Clinical significance of minimal residual disease quantification in adult
patients with standard risk acute lymphoblastic leukemia

Monika Brüggemann1, Thorsten Raff1, Thomas Flohr2, Nicola Gökbuget3, Makoto Nakao2, Jo
Droese5, Silke Lüschen1, Christiane Pott1, Matthias Ritgen1, Urban Scheuring3, Heinz-August
Horst1, Eckhard Thiel4, Dieter Hoelzer3, Claus R. Bartram2 and Michael Kneba1

for the German Multicenter Study Group for Adult ALL

1 Medical Clinic II, University of Kiel; Germany
2 Institute of Human Genetics, University of Heidelberg; Germany
3 Medical Clinic II, Department of Hematology, University of Frankfurt; Germany
4 Medical Clinic III, Charité Universitätsmedizin Berlin, Campus Benjamin Franklin, Germany

Scientific heading: Neoplasia
Running title: MRD in adult standard risk ALL
Word counts: Abstract 198 words
Text 4736 words

The study was supported by the Wilhelm Sander-Stiftung (Grant 2001.074.1), Deutsche
Forschungsgemeinschaft (Grant Kn 422/1-1), Deutsche Krebshilfe (Grant 70-2657-Ho2), and
by the BMBF, Competence Network Leukaemias (Grant No. 01GI 9971).

Specific contributions of all authors to published work
MB participated in the development of PCR assays, was responsible for the data collection
and interpretation, and drafted the manuscript. TR was responsible for molecular MRD
analysis and participated in interpretation of the data and manuscript preparation. TF and MN
did the testing and standardization of the PCR techniques and assessed the data in
Heidelberg. JD, SL, CP and MR optimized the PCR assays and did the MRD analyses in
Kiel. NG is the coordinator of the German Multicenter Study Group for adult ALL (GMALL),
she provided the information on presenting clinical, immunophenotypic and genetic features,
and contributed to preparation of the paper. US organized the collection, storage and
distribution of many follow-up samples. HAH performed morphological analyses of bone
marrow aspirates, ET did the immunophenotyping for classification of the leukemias. DH is
the chairman of the GMALL trials, he was responsible for the treatment protocol design and
the overall clinical conduct of the studies. CRB and MK were responsible for the overall
conduct of the MRD study, and participated in manuscript preparation.

Correspondence: Monika Brüggemann
Medical Clinic II
University of Kiel
Chenmitzstr. 33
24116 Kiel
Germany
tel. +49 431 1697 1266
fax +49 431 1697 1264
email: m.brueggemann@med2.uni-kiel.de
ABSTRACT

Adult patients with acute lymphoblastic leukemia (ALL) who are stratified into the standard risk (SR) group due to the absence of adverse prognostic factors relapse in 40-55% of the cases. To identify complementary markers suitable for further treatment stratification in SR ALL, we evaluated the predictive value of minimal residual disease (MRD) and prospectively monitored MRD in 196 strictly defined SR ALL patients at up to nine time-points during the first year of treatment by quantitative PCR. Frequency of MRD positivity decreased from 88% during early induction to 13% at week+52. MRD was predictive for relapse at various follow-up time-points. Combined MRD information from different time-points allowed definition of three risk groups (p<0.001): 10% of patients with a rapid MRD decline to $<10^{-4}$ or below detection limit at day+11 and day+24 were classified as low-risk and had a 3-year relapse rate (RR) of 0%. A subset of 23% with MRD $\geq 10^{-4}$ until week+16 formed the high-risk group with a 3-year RR of 94% (95% CI 83-100%). The remaining patients whose RR was 47% (31-63%) represented the intermediate-risk group. Thus, MRD quantification during treatment identified prognostic subgroups within the otherwise homogeneous SR ALL population who may benefit from individualized treatment.
Introduction

Investigation of minimal residual disease (MRD) has been proven to be a valuable tool for predicting outcome in childhood acute lymphoblastic leukemia (ALL).\textsuperscript{1-5} In contrast, only a few studies have focused on adult ALL and they were mostly based on patients with heterogeneous risk profiles and different kinds and intensities of treatment.\textsuperscript{6-8} However, monitoring homogeneous patient cohorts at different time-points during therapy might provide additional insight into the nature and clinical relevance of MRD kinetics in adult ALL which is particularly relevant for the large population of standard risk (SR) patients without conventional risk factors. Relapses in this patient group occur in about 40-55\% of cases and cannot be predicted by any known conventional risk factor.\textsuperscript{9-11} In a number of clinical studies this led to a policy of stem cell transplantation (SCT) in first remission\textsuperscript{12-14} causing overtreatment and additional expenses for those patients who are cured by conventional chemotherapy alone. Therefore definition of prognostic factors allowing to discriminate SR patients with poor outcome after standard chemotherapy from those with a favorable prognosis is highly warranted. Currently the most widely used techniques to detect and quantify residual disease in patients with ALL are multiparameter immunophenotypic evaluation of aberrant protein expression\textsuperscript{2; 3; 8} and clone-specific PCR amplification of immunoglobulin (Ig) and T-cell receptor (TCR) gene rearrangements.\textsuperscript{1; 4; 5} Such molecular targets can be identified in more than 90\% of patients with ALL by the use of various PCR primer sets. Besides its large applicability and high sensitivity a main advantage of PCR based assays is the use of DNA as a stable and easy conveyable specimen which is particularly relevant in large multicenter studies. Until now, most molecular studies used semiquantitative or qualitative PCR methods for detection of Ig and TCR gene rearrangements.\textsuperscript{1; 5-7} However, in adult ALL there is a high need for precise quantification to define
discriminating MRD levels at different sampling time-points, as MRD positivity is significantly more frequent in adults than in comparably treated children.\textsuperscript{15} \textsuperscript{16} Therefore, we and others have developed different immune gene real-time quantitative PCR (RQ-PCR) assays to permit accurate MRD quantification during the exponential phase of PCR amplification.\textsuperscript{17} \textsuperscript{21} This allows the quantitative detection of one leukemic cell among 10000 normal cells with a good quality control, standardization and comparability of MRD data. Using these RQ-PCR assays, we conducted a prospective study on the predictive significance of MRD monitoring in adult SR ALL within the German Multicenter Study Group for Adult ALL (GMALL) trial with more than 100 participating centers that recruits more than 60% of the German incidence of adult ALL minimizing the risk of a selection bias and suggesting a high external validity.\textsuperscript{22} Residual disease was sequentially monitored at multiple time-points during therapy of 196 consecutive adults with standard risk ALL who were treated according to the GMALL 06/99 standard risk protocol. In a pilot study with 65 patients logistics and optimal sampling time-points as well as threshold levels were established.
METHODS

Patients

A total of 323 adults with ALL were included in the MRD trial at time of diagnosis. The study was divided into two phases: (1) A pilot phase within the GMALL 05/93 trial (January 1997-September 1999) to establish logistics, validate quantification methods and required sensitivities, define sampling time-points and MRD threshold levels (2) a main phase (October 1999-December 2002) to prospectively prove prognostic significance of MRD in strictly defined adult standard risk ALL patients who were consecutively enrolled in the GMALL 06/99 trial after written informed consent. Approval for these studies was obtained from the institutional review boards at each of the participating institutions.

The MRD pilot study comprised 71 adult ALL patients including 27 patients with high-risk features with available bone marrow samples at diagnosis and at up to six times during the first year of therapy (figure 1).

For the main GMALL 06/99 MRD trial, only standard risk ALL patients as defined by the following criteria were prospectively investigated: (A) absence of the translocations t(4;11)/MLL-AF4 and t(9;22)/BCR-ABL (B1) c-/pre-B-ALL with white blood cell (WBC) count <30000/μl at diagnosis or (B2) cortical- or mature-T-ALL with WBC count <100000/μl, since 2000 restriction to cortical-T-ALL irrespective of WBC count (C) age 15-65 years and (D) achievement of complete remission after phase I of induction therapy. Over the collection period, 291 consecutive GMALL 06/99 patients fulfilled inclusion criteria, bone marrow samples at diagnosis and at up to nine follow-up time-points during first year treatment were prospectively collected from 252 of these patients (87%). The nine bone marrow sampling times were defined according to fixed steps along the treatment protocol: mid-induction I (day+11), end of induction I (day+24), end of induction II (day+44), pre-consolidation...
I (week+11) and II (week+16), pre-reinduction (week+22), pre-consolidation III (week+30) and V (week+41) and end of first year of treatment (week+52) (figure 1). 33/252 patients had to be excluded due to an insufficient amount or quality of DNA in diagnostic and/or follow-up samples, therefore, molecular characterization of TCR-/Ig-rearrangement patterns was finally performed in 219 patients.

**Treatment**

In brief, GMALL 06/99 induction therapy phase I consisted of a six-drug regimen given over a period of three weeks with dexamethasone, cyclophosphamide, vincristine, daunorubicine, asparaginase, and intrathecal methotrexate. After attaining a complete clinical remission, patients received 3 weeks of induction phase II therapy with cyclophosphamide, cytarabine, mercaptopurine and intrathecal methotrexate. Consolidation therapy started at week+11 and consisted of six cycles with alternating combinations of drugs over a period of 40 weeks and was interrupted by a four week reinduction therapy at week+22 (see figure 1). All patients received intrathecal therapy with methotrexate, cytarabine and dexamethasone for one year and cranial irradiation (24 Gy). Mediastinal irradiation was added for patients with residual mediastinal tumor mass after induction therapy (24 Gy). The generally recommended maintenance therapy lasted for additional 12 months and consisted of alternating drug combinations similar to the consolidation cycles with intermediary application of mercaptopurine plus methotrexate.

Treatment elements in the GMALL 05/93 study were similar to the GMALL 06/99 protocol and are described elsewhere.9

**Detection of residual disease**
Mononuclear cells were isolated from bone marrow and stored in liquid nitrogen or at 
-80°C until extraction. Samples were analysed at one of two central laboratories (in 
Heidelberg and Kiel, Germany). Standardization of screening PCR for detection of 
clonal markers was performed in the context of the BIOMED-1 and BIOMED-2 
Concerted Action. Rearrangements of the TCR genes TCRB, TCRG and TCRD, 
and the Ig genes IGH and IGK-Kde were sought by PCR amplification in samples 
obtained at diagnosis. Clonality was confirmed either by heteroduplex analysis or by 
GeneScanning. After sequencing, allele-specific oligonucleotides (ASO) were designed for each MRD 
target on the basis of the sequence data of the junctional regions, using OLIGO 6.3 
software (Cascade, CO, USA). Tests for residual disease were conducted by RQ-
PCR amplification of 500 ng DNA of follow-up samples, using either TaqMan- or 
LightCycler technology. Both systems have been demonstrated to produce 
concordant results with the former standards for molecular MRD quantification, ASO-
PCR and dot blot hybridization. In addition, multicenter quality control of 
quantification data was achieved by interlaboratory tests in the frame of the 
European Study Group on MRD detection in ALL (ESG MRD ALL) under the 
direction of JJM van Dongen, Rotterdam. Fifteen of the patients investigated in the 
pilot phase were quantified by conventional ASO-PCR and dot-blot hybridization, as 
described earlier. MRD levels and individual sensitivity threshold values were 
corrected according to quality of DNA that was checked by amplification of control 
gene segments (albumin or β-actin). Follow-up samples judged MRD negative were 
excluded in case of an amplifiability of less than 10%. The LightCycler measured 
leukemia-specific PCR products, generated by ASO-PCR, at each cycle by staining 
the PCR product with the DNA-binding dye SYBR Green I. For TaqMan RQ-PCR, 
we and others developed target-specific assays with a set of different germline
TaqMan probes (13 TCRB-Jβ, 8 TCRG-Vγ, 1 TCRD-Jδ1, 1 TCRD-Dδ2, 3 TCRD-Vδ, 4 IgH-JH and 1 IGK-Kde)\textsuperscript{17,18,20,21} and ASO primers. If MRD levels, quantified by two or more targets, differed, the higher MRD level was assumed to be more accurate. In case of MRD positivity but levels below the reproducible range, MRD levels were stated as a range between the limits of reproducibility and sensitivity. If no specific PCR product was detectable, this time-point was considered MRD negative irrespective of the PCR-target sensitivity. In total, molecular analysis was performed in 290 patients. In 21/290 patients (7%), no clonal marker was detected, in additional 8/290 cases (3%) clonal markers were not suitable for MRD quantification. Finally, residual disease was evaluated in 261 adult ALL patients (65 within the pilot study 05/93 and 196 within the 06/99 trial). A total of 418 different allele-specific assays were performed for these patients, 108 of them targeting clonal TCRB gene rearrangements, 79 TCRG, 61 TCRD, 126 IgH and 44 IGK-Kde rearrangements. 137 of 261 patients were assayed with two (121/261) or more than two (16/261) PCR targets. In 233/261 cases (89%), sensitivity of at least one assay reached ≤10\textsuperscript{-4}, with a sensitivity limit of 10\textsuperscript{-4} in 77/261 cases (30%), a limit of 5\texttimes10\textsuperscript{-5} in 61/261 cases (23%) and a limit of 10\textsuperscript{-5} in 95/261 cases (36%).

**Statistical analyses**

Distribution of variables between groups was compared using Fisher`s exact test or Chi-square test, the Mann Whitney \textit{U} test was used to estimate significance of differences in continuous parameters. The end point to determine the prognostic significance of variables studied was the disease free survival (DFS), calculated as the interval between first documented complete remission and relapse or end of observation. Patients who underwent SCT in first remission (11 patients of the GMALL 06/99 trial), patients with premature termination of first year treatment (15
GMALL 06/99 patients) and patients who died in complete remission (4 GMALL 06/99 patients) were included but censored at the date of event. Kaplan-Meier estimates were calculated for the time-to-event variables DFS and overall survival (OS). Comparison between curves was performed using the log-rank test, or the log-rank test for linear trend in case of ordered categorical variables. The influence of potential prognostic factors on DFS was estimated with the stepwise Cox proportional hazard model, including age, gender, WBC count, immunophenotype and MRD levels at different time-points.

All the statistical analyses were performed using the GraphPad Prism version 3.02 for windows (GraphPad Software, San Diego, CA, USA) and SPSS version 11.5 (SPSS Inc., Chicago, IL, USA).
RESULTS

Results of the pilot study: definition of sampling time-points and threshold levels

148 first year follow-up samples of 65 ALL patients were evaluable for MRD analysis. Nineteen patients relapsed during the observation period, median follow-up of patients in continuous complete remission was 55 months, estimated 5-year DFS rate was 55.0% (95% confidence interval CI 39.9-70.1%). Ten out of 27 patients with adverse prognostic factors (table 1) were transplanted in first remission, another five patients died during remission.

Percentage of MRD positivity decreased from 71% (20/28) after induction phase I (day+29) to 42% (15/36) before start of consolidation (week+13) and to 30% (8/27) before start of reinduction (week+21). Afterwards follow-up samples of only three patients were MRD positive (one positive sample out of 29 (week+30), 18 (week+46) and 11 (week+52) samples analysed, respectively).

Presence of MRD with a level of at least $10^{-4}$ before start of consolidation (day+29 and/or week+13, figure 1) was associated with a greater risk of relapse: 5-year DFS rate was 32.6% (5.3-59.9%) in 21 patients with MRD $\geq 10^{-4}$ within this period, compared to 71.8% (50.0-93.6%) in 27 patients with low (<$10^{-4}$) or undetectable MRD levels (p=0.003). Frequency of patients with adverse prognostic factors was significantly higher in the MRD high level group (16/21) compared to the group with low or undetectable MRD (6/27), p<0.001. However, multivariate analysis could not be done because of the small number of patients. During consolidation therapy frequency of MRD positive samples and relapse rate was too small for final statements. 6/8 patients with measurable MRD at week+21 relapsed 3 to 34 months after MRD assessment. In contrast, only 2/19 patients without detectable MRD at this time-point relapsed after 33 and 36 months, respectively. Probability of 5-year DFS
was 81.7% (57.6-100%) in MRD negative patients, compared to 0% for MRD positive patients (p<0.001). In conclusion for MRD evaluation of standard risk ALL patients in the GMALL 06/99 trial, (1) a sensitivity limit of $10^{-4}$ was targeted for RQ-PCR assays (2) $10^{-4}$ was defined as crucial threshold MRD level before start of consolidation treatment and (3) additional MRD evaluation time-points during the early course of therapy were defined to gain more insights into MRD kinetics (see figure 1), because residual disease at later time-points dropped below detection limit in almost all patients.

**GMALL 06/99 MRD results for adult standard risk patients**

The 196 standard risk ALL patients were monitored at up to nine time-points during the first year of therapy (figure 1). As the vast majority of all German adult ALL patients are enrolled in the GMALL trials and 87% of all eligible GMALL 06/99 standard risk patients were included into molecular analysis over the collection period, the main potential bias was availability of material and presence of sensitive clonal markers. The investigated standard risk ALL study population was compared to the remaining 95 adult GMALL 06/99 standard risk ALL patients who were not assayed. The clinical features (sex, age, WBC count, immunophenotype and 3-year cumulative incidence of relapse) did not differ between the two groups, with the exception of a higher proportion of cases comprising a T-cell phenotype in the MRD study population (37% vs 22%, p=0.02). Within the observation period, 67 patients relapsed, four patients died during remission, in 15 patients therapy was stopped prematurely during the first year of treatment. In 35/63 cases of medullary relapse bone marrow samples at clinical relapse were available. In 5/52 investigated targets (10%) false-negative results were obtained at relapse probably reflecting changes in the Ig and TCR gene rearrangements during the disease course. This concerned
2/15 (13%) IGH-, 1/4 (25%) IGK-, 1/16 (6%) TCRB-, 0/8 (0%) TCRG- and 1/9 (11%) TCRD-targets and led to a failure to detect the relapse by PCR in 4/35 patients (11%). Estimated 3-year DFS was 52.7% (95% CI 43.5-61.9%), median follow-up period of the patients in continuous complete remission was 30 months.

MRD was detected in 470 of 1196 evaluable follow-up samples, additional 113 samples (8.6% of all samples) were excluded due to insufficient quality or quantity of DNA, particular at day+11 (29/142 samples [20.4%]) and day+44 (22/136 samples [16.2%]), when hypoplasia was frequently found. In 79 MRD negative samples with moderate restriction of quality (control gene amplifiability between 10 and 100%) individual sensitivity thresholds had to be scaled up to >10^{-4} despite of sensitive underlying RQ-PCR assays. In case of detectable MRD with levels <10^{-4}, MRD values were mostly not exactly quantifiable as they were outside the reproducible range of the RQ-PCR assays.\textsuperscript{18,20,21} During induction phase I (day+11), MRD at any level was detectable in the majority of patients (88%). Percentage of MRD positivity decreased to 63% at day+24 and to 39% at week+11. At week+22, MRD was detected in 29% of the standard risk patients, after completion of first year therapy in 13% of patients (figure 2). Of note, the percentage of MRD positivity did not differ significantly in patients with T-lineage ALL compared to B-lineage ALL (p>0.05 for every single time-point). Median MRD levels of MRD positive samples were 7x10^{-2} and 1x10^{-3} during and after induction I (day+11 and day+24, respectively). Afterwards median MRD levels ranged between 3x10^{-4} and 7x10^{-4} in patients with detectable disease.

Figure 3 summarizes the estimated 3-year DFS rates for the different time-points depending on MRD levels. According to the results of the pilot study, a threshold level of 10^{-4} showed the strongest discriminative power after the end of induction I (day+24) until start of consolidation therapy (week+11) and was also used as cut-off
point for week+16 to week+52. During induction I (day+11), MRD loads of greater than or equal to $10^{-2}$ were associated with a 3.2 fold higher incidence of relapse than lower degrees or absence of detectable disease. A tumor load below detection limit (irrespective of target sensitivity) allowed a further substratification of patients with respect to a favorable outcome (see figure 3). After the end of induction I (day+24), the relative risk for relapse in patients with MRD levels of $\geq 10^{-4}$ compared to levels <$10^{-4}$ or below detection limit was 2.4 (95% CI 1.3-4.2). For day+44 and week+11, the accuracy of MRD in defining patients with relapse was reduced, as ongoing treatment reduced MRD in most patients to levels close to $10^{-4}$ irrespective of outcome. MRD information during later course of therapy (week+16 to week+52) narrowed a smaller population of 26% to 9% of patients with persistent disease $\geq 10^{-4}$ and a three- to fivefold increase in relative relapse rates compared to patients with low (<$10^{-4}$) or undetectable MRD (95% CI of relative risks: week+16: 2.3-7.3, week+22: 1.9-7.1, week+30: 1.5-5.9, week+41: 1.5-7.2, week+52: 2.3-11.4). Since only standard risk ALL patients within the GMALL 06/99 trial were investigated, the study population was highly homogeneous concerning classical adverse prognostic factors. The remaining clinical and biological variables (sex, age, WBC count, and B-versus T-lineage ALL) were each tested as single variables in the Cox regression model in addition to MRD. The cut-off points for the quantitative variables WBC count and age were identical to those in table 1. MRD was found to be the only variable that showed a significant impact on outcome.

**Relation between residual disease on day 24 and MRD during consolidation therapy**

Low or undetectable levels of day+24 MRD were related to MRD kinetics during later course of therapy. 105 patients were assayed at day+24, week+16 and beyond
In 47/105 patients, day+24 MRD levels were undetectable or <10^{-4}. MRD remained low or undetectable at week+16 in 45/47 cases (96%), no MRD value ≥10^{-4} was measured from week+16 to week+52 in 42/47 cases (89%).

In contrast, 40/58 patients (69%) with a high tumor load (≥10^{-4}) on day+24 subsequently achieved low (<10^{-4}) or undetectable MRD at week+16. 18/58 patients (31%) showed high level of MRD up to week+16 and 16 of them remained positive >10^{-4} at later time-points during treatment.

**Identification of MRD-based risk-groups**

MRD analyses at single follow-up time-points discriminated standard risk ALL patients with different risks of relapse: Patients with a rapid tumor clearance and low relapse rates were recognized by day+11 MRD assessment, whereas patients with persistent disease ≥ 10^{-4} at week+16 showed an extremely poor outcome. These crucial time-points were therefore adducted to define an MRD-based low- (day+11) and high- (week+16) risk group. However, particularly in case of a reduced amplifiability of an MRD negative follow-up sample or MRD positivity in the range of 10^{-4} it might be important to confirm results by a second MRD test. To enhance accuracy of this risk-stratification, we therefore added day+24 MRD information because this value pre-defined two large populations of patients about the same size with different risks of relapse and generally allowed sensitive and reliable MRD assessment.

This stratification was applicable for 105 patients who did not significantly differ in distribution of sex, age, immunophenotype and DFS from the group of 91 patients without MRD measurement at crucial time-points.
Eleven patients with low (<$10^{-4}$) or undetectable levels of day+11 and day+24 MRD formed the MRD-based low-risk group and had a 3-year DFS and OS rate of 100%. Median follow-up of the patients was 33 (12 to 51) months. Pre-therapeutic clinical features were similar to that of the whole study population. No patient relapsed during the observation period, with one patient being censored after one year, as treatment was prematurely stopped prior to consolidation V due to severe treatment related complications. However, also this patient is still in complete remission after additional 15 months. 24 patients with MRD levels greater than or equal to $10^{-4}$ at both day+24 and week+16 were classified as high-risk. They showed a 3-year DFS rate of only 5.8% (0.0-16.7%) and a 3-year OS rate of 45.1% (22.2-68.2%). Their risk of relapse was increased by the factor 1.4 (1.2 to 1.7) as compared to the whole study population. Until the end of observation, 19 of these patients relapsed, three were censored due to individual treatment modification (in one patient treatment was prematurely stopped, two patients were transplanted in first remission). Only two patients are still in continuous complete remission at 21 and 38 months, respectively. The remaining 70 patients were combined into an intermediate risk group. They had a DFS rate of 53.2% (36.9-69.5%) (p(trend)<0.001) and an OS rate of 69.8% (54.8-84.8%) (p(trend)<0.001; figure 4). The risk of relapse for this MRD-based risk-group did not significantly differ as compared to the whole study population. A further substratification of this patient group adding other time-points and/or MRD threshold levels using the data set acquired within this study was not successful (data not shown).
DISCUSSION

The progress in treatment of adult ALL patients without conventional risk factors has been hampered by the inability to predict relapse after patients achieved a complete remission. Prospective MRD trials in large cohorts of homogeneously treated adult SR patients to define complementary prognostic markers are lacking. Within the GMALL 06/99 trial we were able to demonstrate that sequential monitoring of residual disease is a powerful indicator of treatment outcome. Using the combined information on day+11, day+24 and week+16 MRD, patients with a rapid tumor clearance and favorable outcome were discriminated from those with persistent disease and a particular high relapse rate (see figure 4). Since MRD was monitored prospectively in the GMALL trial that includes more than 60% of the incidence of adult ALL (15-65 years) in Germany and all patients without adverse conventional risk factors were intended to be investigated, a selection bias within the SR-ALL population was effectively minimized. The definition of adverse factors emerged from the more than 3000 adult ALL patients homogeneously treated in the GMALL trials over a period of 20 years. The 3-year relapse rate of 47% in our SR-ALL study population was in keeping with the results of other adult ALL trials, and relapses were not predictable by conventional clinical and biological factors. Therefore, the only known variable potentially influencing outcome in the study population was MRD.

With regard to MRD kinetics, frequency of MRD positivity tended to be higher than reported for childhood ALL. In our study 63% of adult patients had measurable MRD at day+24 and 47% at day+44, whereas in five large prospective studies on childhood ALL, residual disease was detectable in 25% to 58% of patients after four to six weeks of induction therapy. Also at later time-points, percentage of MRD positivity in adult patients was higher. MRD was detected in 23% of cases at
week+30, compared to only 10-13% of pediatric ALL.\textsuperscript{1-5} This probably reflects the higher in vivo drug resistance of adult ALL and was also reported by Mortuza et al.\textsuperscript{7}

MRD quantification during induction (day+11) identified patients with a very rapid molecular response and an excellent prognosis (3-year DFS 92%) in line with reports on childhood ALL.\textsuperscript{27, 28}

Also the extent of residual disease after induction therapy predicted treatment outcome, although patients who reached the condition of MRD clearance down to $10^{-4}$ or below detection limit only at this treatment phase still had a poorer prognosis when compared to those who achieved an early (day+11) profound MRD reduction. Therefore, the same MRD status but different length of time to get it resulted in a different prognosis. The failure to identify a subgroup with an excellent prognosis by post-induction MRD assessment is most likely explained by the PCR sensitivity limit of about $10^{-4}$ to $10^{-5}$ in the present study. Considering that most patients became RQ-PCR negative after induction treatment and the number of residual blasts in these patients might vary between 0 and $10^{8}$ additional subgroups with different MRD kinetics could probably be identified using more sensitive detection techniques.

In our study, post-induction MRD provided important information by identifying patients with an extremely poor prognosis. Single MRD tests at week+16 and week+22 narrowed a population of 38/148 patients (26%) and 25/126 patients (20%), respectively, with a 3-year DFS of 12%. This is in keeping with findings of Vidriales et al.\textsuperscript{8} in an immunophenotypic analysis on 102 adolescent and adult ALL patients who demonstrated a high discriminative power of day+35 MRD but a relapse rate of about 50% even in patients with MRD levels <0.05%. Similarly, Brisco et al.\textsuperscript{15} analyzed MRD in 27 adults by PCR reporting that 8/11 patients (73%) with MRD $>10^{-3}$ relapsed compared to 6/16 (38%) with MRD $<10^{-3}$ after the end of induction (day+22 to +68). Mortuza and colleagues\textsuperscript{7} investigated 85 adult patients with B-lineage ALL.
DFS for patients with detectable MRD 3-5 months and 6-9 months after diagnosis, was 11% and 0%, respectively, as compared to 74% and 80% in MRD-negative patients.

Molecular MRD assessment using Ig-/TCR-gene rearrangements as PCR targets might be hampered by the occurrence of continuing rearrangements. In our study population this is the most likely explanation for false negative PCR results in 11% of the investigated samples obtained at relapse and stresses the importance of the use of two molecular targets for MRD quantification.

Percentage of MRD positivity did not significantly differ in patients with T-lineage ALL compared to B-lineage ALL for every single time-point which was in line with the observation that relapses were not predictable by immunophenotype. For childhood ALL, Willemse et al. identified differences in MRD kinetics between T- and B-lineage ALL with a higher frequency of MRD positivity in T-lineage ALL, but children with T-ALL generally have a poorer prognosis than those with precursor-B-ALL, whereas in adult patients different clinical studies showed a poorer outcome for B-lineage ALL. However, results of different trials are not fully conclusive and comparability of data is hampered, as we exclusively investigated standard risk patients.

In this study we have shown, that molecular MRD quantification in adult ALL is feasible even in large multicenter studies. Early MRD assessment allowed the identification of patients with a high chance for cure by chemotherapy alone, persistent detectable disease during consolidation was associated with a high risk of relapse. Therefore, day+11 and week+16 MRD information was adducted to define an MRD-based low-(day+11) and high-(week+16) risk group. To enhance accuracy for MRD-based risk-stratification, results were approved by a second MRD test (day+24). This stratification was applicable to 105 patients who did not differ in
presenting features from the group of patients without MRD measurement at crucial
time-points. The relatively high percentage of exclusions due to missing samples is
explainable by the fact that this was the first prospective MRD trial within the GMALL
studies checking a considerable number of follow-up time-points and recruiting
patients of more than 100 participating centers. The fraction of evaluable patients
was comparable to the results of the first large prospective MRD studies on pediatric
ALL.\textsuperscript{1-5} However, looking at the distribution of MRD-based risk groups, percentages
substantially differed: In pediatric patients MRD-defined low-risk groups made up for
40\% to 90\% of patients whereas only 5\% to 15\% of patients belonged to the MRD-
based high-risk groups.\textsuperscript{1-3,5} Probably differences would have been even more
pronounced if adult patients with high-risk features were added into our analysis
because median MRD levels in high-risk ALL appear to exceed that of standard risk
ALL\textsuperscript{5, 8, 37} as also the results of the GMALL 05/93 MRD pilot study indicate. The small
size of the MRD-based low-risk group in our study compared to pediatric trials might
reflect differences in biology of the disease, as even SR ALL in adults has a much
poorer prognosis than childhood ALL. In addition, we applied extremely strict criteria
to define the MRD-based low-risk group as specific as possible, accepting a loss of
sensitivity, in order to minimize the risk of relapse for the individual patient being
assigned to the MRD-based low-risk group. The fraction of patients that belongs to
the MRD-based intermediate-risk group is higher compared to childhood ALL.
Although these patients will receive standard therapy further on they might profit from
establishment of patient specific MRD assays as they can easily be monitored during
and after maintenance therapy. Whether a disease recurrence can be identified in
time to allow early intervention prior to a clinical relapse is currently prospectively
analyzed.
The type of the treatment protocol, timing of the follow-up samples and the applied MRD technique might influence the definition of the MRD-based risk-groups. Therefore, precise MRD-threshold levels for risk-group assignment have to be defined carefully for each treatment protocol before MRD-based risk-stratification can be implemented.
PARTICIPATING CENTERS (30 centers selected according to number of patients recruited, listed in alphabetical order)

Berlin- B Dörken, WD Ludwig, U Peters (Charité Universitätsmedizin); Dresden- G Ehninger, R Naumann (Klinikum Carl Gustav Carus); Düsseldorf- R Haas, S Knipp (Universitätsklinikum); Essen- U Dühren, S Mahlmann (Universitätsklinikum), W Heit, KH Baur (Kliniken Essen Süd); Frankfurt- D Hoelzer, N Gökbuget (Universitätsklinikum); Giessen- H Pralle, M Dörner (Universitätsklinikum); Hamburg- N Schmitz, J Rutjes (Allgem. Krankenhaus St. Georg); Hamm- L Balleisen, A Grote-Metke (Evangelisches Krankenhaus); Hannover- A Ganser, H Diedrich (Medizinische Hochschule); Homburg/Saar- M Pfremendschuh, F Hartmann (Universitätsklinikum); Jena- K Höffken, U Wedding (Universitätsklinikum); Karlsruhe- JT Fischer, S Wilhelm (Städt. Klinikum); Kiel- M Kneba, M Lamprecht (UKSH); Köln- M Hallek, P Staib (Universitätsklinikum); Mainz- C Huber, J Beck (Universitätsklinikum); Marburg- A Neubauer, M Jänike (Klinikum Lahnberge); Minden- H Bodenstein, H Lampe (Klinikum Minden); München- C Nerl, T Lipp (Krankenhaus München-Schwabing); C Peschel, F Schneller (Klinikum Rechts d. Isar); W Hiddemann, G Lenz (Universitätsklinikum Grosshadern); Münster- WE Berdel, M Stelljes (Universitätsklinikum); Nürnberg- M Wilhelm, J Neteler (Klinikum Nürnberg Nord); Oldenburg- HJ Illiger, B Metzner (Klinikum Oldenburg); Potsdam- R Pasold, A Gerhardt (Klinikum Ernst von Bergmann); Stuttgart- W Aulitzky, L Leimer (Robert Bosch-Krankenhaus); Tübingen- L Kanz, M Schmalzing (Universitätsklinikum); Ulm- H Döhner, M Schmid (Universitätsklinikum); Wiesbaden- N Frichhofen, C Gerlach (Dr.-Horst-Schmidt-Kliniken).
ACKNOWLEDGEMENTS

We thank the participants of the German Multicenter Study Group for adult ALL for their close collaboration in the MRD study.

We thank R Reutzel and S Hug for their outstanding logistic support. We also thank JU Siebmann for undertaking the statistical analyses. We thank B Brix, L Lorenzen, F Hemken, H Seppelt and N Passow for their excellent technical assistance.
REFERENCES


LEGENDS TO FIGURES

Figure 1 (A) Scheme of GMALL 06/99 SR first year treatment (B) bone marrow sampling time-points in the GMALL 06/99 trial and (C) bone marrow sampling time-points in the GMALL 05/93 MRD pilot trial (time-points synchronized according to the treatment phase).

ADR= adriamycin. CP=cyclophosphamide. DEXA=dexamethasone. DNR=daunorubicine. (HD-)ARAC=(high dose) cytarabine. (HD-)MTX=(high dose) methotrexate. 6MP=mercaptopurine. PEG-ASP=PEG-asparaginase. PRED=prednisolone. TG=thioguanine. VCR=vincristine. VDS=vindesine. VM26=teniposide. VP16=etoposide.

Specifics of drugs and doses used in the main GMALL 06/99 treatment protocol:

**Induction I** (including preinduction treatment): DEXA 10 mg/m² orally (PO) d1-5, 11-14; CP 200 mg/m² intravenously (IV) d1-3*; PEG-ASP 1000 U/m² IV d18**; VCR 2 mg IV d4+11+18; DNR 45 mg/m² IV d4+5+11+12**; MTX 15 mg intrathecally (IT) d1. **Induction II**: CP 1000 mg/m² IV d1+21; ARAC 75 mg/m² IV d3-6, 10-13, 17-20; 6MP 60 mg/m² PO d1-21; MTX 15 mg IT d3+10+17. **Consolidation I**: DEXA 10 mg/m² PO d1-5; VDS 3 mg/m² IV d1; HD-MTX 1500 mg/m² IV d1; VP16 250 mg/m² IV d4+5; HD-ARAC 2x2000 mg/m² IV d5. **Consolidation II**: HD-MTX 1500 mg/m² IV d1+15; PEG-ASP 500 U/m² IV d2+16; 6MP 60 mg/m² PO d1-7, 15-21. **Reinduction I+II**: PRED 3x20 mg/m² PO d1-14; VDS 3 mg/m² IV d1+7; ADR 50 mg/m² IV d1+7; IT triple chemotherapy with MTX 15 mg, ARAC 40 mg, DEXA 4 mg IT d1+15; CP 1000 mg/m² IV d15; ARAC 75 mg/m² IV d17-20, 24-27; TG 60 mg/m² PO d15-28. **Consolidation III**: see Consolidation II. **Consolidation IV**: ARAC 150 mg/m² IV d1-5; VM26 100 mg/m² IV d1-5; IT triple therapy (see above) d1. **Consolidation V**: CP 1000 mg/m² IV d1; ARAC 500 mg/m² IV d1; IT triple therapy (see above) d1. **Consolidation
**VI**: see Consolidation II; IT triple therapy (see above) w52. *CP part of Induction I protocol only until y2000. **Reduced dose intensity in patients > 55 years.

**Figure 2:** Percentage of patients without detectable MRD (bar white), detectable MRD below (bar light gray) and above/equal to (bar dark gray) defined threshold values (10⁻² for day+11, 10⁻⁴ for all other time-points) at the different follow-up time-points. Total number of patients varied at different time-points because sufficient follow-up material was not available from all patients for each time-point (frequency of low quality/quantity DNA was relatively high at time-point day+11 and day+44). * For time-points day+44 to week+52 one (week+16 and week+41), two (week+11, week+22, week+30, week+52) or three (day+44) MRD positive samples could not be assigned to the MRD levels < 10⁻⁴ or ≥ 10⁻⁴ because of values below quantitative detection limit and range between quantitative detection limit and sensitivity limit spanned 10⁻⁴.

**Figure 3:** Probability of disease free survival (DFS) according to MRD results at nine time-points during first year of therapy. Numbers of patients within each group and estimated DFS rates at three years (with 95% confidence interval) are also stated.

**Figure 4:** MRD-based risk-groups: (A) Categorization scheme according to combined MRD results of day+11, day+24 and week+16. (B) Probability of disease free survival (DFS). (C) Probability of overall survival (OS). LR=Low-risk group. IR=Intermediate-risk group. HR=high-risk group.
Figure 1
% of patients with detectable MRD < threshold (light grey)  
or without detectable MRD (white)

% of patients with detectable MRD ≥ threshold (dark grey)  
(threshold: $10^{-2}$ for day 11, $10^{-4}$ for all other time-points)

Figure 2
Figure 3
Figure 4

(A) Diagram showing the distribution of patients based on different criteria and follow-up times.

(B) Graph showing DFS (%) over years with a trend p-value of <0.001.

(C) Graph showing OS (%) over years with a trend p-value of <0.001.
**Tab. 1:** Clinical characteristics and outcome of patients included in the GMALL MRD trials

<table>
<thead>
<tr>
<th>Variable</th>
<th>GMALL 05/93 (MRD pilot trial)</th>
<th>GMALL 06/99</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>65</td>
<td>196</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>43 (66%)</td>
<td>125 (64%)</td>
</tr>
<tr>
<td>female</td>
<td>22 (34%)</td>
<td>71 (36%)</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>years median</td>
<td>35</td>
<td>30</td>
</tr>
<tr>
<td>range</td>
<td>16-66</td>
<td>15-64</td>
</tr>
<tr>
<td>number of patients &lt; 35 years</td>
<td>32 (49%)</td>
<td>121 (62%)</td>
</tr>
<tr>
<td>≥ 35 years</td>
<td>33 (51%)</td>
<td>75 (38%)</td>
</tr>
<tr>
<td><strong>WBC count</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/µl median</td>
<td>17600</td>
<td>8900</td>
</tr>
<tr>
<td>range</td>
<td>1700-3338000</td>
<td>860-731100</td>
</tr>
<tr>
<td>number of patients with WBC count &lt; 30000/µl</td>
<td>36 (60%)</td>
<td>153 (81%)</td>
</tr>
<tr>
<td>≥ 30000/µl</td>
<td>24 (40%)</td>
<td>36 (19%)†</td>
</tr>
<tr>
<td><strong>Immunophenotype</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B-lineage</td>
<td>52 (80%)</td>
<td>124 (63%)</td>
</tr>
<tr>
<td>T-lineage</td>
<td>13 (20%)</td>
<td>72 (37%)</td>
</tr>
<tr>
<td><strong>presence of adverse prognostic factors ‡</strong></td>
<td>27 (42%)</td>
<td>0 (0%)§</td>
</tr>
</tbody>
</table>

* Exact white-blood cell (WBC) count was available for 60/65 GMALL 05/93 patients and 189/196 GMALL 06/99 patients.

† All patients with a WBC count >30000/µl had a T-lineage ALL, as in B-lineage ALL a WBC count > 30000/µl was defined as adverse prognostic factor.

‡ Adverse prognostic factors were: subtype of a pro-B-ALL (n=6), presence of a t(9;22)/BCR-ABL (n=14) or a t(4;11)/MLL-AF4 (n=2) (cytogenetic/molecular data available for 45 patients), WBC count >30000/µl in B-lineage ALL (n=12).

§ Absence of adverse prognostic factors was the precondition for inclusion into GMALL 06/99 MRD SR trial.
Clinical significance of minimal residual disease quantification in adult patients with standard risk acute lymphoblastic leukemia

Monika Bruggemann, Thorsten Raff, Thomas Flohr, Nicola Gokbuget, Makoto Nakao, Jo Droese, Silke Luschen, Christiane Pott, Matthias Ritgen, Urban Scheuring, Heinz-August Horst, Eckhard Thiel, Dieter Hoelzer, Claus R Bartram and Michael Kneba

Information about reproducing this article in parts or in its entirety may be found online at: http://www.bloodjournal.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at: http://www.bloodjournal.org/site/misc/rights.xhtml#reprints

Information about subscriptions and ASH membership may be found online at: http://www.bloodjournal.org/site/subscriptions/index.xhtml

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include digital object identifier (DOIs) and date of initial publication.