Prognostic study of continuous variables (white blood cell count, peripheral blast cell count, haemoglobin level, platelet count and age) in childhood acute lymphoblastic leukaemia. Analysis of a population of 1545 children treated by the French Acute Lymphoblastic Leukaemia Group (FRALLE)

J Donadieu1,2, M-F Auclerc3, A Baruchel4, Y Perel3, P Bordigoni2, J Landman-Parker2, T Leblanc3, G Cornu4, D Sommelet3, G Leverger2, G Schaison3 and C Hill1 for the FRALLE group*

1Département de Biostatistique et d’épidémiologie, Institut Gustave Roussy, rue Camille Desmoulins, Villejuif, 94800 France; 2Service d’Hémato-Oncologie Pédiatrique; Hôpital Trousseau, Paris, France; 3Service d’Hématologie Pédiatrique, Hôpital Saint Louis, Paris, France; 4Unité d’Hémato-Oncologie Pédiatrique, Hôpital Pellegrin, Bordeaux, France; 5Service de Médecine infantile II, Hôpital De Brabois, Vandoeuvre, France; 6Service d’Hématologie Pédiatrique, Clinique Universitaire Saint Luc, Bruxelles, Belgique

Summary
Many cutpoints have been proposed to categorize continuous variables in childhood acute lymphoblastic leukaemia (white blood cell count, peripheral blast cell count, haemoglobin level, platelet count and age), and have been used to define therapeutic subgroups. This variation in the choice of cutpoints leads to a bias called the ‘Will Rogers phenomenon’. The aim of this study was to analyse variations in the relative risk of relapse or death as a function of continuous prognostic variables in childhood ALL and to discuss the choice of cutpoints. We studied a population of 1545 children with ALL enrolled in three consecutive protocols named FRALLE 83, FRALLE 87 and FRALLE 89. We estimated the risk of relapse or death associated with different values of each continuous prognostic variable by dividing the sample into quintiles of the distribution of the variables. As regards age, a category of children under 1 year of age was distinguished and the rest of the population was divided into quintiles. The floated variance method was used to calculate the confidence interval of each relative risk, including the reference category. The relation between the quantitative prognostic factors and the risk was monotonic for each variable, except for age. For the white blood cell count (WBC), the relation is log linear. The risk associated with WBC values in the upper quintile was 1.9 times higher than that in the lower quintile. The peripheral blast cell count correlated strongly with WBC (correlation coefficient: 0.99). The risk increased with the haemoglobin level, and the risk in the upper quintile was 1.3 times higher than that in the lower quintile. The risk decreased as the age increased above one year. The small subgroup of patients (2.5% of the population) under 1 year of age at diagnosis had a risk 2.6 times higher than the reference category. The relation between the quantitative prognostic factors and the risk was monotonic for each variable, except for age. For the white blood cell count (WBC), the relation is log linear. The risk associated with WBC values in the upper quintile was 1.9 times higher than that in the lower quintile. The peripheral blast cell count correlated strongly with WBC (correlation coefficient: 0.99). The risk increased with the haemoglobin level, and the risk in the upper quintile was 1.3 times higher than that in the lower quintile. The risk decreased as the age increased above one year. The small subgroup of patients (2.5% of the population) under 1 year of age at diagnosis had a risk 2.6 times higher than the reference category of patients between 3 and 4.3 years of age. When the risk associated with a quantitative prognostic factor varies monotonously, the selection of a cutpoint is arbitrary and represents a loss of information. Despite this loss of information, such arbitrary categorization may be necessary to define therapeutic stratification. In that case, consensus cutpoints must be defined if one wants to avoid the Will Rogers phenomenon. The cutpoints proposed by the Rome workshop and the NCI are arbitrary, but may represent an acceptable convention. © 2000 Cancer Research Campaign http://www.bjcancer.com

Keywords: acute lymphoblastic leukaemia; childhood; prognostic factors; continuous variable; cutpoint

Prognostic factors can be dichotomous by nature, but are more often either polytomous or continuous. Continuous variables are almost always used in prognostic models after categorization. Categorizing a continuous variable has adverse consequences: it requires the definition of cutpoints and implies a threshold effect: the patients immediately above and below the cut off value are supposed to be less similar than the patients at each end of the same category. Choosing a single cutpoint will result in a loss of statistical power in most instances. But many practical or biological reasons lead researchers to use one or several cutpoints in order to define the therapeutic classification. Different methods have been proposed to select cutpoints. One is called the optimal cutpoint method (Hilsenbeck et al, 1992). It is based on the P value of repeated univariate analyses of the same sample. This method has been rejected (Hill, 1993; Altman et al, 1994) because it artificially decreases the P value. Another possibility is to divide the sample into equal parts (thirds or quintiles for instance) on the basis of the continuous variable under study, and to evaluate the risk in each category. Quantitative prognostic variables in childhood ALL include the white blood cell count, peripheral blast cell count, haemoglobin level, platelet count and age at diagnosis. The prognostic value of these variables was identified more than 30 years ago (Tivey,

Received 17 February 2000
Revised 9 August 2000
Accepted 15 August 2000

Correspondence to: J Donadieu

* Members of the FRALLE group are listed at the end of the paper.
MATERIALS AND METHODS

French acute lymphoblastic leukaemia group (FRALLE) 83, 87 and 89 protocols

FRALLE 83 (May 1983–September 1986) included 818 patients from 59 centres, FRALLE 87 (June 1987–March 1989) 361 patients from 37 centres, and FRALLE 89 (April 1989–January 1992) 586 patients from 35 centres. A brief description of the three protocols can be found in Table 2. The results have been published elsewhere (Donadieu et al, 1998).

Exclusion from the study

Of the 1766 patients included in the three protocols, 221 were excluded from the present analysis: 3 for errors in the initial diagnosis, we used the lowest haemoglobin level and complete blood count. If several blood counts were available at the initial diagnosis may lead to an apparent improvement in the outcome of all the different prognostic risk groups, despite the lack of any global improvement. This is called the Will Rogers phenomenon (Feinstein et al, 1985). The purpose of this study, using a published data set (Donadieu et al, 1998), was to explore systematically the relationship between continuous prognostic variables and outcome in childhood ALL.

Endpoints, cut-off dates for analysis, and follow-up

The chosen endpoint was primary failure, toxic death or relapse at any site. Event-free survival (EFS) was defined as the time between the diagnosis and the time of the event or the time of the last examination when no event occurred. If a patient failed to enter complete remission, the event was recorded as occurring immediately, and EFS was 0. The cut-off date for the analysis was September 30, 1999.

Study variables

We studied quantitative variables with a known prognostic value recorded in the FRALLE data base, i.e. age at diagnosis and the complete blood count. If several blood counts were available at initial diagnosis, we used the lowest haemoglobin level and platelet count before any transfusion, and the highest white blood cell count and peripheral blast cell count.

Table 1 Number and values of cutpoints for age and white blood cell counts (WBC) used in the literature to define prognostic categories in ALL

| Cooperative institution (reference) | Age cutpoints | WBC cutpoints |
|-------------------------------------|---------------|--------------|
|                                    | Number | Values in years | Number | Values in 1000/mm³ |
| French group (Jacquillat et al, 1978) | 3      | 1–15–20        | 1      | 35             |
| FRALLE 83 (Schaison et al, 1990)    | 2      | 2–10           | 3      | 15–50–100      |
| NOPHO group (Gustafsson et al, 1998) | 1      | 1–2–10         | 4      | 10–20–50–100   |
| Japanese group (Hiyoshi et al, 1985) | 4      | 1–4–6–10      | 4      | 5–10–20–50–200 |
| CCG 1972–1975 (Robinson et al, 1980) | 5      | 2–4–6–8–10   | 5      | 5–10–20–50–100 |
| CCG 160 (Bleyer et al, 1986)        | 4      | 1–3–6–10      | 2      | 20–50         |
| BFM before 1981 (Bucsky et al, 1988) | 5      | 0.5–1–1.5–2–10 | NA    |               |
| BFM 81 (Schrapp et al, 1987)        | 4      | 1–2–10–14     | 3      | 10–50–100     |
| BFM 86 (Reiter et al, 1994)         | 3      | 1–6–10        | 4      | 10–20–50–200  |
| DFCI 85–01 (Schorin et al, 1994)    | 3      | 1–2–9         | 3      | 20–50–100     |
| UKALL X (Chessells et al, 1995)      | 2      | 1–10          | 4      | 10–20–50–100  |

Each cutpoint is included in the interval to the right, for instance cutpoint 1 and 10 for age define the intervals age <1, 1 to 9 and 10+. NA: Not available. WBC is indirectly in the model, which included 0.2 Log (peripheral blast cell count +1).
Statistical methods

BMDP® software was used for all statistical analyses. Median follow-up was estimated from a Kaplan–Meier curve in which the last follow-up date was used for the patients who survived as if it was an endpoint, and in which the other patients were censored at the time of death (Schemper and Smith, 1996). The actuarial method was used to estimate survival rates. A Cox semi-parametric model was used to estimate the hazard ratio in each category (Cox, 1972; Collet, 1994). The cutpoints were obtained by dividing the overall sample into quintiles. A category below 1 year of age corresponding to 39 patients (2.5% of the population) was defined, as proposed in the literature (Mastrangelo, 1986b; Smith et al, 1996), and the rest of the population was divided into quintiles. For each cutpoint, the boundary value was included in the higher category. The variance of the hazard ratios was calculated by the floated variance method (Easton et al, 1991), in order to attribute some imprecision to the reference category. In the figures, the hazard ratio for each category (after log transformation for white blood cells) is plotted at the median value of the category on the x axis. Another Cox regression model estimated the trend in the log hazard ratios.

RESULTS

The median follow-up was 111 months in FRALLE 83, 106 months in FRALLE 87, and 86 months in FRALLE 89. Among the 1545 patients, a total of 781 events were recorded during the study period, corresponding to 36 toxic deaths during induction, 42 induction failures, 65 toxic deaths after induction (patients in first remission) and 638 relapses. Complete remissions were obtained after induction in 1467 patients (95%), while 78 patients (5%) failed to enter complete remission after the induction phase.

The 10-year EFS rate was 48.2% (SE 1.3), and the 10-year survival rate was 61.4% (SE 1.3).

Age

The distribution of age was approximately log-normal, with a median of 5.1 years. The risk of events varied with age according to a U-shaped curve. The group of patients under one year of age had a risk 2.6 times higher than that of patients between 3 and 4.3 years of age (the reference quintile). A Cox regression model with a specific hazard ratio for the population under one year of age and a linear trend in the older population was fitted, leading to the estimation:

\[
HR = \exp [1.09 (\text{if age < 1 year}) + 0.042 \times (\frac{\text{age}}{11002} - 3.5) (\text{if age \geq 1 year})]
\]

The risk in the highest quintile was 1.6 times higher than that in the reference quintile (Figure 1).

White blood cell count

The white blood cell count (WBC) had an exponential distribution with a median of 12 000/mm³. The risk of relapse or toxic death increased with increasing WBC (Figure 2). A Cox regression model was fitted with a linear effect of log(WBC):

\[
HR = \exp [0.158 \times (\log (\text{WBC} + 0.1) - \log 2.5)]
\]

where WBC is expressed in 1000/mm³. The risk in the highest quintile was 1.9 times higher than that in the lowest quintile.

Peripheral blast cell count

Peripheral blast cells represent a large proportion of white cells at diagnosis, and their numbers correlate strongly (r = 0.99). Thus, the peripheral blast cell count provides the same prognostic information as the WBC.

Haemoglobin level

The distribution of haemoglobin levels at diagnosis was approximately normal, with a median of 7.7 g dl⁻¹. A Cox regression model was fitted with a linear effect of the haemoglobin level:

\[
HR = \exp [0.037 \times (\text{Hb} - 4.6)]
\]

The risk in the highest quintile was 1.3 times higher than that in the lowest quintile (Figure 3).

Platelet count

The distribution of the platelet count was log-normal, with a median of 56 000/mm³. The hazard ratio of events fell when platelets increased. A Cox regression model was fitted with a linear effect of the platelet count:

\[
HR = \exp [-0.0012 \times (\text{Platelet} - 232)]
\]

The risk in the lowest quintile was 1.2 times higher than that in the highest quintile (Figure 4).
The main result of our study is that, except for age, the relation between the risk of relapse or toxic death and each continuous prognostic factor in childhood ALL is monotonic. The patients under 1 year of age had a markedly increased risk, while the risk for patients above 1 year increased linearly with age. Except for age, the ratio between the risk in the worst quintile and that in the best quintile never exceeded 2.

We performed a univariate analysis for each variable dividing the sample into quintiles. This avoids the optimal cutpoint method and the associated multiple tests. We used a linear or log linear model, which is simpler than fractional polynomials (Royston and Altman, 1994; Sauerbrei, 1999).

When the risk varies monotonously as a function of a continuous variable, any categorization of the variable represents a loss of statistical information, and there is no optimal cutpoint. For instance, the selection of a cutpoint of 50 000 WBC/mm³ for WBC introduces a loss of statistical information, and there is no optimal cutpoint. For example, the selection of a cutpoint of 50 000 WBC/mm³ for WBC is arbitrary, as is the selection of 9 years for age.

The analysis was not adjusted on other known prognostic variables. Here, we present an exploratory study of the relation between the risk of relapse or death and different continuous variables in childhood acute lymphoblastic leukaemia. A multivariate analysis using the same data set showed that age and the white blood cell count were the most significant prognostic variables in childhood acute lymphoblastic leukaemia, whereas the haemoglobin level and the platelet count had a lower prognostic value (Donadieu et al., 1998).

Many cutpoints have been used in the literature (Table 1) but this diversity makes it difficult to compare the results of different protocols. The same patient can be considered as having a low risk or as having a high risk depending on the prognostic classification used. Such classification changes may lead to an apparent improvement in each prognostic risk group, without any overall improvement (Feinstein et al., 1985). This is illustrated in Figure 5 with a teaching example, and in Figure 6 with our data set. A classification change may lead to an apparent improvement in each prognostic risk group, without any overall improvement (Feinstein et al., 1985). This is illustrated in Figure 5 with a teaching example, and in Figure 6 with our data set. A classification change may lead to an apparent improvement in each prognostic risk group, without any overall improvement (Feinstein et al., 1985).
change the values of quantitative variables defining treatment groups at each new analysis or at each new protocol.

Two consensus conferences (the Rome workshop (Mastrangelo, 1986b) and the NCI meeting (Smith et al, 1996)) have proposed cutpoints for age and white blood cells, which represent the most significant continuous variables (Donadieu et al, 1998). They both proposed 1 and 9 years, and 50 000 WBC/mm3 as cutpoints for age and the white blood cell count, respectively. These values are both arbitrary and acceptable.

REFERENCES

Altman DG, Lausen B, Sauerbrei W and Schumacher M (1994) Danger of using “optimal” cutpoints in the evaluation of prognostic factors. J Natl Cancer Inst 86: 829–835

Bleyer WA (1989) Remaining problems in the staging and treatment of childhood lymphoblastic leukemia. Am J Pediatr Hematol Oncol 11: 371–379

Bleyer WA, Sather H, Coccia P, Lukens J, Siegel S and Hammond GD (1986) The staging of childhood acute lymphoblastic leukemia: strategies of the Childrens Cancer Study Group and a three-dimensional technic of multivariate analysis. Med Pediatr Oncol 14: 271–280

Bucsky P, Reiter A, Ritter J, Dopfer R and Riehm H (1988) Die akute lymphoblastische leukämie im Säuglingsalter: Ergebnisse aus fünf multisentrischen therapiestudien ALL-BFM 1970–1986. Klin Padiatr 200: 177–183

Chessells JM, Bailey C and Richards SM (1995) Intensification of treatment and survival in all children with lymphoblastic leukaemia: results of UK medical research council UKALL X. Lancet 345: 143–148

Collet D (1994) Modelling survival data in medical research. Chapman and Hall: London

Cox DR (1972) Regression models and life tables (with discussion). J R Stat Soc 34: 246–275

Donadieu J, Auclerc MF, Baruchel A, Leblanc T, Landman-Parker J, Perel Y, Michel G, Cornu G, Bordignon P, Sonnelet D, Leverger G, Hill C and Schaison G (1998) Critical study of prognostic factors in childhood acute lymphoblastic leukaemia: differences in outcome are poorly explained by the most significant prognostic variables. Br J Haematol 102: 729–739

Easton DF, Peto J and Babiker AGAG (1991) Floating absolute risk: an alternative to relative risk in survival and case control analysis avoiding an arbitrary reference group. Stat Med 10: 1025–1035.

Feinstein AR, Sosin DM and Wells CK (1985) The Will Rogers phenomenon: Stage migration and new diagnosis techniques as a source of misleading statistics for survival in cancer. N Engl J Med 312: 1604–1608

Gustafsson G, Kreuger A, Clausen N, Garwicz S, Kristinsson J, Lie SO, Moe PJ, Perkkio M, Yssing M and Saareninen-Pihkala UM (1998) Intensified treatment of acute childhood lymphoblastic leukaemia has improved prognosis, especially in non-high-risk patients: the Nordic experience of 2648 patients diagnosed between 1981 and 1996. Nordic Society of Paediatric Haematology and Oncology. Acta Paediatrica 87: 1151–1161

Hill C (1993) Valeur pronostique d’une variable continue et point de césure optimal. Bull Cancer 80: 649–652

Hilsenbeck SM, Clark GM and McGuire WL (1992) Why do so many prognostic factors fail to pan out? Breast Cancer Res Treat 22: 197–206

Hiyoshi Y, Fujimoto T, Kuriya N, Otani Y, Ibu K, Yanai M, Sasaki K, Shingaki Y, Yokoyama T and S-K (1985) Prognostic factors in children with acute lymphoblastic leukaemia. Part II: multivariate analysis. Jpn J Clin Oncol 15: 13–23

Jacquillat C, Weil M, Auclerc MF, Chastang C, Flandrin G, Izrael V, Schaison G, Degos L, Boiron M and Bernard J (1978) Prognosis and treatment of acute lymphoblastic leukemia. Cancer Chemother Pharmacol 1: 113–122

Mastrangelo R (1986a) The problem of “staging” in childhood acute lymphoblastic leukaemia: a review. Med Pediatr Oncol 1: 121–123

Mastrangelo R (1986b) Report and recommendations of the Rome workshop concerning post-prognosis ALL in children: Biologic bases for staging, stratification, and treatment. Med Pediatr Oncol 14: 191–194

Reiter A, Schrappe M, Ludwig WD, Hiddemann W, Sauter S, Henze G, Zimmermann M, Lampert F, Havers W, Niethammer D, Odenwald E, Ritter J, Mann G, Welte K, Gadner H and Riehm H (1994) Chemotherapy of 998 unsellected childhood acute lymphoblastic leukemia patients. Results and conclusions of the multicenter trial ALL-BFM 86. Blood 84: 3122–3133

Robison LL, Sather HN, Coccia PF, Nesbit ME and Hammond GD (1980) Assessment of the interrelationship of prognostic factors in childhood acute lymphoblastic leukaemia. Am J Pediatr Hematol Oncol 2: 5–13

Royston D and Altman DG (1994) Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. Applied Statistics 43: 429–467

Sauerbrei W (1999) Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials. J R Statist Soc A 162: 71–94

Schaison G, Ollive D, Leverger G, Vanier JP, De Lumley L, Bancillon A and Cornu G (1990) Treatment of acute lymphoblastic leukaemia: protocol Fralle 83–85. Haematol Blood Transfu 33: 467–472

Schemper M and Smith TL (1996) A note on quantifying follow-up in studies of failure time. Controlled Clin Trials 17: 343–346

Schorin MA, Blattner S, Gelber RD, Tarbell NJ, Donnelly M, Dalton V, Cohen HJ and Sallan SE (1994) Treatment of childhood acute lymphoblastic leukaemia: results of Dana-Farber Cancer Institute/Children’s hospital acute lymphoblastic leukaemia consortium protocol 85–01. J Clin Oncol 12: 740–747

Schrapp M, Beck J, Brandeis WE, Feickert H, Gadner H, Graf N, Havers W, Henze G, Jobke A, Kornhuber B and et al (1987) Treatment of acute lymphoblastic leukaemia in childhood and adolescence: results of the multicenter therapy study ALL-BFM 81. Klin Padiatr 199: 133–150

Smith M, Arthur D, Camitta B, Caroll AJ, Crist W, Gaynon P, Gelber R, Heerema N, Korn EL, Link M, Murphy S, Pui CH, Pullen J, Reaman G, Sallan SE, Sather H, Shuster J, Simon R, Trigg M, Tubergen D, Uckun F and Ungerechter R (1996) Uniform approach to risk-classification and treatment assignment for children with ALL. J Clin Oncol 14: 18–24

Tivey H (1952) Prognosis for survival in the leukemias of childhood. Pediatrics 10: 48–59

Zuelzer WW (1964) Implications of long-term survival in acute stem cell leukaemia of childhood treated with composite cyclic therapy. Blood 24: 477–494
APPENDIX 1 LIST OF THE MEMBERS OF THE FRALLE GROUP

Pautard B, Centre Hospitalier Universitaire, Amiens; Berthou C, Centre Hospitalier Universitaire, Brest; Cornu G, Département d'Hématologie Infantile, Faculté de Médecine de Louvain, Bruxelles, Belgique; Perel Y, Centre Hospitalier Pellegrin, Bordeaux; P Boutard; Centre Hospitalier Universitaire, Caen; Demœcq F, Hôtel Dieu, Centre hospitalier Universitaire, Clermont Ferrand; Lemerle S, Bernaudin F, Centre Hospitalier Inter-Communal de Créteil, Créteil; Couillault G, Centre hospitalier Universitaire, Dijon; Bader-Meunier B, Domergues JP, Tchernia G, Hôpital Bicêtre, Le Kremlin Bicêtre; Piguet C, De Lumley L, Centre hospitalier universitaire, Limoges; Thuret I, Michel G, Hôpital de la Timone, centre Hospitalier Universitaire, Marseille; Bordigoni P, Olive-Sommelet D, Schmitt C, Hôpitaux de Brabois, Centre Hospitalier Universitaire, Nancy; Blanche S, Casanova JL, Debré M, Dupuis S, Fischer A, Quartier P, Thomas C, Hôpital Necker Enfants-Malades, Paris; Auclerc MF, Baruchel A, Berger R, Espérou-Bourdeau H, Leblanc T, Lepage E, Schaison G, Sigaux F, Hôpital Saint-Louis, Paris; Auvrignon A, Donadieu J, El Adjidi N, Landman-Parker J, Leverger G, Rognon C, Tabone MD, Hôpital A Trousseau, Paris; Bergeron C, Edan C, Gandemer V, Le Gall E, Centre Hospitalier Universitaire, Rennes; Vannier JP, Hôpital C Nicole, Rouen; Freycon F, Stephan JL, Centre Hospitalier Universitaire, Saint Etienne; Lamagnères JP, Lejars O, Hôpital de Clocheville, Tours; Hayat M, Pico JL; Institut Gustave Roussy, Villejuif.