Identification of Novel Classes for Patients With Lupus Nephritis Using Two-step Cluster Model

Ting Huang
Central South University

Shasha Xie
Central South University

Liqing Ding
Central South University

Hui Luo (✉ luohuihn@sina.com)
Central South University

Abstract

Objectives: To identify and reclassify the patients in the LN cohort, and to further analyze the prominent clinical features and clinical significance of each cluster of patients.

Methods: This is a cross-sectional study of a cohort of 635 LN patients from the Rheumatology Department of Xiangya Hospital of Central South University. Demographic data, laboratory findings and clinical evaluation system include physician's global assessment and the SLICC/ACR Damage Index were collected. Using two-step cluster analysis, patients with similar clinical property were identified and compared.

Results: Among the 635 LN patients, 599 patients (94.3%) were female. The mean age of the patients were 33.8 ± 10.4 years. Three subgroups were identified by two-step cluster analysis. Cluster 1 included 130 (20.5%) patients, Cluster 2 included 132(20.8%) patients and Cluster 3 included 373 (58.7%) patients. Cluster 3 was the largest group of mild disease activity, patients in this cluster had lower white blood cells, neutrophils, lymphocytes and mean SDI scores compared to those in the other two clusters. Cluster 1 was the smallest group of severe damage, patients in this cluster had multiple positive auto-antibodies, higher SDI scores and lower complement level. Patients of cluster 2 had the highest levels of granulocytes, but the results of other laboratory tests were roughly between the cluster 1 and cluster 3.

Conclusions: This study reclassified three groups of LN patients in a large cohort. Our research shows that the multiple positive ANA antibody may be related to the high SDI score of LN patients. Clinicians can identify patients at different stages through cluster analysis to better implement prognosis.

Introduction

Systemic lupus erythematosus (SLE) is an autoimmune disease involving multiple organs, renal involvement will significantly increase the morbidity and mortality of SLE.[1] Therefore, concerning on the clinical manifestations of SLE patients with renal involvement, called lupus nephritis (LN), is beneficial to actively improve the prognosis of patients.[2]

There are some classic SLE assessment guidelines. The current measure system used to assess damage to SLE is the Systemic Lupus International Collaborating Clinics (SLICC)/American College of Rheumatology (ACR) Damage Index (SDI),[3] and the basic disease assessment system for LN is the 2003 International Society of Nephrology/ Renal Pathology Society (ISN/RPS) classification system,[4, 5] therefore, clinicians could apply targeted treatment strategies for patients based on the above disease assessment classifications.[6] Despite this, 10–30% of patients with LN will progress to ESRD within 15 years of diagnosis.[7] Thence, it indicates that the classification of LN could be improved by incorporating clinical indices and pathological features to evaluate the condition of LN patients.

The main purpose of this study are to use a reasonable clustering method to reclassify the patients in the LN cohort, and to further analyze the prominent clinical features and clinical significance of each cluster

of patients.

Materials And Methods

Patients

A cross-sectional study was conducted at the Rheumatology Department of Xiangya Hospital of Central South University in Changsha, China, between November 2006 and September 2020. All patients met the American College of Rheumatology (ACR) 1997 revised classification criteria for SLE, and these patients were confirmed to be lupus nephritis by renal pathology. This study was approved by the Medical Ethics Committee of Xiangya Hospital of Central South University, according to the requirements, and all recruited patients signed a document indicating their written informed consent. All methods were carried out in accordance with the Declaration of Helsinki.

Data collection

The same evaluations were performed, and data were collected in the present study. These data included the following: demographic data, such as gender and age at recruitment; laboratory findings, including White blood cell (WBC) count, Neutrophils (N) count, Lymphocytes (L) count, Haemoglobin (Hb) concentration, blood platelets (PLT) count, Creatinine (Cr), Estimated Glomerular Filtration Rate (eGFR), Immunoglobulin G (IgG), Immunoglobulin A (IgA), Immunoglobulin M (IgM), C3, C4, antinuclear antibodies (ANA) and antibodies to double-stranded DNA (ds DNA), Smith, SSA, SSB, nucleosome, histone and anti-ribosomal P protein. Clinical evaluation system includes physician's global assessment (PGA) and the SLICC/ACR Damage Index (SDI)[8], both of them were collected from medical records and calculated by an experienced clinician. The PGA, reflecting the clinician's judgement of the disease activity, was scored, prior to reviewing the complement and anti-DNA antibody results.[9]

Statistical analysis

Cluster analysis was carried out by applying Two-step cluster model to identify those groups of LN patients with similar patterns of damage manifestation. The specific analysis principle is as follows: 1. Establish a cluster feature tree, and set the first record of the data onto a leaf node initiated by the root of the tree, which contains all the variable information about this record. Then the distance measurement is used as the similarity criterion. The records of the same tree node have high similarity, and the records of poor similarity will generate new nodes. The established cluster feature tree provides a summary of the variable information of the data set. The distance measurement model uses logarithmic similarity. This likelihood measure assumes that variables obey a certain probability distribution, and assumes that the variables in the clustering model are independent. Furthermore, it is assumed that each continuous variable has a normal (Gaussian) distribution, and each categorical variable has a multinomial distribution. Empirical internal tests have shown that the process is quite robust against violations of the assumptions of independence and distribution. We also tried to understand the extent to which these assumptions conform. Use the bivariate correlation procedure to test the independence of two continuous

variables. Use the cross tab procedure to test the independence of two categorical variables. Use the mean procedure to test the independence between continuous and categorical variables. Use the exploration processes to test the normality of continuous variables. Use the chi-square test procedure to test whether a categorical variable has a specified multinomial distribution. 2. Use to merge clustering algorithm to combine leaf nodes. A set of clustering schemes for different numbers of clusters can be generated. Then, according to the Bayesian Information Criterion (BIC) criteria, various clustering schemes are compared, and the number of clusters is automatically selected to make the clustering scheme optimal.

The two-step cluster analysis automatically determines that the final classification number is 3 categories and the outputs were compared with each other. Results were expressed as mean (S.D.) for continuous variables, and as number of patients (percentages) for binary and categorical variables. The Analysis of Variance (ANOVA) tests was used to compare continuous variables. The chi-squared test was employed to compare the frequencies of categorical variables among the three groups of patients. Statistical significance was concluded when $P < 0.05$. The data were analyzed using SPSS 26.0.

Results

Demographic and clinical characteristics of patients with LN

A total of 635 LN patients were enrolled in the study. The baseline characteristics of patients are described in Table 1. 599 patients (94.3%) were female. The mean age of the patients were 33.8 ± 10.4 years.

Table 1

Demographics and clinical characteristics of LN patients.

Patient characteristic	Value
Age at diagnosis, years	33.8(10.4)
Sex, n (%)	
Male	36(5.7)
Female	599(94.3)
WBC, (10^9/L)	6.7(2.9)
N, (10^9/L)	4.4(2.3)
L, (10^9/L)	1.8(2.0)
Hb, (g/L)	126.1(17.8)
PLT, (10^9/L)	226.6(72.6)
Cr, (μmol/L)	72.9(34.4)
eGFR, mL/min/1.73m^2	96.0(24.0)
IgG, (g/L)	13.7(4.2)
IgA, (g/L)	2.7 (1.2)
IgM, (g/L)	1.0 (0.6)
C3, (mg/L)	712.6(173.4)
C4, (mg/L)	150.4(67.4)
ANA, n (%)	586(92.3)
anti-dsDNA antibodies, n (%)	606(95.4)
anti-Sm antibodies, n (%)	38(6)
anti-RNP antibodies, n (%)	9(1.4)
anti-SSA antibodies, n (%)	32(5)
anti-SSB antibodies, n (%)	14(2.2)

Datas are given as mean (SD), median, or as number and percentage. White blood cell (WBC), Neutrophils (N), Lymphocytes (L), Haemoglobin (Hb), blood platelets (PLT), Creatinine (Cr), Estimated Glomerular Filtration Rate (eGFR), Immunoglobulin G (IgG), Immunoglobulin A (IgA), Immunoglobulin M (IgM), Complement 3 (C3), Complement 4 (C4), antinuclear antibodies (ANA), antibodies to double-stranded DNA (anti-dsDNA antibody), antibodies to Smith (anti-Sm antibodies), antibodies to ribosomal P protein (anti-rRNP antibodies), physician's global assessment (PGA). Systemic Lupus International Collaborating Clinics/ACR Damage Index (SDI).

Patient characteristic	Value
anti-rRNP antibodies, n (%)	5(0.8)
anti-nucleosome antibodies, n (%)	19(3)
anti-histone antibodies, n (%)	8(1.3)
PGA, median (range)	1(0-1.5)
SDI score, mean (S.D.)	0.3(1.2)
Datas are given as mean (SD), median, or as number and percentage. White blood cell (WBC), Neutrophils (N), Lymphocytes (L), Haemoglobin (Hb), blood platelets (PLT), Creatinine (Cr), Estimated Glomerular Filtration Rate (eGFR), Immunoglobulin G (IgG), Immunoglobulin A (IgA), Immunoglobulin M (IgM), Complement 3 (C3), Complement 4 (C4), antinuclear antibodies (ANA), antibodies to double-stranded DNA (anti-dsDNA antibody), antibodies to Smith (anti-Sm antibodies), antibodies to ribosomal P protein (anti-rRNP antibodies), physician's global assessment (PGA). Systemic Lupus International Collaborating Clinics/ACR Damage Index (SDI).	

Cluster analysis

Among the 635 LN patients, three subgroups were identified by two-step cluster analysis and then compared. Cluster 1 included 130 (20.5%) patients, Cluster 2 included 132(20.8%) patients and Cluster 3 included 373 (58.7%) patients. Characteristics within the clusters and P-values for between-cluster comparisons are shown in Table 2.

Table 2
Two-step cluster analysis: demographics and and clinical characteristics.

Patient characteristic	Cluster 1, n = 130 (20.5%)	Cluster 2, n = 132 (20.8%)	Cluster 3, n = 373 (58.7%)	P-values
Age at diagnosis, years	31.3 ± 10.0 [c]	33.3 ± 11.1	34.8 ± 10.1 [a]	0.003
Sex, n (%)				0.002
Male	5(3.8)	16(12.1)	15(4.0)	
Female	125(96.2)	116(87.9)	358(96)	
WBC, (10^9/L)	7.1 ± 3.2[bc]	9.7 ± 2.9[ac]	5.4 ± 1.6[ab]	0.000
N, (10^9/L)	4.6 ± 2.4 [bc]	6.9 ± 2.5 [ac]	3.5 ± 1.3 [ab]	0.000
L, (10^9/L)	2.7 ± 4.1 [bc]	2.0 ± 0.9 [ac]	1.4 ± 0.6 [ab]	0.000
Hb, (g/L)	128.6 ± 16.4 [b]	123.3 ± 25.7 [a]	126.3 ± 17.8	0.051
PLT, (10^9/L)	250.2 ± 74.0 [c]	254.7 ± 75.1 [c]	208.5 ± 65.6 [ab]	0.000
Cr, (μmol/L)	65.3 ± 11.5 [b]	98.6 ± 66.0 [ac]	66.4 ± 11.9 [b]	0.000
eGFR, mL/min/1.73m^2	103.0 ± 20.4 [b]	81.0 ± 33.1 [ac]	98.8 ± 18.6 [b]	0.000
IgG, (g/L)	13.7 ± 4.1 [bc]	10.1 ± 3.0 [ac]	15.0 ± 3.9 [ab]	0.000
IgA, (g/L)	2.8 ± 1.2 [b]	2.2 ± 1.0 [ac]	2.9 ± 1.3 [b]	0.000
IgM, (g/L)	1.3 ± 0.9 [bc]	0.8 ± 0.5 [ac]	1.0 ± 0.5 [ab]	0.000
C3, (mg/L)	685.1 ± 179.5 [b]	805.8 ± 207.8 [ac]	692.6 ± 147.4 [b]	0.000
C4, (mg/L)	130.4 ± 56.5 [b]	199.1 ± 81.2 [ac]	140.1 ± 56.8 [b]	0.000
ANA	92(70.8)[bc]	121(91.7)[a]	373(100)[a]	0.000
anti-dsDNA antibodies	104(80)[bc]	129(97.8)[a]	373(100)[a]	0.000
anti-Sm antibodies	38(29.2)[bc]	0(0)[a]	0(0)[a]	0.000
anti-RNP antibodies	9(6.9)[bc]	0(0)[a]	0(0)[a]	0.000
anti-SSA antibodies	32(24.6)[bc]	0(0)[a]	0(0)[a]	0.000
anti-SSB antibodies	14(10.8)[bc]	0(0)[a]	0(0)[a]	0.000
anti-rRNP antibodies	4(3)[bc]	0(0)[a]	1(0.2)[a]	0.000

Patient characteristic	Cluster 1, n = 130 (20.5%)	Cluster 2, n = 132 (20.8%)	Cluster 3, n = 373 (58.7%)	P-values
anti-nucleosome antibodies	16(12.3)[bc]	3(2.3)[a]	0(0)[a]	0.000
anti-histone antibodies	8(6.2)[bc]	0(0)[a]	0(0)[a]	0.000
PGA, median (range)	1(0-1.1)	1(0.9−1.1)	1(1-1.3)	0.000
SDI score, mean (S.D.)	0.8 ± 1.9[bc]	0.3 ± 1.3[a]	0.2 ± 0.8[a]	0.000

Datas are given as mean (SD), median, or as number and percentage. White blood cell (WBC), Neutrophils (N), Lymphocytes (L), Haemoglobin (Hb), blood platelets (PLT), Creatinine (Cr), Estimated Glomerular Filtration Rate (eGFR), Immunoglobulin G (IgG), Immunoglobulin A (IgA), Immunoglobulin M (IgM), Complement 3 (C3), Complement 4 (C4), antinuclear antibodies (ANA), antibodies to double-stranded DNA (anti-dsDNA antibody), antibodies to Smith (anti-Sm antibodies), antibodies to ribosomal P protein (anti-rRNP antibodies), Systemic Lupus International Collaborating Clinics/ACR Damage Index (SDI).

[a] Significantly different from cluster 1.

[b] Significantly different from cluster 2.

[c] Significantly different from cluster 3.

Cluster 3 was the largest group of mild disease activity. Patients in this cluster had lower mean SDI scores compared to those in the other two cluster. The average values of white blood cells, neutrophils and lymphocytes in this cluster of patients were the lowest, however, the mean values of IgG and IgA were the highest among three clusters. Although all the patients of cluster 3 had positive ANA and anti-dsDNA antibodies, the antinuclear antibody spectrum of this cluster was almost all negative. Cluster 1 was the smallest group of severe damage, patients in this cluster had higher SDI scores. Besides, the complement level of this cluster of patients was the lowest, and there were multiple positive auto-antibodies. Patients of cluster 2 had the highest levels of granulocytes, but the results of other laboratory tests were roughly between the cluster 1 and cluster 3.

In general, we divided LN patients into 3 significantly different categories through two-step cluster analyses. Among them, the overall damage of the first cluster of patients was the most obvious, cluster 3 were the mildest group, while the overall laboratory data of cluster 2 was between cluster 1 and cluster 3. The cluster labels of each class of patient are more intuitively displayed in Fig. 1.

Discussion

This is a cross-sectional study consists of 635 LN patients, Cluster analysis showed that these patients could be divided into three significantly different categories. The general purpose of this research are to

verify a reasonable way to classify clinical patients and further explore how the clusters are different, that is to say, to determine the specific clinical variables that vary patients in different clusters.

The severity of these three clusters of patients was different, and the first type of patients is relatively the most severe: relatively high lymphocyte counts, low complement level, multiple positive autoantibodies, and higher SDI scores. The difference in the number of white blood cells plays an important role in this cluster analysis. According to previous literature research, abnormalities in T and B lymphocytes are the major arms of adaptive immune responses, anomalous complement activation and autoantibody production can contribute to the pathogenesis of SLE and LN.[10–13]

The presence of multiple autoantibody positives in the ANA profile is the main feature of the cluster 1. At present, there are few related studies on ANA antibody profile and specific organ involvement in SLE/LN: Anti-Sm antibodies were associated with neurologic disorder and cardiovascular disease, besides, anti-Sm antibody was also associated with disease activity in SLE patients.[14–16] It had been reported that anti-SSA antibodies with skin involvement and sicca symptoms;[17] SLE patients with anti-SSA antibodies revealed higher frequencies of cutaneous vasculitis, severe mitral valve regurgitation and musculoskeletal involvement.[18, 19] Anti-RNP antibodies are more prevalent in patients with Raynaud's phenomenon and are associated with renal involvement;[20, 21] Another research reported that the combination of anti-RNP/Sm and LA was associated with thrombosis in SLE.[22] Nucleosomes and histone were considered to play a major role in the autoantibody response in LN, and circulating nucleosomes were positively correlated with an active lupus disease.[23, 24] In summary, a variety of positive ANA profile antibodies might be related to the high SDI score of LN patients, and may therefore further affect prognosis of LN.

The two-step cluster analyses adopted in this study reclassified LN patients into 3 clusters, laboratory indicators of these three clusters of patients are significantly different. This suggests that we can classify patients at different periods, so as to find groups of significant differences in treatment response, carry out targeted therapy and improve clinical effectiveness.

Conclusion

In conclusion, this study analyzed and identified different clusters of of LN patients within a large cohort. Our study suggests that positive ANA profile antibodies might be related to the high SDI score of LN patients. Clinicians can reclassify patients at different stages to better implement precise treatment through cluster analysis.

Declarations

Acknowledgement

The authors would like to thank the Department of Rheumatology at Xiangya Hospital for its participation, including patient recruitment.

Authors' contributions

TH performed data analysis and drafted the manuscript. SSX and LQD contributed to data collection and interpretation procedures. HL supervised the study design. All authors read and approved the final manuscript.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

All procedures in our study were approved by the Ethical Committee Group of Xiangya Hospital. Written informed consent was obtained from each participant. All methods were carried out in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Yu F, Haas M, Glassock R, et al. Redefining lupus nephritis: clinical implications of pathophysiologic subtypes[J]. Nature Reviews Nephrology, 2017,13(8):483-495.
2. Meacock R, Dale N, Harrison M J. The humanistic and economic burden of systemic lupus erythematosus : a systematic review[J]. Pharmacoeconomics, 2013,31(1):49-61.
3. Gladman D D, Goldsmith C H, Urowitz M B, et al. The Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index for Systemic Lupus Erythematosus International Comparison[J]. J Rheumatol, 2000,27(2):373-376.
4. Sada K E, Makino H. Usefulness of ISN/RPS classification of lupus nephritis[J]. J Korean Med Sci, 2009,24 Suppl:S7-S10.

5. Markowitz G S, D'Agati V D. The ISN/RPS 2003 classification of lupus nephritis: an assessment at 3 years[J]. Kidney Int, 2007,71(6):491-495.

6. Anders H J, Rovin B. A pathophysiology-based approach to the diagnosis and treatment of lupus nephritis[J]. Kidney Int, 2016,90(3):493-501.

7. Bertsias G K, Tektonidou M, Amoura Z, et al. Joint European League Against Rheumatism and European Renal Association-European Dialysis and Transplant Association (EULAR/ERA-EDTA) recommendations for the management of adult and paediatric lupus nephritis[J]. Ann Rheum Dis, 2012,71(11):1771-1782.

8. Gladman D D, Ibanez D, Urowitz M B. Systemic lupus erythematosus disease activity index 2000[J]. J Rheumatol, 2002,29(2):288-291.

9. Lam G K, Petri M. Assessment of systemic lupus erythematosus[J]. Clin Exp Rheumatol, 2005,23(5 Suppl 39):S120-S132.

10. Yap D, Chan T M. B Cell Abnormalities in Systemic Lupus Erythematosus and Lupus Nephritis-Role in Pathogenesis and Effect of Immunosuppressive Treatments[J]. Int J Mol Sci, 2019,20(24).

11. Devarapu S K, Lorenz G, Kulkarni O P, et al. Cellular and Molecular Mechanisms of Autoimmunity and Lupus Nephritis[J]. Int Rev Cell Mol Biol, 2017,332:43-154.

12. Yap D Y, Lai K N. Pathogenesis of renal disease in systemic lupus erythematosus–the role of autoantibodies and lymphocytes subset abnormalities[J]. Int J Mol Sci, 2015,16(4):7917-7931.

13. Bagavant H, Fu S M. Pathogenesis of kidney disease in systemic lupus erythematosus[J]. Curr Opin Rheumatol, 2009,21(5):489-494.

14. Flechsig A, Rose T, Barkhudarova F, et al. What is the clinical significance of anti-Sm antibodies in systemic lupus erythematosus? A comparison with anti-dsDNA antibodies and C3[J]. Clin Exp Rheumatol, 2017,35(4):598-606.

15. Choe J Y, Lee S S, Kwak S G, et al. Anti-Sm Antibody, Damage Index, and Corticosteroid Use Are Associated with Cardiac Involvement in Systemic Lupus Erythematosus: Data from a Prospective Registry Study[J]. J Korean Med Sci, 2020,35(21):e139.

16. Ahn S S, Jung S M, Yoo J, et al. Anti-Smith antibody is associated with disease activity in patients with new-onset systemic lupus erythematosus[J]. Rheumatol Int, 2019,39(11):1937-1944.

17. Frodlund M, Wettero J, Dahle C, et al. Longitudinal anti-nuclear antibody (ANA) seroconversion in systemic lupus erythematosus: a prospective study of Swedish cases with recent-onset disease[J]. Clin Exp Immunol, 2020,199(3):245-254.

18. Novak G V, Marques M, Balbi V, et al. Anti-RO/SSA and anti-La/SSB antibodies: Association with mild lupus manifestations in 645 childhood-onset systemic lupus erythematosus[J]. Autoimmun Rev, 2017,16(2):132-135.

19. Higuera-Ortiz V, Mora-Arias T, Castillo-Martinez D, et al. Anti-Ro/SSA antibodies are associated with severe mitral valve regurgitation in patients with systemic lupus erythematosus[J]. Mod Rheumatol, 2017,27(3):476-480.

20. ter Borg E J, Groen H, Horst G, et al. Clinical associations of antiribonucleoprotein antibodies in patients with systemic lupus erythematosus[J]. Semin Arthritis Rheum, 1990,20(3):164-173.

21. Bastian H M, Roseman J M, McGwin G J, et al. Systemic lupus erythematosus in three ethnic groups. XII. Risk factors for lupus nephritis after diagnosis[J]. Lupus, 2002,11(3):152-160.

22. Zamora-Medina M, Hinojosa-Azaola A, Nunez-Alvarez C A, et al. Anti-RNP/Sm antibodies in patients with systemic lupus erythematosus and its role in thrombosis: a case-control study[J]. Clin Rheumatol, 2019,38(3):885-893.

23. Amoura Z, Chabre H, Koutouzov S, et al. Nucleosome-restricted antibodies are detected before anti-dsDNA and/or antihistone antibodies in serum of MRL-Mp lpr/lpr and +/+ mice, and are present in kidney eluates of lupus mice with proteinuria[J]. Arthritis Rheum, 1994,37(11):1684-1688.

24. Williams R J, Malone C C, Meyers C, et al. Detection of nucleosome particles in serum and plasma from patients with systemic lupus erythematosus using monoclonal antibody 4H7[J]. J Rheumatol, 2001,28(1):81-94.
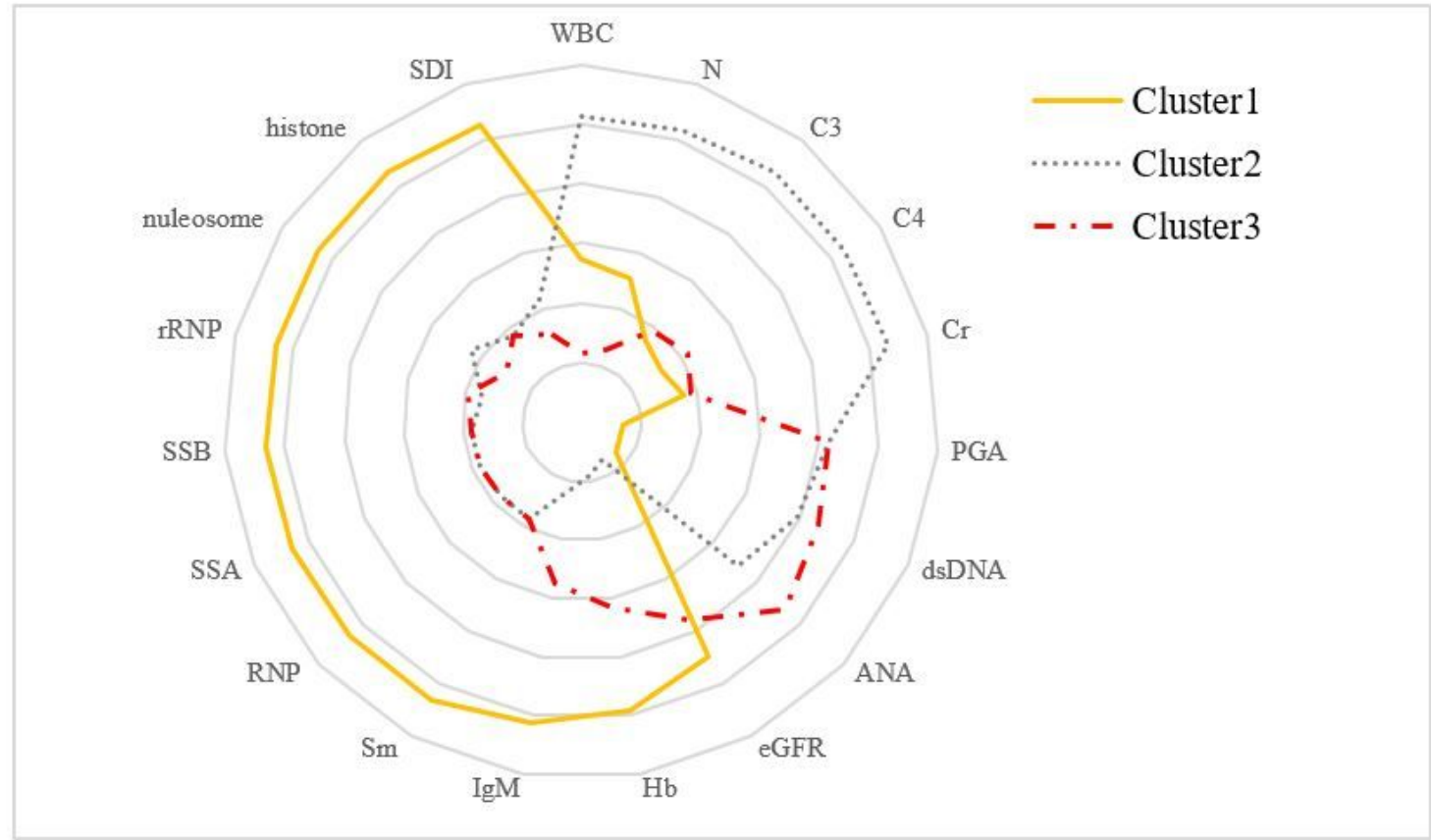
Figures

Figure 1

Labels of important clinical variables for each cluster of patients.