*Haematologica 1994; 79:233-240*

# STUDY OF PROGNOSIS IN ACUTE MYELOID LEUKEMIAS (AML) BY CLUSTER ANALYSIS

***Gian Matteo Rigolin, Franca Fagioli, Romedio Spanedda, Gianluigi Scapoli, Francesco Lanza, Antonio Cuneo, Paolo Tomasi, Gianluigi Castoldi***

*Institute of Hematology, University of Ferrara, Italy*

**ABSTRACT**

***Background.*** **Cluster analysis is particularly effective in detecting homogeneous subgroups among large series of observations. We applied this relatively uncommon approach to the study of prognosis in 137 patients affected by acute myeloid leukemia (AML).**

***Methods and Results.*** **Employing simple presentation parameters (age, WBC, splenomegaly, hepatomegaly) we used cluster analysis to define 3 groups with different overall survival (p=0.0019). This classification was obtained following a rescaling of the variables and principal component analysis. Validation was performed through random definition of a control group. With the same variables, univariate analysis demonstrated age was the only prognostic factor, while Cox's model was not significant.**

***Conclusions.*** **In our series cluster analysis allowed a better definition of prognosis than Cox's analysis. Since the 3 groups are well identifiable, each patient can be rapidly classified and his allocation confirmed by discriminant functions. For cluster 2 we were able to project a possible myelodysplastic evolution, while cluster 3 was more frequently associated with a monocytic blastic component. We think that cluster analysis deserves consideration as an alternative statistical approach in the analysis of large series of data; its usefulness lies in its power to define homogeneous prognostic or biologic subgroups and to elaborate further hypotheses for new studies.**

*Key words: cluster analysis, acute myeloid leukemias, overall survival, prognosis*



The definition of clinical and laboratory features affecting prognosis is one of the main aims of statistical analysis on a large series of patients. The most popular methods applied for this purpose are univariate analysis and Cox's proportional hazard model.[1]

Cluster analysis methods,[2,3] using a mathematical algorithm, are particularly appropriate for detecting subgroups of observations with similar features in relatively large series. This approach is useful since it enables us to define different disease entities, different etiologies or different therapeutic or prognostic outcomes on an adequate series of patients. The advantage to these methods lies in their power to define subgroups involving a duality between variables and subjects. The aim of this work was to analyze retrospectively 160 AML patients referred to the Institute of Hematology of Ferrara, using cluster analysis to delineate subgroups of patients with particular features and different prognoses.

## *Patients and methods*

One hundred sixty *de novo* AML patients, aged 14-86 years, consecutively referred to the Institute of Hematology of Ferrara between January, 1979 and April, 1992 were retrospectively studied. At diagnosis all patients, classi-

*Correspondence: Dr. Gian Matteo Rigolin, Institute of Hematology, University of Ferrara, via Savonarola 9, 44100 Ferrara, Italy. Fax: international +39.532.212142.*
*Acknowledgments: this work was supported by MURST (60%) and CNR Rome.*
*Received February 2, 1994; accepted April 7, 1994.*

fied in accordance with FAB criteria,[4,5] underwent clinical and laboratory studies that included physical examination, hemoglobin, white blood cell (with differential) and platelet determinations, evaluation of kidney, liver and heart functions. Two protocols, described elsewhere,[6] were applied for induction: the first regimen consisted of one course of DAT3/7 (daunorubicin, cytosine arabinoside and thioguanine) followed by 3 DAT1/5 each after a 15-day interval; the second consisted of 2 VAE (vindesine, cytosine arabinoside and etoposide) alternating DAT3/7 and DAT1/5 after a 2-week interval. Conventional dosages were reduced in the presence of kidney, heart or liver impairment or advanced age. In some patients daunorubicin was substituted by epirubicin or idarubicin or mitoxantrone. Patients in complete remission at the end of induction received monthly alternating courses of AT1/5, AT1/5 or VAE, DAT1/5 as maintenance therapy. Treatment was stopped after a 2-year event-free period. Patients with low performance status received low-dose cytosine arabinoside induction and maintenance regimens. Eleven patients, aged 18-34, underwent bone marrow transplantation.

*Statistical analysis*

Overall survival (OS) and relapse-free survival (RFS) were computed according to the Kaplan-Meier technique:[7] OS was determined from the time of diagnosis, while RFS was computed from the time of first bone marrow aspiration in complete remission (CR). Patients who underwent bone marrow transplantation did not drop out of the study. Survival and remission durations were compared by the log-rank test.[8] Cox's hazard proportional model (Maximum Log-Likelihood Estimate) was used to determine multifactorial effects of presentation parameters on survival and CR durations.[9] The chi-square test was used throughout; *k-means* Hartgan's algorithm[10] was used in cluster analysis. Data were rescaled using a rank transformation (high tie). Principal component analysis was applied in order to extract the real variables describing our group of patients; this vectorial analysis procedure allows exclusion of those variables whose weight is not relevant to the definition of the population. Validation was obtained with a cohort of subjects after random division of the patients into two groups: one (70% of subjects) for cluster definition, the other (30% of patients) as control group. The grouped and ungrouped patients were then classified with discriminant analysis to verify the goodness of cluster definition. At the end of Cox's analysis, the patients were divided into three groups according to the numeric value of the prognostic score. These groups were then compared with a log-rank test for their prognostic impact.

The SPSS statistical package was used for elaborating the data. Data were analyzed in May, 1993.



***Results***

Of the 160 patients considered for this study, 137 were valuable for statistical analysis; 13 who received low-dose cytosine arabinoside induction treatment were excluded because this regimen was not comparable with that applied to the great majority of patients; 4 cases were excluded for major protocol violations, 4 for inadequate follow-up documentation, 2 for refusal of treatment.

Table 1 summarizes the principal features of the 137 patients at presentation. Five-year OS of the 137 patients was 5.7% (standard error, SE 2.3%), with a median OS of 3 months. Forty-six (33.6%) patients achieved CR with a 5-year RFS of 20.4% (SE 7.0%) and a median RFS of 14 months. The impact of age on survival was studied by stratifying patients in three groups (≤35 years; 36-70; ≥70): the 5-year OS of the three groups was, respectively, 0%, 8.7% (SE 4.6%), 3.4% (SE 2.7%)($p=0.015$), while median OS was 11, 5 and 2 months, respectively.

The 5-year OS of patients who achieved and those who did not achieve CR were, respectively, 17.5% (SE 6%) vs 0% ($p<0.0001$) and median OS was 16 vs 2 months. No other factors affected OS or RFS in univariate analysis. Following principal component analysis, four variables (age at diagnosis, splenomegaly, hepatomegaly, white blood cell count) were tested in Cox's proportional hazard model for their impact on

Table 1. Clinical-laboratory features at presentation.

| | |
|---|---|
| n° pts | 137 |
| M/F | 75/62 |
| median age (yr) | 62 (14-86) |
| splenomegaly (n° pts) | 55 (40.1%) |
| hepatomegaly (n° pts) | 80 (58.4%) |
| Hb (median) g/dL | 9.0 (4.4-15.7) |
| WBC $10^9$/L | 7.3 (0.2-371.3) |
| Plts $10^9$/L | 72 (7-658) |
| M1 n° pts (%) | 12 (8.6%) |
| M2 n° pts (%) | 32 (22.9%) |
| M3 n° pts (%) | 13 (9.3%) |
| M4 n° pts (%) | 43 (30.9%) |
| M5 n° pts (%) | 29 (20.7%) |
| M6 n° pts (%) | 6 (4.3%) |
| M7 n° pts (%) | 2 (1.4%) |
| CR n° pts (%) | 46 (33.6%) |
| induction deaths pts (%) | 49 (35.7%) |
| relapse n° pts (%) | 31 (67.4%) |
| mean follow-up (months) | 12 (1-128) |

survival and RFS: Cox's regression, considering all the four variables together, was not significant for OS (chi-square 9.03, 4 df; p=0.07), while the model was significant (chi-square 12.8, p=0.015, 4 df) regarding RFS. Hartgan's algorithm, with the same variables tested in Cox's analysis, was used to perform cluster analysis. We identified 3 different groups, the clinical and laboratory features of which are presented in Table 2 and Figure 1. The 3 groups were compared for OS (Figure 2): the 5-year OS was, respectively, 4.9% (SE 3.0%), 0%, 6.4% (SE 4.0%), with a median OS of 6, 1 and 4 months, respectively (p=0.0019). Five-year RFS for the three clusters was 13.6% (SE 7.7%), 0%, 24 (SE 14%), respectively, while median RFS was 8, 23 and 18 months (p=0.010: Tarone-Ware statistic). To validate cluster analysis all patients were divided by means of a random choice procedure into two groups: one (99 pts.) for cluster determination (Table 3; Figure 3), the other to verify by discriminant analysis the distribution of the *ungrouped subjects* in the predefined clusters. Correct allocation of patients was attained in 85% of cases (Figure 4). The OSs of the 3 predefined groups were significantly different (p=0.003), and the three classes of patients obtained considering prognostic score values were also significantly different (p=0.03). A comparison between the cluster analysis results and Cox's analysis was made by defining 3 groups on the basis of the prognostic score and observing the outcomes of the two approaches in both univariate analysis and in Cox's model (categorical variables). Within clusters, stepwise Cox's analysis demonstrated a prognostic impact for splenomegaly in cluster 1 (chi-square 5.246, 1 df, p=0.025), and for hepatomegaly in cluster 3 (chi-square=5.071, 1 df, p=0.023).



Table 2. Clinical-laboratory aspects of the 3 cluster analysis groups.

| *cluster* | *1* | *2* | *3* |
|---|---|---|---|
| n° pts | 60 | 41 | 36 |
| M/F | 30/30 | 25/16 | 20/16 |
| M1 n° pts (%)* | 6 (10%) | 4 (9.8%) | 2 (5.6%) |
| M2 n° pts (%)* | 15 (25%) | 9 (22%) | 7 (19.4%) |
| M3 n° pts (%)* | 8 (13.3%) | 2 (4.9%) | 2 (5.6%) |
| M4 n° pts (%)* | 21 (35%) | 12 (29.3%) | 13 (36.1%) |
| M5 n° pts (%)* | 9 (15%) | 14 (17.7%) | 10 (27.8%) |
| M6 n° pts (%)* | 0 | 5 (12.3%) | 1 (2.8%) |
| M7 n° pts (%)* | 1 (1.7%) | 0 | 1 (2.8%) |
| induction deaths (%)° | 11 (18.5%) | 21 (50.2%) | 11 (30.5%) |
| 5-yr OS (SE)# | 4.9 (3.0%) | 0% | 6.4 (4.0%) |
| pts alive (%) | 5 (18.3%) | 1 (2.4%) | 2 (5.6%) |
| CR n° pts (%)^ | 29 (48.3%) | 3 (7.3%) | 12 (33.3%) |

**p=ns; °p=0.0023; #p=0.0019; ^ p=0.0008*

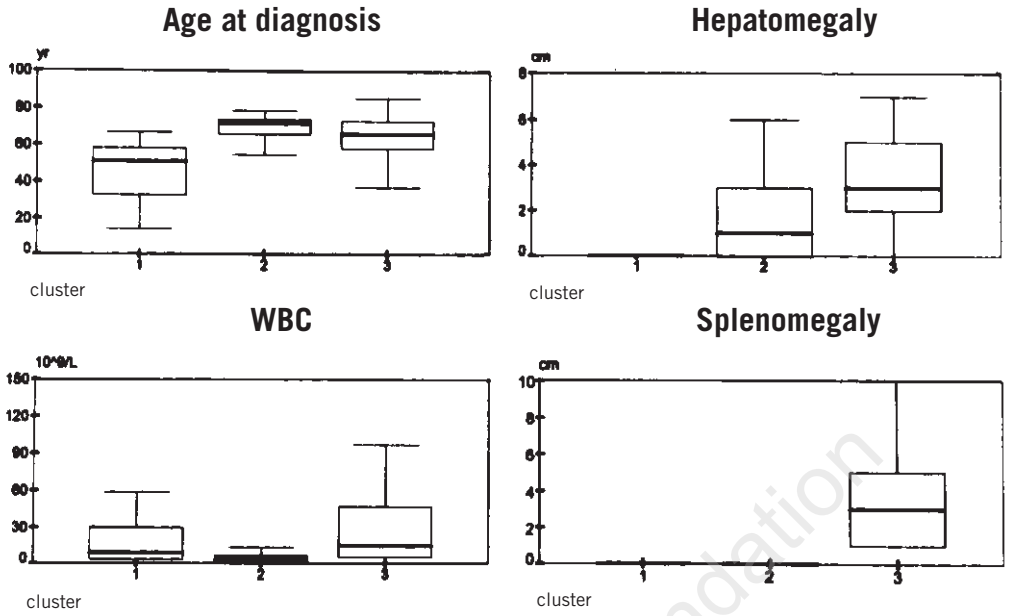


Figure 1. Box plots of the 4 classifying variables in the three clusters of patients.



***Discussion***

The definition of prognostic groups in a large series of patients is the main aim of statistical analysis. The most popular methods applied for

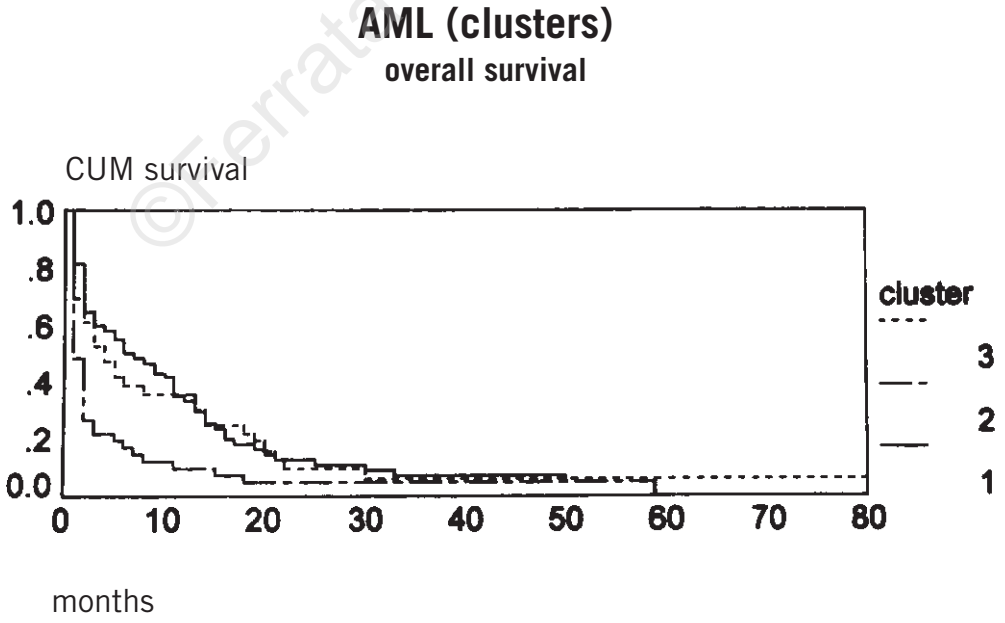


Figure 2. Cluster analysis and survival Log-rank test statistic: overall p=0.0019; cluster 1 vs cluster 2, p=0.0008; cluster 1 vs cluster 3 p=0.071; cluster 2 vs cluster 3, p=0.0018.

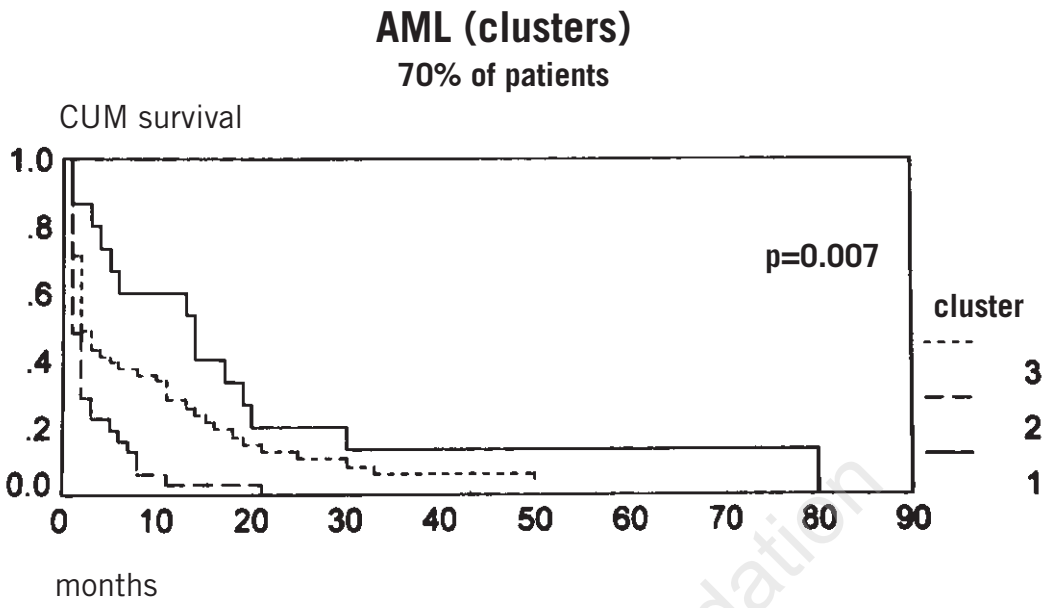


Figure 3. Cluster analysis and survival Log-rank test statistic in the 99 randomly selected patients: overall p=0.007; cluster 1 vs cluster 2, p=0.0001; cluster 1 vs cluster 3, p=0.11; cluster 2 vs cluster 3, p=0.0018.



this purpose are univariate analysis and multivariate analysis via Cox's proportional hazard model. Two limitations to this approach are the heterogeneity of the population analyzed and the observation that a parameter that could be prognostically important for one group of patients might not be decisive in another. Cluster analysis methods, which consider a duality between variables and subjects, are particularly effective in the definition of distinct subgroups of patients with similar features; in this sense they are more flexible than other methods. These groups may represent different diseases, different etiopathologic mechanisms, different probabilities of response to therapies or different prognoses.

As recently reviewed in the literature, very few studies have applied cluster analysis as a statistical method. In particular, to our knowledge this approach had never been used in the study of AML prognosis. Our application could be a model for an alternative statistical approach to subgrouping AML patients with a view to defining prognosis or other biological features.

Using this relatively uncommon approach we were able to distinguish 3 groups of patients with different prognostic outcomes and CR rates on the basis of 4 simple variables. An interesting aspect of our analysis is the comparison of results obtained using conventional methods (univariate analysis and Cox's hazard model) with those using cluster analysis. Age at diagnosis was delineated as the only variable affecting survival in univariate analysis, while in Cox's model all four variables together were not significant.

These results in our opinion suggest that, when studying prognosis or other biological aspects in a large series of patients, cluster analysis could be a very useful tool for analyzing data in order to separate distinct groups of subjects with similar features. Besides its prognostic impact, this and other possible classifications could be used to formulate operative hypotheses for new studies.

In particular, from our analysis of hematological parameters in the 3 groups studied we can hypothesize that group 2, with its lower median white blood cell count, tendency to low hemoglobin levels, lower mean peripheral blast cell percentage (40% vs 60%) and five M6 patients, represents a cluster with a high percentage of

Table 3. Clinical-laboratory features of the 3 clusters obtained with 99 (randomly selected) patients.

| *cluster* | *1* | *2* | *3* |
|---|---|---|---|
| n° pts | 15 | 31 | 53 |
| M/F | 4/11 | 21/10 | 30/23 |
| age (median) | 57 (19-73) | 73.5 (59-86) | 54.1 (16-86) |
| Hb g/dL | 7.6 (6.5-12) | 9.1 (6-14) | 9.0 (5.2-13.7) |
| WBC $10^9$/L | 51 (6-371) | 2.7 (0.7-50) | 6.1 (0.2-122) |
| Blasts % | 78.2 (35-95) | 40 (5-59) | 48 (10-67) |
| Plt $10^9$/L | 58 (20-243) | 50 (9-227) | 81 (17.395) |
| hepatomegaly n° pts | 13 (87%) | 30 (93%) | 12 (23%) |
| splenomegaly n° pts | 13 (87%) | 14 (52%) | 9 (17%) |
| M1 n° pts (%)* | 2 (13.3%) | 4 (12.9%) | 3 (5.7%) |
| M2 n° pts (%)* | 2 (13.3%) | 5 (16.1%) | 13 (24.5%) |
| M3 n° pts (%)* | 1 (6.7%) | 1 (3.2%) | 6 (11.3%) |
| M4 n° pts (%)* | 6 (40.7%) | 11 (35.5%) | 19 (35.8%) |
| M5 n° pts (%)* | 3 (20%) | 6 (19.4%) | 10 (18.9%) |
| M6 n° pts (%)* | 0 | 6 (19.4%) | 1 (1.9%) |
| M7 n° pts (%)* | 1 (6.7%) | 0 | 1 (1.9%) |
| induction deaths (%)^ | 2 (13.3%) | 16 (51.6%) | 15 (28.3%) |
| 5-yr OS (SE)§ | 13.3 (8%) | 0% | 3.3 (2%) |
| pts alive (%) | 1 (6.7%) | 0 | 4 (7.5%) |
| CR n° pts (%)° | 7 (46.7%) | 2 (6.5%) | 20 (37.7%) |
| relapses | 6 | 1 | 13 |

*p=ns; ^p=0.01; °p=0.001; §p=0.007

patients with dysplastic hematopoiesis. In fact, some of these parameters tend to fulfill criteria for the presence of dysplastic hematopoiesis.[11] On the other hand, group 3 (characterized by hepatosplenomegaly) could be considered a subset of patients with a higher incidence of monocyte lineage involvement. This cluster does indeed show a higher incidence of M4 and M5, in which organomegaly is often associated with a peculiar pattern of presentation. This cluster is also characterized by a higher WBC count, which was correlated with a loss of HLA class I antigens.[12]

The definition of more homogeneous subgroups could also represent a useful step in the development of further analyses on subsets of patients.[13] Within clusters it is also possible to identify new prognostic implications. In clusters 1 and 3 Cox's analysis demonstrated that splenomegaly and hepatomegaly, respectively, had further prognostic impact.

Another possible advantage to this approach is the fact that it utilizes simple variables included in the routine approach to all patients. Each cluster is well identifiable for its clinical and laboratory features; moreover, using discriminant functions, we can predict the probability that new patients will belong to one or other group, thus facilitating the definition of different therapeutic strategies. The patients in the control group were in fact correctly classified in 85% of cases.

We also decided to perform cluster analysis with age at diagnosis as the only subgrouping variable, to test if this alone could justify the statistical significance of our results. The 3 groups resulting from this approach were characterized by completely different median ages from those obtained in the previous cluster analysis approach (Table 4). The OSs of the corresponding 3 groups differ (p=0.015 vs p=

Table 4. Cluster analysis with age at diagnosis as only subgrouping variable.

| *cluster* | *1* | *2* | *3* |
|---|---|---|---|
| n° pts | 23 | 73 | 41 |
| median age | 24 (14-42) | 52 (43-59) | 69 (61-84) |
| 5-yr OS (SE)* | 0 | 3.4% (2.8%) | 7.8% (4.3%) |

* p=0.015

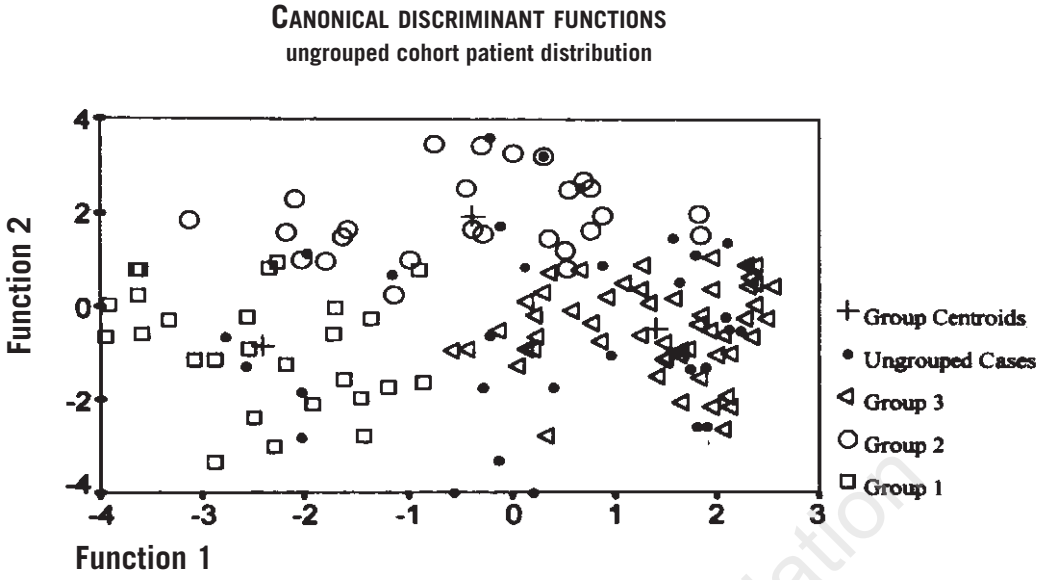


Figure 4. Scatterplot of the discriminant function values in the classifying and control groups of patients obtained with a random procedure.



0.0019 for the previous analysis), indicating the importance of the other parameters in this subgrouping and an even higher statistical significance for the multi-variable approach. The 3 groups defined by cluster analysis using age at diagnosis as the only variable are very similar to those obtained with age stratification in univariate analysis. These observations confirm the validity of cluster analysis in elaborating data and demonstrate its usefulness in analyzing the prognostic impact of single variables as well.

The CR rate and the OS of this series of patients might appear dissimilar from those reported in international studies.[14-16] We have to stress that we decided to enroll patients over 65 years old who are usually excluded in international survival reports. This decision had two motivations: first, to define the real leukemic population referred to our Institute with its OS and RFS and, second, to demonstrate the usefulness of cluster analysis in subgrouping observations. Our series of patients was characterized by a higher median age (62 years) than that of other reports: 56.4% of our patients were over 60, 75% over 50 years of age. Only 46 patients achieved CR but 5-year RFS of this group was exactly within international standards. When considering survival reports, one of the main problems is the comparability of the patients involved.

In conclusion, we propose cluster analysis as a possible and very powerful method for analyzing a large series of observations. This approach, which defines groups of subjects with similar features, could be very useful for examining data with a view to defining prognostic groups or other clinical-biological implications.

## *References*

1. Sanz GF, Sanz MA, Vallespi T, et al. Two regression models and a scoring system for predicting survival and planning treatment in myelodysplastic syndromes: a multivariate analysis of prognostic factors in 371 patients. Blood 1989; 74:395-408.
2. Vogt W, Nagel D. Cluster analysis in diagnosis. Clin Chem 1992; 38:182-98.
3. Barrai I. Metodi di regressione e classificazione in biometria. Bologna: Edagricola, 1984.
4. Bennett J, Catowsky D, Daniel MT, et al. Proposal for the classification of the acute leukemias. Br J Haematol 1976; 33:451-8.
5. Castoldi GL, Liso V, Fenu S, Vegna L, Mandelli F. Reproducibility of the morphological diagnostic criteria for acute myeloid leukemia; the GIMEMA group experience. Ann Hematol 1993; 66:171-4.

6. Bandini G, Baccarani M, Cavazzini G, et al. Treatment of acute non lymphocytic leukemia. Results of a prospective trial of daunorubicin, arabinosylcytosine and 6-thioguanine (DAT) vs DAT plus vindesine and etoposide. Haematologica 1987; 72:431-7.
7. Kaplan EL, Meier R. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958; 53:457-81.
8. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomised clinical trials requiring prolonged observations of each patient. Br J Cancer 1977; 35:1-39.
9. Cox DR. Regression models and life tables. J R Stat Soc 1972; 34:187-220.
10. Hartgan JA. Clustering algorithms. New York: Wiley & Sons, 1985.
11. Brito-Babapulle F, Catowsky D, Galton AG. Clinical and laboratory features of de novo acute myeloid leukemia with trilineage myelodysplasia. Br J Haematol 1987; 66:445-50.
12. Savoia P, D'Alfonso A, Peruccio D, et al. Loss of surface HLA class I molecules in leukemic myeloblasts is correlated with an increased leukocyte concentration at the onset. Haematologica 1992; 77:127-9.
13. Mirto S, Santoro A, Barbata G, et al. ANLL patients with normal karyotype are not a homogeneous prognostic group. Haematologica 1992; 77:484-6.
14. Gorst DW, Johnson KW. Survival in adult leukemia. Clin Lab Haematol 1992; 14:99-108.
15. Cassileth PA, Lynch E, Hines JD, et al. Varying intensity of postremission therapy in acute myeloid leukemia. Blood 1992; 79:1924-30.
16. Tashiro S, Kyo T, Tanaka K, et al. The prognostic value of cytogenetic analysis in patients with acute non-lymphocytic leukemia treated with the same intensive chemotherapy. Cancer 1992; 70:2809-15.