



Published in final edited form as:

Pediatr Blood Cancer. 2011 October ; 57(4): 569–577. doi:10.1002/pbc.22944.

Augmented Therapy Improves Outcome For Pediatric High Risk Acute Lymphocytic Leukemia: Results Of Children’s Oncology Group Trial P9906

W. Paul Bowman, MD¹, Eric L Larsen, MD², Meenakshi Devidas, PhD³, Stephen B. Linda³, Laurie Blach, MD⁴, Andrew J Carroll, PhD⁵, William L Carroll, MD⁶, D. Jeanette Pullen, MD⁷, Jonathan Shuster, PhD⁸, Cheryl L Willman, MD⁹, Naomi Winick, MD¹⁰, Bruce M Camitta, MD¹¹, Stephen P Hunger, MD¹², and Michael J Borowitz, MD PhD¹³

¹Cook Children’s Hospital, Fort Worth TX

²Maine Children’s Cancer Program, Scarborough ME

³Children’s Oncology Group and the College of Medicine, University of Florida, Gainesville FL

⁴Mount Sinai Comprehensive Cancer Center, Miami Beach FL

⁵University of Alabama at Birmingham, Birmingham, AL

⁶New York University Cancer Institute, New York NY

⁷University of Mississippi School of Medicine, Jackson MS

⁸College of Medicine, University of Florida, Gainesville FL

⁹University of New Mexico School of Medicine, Albuquerque NM

¹⁰University of Texas Southwestern School of Medicine, Dallas, TX

¹¹Midwest Children’s Center for Cancer and Blood Disorders, Medical College of Wisconsin, Milwaukee WI

¹²University of Colorado Denver School of Medicine, and The Children’s Hospital, Aurora, CO

¹³Department of Pathology, Johns Hopkins Medical Institutions, Baltimore MD

Children’s Oncology Group, Arcadia, CA

Abstract

Background—The augmented BFM regimen improves outcome for children with NCI high acute lymphoblastic leukemia (ALL). Patient age, sex, and presenting white blood cell count (WBC) can be used to identify a subset of approximately 12% of children with B-precursor ALL that had a 5-year continuous complete remission (CCR) rate of only about 50% on earlier Pediatric Oncology Group (POG) trials.

Procedures—Children’s Oncology Group trial P9906 evaluated a modified augmented BFM regimen in 267 patients with particularly high risk B-precursor ALL. Minimal residual disease (MRD) was assessed in blood at day 8 and in marrow at day 29 of Induction and correlated with outcome.

Results—The 5-year CCR probability for patients in P9906 was significantly better than that observed for similar patients on POG trials 8602/9006 ($62.2 \pm 3.7\%$ versus $50.6 \pm 2.4\%$; $p=0.0007$) but similar to POG 9406 ($63.5 \pm 2.4\%$; $p=0.81$). Interim analysis showed poor central nervous system (CNS) control, especially in patients with initial WBC $\geq 100,000/\mu\text{l}$. Day 29 marrow MRD positive ($\geq 0.01\%$) vs. negative patients had 5 year CCR rates of $37.1 \pm 7.4\%$ vs. $72.6 \pm 4.3\%$; day 8 blood MRD positive vs. negative patients had 5 year CCR rates of $57.1 \pm 4.6\%$ vs. $83.6 \pm 6.3\%$. End induction marrow MRD predicted marrow but not CNS relapse. In multivariate analysis, day 29 MRD $> 0.01\%$, initial WBC $\geq 100,000/\mu\text{l}$, male gender, and day 8 blood MRD $> 0.01\%$ were significant prognostic factors.

Conclusions—Augmented BFM therapy improved outcome for children with higher risk ALL. Day 8 blood and day 29 marrow MRD were strong prognostic factors in these patients.

Keywords

Acute lymphocytic leukemia; Phase III clinical trial; Prognostic factors; Minimal residual disease

INTRODUCTION

Despite advances in the treatment of children with acute lymphocytic leukemia (ALL), some patients still succumb to their disease[1–6]. Recent efforts have focused on defining prognostic factors so that novel treatment strategies can be offered to patients with a poor prognosis. A number of clinical and laboratory variables evident within four weeks of diagnosis, are predictive of a poor outcome, including: NCI high risk [based on age and white blood cell (WBC) count], genotype (Philadelphia chromosome or hypodiploidy) and a slow response to induction chemotherapy[7–9]. Recent efforts have focused on defining additional prognostic factors so that intervention with novel treatment strategies can be offered to patients with a poor prognosis. Among these, the presence of minimal residual disease (MRD) early in therapy has been associated with an adverse prognosis[10–17].

The “augmented BFM” (ABFM) regimen improves the outcome of NCI high risk ALL patients with either a slow or rapid early response to one week of multiagent chemotherapy as assessed by bone marrow (BM) morphology.[18,19] Children’s Oncology Group (COG) study P9906, was designed to test the administration of a modified ABFM regimen in a subgroup of patients with B-precursor ALL at particularly high risk of treatment failure defined by an algorithm developed by Shuster et al [20] based on age, WBC count and sex. These patients comprise approximately 12% of children and adolescents with B precursor ALL and had a predicted 4-year event-free survival of 44% based on earlier Pediatric Oncology Group (POG) studies[20–22]. A further objective of this study was to investigate the prognostic significance of MRD. Our results demonstrate that administration of ABFM therapy significantly improved outcome for this selected high risk subset compared to similar patients treated on some earlier POG trials. However, there was a high rate of isolated central nervous system (CNS) relapse, necessitating modification of therapy for the subset of patients with initial WBC $100,000/\mu\text{l}$ or higher. Patients who were MRD-positive by flow cytometry at day 8 (blood) or day 29 (BM) of Induction therapy had worse outcomes, directly related to the MRD burden. Finally, day 29 BM MRD was highly predictive of marrow, but not CNS relapse.

MATERIALS AND METHODS

Patients

The National Cancer Institute and institutional review boards of all participating POG institutions approved this study. Written informed consent was obtained from guardians or

parents according to the guidelines of the National Institutes of Health. Patients with B precursor ALL aged 1 to 21.99 years first enrolled on the COG P9900 classification/induction study, and received Induction chemotherapy according to NCI risk group. All patients ultimately enrolled in this study received a 4 drug (prednisone, vincristine, asparaginase, daunorubicin) Induction. Patients with an M2 (5–25% blasts) marrow at end Induction (day 29) received 2 additional weeks of therapy with the same agents. Patients who had <5% marrow blasts at day 29 or day 43 were eligible to participate in post-Induction trials for low (P9904), standard (P9905) or high risk ALL (P9906). Patients eligible for the study reported herein (P9906) either met the Shuster age/sex/WBC criteria for higher risk[20,23] (Supplementary Table 1), or they had CNS3 leukemia (5 or more WBC/microliter with blasts present on initial cerebrospinal fluid examination), testicular leukemia, or an *MLL* translocation. Patients with a Philadelphia chromosome or hypodiploidy (DNA index <0.81 or <45 chromosomes) were not eligible. Patients with the favorable genetic features of *ETV6-RUNX1* (formerly *TEL-AML1*) fusion or trisomies of both chromosomes 4 and 10 were not eligible unless they had CNS3 or testicular leukemia. Diagnostic immunophenotyping, cytogenetic and molecular genetic studies needed to determine information required for eligibility for P9906 were performed at reference laboratories as described previously. [23]

Treatment

Post-induction therapy is summarized in Table 1. This study was the first time that the augmented Berlin Frankfurt Muenster (ABFM) regimen developed by the Children's Cancer Group (CCG) was used in a POG trial. Because of this, there were a few changes from the baseline ABFM regimen used in CCG 1882 and 1961.[18,19] These changes included: (1) treatment duration was 130 weeks for all patients rather than 112 weeks for girls and 164 weeks for boys; (2) the prednisone dose during Induction was 40 mg/m²/day rather than 60 mg/m²/day; (3) the total daunomycin dose during induction therapy was 90 mg/m² (30 × 3) compared to 100 mg/m² (25 × 4); (4) patients received 6 doses of L-asparaginase at 10,000 units/m² over 2 weeks, compared to 9 doses at 6,000 units/m² over 3 weeks during Induction therapy; (5) during post-Induction therapy native *E.coli* asparaginase was used as had been done in CCG 1882, but different from pegylated asparaginase used in CCG 1961; (6) during maintenance therapy, dexamethasone 6 mg/m²/day was administered for five days every four weeks compared to prednisone 40 mg/m²/day for five days every four weeks; (7) in CCG 1882 and 1961 slow responding patients (bone marrow >25% blasts at day 7 of Induction therapy) received 1800 cGy cranial irradiation during consolidation those with CNS3 disease received 2400 cGy cranial plus 600 cGy spinal irradiation; originally study P9906 gave radiation only to patients with CNS3 disease (1800 cGy cranial) and was delayed until the first cycle of maintenance, approximately 47 weeks from treatment initiation; (8) CNS prophylaxis was intrathecal (IT) methotrexate alone at two week intervals during the eight week consolidation phase, compared to weekly for four doses in the original ABFM regimen.

In May, 2004, after accrual was completed, the P9906 protocol was amended due to a high rate of isolated CNS relapse as described below. At that time, it was recommended that patients with an initial WBC count of ≥100,000/μl who were within 24 months of initial diagnosis should receive 1800 cGy cranial irradiation.

Minimal residual disease

Peripheral blood (PB; day 8) and BM (day 29) specimens were sent to a single central reference laboratory for determination of MRD via 4-color flow cytometry as previously described.[17,23] Sensitivity of 0.01% was achieved in nearly all cases. In a few cases either cell numbers were limiting or the phenotype was not informative and overlapped

substantially with that of normal B cells so that sensitivity was limited to 0.1%; such cases were excluded from analyses that used an MRD threshold of 0.01%. The MRD results were indeterminate in 1.9% of cases at end induction, and no results obtained because no sample was provided in an additional 1.9%. Day 8 PB samples were not provided in 9% of cases but satisfactory results were obtained on all but 4 specimens that were received. No alterations in therapy were made based on MRD results, which were not provided to patients or treating physicians.

Statistical analyses

Patients were enrolled on P9906 after achieving complete remission (CR) at the end of Induction. The primary endpoint was continuous complete remission (CCR), computed as the time from enrollment on P9906 to either failure (relapse, second malignancy, or death) or, for non-failures, last contact. Actuarial comparisons were conducted using the log-rank test. Survival curves were estimated using the Kaplan-Meier method and standard errors of Peto et al.[24] Results are expressed as estimate \pm standard error. Cumulative incidence rates for isolated CNS relapse and marrow relapse were computed using the cumulative incidence function for competing risks, and comparisons were conducted using the *K*-sample test.[25] Multivariate analysis of risk factors for outcome was performed using the Cox regression model.[26]

The P9906 study was designed to determine if augmented BFM therapy would produce better results in this patient cohort as compared with historical controls from POG 8602 (1986–1991) [21] and POG 9006 (1991–1994)[22], selected to match eligibility requirements on P9906. Results from the POG 9406 high risk ALL study, which enrolled patients from 1994–1999[6], were not available at the time COG P9906 was designed, but have been included in outcome comparisons in this report. Since neither FISH nor molecular studies were performed on the historical studies to determine *ETV6-RUNX1* fusion, trisomies 4 and 10 (DT), Philadelphia chromosome (Ph) status, or presence of *MLL* translocations, the presumed status of these chromosomal abnormalities was inferred from the karyotype string. Patients with Ph+, DNA index < 0.81 , age ≥ 22.0 yrs, or failure to achieve CR were excluded from the historical control cohorts. Among the remaining patients, those not known to be definitively *ETV6-RUNX1*+ or DT+ who were *MLL*+, CNS3 with testicular disease at diagnosis, or met the Shuster age/sex/WBC high risk criteria were included in the historical control cohorts. The rate of CCR for the patients in the POG 8602/9006 (n=457) and 9406 (n=437) comparison groups was computed as time from date of CR to failure or last contact.

RESULTS

COG P9906 accrued 276 patients with B-Precursor ALL between 3/15/00 and 4/25/03. Data current as of April 2008 were used in this report. Though no formal analysis was performed after this, review of data in Jan 2010 identified only 3 additional first events (1 bone marrow relapse, 1 isolated CNS relapse, and one second malignancy) that do not affect the conclusions shown. Nine patients were either ineligible or inevaluable for study. Clinical features of the remaining 267 patients on P9906 are presented in Table 2. During the same time period, 55 additional patients with clinical characteristics that would have made them eligible for P9906 were enrolled on the COG P9900 classification study and achieved complete remission but did not enroll on P9906. There was no statistically significant difference in clinical characteristics including race, gender, CNS or testicular disease, age, white blood cell count, NCI risk group or frequency of *MLL* rearrangements between the 267 patients who enrolled on P9906 and the 55 who did not.

Outcomes

The five-year CCR rate for patients enrolled in P9906 was $62.2 \pm 3.7\%$, which was superior to that for similar patients who had been treated on P8602/P9006 ($50.6 \pm 2.4\%$; $p=0.0007$) [21,22]. However, the P9906 CCR rate was similar to that seen on the immediately preceding study P9406[6] ($63.5 \pm 2.4\%$, $p=0.81$) (Figure 1). Sites of relapse in P9906 are shown in Table 2: 66.3% (63/95) of relapses in P9906 occurred in marrow with or without concurrent extramedullary sites vs. 74.2% (112/151) in P9406 ($p=.196$) Similar differences were seen in 5-year overall survival rates when outcomes of P9906 were compared with P8602/P9006 ($81.0 \pm 2.9\%$ vs. $68.6 \pm 2.2\%$; $p=0.0053$) or 9406 ($76.1 \pm 2.1\%$; $p=0.7$)

Interim analysis conducted in early 2004, several months after P9906 was closed to patient accrual, indicated a higher rate of isolated CNS relapse (3-year cumulative incidence rate: $12.2 \pm 6.8\%$) compared to the most recent predecessor study P9406 ($5.7 \pm 1.5\%$) ($p < 0.01$). Because of this, further exploratory analyses were conducted to determine whether or not the risk of isolated CNS relapse was limited to certain patient subgroups. These analyses showed that P9906 patients with initial WBC $\geq 100,000/\mu\text{l}$ had an isolated CNS relapse rate of $28.2 \pm 11.0\%$ vs. $4.8 \pm 5.6\%$ ($p < 0.0001$) for those with initial WBC $< 100,000/\mu\text{l}$. Based on these findings, P9906 was amended and prophylactic cranial irradiation therapy (1800 cGy/12 fractions/16 days) was recommended for patients < 2 years from study entry with initial WBC $\geq 100,000/\mu\text{l}$ who had not been CNS3 at diagnosis. Twenty-two patients met these criteria, 17 of whom received irradiation. Five of the 17 irradiated patients subsequently relapsed (3 marrow; 1 isolated CNS; 1 marrow + CNS), as did one of the 5 patients (marrow) that did not receive irradiation. In this final study analysis, the 5-year cumulative incidence rates of isolated CNS relapse for patients with WBC $< 100,000/\mu\text{l}$ (N=175) compared to those with higher WBC (N=91) were $5.8 \pm 1.8\%$ vs. $19.8 \pm 4.2\%$ ($p=0.0004$), respectively. Overall median time to isolated CNS relapse was 15.7 months. Seven of the 28 isolated CNS relapses occurred before week 47, which was the time that patients with CNS3 disease at diagnosis were to receive cranial irradiation. Only two of the 28 CNS relapses occurred in CNS3 patients, one of whom relapsed before week 47. Overall, there was no significant difference ($p=0.71$) in cumulative incidence rates of isolated CNS relapse between P9906 ($10.6 \pm 1.9\%$ at 5 years) and the P8602/P9006 studies ($9.9 \pm 1.4\%$ at 5 years). However, the 5-year rate of $19.8 \pm 4.2\%$ for patients with WBC $> 100,000/\mu\text{l}$ was significantly higher than that seen on study P9406 ($4.4 \pm 1.6\%$; $p < 0.0001$), and trended higher than that seen on P8602/9006 ($13.3 \pm 2.4\%$; $p=0.15$). Four year post-relapse survival was $52.4 \pm 12.1\%$ in the patients on P9906 with isolated CNS relapse after a variety of salvage regimens.

Toxicities

The spectrum of side effects was similar to those in previous studies with ABFM therapy[18,19]. Cytopenias and infections were the most common toxicities. Allergic reactions to asparaginase occurred in slightly more than a fourth of patients, with 13% of patients having grade 3 or 4 toxicities. These allergic reactions were most commonly seen in consolidation. The availability of Erwinia asparaginase was limited throughout this study. For patients with allergic reactions to E coli asparaginase, PEGylated asparaginase was substituted until sensitivity developed. One quarter of the patients had cessation of all asparaginase products prior to completion of the second interim maintenance phase, which is similar to rates observed on the contemporary CCG 1961 trial for patients with high risk ALL[19].

Symptomatic avascular necrosis (AVN) was reported in 14 patients (11.2 to 20.4 years of age). However, because data on AVN was not collected prospectively, this may be an

underestimate of the true incidence. Maintenance dexamethasone was discontinued in patients with AVN.

Four patients (1.5%) died in remission. Causes were: infection (2) (week 10, fungal and bacterial meningitis; week 15, septic shock), hemophagocytic lymphohistiocytosis (week 45); and multiple cerebral infarcts (week 30).

Effect of minimal residual disease on outcome

Satisfactory MRD results were obtained at a sensitivity of 0.01% from 238 patients (89.1%) in day 8 blood and in 245 patients (91.8%) in the end of Induction bone marrow samples; 222 patients (83.1%) had satisfactory MRD results for both day 8 blood and end induction marrow. Figure 2a illustrates CCR rate by day 29 marrow MRD, comparing patients who are negative to those expressing increasing degrees of positivity. Patients who were MRD positive at any level $>0.01\%$ had an extremely poor outcome, with a 5-year CCR rate of $37.1\pm 7.4\%$. Patients who were MRD negative had a better, though still not excellent outcome, with a 5-year CCR rate of $72.6\pm 4.3\%$ ($p < 0.0001$). However, in spite of its prognostic value, 46.8% of all events occurred in the 65.7% of patients (161/245) who were day 29 MRD negative.

Day 8 blood MRD was somewhat better at identifying a good risk group of patients (Figure 2b). Patients with increasing levels of day 8 blood MRD had increasingly poor outcomes, with the small group (9%) of patients with MRD $>10\%$ having a $25.0\pm 12.5\%$ 5-year CCR rate. Patients who were MRD positive at any level ($>0.01\%$) in day 8 blood had a $57.1\pm 4.6\%$ 5-year CCR rate compared to $83.6\pm 6.3\%$ for those who were negative ($p=0.0021$). Only 9.2% of events occurred in the 21% (49/238) of patients who were day 8 blood MRD negative.

In stepwise multivariate main effects Cox regression analysis WBC $\geq 100,000/\mu\text{l}$ (hazard ratio (HR) 2.7, $p < 0.0001$), day 29 bone marrow MRD positivity (HR 2.38, $p=0.0001$), male gender (HR 1.84, $p=0.02$), and day 8 blood MRD positivity (HR 2.19, $p=0.045$) contributed to poor prognosis. The independent (test of interaction, $p=0.75$) prognostic effect of MRD and WBC is shown in Figure 3.

Figure 4 demonstrates that end Induction marrow MRD was a robust predictor of marrow relapse but was not significantly associated with extramedullary (mainly CNS) relapse. The cumulative incidence of marrow relapse at 5 years was $13.0\pm 2.8\%$ among day 29 marrow MRD-negative patients ($<0.01\%$) vs. $48.8\pm 5.8\%$ among patients that were MRD positive ($p < 0.0001$). By contrast, the risk of non-marrow relapse at 5 years was identical in day 29 marrow MRD positive ($12.4\pm 3.7\%$) and MRD negative patients ($12.5\pm 2.6\%$, $p = 0.9170$)

DISCUSSION

Shuster identified an algorithm based on age/sex/WBC that identified a subset of about 12% of pediatric B-precursor ALL patients that had a very poor outcome on prior POG clinical trials, with 5-year CCR rates of only about 50% for children enrolled on the P8602 and P9006 studies between 1986–1994.[20–22] The P9906 study was designed to determine if augmented BFM therapy, which significantly improved the outcomes of high risk patients with a poor (CCG 1882, [18]) or good (CCG 1961, [19]) early marrow response, would improve outcome for this high risk patient subset. At the time that P9906 opened, the P9406 study had just completed accrual but outcomes were not available. The results of the P9906 study show that augmented BFM therapy produces significantly better outcomes for this high risk patient cohort than P8602 (which used 6 courses of intermediate dose (ID; 1 gram/ m^2) methotrexate with either intensive asparaginase treatment or high dose cytarabine) or

P9006, (which used 12 courses of ID MTX with intravenous 6-mercaptopurine or rotating drug pairs). The 5-year CCR rate on P9906 was not significantly different from that on P9406, although there was a non-significant trend towards improved 5-year overall survival on P9906. The P9406 study was a 2x2 study that tested 1 vs. 2.5 grams/m² of methotrexate and a teniposide/cytarabine combination versus high dose cytarabine alone. Neither experimental intervention led to a significant improvement in outcome on P9406, and the high dose cytarabine regimen was associated with significantly more toxicity[6]. Given these factors and the outstanding results obtained with augmented BFM therapy in other high risk ALL trials [18,19], the COG has continued to use this regimen for patients with NCI high risk ALL.[20–22]

Care must be taken in drawing conclusions from comparisons with historical controls because these may be biased by differences in patient populations. Patients enrolled on POG 8602, P9006, and P9406 were not exactly comparable to those on P9906 because the earlier studies did not have comprehensive studies to detect prognostically significant genetic lesions used for treatment stratification in contemporary COG regimens. However, this is unlikely to affect our conclusions. The most important poor prognosis lesions (Ph+ and hypodiploidy) would have been detected in most patients by the centrally performed cytogenetics used in the P8602/P9006/P9406 studies. Moreover, the favorable cytogenetic lesions DT and *ETV6-RUNX1*, which excluded patients from participation in P9906, are significantly more common than the unfavorable lesions and *ETV6-RUNX1* fusion is almost never detected by standard karyotype analysis, so that any bias would tend to make CCR higher in the historical control trials by including patients with undetected favorable genetic features. Other investigators have employed a variety of strategies to intensify therapy in high risk patients[27–29]; direct comparison with our results is difficult because our study was conducted on a highly selected poor risk population.

Interim analysis of P9906 revealed an unexpectedly high rate of CNS relapse. The 5-year CNS relapse rate of 10.6±1.9% was significantly higher than that on CCG 1961 (4.2±1.1%) which utilized a similar BFM backbone[19]. The reasons for this are unclear. One possible reason is patient selection; this specific group of patients selected as particularly high risk for relapse has never been treated in a uniform manner as done on this study. Other reasons might be therapy-related. This was the first use of the augmented BFM regimen in POG centers, and the treatment employed had several modifications from the original ABFM regimen including a lower dose of prednisone during Induction (40 vs. 60 mg/m²/day) and the delivery of the four doses of intrathecal (IT) methotrexate over eight rather than four weeks during Consolidation. Once we identified the very high rate of CNS relapse in patients with >100,000 WBC/ul we modified the protocol to provide CNS radiation to those patients, though it is not clear that the lack of CNS radiation per se was the reason for relapse; it has recently been shown that CNS relapses can be prevented with adequate chemotherapy without radiation even in high risk patients[30]. Finally, given the modest sample size of this trial it is also possible that this unexpected finding was due to chance. Whatever the explanation, our results indicate that even small changes from previously established therapeutic protocols may have significant effects on outcome when used in a different context.

The treatments on COG P9906 were well tolerated. The toxicity profile was similar to the previously reported intensive augmented BFM based therapy.[18,19], Similar to the experience in CCG 1882 and 1961, there was a relatively high rate of allergic reactions to asparaginase. Most of these events occurred during the consolidation and interim maintenance phases, during which patients were not receiving concurrent steroid therapy. There were 4 deaths related to treatment (1.5%), which was similar to that seen in the POG 9006 [22] (5/470; 1.1%) and CCG 1961 [19] (24/1299; 1.8%) studies for high risk ALL.

The level of MRD present both in blood at day 8 and in bone marrow at day 29 of Induction identified groups of patients with very different rates of CCR. Our findings agree with considerable evidence that MRD assessment during early phases of treatment is a powerful prognostic indicator in ALL, [12,13,15,17,31–35] We used flow cytometry to assess MRD, with a sensitivity of detection of .01% in the vast majority of patients. Other studies have used either flow cytometry, or polymerase chain reaction (PCR) amplification of clonal antigen receptor gene rearrangements. Although PCR often achieves a sensitivity of better than 0.01% [35–37], Neale et al have demonstrated that there is excellent correlation between PCR and flow cytometry in measuring MRD [38]. The use of flow cytometry for MRD measurements has become increasingly popular because of sensitivity, rapid turnaround, lack of requirement for patient- specific reagents, and affordability. [39]

Day 29 bone marrow MRD and WBC count $\geq 100,000/\mu\text{l}$ were strong independent predictors of poor outcome. Patients with both WBC count $\geq 100,000/\mu\text{l}$ and day 29 MRD $>0.01\%$ had only a $27.7 \pm 11.8\%$ 5-year CCR rate. By contrast, those patients with WBC $< 100,000/\mu\text{l}$ who were MRD negative had a relatively good 5-year CCR rate of $79.7 \pm 4.8\%$. It was recently shown that *IKZF1* deletions or mutations, *JAK* mutations, or translocations/focal interstitial deletions that lead to over-expression of *CRLF2* (all of which are strongly associated with one another) in this cohort was strongly associated with adverse prognosis independent of leukocyte count and MRD [40–42], but we did not include these factors in our model. Finally, the difference in outcome between day 29 MRD positive and negative groups is solely accounted for by marrow events; MRD status was not significantly associated with CNS relapse. This latter observation has important implications for how MRD might be used as a surrogate marker to evaluate the efficacy of treatment interventions. It seems reasonable to assume that an intervention provided early in therapy that reduces the rate of marrow relapse will also be associated with reductions in MRD burden. However, our results suggest that an intervention that reduces CNS but not marrow relapses might improve overall EFS but not show any effect on bone marrow MRD.

Day 8 blood MRD was also highly prognostic of outcome, and maintained its significance in multivariate analysis. Although the few patients with very high ($>10\%$) day 8 MRD had a dismal $25.1 \pm 12.5\%$ 5-year CCR rate, this measurement appeared particularly useful at identifying good risk patients; while only 21% of patients were day 8 PB MRD negative, these otherwise higher-risk patients had a 5-year CCR rate of $83.6 \pm 6.3\%$. Other investigators have also shown that measurements of MRD during induction can provide useful information and especially identify patients with very favorable prognosis [32,34] though these studies used bone marrows and not blood; there have been few measurements of blood MRD in other studies, except for patients with T-ALL, where blood has been shown to be a reasonable surrogate for marrow [43,44].

In summary, this study has several important implications for future ALL trials. First, these clinical age/WBC/sex criteria clearly identify a small subset ($\sim 12\%$) of pediatric ALL patients with a very poor outcome, even with highly effective contemporary ALL therapy. Second, ABFM improves outcome for these patients as compared to older POG anti-metabolite regimens, and has been carried forward in contemporary COG clinical trials. Third, even among this particularly high risk group of patients, Day 29 MRD is highly correlated with marrow relapse and associated with an extremely poor outcome, suggesting that patients with these higher-risk age/sex/WBC criteria who also have marrow MRD $>0.01\%$ at the end of induction therapy should be considered for novel treatment strategies in first remission. While there are fewer CNS than marrow relapses, and therefore more limited statistical power, the nearly identical CNS relapse rates between MRD positive and negative patients suggest that day 29 marrow MRD levels have minimal, if any, association

with subsequent CNS relapse so that CNS-directed therapy should be guided by factors other than the presence of marrow MRD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by Grants R01 CA86011, CA29139, and U10 CA98543 from NCI. SPH is the Ergen Family Chair in Pediatric Cancer. Presented in part at the 2005 American Society of Hematology Meeting, Atlanta GA

A complete listing of grant support for research conducted by CCG and POG before initiation of the COG grant in 2003 is available online at: <http://www.childrensoncoogygroup.org/admin/grantinfo.htm>

References

1. Silverman LB, Stevenson KE, O'Brien JE, et al. Long-term results of Dana-Farber Cancer Institute ALL Consortium protocols for children with newly diagnosed acute lymphoblastic leukemia (1985–2000). *Leukemia*. 2010; 24:320–334. [PubMed: 20016537]
2. Conter V, Arico M, Basso G, et al. Long-term results of the Italian Association of Pediatric Hematology and Oncology (AIEOP) Studies 82, 87, 88, 91 and 95 for childhood acute lymphoblastic leukemia. *Leukemia*. 2010; 24:255–264. [PubMed: 20016536]
3. Gaynon PS, Angiolillo AL, Carroll WL, 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; 24:285–297. [PubMed: 20016531]
4. Moricke A, Zimmermann M, Reiter A, et al. Long-term results of five consecutive trials in childhood acute lymphoblastic leukemia performed by the ALL-BFM study group from 1981 to 2000. *Leukemia*. 2010; 24:265–284. [PubMed: 20010625]
5. Pui CH, Pei D, Sandlund JT, et al. Long-term results of St Jude Total Therapy Studies 11, 12, 13A, 13B, and 14 for childhood acute lymphoblastic leukemia. *Leukemia*. 2010; 24:371–382. [PubMed: 20010620]
6. Salzer WL, Devidas M, Carroll WL, et al. Long-term results of the pediatric oncology group studies for childhood acute lymphoblastic leukemia 1984–2001: a report from the children's oncology group. *Leukemia*. 2010; 24:355–370. [PubMed: 20016527]
7. Gaynon PS, Desai AA, Bostrom BC, et al. Early response to therapy and outcome in childhood acute lymphoblastic leukemia: a review. *Cancer*. 1997; 80:1717–1726. [PubMed: 9351539]
8. Schultz KR, Pullen DJ, Sather HN, 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; 109:926–935. [PubMed: 17003380]
9. Smith M, Arthur D, Camitta B, et al. Uniform approach to risk classification and treatment assignment for children with acute lymphoblastic leukemia. *J Clin Oncol*. 1996; 14:18–24. [PubMed: 8558195]
10. Bjorklund E, Mazur J, Soderhall S, et al. Flow cytometric follow-up of minimal residual disease in bone marrow gives prognostic information in children with acute lymphoblastic leukemia. *Leukemia*. 2003; 17:138–148. [PubMed: 12529671]
11. Cave H, van der Werff ten Bosch, Suciu S, et al. Clinical significance of minimal residual disease in childhood acute lymphoblastic leukemia. European Organization for Research and Treatment of Cancer--Childhood Leukemia Cooperative Group. *N Engl J Med*. 1998; 339:591–598. [PubMed: 9718378]
12. Coustan-Smith E, Sancho J, Hancock ML, et al. Clinical importance of minimal residual disease in childhood acute lymphoblastic leukemia. *Blood*. 2000; 96:2691–2696. [PubMed: 11023499]

13. Dworzak MN, Froschl G, Printz D, et al. Prognostic significance and modalities of flow cytometric minimal residual disease detection in childhood acute lymphoblastic leukemia. *Blood*. 2002; 99:1952–1958. [PubMed: 11877265]
14. Okamoto T, Yokota S, Katano N, et al. Minimal residual disease in early phase of chemotherapy reflects poor outcome in children with acute lymphoblastic leukemia--a retrospective study by the Children's Cancer and Leukemia Study Group in Japan. *Leuk Lymphoma*. 2002; 43:1001–1006. [PubMed: 12148878]
15. Panzer-Grumayer ER, Schneider M, Panzer S, et al. Rapid molecular response during early induction chemotherapy predicts a good outcome in childhood acute lymphoblastic leukemia. *Blood*. 2000; 95:790–794. [PubMed: 10648387]
16. van Dongen JJ, Seriu T, Panzer-Grumayer ER, et al. Prognostic value of minimal residual disease in acute lymphoblastic leukaemia in childhood. *Lancet*. 1998; 352:1731–1738. [PubMed: 9848348]
17. Borowitz MJ, Devidas M, Hunger SP, et al. Clinical significance of minimal residual disease in childhood acute lymphoblastic leukemia and its relationship to other prognostic factors: a Children's Oncology Group study. *Blood*. 2008; 111:5477–5485. [PubMed: 18388178]
18. Nachman JB, Sather HN, Sensel MG, et al. Augmented post-induction therapy for children with high-risk acute lymphoblastic leukemia and a slow response to initial therapy. *N Engl J Med*. 1998; 338:1663–1671. [PubMed: 9614257]
19. Seibel NL, Steinherz PG, Sather HN, 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; 111:2548–2555. [PubMed: 18039957]
20. Shuster J, Camitta B, Borowitz MJ, et al. Identification of newly diagnosed children with acute lymphocytic leukemia at high risk for relapse. *Cancer Res Therapy and Control*. 1999; 9:101–107.
21. Harris MB, Shuster JJ, Pullen J, et al. Treatment of children with early pre-B and pre-B acute lymphocytic leukemia with antimetabolite-based intensification regimens: a Pediatric Oncology Group Study. *Leukemia*. 2000; 14:1570–1576. [PubMed: 10995002]
22. Lauer SJ, Shuster JJ, Mahoney DH Jr, et al. A comparison of early intensive methotrexate/mercaptopurine with early intensive alternating combination chemotherapy for high-risk B-precursor acute lymphoblastic leukemia: a Pediatric Oncology Group phase III randomized trial. *Leukemia*. 2001; 15:1038–1045. [PubMed: 11455971]
23. Borowitz MJ, Pullen DJ, Sather HN, et al. Minimal residual disease detection in childhood precursor-B-cell acute lymphoblastic leukemia: relation to other risk factors. A Children's Oncology Group study. *Leukemia*. 2003; 17:1566–1572. [PubMed: 12886244]
24. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. *Br J Cancer*. 1977; 35:1–39. [PubMed: 831755]
25. Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Annal Stat*. 1988; 16:1141–1154.
26. Cox DR. Regression models and life tables. *J R Stat Soc B*. 1972; 34:187–220.
27. Pui CH, Sandlund JT, Pei D, et al. Improved outcome for children with acute lymphoblastic leukemia: results of Total Therapy Study XIII B at St Jude Children's Research Hospital. *Blood*. 2004; 104:2690–2696. [PubMed: 15251979]
28. Pui CH, Boyett JM, Rivera GK, et al. Long-term results of Total Therapy studies 11, 12 and 13A for childhood acute lymphoblastic leukemia at St Jude Children's Research Hospital. *Leukemia*. 2000; 14:2286–2294. [PubMed: 11187920]
29. Silverman LB, Gelber RD, Dalton VK, et al. Improved outcome for children with acute lymphoblastic leukemia: results of Dana-Farber Consortium Protocol 91-01. *Blood*. 2001; 97:1211–1218. [PubMed: 11222362]
30. Pui CH, Campana D, Pei D, et al. Treating childhood acute lymphoblastic leukemia without cranial irradiation. *N Engl J Med*. 2009; 360:2730–2741. [PubMed: 19553647]
31. Bruggemann M, Raff T, Flohr T, et al. Clinical significance of minimal residual disease quantification in adult patients with standard-risk acute lymphoblastic leukemia. *Blood*. 2006; 107:1116–1123. [PubMed: 16195338]

32. Coustan-Smith E, Sancho J, Behm FG, et al. Prognostic importance of measuring early clearance of leukemic cells by flow cytometry in childhood acute lymphoblastic leukemia. *Blood*. 2002; 100:52–58. [PubMed: 12070008]
33. Nyvold C, Madsen HO, Ryder LP, et al. Precise quantification of minimal residual disease at day 29 allows identification of children with acute lymphoblastic leukemia and an excellent outcome. *Blood*. 2002; 99:1253–1258. [PubMed: 11830473]
34. Basso G, Veltroni M, Valsecchi MG, et al. Risk of relapse of childhood acute lymphoblastic leukemia is predicted by flow cytometric measurement of residual disease on day 15 bone marrow. *J Clin Oncol*. 2009; 27:5168–5174.0. [PubMed: 19805690]
35. Conter V, Bartram CR, Valsecchi MG, 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; 115:3206–3214. [PubMed: 20154213]
36. Van Der Velden V, Panzer-Grumayer ER, Cazzaniga G, et al. Optimization of PCR-based minimal residual disease diagnostics for childhood acute lymphoblastic leukemia in a multi-center setting. *Leukemia*. 2007; 21:706–713. [PubMed: 17287857]
37. Bruggemann M, Schrauder A, Raff T, et al. Standardized MRD quantification in European ALL trials: proceedings of the Second International Symposium on MRD assessment in Kiel, Germany, 18–20 September 2008. *Leukemia*. 2010; 24:521–535. [PubMed: 20033054]
38. Neale GA, Coustan-Smith E, Stow P, et al. Comparative analysis of flow cytometry and polymerase chain reaction for the detection of minimal residual disease in childhood acute lymphoblastic leukemia. *Leukemia*. 2004; 18:934–938. [PubMed: 15029212]
39. Robillard N, Cave H, Mechinaud F, et al. Four-color flow cytometry bypasses limitations of IG/TCR polymerase chain reaction for minimal residual disease detection in certain subsets of children with acute lymphoblastic leukemia. *Haematologica*. 2005; 90:1516–1523. [PubMed: 16266899]
40. Mullighan CG, Su X, Zhang J, et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *N Engl J Med*. 2009; 360:470–480. [PubMed: 19129520]
41. Harvey RC, Mullighan CG, Chen IM, et al. Rearrangement of CRLF2 is associated with mutation of JAK kinases, alteration of IKZF1, Hispanic/Latino ethnicity, and a poor outcome in pediatric B-progenitor acute lymphoblastic leukemia. *Blood*. 2010; 115:5312–5321. [PubMed: 20139093]
42. Mullighan CG, Zhang J, Harvey RC, et al. JAK mutations in high-risk childhood acute lymphoblastic leukemia. *Proc Natl Acad Sci U S A*. 2009; 106:9414–9418. [PubMed: 19470474]
43. Coustan-Smith E, Sancho J, Hancock ML, et al. Use of peripheral blood instead of bone marrow to monitor residual disease in children with acute lymphoblastic leukemia. *Blood*. 2002; 100:2399–2402. [PubMed: 12239148]
44. Van Der Velden V, Jacobs DC, Wijkhuijs AJ, et al. Minimal residual disease levels in bone marrow and peripheral blood are comparable in children with T cell acute lymphoblastic leukemia (ALL), but not in precursor-B-ALL. *Leukemia*. 2002; 16:1432–1436. [PubMed: 12145681]

9906 vs 8602/9006 and 9406

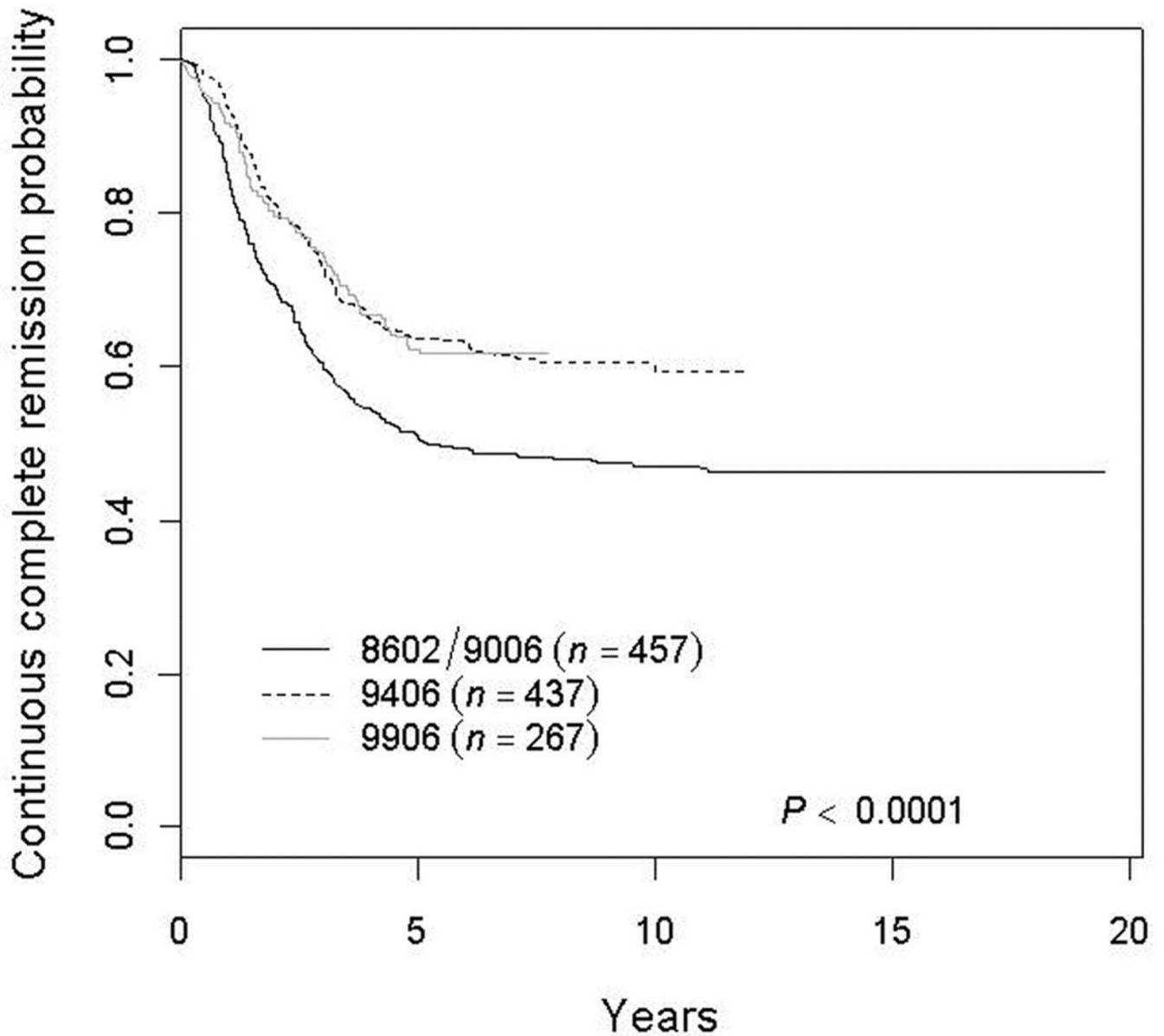
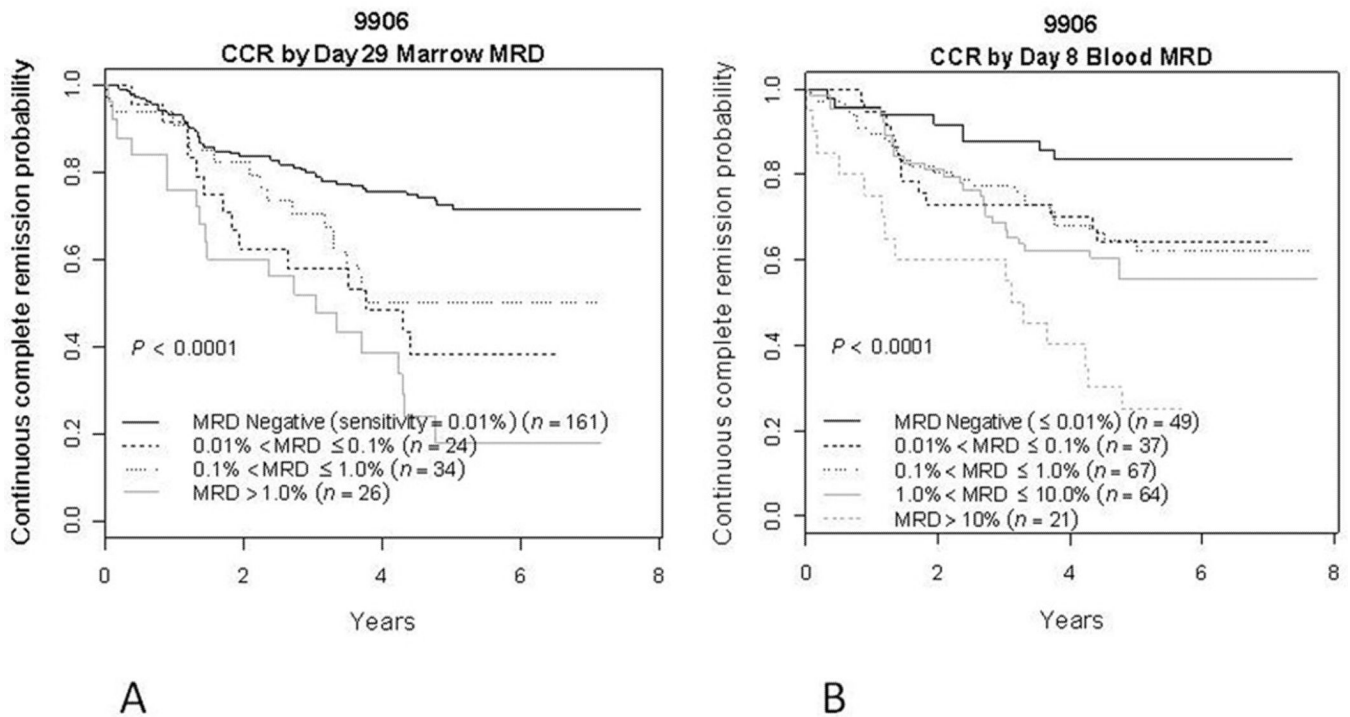


Figure 1.

Kaplan-Meier curves comparing the continuous complete remission rates of patients treated on COG P9906 with that of historical controls (similarly-defined subset of high risk patients treated on prior POG protocols 8602, 9006, and 9406). Study 8602 enrolled 1933 patients with ALL, 251 of which had the comparable high risk characteristics of patients in 9906; study 9006 enrolled 614 patients with high risk ALL, with 235 having matching characteristics; and study 9406 enrolled 786 high risk patients with 437 with comparable characteristics. The 5 y CCR rate on 9906 was $62.2 \pm 3.7\%$ compared to $50.6 \pm 2.4\%$ for the relevant comparison group on 8602/9006 and $63.5 \pm 2.4\%$ on 9406.

**Figure 2.**

Kaplan-Meier curves displaying the outcome of patients based on the level of bone marrow MRD at day 29 of induction therapy (A) and on the level of blood MRD at day 8 of induction therapy. (B) A. Patients with <0.01% marrow MRD had a $72.6 \pm 4.3\%$ 5 year CCR rate while those with increasing levels of MRD generally showed an increasingly bad outcome. The 5 year CCR of all patients positive at 0.01% or greater at day 29 was $37.1 \pm 7.4\%$. B. Patients with <0.01% day 8 blood MRD had a $83.6 \pm 6.3\%$ 5 year CCR rate while those with increasing levels of MRD showed an increasingly bad outcome, with those patients with $1\% < \text{MRD} \leq 10\%$ in the blood at day 8 having a $55.6 \pm 10.0\%$ 5 year CCR rate and those with $>10\%$ having a 5 y CCR rate of $25.1 \pm 12.5\%$. Overall, 80% of patients had blood MRD positive at 1% or greater at day 8, and 5 year CCR rate was $57.1 \pm 4.6\%$ for these patients.

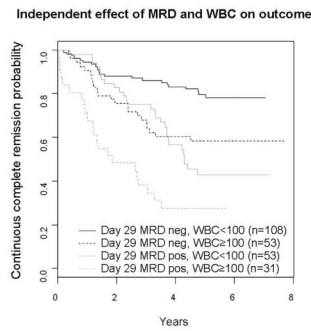


Figure 3.

Independent effect of day 29 MRD and that of initial white blood cell count on the outcome of patients. The 5 y CCR rate for MRD negative patients with $WBC < 100,000/\mu l$ was $79.7 \pm 4.8\%$, while those who were MRD positive with a $WBC \geq 100,000/\mu l$ was $27.7 \pm 11.8\%$. Patients with $< 100,000 WBC/\mu l$ who were MRD positive, or those with $WBC \geq 100,000/\mu l$ and MRD negative had intermediate outcomes with $42.7 \pm 9.3\%$ and $58.4 \pm 8.2\%$ 5 y CCR respectively ($p < 0.0001$). In subset analyses, elevated WBC was associated with significantly worse outcome in both the MRD negative ($p = 0.0049$) and MRD positive patients ($p = 0.0063$). Similarly, MRD positivity was associated with significantly worse outcome in both the lower ($p < 0.0001$) and higher ($p = 0.027$) WBC patients.

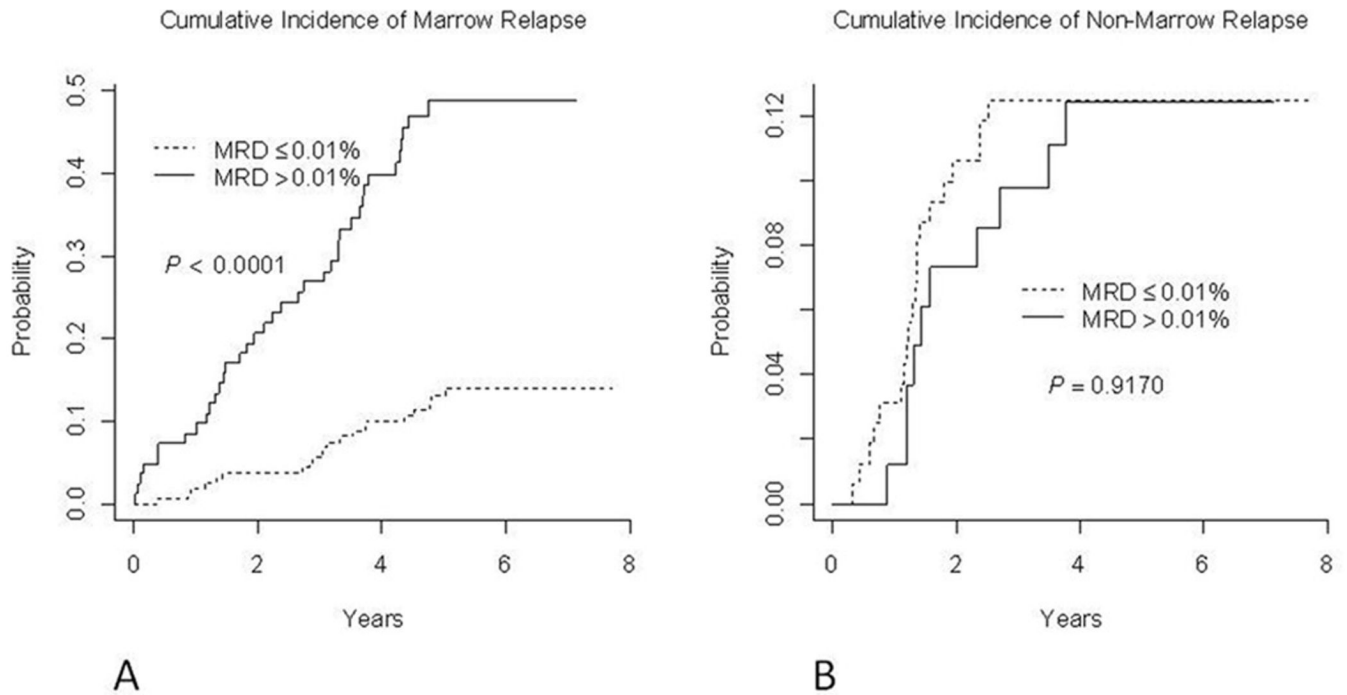


Figure 4. Cumulative incidence of marrow (A) and non-marrow (B) relapse of patients based on end of induction MRD status. Patients with at least .01% marrow MRD had a nearly 4 fold increased risk of marrow relapse at 5 years, at $48.8 \pm 5.8\%$ compared to $13.0 \pm 2.8\%$ for those who were MRD negative. The 5 year cumulative incidence of non marrow relapse was 12% in both the MRD positive and negative groups.

Table I

Treatment scheme for P9906

TREATMENT PHASE/DRUG	DAILY DOSE	DAYS GIVEN
4-DRUG INDUCTION (weeks 1–4)		
PREDNISONE (PO)	40 mg/m ² (MAX 60 mg)	1–28
VINCRIStINE (IV)	1.5 mg/m ² (MAX 2 mg)	1, 8, 15, 22
E COLI L-ASPARAGINASE (IM)	10,000 IU/m ²	2, 5, 8, 12, 15, 19
DAUNORUBICIN (IV)	30 mg/m ²	8, 15, 22
METHOTREXATE (IT)	Age adjusted doses	1, 8; ALSO 15, 22 if CNS 2 or 3
CONSOLIDATION (weeks 6–14)		DAYS GIVEN
METHOTREXATE (IT)	Age adjusted doses	1, 15, 29, 43
6-MERCAPTOPYRINE (PO)	60 mg/m ² /day	1–14, 29–42
CYCLOPHOSPHAMIDE (IV)	1,000 mg/m ²	1, 29
CYTOSINE ARABINOSIDE (SC or IV)	75 mg/m ² /day	2–5, 9–12, 30–33, 37–40
VINCRIStINE (IV)	1.5 mg/m ² (MAX 2 mg)	15, 22, 43, 50
E COLI L-ASPARAGINASE (IM)	6,000 IU/m ²	15, 17, 19, 22, 24, 26, 43, 45, 47, 50, 52, 54
INTERIM MAINTENANCE (weeks 15–22; repeat on weeks 31–38)		DAYS GIVEN
METHOTREXATE (IV)	100 mg/m ² ; escalate by 50 mg/m ² if no toxicities	1, 11, 21, 31, 41
METHOTREXATE (IT)	Age adjusted doses	1, 31
VINCRIStINE (IV)	1.5 mg/m ² (MAX 2 mg)	1, 11, 21, 31, 41
E COLI ASPARAGINASE (IM)	15,000 IU/m ²	2, 12, 22, 32, 42
DELAYED INTENSIFICATION (weeks 23–30; repeat weeks 39–46)		DAYS GIVEN
DEXAMETHASONE (PO)	10 mg/m ²	1–7, 15–21
VINCRIStINE (IV)	1.5 mg/m ² (MAX 2 mg)	1, 8, 15, 43, 50
E COLI ASPARAGINASE (IM)	6,000 IU/m ²	4, 6, 8, 11, 13, 15, 43, 45, 47, 50, 52, 54
DOXORUBICIN (IV)	25 mg/m ²	1, 8, 15
METHOTREXATE (IT)	Age adjusted doses	1, 29

TREATMENT PHASE/DRUG	DAILY DOSE	DAYS GIVEN
THIOGUANINE (PO)	60 mg/m ²	29–42
CYTOSINE ARABINOSIDE (SC or IV)	75 mg/m ²	30–33, 37–40
CYCLOPHOSPHAMIDE (IV)	1,000 mg/m ²	29
CONTINUATION (weeks 47–130; repeated 12 week courses)		DAYS GIVEN
METHOTREXATE	20 mg/m ²	8, 15, 22, 29, 36, 43, 50, 57, 64, 71, 78
6-MERCAPTOPYRINE ((PO)	75 mg/m ²	1–84
IT METHOTREXATE	Age adjusted doses	1 (if patient received cranial XRT, omit IT and give PO dose instead)
VINCRIStINE	1.5 mg/m ² (MAX 2 mg)	1, 29, 57
DEXAMETHASONE	6 mg/m ²	1–5, 29–33, 57–61
CNS IRRADIATION (weeks 47–48)	1800 cGy in 10 fractions for patients with CNS 3 at diagnosis; after May 2004 , recommended for all patients with an initial WBC ≥ 100,000/ul who were within 24 months of diagnosis	

Table II

Summary of Patient Characteristics at diagnosis

Patient Characteristics	N
Gender	
Male	185
Female	82
Race	
White	155
Hispanic or Latino	65
Black or African American	20
Native Hawaiian or other Pacific Islander	2
Asian	10
American Indian or Alaska native	5
Other	5
Unknown	5
CNS status	
CNS 1	208
CNS 2	28
CNS 3	31
Testicular disease	
No	188
Yes	6
NA	73
Age	
<10 yrs	88
10+ yrs	179
WBC	
<50,000/ μ l	149
50,000 – 100,000/ μ l	26
\geq 100,000/ μ l	92
NCI risk group	
Standard risk	27
High risk	240
Congenital abnormality	
Unknown	65
None	192
Down syndrome	8
Other	2
MLL abnormality	
Negative	243

Patient Characteristics	N
Positive	24
Event type	
Died in remission	4
None (continuous complete remission)	168
Relapse: marrow	52
Relapse: isolated CNS	28
Relapse: marrow + CNS	8
Relapse: other - bones (1), marrow+lymph nodes (1), marrow+testes (2), retina (1), testes (1)	6
Relapse: unknown site	1