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

Conter, V; Bartram, C R; Valsecchi, M G; Schrauder, A; Panzer-Grümayer, R; Möricke, A; Aricò, M; Zimmermann, M; Mann, G; De Rossi, G; Stanulla, M; Locatelli, F; Basso, G; Niggli, F; Barisone, E; Henze, G; Ludwig, W D; Haas, O A; Cazzaniga, G; Koehler, R; Silvestri, D; Bradtke, J; Parasole, R; Beier, R; van Dongen, J J M; Biondi, A; Schrappe, M
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

Abstract

The Associazione Italiana di Ematologia Oncologia Pediatrica and the Berlin-Frankfurt-Münster Acute Lymphoblastic Leukemia (AIEOP-BFM ALL 2000) study has for the first time introduced standardized quantitative assessment of minimal residual disease (MRD) based on immunoglobulin and T-cell receptor gene rearrangements as polymerase chain reaction targets (PCR-MRD), at 2 time points (TPs), to stratify patients in a large prospective study. Patients with precursor B (pB) ALL (n = 3184) were considered MRD standard risk (MRD-SR) if MRD was already negative at day 33 (analyzed by 2 markers, with a sensitivity of at least 10^{-4}); MRD high risk (MRD-HR) if 10^{-3} or more at day 78 and MRD intermediate risk (MRD-IR): others. MRD-SR patients were 42% (1348): 5-year event-free survival (EFS, standard error) is 92.3% (0.9). Fifty-two percent (1647) were MRD-IR: EFS 77.6% (1.3). Six percent of patients (189) were MRD-HR: EFS 50.1% (4.1; P < .001). PCR-MRD discriminated prognosis even on top of white blood cell count, age, early response to prednisone, and genotype. MRD response detected by sensitive quantitative PCR at 2 predefined TPs is highly predictive for relapse in childhood pB-ALL. The study is registered at http://clinicaltrials.gov: NCT00430118 for BFM and NCT00613457 for AIEOP.
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

CME article

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

*Valentino Conter,1,2 *Claus R. Bartram,3 Maria Grazia Valsecchi,4 André Schrauder,5 Renate Panzer-Grümayer,6 Anja Mörcke,8 Maurizio Aricò,7 Martin Zimmermann,8 Georg Mann,6 Giulio De Rossi,9 Martin Stanulla,8 Franco Locatelli,10 Giuseppe Basso,11 Felix Niggi,12 Elena Barisone,13 Günter Henze,14 Wolf-Dieter Ludwig,15 Oskar A. Haas,6 Giovanni Cazzaniga,16 Rolf Koehler,3 Daniela Silvestri,4 Jutta Bradtke,17 Rosanna Parasole,18 Rita Beier,8 Jacques J. M. van Dongen,19 Andrea Biondi,1,16 and Martin Schrappe5

1Department of Pediatrics, University of Milano-Bicocca, Ospedale S Gerardo, Monza, Italy; 2Department of Pediatrics, Ospedali Riuniti, Bergamo, Italy; 3Institute of Human Genetics, Ruprecht-Karls University, Heidelberg, Germany; 4Medical Statistics Unit, Department of Clinical Medicine and Prevention, University of Milano-Bicocca, Milan, Italy; 5Department of Pediatrics, University Medical School Schleswig-Holstein, Campus Kiel, Kiel, Germany; 6Children’s Cancer Research Institute and St Anna Kinderspital, Vienna, Austria; 7Department of Pediatric Hemato-Oncology, Ospedale Meyer, Firenze, Italy; 8Department of Pediatric Hematology and Oncology, Medical School Hannover, Hannover, Germany; 9Department of Pediatric Hemato-Oncology, Ospedale Bambin Gesù, Rome, Italy; 10Pediatric Hematology-Oncology, Fondazione, Istituti di Ricovero e Cura a Carattere Scientifico (IRCCS) Policlinico San Matteo, Università di Pavia, Italy; 11Pediatric Hemato-Oncology, University of Padua, Ospedale Policlinico, Padova, Italy; 12Pediatric Oncology, University Children’s Hospital Zürich, Zürich, Switzerland; 13Department of Pediatric Hemato-Oncology, Ospedale Infantile Regina Margherita, Turin, Italy; 14Pediatric Hematology and Oncology, Virchow Hospital, Charité, Berlin, Germany; 15Hematology/Oncology, Robert-Rössle-Klinik at the HELIOS Klinikum, Charité, Berlin, Germany; 16Centro Ricerca M Tettamanti, Clinica Pediatrica Università Milano-Bicocca, Monza, Italy; 17Oncogenetic Laboratory, Pediatric Hematology and Oncology, Justus-Liebig-University, Giessen, Germany; 18Department of Pediatrics Hemato-Oncology, Ospedale Pausillipon, Napoli, Italy; and 19Department of Immunology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

The Associazione Italiana di Ematologia Oncologia Pediatrica and the Berlin-Frankfurt-Münster Acute Lymphoblastic Leukemia (AIEOP-BFM ALL 2000) study has for the first time introduced standardized quantititative assessment of minimal residual disease (MRD) based on immunoglobulin and T-cell receptor gene rearrangements as polymerase chain reaction targets (PCR-MRD), at 2 time points (TPs), to stratify patients in a large prospective study. Patients with pre-cursor B (pB) ALL (n = 3184) were considered MRD standard risk (MRD-SR) if MRD was already negative at day 33 (analyzed by 2 markers, with a sensitivity of at least 10−4); MRD high risk (MRD-HR) if 10−3 or more at day 78 and MRD intermediate risk (MRD-IR): others. MRD-SR patients were 42% (1348): 5-year event-free survival (EFS, standard error) is 92.3% (0.9). Fifty-two percent (1647) were MRD-IR: EFS 77.6% (1.3). Six percent of patients (189) were MRD-HR: EFS 50.1% (4.1; P < .001). PCR-MRD discriminated prognosis even on top of white blood cell count, age, early response to prednisone, and genotype. MRD response detected by sensitive quantitative PCR at 2 predefined TPs is highly predictive for relapse in childhood pB-ALL. The study is registered at http://clinicaltrials.gov: NCT00430118 for BFM and NCT00613457 for AIEOP. (Blood. 2010;115(16):3206-3214)
Over the past 3 decades, remarkable advances have been achieved in the treatment of acute lymphoblastic leukemia (ALL) in children.1-9 However, significant challenges still remain. In fact, although current treatment strategies result in long-term remission for nearly 80% of children with ALL, the remaining 20% ultimately relapse, and cure rate after relapse is approximately 25% to 40%. Moreover, some subgroups of children who now receive intensive therapy are likely to be overtreated and may well be cured using less intensive regimens, resulting in reduced toxicity and fewer long-term side effects.10-12 Besides risk factors associated with the patient (eg, sex, age at diagnosis) and the disease (eg, white blood cell count at diagnosis, immunophenotype, structural and numeric chromosomal aberrations), measurement of in vivo treatment effectiveness has proven to be most important in predicting patient outcome and risk of relapse.4,13-18

Over the past 10 to 15 years, several techniques have been developed to complement and refine morphology in assessing response to treatment, including immunologic or molecular markers, fluorescent in situ hybridization, in vitro drug response, and colony assays.19-24 This technologic advancement led to introducing the concept of minimal residual disease (MRD), which has challenged the conventional definition of “remission.”25

Several studies on childhood and adult ALL, based mainly on immunoglobulin (Ig) and T-cell receptor (TCR) gene rearrangements as polymerase chain reaction (PCR) targets for monitoring MRD, have indicated that the detection of residual disease during initial cytotoxic treatment can predict outcome.13,17,21 Groups collaborating in the International Berlin-Frankfurt-Münster Study Group (I-BFM-SG) have pioneered the evaluation of MRD by PCR in childhood ALL. MRD assessment at days 33 and 78 of treatment by 2 clonal Ig/TCR markers, with sensitivity of at least 10⁻⁴, identified 3 risk groups. MRD-based stratification was superior to that based on other clinically relevant risk factors, including age, blast count at diagnosis, and immunophenotype.13

The collaborative prospective Associazione Italiana di Ematologia Pediatrica and Berlin-Frankfurt-Münster (AIEOP-BFM ALL) 2000 study of childhood ALL (carried out in Austria, Germany, Italy, and Switzerland) was based on these results. The laboratory requirements and feasibility results of the MRD-directed risk stratification using real-time, quantitative PCR analysis of Ig/TCR gene rearrangements, as applied in this large multicenter prospective trial, have been recently published.19 In this paper, we report on the clinical results, and the predictive value for relapse, of the quantitative assessment of MRD in childhood B-cell precursor ALL (pB-ALL), treated with the AIEOP-BFM ALL 2000 protocol. Results in T-ALL will be reported separately.

Methods

Patients

From July 1 (September 1 for AIEOP) 2000 to July 31 (June 30 for BFM) 2006, a total of 4741 patients with Philadelphia chromosome-negative (Ph⁻) ALL, aged between 1 and 18 years (infants younger than 1 year were eligible for a separate protocol), were diagnosed in one of the 127 participating study centers and were eligible for the study AIEOP-BFM ALL 2000. Of these patients, 4016 had Ph⁻ pB-ALL and 3184 (79.3%) were stratified by MRD. Lack of material or of sensitive MRD markers (n = 772), early death (n = 59), or relapse before day 78 (n = 1) precluded the use of MRD for final patient stratification in the remaining 832 (20.7%) patients.19 Because a new specific protocol was opened for Ph⁺ ALL subtype from 2004 onward, only 79 Ph⁺ ALL patients were recruited in AIEOP-BFM ALL 2000; 54 were stratified by MRD and are here described as a separate additional subgroup.

Diagnostic studies

Diagnosis of ALL was performed using cytomorphology (French-American-British criteria) and cytochemistry when 25% or more lymphoblastic cells were present in the bone marrow. Flow cytometric immunophenotyping was performed according to consensus protocols based on the guidelines proposed by the European Group for the Immunological Characterization of Leukemias.26 Presence of TEL/AML1, BCR/ABL, and MLL/AF4 fusion transcripts was screened as previously described.27 Complete remission (CR) was defined as the absence of physical signs of leukemia or detectable leukemia cells on blood smears, a bone marrow with active hematopoiesis and fewer than 5% leukemia blast cells, and normal cerebrospinal fluid. Patients who did not achieve CR at the end of induction phase IA were treated with phase IB of protocol I, and 3 subsequent high-risk (HR) blocks; resistance was defined as failure to achieve CR by the end of the third HR block (supplemental Figure 1 and supplemental Table 1 for protocol details, available on the Blood Web site; see the Supplemental Materials link at the top of the online article).

Molecular marker identification and MRD analysis

The logistics of the study, cell sample isolation, and identification of the markers for MRD evaluation have been recently reported.19 Briefly, DNA samples obtained at diagnosis were screened by PCR amplification using the BIOMED-1 primer sets for Ig kappa deleting element gene rearrangements, IGK-Kde (Vk-Kde, intron-Kde), complete and incomplete TCR delta (TCRD; VD-(DJ)-JC1, DJ2-JC1, DJ3-JC1), and TCR gamma (TCRG; VG-j-g; Jg1.1/2.1) rearrangements.26 Complete and incomplete IGH rearrangements (VH-(DH)-1, DH-JH) were identified using 5 VH and 7 DH family primers in combination with 1 JH consensus primer,29,30 whereas for incomplete and complete TCR beta (TCRB; Dβ-Jβ and Vβ-Dβ-Jβ) and IGH (Vc-Jc) rearrangements, the respective BIOMED-2 multiplex PCR primer sets were used.31

Junctional regions of clonal PCR products were sequenced, and patient-specific junctional sequences of potential PCR-MRD targets were identified.19 Allele-specific oligonucleotide primers were designed complementary to the junctional region sequence of each target, either manually or using the Primer Express software (Applied Biosystems). PCR-MRD targets were tested for specificity and sensitivity with the aim for each patient to select 2 targets with a sensitivity of at least 10⁻⁴ and a quantitative range of at least 10⁻⁴ for one target and at least 5 × 10⁻⁴ for the second target.32 Quantitative reverse-transcription–PCR analysis was performed and interpreted according to the guidelines developed within the European Study Group for MRD detection in ALL (ESG-MRD ALL).32
Relapse in Treating Young Patients With Acute Lymphoblastic Leukemia” with the protocol identification number NCT00430118 for BFM and NCT00613457 for AIEOP.

**Results**

The cohort of 4016 children between 1 and 18 years of age with newly diagnosed Ph$^-$pB-ALL, enrolled in the AIEOP-BFM ALL 2000 study, had a 7-year EFS (standard error [SE]) and survival (SE) of 80.4% (0.9) and 91.8% (0.5), respectively.

The 3184 children stratified by MRD had, by definition, completed induction phase IA (TP1) and induction consolidation phase IB (TP2). This report focuses only on these patients, alive at TP2 (day +78).

Their overall 7-year estimates for EFS (SE) and survival (SE) were 80.7% (1.0) and 92.8% (0.6), respectively, with a median follow-up of 4.0 years (supplemental Figure 2). The 7-year EFS and survival were 77.4% (1.6) and 92.0% (0.9) for 1329 AIEOP patients vs 83.0% (1.3) and 93.4% (0.7) for 1855 BFM patients, respectively. The difference in EFS was consistent in all subgroups of patients stratified by MRD, and the impact of PCR-MRD was found to be the same in AIEOP and BFM centers. The presented analyses will therefore concern the whole cohort of patients recruited by the 2 groups and stratified by MRD.

Among the 3184 patients, 42% were MRD-SR, 52% MRD-IR, and 6% MRD-HR; they had a significantly different outcome, with 5-year EFS estimates of 92.3% (0.9), 77.6% (1.3), and 50.1 (4.1), respectively (Figure 1A). The 5-year survival estimates were

---

**Risk group definitions and final stratification**

Patients were defined as MRD standard risk (MRD-SR) if MRD was found to be negative at both days 33 (time point 1 [TP1]) and 78 (TP2), using at least 2 molecular markers with sensitivity of $10^{-4}$ or less. If MRD levels differed between the 2 markers, the highest MRD level was chosen for the final MRD assessment. Patients were considered to be MRD intermediate risk (MRD-IR) when MRD was positive at 1 or both TPs but at a level of less than $10^{-3}$ at TP2 with at least 2 markers. Patients with MRD $10^{-3}$ or more at TP2 were defined MRD high risk (MRD-HR), independent of the sensitivity and the number of markers. Patients with prednisone poor response (PPR; $>1000$ leukemic blasts per microliter in the peripheral blood on day 8) or failure to achieve remission (ie, persistent extramedullary disease) after induction phase IA (induction failure) or positivity for MLL/AF4 fusion transcript were treated in the HR arm irrespective of their MRD results. If MRD evaluation was not available, patients were assigned to the IR group or, based on clinical parameters, to the HR group; these patients are not part of this report.

**Treatment protocol**

All enrolled patients were treated according to the AIEOP-BFM ALL 2000 study protocol: treatment outlines, details, and differences between AIEOP and BFM are shown in supplemental Figure 1 and supplemental Table 1. The ethics committee of each participating organization approved the study protocol. Treatment outlines, details, and differences between AIEOP and BFM are shown in supplemental Figure 1 and supplemental Table 1. The ethics committee of each participating organization approved the study protocol.

**Induction and consolidation phase.** All patients underwent 7 days of prephase with steroid therapy (prednisone) and 1 intrathecal dose of methotrexate (intrathecal MTX), followed by induction phase IA and induction consolidation phase IB; from day 8, patients were randomized to continue steroid treatment with either prednisone (60 mg/m² per day) or dexamethasone (10 mg/m² per day) until day 28 with subsequent tapering of dose in 1 week.

**Protocol M and reinduction phases.** SR and IR patients received 4 cycles of high-dose MTX (HD-MTX, 5 g/m²); AIEOP patients with non–T-ALL and without central nervous system (CNS) or testicular involvement at diagnosis received 2 g/m². At the beginning of the reinduction phase, patients were randomized to receive either protocol II or reduced-intensity protocol III in SR group, or protocol II versus reduced-intensity protocol III given twice in the IR group. HR patients were randomized to receive 3 blocks of non-cross-resistant drugs followed by protocol III given 3 times versus 3 blocks followed by protocol II given twice in the AIEOP group, or 6 blocks followed by protocol II in the BFM group.

**Maintenance therapy.** Maintenance therapy consisted of daily 6-mercaptopurine together with weekly MTX until 24 months from diagnosis.

**CNS-directed therapy.** CNS-directed therapy consisted of repeated intrathecal MTX administration during each treatment phase. Only patients who were treated in AIEOP centers and who did not undergo irradiation received MTX treatment also during the continuation phase. Cranial radiotherapy was given (dosage by age; supplemental Table 1) to patients at HR or with CNS involvement at diagnosis.

**Statistical analysis**

Event-free survival (EFS) and survival times were calculated from date of diagnosis to date of event, which, for EFS, was resistance, relapse, death, or second neoplasm, whichever occurred first, and, for survival, death from any cause. Of note, stratification by MRD is possible, by definition, only on the subpopulation of patients who are failure-free by TP2. EFS and survival curves were estimated according to Kaplan-Meier with Greenwood standard error, always indicated in parentheses, and compared according to log-rank test. Cumulative incidence curves for relapse were estimated adjusting for competing risks of other events and were compared with the Gray test. The Cox model, after stratification by group (AIEOP and BFM), was used to analyze the prognostic role of PCR-MRD in terms of cause-specific hazard of relapse. A 1-step multivariate model that included the most frequently used conventional prognostic features was applied. The proportional hazard assumption was tested by graphic checks. Tests were 2-sided, with .05 significance level. Analyses were carried out using SAS 9.1.

The study is registered at the US National Institutes of Health website http://clinicaltrials.gov as “Combination Chemotherapy Based on Risk of
Institute criteria, 34 were at standard risk (MRD-SR, 45.3%; IR, 9.8%) confirms this observation.

The distribution by MRD among patients who, by National Cancer HR, 50.5%; HR, 4.3%) or at high risk (MRD-SR, 35.5%; IR, 54.7%; HR, high risk; and CR, continuous complete remission.

50.5% (4.3) for those with MRD-IR shows that with WBC count 100 such patients had an unfavorable prognosis (Figure 3A-B). Furthermore, early MRD negativity among patients at medium risk by BFM 95 stratification criteria (Figure 3B) identified a major subset (42.4%) with excellent prognosis. Interestingly, patients at high risk by BFM 95 criteria (Figure 3C), most of them with PPR, also could be clearly stratified by MRD, and 15.6% with early clearance of residual disease had an excellent outcome.

Superiority of PCR-MRD over conventional stratification criteria can also be shown by evaluating its impact in patients of the AIEOP-BFM ALL 2000 study after stratifying them according to the BFM 95 criteria (based on age, WBC count, early morphologic response, and clonal translocations; Table 2). Patients (3% and 4.8%, respectively) at standard and medium risk by BFM 95 criteria were MRD-HR, and these patients had an unfavorable prognosis (Figure 3A-B).

97.8% (0.5), 93.4% (0.7), and 60.8 (4.0), respectively. Types of event are reported in Table 1; 43 patients died in CR (17 of them after stem cell transplantation [SCT] in first CR) and 22 (5-year cumulative incidence 0.8% [0.2]) had a secondary neoplasm. Most events were relapses and the 5-year cumulative incidence was markedly different according to the group of risk by MRD: 6.0% (0.8) for MRD-SR, 21.0% (1.2) for MRD-IR, and 34.9% (3.8) for MRD-HR patients (Figure 1B).

The distribution of disease and patient characteristics by PCR-MRD stratification is shown in Table 2. Patients with hyperleukocytosis (white blood cell [WBC] count ≥ 100 x 10^9/L) or PPR had higher probability of being MRD-HR. However, the MRD-SR group comprised rather large proportions of patients from subsets conventionally considered at high risk, namely those with WBC count 100 x 10^9/L or higher and age 10 years or older. With MRD according to age, WBC count, early morphologic response, and clonal translocations; Table 2).

MRD diagnostics discriminates prognostic subgroups on top of classical HR criteria. Figure 2 shows the prognostic value of MRD in subgroups defined by genetic features. In TEL/AML1-positive patients (n = 762; 25.9% of pB-ALL), high levels of MRD at TP2 were rarely detected (n = 10; 1.3%); yet MRD was able to stratify the remaining patients in 2 large subgroups (57.7% MRD-SR; 40.9% MRD-HR) with significantly different outcomes (Figure 2A). Similar results were obtained in patients with favorable DNA index (≥ 1.16 and < 1.6): MRD-HR patients were rare (1.4%); MRD-SR and IR patients had significantly different outcome (Figure 2B).

Fifty-four Ph+ ALL patients treated in our study and stratified by MRD are here analyzed separately, as it is of major interest to evaluate whether MRD response discriminates outcome also in these patients who are generally considered to have a dismal prognosis. Ph+ patients enrolled in the AIEOP-BFM ALL 2000 study were allocated to the high-risk treatment arm and had a markedly different outcome according to MRD (Figure 2C). In more detail, 8 patients (14.8%) were MRD-SR; 6 of them remained in continuous complete remission (CCR) (3 after SCT), 1 died in CCR, and 1 relapsed at 2.7 years from diagnosis; 24 (44.4%) were MRD-HR: 16 remained in CCR (13 after SCT), 1 died in CCR, 2 died because of treatment-related complications after SCT, and 5 relapsed after 0.6 to 5.1 years; 22 (40.8%) were MRD-HR and only 4 remained in CCR (all after SCT), 1 died in CCR, 6 died from treatment related mortality after SCT and 11 relapsed. Of note, MRD identified, even within the Ph+ ALL prednisone good-response subgroup (n = 41), patients at very high risk of relapse (MRD-HR, n = 11, 5 relapses, 4 deaths).

Table 1. Distribution of events in pB ALL Ph− patients according to MRD classification

<table>
<thead>
<tr>
<th>MRD-SR</th>
<th>MRD-IR</th>
<th>MRD-HR</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>%</td>
<td>No.</td>
<td>%</td>
</tr>
<tr>
<td>Total</td>
<td>1348</td>
<td>42.3</td>
<td>1647</td>
</tr>
<tr>
<td>Resistant*</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Relapses</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bone marrow</td>
<td>32</td>
<td>2.4</td>
<td>168</td>
</tr>
<tr>
<td>Central nervous system</td>
<td>9</td>
<td>0.6</td>
<td>29</td>
</tr>
<tr>
<td>Testis</td>
<td>7</td>
<td>0.5</td>
<td>21</td>
</tr>
<tr>
<td>Bone marrow + other</td>
<td>12</td>
<td>0.9</td>
<td>41</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>0.1</td>
<td>7</td>
</tr>
<tr>
<td>Death in CCR</td>
<td>10</td>
<td>0.7</td>
<td>13</td>
</tr>
<tr>
<td>After chemotherapy</td>
<td>10</td>
<td>0.7</td>
<td>10</td>
</tr>
<tr>
<td>After stem cell transplantation</td>
<td>0</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Second malignant neoplasm†</td>
<td>10</td>
<td>0.7</td>
<td>9</td>
</tr>
<tr>
<td>CCR</td>
<td>1267</td>
<td>94.1</td>
<td>1359</td>
</tr>
</tbody>
</table>

Only patients alive by TP2 are included in this cohort and consequently induction deaths in protocols IA and IB are not included.

ALL indicates acute lymphoblastic leukemia; MRD, minimal residual disease; SR, standard risk; IR, intermediate risk; HR, high risk; and CCR, continuous complete remission.

*Resistant patients are those who did not achieve CR by the end of the third HR block of chemotherapy.

†Acute myeloid leukemia (n = 10; 1.3%); yet MRD was able to stratify the distribution of relapses in pB-ALL are shown in supplemental Table 2. Patients (3% and 4.8%, respectively) at standard and medium risk by BFM 95 criteria were MRD-HR, and these patients had an unfavorable prognosis (Figure 3A-B).

Superiority of PCR-MRD over conventional stratification criteria can also be shown by evaluating its impact in patients of the AIEOP-BFM ALL 2000 study after stratifying them according to the BFM 95 criteria (based on age, WBC count, early morphologic response, and clonal translocations; Table 2). Patients (3% and 4.8%, respectively) at standard and medium risk by BFM 95 criteria were MRD-HR, and these patients had an unfavorable prognosis (Figure 3A-B).

Furthermore, 49.7% of patients at standard risk by BFM 95 criteria (Figure 3A) presented a slower MRD clearance (MRD-IR) and had a significantly inferior outcome compared with MRD-SR patients (P < .001). Conversely, early MRD negativity among patients at medium risk by BFM 95 stratification criteria (Figure 3B) identified a major subset (42.4%) with excellent prognosis. Interestingly, patients at high risk by BFM 95 criteria (Figure 3C), most of them with PPR, also could be clearly stratified by MRD, and 15.6% with early clearance of residual disease had an excellent outcome.

The dynamics of MRD clearance at TP1 and TP2 and its relation to the distribution of relapses in pB-ALL are shown in supplemental Table 2. By single time point, 44% of the patients (1399/3176) were MRD (2470/3176) at TP2, with a 10.6% (0.8) 5-year cumulative incidence of resistant* 0 0 3 1.6 3 0.1
and age between 1 and 6 years and no T-ALL. In the multivariate analysis (Table 3), except for the increased risk of relapse in PPR patients, which is not significant (hazard ratio 40.7% (3.0) in patients with MRD still positive but at a level less than 5-year cumulative incidence of relapse of 20.7% (2.9), compared with SR levels.

The multivariate analysis confirms the prognostic value of PCR-MRD and shows that WBC count, TEL/AML1 status, and DNA index retain independent significant impact on the hazard of relapse. Interestingly, PPR is now associated with a borderline significant lower risk of relapse. For the interpretation of these results, many aspects should be considered. First, PPR (compared with prednisone good responders [PGR]) is still related to a higher probability of having PCR-MRD HR levels (Table 2), which explains the higher relapse rate in univariate analysis (Table 3). However, Kaplan-Meier analysis (Table 2) shows that PPR completely loses its adverse prognostic value if compared with PGRs who have the same PCR-MRD levels (SR, IR, HR). This is not true, for instance, for higher WBC counts, which, within the same age group, clearly plays a role in improving these patients' outcome.

### Discussion

AIEOP-BFM ALL 2000 is the first, large-scale, prospective clinical trial in ALL in which PCR-MRD measured at 2 time points, days 33 and
due to an intensification of treatment as already shown in AIEOP ALL 95 and ALL-BFM 95 studies.6,13,37,39

Besides MRD-HR patients, those presenting the chromosomal translocations t(4;11) and t(9;22), or PPR, or not reaching CR at the end of induction were treated as HR regardless of the PCR-MRD findings. Postinduction treatment in other patients (SR or IR) was determined only by MRD response. In this context, MRD measurement by PCR retained its ability to discriminate prognosis in pB-ALL. This was consistently observed in each subgroup defined by the most relevant initial presenting features, either favorable or unfavorable, such as WBC count, age, and prednisone response. PCR-MRD discriminated distinct prognostic groups also in the 2 largest genetic subsets of pediatric pB-ALL, that is, the TEL/AML1-positive patients (A), in 487 patients with favorable DNA index (≥1.16 and <1.6; B), and in 54 Ph+ patients (C).

Figure 2. Prognostic impact of PCR-MRD in 762 TEL/AML1-positive patients (A), in 487 patients with favorable DNA index (≥1.16 and <1.6; B), and in 54 Ph+ patients (C).

Figure 3. Prognostic impact of PCR-MRD in 3184 pB-ALL patients (Ph-) within subgroups according to ALL-BFM 95 criteria: standard (1263 patients; A), medium (1684 patients; B), and high (237 patients; C).
associated with a very poor prognosis. AML1-positive subgroup and the hyperdiploid subgroup, and in the rare Ph+ ALL subtype, which is generally considered to be associated with a very poor prognosis.

Data presented here on a very large cohort of pB-ALL patients (n = 3184) clearly show the superiority of PCR-MRD over more conventional stratification criteria in identifying 3 risk groups, with very different prognosis. Subgroups of patients who would be qualified as standard, medium, or high risk by BFM-95 stratification criteria (WBC count, age, response to prednisone or to induction phase IA, and the t(4;11) or t(9;22) chromosomal translocations) can easily be discriminated into 3 new risk groups by MRD measured at 2 time points. The same pattern is observed within ALL National Cancer Institute subgroups.

Results reported here confirm that, in pB-ALL, MRD negativity at the end of induction (TP1) is the strongest predictor for excellent outcome (5-year EFS > 90%). Therefore, in the future, it might be difficult to plan randomized studies to improve outcome in MRD-SR patients.

Interestingly, among 1358 MRD-SR patients, only 61 relapsed, and 29 of them presented with an extramedullary either isolated (n = 17) or combined (n = 12) relapse, which may be, especially for the isolated ones, not likely to be predicted by bone marrow MRD response. High levels of MRD at TP2 are predictive for poor outcome (5-year EFS < 50%), with a 5-year cumulative incidence of relapse (SE) of 27.8% (4.7) and 43.5% (6.4) in patients presenting PCR-MRD levels of 10^{-3} or more than 10^{-3}, respectively (supplemental Table 2). Interestingly, however, whereas among patients presenting high MRD levels at TP1, those with no PCR-MRD detectable at TP2 had a favorable outcome, those with PCR-MRD still present at TP2 had a marked increase in the risk of relapse (Figure 4). These findings show that TP2 PCR-MRD levels are very important in tailoring treatment in patients with high PCR-MRD levels at TP1. In the multivariate analysis (Table 3), PCR-MRD is, by far, the most relevant factor discriminating prognosis. It overshadows the importance of other known prognostic features, although WBC count at diagnosis, TEL/AML1 status, and DNA index retain independent value. Noticeably, PPR appears associated with a decreased hazard of relapse, which makes sense only considering that all PPR patients were allocated to HR treatment, regardless of any other features. Although the MRD prognostic value is also therapy dependent, in our opinion these results may provide some hints for other studies, too. A favorable prognosis in all patients with fast clearance of MRD (ie, MRD negative after 5 weeks of therapy) can be demonstrated independent of their non–MRD-related risk features. This suggests that in patients undergoing relatively intensive treatment, such as that of the AIEOP-BFM ALL 2000 study, SCT may not be indicated even in the presence of any other combination of adverse prognostic features (or risk factors) if the MRD response proved to be favorable. On the other hand, for patients with poor MRD response (ie, MRD ≥ 10^{-3} after 2 months of therapy) despite rather favorable non-MRD risk criteria, treatment intensification also comprising SCT may be indicated to compensate for the MRD-derived high relapse risk. MRD response in pB-ALL could thus

Table 3. Results of the univariate and multivariate analyses

<table>
<thead>
<tr>
<th>Age, y</th>
<th>Univariate analysis</th>
<th>Multivariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hazard ratio</td>
<td>95% CI</td>
</tr>
<tr>
<td>1-9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10-17</td>
<td>1.68</td>
<td>1.32-2.14</td>
</tr>
<tr>
<td>WBC count, ×10^9/L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower than 50</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>50-100</td>
<td>1.60</td>
<td>1.16-2.21</td>
</tr>
<tr>
<td>100 or higher</td>
<td>2.41</td>
<td>1.73-3.34</td>
</tr>
<tr>
<td>TEL/AML1*</td>
<td>Negative</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>1</td>
</tr>
<tr>
<td>DNA index</td>
<td>Less than 1.16 or 1.6 or more</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.16 or more and less than 1.6</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>Missing</td>
<td>0.80</td>
</tr>
<tr>
<td>Response to prednisone</td>
<td>Prednisone good responder</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Prednisone poor responder</td>
<td>1</td>
</tr>
<tr>
<td>PCR-MRD</td>
<td>MRD-SR</td>
<td>3.83</td>
</tr>
<tr>
<td></td>
<td>MRD-IR</td>
<td>9.47</td>
</tr>
<tr>
<td></td>
<td>MRD-HR</td>
<td>1</td>
</tr>
</tbody>
</table>

* Cox model stratified by group (AIEOP and BFM) on hazard of relapse in 2927 pB-ALL (Ph+) patients (with 354 relapses) for whom information on most covariates was available.

CI indicates confidence interval; WBC, white blood cell; PCR-MRD, polymerase chain reaction minimal residual disease; SR, standard risk; IR, intermediate risk; and HR, high risk.
Indeed be crucial in designing an algorithm on some therapeutic elements: we think, for example, that all CNS-negative patients who do not have a poor MRD response could be spared CNS radiotherapy. Treatment intensification could also be planned for patients with slow clearance of their disease (ie, MRD still detectable after 2 months of therapy) particularly if levels of MRD were high 4 or 5 weeks after diagnosis. Reduction of treatment intensity instead, in our opinion, should be performed only under controlled conditions. Such MRD-based risk algorithms are, however, still to some extent treatment dependent, and cannot universally be applied in every available ALL treatment regimen.

Measurement of MRD is widely applied in contemporary childhood ALL studies. The technologies used may be either molecular or flow cytometric (FCM) and these 2 approaches may differ for sensitivity, specificity, costs, reproducibility, and feasibility.45 Standardization, reproducibility, and quality control among different laboratories are absolutely necessary for both methodologies.32

In our study, we very strict criteria for patient stratification (at least 2 targets with a sensitivity of at least 10^-4),19,22,32 However, if we analyze data of patients excluded from the present report because they had only 1 sensitive marker available (n = 454), results are very similar to those of patients stratified by 2 sensitive markers, with a 5-year EFS of 94.1% (1.6) and of 83.3% (3.3), respectively, in patients who were MRD-SR (n = 294) or IR (n = 160) by 1 marker. Only 1 sensitive PCR-MRD marker may thus be adequate and may allow stratification of more than 90% of the patients. Accordingly, in the AIEOP-BFM ALL 2009 study, 1 marker will be considered sufficient to stratify patients by PCR-MRD in patients lacking a second one.

The choice of the MRD methodology for a childhood ALL study depends largely on the aims of the study and resources available. The 2 methodologies (eg, molecular biology and flow cytometry) could in fact also be used in a complementary fashion for different aims within the same study. As recently shown in a study conducted by AIEOP,46 measurement of FCM-MRD on bone marrow samples collected on day 15, as already suggested by PCR-MRD findings at day 15,47 may allow identification of very early responders (< 0.1% blast cells), who may benefit from treatment reduction, as well as a small subset of patients with high MRD levels (≥ 10% blast cells) who have a poor prognosis independent of PCR-MRD response and could thus benefit from treatment intensification. Accordingly, in the AIEOP-BFM ALL 2009 study, these 2 groups of patients will be eligible for a randomized treatment reduction study and for HR treatment, respectively.

The main finding of the present study is that MRD response in pB-ALL detected by highly sensitive and well-standardized PCR technique is highly predictive of relapse, thus reducing markedly the importance of conventional prognostic factors such as age, WBC count at diagnosis, genetic abnormalities, and also prednisone response. Yet, further stratification in prognostic subgroups for tailoring treatment intensity can be considered. In this regard, future AIEOP-BFM strategy will adopt more extensively the concept of MRD early response, combining FCM- and PCR-MRD evaluation, also taking biologic subgroups into consideration.

Acknowledgments

We thank the participants of the ESG-MRD-ALL for fruitful discussion on standardization and quality control of the MRD diagnostics. For AIEOP: We thank Marco A. Citterio, AIEOP reference laboratories (G.B. and A.B.) for their contribution. We thank in particular Prof G. Masera for his essential contribution to the design of the study and for the thoughtful supervision of the conduction of the study in Italy.

This work was supported by Comitato M. L. Verga and Fondazione Tettamanti, Fondazione Città della Speranza, Fondazione Cariparo and Grant Ric. Corrente OBG 2006/02/R/001822 (G.D.R.), Associazione Italiana per la Ricerca sul Cancro (AIRC; A.B., M.G.V.), Fondazione Cariplo (A.B.), and Ministero dell’Istruzione, Università e Ricerca (MIUR; A.B.).

For BFM: We thank the partners in the reference laboratories and in the central MRD laboratories, all the technicians for their expert work in cytology and MRD diagnostics, and the data managers for their careful study conduct.

This work was supported by Deutsche Krebshilfe (grant no. 50-2698 Schr1 and grant no. 50-2410 Ba7), by the Competence Network Pediatric Hematology and Oncology (KPOH), which was funded by the Federal Ministry of Research (BMBF) and Oncosuisse/Krebsforschung Schweiz (grant OCS 1230-02-2002), and the St Anna Kinderkrebsforschung Austria.

Authorship

Contribution: V.C. planned the study and contributed to study coordination for AIEOP; C.R.B. was responsible for diagnosis and PCR-MRD analyses for BFM Germany; M.G.V. was the study statistician for AIEOP; A.S. was responsible for diagnosis and PCR-MRD analyses for BFM Germany; R.P.-G. was responsible for diagnosis and PCR-MRD analyses for BFM Austria; A.M. contributed to study coordination and data revision for BFM Germany; M.Z. was the study statistician for BFM; G.M. planned and coordinated the study in Austria; M.A., G.D.R., M.S., F.L., E.B., G.H., and R.P. contributed to planning the study; G.B. was responsible for diagnosis and PCR-MRD analyses for AIEOP; F.N. planned the study and coordinated the study in Switzerland; W.-D.L. was responsible for analysis of immunophenotype and DNA index for BFM Germany; O.A.H. was responsible for cytogenetics, DNA index, and fluorescent in situ hybridization data in Austria; G.C. was responsible for diagnosis and PCR-MRD analyses for AIEOP; F.N. was responsible for analysis of fusion genes for BFM Germany; R.B. contributed to study coordination for BFM Germany; J.J.M.V.D. organized and supervised the standardization and quality control of the PCR-MRD diagnostics, including the development of guidelines for data interpretation; A.B. was responsible for diagnosis and PCR-MRD analyses for AIEOP; M.S. planned the study and coordinated the study in Germany; V.C., M.G.V., M.S., and A.B. wrote the paper, and all other authors reviewed the paper; all authors participated in the protocol development, study supervision, and data interpretation stages of this study, and have seen and approved the final version.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

A list of the AIEOP-BFM ALL 2000 study group members appears in the online supplemental Appendix.

Correspondence: Andrea Biondi, University of Milano Bicocca, M Tettamanti Research Center, Via Pergolesi, 33, 20052 Monza, MI; e-mail abiondi.unimib@gmail.com; andrea.biondi@unimib.it.
References