

CLINICAL SIGNIFICANCE OF MINIMAL RESIDUAL DISEASE IN CHILDHOOD ACUTE LYMPHOBLASTIC LEUKEMIA

HÉLÈNE CAVÉ, PH.D., JUTTE VAN DER WERFF TEN BOSCH, M.D., STEFAN SUCIU, M.S., CHRISTINE GUIDAL, M.S., CHRISTINE WATERKEYN, M.S., JACQUES OTTEN, M.D., MARLEEN BAKKUS, PH.D., KRIS THIELEMANS, M.D., BERNARD GRANDCHAMP, PH.D., M.D., AND ETIENNE VILMER, M.D., FOR THE EUROPEAN ORGANIZATION FOR RESEARCH AND TREATMENT OF CANCER—CHILDHOOD LEUKEMIA COOPERATIVE GROUP

ABSTRACT

Background and Methods The implications of the detection of residual disease after treatment of acute lymphoblastic leukemia (ALL) are unclear. We conducted a prospective study at 11 centers to determine the predictive value of the presence or absence of detectable residual disease at several points in time during the first six months after complete remission of childhood ALL had been induced. Junctional sequences of T-cell-receptor or immunoglobulin gene rearrangements were used as clonal markers of leukemic cells. Residual disease was quantitated with a competitive polymerase-chain-reaction (PCR) assay. Of 246 patients enrolled at diagnosis and treated with a uniform chemotherapy protocol, 178 were monitored for residual disease with one clone-specific probe (in 74 percent) or more than one probe (in 26 percent). The median follow-up period was 38 months.

Results The presence or absence and level of residual leukemia were significantly correlated with the risk of early relapse at each of the times studied ($P < 0.001$). PCR measurements identified patients at high risk for relapse after the completion of induction therapy (those with $\geq 10^{-2}$ residual blasts per 2×10^5 mononuclear bone marrow cells) or at later time points (those with $\geq 10^{-3}$ residual blasts). Multivariate analysis showed that as compared with immunophenotype, age, risk group (standard or very high risk), and white-cell count at diagnosis, the presence or absence and level of residual disease were the most powerful independent prognostic factors.

Conclusions Residual leukemia after induction of a remission is a powerful prognostic factor in childhood ALL. Detection of residual disease by PCR should be used to identify patients at risk for relapse and should be taken into account in considering alternative treatment. (N Engl J Med 1998;339:591-8.)

©1998, Massachusetts Medical Society.

DESPITE advances in the treatment of childhood acute lymphoblastic leukemia (ALL), the risk of relapse remains about 30 percent. Studies have shown that the presence or absence of residual disease, as assessed by the polymerase-chain-reaction (PCR) assay, can serve as a prognostic factor in patients with ALL,¹⁻⁴ and threshold levels of residual leukemic cells have been proposed for predicting relapse.⁵⁻¹⁰ However, many of the studies have been retrospective anal-

yses^{4,5,8} or have involved a small number of patients who were sometimes treated with different protocols. Furthermore, the course of residual disease has varied considerably in some studies.^{4,6-10}

In a pilot study, we validated the method of quantitating residual blasts in the marrow with a competitive PCR assay.¹¹ This method uses the rearranged T-cell-receptor or immunoglobulin heavy-chain genes of the leukemic blasts as clonal markers. We found that quantitation of residual leukemia during the first months of remission can help identify patients who are likely to have a relapse.¹¹ We undertook this prospective study to extend our preliminary findings.

METHODS**Treatment**

We used the Berlin-Frankfurt-Munster (BFM) treatment protocol with minor modifications (European Organization for Research and Treatment of Cancer [EORTC] protocol 58881).¹² In brief, after one week of treatment with prednisolone and one intrathecal injection of methotrexate, induction therapy was begun. It consisted of a five-drug regimen given over a period of four weeks (daily prednisolone, weekly vincristine and daunorubicin, asparaginase twice weekly, and intrathecal methotrexate on days 8 and 22). After the completion of this treatment, patients who had had more than 1000 blasts per cubic millimeter of blood at the end of the first week of prednisolone treatment, those who did not have a complete remission, and those with the $t(4;11)$ or $t(9;22)$ translocation were classified as having a very high risk of relapse. All other patients were classified as having a standard risk.

Patients with a standard risk of relapse received four weeks of consolidation therapy, consisting of daily mercaptopurine, four four-day courses of cytarabine, and cyclophosphamide on days

From the Laboratoire de Biochimie Génétique and the Service d'Hématologie-Immunologie, Hôpital Robert Debré, Paris (H.C., C.G., B.G., E.V.); the Department of Physiology, Vrije Universiteit Brussel, Brussels, Belgium (J.W.B., M.B., K.T.); the European Organization for Research and Treatment of Cancer Data Center, Brussels, Belgium (S.S., C.W.); and the Akademisch Ziekenhuis, Vrije Universiteit Brussel, Brussels, Belgium (J.O.). Address reprint requests to Dr. Vilmer at the Service d'Hématologie-Immunologie, Hôpital Robert Debré, 48 Blvd. Serurier, 75019 Paris, France.

Other authors were Brigitte Nelken and Martine Fournier (Centre Hospitalier Universitaire, Lille, France), Patrick Boutard and Emmanuel Lebrun (Centre Hospitalier Universitaire, Caen, France), Françoise Méchinaud and Richard Garand (Centre Hospitalier Régional Hôtel Dieu, Nantes, France), Alain Robert and Nicole Dastugue (Centre Hospitalier Universitaire, Toulouse, France), Emmanuel Plouvier and Evelyne Racadot (Centre Hospitalier Régional, Besançon, France), Alice Ferster (Centre Hospitalier Universitaire, Reine Fabiola, Brussels, Belgium), Jan Gyselincx (Algemeen Kinderziekenhuis, Antwerp, Belgium), Odile Fenneteau and Michel Duval (Hôpital Robert Debré, Paris), Gabriel Solbu (European Organization for Research and Treatment of Cancer Data Center, Brussels, Belgium), and Anne-Marie Manel (Centre Hospitalier Régional, Lyons, France).

1 and 28. This consolidation phase was followed by an eight-week course of daily mercaptopurine and four courses of high-dose methotrexate (interval therapy). A delayed intensification phase consisted of dexamethasone (for 3 weeks), four weekly injections of vincristine and doxorubicin, and four injections of asparaginase (protocol IIA), followed by daily thioguanine (for 14 days), one injection of cyclophosphamide, two courses of cytarabine, and one intrathecal injection of methotrexate (protocol IIB). The duration of treatment from the start of induction therapy to the completion of the delayed intensification phase was about 27 weeks. Delayed intensification therapy was followed by maintenance treatment consisting of daily mercaptopurine and weekly methotrexate. The total duration of treatment was two years.

Patients at very high risk for relapse received intensified consolidation therapy of six weeks' duration, consisting of cyclophosphamide, high-dose methotrexate and cytarabine, asparaginase, and oral mercaptopurine, followed by two series of three chemotherapeutic courses according to the BFM relapse protocol.¹³

Detection of Residual Disease

Bone marrow mononuclear cells were counted, lysed, and stored at -20°C until analysis. Rearrangements of the T-cell-receptor genes *TCR γ* (V1-J1,2; V1-JP1,2; and V9-J1,2) and *TCR δ* (V2-D3, V1-J1, D2-D3, and V2-J1) were sought in samples obtained at the time of diagnosis.^{11,14} If none of these rearrangements were detected, rearrangements of the gene for immunoglobulin heavy chain (*IgH*) were sought with the use of consensus FRIII and JH primers.¹⁵ The presence or absence of such clonal markers was determined after polyacrylamide-gel electrophoresis. Discrete bands of PCR products corresponding to clonal rearrangements were sequenced. An oligonucleotide probe specific for the junctional sequence was synthesized for each rearrangement. Tests for residual disease were conducted by PCR amplification of 2×10^5 mononuclear bone marrow cells in samples obtained during remission, with the use of the primer set corresponding to the T-cell or B-cell clonal rearrangement identified at the time of diagnosis. PCR products were dot-blotted and hybridized to the radiolabeled clone-specific probe.¹¹

All frozen samples obtained at all time points from a given patient were run at the same time. The specificity of detection was checked for each probe on at least two different polyclonal samples. The sensitivity of each probe was assessed by testing serial dilutions of the patient's blasts in a mixture of polyclonal marrow mononuclear cells. The median level of detection was 5×10^{-5} (i.e., 5 blasts per 100,000 normal mononuclear cells). For the statistical analyses, the results were considered negative only if the level detected was less than 1.5×10^{-4} .

We used a competitive PCR assay to quantitate residual blasts. Amplification was carried out as for blast detection in the presence of 100 copies of internal standard in each PCR sample. Internal standards consisted of DNA from monoclonal cells that had a rearrangement involving the same genomic segments as the patient's blasts but with a distinct junctional sequence. Each PCR series included serial dilutions of the patient's blasts. PCR products were hybridized in duplicate with clone-specific probes corresponding to the patient's blasts and to the internal standard. The ratio of the radioactivity of the two probes was calculated. A calibration curve was drawn from the results obtained with the serial dilutions. In samples obtained during remission, the number of blasts per 100,000 mononuclear marrow cells was derived from the calibration curve. Replicate assays gave results with a standard deviation of 15 to 30 percent of the mean value. When two different markers (e.g., one *TCR δ* V2-D3 rearrangement and one *TCR γ* V9-J1 rearrangement) were analyzed to quantitate residual leukemic cells in the same sample, the results were closely correlated (correlation coefficient for 28 samples, 0.94).

Study Design

The 11 centers participating in the study enrolled all their patients at the time of diagnosis, after obtaining informed consent.

Bone marrow samples from all patients were obtained at the time of diagnosis and at the end of induction therapy. In the standard-risk group, samples were obtained after consolidation, interval, and delayed intensification treatments. In the very-high-risk group, samples were obtained on completion of intensified consolidation therapy, which was given after induction therapy. Bone marrow samples were analyzed at one of two reference laboratories (in Brussels, Belgium, and in Paris) for the detection of residual disease. Quantitative analysis was performed at a single laboratory (in Paris) for all samples in which residual leukemia was detected. Personnel at both laboratories were unaware of the patients' status at the time the samples were assayed. Clinical and molecular data were centralized at the EORTC data center, where the statistical analysis was performed.

Patients

A total of 246 children with ALL were enrolled in the study at the time of diagnosis. The enrollment period started in November 1989 (at 1 center) or July 1993 (at 10 centers) and ended in March 1996. Four patients were excluded because they did not have a remission, defined by the detection of fewer than 5 percent blasts in bone marrow smears, which were independently reviewed by two cytologists. At least three bone marrow samples from each patient were studied, with the exception of patients who had relapses before delayed intensification therapy was completed. Sixteen patients were excluded because fewer than three follow-up bone marrow samples were obtained.

Among the remaining 226 patients, no gene rearrangements were detected in 25 (11 percent), whereas at least one rearrangement was detected in 201 (89 percent). *TCR δ* was rearranged in 115 of 188 patients (61 percent) with B-lineage ALL and in 14 of 38 (37 percent) with T-lineage ALL. *TCR γ* was rearranged in 108 patients (57 percent) with B-lineage ALL and in 32 (84 percent) with T-lineage ALL. An *IgH* rearrangement was detected in 12 of 32 patients (38 percent) with B-lineage ALL in whom no *TCR δ* or *TCR γ* rearrangement was detected.

In 23 of 201 patients with at least one rearrangement, no probe could be obtained because of an oligoclonal pattern of rearrangement or biallelic rearrangements that were unsuitable for good electrophoretic separation. At least one clone-specific probe was available for 178 patients, who formed the study group. Residual disease was evaluated with the use of a single probe in 132 patients (74 percent) and with two or more probes in 46 patients (26 percent). One or more *TCR δ* probes were used alone in 64 patients (36 percent), one or more *TCR γ* probes were used alone in 83 patients (47 percent), both *TCR δ* and *TCR γ* probes were used in 19 patients (11 percent), and one or more *IgH* probes were used alone in 12 patients (7 percent). In 18 patients, residual disease was detected but could not be quantitated because the samples were inadequate. Data for the 25 patients in the pilot study¹¹ were included in the statistical analysis.

The comparability of our patients with the general population of children with ALL was evaluated by comparing the outcomes in our group with those among the 654 children treated during the same period in the Childhood Leukemia Cooperative Group centers that did not participate in the study of residual disease (Table 1).

Statistical Analysis

The principal end point used to determine the prognostic value of the presence or absence of residual disease was the relapse-free interval, which was calculated as the interval from the time of a given assessment of residual disease until the date of the first relapse. Actuarial curves were computed according to the Kaplan-Meier method.¹⁶ The prognostic value of the variables studied was assessed with the use of the log-rank test,¹⁷ or the log-rank test for linear trend in the case of ordered categorical variables. The relative risk represented the ratio of the daily risk of relapse in patients with residual disease to the risk in patients without residual disease or the ratio of the daily risk of relapse in those with

TABLE 1. PROGNOSTIC FACTORS AND OUTCOMES IN THE STUDY GROUP AND AMONG PATIENTS AT OTHER CENTERS.*

VARIABLE	STUDY GROUP			PATIENTS AT OTHER CENTERS		
	NO. OF PATIENTS (%)	RELATIVE RISK†	P VALUE‡	NO. OF PATIENTS (%)	RELATIVE RISK†	P VALUE‡
Total	178 (100)			654 (100)		
Age (yr)			0.10			<0.001
0-1	10 (6)	1.7		70 (11)	2.4	
2-9	145 (81)	1.0		470 (72)	1.0	
10-15	23 (13)	2.2		114 (17)	2.3	
Risk group			0.07			<0.001
Standard risk	162 (91)	1.0		563 (86)	1.0	
Very high risk	16 (9)§	2.2		91 (14)	2.9	
Immunophenotype			0.01			0.50
B-lineage	149 (84)	1.0		565 (86)	1.0	
T-lineage	29 (16)	2.4		89 (14)	1.2¶	
White-cell count			0.05			<0.001
<10,000/mm ³	76 (43)	1.0		310 (47)	1.0	
10,000-99,000/mm ³	81 (46)	1.4		261 (40)	2.0	
≥100,000/mm ³	21 (12)	2.6		83 (13)	4.7	
Karyotype			0.08			0.40
Normal	55 (31)	3.0		155 (24)	1.9	
Hyperdiploidy (>50 chromosomes)	33 (19)	1.0		49 (7)	1.0	
Other	50 (28)	1.8		284 (43)	1.8	
Not available	40 (22)	—		166 (25)	—	
Outcome						
Continuous remission	140 (79)			531 (81)		
Marrow relapse	22 (12)			63 (10)		
Extramedullary relapse	9 (5)			20 (3)		
Combined relapse	7 (4)			25 (4)		
Death during remission	0 (0)			15 (2)		

*Percentages may not sum to 100 because of rounding. The patients at other centers were treated with the same chemotherapeutic regimen as the study group, but these centers did not participate in the analysis of residual disease.

†The relative risk was estimated by calculating the ratio of observed to expected relapses.

‡P values were determined by the log-rank test. Because of the small number of relapses in the study group (38), the P value did not reach statistical significance for several comparisons.

§Fifteen of the patients had poor responses to the corticosteroid therapy, and one had the t(9;22) translocation.

¶The prognosis was slightly worse for the T-lineage group, particularly because of their higher relapse rate during therapy.

||The instantaneous risk of relapse was similar in the two groups of patients (relative risk in the study group as compared with the patients at other centers, 1.19). For the analysis of outcomes, data for patients who died while in remission were censored at the time of death.

a high level of residual disease to the risk in those with a low level. This was estimated by calculating the ratio of observed to expected relapses¹⁸ with the use of log-rank computations.

The prognostic significance of other variables measured at the time of diagnosis was determined in the same way. Subgroups of patients were defined according to classic prognostic factors (Table 1). The stratified log-rank test was used to determine the prognostic value of residual disease as compared with other prognostic factors, and the corresponding stratified relative risks were computed. Cytogenetic characteristics were not included in the stratified analyses because of the high percentage of patients in whom they could not be evaluated (22 percent). The Cox regression model¹⁹ was used to determine the most significant independent prognostic factors. The stratified Cox regression model¹⁹ was used to determine the prognostic value of residual disease as compared with that of immunophenotype. This method provided an estimate of the relative risk and 95 percent confidence interval.

RESULTS

Of the 246 patients enrolled in the study at the time of diagnosis, 178 (72 percent) were monitored

for residual disease. These 178 patients were similar to the 654 patients in the nonparticipating centers with regard to the time to relapse and the distribution of prognostic factors (Table 1). The effects of prognostic factors were also similar in the two groups, except for immunophenotype. The duration of remission in the group of 68 patients who were enrolled at the time of diagnosis but were subsequently excluded from the study was similar to that in the 178 patients who remained in the study. The median follow-up period was 38 months.

Residual Disease during the First Six Months of Remission

After the completion of induction therapy, 42 percent of the patients had residual leukemia (Table 2). Residual blasts were detected in 38 percent of the standard-risk group and 66 percent of the very-high-risk group; they were detected in 36 percent of

TABLE 2. RISK OF RELAPSE ACCORDING TO THE PRESENCE OR ABSENCE AND LEVEL OF RESIDUAL DISEASE AT DIFFERENT TIMES AFTER INDUCTION OF REMISSION.

RESIDUAL DISEASE*	NO. OF PATIENTS/ TOTAL No. (%)†	NO. WITH RELAPSES	RELATIVE RISK‡
After induction therapy			
Absent	88/151 (58)	7	1.0
Present	63/151 (42)	25	5.7
Residual blasts			
<10 ⁻³	109/133 (82)	10	1.0
10 ⁻³ to <10 ⁻²	9/133 (7)	2	2.3
≥10 ⁻²	15/133 (11)	11	16.0§
After consolidation therapy			
Absent	95/127 (75)	8	1.0
Present	32/127 (25)	15	7.3
Residual blasts			
<10 ⁻³	110/118 (93)	11	1.0
≥10 ⁻³	8/118 (7)	6	15.3
After interval therapy			
Absent	108/130 (83)	11	1.0
Present	22/130 (17)	11	7.3
Residual blasts			
<10 ⁻³	118/127 (93)	13	1.0
≥10 ⁻³	9/127 (7)	7	18.5
After delayed intensification therapy			
Absent	107/123 (87)	14	1.0
Present	16/123 (13)	10	9.2
Residual blasts			
<10 ⁻³	114/119 (96)	17	1.0
≥10 ⁻³	5/119 (4)	5	22.0

*The number of residual blasts is expressed as the ratio of residual blasts to 2×10^5 mononuclear bone marrow cells.

†The numbers of patients studied were different at different points in time for two reasons: samples were not obtained from all patients at each time point, and at the end of induction therapy, the study group consisted of patients at standard risk and those at very high risk, whereas the study group consisted only of patients at standard risk at the other time points.

‡The relative risk was estimated by calculating the ratio of observed to expected relapses. For all comparisons, $P < 0.001$ by the log-rank test.

§The relative risk was 14.2 for patients who had $\geq 10^{-2}$ residual blasts as compared with those who had $< 10^{-2}$ residual blasts.

the group with B-lineage ALL and 90 percent of the group with T-lineage ALL. In the standard-risk group, 25 percent of the patients had detectable residual disease after consolidation therapy, 17 percent after interval therapy, and 13 percent after delayed intensification therapy.

Table 3 shows the predictive value of a change in status with respect to detectable residual disease at two different points in time. The relative risk of relapse was 4.9 for the patients who had residual disease initially but did not have residual disease after consolidation therapy and 15.0 for the patients with persistent residual disease after consolidation therapy, as compared with the patients in whom residual disease was undetectable after both induction and consolidation therapy. The results were similar at subsequent time points. During the first six months after induction therapy, no patient who initially had no detectable disease subsequently had detectable disease. However, six patients in whom the level of

residual disease increased during the first 6 months had a relapse 1 to 15 months later.

Predictive Value of Residual Disease at the End of Induction Therapy

The patients with detectable residual disease at the end of induction therapy, including those with a standard risk of relapse and those with a very high risk, had a significantly shorter time to relapse than the patients with undetectable residual disease ($P < 0.001$ by the log-rank test), and the instantaneous risk of relapse was 5.7 times as high in the patients with residual disease as in those without residual disease (Fig. 1 and Table 2). To determine whether the level of residual disease could be used to predict relapse, we assigned the patients to one of three subgroups according to the concentration of residual leukemic cells in marrow samples: less than 10^{-3} , 10^{-3} or more but less than 10^{-2} , and 10^{-2} or more. The probability of relapse increased with the level of residual disease. Patients with 10^{-2} or more residual blasts had a shorter time to relapse than those with lower levels of blasts ($P < 0.001$ by the log-rank test) (Fig. 1), and the relative risk of relapse was 16 times as high in the patients with 10^{-2} or more blasts as in those with less than 10^{-3} blasts.

The duration of survival was also related to the presence or absence and level of residual disease. Patients with residual disease had a risk of death that was 10 times that in patients without detectable residual disease. The quantitative analysis was even more predictive: the risk of death was 24 times as high for patients with 10^{-2} or more residual leukemic cells as for patients with lower levels of residual disease (data not shown).

Predictive Value of Residual Disease at Later Times

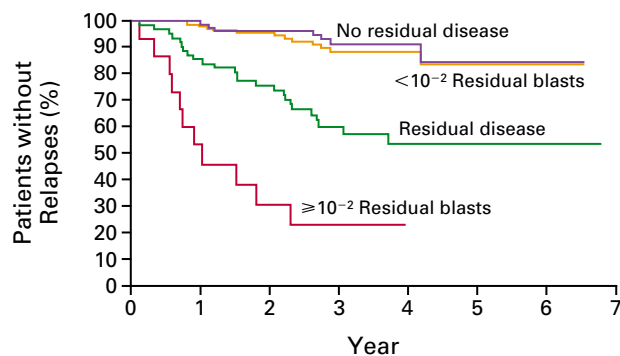
At three different time points, the time to relapse in the standard-risk group was significantly shorter for patients with detectable residual disease than for those without detectable residual disease ($P < 0.001$ by the log-rank test) (Fig. 2). The risk of relapse was significantly higher in patients with detectable disease than in those with undetectable disease (7.3 times as high after consolidation and interval therapy and 9.2 times as high after delayed intensification therapy) (Table 2). Quantitative assessment of residual leukemia allowed further discrimination: a value at or above a threshold of 10^{-3} residual leukemic cells was highly predictive of relapse at all three time points ($P < 0.001$ by the log-rank test) (Fig. 2). The risk of relapse increased by a factor of 15.3 to 22.0 in patients with 10^{-3} or more residual blasts (representing 4 to 7 percent of the patients), as compared with those with fewer than 10^{-3} residual blasts (Table 2). The risk of death was increased by a factor of approximately 25 in patients with 10^{-3} or more residual blasts at each time point (data not shown).

TABLE 3. RELATIVE RISK OF RELAPSE ACCORDING TO THE PRESENCE OR ABSENCE OF RESIDUAL DISEASE AT TWO TIME POINTS.

RESIDUAL DISEASE*	AFTER INDUCTION THERAPY, AFTER CONSOLIDATION THERAPY			AFTER INDUCTION THERAPY, AFTER INTERVAL THERAPY			AFTER CONSOLIDATION THERAPY, AFTER INTERVAL THERAPY		
	NO. OF PATIENTS	NO. WITH RELAPSES	RELATIVE RISK†	NO. OF PATIENTS	NO. WITH RELAPSES	RELATIVE RISK†	NO. OF PATIENTS	NO. WITH RELAPSES	RELATIVE RISK†
Absent, absent	73	3	1.0	78	4	1.0	91	7	1.0
Present, absent	15	3	4.9	23	5	4.1	8	2	3.0
Present, present	32	15	15.0	22	11	14.0	22	11	9.6

*During the first six months, residual disease was not detected in any of the patients who did not previously have detectable residual disease.

†The relative risk was estimated by calculating the ratio of observed to expected relapses. For all comparisons, $P < 0.001$ by the log-rank test for linear trend.



	NO. OF RELAPSES	NO. OF PATIENTS AT RISK						
No residual disease	7	88	86	74	44	18	1	1
Residual disease	25	63	53	42	25	10	4	1
$< 10^{-2}$ Residual blasts	12	118	115	98	58	25	2	1
$\ge 10^{-2}$ Residual blasts	11	15	7	4	3	0	0	0

Figure 1. Kaplan–Meier Estimates of the Relapse-free Interval According to the Presence or Absence and Level of Residual Disease in Patients with a First Complete Remission of ALL at the End of Induction Therapy.

$P < 0.001$ for the comparison between patients with residual disease and those without residual disease and for the comparison between patients with $\ge 10^{-2}$ residual blasts and those with $< 10^{-2}$ residual blasts. Nine of the 15 patients with a high level of residual disease ($\ge 10^{-2}$ blasts) died, as compared with only 4 of the 118 with a lower level of residual disease ($< 10^{-2}$ blasts). The numbers of patients shown below the graph are the numbers at standard or very high risk for whom bone marrow samples were available. In 18 patients, residual disease was detected but was not quantitated.

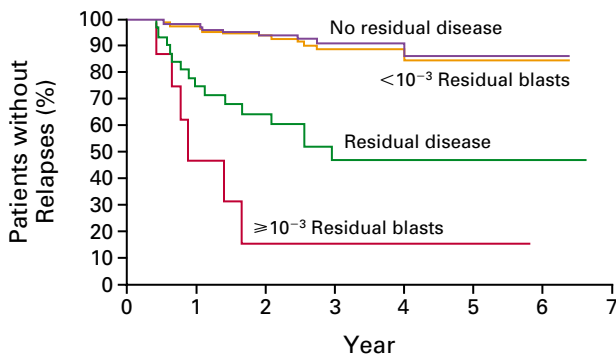
In the very-high-risk group, patients without detectable residual disease after intensified consolidation therapy had a lower probability of relapse than those with detectable disease ($P = 0.03$ by the log-rank test).

Predictive Value of Residual Disease after Stratification for Other Factors

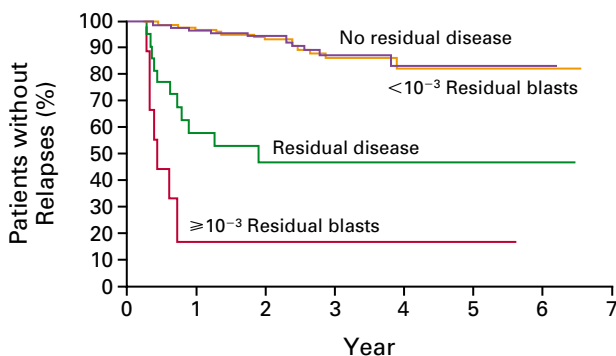
The T-lineage immunophenotype, a white-cell count of 100,000 per cubic millimeter or higher, an age of 10 to 15 years, and assignment to the very-high-risk group (which accounted for 9 to 16 percent of the patients monitored for residual disease)

were associated with the poorest outcome, with a relative risk of relapse ranging from 2.18 to 2.58 (Table 1). Bivariate analyses showed that the presence or absence and level of residual disease at different time points remained significant prognostic factors after stratification for white-cell count, immunophenotype, risk group, and age (Table 4). With the use of the stratified log-rank method, the estimated relative risk of relapse was about 5 for the patients with residual disease and more than 5 for those with 10^{-2} or more residual blasts after induction and 10^{-3} or more subsequently (Table 4).

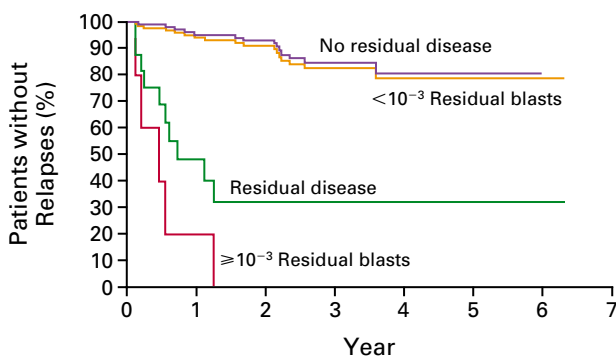
In a Cox model, residual disease remained the



	NO. OF RELAPSES	NO. OF PATIENTS AT RISK						
No residual disease	8	95	92	76	46	20	3	2
Residual disease	15	32	23	17	10	3	3	1
$<10^{-3}$ Residual blasts	11	110	106	88	54	21	4	2
$\geq 10^{-3}$ Residual blasts	6	8	3	1	1	1	1	0



	NO. OF RELAPSES	NO. OF PATIENTS AT RISK						
No residual disease	11	108	101	85	48	14	2	2
Residual disease	11	22	12	8	5	3	2	1
$<10^{-3}$ Residual blasts	13	118	111	91	52	16	3	3
$\geq 10^{-3}$ Residual blasts	7	9	1	1	1	1	1	0



	NO. OF RELAPSES	NO. OF PATIENTS AT RISK						
No residual disease	14	107	97	76	37	16	3	0
Residual disease	10	16	7	3	2	1	1	1
$<10^{-3}$ Residual blasts	17	114	101	78	39	17	4	1
$\geq 10^{-3}$ Residual blasts	5	5	1	0	0	0	0	0

TABLE 4. RELATIVE RISK OF RELAPSE ACCORDING TO THE PRESENCE OR ABSENCE AND LEVEL OF RESIDUAL DISEASE AND OTHER PROGNOSTIC FACTORS.*

RESIDUAL DISEASE†	RISK GROUP	AGE	WHITE-CELL COUNT	IMMUNOPHENOTYPE	IMMUNOPHENOTYPE (COX MODEL)
relative risk‡					
After induction therapy					
Present versus absent	5.2	5.2	4.8	4.3	5.3 (2.2–12.6)
≥10 ⁻² versus <10 ⁻² residual blasts	12.8	10.7	8.1	5.7	10.6 (3.9–28.7)
After consolidation therapy					
Present versus absent	—	6.5	6.2	5.3	6.1 (2.5–14.9)
≥10 ⁻³ versus <10 ⁻³ residual blasts	—	14.4	5.6	7.8	11.2 (3.6–34.7)
After interval therapy					
Present versus absent	—	6.5	5.3	5.4	6.5 (2.7–15.6)
≥10 ⁻³ versus <10 ⁻³ residual blasts	—	12.5	6.1	9.2	—§
After delayed intensification therapy					
Present versus absent	—	7.8	5.5	5.2	7.9 (3.1–19.8)
≥10 ⁻³ versus <10 ⁻³ residual blasts	—	22.3	5.1	6.0	14.7 (3.2–64.7)

*The cutoff points for categorical variables were identical to those in Table 1 except for age (2 to 9 years vs. others); upper and lower age groups (<2 years and ≥10 years) were pooled to obtain a larger group with a poor prognosis. Analysis according to risk group (standard or very high risk) was carried out for all patients at the end of induction therapy. For later time points, stratified and regression analyses were performed only for patients at standard risk.

†The number of residual blasts is expressed as the ratio of residual blasts to 2×10⁵ mononuclear bone marrow cells.

‡For all four prognostic factors, the relative risk was estimated by calculating the ratio of observed to expected differences. For immunophenotype, the relative risk was also estimated with the stratified Cox model, as shown in the right-hand column, with 95 percent confidence intervals in parentheses.

§The relative risk was not estimated because of problems with the convergence of values.

most important prognostic factor, followed by either immunophenotype or white-cell count (data not shown). Since the immunophenotype was a significant prognostic factor and was closely correlated with the level of residual disease after induction therapy, the Cox model was stratified according to immunophenotype to assess the relative prognostic importance of the subsequent evaluations of residual disease (Table 4). The relative risks calculated with this model were higher than those based on the ratio of observed to expected relapses. However, the 95 percent confidence intervals for the Cox-model relative risks spanned the ratios of observed to expect-

ed relapses, indicating that the two methods give consistent results (Table 4). The lower limits of these confidence intervals were markedly higher than 1, confirming that the presence or absence and level of residual disease were important independent prognostic factors.

DISCUSSION

We found that the use of PCR to detect small numbers of leukemic cells remaining in the bone marrow after the induction of a remission by chemotherapy can predict relapse. Since the clinical outcome was similar in the 178 patients who were analyzed for residual disease and the 654 patients registered in the same trial but not enrolled in the study of residual disease, we believe that our conclusions have general applicability to patients with ALL.

Residual disease was detected in about 40 percent of patients after the completion of induction therapy. After consolidation and interval treatment, the proportion of patients with detectable residual disease decreased. Delayed intensification therapy had a limited effect on eliminating residual disease, perhaps because of resistance to chemotherapy. Our results differ from those of studies showing that leukemic cells persist in most patients during the first six months of treatment.^{4,7,9,10,20} However, there are differences in the sensitivity of detection methods. For

Figure 2. Kaplan–Meier Estimates of the Relapse-free Interval in Patients with ALL at Standard Risk, According to the Presence or Absence and Level of Residual Disease after Consolidation Therapy (Top Panel), Interval Therapy (Middle Panel), and Delayed Intensification Therapy (Bottom Panel).

P<0.001 for the comparison between patients with residual disease and those without residual disease and for the comparison between patients with ≥10⁻³ residual blasts and those with <10⁻³ residual blasts. The majority of patients with ≥10⁻³ blasts died: four of eight after consolidation therapy, six of nine after interval therapy, and five of five after delayed intensification therapy. For each point in time, the numbers of patients shown below the graph are the numbers at standard risk for whom bone marrow samples were available. In 18 patients, residual disease was detected but was not quantitated.

example, in the study by Roberts et al.,²⁰ a detection level of about 5×10^{-6} was achieved by testing a large number of cells,²¹ thus resulting in a longer period during which residual leukemia could be detected. Our study also differs from the work of Roberts et al. with regard to the statistical methods used and the characteristics of the patients. Their small sample (25 patients) may not be representative; none of their patients, for example, had an early relapse. In our view, the numbers of measurements made in the early stages of remission in the study by Roberts et al.²⁰ were too small for an accurate assessment of the predictive value of the level of residual disease.

In our study, the risk of relapse was markedly increased in patients with 10^{-2} or more residual leukemic cells at the end of induction therapy. This cutoff value was below the limit of detection with conventional microscopical examination of bone marrow by two experienced cytologists. For the additional three time points, a value at or above a cutoff of 10^{-3} leukemic blasts was highly predictive of relapse.

The presence or absence and level of residual leukemic cells were predictive of survival at all four time points. Our results show that residual leukemia is especially predictive of relapses during therapy. Under these circumstances, the prognosis is particularly poor.

Stratified and multivariate analyses showed that the presence or absence and level of residual disease remained significant prognostic factors even when other known prognostic factors, such as immunophenotype, white-cell count, age, and risk group, were taken into consideration. However, the occurrence of individual relapses was not always predicted. The failure to predict relapses in patients without detectable residual disease was probably not due to inadequate sensitivity of the PCR assay (median level of detection, 5×10^{-5}), since the relapse rate was similar in the group of patients with a level of residual disease below 10^{-3} and in those without detectable disease. It is likely that these relapses resulted from the emergence of a malignant subclone with resistance to chemotherapy.

If the detection of residual leukemia is to be used in clinical practice to identify patients with a high probability of early relapse, two conditions must be met: the analysis should be performed as early as possible, and the laboratory technique should be simple and rapid so that treatment can be tailored to the adjusted assessment of risk. In this respect, the thresholds that we found predicted relapse can now be reached with simpler and more rapid techniques, such as use of the leukemia-associated immunophenotype to detect residual cells²² or fluorescence-PCR analysis of gene rearrangements.²³

Supported by grants from the Association pour la Recherche sur le Cancer and the National Cancer Institute (5U10-CA11488-23 through 5U10-CA11488-28). Drs. van der Werff ten Bosch and Bakkus are the recipients of research awards from the Belgian National Fund for Scientific Research.

The views expressed in this article are solely those of the authors and do not represent the official views of the National Cancer Institute.

REFERENCES

1. d'Auriol L, Macintyre E, Galibert F, Sigaux F. In vitro amplification of T cell γ gene rearrangements: a new tool for the assessment of minimal residual disease in acute lymphoblastic leukemias. *Leukemia* 1989;3:155-8.
2. Macintyre EA, d'Auriol L, Duparc N, Leverger G, Galibert F, Sigaux F. Use of oligonucleotide probes directed against T cell antigen receptor gamma delta variable-(diversity)-joining junctional sequences as a general method for detecting minimal residual disease in acute lymphoblastic leukemias. *J Clin Invest* 1990;86:2125-35.
3. Hansen-Hagge T, Yokota S, Bartram CR. Detection of minimal residual disease in acute lymphoblastic leukemia by in vitro amplification of rearranged T-cell receptor δ chain sequences. *Blood* 1989;74:1762-7.
4. Yamada M, Wasserman R, Lange B, Reichard BA, Womer RB, Rovera G. Minimal residual disease in childhood B-lineage lymphoblastic leukemia: persistence of leukemic cells during the first 18 months of treatment. *N Engl J Med* 1990;323:448-55.
5. Campana D, Pui CH. Detection of minimal residual disease in acute leukemia: methodologic advances and clinical significance. *Blood* 1995;85:1416-34.
6. Wasserman R, Galili N, Ito Y, et al. Residual disease at the end of induction therapy as a predictor of relapse during therapy in childhood B-lineage acute lymphoblastic leukemia. *J Clin Oncol* 1992;10:1879-88.
7. Potter MN, Steward CG, Oakhill A. The significance of detection of minimal residual disease in childhood acute lymphoblastic leukaemia. *Br J Haematol* 1993;83:412-8.
8. Brisco MJ, Condon J, Hughes E, et al. Outcome prediction in childhood acute lymphoblastic leukaemia by molecular quantification of residual disease at the end of induction. *Lancet* 1994;343:196-200.
9. Yokota S, Hansen-Hagge TE, Ludwig WD, et al. Use of polymerase chain reactions to monitor minimal residual disease in acute lymphoblastic leukemia patients. *Blood* 1991;77:331-9.
10. Nizet Y, Van Daele S, Lewalle P, et al. Long-term follow-up of residual disease in acute lymphoblastic leukemia patients in complete remission using clonogenic IgH probes and the polymerase chain reaction. *Blood* 1993;82:1618-25.
11. Cavé H, Guidal C, Rohrlach P, et al. Prospective monitoring and quantitation of residual blasts in childhood acute lymphoblastic leukemia by polymerase chain reaction study of δ and γ T-cell receptor genes. *Blood* 1994;83:1892-902.
12. Reiter A, Schrappe M, Ludwig W-D, et al. Chemotherapy in 998 unselected childhood acute lymphoblastic leukemia patients: results and conclusions of the multicenter trial ALL-BFM 86. *Blood* 1994;84:3122-33.
13. Henze G, Buchmann S, Fengler R, Hartmann R. The BFM relapse studies in childhood ALL: concepts of two multicenter trials and results after 2 1/2 years. *Haematol Blood Transfus* 1987;30:147-55.
14. Breit T, Wolvers-Tettero LI, Hählen K, van Wering ER, van Dongen JJ. Limited combinatorial repertoire of $\gamma\delta$ T-cell receptors expressed by T-cell acute lymphoblastic leukemias. *Leukemia* 1991;5:116-24.
15. Trainor KJ, Brisco MJ, Wan JH, Neoh S, Grist S, Morley AA. Gene rearrangement in B- and T-lymphoproliferative disease detected by the polymerase chain reaction. *Blood* 1991;78:192-6.
16. Peto J. The calculation and interpretation of survival curves. In: Buyse ME, Staquet MJ, Sylvester RJ, eds. *Cancer clinical trials: methods and practice*. Oxford, England: Oxford University Press, 1984:361-80.
17. Breslow N. Comparison of survival curves. In: Buyse ME, Staquet MJ, Sylvester RJ, eds. *Cancer clinical trials: methods and practice*. Oxford, England: Oxford University Press, 1984:381-406.
18. 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.
19. Cox DR. Regression models and life-tables. *J R Soc Stat [B]* 1972;34:187-220.
20. Roberts WM, Estrov Z, Ouspenskaia MV, Johnston DA, McClain KL, Zipf TE. Measurement of residual leukemia during remission in childhood acute lymphoblastic leukemia. *N Engl J Med* 1997;336:317-23.
21. Ouspenskaia MV, Johnston DA, Roberts WM, Estrov Z, Zipf TE. Accurate quantitation of residual B-precursor acute lymphoblastic leukemia by limiting dilution and a PCR-based detection system: a description of the method and the principles involved. *Leukemia* 1995;9:321-8.
22. Coustan-Smith E, Behm FG, Sanchez J, et al. Immunological detection of minimal residual disease in children with acute lymphoblastic leukaemia. *Lancet* 1998;351:550-4.
23. Shiach CR, Evans PA, Short MA, Bailey CC, Lewis IJ, Kinsey SE. Detection and accurate sizing of PCR product by automated scanning: improved detection of immunoglobulin gene rearrangements in ALL. *Br J Haematol* 1993;85:431-3.