Quantitative analysis of minimal residual disease by PCR in childhood acute lymphoid leukemia

de Haas, V.

Citation for published version (APA):
Chapter 8

SUMMARY, GENERAL DISCUSSION AND FUTURE PERSPECTIVES
PCR-based detection of minimal residual disease (MRD) in children with acute lymphoblastic leukemia (ALL) at different time-points during follow-up provides information on the effectiveness of treatment. Our aim was to study whether patients with high or low risk for relapse can be recognized early. The ultimate goal is to stratify treatment, using PCR-based MRD detection. Therefore, we focused on four questions: Is it possible to quantify MRD in children with ALL and can this add to therapy-stratification and outcome of patients? What is the preferable technical approach? Which time-point(s) is (are) the most informative? And finally, how can MRD detection be implemented into routine patient care? The work presented in this thesis provides support for clinical application of PCR-based MRD. Secondly, information on the early changes of leukemic cells and the responsiveness on treatment is provided.

**Summary of results of PCR-based MRD-detection in ALL patients**

The junctional regions of rearranged immunoglobulin heavy chains (IGH) and T-cell receptor (TCR) genes can be used as patient-specific PCR targets for the detection of MRD in childhood ALL, but the major drawback of these targets is their instability. It has been demonstrated that it is very important to select for stable monoclonal targets in case of oligoclonality in patients with B-precursor ALL. Studies from Steenbergen\(^47\) and Green\(^59\) have already focused on the mechanisms of oligoclonality and the questions whether oligoclonality can hamper MRD detection later in the disease. Ongoing rearrangements of immunoglobulin (Ig) and TCR genes can result in subclone formation. The ALL cells at relapse can contain antigen receptor rearrangement(s) differing from the rearrangements found in the major clone(s) at diagnosis. We investigated whether oligoclonality also hampers MRD detection in the beginning of the disease. By determining precisely the relative frequency of the various oligoclonal Ig gene rearrangements at diagnosis and at the end of induction therapy, we were able to calculate the reduction factor of each Ig rearrangement in response to induction therapy. We noted that the various subclones at diagnosis were strongly differing in their response to therapy (chapter 2). Moreover, in the two relapsing patients, the clones detected at relapse, showed the slowest regression during initial treatment or even became detectable for the first time during induction therapy.

Also in seemingly monoclonal patients, subclones that are less responsive to therapy may be present at diagnosis. It is possible that relapsing patients, who are monoclonal at diagnosis, have resistant subclones already at initial diagnosis in which ongoing Ig and/or
TCR rearrangements have occurred. These findings have major implications for identification of targets at diagnosis and detection of MRD in the first weeks of treatment because different PCR targets must be applied in each patient to minimize the risk that these small subclones are missed during PCR follow-up.

Patients at risk of relapse are clinically identified as patients with high white blood cell count (> 50.10^9 cells/l), specific karyotype [(t(4;11); t(9;22)), age (<1 and > 10 years), mediastinal enlargement and immunophenotype (T-lineage ALL). Several clinical studies have shown that chromosomal aberrations in ALL can be used for risk-group classification. For instance, the t(9;22) and t(4;11) aberrations are associated with poor prognosis whereas the t(12;21) is associated with good prognosis. The t(12;21)(p13;q23) translocation involves the fusion between the TEL and AML1 gene, resulting in a TEL-AML1 fusion gene transcript. We focused on the TEL-AML1 fusion gene as this translocation is present in 25% of children with B-precursor ALL. Initially, reports even stated that patients with this translocation had such a good outcome that it might be questioned whether MRD detection was needed as an additional prognostic factor in this subgroup of ALL patients. Nevertheless, the presence of TEL-AML1 in several relapsing patients, urged to re-evaluate this aspect. Studying the reduction of MRD by PCR analysis in relapsed TEL-AML1-positive patients, our results show that MRD level at the end of induction has a higher impact as compared to the presence of TEL-AML1 (chapter 3). Therefore, analysis of MRD is still important in these patients. In this context, RT-PCR analysis of TEL-AML1 fusion transcript might be a simpler and less laborious alternative approach for the genomic RQ-PCR on antigen receptor gene rearrangements. However, it was unknown how stable the mRNA is and whether the number of transcripts per leukemic cell remains constant during follow-up. Hence, by comparison of TEL-AML1 RQ-PCR versus Ig/TCR RQ-PCR, it was shown that both targets have the same level of MRD in the TEL-AML1-positive ALL patients (chapter 4). Consequently, one can assume that within a patient no major variation occurs in the number of fusion gene transcripts per leukemic cell at different disease stages. This finding suggests that the Ig/TCR RQ-PCR can be replaced by the TEL-AML1 RQ-PCR. This will have major consequences for MRD analysis by PCR, as this target is available for a large number (25%) of patients and is easier to use in MRD analysis than antigen receptor gene PCR.

Our study on kinetics of reduction of MRD by PCR already at two weeks after start of induction therapy demonstrates that this early time-point, in addition to results at later time-points, has important consequences for outcome and may be highly informative. We were able to perform an accurate quantification by RQ-PCR and our results indicate that it might even be possible to make a clinical decision upon this single time-point at day 15 (chapter 5). With this information, it will be possible to start modification, either
intensification or decreasing of therapy, early during the course of the disease. Therefore, it has to be investigated in prospective studies whether indeed a clear cut-off level will be found.

The clinical response of leukemic cells to chemotherapy can be predicted in vivo by measuring kinetics of MRD and in vitro by drug sensitivity assays. Both the PCR technique and the MTT assay have individually shown their significant value to predict outcome and might be used to modify treatment. To see whether both techniques identify the same group of patients and measure the same relative therapy resistance, we analyzed 20 patients using the MTT assay as well as PCR-based MRD detection (Chapter 6). There was no significant correlation between the MTT-assay and the level of MRD at the end of induction therapy. However, for prednisolone resistance, a trend was seen for correlation between the MTT assay and MRD detection. This is in concordance with previous studies, where a poor in vivo clinical response to initial prednisolone monotherapy was shown to be associated with poor prognosis. Therefore, we hypothesize that this correlation reflects the strong influence of prednisolone on the initial treatment response, which can be used to determine the risk for relapse. Because both the MTT-assay and MRD detection have been shown to predict relapse-risk in patients and because our results suggest that both assays might recognize different groups of patients, we expect that a combination of both tests may give a better correlation with disease outcome. For example, a patient with intermediate MRD level and a resistant pattern in the MTT test, might be stratified as high risk. Vice versa, a patient with intermediate MRD level and a sensitive pattern in the MTT test, might be stratified as low risk. As our study was only a technical comparison of both assays and not a study of the prognostic significance of both tests, we can not make any conclusions on the correlation between both tests and outcome. A larger prospective study, in which both techniques are performed and evaluated with regard to outcome, is warranted to analyze the profit of the combination of both MRD and MTT.

Finally, we studied the occurrence and early diagnosis of a central nervous system (CNS) relapse in ALL patients. Diagnostics of CNS relapse are normally performed by cytomorphology and flow cytometry. Few studies reported about detection of MRD in the cerebrospinal fluid (CSF) by PCR. Application of PCR on CSF samples from patients with or without CNS involvement, did show that PCR is a specific and sensitive method for early detection of leukemic cells in the CSF as compared with the conventional detection techniques (chapter 7). As data were obtained from a small group, this study may serve as a basis for a more extended study to test if it is worthwhile to perform PCR-based MRD analysis in CSF in a routine setting. Considering the advantages of routinely used RQ-PCR for MRD detection in the bone marrow, it seems feasible to implement PCR-based MRD detection of CSF for diagnostic purposes. Nevertheless, it should also be
General Discussion

considered whether these results will have clinical relevance and earlier start of therapy will indeed lead to better outcome.

Pathogenesis and biology of childhood ALL

It is still not clear what the etiology of leukemia is. Key issues are the time frame of ALL development and whether the consistent molecular abnormalities observed, are the product of primary or secondary etiologic events. The cause of childhood acute lymphoblastic leukemia lies probably in a sequential combination of (prenatal and postnatal) events. Molecular insights into childhood leukemia have provided an opportunity to investigate the biology of the disease more clearly. In particular, the consistent stable chromosomal abnormalities and unique antigen receptor gene rearrangements provide clonotypic markers for tracking the leukemic cells.

Given the young age of most children who have ALL and the latency expected for clonal evolution of cancer, it was postulated that the disease originates - at least in some patients - in utero. This was first demonstrated by sequence analysis of junctional regions of antigen receptor gene rearrangements. N-nucleotides are frequently missing at the D-J\(_H\) junctional region in childhood B-precursor ALL with disease onset below the age of three years.\(^{125}\) The enzyme responsible for the addition of N-nucleotides, terminal deoxynucleotidyl transferase (TdT), is developmentally regulated and in humans not expressed until the second trimester of fetal development. Therefore, it is suggested that presence of rearrangements with absent N regions in some patients, are suspect for leukemia generated from early fetal life.\(^{126;127}\) Secondly, the group of Greaves\(^{85;128}\) showed the presence of the TEL-AML1 fusion gene transcript in the blood-spot taken within five days after birth, from patients who later on developed leukemia. This is direct evidence that this disease can originate in utero. Thirdly, their molecular studies of identical twins with concordant ALL suggest a prenatal origin for acute lymphoblastic leukemia. Molecular analysis of genotypic fusion sequences by PCR in two twin-pairs with either the MLL-AF4- or the TEL-AML1 fusion transcript, revealed in each pair of twins the same unique clonotypic breakpoint. Most of these studies focus on the presence of gene fusion transcripts and in a minority of cases rearrangements of the Ig or TCR gene were studied. A recent Japanese study\(^ {129}\) described the presence of leukemic Ig and TCR rearrangements in the neonatal blood spots of infants and children with B-precursor ALL. The same clonotypic Ig or TCR sequences were detected in the bloodspots of both infant ALL and in two out of five cALL patients.

The question remains whether the oncogenic transformation occurs at this early stage, or later on by postnatal secondary and complementary mutations. Retrospective analysis of one twin with clonotypic TEL-AML1 primers, showed that the presumptive
preleukemic clone was present in this twin already years before clinical diagnosis. This demonstrates that the latency of ALL can be variable and protracted. The modest level of leukemia concordance in cALL (estimated 5% in monozygotic twins) suggests that fetal gene-fusion is insufficient for clinical development of cALL and that some additional genetic event(s) or exposure are required postnatally to promote evolution of the fetal preleukemic clone to frank malignancy. Considering the genotypic fusion sequences that were tracked back in utero, the initial mutation takes place in a B-progenitor cell and causes the persistence of a preleukemic cell clone. Over time, preleukemic cells might acquire additional genetic alterations, which eventually can lead to clinically overt leukemia.

But probably there is also another group of patients whose leukemia developed after birth, and in some cases just before the occurrence of clinical leukemia. We have not been able to perform these studies, although we examined preleukemic bone marrow cells of four patients who were diagnosed with cALL 6 months after diagnosis of unexplained aplastic anemia. In all four patients (aged from 8 months to 4 years) no leukemic-specific antigen receptor gene rearrangements could be found in the BM samples, taken 2 to 6 months before morphologic diagnosis. There are two possible explanations for these findings. Firstly, in these children ALL may have been initiated very shortly before clinical diagnosis. Secondly, the leukemia was initiated earlier but the sensitivity of the PCR was inadequate, with too few preleukemic cells in the bone marrow at the early time point. Not many studies with PCR analysis have been described on this subject. Knulst et al. performed immunological marker analysis in the hypoplastic BM of four children at 18 to 68 days before diagnosis of ALL and suspected ALL based on these immunological data.

In 1996, Steenbergen proposed in his thesis two possible pathways for the pathogenesis of relapse in children with ALL. After diagnosis, therapy may eradicate all leukemic cells and cure the patient, whereas new additional mutations may result in a therapy resistant clone. Also, it is possible that one of the clones present at diagnosis, is persisting during treatment, resulting in a renewed malignant dedifferentiation later in the disease. The findings of this thesis are more in favour of the latter. Quantification of MRD has given the opportunity to get more information about the kinetics of the subclones in early disease. We demonstrated that, at least in oligoclonal ALL, the relapse-causing subclones already show a different biological behaviour at the very beginning of the disease and that there is no selection during therapy due to ongoing genetic events (chapter 2). This may imply that one or more subclones exist that are hardly responsive to the treatment used. This is also illustrated by the observation that two rearrangements, present in the relapse sample, were not detectable at initial diagnosis but were detectable at the end of induction therapy. Presumably these rearrangements, present in a small leukemic subclone, were below the detection limit at diagnosis. This observation adds to our hypothesis that the
increased risk of relapse in patients with slow regression of MRD during induction therapy may be caused by therapy-resistant leukemic cells already present at diagnosis. It is very well possible that this specific group of patients cannot be treated by the current treatment and need to be treated in alternative ways.

**Clinical relevance of MRD and Future Perspectives**

Although it is tempting to try to understand the biologic mechanism behind the leukemic process, our main aim was to identify the patient at high/low risk of relapse and to define the criteria for good clinical management. The main improvement in MRD studies over the last few years, is the fact that quantification of MRD enables stratification of patients more accurately and therefore gives opportunities to adapt treatment to these results. It has become clear that not the presence of residual disease alone, but the *level* of MRD is the most informative. Also the findings of this thesis underscore the importance of quantification.

*What will be the preferable approach for MRD detection?* Both immunological and molecular techniques have been used to measure MRD. Both approaches have their advantages and disadvantages. Flow cytometric detection MRD in acute leukemia’s mainly relies on the identification of minor cell populations that can be discriminated from normal (precursor) cells on the basis of immunophenotypic aberrances observed at diagnosis. A minimal sensitivity of $10^3$ can be reached and it is expected that new developments with four and five-colour cytometry may increase sensitivity to $10^4$. Flow cytometry is easier to perform, is better suited for quantitative measurements and will give results within a shorter time-span than the PCR analysis, although there are false-negative results, due to phenotypic changes and the interpretation is more difficult. The PCR is more sensitive ($10^5$) and both antigen receptor gene rearrangements and chromosomal aberrations can be used as targets. However, specific chromosomal aberrations are only present in a small number of patients and clonal evolution may also give false-negative results. Also, characterisation of antigen receptor gene rearrangement and development of suitable primers may take some time. As the independent prognostic value of MRD is clear, it has to be decided which technique will preferably be used in the near future. Technical progress may allow the study of all patients with one single method, but it was proposed by several groups (DCLSG, St Jude’s Children’s Research Hospital) to use the PCR results in tandem with results from flow cytometry. The use of two entirely independent features of leukemic cells to monitor MRD should minimize the chance of false-negative results. Moreover, it will be possible to study almost all patients with one or two methods, since only in a few cases neither of the techniques can be used. Also, it will
be dependent on the experience of each centre which technique will be used. Our study has focused on the quantification of MRD by PCR and therefore our conclusions fit better in this context. With the RQ-PCR analysis of MRD, new possibilities arise to perform an accurate and efficient quantification of MRD on a routine basis. In contrast with the very laborious and time-consuming technique of limiting dilution, this method is fast and we did show that this technique gives reproducible results. It has been demonstrated that it is possible to use germline primer and probe combinations, homologous to the leukemia-specific rearrangement. This will reduce costs of MRD analysis by RQ-PCR. Both our results and the possibility of using TEL-AML1 transcript for quantification in TEL-AML1-positive patients will facilitate application of the PCR-based detection of MRD.

*Which time-point is the most informative?* Information about the kinetics of tumor load reduction has allowed the recognition of three MRD-based risk groups with significant differences in relapse rate. Van Dongen *et al.* showed that the combined measurement of MRD at the end of induction and at the start of consolidation therapy is of prognostic significance. A high degree of MRD \((10^2)\) at the end of induction therapy was associated with a thirty-fold higher relapse rate in comparison with a low degree of MRD \((10^4)\). These findings were obtained with a semi-quantitative PCR method. In our study of patients (chapter 5) using real-time quantitative PCR (RQ-PCR), we were able to show that quantitative MRD results at day 15 are highly informative. Therefore, it might be possible to discriminate between good- and poor-risk patients by precise quantification of the MRD level at this early time-point and at the end of induction. With the quantitative RQ-PCR, it is possible to differentiate between patients with relatively high tumor loads, and our results show that this accurate quantification is necessary for the identification for patients at high risk for relapse. Our preliminary results even suggest that it is possible to discriminate patients upon MRD results of day 15 alone. Quantitative PCR analysis will be most interesting for the intermediate group with MRD levels at the end of induction from \(10^2\) to \(10^4\). Possibly, accurate quantification will result in a better recognition of patients with slow regression within this group. Compared to the high-risk group the percentage of failures is relatively low, but the sizeable medium risk-group accounts for the largest number of unpredictable relapses. Stratification and change of treatment in the children who are likely to relapse, may have the largest consequences for improvement of survival. Yet, it is not known whether treatment stratification or whether early change of treatment will lead to higher survival.

*Implementation into routine clinical care?* MRD quantification can be used for two approaches. On the one hand, detection of MRD early during treatment to measure the
response on therapy and on the other hand MRD detection at relevant time points during follow-up to detect a molecular relapse before clinical relapse presents.

Early during treatment MRD quantification can be helpful for identification of four different patient groups. First it can be used for further intensification of treatment for patients already recognized at diagnosis as high risk, in order to improve survival. However, effective therapy (including BMT) is still lacking for some patients. It is a matter of debate whether intensification of treatment will improve survival, as results with the present world-wide used treatment protocols are already very promising for childhood ALL. After many years of clear improvement of survival, due to better treatment protocols, the last ten years a plateau is seen in survival curves. This may imply that a further improvement of prognosis can only be achieved by a more individualized, tailor-made treatment for the patients that are included in the high-risk group, based on results of MRD detection. But it is also possible that the leukemia in some patients is drug-resistant and that they can not be saved by earlier intensification of treatment or that novel treatment methods are indicated. This is best illustrated in our oligoclonal patients (chapter 2), since we showed that therapy resistant subclones were already present at diagnosis or emerged very shortly thereafter. In this latter scenario, quantification of MRD can only help to select a group of patient that needs alternative treatment.

The second group of patients are the patients that are identified as high-risk at diagnosis, based on the current prognostic factors, but who have low-risk MRD levels at the end of induction. It is the question whether in these patients the initially started high-risk treatment should be continued or whether it is safe to reduce treatment to a low-risk treatment.

A third group of patients is the group that is wrongly identified and treated as non-high risk. These patients are, based upon the currently used prognostic factors identified as non-high risk but belong, according to MRD level, to the intermediate- or high-risk group. In the future, therapy stratification for these patients can lead to a more intensive treatment and hopefully a better prognosis.

Finally, another consequence of MRD detection will be that therapy-stratification might result in less treatment in those patients presumed to have a very good outcome. This group of low-risk patients may be overtreated and suffers from unnecessary late side effects of treatment. Treatment can be changed in two ways. Firstly, therapy can be given in a shorter period, i.e. 18 months instead of 24 months, for those children that have low MRD levels in the early treatment stage and that are staying in continuous complete remission during treatment. Secondly, it is possible to give a less intensive treatment by changing the dosage and frequency of certain drugs. In this, it will be interesting to find out in clinical studies whether it is safe to reduce the intensity of the most toxic agents. Simplified
protocols, applied in developing countries, already showed that it might be possible to give shorter and less intensive treatment. Also in the past in the Netherlands a considerable number of the children with ALL survived with less intensive treatment protocols.\textsuperscript{16;134;135} But it is understandable that clinicians have reasonable doubt to change a successful treatment protocol.

The detection of MRD during long-term follow-up is a second possible application. Initial studies\textsuperscript{3;8;43} studied MRD in childhood ALL during follow-up and showed that reoccurrence of MRD after PCR negativity in the course of the disease was always followed by overt relapse. However, so far no clinical studies have been started in which reinduction therapy is initiated guided by PCR results. Although the effectiveness of therapy might be increased by starting therapy at lower levels of tumorload, much scepticism still exists among clinicians to start therapy in children who are seemingly healthy. Thus examination of follow-up bone marrow samples is not a major aim anymore.

Future studies will have to focus on the implementation of MRD detection into routine diagnostics using validated uniform criteria. MRD cut-off levels and the risk-group differentiation, as proposed in the study by van Dongen \textit{et al.},\textsuperscript{7} have to be assessed for each treatment protocol. It is important to perform MRD studies on the currently used protocols to identify the cut-off levels that make a reasonable differentiation of risk-groups. For the Netherlands, this study is now ongoing in a collaboration between Rotterdam, Amsterdam and the DCLSG.

At initial diagnosis a first rough division between low- and high-risk patients can be made upon strong risk factors like prednisolone response, age, WBC and the presence of certain translocations. In this early phase of treatment also be a for the MTT-assay as results are available within 4 days post diagnosis. Within a few weeks after diagnosis, the initial risk-group classification can be refined to a final risk-group stratification by results of kinetics of MRD. It is interesting to speculate which decision has to be made, if a patient has conflicting risk factors, i.e. high leucocyte count with low levels of MRD. We believe that the MRD level is the strongest predictor of outcome, as it gives information on the kinetics of the disease.

It will be necessary to focus on understanding the pathogenesis of ALL better in order to develop innovative therapeutic approaches. Whereas in vitro drug-resistance assays may not be additionally helpful in identifying all patients at low/high risk, these assays are interesting to study drug resistance mechanisms and ways to circumvent them. Ultimately, new treatment strategies probably need to target specific genetic lesions to improve outcome for all children with ALL. To this end, developments with cDNA-micro array and SAGE are of interest.
In conclusion, our study has contributed to PCR-based MRD detection and has shown possibilities to apply this approach in routine diagnostics. It also illustrates pitfalls of MRD detection by PCR. The main improvement of the last years is the accurate quantification of MRD. This will contribute to a better division into risk groups. What we do require is a consensus as to the most efficient and cost-effective way of defining slow response and/or persistent residual disease, and when we should adjust therapy in response to such disease persistence. Prior aim in all will be a better individualized risk-adapted treatment and early recognition of eventual relapse, finally leading to cure of all children without unnecessary late-effects.