Statistical Hierarchy of Diagnostic Criteria for Chronic Myeloid Leukemia

DESPINA CALAMAR POPOVICI[1,3#], IOANA IONITA[1,3]*, CLAUDIU IONITA[2,4], ANDRADA MARINITA[1], RADU DUMITRU MOLERIU[5#], HORTENSIA IONITA[1,3], DANIELA IACOB[6]*, VERONICA DANIELA CHIRIAC[7#], IZABELLA PETRE[7]

[1]University of Medicine and Pharmacy Victor Babes Timisoara, Department of Hematology, 2 Eftimie Murgu Sq., 30041, Timisoara, Romania

[2]University of Medicine and Pharmacy Victor Babes Timisoara, Department of Surgery, 2 Eftimie Murgu Sq., 30041, Timisoara, Romania

[3]Clinical County Hospital Timisoara – Department of Hematology, 5 Gheorghe Dima Str., 300079, Timisoara, Romania

[4]Clinical County Hospital Timisoara – Department of Surgery, 5 Gheorghe Dima Str., 300079, Timisoara, Romania

[5] West University of Timisoara, Faculty of Mathematics and Computer Science, Department of Mathematics, 4 Vasile Parvan Str., 300223, Timisoara, Romania

[6] Victor Babes University of Medicine and Pharmacy Timisoara, Department XIII, Discipline of Neonatology, 2 Eftimie Murgu Sq., 30041, Timisoara,Romania

[7]Victor Babes University of Medicine and Pharmacy Timisoara, Department of Obstetrics and Gynecology, 2 Eftimie Murgu Sq., 30041, Timisoara, Romania

The rational design of treatment is affected by two contradictory requirements: for diagnosis it is necessary to take into account all the factors of influence in the disease system, which requires long time and additional costs and also It is necessary to minimize the number of clinical determinations or the optimization of the medical act. A solution for eliminating this contradiction would be to reduce the number of parameters. The purpose of this study is to reduce the number of parameters by selecting factors with major influence in the course of the disease .This study began with the inventory of the diagnostic factors registered in the clinical observation sheet of 140 patients diagnosed with chronic myelogenous leukemia at the Hematology Department of Timisoara during January 2006 - January 2016. The statistical method of rank correlation used is a very effective solution of the problem, applicable in the medical field. The statistical hierarchy of the diagnostic criteria determines their ranks and allows focusing attention on the most important ones. The presented statistical study creates the theoretical and practical premises for the treatment to be differentiated, the disease to be recognized in a period as early as possible.

Keywords: chronic myeloid leukemia, statistical hierarchy

One of the current difficulties in medical research is the large number of factors of influence that affect the various diseases and the fact that they are not reproducible. A solution to this problem would be to reduce the number of factors by selecting factors with major influence in the course of the disease and eliminating those insignificant from a statistical point of view and focusing only on the most important ones. Opportunities in this regard offer several methods of statistical research: the random balance method, the full or fractional factorial experiment method, the correlation method of rank [1-5]

In this study we have chosen the method of correlation of rank which aims to establish the criteria for the diagnosis of myeloid leukemia, to prioritize these criteria, in order of importance attributed to them by specialists, to highlight the important ones and to eliminate those less important. For this, n = 12 specialists from the field, doctors, teachers of the Victor Babes University of Medicine and Pharmacy Timi°oara were selected, working as specialists at the Clinical County Hospital Timisoara - Hematology Department.

The purpose of this study is to reduce the number of parameters by selecting factors with major influence in the course of the disease and eliminating those insignificant from a statistical point of view and then concentrating attention only on those important in order to reasonably conceive a treatment based on an objective diagnosis.

Results and didscussions

Material and methods

In the study that led to the design of this work, an inventory of diagnostic factors enrolled in the clinical observation sheet of 140 patients diagnosed with chronic myeloid leukemia at the Hematology Department in Timisoara between January 2006 and January 2016 was started. The statistical method of correlation the rank used in this study for the hierarchy of significant diagnostic factors is also known as the psychological experiment method, it is little applied in Romania, but it is a particularly effective solution to the problem, applicable in the medical field.

Results and discussions

In order to achieve the proposed objective, several stages were required.

In the first phase, the survey organizer, based on previous experience and sources mentioned, determines which diagnostic criteria are taken into account, taken from the patient's records. A number of 12 criteria have been established, Xj, j = 1 ... m m = 12. Inquiry form is prepared according to the model shown below, which is handed to the specialists Si, i =1 until min. 12 together with the purpose of the investigation and the instructions for completing the forms. Specialists are asked to give each diagnostic criterion a score in descending order of importance that is to order the criteria. Table 1 presents

* email: mdioanaionita@yahoo.com; danielariacob@yahoo.com # Authors with equal contribution

Crit. / Spec.	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
S_1	11	1	8	9	4	5	6	7	10	3	2	5
S_2	5	1	10	10	5	8	4	9	7	3	2	6
S_3	11	1	9	8	4	5	6	7	10	3	2	5
S_4	9	5	6	6	2	7	1	8	5	2	3	4
S_5	6	2	2	2	3	4	1	3	5	3	4	5
S_6	7	1	4	1	2	5	1	3	6	2	2	2
S_7	12	1	3	6	2	8	7	11	9	10	4	5
S_8	3	7	6	2	5	11	1	8	10	4	12	9
S_9	6	1	6	6	1	7	1	3	5	2	3	4
S_{10}	7	1	2	4	6	8	5	10	12	9	3	11
S_{11}	7	1	2	4	6	8	5	10	12	9	3	11
S_{12}	10	4	6	7	3	11	1	5	9	6	2	8
	94	26	64	65	43	87	39	84	100	56	42	75
$\sum a_{ij}$	94	26	64	65	43	87	39	84	100	56	42	75
Place	**XI**	**I**	**VI**	**VII**	**IV**	**X**	**II**	**IX**	**XII**	**V**	**III**	**VIII**

Table 1
PRIMARY HIERARCHY OF DIAGNOSTIC CRITERIA FOR CML

the hierarchies proposed by the 12 specialists, in the final lines of the table, highlighting the sum of the ranks awarded by the specialists, Σa_{ij} as well as the resulting primary (natural) hierarchy, I, II. ..XII.

In the next phase, the degree of consensus of the experts, by comparing the actual frequency of the average frequency of the views to those skilled in the views, by using the criterion (chi square). If χ_c^2 calculated (with mathematical formulas) is higher than χ_c^2 in the table (table 3), then the opinions of the specialists agree, the differences of opinion being not significant. Otherwise, the differences of opinion are statistically significant and the opinions of those specialists who are the furthest away from the average opinion can be dropped. Reconciliation of opinion is reconsidered and, if satisfied after eliminating the inappropriate ones, the next step is taken.

Using the formulas: 1, 2, 3 from table 2 is constructed, and with the data from table 2 and relations 4, 6 are computed χ_c^2 and compared with χ_c^2.

As $\chi_c^2 = 61.96$ is higher than $\chi_t^2 = 19.7$, for f = 12 - 1 = 11 degrees of freedom and confidence level $\alpha = 0.05$, the hypothesis of the consistency of expert opinion is checked.

Investigation form with diagnostic criteria for CML:

Diagnostic criteria Xj, j = 1 ... m , m = 12.

X1 - the Sokal index
X2 - number of white blood cells
X3 - the presence of myeloblasts and promyelocytes in peripheral blood
X4 - the presence of myeloblasts and promyelocytes in the bone marrow
X5 - spleen size
X6 - the G / E ratio
X7 - the presence of the Philadelphia chromosome
X8 - F.A.L.
X9 - Anemia
X10 - the presence of basophils in peripheral blood
X11 - left shift deviation of leukocyte formula
X12 - Platelet count

Formula 1:

$$\Delta_j = \left| \sum_{i=1}^{m} a_{ij}\delta \; - \frac{1}{n} \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}\delta \right|$$

δ = weighting coefficient ; a_{ij} = sum of the column ranges ;

n = criteria (12) , m = specialists (12).

Formula2:

$$a_j = \frac{1}{m} \sum_{i=1}^{m} a_{ij}\delta, \; M_j = \frac{\sum_{j=1}^{n} a_{ij}\delta}{\sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}\delta}$$

Formula 3:

$$R\,a_j = a\,ij_{max} - a\,ij_{min}$$

Formula 4:

$$\chi_c^2 = m\,(n-1)\,W$$

Formula 5:

$$W = \frac{12m\sum_{j=1}^{n}\Delta_j^2}{\left[m\left(n^3-n\right)-\sum_{i=1}^{m}T_i\right]\left(\sum_{i=1}^{m}\delta_i\right)^2}$$

Formula 6:

$$T_i = \sum_{j=1}^{n}\left(t_j^3 - t_j\right); \; \sum_{i=1}^{n} T_i = 120$$

$$W = \frac{12\cdot 12\cdot 21622{,}128}{\left[12\left(12^3-12\right)-120\right]\cdot 18^2} = \frac{12\cdot 12\cdot 21622{,}128}{20472\cdot 324} =$$

$$= 0{,}4694135 \;\; \text{(Coefficient of concordance)}$$

$$\chi_c^2 = 12\cdot(12-1)\cdot 0{,}4694135,, = 61{,}962582$$

$$\chi_t^2 = 19{,}7\;(tab.3)\;;\; \text{Confidence level} = 0{,}05$$

Interpretation of results and decision-making can be done based on graphical representations in different forms. Histograms are most often used. Such a graphical representation (figure1) allows the hierarchy of diagnostic

Table 2
MATHEMATICAL RELATIONS FOR CALCULATING χ^2

x_j	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
Δ_j	55.325	71.82	5.085	6.065	37.365	42.97	52.45	35.79	63.715	13.53	35.385	23.9
Δ_j^2	3060.8556	5158.1124	25.857225	36.784225	1396.1432	1846.4209	2751.0025	1280.9241	4059.6012	183.0609	1252.0982	571.21
a_j	14.360416	3.765	9.32625	9.2445833	6.63625	13.330833	5.3791666	12.7325	15.059583	8.6225	6.80125	11.741666
n_j	9	6	9.5	9.5	4	6.5	6	5.5	5,5	7.5	10	6

Table 3
CRITERIA χ^2_t

Freedom degrees f	Level of confidence α				Grades of freedom f	Level of confidence α			
	0.50	0.10	0.05	0.01		0,50	0.10	0.05	0.01
1	0.45	2.71	3.84	6.64	14	13.3	21.1	23.7	29.1
2	1.39	4.61	5.99	9.21	15	14.3	22.3	25.0	30.6
3	2.37	6.25	7.81	11.3	16	15.3	23.5	26.3	32.0
4	3.36	7.78	9.49	13.3	17	16.3	24.8	27.6	33.4
5	4.35	9.24	11.1	15.1	18	17.3	26.0	28.9	34.8
6	5.35	10.6	12.6	16.8	19	18.3	27.2	30.1	36.2
7	6.35	12.0	14.1	185	20	19.3	28.4	31.4	37.6
8	7.34	13.4	15.5	20.1	21	20.3	29.6	32.7	38.9
9	8.34	14.7	16.9	21.7	22	21.3	30.8	33.9	40.3
10	9.34	16.0	18.3	23.2	23	22.3	32.0	35.2	41.6
11	10.3	17.3	19.7	24.7	24	23.3	33.2	36.4	43.0
12	11.3	18.5	21.0	26.2	25	24.3	34.4	37.7	44.3
13	12.3	19.8	22.4	27.7	26				

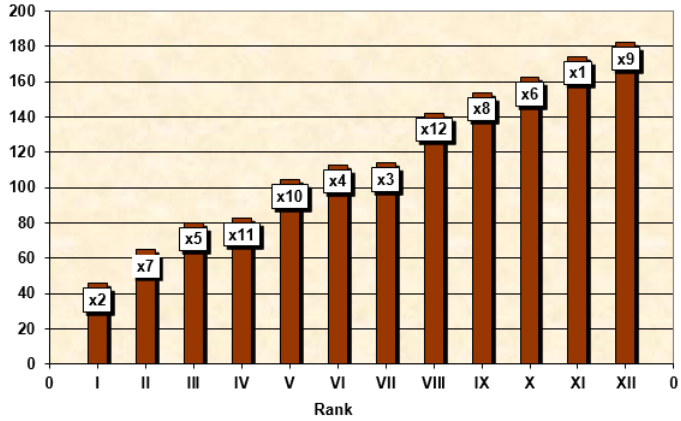

Fig. 1. Hierarchy (rank) of diagnostic criteria in CML in order of importance

criteria to be visualized, but also the intensity of the preferences to be compared by comparing the scores obtained by each diagnostic criterion with the help of the statistical criterion K. The calculated value, Kc is compared to the table value Kt. If for a number of ν criteria, Kc < Kt then all the criteria are of the same rank.

As shown in figure 1 after the statistical processing of the first five places, the following criteria are placed: X2 = number of leukocytes (I), X7 = presence of chromosome Ph (II), X5 = splenomegaly (spleen size) (III), X11 = left shift deviation of leucocytes (IV) , X10 = presence of basophils in peripheral blood (V). The other factors are of lesser importance but not negligible. The confidence level of the statistical analysis is over 95% ($p < 0.05$) for a significance threshold $\alpha = 0.05$ ($\alpha = 1 - p$).

The statistical hierarchy of the diagnostic criteria more objectively determines their ranks and allows focusing attention on the most important ones, and monitoring their remission. Establishing a lifetime prediction is possible by making a right decision based on multiple therapeutic options. For these reasons, knowing the importance of the rank of each criterion in disease progression established by this statistical study is very important [6,7]. In medicine, as in other areas of science, the law of process interdependence is well known. That is why for the complete knowledge of the phenomena it is necessary to interpret them from the point of view of mutual relations. Statistical correlation means the link of dependence, causality between two or more variables of phenomena, systems [8-12].

Table 4
THE MAIN DIAGNOSTIC CRITERIA IN CML, STATISTICALLY DETERMINED

Rank	Criteria (variable)	Defining elements of diagnostic criteria
I	Number of white blood cells	- granulated leukocytes comprising neutrophils, eosinophils and basophils, - monocellular, including lymphocytes and monocytes
II	The presence of chromosome Ph	-Ph1, abnormally small chromosome derived from a small acrocentric chromosome (cr.21 or 22) - occurs in the leucocyte cultures of many patients affected by CML.
III	Splenomegaly	- myeloproliferative syndrome characterized by erythroblastic metaplasia of the spleen (liver) and a progression to myelofibrosis, - thrombocytopenia - low platelet count, - hemodilution, erythrocyte depletion, - haemolytic anemia - disintegration of oxygen-carrying erythrocytes via Hb, - leukopenia - reducing the number of white blood cells, - sometimes (10%) adenopathies and hepatomegaly.
IV	Deviation to the left of leukocyte formula	- increased percentage of younger neutrophils in the blood. - Neutrophils are mature leukocytes with a core having 3 to 5 bonded lobes.
V	The presence of basophils in peripheral blood	- basophil is a leukocyte that contains in the cellular substance (cytoplasm), large granules which are colored with basic dyes (the nuclei are colored), - Basophilia is the abnormal increase in basal blood leukocytes.

The correlation method is a method of investigating the total or partial interdependence of variables that define a system, and therefore the method can be transplanted into CML as leukemic disease as a complex system. Correlations can be direct or indirect. In the case of direct or inverse functional correlations, any modification of an X size necessarily occurs in the modification of another Y-size, in the same sense or opposite to the same quantity or in different proportions.

In biomedical phenomena we rarely find functional bonds. Here, even if two or more variables are interrelated, increasing or decreasing them in the same or opposite direction does not necessarily occur with the same amount and uniformity [13-15]. It was seen that after the processing of the statistical material, indicators were shown which were presented graphically, illustratively. From table 4 we can see that between their parameters and their ranks there are functional links in the sense that many of them are defined by the same blood elements with direct or inverse correlations. The effectiveness of the diagnostic and therapeutic medical act in the CML cannot be accomplished beyond the minimum qualification of information on the rank occupied by different diagnostic criteria and without knowing these correlation links.

The analysis on samples of patients, following the treatments applied, indicates the existence of this type of statistically established correlations [16-19].

Conclusions

The presented statistical study creates the theoretical and practical premises for the treatment to be differentiated, the disease to be recognized in a period as early as possible, closer to its clinical onset. In this way, it is possible that on the one hand the number of patients who can reach irreversible advanced stages may be reduced, and on the other hand it favors the detection of the disease in a phase useful for differentiated, medical or chemotherapeutic treatment that results in a remission or a postponement of metamorphosis in a treatment-resistant phase.

References

1.IONITA, H. et al. Clinical Hematology. Ed Victor Babes Timisoara 2015, p. 250.
2.NICHICI, A., CICALA, E. Processing of experimental data. Polytechnic University Timi°oara 1996, p. 34.
3.POP, O., NISTOR, F. General Epidemiology (Chapter Statistics) Ed. Helicon, Timi°oara, 1996, p. 52.
4.WILLIAMS HEMATOLOGY**, 9th edition,** Ed. Mc Graw – Hill Education USA, 2016, p.1023.
5.WINTROBE'S CLINICAL HEMATOLOGY, 13th Edition, Ed. Lippincott Williams & Wilkins, 2014, p. 645.
6.KANTARJIAN, H.M., GILES, F., BHALLA, K.N. et al. Blood, **117**, 2011, p. 1141.
7.HASFORD, J., BACCARANI, M., HOFFMANN, V. et al. Blood, **118**, 2011, p. 686.
8.*** PDQ Adult Treatment Editorial Board. Chronic Myelogenous Leukemia Treatment (PDQ®): Health Professional Version. 2017.
9.SAWYERS, CL. N Engl J Med, **340**(17), 1999, p. 1330.
10.DRUKER, B.J., SAWYERS, C.L., KANTARJIAN, H. et al. N Engl J Med, **344**(14), 2001, p. 1038.
11.KANTARJIAN, H., SAWYERS, C., HOCHHAUS, A. et al. N Engl J Med, **346**(9), 2002 p. 645.
12.MERX, K., MULLER, M.C., KREIL, S. et al. Leukemia, **16**(9), 2002 p. 1579.
13.TALPAZ, M., SILVER, R.T., DRUKER, B.J. et al. Blood **99**(6), 2002 p. 1928.
14.KANTARJIAN, H.M., CORTES, J.E., O'BRIEN, S. et al. Blood, **101**(1), 2003, p. 97.
15.SHAH, N.P., TRAN, C., LEE, F.Y. et al. Science, **305**(5682), 2004, p. 399.
16.HOCHHAUS, A., LARSON, R.A., GUILHOT, F., RADICH, J.P., BRANFORD, S., HUGHES, T.P. et al. N Engl J Med. **376**(10), 2017 p. 917.
17.*** 'Imatinib Changed Everything': The Future Is Now More Hopeful. Medscape Medical News..March 9, 2017; Accessed: March 10, 2017.
18.ISLAMAGIC, E., HASIC, A., KURTOVIC, S. et al. Clin Lymphoma Myeloma Leuk. 17(4), 2017, p. 238.
19.CASTELLINO, A.M. Medscape Medical News. December 5, 2017; Accessed: March 10, 2017.

Manuscript received: 28.03.2017