Prediction of outcome by early bone marrow response in childhood acute lymphoblastic leukemia treated in the ALL-BFM 95 trial: differential effects in precursor B-cell and T-cell leukemia

Melchior Lauten,[1] Anja Möricke,[2] Rita Beier,[3] Martin Zimmermann,[3] Martin Stanulla,[2] Barbara Meissner,[2] Edelgard Odenwald,[3] Andishe Attarbaschi,[4] Charlotte Niemeyer,[5] Felix Niggli,[6] Hansjörg Riehm,[3] and Martin Schrappe[2]

[1]Pediatric Hematology and Oncology, University Hospital Schleswig-Holstein, Lübeck Campus, Germany; [2]Department of Pediatrics, University Hospital Schleswig-Holstein, Kiel Campus, Germany; [3]Pediatric Hematology and Oncology, Hannover Medical School, Hannover, Germany; [4]St. Anna Children's Hospital, Vienna, Austria; [5]Pediatric Hematology and Oncology, University Hospital Freiburg, Germany; and [6]Kinderspital, University Hospital Zürich, Switzerland

ABSTRACT

Background
In the ALL-BFM 95 trial for treatment of acute lymphoblastic leukemia, response to a prednisone pre-phase (prednisone response) was used for risk stratification in combination with age and white blood cell count at diagnosis, response to induction therapy and specific genetic high-risk features.

Design and Methods
Cytomorphological marrow response was prospectively assessed on Day 15 during induction, and its prognostic value was analyzed in 1,431 patients treated on ALL-BFM 95.

Results
The 8-year probabilities of event-free survival were 86.1%, 74.5%, and 46.4% for patients with M1, M2, and M3 Day 15 marrows, respectively. Compared to prednisone response, Day 15 marrow response was superior in outcome prediction in precursor B-cell and T-cell leukemia with, however, a differential effect depending on blast lineage. Outcome was poor in T-cell leukemia patients with prednisone poor-response independent of Day 15 marrow response, whereas among patients with prednisone good-response different risk groups could be identified by Day 15 marrow response. In contrast, prednisone response lost prognostic significance in precursor B-cell leukemia when stratified by Day 15 marrow response. Age and white blood cell count retained their independent prognostic effect.

Conclusions
Selective addition of Day 15 marrow response to conventional stratification criteria applied on ALL-BFM 95 (currently in use in several countries as regular chemotherapy protocol for childhood acute lymphoblastic leukemia) may significantly improve risk-adapted treatment delivery. Even though cutting-edge trial risk stratification is meanwhile dominated by minimal residual disease evaluation, an improved conventional risk assessment, as presented here, could be of great importance to countries that lack the technical and/or financial resources associated with the application of minimal residual disease analysis.

Key words: pediatric acute lymphoblastic leukemia, early treatment response, clinical trial.

Citation: Lauten M, Möricke A, Beier R, Zimmermann M, Stanulla M, Meissner B, Odenwald E, Attarbaschi A, Niemeyer C, Niggli F, Riehm H, and Schrappe M. Prediction of outcome by early bone marrow response in childhood acute lymphoblastic leukemia treated in the ALL-BFM 95 trial: differential effects in precursor B-cell and T-cell leukemia. Haematologica 2012;97(7):1048-1056. doi:10.3324/haematol.2011.047613

ML and AM contributed equally to this manuscript.

Acknowledgments: we thank all patients and their families who participated in this trial, and the physicians, nurses and study nurses of all the hospitals involved for their input in performing this study. We thank N Götz, D Janousek, U Meyer, I Krämer and K Mischke for data management.

Funding: the clinical trial was supported by the Deutsche Krebshilfe, Bonn, Germany (50-2614-Ri 6; H.R.). This work was also supported by the Madeleine-Schickedanz-Kinderkrebsstiftung, Fürth, Germany, which provided a fellowship to ML.

Manuscript received on May 31, 2011. Revised version arrived on January 5, 2012. Manuscript accepted January 11, 2012.

Correspondence:
Anja Möricke, MD, University Hospital Schleswig-Holstein, Campus Kiel, Department of Pediatrics Schwanenweg 20, 24105 Kiel, Germany.
Phone: international +49.431.5974028.
Fax: international +49.431.5974034.
E-mail: a.moericke@pediatrics.uni-kiel.de

The online version of this article has a Supplementary Appendix.

Introduction

Early reduction of malignant cell load is known to be of major importance for the prediction of treatment outcome in solid and hematologic tumors.[1-4] In 1983, the Berlin-Frankfurt-Münster (BFM) study group started to evaluate the early treatment response to a prednisone pre-phase (prednisone response, PR) as a predictive factor for treatment outcome by measuring the peripheral blast count on Day 8 of treatment.[5] Since then, the PR has consistently been found to be one of the strongest independent prognostic factors for the prediction of treatment outcome in ALL-BFM studies.[6]

In the 1970s, the Children's Cancer Study Group (CCG) started to evaluate early bone marrow response during multi-agent induction treatment and demonstrated the predictive value of early marrow response in terms of remission achievement and ultimate outcome.[7-9] In the following trials, the CCG generated many data on the prognostic importance of marrow response on Days 7 and 14, the combined impact of the two evaluation points, and the differential effect in patients at standard or high risk according to the NCI/Rome criteria.[10-13] Based on these results, early marrow response has become an integral part of risk stratification in the successive CCG and contemporary Children's Oncology Group (COG) ALL treatment regimens.[14-19] The St. Jude Total Therapy Study Group showed that even the persistence of low percentages (1-4%) of BM lymphoblasts on Day 15 (corresponding to Day 22 of the BFM protocol without prednisone prephase) and Days 22 to 25 of induction therapy was associated with a significantly poorer 5-year event free survival rate compared to patients without detectable BM blasts (40±6% *vs.* 78±2%).[20]

In the ALL-BFM 95 trial, cytomorphological response in BM on Day 15 (BMd15) was prospectively assessed without being used for risk stratification. In the present study, the prognostic value of BMd15 in ALL-BFM 95 was evaluated in comparison and combined with PR, cytomorphological BM response to induction therapy (Day 33), age and white blood cell count (WBC) at diagnosis; all factors included in the ALL-BFM 95 risk stratification. The aim was to refine the risk criteria used in ALL-BFM 95 without using modern minimal residual disease (MRD) techniques that might be not available in less affluent countries because of cost.

Design and Methods

Patients

From the 2,169 patients eligible for the ALL-BFM 95 study, 1,431 patients had assessable information on BM morphology on Day 15. These patients were included in the current study. Informed consent was obtained from the parents or guardians of each patient. Data were managed in the ALL-BFM study center. The trial was approved by the ethics committee of the Hanover Medical School, Germany. Treatment regimen and outcome of the ALL-BFM 95 trial have been recently described.[21]

Response and relapse criteria

PR was defined by the absolute number of leukemic blasts/μL in the peripheral blood after seven days of prednisone treatment and one intrathecal (IT) dose of methotrexate, regardless of the initial leukemic blast count. Prednisone good responders (PGR) were characterized by less than 1,000 blasts/μL, whereas prednisone poor responders (PPR) showed 1,000 blasts/μL or more on Day 8 of treatment.[5] Response in BM was evaluated on Days 15 and 33 of induction treatment and was categorized as M1 (<5%), M2 (5 to <25%), and M3 (≥25% lymphoblasts). Complete remission (CR) was defined as M1 BM on Day 33 of induction therapy, the absence of leukemic blasts in blood and CSF, and no evidence of local disease. Relapse was defined as recurrence of 25% lymphoblasts or over in BM or local leukemic infiltrates at any site. Both PR and BM evaluation were reviewed centrally in two reference laboratories.

Risk stratification

Patients were stratified into three risk groups according to the following criteria:

HR: PPR, and/or no CR on Day 33, and/or evidence of t(9;22) (or BCR/ABL), and/or evidence of t(4;11) (or MLL/AF4);

MR: No HR criteria, and initial WBC ≥20×10^9/L and/or age at diagnosis <1 or ≥6 years, and/or T-ALL;

SR: No HR criteria, and initial WBC <20×10^9/L and age at diagnosis ≥1 and <6 years, and no T-ALL.

CNS status was not a stratification criterion.

Statistics

Event-free survival was defined as the time from diagnosis to the date of last follow up in complete remission or first event. Events were resistance to therapy (non-response), relapse, secondary malignant neoplasm (SMN) or death from any cause. Failure to achieve remission due to early death or non-response was considered as event at time zero. Patients lost to follow up were censored at the time of their withdrawal. The Kaplan-Meier method[22] was used to estimate survival rates; differences were compared with the two-sided log rank test.[23] Differences in the distribution of individual parameters among patient subsets were analyzed using the χ^2 test for categorized variables. All P values were two-sided and $P<0.05$ was considered statistically significant. Cox's proportional hazards model was used to obtain the estimates and the 95% confidence interval of the relative risk for prognostic factors.[24] The results of the ALL-BFM 95 trial were updated in August 2008.

Statistical analyses were performed using the SAS statistical program (SAS-PC, Version 9.1, SAS Institute Inc., Cary, NC, USA, and IBM SPSS statistics, version 15).

Results

BM puncture on Day 15 of induction therapy Protocol I was performed in 1,696 (78%) of the 2,169 patients of the ALL-BFM 95 trial. In 1,431 (84%) of the 1,696 BM punctures the BM smears were eligible for evaluation and could be included in the present study. Characteristics of these patients and of the patients who could not be included due to missing BMd15 data are shown in the *Online Supplementary Table S1*. Patients who were not included due to non-representative BMd15 had a lower rate of PPR, presented less often with hyperleukocytosis, were less often BCR/ABL positive and included a lower rate of high-risk patients compared to those patients included in the study. However, the rate of complete remission on Day 33 was higher in patients not included in the study (*Online Supplementary Table S1*).

The estimated probability of 8-year EFS (8y-pEFS) of all patients included in this study was 78.8±0.9%.

PR was evaluable in 1,419 of the 1,431 (99%) patients

analyzed; 1,280 (90%) patients showed PGR, 139 (10%) patients were defined as PPR. The 8y-pEFS was 81.3±0.9% for patients with PGR and 55.1±3.7% for patients with PPR (P<0.01). BMd15 characterized three distinct risk groups. The 8y-pEFS of these groups was 86±1%, 74±2%, and 46±4% for the patients with M1, M2 and M3 marrow, respectively (Table 1). BM response on Day 33 (BMd33) could be assessed in 1,415 of 1,431 patients. Only 42 of these patients did not achieve BM remission on Day 33 (NRd33) and had an 8y-pEFS of 36.3±6.9%. Of these patients, 38 (90%) had an M3 and 4 patients an M2 BMd15. Among all patients with M3 BMd15, 8y-pEFS was 52.5±4.2% (n=146) for those patients who achieved complete cytomorphological remission (CR) by Day 33 and 25.4±7.2% (n=38) for those who did not (P<0.001).

The results of the BMd15 subgroups according to various patients' characteristics are presented in Table 1. Poor response in BMd15 was significantly associated with T-ALL (P<0.001) and the known high-risk features were adolescent age (P<0.001), hyperleukocytosis (P<0.001), BCR/ABL (P=0.003), CNS involvement (P=0.005), PPR (P<0.001), and NRd33 (P<0.001). A significant prediction of prognosis by BMd15 could be seen for all subgroups analyzed except for the small subgroup of patients with initial CNS involvement.

The cut-off values characterizing M1, M2 and M3 for the distinction of BMd15 subgroups are internationally recognized. However, each subgroup includes patients with a wide range of BM blasts. Therefore, in addition to the traditional M1, M2 and M3 categories, we analyzed patients within narrower ranges of blasts. Results are shown in Figure 1 and indicate that the steady increase in BMd15 blasts proceeds parallel to a steady decrease in 8y-

Table 1. Patient's characteristics and treatment outcome according to cytomorphological response in bone marrow on day 15.

	Bone marrow day 15						
	M1		M2		M3		
Variable	N[1] (%)	8y-pEFS, % (SE)	N[1] (%)	8y-pEFS, % (SE)	N[1] (%)	8y-pEFS, % (SE)	P[2]
All	880 (100)	86.1 (1.2)	365 (100)	74.5 (2.3)	186 (100)	46.4 (3.7)	<0.001
Gender							
Male	502 (57.0)	86.1 (1.6)	212 (58.1)	74.2 (3.1)	107 (57.5)	44.6 (4.9)	<0.001
Female	378 (43.0)	86.1 (1.8)	153 (41.9)	74.9 (3.5)	79 (42.5)	48.8 (5.7)	<0.001
Age at diagnosis (years)							
<1[3]	9 (1.0)	66.7 (15.7)	11 (3.0)	27.2 (13.7)	6 (3.2)	16.7 (15.2)	0.024
1-<6	519 (59.0)	89.3 (1.4)	214 (58.6)	80.9 (2.7)	75 (40.3)	54.2 (5.8)	<0.001
6-<10	183 (20.8)	86.3 (2.6)	72 (19.7)	81.8 (4.6)	39 (21.0)	48.3 (8.1)	<0.001
≥10	169 (19.2)	75.7 (3.7)	68 (18.6)	50.9 (7.3)	66 (35.5)	40.2 (6.1)	<0.001
Initial WBC (×10^9/L)							
<20	603 (68.5)	87.1 (1.4)	217 (59.5)	76.7 (2.9)	72 (38.7)	46.6 (6.2)	<0.001
20-<50	128 (14.5)	90.1 (2.7)	70 (19.2)	70.9 (5.5)	38 (20.4)	54.8 (8.1)	<0.001
50-<100	70 (8.0)	80.7 (4.8)	35 (9.6)	77.0 (7.1)	21 (11.3)	42.9 (10.8)	<0.001
≥100	79 (9.0)	70.9 (5.1)	43 (11.8)	67.3 (7.2)	55 (29.6)	41.3 (6.7)	<0.001
Immunophenotype							
T-ALL	107 (12.6)	86.9 (3.3)	40 (11.2)	78.8 (6.7)	47 (25.4)	45.0 (7.7)	<0.001
pB-ALL	741 (87.4)	86.5 (1.3)	317 (88.8)	74.2 (2.5)	138 (74.6)	47.1 (4.3)	<0.001
CNS involvement							
Negative	811 (92.5)	87.4 (1.2)	334 (92.0)	75.1 (2.4)	155 (83.8)	48.7 (4.1)	<0.001
TLP+[4]	45 (5.1)	75.4 (6.5)	20 (5.5)	68.8 (10.7)	20 (10.8)	40.0 (11.0)	0.021
Positive	21 (2.4)	61.9 (10.6)	9 (2.5)	55.6 (16.6)	10 (5.4)	30.0 (14.5)	0.296
TEL/AML1							
Positive	104 (23.7)	90.7 (3.0)	37 (21.0)	94.6 (3.7)	15 (15.2)	69.6 (12.7)	0.030
Negative	335 (76.3)	84.5 (2.0)	139 (79.0)	65.8 (4.1)	84 (84.8)	37.5 (5.3)	<0.001
BCR/ABL							
Positive	14 (1.8)	42.9 (13.2)	9 (2.6)	11.1 (10.5)	11 (6.4)	9.1 (8.7)	0.044
Negative	768 (98.2)	86.6 (1.3)	332 (97.4)	76.5 (2.4)	162 (93.6)	50.0 (4.0)	<0.001
Risk group (ALL-BFM 95)[5]							
Standard	342 (38.9)	91.6 (1.5)	114 (31.2)	84.6 (3.5)	18 (9.7)	66.7 (11.1)	<0.001
Intermediate	494 (56.1)	84.3 (1.7)	192 (52.6)	73.3 (3.3)	74 (39.8)	52.5 (6.0)	<0.001
High	44 (5.0)	63.6 (7.3)	59 (16.2)	58.8 (6.5)	94 (50.5)	38.0 (5.0)	0.001
Prednisone response							
Good	847 (97.1)	86.4 (1.2)	321 (88.4)	75.1 (2.5)	112 (60.9)	50.9 (4.8)	<0.001
Poor	25 (2.9)	76.0 (8.5)	42 (11.6)	68.3 (7.3)	72 (39.1)	43.1 (5.8)	0.001
Remission Day 33							
No	-	-	4 (1.1)	100 (0.0)	38 (20.7)	25.4 (7.2)	0.013
Yes	871 (100)	86.3 (1.2)	356 (98.9)	74.7 (2.3)	146 (79.3)	52.5 (4.2)	<0.001

[1]Data given refer to patients with successful investigation of the respective criterion. [2]The P value (log rank test) refers to the comparison of BMd15 groups within the subgroups of presented patient's characteristics. [3]Patients treated in the Interfant-99 pilot study were excluded. [4]TLP+ indicates traumatic lumbar puncture with evidence of leukemic blasts in cerebral spinal fluid. [5]According to the risk criteria of trial ALL-BFM 95.

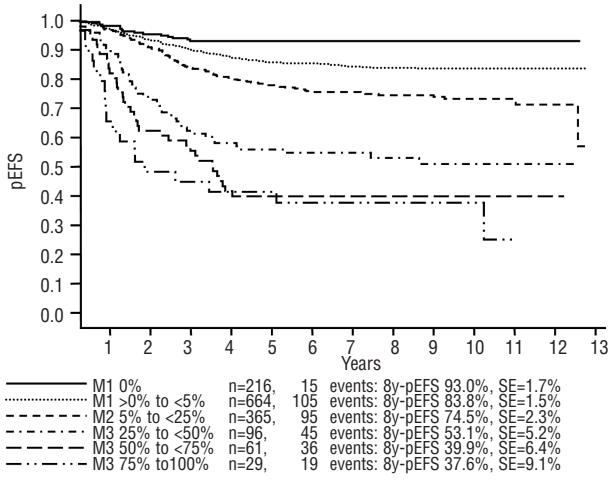

Figure 1. **Kaplan-Meier estimate of event-free survival according to percentage of blast counts in the bone marrow on Day 15. Log rank test (pair-wise comparisons): M1 (0%) *vs.* M1 (>0-<5%) *P*=0.001; M3 (25-<50%) *vs.* M3 (50-<75%) *P*=0.102; M3 (25-<50%) *vs.* M3 (75-100%) *P*=0.025; M3 (50-<75%) *vs.* M3 (75-100%) *P*=0.381. 8y-pEFS indicates probability of event-free survival at 8 years; SE: standard error.**

pEFS. Interestingly, there is a clear distinction in pEFS between 0% and over 0% to less than 5% (M1 category), between 25% to less than 50% and 50% or over BM blasts on Day 15 (M3 category) (Figure 1). Of those patients with 75% blasts or more on Day 15, 55% nonetheless reached CR on Day 33 and had an 8y-pEFS of 42.9±12.6%. Among those patients with 25% to less than 50% blasts on Day 15, 92% reached CR on Day 33 with an 8y-pEFS of 57.5±5.5%.

Detailed analyses revealed marked differences within immunophenotypic subgroups. Results are, therefore, shown for each subgroup separately.

pB-ALL

In pB-ALL (n=1,196; 1,187 patients with evaluable PR), patients with PGR had an 8y-pEFS of 80.2±1.2% (n=1,109) and PPR patients an 8y-pEFS of 58.4±5.6% (n=78) (*P*<0.001). Patients in the BMd15 subgroups had an 8y-pEFS of 86.5±1.3% (n=741), 74.2±2.5% (n=317), and 47.1±4.3% (n=138) for the M1, M2 and M3 groups, respectively. Though a smaller group of pB-ALL patients could be identified by PPR compared to M3 BMd15 (6.6% *vs*. 11.5%), the EFS of the patients in the M3 BMd15 group was even worse than the EFS of PPR patients, showing the better prognostic discriminative value of M3 BMd15. This was also reflected in the distribution of events: 12.7% of all events in pB-ALL (n=32 of 256) clustered in the PPR group, whereas 28.5% were detected in the BMd15 M3 group (n=72 of 256) (Figure 2A and B). Sensitivity of PR to predict poor BM response on Day 15 or Day 33 was low as only 27.9% of patients with M3 BMd15 and 56.7% of patients with NRd33 had shown PPR before. BMd15 allowed a clear separation of three different risk groups for patients with M1, M2 and M3 marrow within the subgroups of PGR and PPR patients (Figure 2A and B). There was no statistical difference in pEFS between patients in the same BMd15 subgroup when analyzed according to PR (Table 2).

Age and WBC as well as NCI risk criteria[25,26] and risk group criteria of the ALL-BFM 95 trial (both using age at diagnosis and initial WBC) showed an additional prognostic value when analyzed in combination with BMd15 (Table 2).

Univariate results were confirmed by a multivariate Cox's regression analysis including NCI risk criteria, PR, BMd15 and BMd33 as covariates. In this analysis, PR lost its prognostic significance whereas the NCI risk criteria, as well as BM response on Days 15 and 33, retained significance (Table 3).

T-ALL

In T-ALL (n=194; 191 patients with evaluable PR), PGR patients had an 8y-pEFS of 84.6±3.3% (n=130) and patients with PPR had an 8y-pEFS of 54.1±6.4% (n=61). The 8y-pEFS of patients with M1, M2 and M3 BMd15 was 86.9±3.3% (n=107), 78.8±6.7% (n=40), and 45.0±7.7% (n=47), respectively. Sensitivity of PR to predict poor BM response on Day 15 or Day 33 was better in T-ALL than in pB-ALL: 72.3% of the patients with M3 BMd15 and 81.8% of the patients with NRd33 had shown PPR before.

Among the patients with PPR, BMd15 was not able to characterize subgroups with significantly different outcomes (Figure 2D). In PGR, however, outcome of patients with BMd15 M3 was significantly worse (M3, 8y-pEFS 43.1±14.7%) than the M1 and M2 subgroup with similarly favorable results (M1: 8y-pEFS 91.1±3.0%; M2: 8y-pEFS 83.4±7.7%) (Figure 2C). The prognostic relevance of the PR within the BMd15 subgroups in T-ALL is illustrated by the reverse analysis in Table 2. Within the BMd15 M1 subgroup, patients with significantly worse pEFS could be identified through PPR whereas no difference in outcome was shown within the BMd15 M3 subgroup. Within the small BMd15 M2 subgroup, the difference between PGR and PPR did not reach statistical significance. Thus, by combining PR and BMd15, T-ALL patients can be stratified into two distinct risk groups: one including the patients with PGR plus M1 or M2 BMd15 (n=120, 8y-pEFS 89.5±2.9%) the other including all patients with PPR and/or M3 BMd15 (n=74, 8y-pEFS 52.1±5.9%) (*P*<0.001) (Figure 3).

NCI risk criteria had a borderline significant prognostic value in patients with M1 BMd15 but showed no statistic significance in patients with M2 or M3 BMd15 (Table 2).

Consistent with these results, multivariate Cox's regression analysis including NCI risk criteria, PR, BMd15 and BMd33 as covariates revealed BMd15 M3 as the strongest independent adverse risk factor, and also marginal significance for PPR and NCI-HR (Table 3).

Discussion

For more than 20 years, cytomorphological response has been the leading criterion for stratifying patients into risk groups within the ALL-BFM trials. Since the ALL-BFM 86 trial, cytomorphological response has been estimated very early during induction treatment using the PR as criterion for risk stratification. Cytomorphological treatment response in the BM, however, was evaluated only at the end of induction treatment (Day 33). Poor cytomorphological response at either response evaluation point qualified a patient for high-risk treatment.[6,27,28]

The prognostic significance of early reduction of

leukemic blasts in BM at different time points during induction treatment was shown in a number of pediatric ALL trials[10] and was implemented as risk stratification criterion in various international trials.[12,13,29-34] Specificity of response evaluation might, nevertheless, vary depending on the time of response evaluation with regard to the therapy and the composition of the treatment.[13,35,36]

With the aim of prospectively assessing the prognostic value of an early cytomorphological response evaluation in the BM, a BM puncture on Day 15 of induction treatment was performed in addition to the evaluation of PR and BMd33 in ALL-BFM 95.[21] However, whereas the very easy sampling and evaluation of the peripheral blood samples on Day 8 provided assessable PR samples for nearly all patients, 15.6% of the BM aspirates on Day 15 could not be assessed due to non-representative BM morphology.

Overall, the prediction of treatment outcome was possible with each of the three response parameters PR, BMd15 or BMd33. BMd15 allowed a better prediction of outcome than PR in pB-ALL as well as T-ALL but the additional prognostic value of PR depended on the immunophenotype. In pB-ALL, BMd15 could identify three distinct risk groups, and the PR had no significant additional effect in patients stratified by BMd15. Biologically, this seems highly plausible considering the fact that the PR is measured after the administration of seven days of prednisone and one IT dose of MTX, while the evaluation of the BM on Day 15 reflects the response to 14 days of prednisone, one dose of vincristine, daunorubicin and asparaginase, respectively, and two doses of IT MTX. This might also indicate that, in pB-ALL, resistance to prednisone can be compensated by high sensitivity to other chemotherapeutic drugs and that high sensitivity to prednisone can be overridden by resistance to other agents. These biological considerations seem to be less applicable for T-ALL patients. Our data indicate that, in the end, resistance to prednisone (i.e. PPR) in T-ALL could not be overcome by the subsequent chemotherapy even in those patients who apparently had a reasonable response in the later course of induction treatment as reflected by M1 or M2 BMd15. The reliability of these data might be weakened due to the small patient numbers remaining in the T-ALL subgroups in this analysis. However, the results are supported by the recently

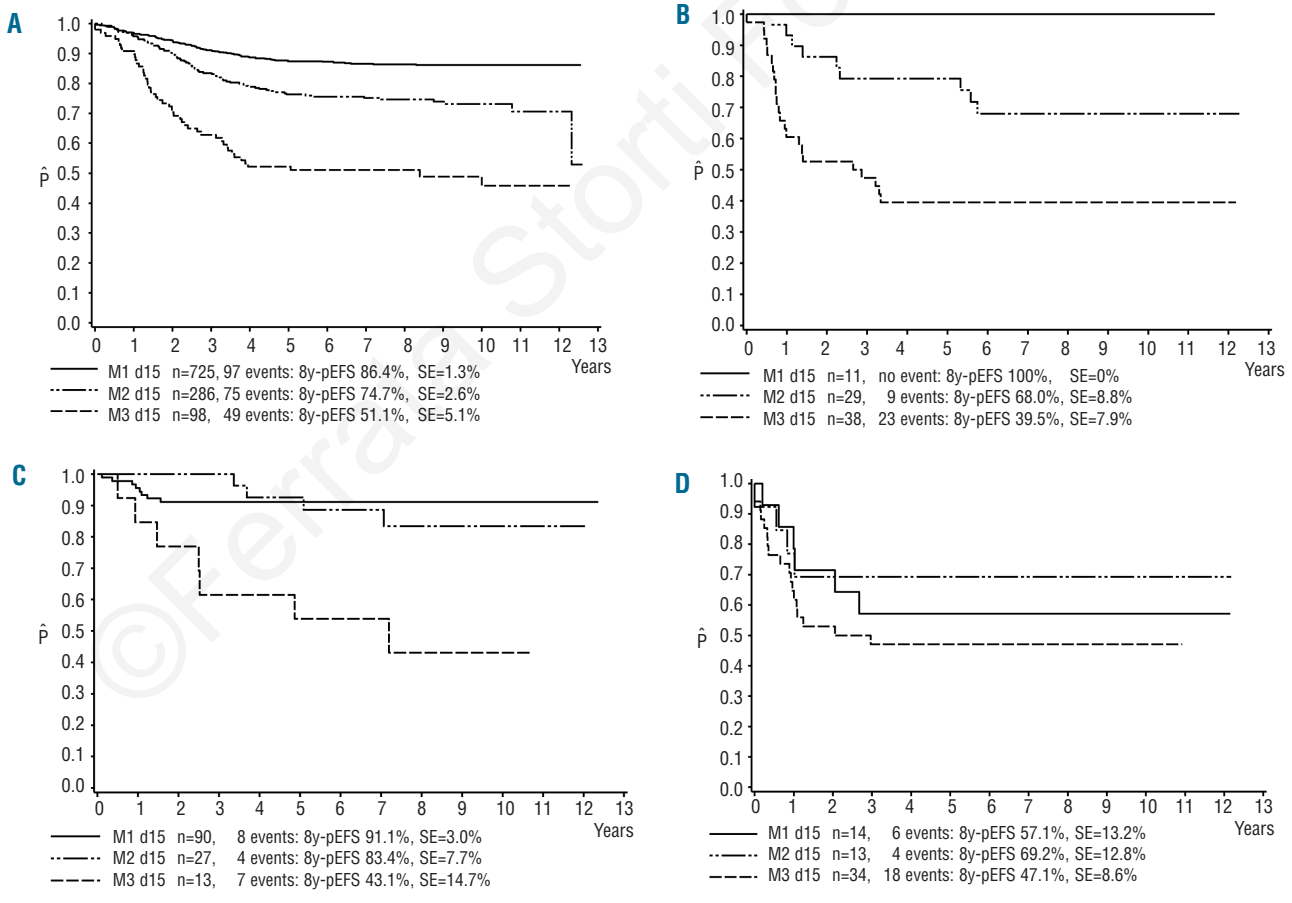

Figure 2. Kaplan-Meier estimate of event-free survival according to bone marrow response on Day 15 stratified by immunophenotypic lineage and prednisone response. For definition of bone marrow M1, M2, and M3 see the Design and Methods section. (A) pB-ALL, prednisone good-response; log rank test (pair-wise comparisons): all *P* values <0.001. (B) pB-ALL, prednisone poor-response; log rank test: M1 *vs.* M2 *P*=0.042; M2 *vs.* M3 *P*=0.007; M1 *vs.* M3 *P*=0.002; (C) T-ALL, prednisone good-response; log rank test: M1 *vs.* M2 *P*=0.44; M2 *vs.* M3 *P*=0.005; M1 *vs.* M3 *P*<0.001; (D) T-ALL, prednisone poor-response; log rank test: M1 *vs.* M2 *P*=0.65; M2 *vs.* M3 *P*=0.22; M1 *vs.* M3 *P*=0.42. 8y-pEFS indicates probability of event-free survival at 8 years; SE, standard error.

published data on the prognostic impact of MRD in the AIEOP-BFM ALL 2000 trial.[37] In this study, the PR in T-ALL also retained prognostic value (although this had only borderline significance) when analyzed in a multivariate model including the MRD risk groups. In contrast, in pB-ALL, PPR completely lost its adverse prognostic value if compared with PGR patients with the same PCR-MRD levels.[38]

Our data may suggest that the PR could be omitted as a stratification parameter for patients with pB-ALL. However, in the ALL-BFM 95 trial, the good outcome of pB-ALL patients with PPR and subsequently good BM response on Day 15 was achieved with an intensified high-risk treatment. Whether these results could be reproduced with less intensive treatment, remains unclear. Therefore, we think that the omission of the PR as a risk stratification parameter should not be considered for the moment.

In clinical practice, the question often arises as to whether an early change or intensification of treatment is

Table 2. Treatment outcome in pB-ALL and T-ALL by age and WBC at diagnosis, different risk group classifications and treatment response.

	Bone marrow Day 15								
	M1			M2			M3		
Variable	N[1] (%)	8y-pEFS, % (SE)	P[1]	N (%)	8y-pEFS, % (SE)	P[1]	N (%)	8y-pEFS, % (SE)	P[1]
pB-ALL									
Age (years)									
<1[5]	8 (1.1)	62.5 (17.1)		10 (3.2)	20.0 (12.6)		6 (4.3)	16.7 (15.2)	
1-<10	312 (82.6)	88.8 (1.3)	<0.001[4]	253 (79.8)	81.8 (2.5)	<0.001[4]	89 (64.5)	51.3 (5.3)	0.133[4]
≥10	121 (16.3)	75.7 (4.3)		54 (17.0)	46.3 (7.9)		43 (31.2)	43.6 (7.7)	
Initial WBC (×10⁹/L)									
<50	648 (87.4)	88.2 (1.3)	<0.001	262 (82.6)	75.7 (2.7)	0.160	92 (66.7)	53.7 (5.3)	0.015
≥50	93 (12.6)	74.8 (4.5)		55 (17.4)	67.2 (6.3)		46 (33.3)	33.9 (7.1)	
Risk group (ALL-BFM 95)[2]									
Standard	327 (44.1)	91.8 (1.5)	<0.001	109 (34.4)	84.8 (3.5)	<0.001	17 (12.3)	70.6 (11.1)	0.002
Intermediate	384 (51.8)	83.7 (1.9)		162 (51.1)	72.4 (3.6)		63 (45.7)	52.9 (6.4)	
High	30 (4.0)	66.7 (8.6)		46 (14.5)	55.9 (7.4)		58 (42.0)	33.9 (6.3)	
NCI/Rome risk group[3]									
Standard	537 (73.3)	90.6 (1.3)	<0.001	213 (69.4)	81.7 (2.7)	<0.001	59 (44.7)	58.8 (6.5)	0.016
High	196 (26.7)	76.0 (3.2)		94 (30.6)	62.8 (5.2)		73 (55.3)	40.7 (5.8)	
Prednisone response									
Good	725 (98.5)	86.4 (1.3)	0.20	286 (90.8)	74.7 (2.6)	0.51	98 (72.1)	51.1 (5.1)	0.088
Poor	11 (1.5)	100 (0.0)		29 (9.2)	68.0 (8.8)		38 (27.9)	39.5 (7.9)	
Remission Day 33									
No	0 (0.0)	-	-	3 (1.0)	-	-	28 (20.6)	23.6 (8.2)	<0.001
Yes	733 (100)	86.8 (1.3)		310 (99.0)	74.3 (2.5)		108 (79.4)	54.0 (4.8)	
T-ALL									
Age (years)									
<1[5]	0 (0.0)	-		1 (2.5)	-		0 (0.0)	-	
1-<10	67 (62.6)	88.1 (4.0)	0.65	26 (65.0)	80.6 (7.8)	0.76	24 (51.1)	56.8 (10.5)	0.126
≥10	40 (37.4)	85.0 (5.6)		13 (32.5)	67.7 (17.1)		23 (48.9)	34.2 (10.0)	
Initial WBC (×10⁹/L)									
<50	55 (51.4)	92.7 (3.5)	0.069	18 (45.0)	75.7 (10.7)	0.89	17 (36.2)	20.2 (15.5)	0.46
≥50	52 (48.6)	80.7 (5.5)		22 (55.0)	81.8 (8.2)		30 (63.8)	53.3 (9.1)	
Risk group (ALL-BFM 95)[2]									
Standard	-	-		-	-		-	-	
Intermediate	93 (86.9)	91.4 (2.9)	<0.001	27 (67.5)	83.4 (7.7)	0.168	11 (23.4)	50.9 (16.3)	0.26
High	14 (13.1)	57.1 (13.2)		13 (32.5)	69.2 (12.8)		36 (76.6)	44.4 (8.3)	
NCI/Rome risk group[3]									
Standard	32 (29.9)	96.6 (3.1)	0.054	12 (30.8)	82.5 (11.3)	0.62	6 (12.8)	41.7 (30.4)	0.26
High	75 (70.1)	82.6 (4.4)		27 (69.2)	76.0 (8.7)		41 (87.2)	43.8 (7.8)	
Prednisone response									
Good	90 (86.5)	91.1 (3.0)	<0.001	27 (67.5)	83.4 (7.7)	0.168	13 (27.7)	43.1 (14.7)	0.58
Poor	14 (13.5)	57.1 (13.2)		13 (32.5)	69.2 (12.8)		34 (72.3)	47.1 (8.6)	
Remission Day 33									
No	-	-	-	1 (2.6)	-	-	10 (21.3)	30.0 (14.5)	0.129
Yes	106 (100)	86.8 (3.3)		38 (97.4)	80.3 (6.8)		37 (78.7)	49.3 (8.7)	

[1]The P value (log rank test) refers to comparison within BMd15 subgroups. [2]Risk groups according to the risk criteria of ALL-BFM 95 trial. [3]NCI/Rome standard risk, age one year or older and less than ten years and WBC less than 50x10⁹/L. NCI/Rome high risk, age ten years or older or WBC 50x10⁹/L or higher; infants les than one year were excluded from the NCI definition; [4]The P value (log rank test) refers to comparison of the age groups one to less than ten years vs. ten years and older. [5]Patients treated in the Interfant-99 pilot study were excluded.

reasonable in patients with poor early response. Our data show that patients with M3 marrow on Day 15 still have a good chance of achieving CR by end of induction. Among those patients with 25% to less than 50% BM blasts on Day 15, 92% achieved CR (71 of 77 pB-ALL and 15 of 17 T-ALL patients). Even a fraction of patients with 75% or over BM blasts on Day 15 reached CR by this time point (8 of 17 pB-ALL patients and 8 of 12 T-ALL patients). This suggests that if the aim is just to achieve remission there is no strong evidence for the need for alternative ALL treatment at this point. However, treatment results of these patients are poor and might be improved by early treatment intensification.

In the stratification of T-ALL patients, combining BMd15 with PR added significant value to the response parameters alone and allowed stratification into two widely separated risk groups, the better of them with an excellent 8y-pEFS of almost 90% and another poor risk group with an 8y-pEFS of nearly 50%, including 74% of all T-ALL events.

In pB-ALL, in contrast, the use of the PR in addition to BMd15 failed to improve the discrimination between risk groups obtained through BMd15 alone. Yet the combination of BMd15 with the ALL-BFM 95 risk criteria or the NCI criteria, both using age and initial WBC, gave an added prognostic value. In the COG (or former CCG) protocols, the combination of NCI risk criteria with early (Day 7 and Day 14) marrow response has been used for risk stratification for many years.[12,39] The ALL IC-BFM study group introduced cytomorphological BM response on Day 15 in the non-MRD-based protocol ALL IC-BFM 2002 for a risk stratification system which was based on the ALL-BFM 95 criteria, but shifted the patients to a higher risk group in the case of an M3 BMd15.[40] ALL IC-BFM 2002 was performed in countries which did not have access to MRD diagnostics, mainly due to economic concerns.[40] For these countries, optimization of risk stratification by the intelligent use of clinical parameters and cytomorphological response evaluation is worthwhile. However, the prognostic relevance of cytomorphological response must always be interpreted in the context of the specific chemotherapy regimen administered. Therefore, we should approach transferring our data onto other and, in particular, less intensive treatment regimens with caution.

Results of the ALL IC-BFM 2002 study have not yet been published. It will be interesting to see whether the results of the current study, which were generated in a setting with centralized cytomorphology services and, therefore, a high level of staff continuity, can be reproduced in a setting with decentralized cytomorphology services.

Polymerase chain reaction[41] and flow cytometry[42] have been shown to detect MRD and help discriminate between patients with a differential response at later time points when patients have already reached morphological remission. Recently, the AIEOP-BFM group published data on a total of 3,648 ALL patients (pB-ALL: n=3184; T-ALL: n=464).[37,38] In these studies, the 5y-pEFS of patients already MRD negative at end of induction (MRD-SR) were 92.3±0.9% (pB-ALL) and 93.0±3.0% (T-ALL), respectively. Considering together all pB-ALL and T-ALL patients, this group made up 39.0% of all patients. An equally good pEFS of 93.0±1.7% (8y-pEFS) was achieved in our study in those patients with 0% blasts in the BM on Day 15. However, this group made up only 15.1% of the study population, showing that the PCR-MRD technique is able to allocate more patients for less intensive treatment.

To summarize, in the context of the ALL-BFM 95 treatment, in pB-ALL the combination of BM response on Day

Table 3. Multivariate Cox's regression analysis of cytomorphological response and NCI risk criteria, shown separately for pB-ALL and T-ALL.

Variable	N*	RR	95%-Confidence interval	*P* (Wald)
pB-ALL				
NCI/Rome risk group				
Standard	796	1		
High	353	2.17	1.65-2.85	<0.001
Prednisone response				
Good	1079	1		
Poor	70	0.92	0.59-1.43	0.719
Bone marrow on Day 15				
M1	720	1		
M2	301	1.96	1.44-2.67	<0.001
M3	128	3.93	2.74-5.64	<0.001
Remission on Day 33				
Yes	1122	1		
No	27	2.14	1.25-3.67	0.006
T-ALL				
NCI/Rome risk group				
Standard	47	1		
High	141	2.25	0.87-5.80	0.092
Prednisone response				
Good	129	1		
Poor	59	1.99	0.97-4.06	0.059
Bone marrow on Day 15				
M1	103	1		
M2	38	1.17	0.47-2.94	0.738
M3	47	2.93	1.32-6.47	0.008
Remission on Day 33				
Yes	177	1		
No	11	1.63	0.69-3.85	0.264

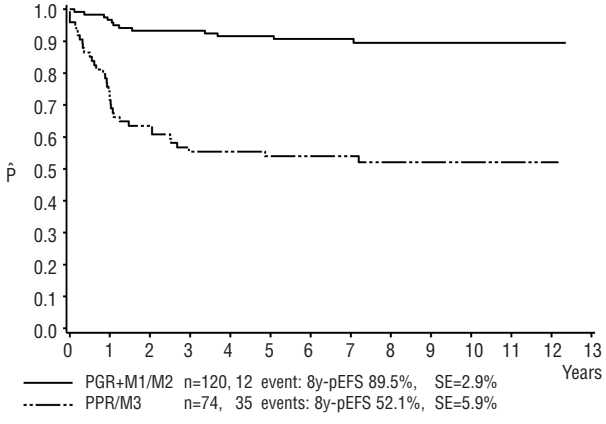

Figure 3. Kaplan-Meier estimate of event-free survival in T-ALL comparing patients with prednisone good-response plus M1 or M2 marrow on Day 15 (PGR+M1/M2) with patients with prednisone poor-response and/or M3 marrow on Day 15 (PPR/M3). Log rank test *P*<0.001. 8y-pEFS indicates probability of event-free survival at 8 years; SE, standard error.

15 with the ALL-BFM 95 risk criteria allows a more subtle definition of risk groups. The PR, included as high-risk stratification criterion in ALL-BFM 95, completely lost its significance in combination with BMd15. In T-ALL, BMd15 was also a better predictor of outcome than PR, though within the subgroup of patients with M1 (and possibly also M2) BMd15, the PR added an important prognostic effect.

Today, the ALL-BFM 95 protocol is regularly used as a chemotherapy protocol for childhood ALL in several countries. Our data demonstrate that the inclusion of BMd15 crucially improves the ALL-BFM 95 risk stratification in the context of the ALL-BFM 95 therapy. This is of particular interest in less affluent countries where limited economic resources mean expensive laboratory techniques cannot be used.

Authorship and Disclosures

The information provided by the authors about contributions from persons listed as authors and in acknowledgments is available with the full text of this paper at www.haematologica.org.

Financial and other disclosures provided by the authors using the ICMJE (www.icmje.org) Uniform Format for Disclosure of Competing Interests are also available at www.haematologica.org.

References

1. Felice MS, Zubizarreta PA, Alfaro EM, Sackmann-Muriel F. Childhood acute lymphoblastic leukemia: prognostic value of initial peripheral blast count in good responders to prednisone. J Pediatr Hematol Oncol. 2001;23(7):411-5.
2. Haw R, Sawka CA, Franssen E, Berinstein NL. Significance of a partial or slow response to front-line chemotherapy in the management of intermediate-grade or high-grade non-Hodgkin's lymphoma: a literature review. J Clin Oncol. 1994;12(5): 1074-84.
3. Keating MJ, Smith TL, Gehan EA, McCredie KB, Bodey GP, Spitzer G, et al. Factors related to length of complete remission in adult acute leukemia. Cancer. 1980;45(8):2017-29.
4. Lauten M, Stanulla M, Zimmermann M, Welte K, Riehm H, Schrappe M. Clinical outcome of patients with childhood acute lymphoblastic leukaemia and an initial leukaemic blood blast count of less than 1000 per microliter. Klin Padiatr. 2001; 213(4):169-74.
5. Riehm H, Reiter A, Schrappe M, Berthold F, Dopfer R, Gerein V, et al. [Corticosteroid-dependent reduction of leukocyte count in blood as a prognostic factor in acute lymphoblastic leukemia in childhood (therapy study ALL-BFM 83)]. Klin Padiatr 1987;199 (3):151-60.
6. Schrappe M, Reiter A, Zimmermann M, Harbott J, Ludwig WD, Henze G, et al. Long-term results of four consecutive trials in childhood ALL performed by the ALL-BFM study group from 1981 to 1995. Berlin-Frankfurt-Munster. Leukemia. 2000; 14(12):2205-22.
7. Gaynon PS, Bleyer WA, Steinherz PG, Finklestein JZ, Littman P, Miller DR, et al. Day 7 marrow response and outcome for children with acute lymphoblastic leukemia and unfavorable presenting features. Med Pediatr Oncol. 1990;18(4):273-9.
8. Miller DR, Leikin S, Albo V, Vitale L, Sather H, Coccia P, et al. Use of prognostic factors in improving the design and efficiency of clinical trials in childhood leukemia: Children's Cancer Study Group Report. Cancer Treat Rep. 1980;64(2-3):381-92.
9. Miller DR, Leikin S, Albo V, Sather H, Karon M, Hammond D. Prognostic factors and therapy in acute lymphoblastic leukemia of childhood: CCG-141. A report from childrens cancer study group. Cancer. 1983;51(6):1041-9.
10. Gaynon PS, Desai AA, Bostrom BC, Hutchinson RJ, Lange BJ, Nachman JB, et al. Early response to therapy and outcome in childhood acute lymphoblastic leukemia: a review. Cancer. 1997;80(9):1717-26.
11. Gaynon PS. Prognostic factors in acute lymphoblastic leukemia. J Pediatr Hematol Oncol. 2000;22(5):403-4.
12. Schultz KR, Pullen DJ, Sather HN, Shuster JJ, Devidas M, Borowitz MJ, et al. Risk- and response-based classification of childhood B-precursor acute lymphoblastic leukemia: a combined analysis of prognostic markers from the Pediatric Oncology Group (POG) and Children's Cancer Group (CCG). Blood. 2007;109(3):926-35.
13. Steinherz PG, Gaynon PS, Breneman JC, Cherlow JM, Grossman NJ, Kersey JH, et al. Cytoreduction and prognosis in acute lymphoblastic leukemia - the importance of early marrow response: report from the Childrens Cancer Group. J Clin Oncol. 1996;14(2):389-98.
14. Matloub Y, Bostrom BC, Hunger SP, Stork LC, Angiolillo A, Sather H, et al. Escalating intravenous methotrexate improves event-free survival in children with standard-risk acute lymphoblastic leukemia: a report from the Children's Oncology Group. Blood 2011;118(2):243-51.
15. Nachman J, Sather HN, Gaynon PS, Lukens JN, Wolff L, Trigg ME. Augmented Berlin-Frankfurt-Munster therapy abrogates the adverse prognostic significance of slow early response to induction chemotherapy for children and adolescents with acute lymphoblastic leukemia and unfavorable presenting features: A report from the children's cancer group. J Clin Oncol. 1997;15 (6):2222-30.
16. Seibel NL, Steinherz PG, Sather HN, Nachman JB, DeLaat C, Ettinger LJ, et al. Early postinduction intensification therapy improves survival for children and adolescents with high-risk acute lymphoblastic leukemia: a report from the Children's Oncology Group. Blood. 2008;111(5):2548-55.
17. Bowman WP, Larsen EL, Devidas M, Linda SB, Blach L, Carroll AJ, et al. Augmented therapy improves outcome for pediatric high risk acute lymphocytic leukemia: results of Children's Oncology Group trial P9906. Pediatr Blood Cancer. 2011;57 (4):569-77.
18. Lange BJ, Bostrom BC, Cherlow JM, Sensel MG, La MK, Rackoff W, et al. Double-delayed intensification improves event-free survival for children with intermediate-risk acute lymphoblastic leukemia: a report from the Children's Cancer Group. Blood. 2002;99(3):825-33.
19. Matloub Y, Lindemulder S, Gaynon PS, Sather H, La M, Broxson E, et al. Intrathecal triple therapy decreases central nervous system relapse but fails to improve event-free survival when compared with intrathecal methotrexate: results of the Children's Cancer Group (CCG) 1952 study for standard-risk acute lymphoblastic leukemia, reported by the Children's Oncology Group. Blood. 2006;108(4):1165-73.
20. Sandlund JT, Harrison PL, Rivera G, Behm FG, Head D, Boyett J, et al. Persistence of lymphoblasts in bone marrow on day 15 and days 22 to 25 of remission induction predicts a dismal treatment outcome in children with acute lymphoblastic leukemia. Blood. 2002;100(1):43-7.
21. Möricke A, Reiter A, Zimmermann M, Gadner H, Stanulla M, Dordelmann M, et al. Risk-adjusted therapy of acute lymphoblastic leukemia can decrease treatment burden and improve survival: treatment results of 2169 unselected pediatric and adolescent patients enrolled in the trial ALL-BFM 95. Blood. 2008;111(9):4477-89.
22. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958;53:457-81.
23. Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemother Rep. 1966;50(3):163-70.
24. Cox DR. Regression models and life tables. J R Stat Soc. 1972;34:187-220.
25. Smith M, Bleyer A, Crist W, Murphy S, Sallan SE. Uniform criteria for childhood acute lymphoblastic leukemia risk classification. J Clin Oncol. 1996;14(2):680-1.
26. Smith M, Arthur D, Camitta B, Carroll AJ, Crist W, Gaynon P, et al. Uniform approach to risk classification and treatment assignment for children with acute lymphoblastic leukemia. J Clin Oncol. 1996;14(1):18-24.
27. Dördelmann M, Reiter A, Borkhardt A, Ludwig WD, Gotz N, Viehmann S, et al. Prednisone response is the strongest predictor of treatment outcome in infant acute lymphoblastic leukemia. Blood. 1999;94(4): 1209-17.
28. Reiter A, Schrappe M, Ludwig WD, Hiddemann W, Sauter S, Henze G, et al. Chemotherapy in 998 unselected childhood acute lymphoblastic leukemia patients. Results and conclusions of the multicenter trial ALL-BFM 86. Blood. 1994;84(9):3122-33.
29. Gaynon PS, Angiolillo AL, Carroll WL, Nachman JB, Trigg ME, Sather HN, et al.

Long-term results of the children's cancer group studies for childhood acute lymphoblastic leukemia 1983-2002: a Children's Oncology Group Report. Leukemia. 2010;24(2):285-97.
30. Vora A, Mitchell CD, Lennard L, Eden TO, Kinsey SE, Lilleyman J, et al. Toxicity and efficacy of 6-thioguanine versus 6-mercaptopurine in childhood lymphoblastic leukaemia: a randomised trial. Lancet. 2006;368(9544):1339-48.
31. Gajjar A, Ribeiro R, Hancock ML, Rivera GK, Mahmoud H, Sandlund JT, et al. Persistence of circulating blasts after 1 week of multiagent chemotherapy confers a poor prognosis in childhood acute lymphoblastic leukemia. Blood. 1995;86(4):1292-5.
32. Chessells JM, Bailey C, Richards SM. Intensification of treatment and survival in all children with lymphoblastic leukaemia: results of UK Medical Research Council trial UKALL X. Medical Research Council Working Party on Childhood Leukaemia. Lancet. 1995;345(8943):143-8.
33. Oudot C, Auclerc MF, Levy V, Porcher R, Piguet C, Perel Y, et al. Prognostic factors for leukemic induction failure in children with acute lymphoblastic leukemia and outcome after salvage therapy: the FRALLE 93 study. J Clin Oncol. 2008;26(9):1496-503.
34. Arico M, Basso G, Mandelli F, Rizzari C, Colella R, Barisone E, et al. Good steroid response in vivo predicts a favorable outcome in children with T-cell acute lymphoblastic leukemia. The Associazione Italiana Ematologia Oncologia Pediatrica (AIEOP). Cancer. 1995;75(7):1684-93.
35. Schrappe M, Reiter A, Riehm H. Cytoreduction and prognosis in childhood acute lymphoblastic leukemia. J Clin Oncol. 1996;14(8):2403-5.
36. Thyss A, Suciu S, Bertrand Y, Mazingue F, Robert A, Vilmer E, et al. Systemic effect of intrathecal methotrexate during the initial phase of treatment of childhood acute lymphoblastic leukemia. The European Organization for Research and Treatment of Cancer Children's Leukemia Cooperative Group. J Clin Oncol. 1997;15(5):1824-30.
37. Schrappe M, Valsecchi MG, Bartram CR, Schrauder A, Panzer-Grümayer R, Möricke A, et al. Late MRD response determines relapse risk overall and in subsets of childhood T-cell ALL: results of the AIEOP-BFM-ALL 2000 study. Blood. 2011;118(8):2077-84.
38. Conter V, Bartram CR, Valsecchi MG, Schrauder A, Panzer-Grümayer R, Möricke A, et al. Molecular response to treatment redefines all prognostic factors in children and adolescents with B-cell precursor acute lymphoblastic leukemia: results in 3184 patients of the AIEOP-BFM ALL 2000 study. Blood. 2010;115(16):3206-14.
39. Gaynon PS, Trigg ME, Heerema NA, Sensel MG, Sather HN, Hammond GD, et al. Children's Cancer Group trials in childhood acute lymphoblastic leukemia: 1983-1995. Leukemia. 2000;14(12):2223-33.
40. Fronkova E, Mejstrikova E, Avigad S, Chik KW, Castillo L, Manor S, et al. Minimal residual disease (MRD) analysis in the non-MRD-based ALL IC-BFM 2002 protocol for childhood ALL: is it possible to avoid MRD testing? Leukemia. 2008;22(5):989-97.
41. van Dongen JJM, Seriu T, Panzer-Grümayer ER, Biondi A, Pongers-Willemse MJ, Corral L, et al. Prognostic value of minimal residual disease in acute lymphoblastic leukaemia in childhood. Lancet. 1998;352(9142):1731-8.
42. Campana D, Coustan-Smith E. Minimal residual disease studies by flow cytometry in acute leukemia. Acta Haematol. 2004;112(1-2):8-15.