distinguishing low WT1 in leukemia cells from background from normal cells. Thus, the most consistent use of WT1 is as a prognostic tool at diagnosis. Several groups have reported that WT1 levels at diagnosis are associated with prognosis (low, good; high, bad) (87). In some studies, the increase in WT1 during therapy is associated with subsequent relapse (85) and correlated to other disease markers, such as leukemia-specific fusion transcripts (86). Other groups, however, have found WT1 levels to be important at distinct intervals of treatment (88).

The most common genetic mutations found in AML involve FLT3 and NPM, both occur in ~25% or greater of cases. Theoretically, these may be targets of PCR-based assays but not easily. In FLT3, the most common alteration is an internal tandem duplication that causes a 5–100 base pair addition to the juxtamembrane domain. Primers can be made that allow for leukemia-specific amplification for the specific patient (89–91). However, the FLT3 mutation is not terribly stable, and cases relapse without the FLT3 mutation ~10–20% of the time (92, 93). Mutation-specific PCR reactions have been constructed for the most common point mutations in NPM1, many potential mutations exist (94). It is unlikely that PCR mutation detection in these subsets of AML will replace the faster and near-equally sensitive flow cytometry.

**MRD in AML after transplantation**

There are few studies documenting the prevalence and significance of MRD in AML following HSCT. This in part
is due to the fact that the most common translocations found in AML [the t(15;17), t(8;21), and inv. (16)] are associated with intermediate or good outcomes with conventional chemotherapy and thus are underrepresented in the transplant population.

The significance of MRD following transplantation depends largely on the specific genetic subtype of AML and generally correlates to the ‘rules’ seen in the chemotherapy setting. Thus, in t(15;17) AML, the detection of the PML/RARα transcript following HSCT is highly correlated to relapse (95, 96). In t(8;21), the detection of AML1/ETO remarkably often persists after allogeneic transplantation, without subsequent relapse (66–68). There are too few cases of inv (16) studied in the transplant setting to make a firm conclusion about its association with relapse (65, 97).

The tumor suppressor gene WT1 has been used as a marker for MRD following HSCT (98). The results have been mixed. Elmaagachi et al. studied 38 patients after HSCT, and 14/38 were positive for elevated WT1 at least once after transplantation. Of the 14 MRD-positive patients, 7 relapsed compared with five relapses in the 34 patients without MRD. Elevated WT1 was in concordance with the persistence of translocation markers in 70% of cases. However, other studies (99) have not found WT1 levels predictive of relapse following HSCT. Another possibility is using WT1 levels before transplantation as a measure of disease burden to predict subsequent response. Osborne et al. measured WT1 transcript levels in BM harvests and correlated it to the outcome after autologous transplantation. The median RFS for patients with high WT1 levels was 10.5 months, while patients with low levels have not reached the median RFS at a median follow-up of 94 months (100).

WT1 is being investigated as a potential target for immunotherapy, and WT1 levels thus may suggest particular leukemia more susceptible to this therapeutic approach. Ogawa et al. (101) studied 72 patients after HSCT (50 AML, 15 ALL, and 7 CML), and all 20 patients who relapsed showed elevated WT1 level at the time of relapse. After relapse, the patients were treated with some form of immunomodulation (DLI or discontinuation of immunosuppressive therapy). The calculated WT1 doubling time that occurred with relapse for patients was significantly longer for patients who responded to immunomodulation (a median doubling time of 26 days) than for patients who did not (a median of 6 days). It is unclear if the same information in regard to the kinetics could be also produced by serial flow cytometric studies, however.

Last, the measurement of the MRD burden of transplantation products may also be used to guide patient-tailored purging procedures (102). An MRD level of less than 0.05% in the stem cell products, as measured by flow cytometry, was found to be associated with an excellent RFS of 100% at 12 months. In an autologous transplant setting the MRD status before transplantation, measured by multidimensional flow cytometry, predicted reliably outcome (103). Patients (n = 19) with residual leukemia cells <3.5 × 10−4, were assigned to low-risk MRD group, and those (n = 12) with ≥3.5 × 10−4 cells to high-risk MRD group. One hundred percent of patients with high-risk disease relapsed vs 26% of those with low-risk disease relapsed (P = 0.0004). The median RFS after transplantation was 7 and 48 months for the high-risk and low-risk MRD group, respectively (P = 0.007). Increasing MRD levels were shown in three of the five patients with relapsed disease by longitudinal flow cytometry monitoring. This report suggested that MRD can be used for predicting outcome in patients receiving autologous transplantation and impending relapse can be detected by sequential MRD monitoring.

**MRD detection in CML**

The power and potential of MRD detection to predict relapse and guide therapy have been led by the studies on CML. The clinical utility of MRD detection of the chimeric bcr–abl transcript was first shown in the HSCT setting and has lately been used in studies of the tyrosine kinase inhibitors, especially imatinib.

**MRD following imatinib therapy**

In CML, QPCR for the bcr–abl defines a population of patients in cytogenetic remission who have a prolonged course of remission. Quantitative detection of patients with CML treated with tyrosine kinase inhibitors may help in guiding therapy in regard to the depth of disease burden undetectable by cytogenetic testing.

The landmark International Randomized study of Interferon and ST1571 (IRIS) trial compared imatinib mesylate (IM) with the ‘standard’ non-transplant therapy of IFN and Ara-C in newly diagnosed patients with chronic phase (CP) CML. At the time of study publication, >70% of patients with IM achieved a complete cytogenetic remission (CCR) compared with only 7% in the IFN and Ara-C arm (104, 105). bcr–abl detection by QPCR was used to examine the patients who achieved a CCR (106). Patients in CCR after imatinib therapy had a statistically significantly greater reduction in bcr–abl level than those on the IFN/Ara-C arm. A three-log reduction in bcr–abl copy number from baseline (a so-called ‘major molecular response’ or MMR) was obtained in 39% of imatinib-treated patients after 12 months of therapy compared with only 2% in the IFN/Ara-C arm. The depth of molecular response at 12 months was associated with progression-free survival. Thus, if patients did not achieve a CCR, the risk for progression in patients without a CCR was ~25% compared with those who achieved a CCR and had a <3-log reduction (~10%) or a ≥3-log reduction in bcr–abl by 12 months (~0%). Undetectable
bcr–abl (defined as quantitatively undetectable bcr–abl as well as a PCR-negative test by a more sensitive nested reverse transcription polymerase chain reaction (RT-PCR) that was confirmed by another laboratory in the study) was unusual (<5% of cases at the 12 month mark).

An important and often misunderstood feature of the IRIS molecular study was the use of log reduction as a measure of bcr–abl response, as opposed to actual bcr–abl copy number, or the ratio of bcr–abl to a control gene. The analysis was not done using each patient’s baseline bcr–abl value. Rather, each laboratory established its own baseline on the same set of 30 baseline samples. Subsequent individual samples had bcr–abl/bcr values calculated, and these values compared with that laboratories baseline value yielded the log reduction for that sample. This method had two major advantages. In regard to the IRIS study, it standardized the assay across the three laboratories and allowed other laboratories to compare results, as long as each laboratory had a baseline value since the log reduction did not require the use of the same control gene from laboratory to laboratory. The recent expert consensus on bcr–abl advocates the use of an International Scale based on the log reduction principal (107). Notably, all bcr–abl molecular testing in the IRIS trial was performed on PB samples. Similar to posttransplant MRD data, there is excellent correlation between RT-PCR bcr–abl levels between BM and PB (105, 108, 109). Thus, if a patient is in a CCR, molecular testing for bcr–abl can be confidently performed on PB.

Subsequent studies have validated the clinical utility of bcr–abl testing after IM therapy (110, 111). The Hammersmith group studied a more heterogeneous group of patients with CP CML who achieved a CCR on imatinib (111). Thirty-two of 106 patients studied had a persistent decline in bcr–abl and, predictably, none had relapsed disease. Twenty-six patients showed an increase in bcr–abl levels after initial response, with subsequent relapse. However, 42 patients achieved a ‘plateau’ bcr–abl level without quite reaching an MMR and remained in CCR. This study is in contrast with a study of 280 cases with CML from MD Anderson, which reported a 5% loss of CCR in MMR cases compared with a 37% loss of CCR in cases that did not achieve a MMR (110). However, in both studies compared with the IRIS study, cases with CML in these studies included patients at various times from diagnosis and also included patients who had previously received IFN therapy.

The rate of bcr–abl mRNA decline after IM therapy predicts long-term response, as bcr–abl decline after the initial 2–3 months of IM predicts of subsequent response (110, 112–114). Thus, patients who do not achieve at least a one-log reduction in bcr–abl after 3 months of therapy have a lower probability of achieving a CCR and an MMR or a higher probability of developing resistance (110, 112). Thus, patients with CML treated with IM who achieve a <1-log reduction in bcr–abl after 3 months have a 13% chance of ever achieving an MMR compared with >70% in patients with a better response (112). Similarly, a separate study found that only 80% of patients with <1-log bcr–abl response at 3 months achieved a CCR compared with 100% with patients who achieved a greater reduction in bcr–abl (110). Thus, frequent and early monitoring may help define patients who need early consideration for alternative therapies such as transplantation or other experimental protocols.

**MRD detection after transplantation**

A large number of studies have studied the interaction of MRD and relapse in CML after transplantation. A number of robust observations have been made, which illustrate the complexity of interpreting MRD results. First, at least in CML, the molecular detection and quantification of bcr–abl may be performed on either PB or BM specimens (52). While QPCR is routine today, it should be noted that in CML, the qualitative detection (yes or no) of bcr–abl after HSCT is associated with relapse and varies with the type of transplant and the time from transplantation (52, 115–120).

The highest risk of relapse associated with bcr–abl detection appears to be associated ‘early’ (<12 months after HSCT) after transplantation (115, 119, 121). In one study of 346 patients after transplantation, a positive bcr–abl assay 6–12 months after transplantation was associated with a 42% risk of relapse as opposed to a 3% risk of relapse in PCR-negative patients. Tests earlier than 3 months after HSCT were not strongly associated with relapse. Moreover, MRD can be detected by QPCR for years after HSCT. Bcr–abl has been detected in 25–50% of patients ≥3 years after transplantation, with subsequent relapse rates of ~10–20% (118, 121–123). Patients in CR for >10 years after HSCT have been found to be persistently MRD positive (124).

The type of transplant (allogeneic, T-cell depleted, or unrelated) affects the association of MRD and relapse. This is presumably due to the immunologic effect associated with an allograft (‘graft vs leukemia’, or GVL). Transplants from unmanipulated unrelated donor grafts had a lower risk of relapse compared with matched related transplants. Patients with T-cell depletion have a very strong association of MRD detection and subsequent relapse (118, 125). Thus, while T-cell depletion decreases graft-vs-host disease, it clearly does so at the cost of the associated GVL effect.

The quantification of bcr–abl by QPCR further refines the predictive value of MRD detection in CML (123, 126–131). Thus, low residual bcr–abl after transplantation was associated with a very low risk of relapse (1%) compared with 75% relapse rate in patients with increasing or persistently high bcr–abl levels (132). Olavarria et al. studied 138 patients with CML early (3–5 months) after
transplantation and found that bcr-abl level was highly correlated to relapse (130). Patients with undetectable bcr-abl had a 9% risk of relapse compared with those with a ‘low’ (<100 bcr-abl transcripts/μg RNA) or ‘high’ level of transcripts (>100 copies bcr-abl/μg), who had a cumulative relapse rate of 30 and 74%, respectively. In addition, in a study of 379 patients with CML ‘late’ (>18 months) after HSCT (123), 90 patients (24%) had at least one positive disease (hazard ratio of relapse = 19). An increasing bcr-abl reliably portended relapse.

**MRD in chronic lymphocytic leukemia**

Other B-lineage hematologic malignancies [chronic lymphocytic leukemia (CLL), multiple myeloma, and NHL] can be studied by similar strategies used to study MRD in ALL. Thus, IgH V-D-J rearrangements can be used in most lymphoid malignancies to detect MRD. Moreover, CLL has a phenotype that is detectable by flow cytometry with excellent sensitivity. In head-to-head comparisons of PCR and flow cytometry, PCR has been found to be slightly more sensitive, with approximately 10% of flow cytometry negative cases being found to be PCR positive for clonal rearrangements. Obviously, however, this increased sensitivity must be weighed against the relative ease and speed of flow cytometry compared with the more complicated and costly PCR technique.

The clinical significance of MRD in CLL appears to be tied to the treatment. This appears to be because of the sensitivity of CLL to the allograft ‘GVL’ effect. Thus, in cases with CLL treated with either anti-CD52 antibody-based therapy (CAMPATH) or autologous transplantation, the presence of MRD (by either flow cytometry or PCR) is very strongly associated with subsequent relapse. Rawstron et al. found a strong correlation of quantitative level of CLL in PB and BM (133). In 25 cases who achieved a CR after CAMPATH-1H antibody or autologous transplant, 19 without MRD at the time of CR, and their EFS was >90%; in contrast, all six patients with MRD at the time of CR subsequently relapsed. Similar findings were found by Esteve et al. who found that four of five patients who had MRD while in CR following autologous transplantation eventually relapsed compared to two of nine without MRD (134).

By contrast, MRD negativity allogeneic transplantation is strongly associated with a very low relapse risk, but many patients who are MRD positive do not relapse but remain low-level MRD positive. Of the 12 patients who had an allogeneic transplantation, only one of eight cases had MRD at last contact. Some cases took up to 4 years to clear MRD and stayed MRD free for another year of follow-up. This long-term clearance of MRD has also been seen in other studies of MRD following allogeneic transplantation and after nonmyeloablative ‘mini’ transplants, where some patients who became MRD negative 6 months after transplant of an allogeneic-related immune system (134, 135). This is reminiscent of the situation with CML, as noted above. It remains to be seen if the allogeneic effect has a time limit, that is when the graft finally becomes tolerant of the host, if the graft will also become tolerant of the residual disease.

**Dormancy**

Not all patients with detectable MRD progress to relapse. As noted above, this has been observed in t(8;21) AML following conventional chemotherapy and in CML following HSCT (11, 136). In t(8;21) AML, the AML1/ETO mRNA can be detected in a minority of lymphoid and erythroid colonies taken from remission samples, suggesting that the disease involves a primitive leukemia progenitor cell (137). As such, some MRD in this disease may arise from nonleukemia lymphoid cells. Furthermore, another example of dormancy is the observation that late relapses in pediatric ALL (who relapse >10 years after CR) have the same clonal IgH V-D-J gene rearrangement as at diagnosis (138). How can such examples of dormancy be explained? Perhaps prolonged immune surveillance, most relevant after HSCT, can control the expansion of a leukemia clone. As in the example of t(8;21) noted above, the MRD assay may detect signal in a cell lineage not involved in the leukemia. Last, a residual leukemia cell may be ‘preleukemia’ not having all the mutations necessary for malignancy or ‘postleukemia’ whereby additional mutations have rendered the clone relatively inert. Regardless, it is likely that the clinical definition of ‘cure’ does not mean the elimination of all cells involved in the leukemia process. The further study of dormancy may clarify what it takes to cure leukemia.

**Conclusion**

The detection of MRD is strongly associated with relapse. MRD monitoring is being melded into the design of clinical trials and is likely to improve the speed and efficacy of these trials. Many research avenues bear further investigation. The treatment of MRD in almost all settings is still in investigation and cries out for large-scale studies. The recognition of ‘dormancy’ may allow us to adopt therapies to achieve that state, rather than persist in high-dose therapy designed to eliminate leukemia, which may be impossible. In many leukemia subtypes, there are few good MRD markers. Further research (for example gene expression and proteomic studies) may identify new markers for MRD assays. In the future, the detection of MRD will play a pivotal role in patient management. The support of clinical studies to clarify the role of MRD detection is essential.
Significance of MRD in acute and chronic leukemia

N.-G. Chung et al.

References

11. Kusec R, Laczika K, Knobl P et al. AML1/ETO fusion mRNA can be detected in bone blood samples of all patients with t(8;21) acute myeloid leukaemia after chemotherapy or autologous bone marrow transplantation. Leukemia 1994: 8: 735–9.


Significance of MRD in acute and chronic leukemia

N.-G. Chung et al.


57. van der Velden VH, Jacobs DC, Wijkhuis AJ et al. Minimal residual disease levels in bone marrow and peripheral blood are comparable in children with T cell acute lymphoblastic leukemia (ALL), but not in precursor-B-ALL. *Leukemia* 2002: 16: 1432–6.


98. Elmaagachi AH, Beelen DW, Trenschel R, Schaefer UW. The detection of wt-1 transcripts is not associated with a increased leukemic relapse rate in patients with acute leukemia after allogeneic bone marrow or peripheral blood stem cell transplantation. *Bone Marrow Transplant* 2000: **25**: 91–6.


100. Osborne D, Frost L, Tobal K, Liu Yin JA. Elevated levels of WT1 transcripts in bone marrow harvests are associated with a high relapse risk in patients autografted for acute myeloid leukemia. *Bone Marrow Transplant* 2005: **36**: 67–70.


133. Rawstron AC, Kennedy B, Evans PA et al. Quantitation of minimal disease levels in chronic lymphocytic leukemia using a sensitive flow cytometric assay improves the prediction of outcome and can be used to optimize therapy. *Blood* 2001: **98**: 29–35.


