Clinical Medicine Research
2019; 8(3): 56-62
http://www.sciencepublishinggroup.com/j/cmr
doi: 10.11648/j.cmr.20190803.11
ISSN: 2326-9049 (Print); ISSN: 2326-9057 (Online)

Classification of Incident Types of Hematologic Malignancy Using Discriminant Analysis at Kinshasa University Clinics, DR Congo

Mireille Solange Nkanga Nganga[1], Benjamin Longo-Mbenza[2], Fons Verdonck[3], Branly Kilola Mbunga[4], Aimé Tshiyamu Mbaya[1], Christian Lueme Lokatola[2], Teke Apalata[2], Jean-Marie Ntumba Kayembe[1], Georges Lelo Mvumbi[1], Blaise Matondo Ma Nzambi Sumbu[1], Alain Nzonzila Nganga[5], David Muballe[2], Cecile Roth Laure Miakassissa Mapapa[6], Paul Roger Kazadi Beia[7], Aurore Cecilia Orphée Mbombo Beia[8], Donatien Nzongola Nkasu Kayembe[1]

[1]Faculty of Medicine, Department of Medical Biology, University of Kinshasa, Kinshasa, DR Congo

[2]Faculty of Medicine, Department of Research, Walter Sisulu University, Mthatha, South Africa

[3]UZ Gasthuisberg/ KU Leuven, Department of Education, Leuven, Belgium

[4]Public Health School, Department of Nutrition, University of Kinshasa, Faculty of Medicine, Kinshasa, DR Congo

[5]Saint Joseph Hospital, Department of Reanimation, Kinshasa, DR Congo

[6]Service of Pediatry, Department of Neonatology, University Hospital Center of Brazzaville, Brazzaville, Republic of Congo

[7]OVD Hospital Centre, Department of Health, Kinshasa, DR Congo

[8]Faculty of Medicine, Biomedical School of Medicine, University of Kinshasa, Kinshasa, DR Congo

Email address:

mireille.nganga@unikin.ac.cd (M. S. N. Nganga), mnganga2002@yahoo.fr (M. S. N. Nganga)

To cite this article:

Mireille Solange Nkanga Nganga, Benjamin Longo-Mbenza, Fons Verdonck, Branly Kilola Mbunga, Aimé Tshiyamu Mbaya, Christian Lueme Lokatola, Teke Apalata, Jean-Marie Ntumba Kayembe, Georges Lelo Mvumbi, Blaise Matondo Ma Nzambi Sumbu, Alain Nzonzila Nganga, David Muballe, Cecile Roth Laure Miakassissa Mapapa, Paul Roger Kazadi Beia, Aurore Cecilia Orphée Mbombo Beia, Donatien Nzongola Nkasu Kayembe. Classification of Incident Types of Hematologic Malignancy Using Discriminant Analysis at Kinshasa University Clinics, DR Congo. *Clinical Medicine Research*. Vol. 8, No. 3, 2019, pp. 56-62. doi: 10.11648/j.cmr.20190803.11

Received: November 29, 2018; **Accepted**: January 29, 2019; **Published**: July 22, 2019

Abstract: The objective of this study was to identify important biomarker differences between absence of HM and expected morphopathologic types of HM. A retrospective analysis study of adult patients aged ≥ 20 years was managed by cytologic aspects such as normal myelogram vs. HM types between 2009 and 2015. Out of 105 patients, 63 (60%) experienced incident HM while 42, 14, 18, 10, 10, 6, and 5 patients had normal myelogram, multiple myeloma (MM), acute myeloid leukaemia (AML), myelodysplastic syndromes (MDS), chronic myeloid leukaemia (CML), acute myeloid leukaemia (CLL) and acute lymphoid leukaemia (ALL), respectively. In Discriminant Analysis (DA), only levels of transfusion, Hb, and WCC discriminated significantly (Wilks lambda =0.159; $P < 0.0001$) the study groups through Function 1 [Eigen value (EV) = 2.591; cumulative variance (CV) = 78, 7% and Canonical correlation (CC) = 0.849], Function 2 (EV = 0.619; CV = 97.5%; CC = 0.618), and Function 3 (EV = 0.081; CV = 100%; CC = 0.274). The highest Mahalanobis distance (Min D Squared = 0.162) was observed between CML and MDS. For early diagnosis, precise medicine, and good practice in hematologic oncology, DA separated CML, MDS, MM, AML, CLL, and ALL from normal myelogram in Congolese patients.

Keywords: Hematologic Malignancy Types, Classification, Statistics, Congolese Patients

1. Introduction

Globally malignant tumors are becoming a leading cause of morbidity, mortality and disability in both rich and disadvantaged countries [1-5]. Specifically, hematologic malignancy sub-types are emerging whilst their epidemiologic features [5-8] and diagnosis remain challenging in low-income settings such as Asia and sub-Saharan Africa [8-11]. Lymphoid types found in 70.96% of patients [12] define the burden of hematologic malignancy classification. However, there are no comprehensive classifications of hematologic malignancy sub-types among patients managed in Kinshasa University Clinics (KUC), a tertiary academic hospital in Democratic Republic of the Congo (DRC). Therefore, the objective of this study was to identify independent variables, which are significantly associated with hematologic malignancies using discriminant analysis (DA) among black patients diagnosed with anemia at KUC.

2. Materials and Methods

This was a retrospective analysis study of adult patients aged ≥ 20 years diagnosed with anemia. Patients were investigated further using cytology in order to either confirm or rule out the presence of an underlying HM subtype. The study was conducted between 2009 and 2015 at KUC, Department of Chemical pathology. Because there was insufficient and inaccurate diagnosis as well as other information of incident HM from African literature [1, 13].

For the sample size (n), the total number of patients was $NI=4x(Zx)^2$ x IC $(1\text{-}IC)/w^2$ where 2x=1.96. The standard normal deviation for a two-sided α=0.05, 1-0.05=95% Confidence level, and W=total with of confidence level =0.20. Then the total number of patients with anemia required was Ni=96 rounded to 100 (+20%) of potential misses = 120.

The variables of interest were gender, age, year of initial diagnosis, outcome (fatal or non-fatal), clinical features, blood counts, peripheral blood films and bone narrow morphology including cytochemical staining techniques. Results from the latter tests were red and interpreted by two independent senior pathologists in the department. In cases where discrepant findings were observed between the two pathologists, a third opinion was sought.

2.1. Definitions

Anemia was defined by hemoglobin < 12 mg/dL in men and hemoglobin < 11mg/dL in women. Blood variables included hematologic parameters such as hemoglobin hematocrit, white Cell Count (WCC), platelets, and erythrocyte rate sedimentation (ERS) at 1 hour (1H). Hematological malignancy morphological sub-types were diagnosed according to French America British (FAB) Classification (2008) and supplemented by WHO criteria for hematologic cancers (2008).

2.2. Statical Analysis

Continuous variables were expressed as means ± standard deviation (SD) when normally distributed, while categorical variables were presented as frequency (count=n) and proportions (%).

In univariate analysis, student's t-test was calculated to assess differences between 2 groups and analysis of variance (ANOVA) to compare means between ≥3 groups. Multiple comparisons of means were computed using Post Hoc Bonferroni pair means at considering TYPE I error rate of 0.05. Chi-Square test was used to compare percentages of categorical variables between groups.

In multivariate analysis, DA was used as the model of the conventional classification techniques at discriminating a single categorical variable using multiple attributers such as normal myelogram and different sub-types of hematological cancers. DA used canonical variables that would maximally differentiate (classify) group membership within patients with anemia. The important underlying assumptions of DA were stated as follow: (i) each predictive variable was normally distributed; (ii) there must be homogeneity of covariance between sub-types and normal myelogram; (iii) there must be at least 2 groups with each sub-group belonging to only one group so that the groups were mutually exclusive and collectively exhaustive; (iv) the groups should be characterized before collecting the data; (v) the predictive variables considered to separate the groups should classify quite clearly between the groups so that each category overlap was clearly non-existent or minimal; (vi) and groups sizes of the dependents should not be grossly different and should be at least 5 times the number of independent variables.

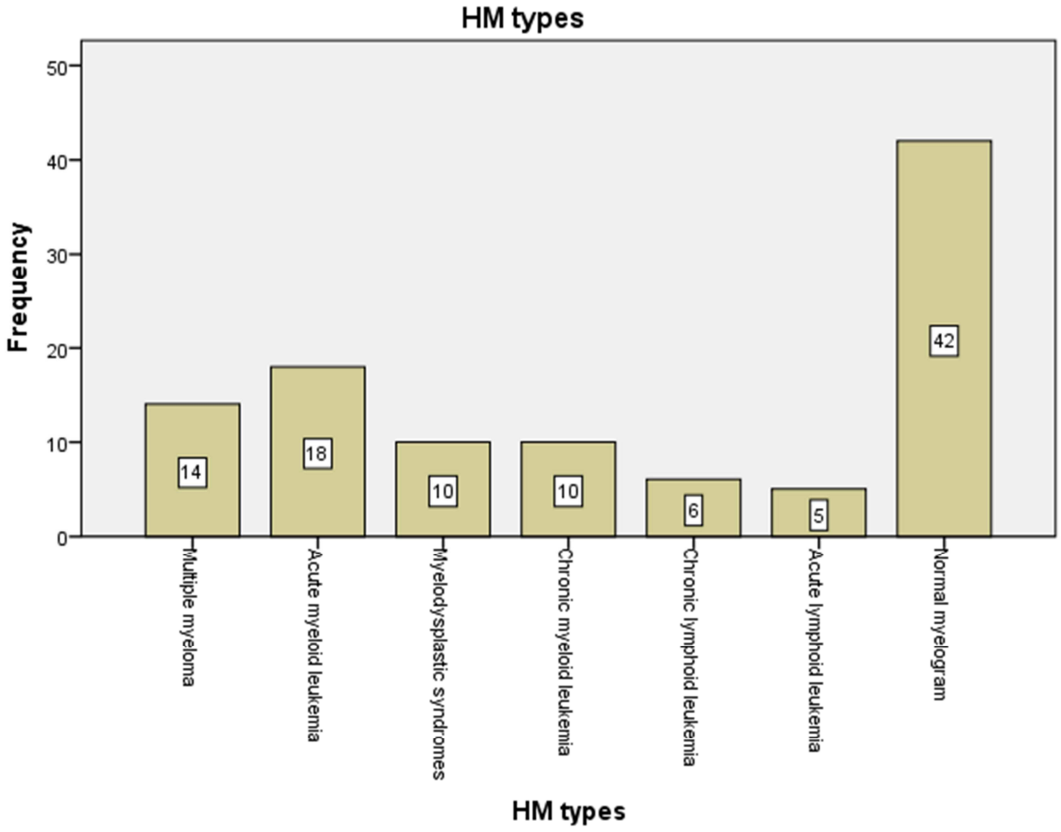
The box's test of Equality for covariance Matrices was considered to check the assumption for homogeneity of covariance across the categories.

Mahalanobis distances were computed at supporting the classification of canonical variates into distinct sub-types and normal myelogram among groups' centroids at determining the degree of segregation with each Wilk's Lambda value closer to zero being the evidence for well-separated groups. Thus, a multi-dimensional generalization of the idea of measuring the distance between a point P and a distribution D mean (how many standard deviations away from P).

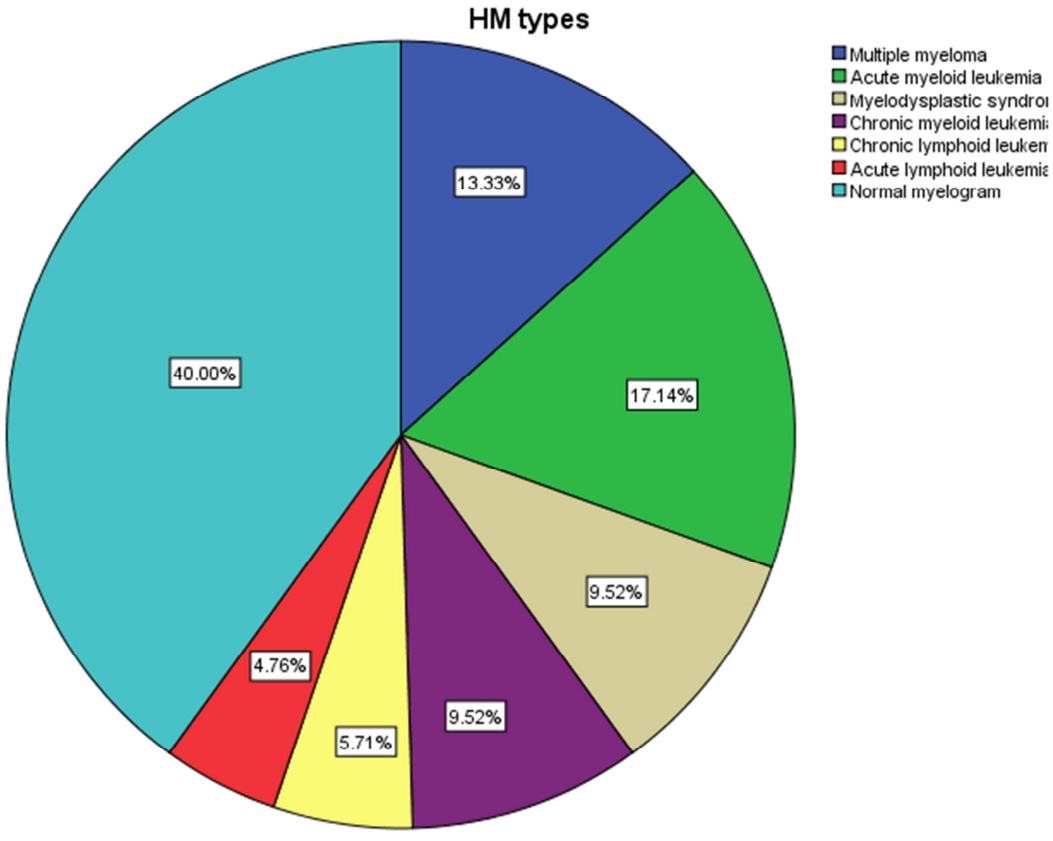
A P-value < 0.05 was considered with significant differences. All analyses were computes using Social Package for Social Sciences (SPSS) version 22.0.

3. Results

In total, 105 patients were diagnosed with anemia. Some patients had an underlying infectious syndrome (fever, leucopenia, and hyperleucocytosis) whilst HM was suspected of being the underlying cause of anemia among the rest. All patients were managed at CUK. Figure 1 AB depicts the frequency and proportions (%) of HM types/subtypes such as MM, AML, MDS, CML, and CLL, ALL. Of 105 patients, HM was diagnosed in 63 (60%) patients. Among HM types observed, the highest extents were in AML (28.6% n = 18) and MM (22.2% n = 14). The rest of HM types included 8% (n = 5)-16% (n = 10) of the rest of HM types such ALL, CLL, CML and MDS.



A



B

Figure 1. *Distribution of frequency (A) and proportion (B) of the study population with HM types and normal bone marrow.*

Proportions of HM types did not vary (P > 0.05) between males and females (results not shown). The median age was 50 years for HM patients.

Table 1 present the comparison of means values of conventional markers across the management of HM types using ANOVA following a univariate analysis. As compared to cases without HM, all patients with HM were of old age irrespective of the type of HM. The lowest hemoglobin level was found in patients with AML, the highest number of previous blood transfusion was determined in MM cases, and as expected, the highest levels of white cell counts (WCC) were determined in patients with CML. Paradoxically, WCC were elevated in patients diagnosed with CLL but showed to be the lowest in cases with ALL and AML. The highest levels of ESR were found in patients with MM.

Furthermore, Post Hoc test Bonferroni did not show significant differences for the majority of markers between HM groups. However, this Post Hoc test demonstrated significant delineating means of each HM type in comparison with counterpart levels from bone marrow normality group.

In multivariate DA, tables 2-6 summarized tests of equality group means with Wilks Lambda, boxes test of equality of covariance matrices, summary of canonical discriminant functions, canonical discriminant function, and classification function coefficients. For caution, first 3 canonical discriminant functions were used in the analysis showed in the table 4.

After DA, only numbers of previous transfusion, Hb, and WCC discriminated signicantly (Wilks lambda= 0.159; P< 0.0001) the study groups through Function 1 [Eigen value (EV) = 2.591; cumulative variance (CV) = 78.7% and canonical correlation (CC) = 0.849], and Function 2 (EV+0.619; CV=97.5%; CC= 0.274). The highest Mahalanobis distance (Min D Squared= 0.162) was observed between CML and MDS.

Figure 2 presents highest Mahalanobis distances to discriminate HM types using discriminant score for each subject from the centroid group.

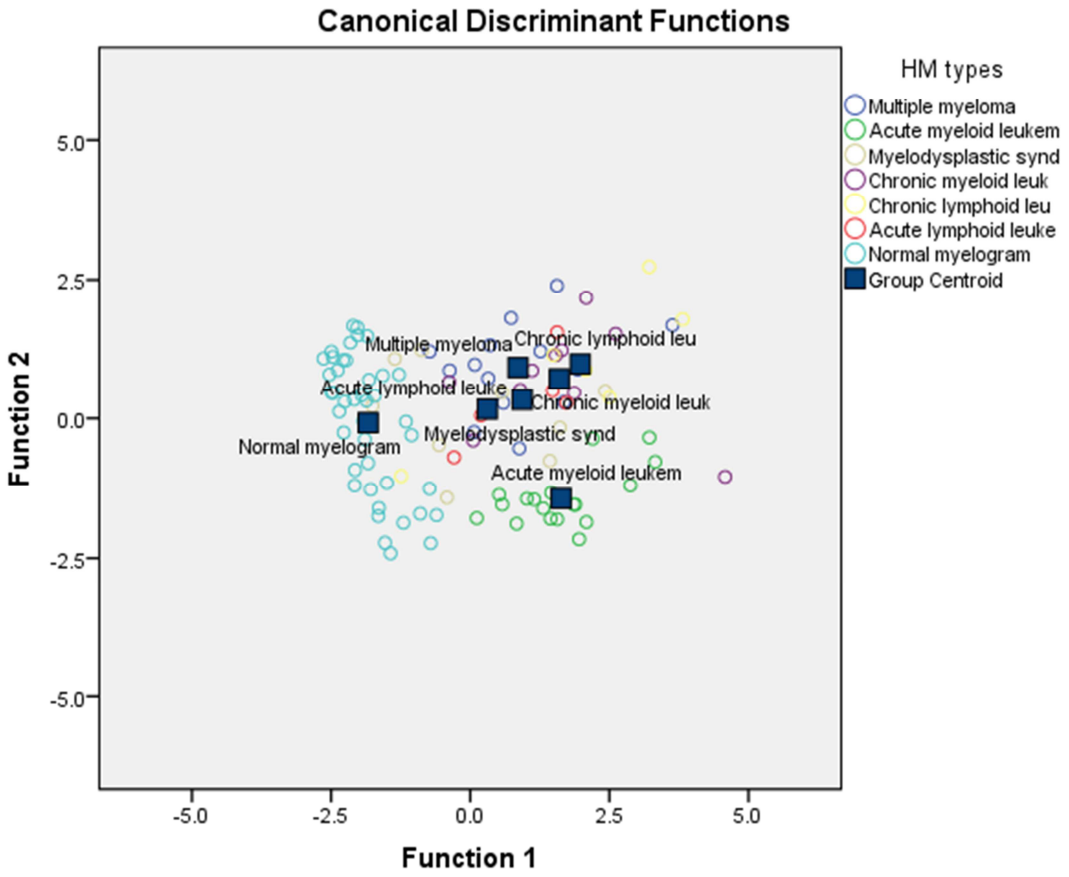


Figure 2. *Highest Mahalanobis distances to discriminate HM types using discriminant score for each subject by the centroid.*

Table 1. *Comparison of mean values of age, therapeutic, and conventional biomarkers across HM subtypes.*

nIndependent Variables	MM 14 X±SD	AML 18 X±SD	MDS 10 X±SD	CML 10 X±SD
Age (years)	60.8±12.1	65.1±14.6	65.5±10.9	53.5±20.6
Blood Transfusion (number)	9.3±2.4	9.6±2.8	8±3.9	10.3±4.1
Hb (g/dL)	8.6±1.8	3.6±0.8	7.3±2	8±1.7
White cell count (mm^3)	22299.3±31759	24509±11351	11449±10.000	34536±28552
PLT (mm^3)	2475535±135112	130361±15825	218750±1392274	155956±70762
ERS ($mm/1^{er}H$)	105±12	134±20	79±53	95±59

Table 1. *Continued.*

nIndependent Variables	CLL 6 X±SD	ALL 5 X±SD	Absence 42 X±SD	Anova p
Age (years)	59.5±19	54.2±24	34.5±13.8	<0.0001
Blood Transfusion (number)	12.8±5.1	8.8±4	2.2±1.1	<0.0001
Hb (g/dL)	7.9±1.8	7.6±1.1	8.2±2.6	<0.0001
White cell count (mm^3)	16998±15.000	25960±15.968	7029±4053	<0.0001
PLT (mm^3)	289166±201974	127557±28822	244107±151751	<0.009
ERS (mm/1^{er}H)	80±47	140±19	111±44	<0.013

Table 2. *Tests of Equality Group Means.*

Variables	Wilks' Lambda	F	dfl	df2	P
Age (years)	0.645	9.004	6	98	<0.0001
Blood Transfusion (number)	0.333	32.698	6	98	<0.0001
Hb (g/dL)	0.573	12.151	6	98	<0.0001
White cell count (/mm^3)	0.752	5.381	6	98	<0.0001
PLT (/mm^3)	0.844	3.024	6	98	<0.009
ERS (mm/1^{er}H)	0.850	2.881	6	98	<0.013

Table 3. *Box's Test of Equality of Covariance Matrices: Log Determinants.*

HM types	Rank	Log Determinant
MM	3	23.127
AML	3	20.205
MDS	3	23.169
CML	3	23.577
CLL	3	23.135
ALL	3	18.715
Absence	3	22.912

Table 4. *Summary of canonical discriminant Functions.*

Eigenvalues

Function n	Eigen values	% of variance	Cumulative%	Canonical correlation
1	2.591	78.7	78.7	0.849
2	0.619	18.8	97.5	0.618
3	0.081	2.5	100.0	0.274

Table 5. *Canonical Discriminant Function.*

Unstandardized coefficients

	Function		
	1	2	3
Blood transfusion (number)	0.323	0.116	-0.113
Hb (g/dL)	-0.125	0.475	0.023
White cell count (mm^3)	0.000	0.000	0.000
Constant	-1.719	-4.209	0.296

Table 6. *Classification Function Coefficients.*

	HM types						
	MM	AML	MDS	CML	CLL	ALL	Absence
Blood transfusion (number)	1.403	1.389	1.187	1.564	1.858	1.335	0.404
Hb (g/dL)	2.121	0.906	1.823	1.943	1.995	1.841	1.985
White cell count (/mm^3)	7.425E-005	9.974E-005	3.678E-005	0.000	6.720-005	8.967E-005	2.967E-005
Constant	-18.453	-11.457	-13.577	-19.917	-22.299	-15.943	-10.528

4. Discussion

This is the first Sub-Saharan study to use Anova and discriminant function analysis at distinguishing among HM subtypes based on some conventional hematologic and therapeutic markers at CUK, DRC.

4.1. Univariate Analysis

The present study showed significant associations between aging, decrease in hemoglobin but marked increase in both blood transfusion, white cell count, platelet count, ESR and HM. However, sex did not impact on these incident HM types.

Thus, the present study confirmed the role of longevity on HM epidemics in both developed countries [9, 14], and low and middle-income countries including African countries such as DRC [8, 15]. Concurrent presence of anemia and aging is explained by emerging cancers well reported by previous different Authors [16, 17]. Aging itself is a strong oxidant process [18] There is also evidence about insignificant relationship of aging and the pathobiology of the clonal myeloid diseases (leukemias) [19].

This study conformed highest proportion of AML and MM as previously reported in Kinshasa, RDC [15] and in the other parts of Africa [20].

4.2. Multivariate Analysis

The present study showed that multivariate mathematical functions obtained by combining numbers of previous blood transfusion, hemoglobin levels and WCC were capable of discriminating different types of HM including MM, AML, MDS, CML, CLL, ALL as well as patients with normal bone marrow findings (absence of HM) with a dummy-coded dependent variable.

4.3. Explanation and Prediction

In this study, DA derived which set of conventional markers most and significantly related to these HM subtypes and way for predicting HM subtype membership.

Thus, the most discriminant functions derived were one less than 7 groups in the dependent variable HM or bone marrow normality of independent variables (blood transfusion, hemoglobin and WCC).

4.3.1. Function by Function

The first discriminant function (cumulative variance = 78.7%) derived from the data explained most of the between-group variances; the second discriminant function (cumulative variance = 97.5%) explained the next largest piece of variance and as did the third discriminant function (cumulative variance = 100%). This means that these functions were not correlated with each other. Furthermore, all the markers were entered in DA models using a stepwise strategy starting with the most discriminating markers.

4.3.2. Eigenvalues

Eigenvalues and their related correlations were applied to judge the most discriminating diagnostic indicators (previous blood transfusion, hemoglobin levels and WCC) and hematologic markers (diagnosis and severity markers).

Indeed, Eigenvalues represented the amount of variance explained by a discriminant function: Eigen value= 2.591 for Function 1, Eigenvalue = 0.619 for Function 2, and Eigenvalue= 0.081 for Function 3.

In this multiple regression, each marker was weighted when those weights produced a discriminant score for each subject as compared to the centroid group (mean of the discriminant scores of give subtype).

Both Factor analysis and DA used Principal Component's analysis on a matrix of indices rather than among markers. Then, Varimax Rotation was used to increase the interpretability of obtained functions in this study.

4.3.3. Wilks' Lambda

It was necessary to use Wilks' lambda at measuring the correlation between all markers (independent variables) and HM subtypes (dependent variable) in the present study. The scores helped the members of each HM subtype to be classified as the process was well functioning with correct and incorrect percentages.

DA produced raw coefficient (like b in multiple linear regression), standard coefficient (like betas), and structured coefficients (like in canonical correlation). The raw coefficient served to calculate scores for each marker. The standardized coefficients were used to express the relative importance of the independent markers with they were correlated. However, like betas, standardized coefficients, unstable, were interpreted with caution.

Structure coefficients or loadings represented the association between the discriminant score for each marker and the scores on the original variables. The square of those coefficients was the proportion of variance in a particular marker explained by the discriminant functions.

Structure coefficients of > 0.30 were considered meaningful [21].

The discriminant function estimations were very sensitive to the assumption of normality. Hosmet and Zemeshow calculated the magnitude of overestimated association.

4.4. Implication for Chemical Pathology and Perspectives for Public Health

The present findings will impact routine practice, training, capacity building, and research related to HM subtypes at CUK, RDC, a country facing complex social, economic, and political crises. Therefore, integrated and collaborative research, advocacy and prevention, are needed between Congolese universities, African states and international organizations (American Society of Hematology, NCI, West Africa College of Physician, Africa Organizations) for Research and training in Cancer related to hemato-oncology [1].

Indeed, RDC and other low and middle-income countries do experience diagnostic challenges since their capacity in hematopathology remains limited, and pathologist scarcity is of major concern due to critical health professional shortages [1].

In fact, research on HM subtypes will be deeply carried to improve early diagnosis and treatment of HM at CUK, DRC.

4.5. Strength and Limitations

Use of discriminant analysis was the strength of the present study with calculated sample size. However; this study was limited to some degree. The lack of cytogenetic, detailed clinical information, poverty, were many of these limitations. The present findings cannot be generalized to other Congolese hospitals as well as to the general Congolese population.

5. Conclusion

This is the first African study with aging and emerging HM subtypes among anemic patients with highest proportion of acute myeloid leukemia and multiple myeloma. For early diagnosis, precise medicine, and good practice in hematologic oncology are warranted. As a result, DA is found promissing in separating CML, MM, CLL, MDS, AML, and ALL from normal myelogram in Congolese patients.

References

[1] Satish Gopal, William A. Wood, Stephanie J. Lee, and al. Meeting the challenge of hematologic malignancies in sub-Saharan Africa. BLOOD, 2012; 119 (22): 5078-5087.

[2] OMS. Cancer. Aide-mémoire n 297. Mars 2017.

[3] Pierre Aubry, Bernard-Alex Gaüzère. Les cancers dans les pays en développement Actualités 2017. www.medecinetropicale.com. 16/01/2018.

[4] Facts 2015-2016 The incidence, prevalence and mortality data in Facts 2015-2016 Cancers Facts and Figures 2016 Atlanta, GA.

[5] Olaleye O., Ekrikpo U. Epidemiology of Cancers in Sub-Saharan Africa. In: Adedeji O. (eds) Cancer in Sub-Saharan Africa. Springer, 2017; 3-19.

[6] Schonfeld SJ, Erdmann F, Wiggill T, Singh E, Kellett P, and Babb C, Schüz J. Hematologic malignancies in South Africa-2006: Analysis of data reported to the National Cancer Registry. Cancer Med 2016; 5 (4): 728-738.

[7] Marshall A. Litchman Battling the hematological Malignancies: The 200 years' War. The oncologist. 2008; 13 (2): 126-138.

[8] Mohammad Sorowar Hossain, Mohd S Iqbal, Mohiuddin Ahmed Khan and al. Diagnosed hematological malignancies in Bangladesh - a retrospective analysis of over 5000 cases from 10 specialized hospitals BMC Cancer 2014, 14: 438.

[9] Mounia Elidrissi Errahhali, Manal Elidrissi Errahhali, Redouane Boulouiz, Meryem Ourzame and Mohammed Bellaoui Distribution and features of hematological malignancies in Easten Morroco: a retrospective multicenter study over 5 years BMC Cancer 2016, 159.

[10] Makani j, Roberts DJ. Hematology in Africa. Hematol Oncol Clin North Am 2016 Apr; 30 (2): 457-75.

[11] Abel GA, Klepin HD. Frailty and the management of hematologic malignancies. Blood. 2018; 131 (5): 515-524.

[12] D A Diallo, L. S Sissoko, Y Sissoko et al. Epidémiologie actuelle des hémopathies malignes dans les services d'hématologie oncologie médicale et de médecine interne de l'hôpital du Point G, Bamako, Mali, Mali Médical 2005 T XX N 4.

[13] Mashinda KD, Kayembe KP, and Mapatano MA. Prévalence du cancer en République Démocratique du Congo: données anatomopathologiques recueillies aux Cliniques Universitaires de Kinshasa et à l'Hôpital Général de Référence de Kinshasa Ann. Afr. Med. 2012; 6 (4): 1087-1093.

[14] Hassan M, Abedi-Valugerdi M. Hematologic malignancies in elderly patients. Haematologica. 2014; 99 (7): 1124-7.

[15] Mufuta N JP, Mbayo K, Kayembe NZ, Gini EK, and Mbuyi M. Cytologie des hémopathies malignes dans deux formations médicales de Kinshasa /DRCongo Ann. Afr. Med, 2013; 6 (4): 1499-1505.

[16] David P. Steensma. New challenges in evaluating anemia in older persons in the era of molecular testing. ASH Education Book, 2016; 2016 (1): 67-73.

[17] Dominique Bron, Lionel Ades, Tamas Fulop Valentin Goede, and Reinhar Stauder. Aging and Blood Disorders: New Perspectives, New Challenges. Haematologica April 2015 100: 415-417.

[18] Liguori I, Russo G, Curcio F, et al. Oxidative stress, aging, and diseases. Clin Interv Aging. 2018; 13: 757-772.

[19] Angelica de Souza Batista Maia Alves, Fernanda Barbi Bataglia, Luciene de Oliveira Conterno, Rosimeire Segato, and Spencer Luiz Marques Payão. Epidemiological and cytogenetic profiles of patients with hematological malignancies and their relationship with aging. Hematol Transfus Cell Ther. 2018 Jul-Sep; 40 (3): 200–206.

[20] Ouédraogo SM, Hien F, Bazié W, Millogo A, Drabo YJ Place des Hémopathies Malignes en Service de Médecine Interne du CHU de Souro Sanou (Burkina Faso) Mali Médical 2011 Tome XXVI (3): 17-21.

[21] Elazar J. Pedhazur. Multiple regression in behavioral research: explanation and prediction. Fort Worth: Harcourt Brace College Publishers, c1997.