

Original article

Sickle cell anemia: hierarchical cluster analysis and clinical profile in a cohort in Brazil



Valéria de Freitas Dutra *, Thais Priscila Biassi , Maria Stella Figueiredo 

Escola Paulista de Medicina, Universidade Federal de São Paulo (EPM/UNIFESP), São Paulo, SP, Brazil

ARTICLE INFO

Article history:

Received 2 June 2021

Accepted 17 August 2021

Available online 9 December 2021

Keywords:

Sickle cell anemia

Hierarchical clusters

Hemoglobin S

ABSTRACT

Introduction: Sickle cell anemia is a monogenic disorder caused by a mutation in the β -hemoglobin gene, resulting in sickle hemoglobin that can polymerize. Presentation and clinical course have significant inter-individual variability and classifying these patients for severity is a challenge.

Methods: We applied hierarchical clusters with 10 routine laboratory tests to understand if this grouping could be associated with clinical manifestations. We included 145 adult homozygous patients (SS) at an outpatient clinic in a retrospective study.

Results: We found five clusters by counting those that had been differentiated by unconjugated bilirubin, reticulocytes, LDH, leukocytes, lymphocytes and monocytes. When comparing groups to clinical findings, the clusters were different only for liver abnormality. Cluster 3 had the lower median of reticulocytes, LDH, leukocytes, lymphocytes and monocytes and a higher percentage of patients under treatment. Clusters 4 and 5 had higher frequencies of liver impairment and higher medians of reticulocytes, LDH, leukocytes, lymphocytes and monocytes. Hemolysis and inflammation seemed to influence the grouping.

Conclusion: In our study, cluster analysis showed five groups that exhibited different degrees of inflammation and hemolysis. When comparing clinical data, the result was different only for the criteria of liver abnormality.

© 2021 Associação Brasileira de Hematologia, Hemoterapia e Terapia Celular. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Corresponding author at: Escola Paulista de Medicina, Universidade Federal de São Paulo (EPM/UNIFESP), R Dr. Diogo de Farias, 824, 04037-002 Vila Clementino, São Paulo, SP, Brazil.

E-mail address: valeriadutra@yahoo.com (V.d.F. Dutra).

<https://doi.org/10.1016/j.htct.2021.08.015>

2531-1379/© 2021 Associação Brasileira de Hematologia, Hemoterapia e Terapia Celular. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Sickle cell anemia (SCA) is a monogenic disorder caused by a mutation in the β -hemoglobin gene (*HBB*), *glu6val*, resulting in abnormal hemoglobin (HbS) that can polymerize.¹ The most common genotype is the homozygous HbS. Despite an identical HbS genotype, the clinical presentation is heterogeneous and classifying these patients is difficult, especially because the disease does not have a specific severity marker. Predicting the clinical phenotype could offer better treatment and prognosis.²

The SCA is a public health problem due to the high costs of hospitalization and comorbidities.³ It is particularly true in Brazil, where a Unified Health System finances the costs and where preventive measures could significantly reduce the disease's economic burden. Data from the World Health Organization estimates that 275,000 newborns around the world are diagnosed with sickle cell disease each year,⁴ being around 1000 newborns/year in Brazil,⁵ where neonatal screening has been obligatory since 2001.⁶

This is a systemic disorder, associated with acute episodes and chronic conditions, with progressive organ damage⁷ and with only one drug, hydroxyurea, approved for use in Brazil. Low-cost biomarkers could help developing countries with patient risk stratification and follow-up.

Hierarchical clustering is a multivariate statistical technique that could be helpful in grouping SCA patients. This is an algorithm to group similar characteristics, revealing subgroups with heterogeneous data. It is one of the most used clustering techniques in bioinformatics and it has been already used for many clinical purposes,^{8–10} including diagnosis and artificial intelligence development.¹¹ The result or the classifications can be pictured as hierarchical trees, the dendrograms. To form the clusters, this method uses the similarities or dissimilarities, also called distance (Euclidean distance, squared Euclidean distance, city-block distance, power distance, etc.). There are different techniques used, for example, single linkage, complete linkage, average, centroid, and Ward's method; however, comparative studies showed that Ward's seems to be better than the others.¹²

Du et al., 2018¹³ proposed a system-type approach to discover profiles of multiple, common biomarkers that correlate

with morbidity and mortality for sickle cell disease. Based on this, we designed a study with cluster analysis in a sample of 145 HbSS patients using 10 routine, low-cost laboratory tests.

Materials and methods

Biomarker inclusion and data

Clinical data and laboratory tests were collected in steady-state (at least 3 months without crisis) and at an outpatient clinic, retrospectively, from 2008 to 2018. To eliminate the bias of different genotypes, we only included the homozygous SS, totalizing 145 adults. Since cluster analysis requires complete data on all patients, we excluded 13 patients. Laboratory tests were included based on availability, cost and routine use. All of the tests were quantitative variables: counting hemoglobin, leukocytes, lymphocytes and monocytes and unconjugated bilirubin, creatinine, ferritin, lactate dehydrogenase (LDH), uric acid and reticulocytes.

Hierarchical clustering and statistics

We applied hierarchical clustering with the squared Euclidean distance and Ward's method (SPSS 22.0). As a result, we found five final clusters, represented in the dendrogram (Figure 1). Age and biomarkers' median values were compared by performing a non-parametric test, the Kruskal-Wallis, followed by the *post hoc* Dunn multiple comparisons test. The Chi-Square test was performed for clinical data association. A *p*-value less than 0.05 was statistically significant.

Clinical data

To define clinical data, we used:

1. Liver abnormality: conjugated bilirubin greater than 0.6 (median obtained in a previous analysis, [¹⁴, based on¹⁵]) or palpable liver on physical examination or liver abnormality in ultrasound or tomography;
2. Nephropathy: albuminuria (higher than 30 mg in an isolated sample or a 24h proteinuria higher than 150mg) or an

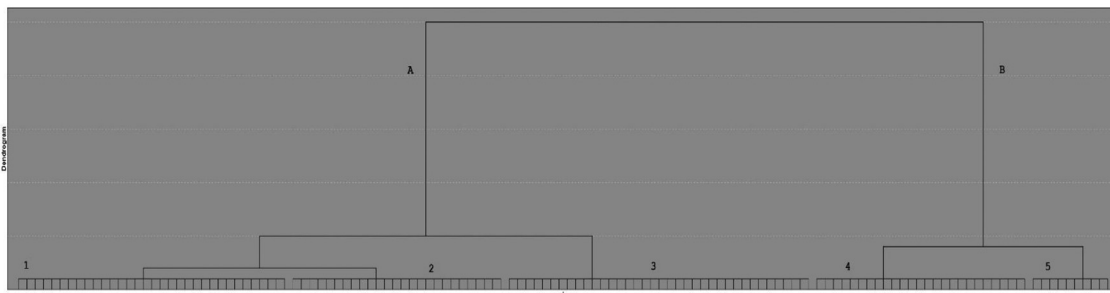


Figure 1 – Dendrogram of homozygous SS patients. Ward's Method.

The figure shows the five final clusters, represented by numbers 1 to 5. Clusters in the same branch are more similar and the shorter the branch, the greater the similarity. The dendrogram shows two major branches: A, with clusters 1, 2 and 3 and B, with 4 and 5.

Table 1 – Biomarker median values in our total sample (n = 145).

Biomarker	Median	Range (Min. -Max.)	Reference values
Hemoglobin (g/dL)	9.0	5 - 13	12 - 15.5
Reticulocytes (x10 ⁹ /L)	202.5	67.2 - 592.0	16.0 - 70.0
LDH (u/L)	489.0	141 - 2,673	< 250
Leucocytes (x10 ⁹ /L)	8.8	3.5 - 20.7	3.5 - 10.5
Lymphocytes (x10 ⁹ /L)	2.9	1.0 - 8.5	0.9 - 2.9
Monocytes (x10 ⁹ /L)	0.7	1.6 - 2.5	0.3 - 0.9
Unconjugated bilirubin (mg/dL)	1.58	0.21- 20.0	0.1 - 0.6
Creatinine (mg/dL)	0.59	0.3 - 9.52	0.5 - 0.9
Ferritin (ng/mL)	464.5	11 - 8,000	M: 30 - 400 F:13 - 150
Uric acid (mg/dL)	4.95	2 - 12	2.4 - 5.7

LDH, Lactate dehydrogenase; M, male; F, female.

elevated creatinine level (higher than the reference value-[Table 1](#));

- Osteonecrosis at any site, confirmed by a radiological examination (radiography or nuclear magnetic resonance);
- Sickle cell acute pain: at least one episode of acute pain crisis in life, refractory to the use of simple analgesics, such as acetaminophen or dipyron, with medical care assistance;
- Acute chest syndrome (ACS) characterized in medical records as pain, fever, desaturation and change in the radiological exam, at least one episode throughout life;
- Thrombotic manifestations: included peripheral venous thrombosis (PVT), pulmonary thromboembolism (PE) or cerebral venous thrombosis (CVT), at least one episode in life;
- Leg ulcer: previous or active, with or without hydroxyurea use;
- Stroke: ischemic or hemorrhagic, confirmed by computed tomography or magnetic resonance.

Treatment

We included patients with absent treatment, under hydroxyurea use (according to a Brazilian protocol: a minimal dose of 15 mg/kg up to the maximum tolerated dose) and receiving chronic manual blood transfusion.¹⁶

Ethical statement

The Ethics Committee (CAAE 70891517.9.0000.5505) approved this study and all participants provided written informed consent.

Results

The median age of participants was 33 years and subjects ranged in age from 19 to 76 years. Only 3.8% (5) were over 60 years. Women represented most (57.9%) of the patients. The majority (57.2%) were in treatment with hydroxyurea (HU) and 17.2% were in chronic transfusion. Of the 25.6% (37) without treatment, 21.6% (8) have had a treatment prescription (blood transfusion or hydroxyurea), but there was no therapeutic adherence.

The occurrence of clinical manifestations was 78.6% for painful episode, 53.8% for ACS, 40.7% for nephropathy, 27.6% for osteonecrosis, 22.8% for chronic hepatic disease, 20.0% for leg ulcers, 19.3 % for stroke and 9.0% for embolic manifestations. Despite not having included this data, at least one episode of priapism occurred in 44.3% of the men. [Table 1](#) shows the biomarker median values in our sample.

As mentioned, cluster analysis included 132 patients. This tool determined five clusters based on ten circulating biomarkers ([Figure 1](#)). [Table 2](#) shows this distribution and biomarker medians in each one.

Age, hemoglobin, ferritin, creatinine and uric acid showed no difference among the groups. There was a statistical difference for the following variables: leukocytes, lymphocytes and monocytes counts, reticulocytes, LDH and unconjugated bilirubin. Cluster 3 presented the lowest medians, when considering all these variables. When considering only reticulocytes, cluster 3 (median = 123,0) was different from clusters 1, 2, 4 and 5. Cluster 1 (median = 180,0) was also different from the others. Cluster 2 (median = 237,0) was different from 1, 3 and 5. Cluster 4 (median= 327,0) was different from 1 and 3 and cluster 5 was different from 1, 2 and 3. When taking into account the LDH, cluster 3 differed from clusters 4 and 5. Unconjugated bilirubin showed differences from cluster 3 to clusters 1, 2 and 4. Leukocyte analyses showed that cluster 1 was different from clusters 4 and 5, while cluster 3 differed from clusters 2, 4 and 5. For the lymphocyte count, cluster 3 was different from clusters 2, 4 and 5. Monocyte count analyses showed that cluster 1 differed from cluster 5 and cluster 3 differed from clusters 4 and 5.

To evaluate if the groups were different for clinical manifestations, we performed a Chi-Square test, as shown in [Table 3](#). As a result, the only statistical significance was found in liver abnormality. To show these differences, [Figure 2](#) presents the distribution of liver impairment among the five groups. Curiously, clusters 4 and 5, which presented the highest frequencies of liver abnormality (38.5% and 40%, respectively), are also the groups with the highest counts of reticulocytes, LDH, leukocytes, lymphocytes and monocytes. Moreover, as shown in [Figure 1](#), they are in the same major branch in the hierarchical cluster (represented as B). A point to be highlighted is that these two clusters also have a lower frequency of patients in treatment.

Table 2 – Age and biomarkers medians in each cluster (n=132).

Biomarkers	Cluster 1 (n = 33)	Cluster 2 (n = 26)	Cluster 3 (n = 37)	Cluster 4 (n = 26)	Cluster 5 (n = 10)	p-value
Age (years)	32.0 (24.5 - 39.5)	32.5 (23.7 - 39.0)	35.0 (29.0 - 45.0)	31.5 (25.7 - 36.2)	27.5 (24.5 - 38.5)	0.145
Hemoglobin (g/dL)	9.0 (8.0 - 11.0)	9.0 (8.0 - 10.0)	9.0 (8.0 - 10.0)	9.0 (8.0 - 10.0)	9.0 (7.7 - 9.2)	0.633
Reticulocytes (x10 ⁹ /L)	180.0 (165.0 - 194.5) ^{bcde}	237.0 (216.3 - 263.0) ^{ace}	123.0 (104.5 - 134.7) ^{abde}	327.0(305.0 - 357.2) ^{ac}	455.5 (423.7 - 489.5) ^{abc}	<0.0001
LDH (u/L)	498.0 (375.0 - 729.0)	474.0 (324.0 - 663.8)	388.0 (317.5 - 503.5) ^{de}	651.5 (414.5 - 790.8) ^c	736.0 (543.5 - 907.8) ^c	0.003
Leukocytes (x10 ⁹ /L)	8.6 (5.6 - 9.9) ^{de}	9.2 (7.3 - 11.4) ^c	6.7 (5.2 - 8.6) ^{bde}	12.0 (9.9 - 13.4) ^{ac}	13.4 (9.6 - 16.5) ^{ac}	<0.0001
Lymphocytes (x10 ⁹ /L)	2.8 (2.0 - 3.5)	3.4 (2.5 - 4.1) ^c	2.0 (1.7 - 2.9) ^{bde}	3.5 (2.3 - 4.4) ^c	3.6 (2.9 - 4.2) ^c	0.0008
Monocytes (x10 ⁹ /L)	0.65 (0.38 - 0.99) ^e	0.77 (0.53 - 1.0)	0.61 (0.32 - 0.76) ^{de}	0.97 (0.70 - 1.2) ^c	1.13 (0.92 - 1.6) ^{ac}	<0.0001
Unconjugated bilirubin (mg/dL)	1.9 (1.1 - 2.8) ^c	2.0 (0.8 - 2.9) ^c	0.7 (0.4 - 1.3) ^{abd}	1.8 (1.2 - 2.8) ^c	1.6 (1.0 - 2.8)	<0.0001
Creatinine (mg/dL)	0.55 (0.46 - 0.71)	0.56 (0.50 - 0.65)	0.65 (0.53 - 0.95)	0.60 (0.45 - 0.74)	0.50 (0.42 - 0.83)	0.110
Ferritin (ng/mL)	338 (203 - 868)	614 (251 - 1,133)	560 (242 - 1,246)	517 (185 - 1,660)	343 (169 - 1,777)	0.840
Uric acid (mg/dL)	5.1 (3.7 - 5.7)	4.6 (3.8 - 6.4)	5.0 (3.5 - 6.5)	4.8 (4.0 - 6.7)	6.0 (3.1 - 8.0)	0.937

LDH – lactate dehydrogenase. Kruskal-Wallis test. Data are presented as the median (interquartile range, P25 - P75 percentile). Each letter represents a specific cluster, where: a = cluster 1, b = cluster 2, c = cluster 3, d = cluster 4 and e = cluster 5. (Dunn's test; differences with $p < 0.05$).

Discussion

The SCA is responsible for severe comorbidities¹ and premature deaths.¹⁷ Despite having a known molecular basis, it possibly involves other pathophysiologic mechanisms, such as inflammation¹⁸ and thrombosis.¹⁹ Phenotypic variability is common and the current literature lacks clinical stratification data, impairing the reproducibility of basic studies or even equivalence in clinical trials.²⁰ Thus, the development of tools, as grouping with hierarchical clusters, can contribute to a better knowledge of the disease.

As the quality of assistance has improved, patients are living longer, with an enhanced life expectancy. In the sixties, the disease was a childhood problem. In the following decade, life expectancy was 14 years with few patients living up to 30 years.²¹ In the eighties, prophylactic therapy with oral penicillin decreased the morbidity and mortality associated with pneumococcal septicemia.²² In 1994, a prospective study from 1978 to 1988 reported a life expectancy of 42 years for males and 48 years for females, a decrease of about 25 to 30 years, when compared to the control consisting of African Americans.¹⁷ Two decades later, another study (1999 - 2009) reproduced similar results.²³

Table 3 – Distribution of gender, treatment and clinical manifestations by clusters (n = 132).

	Cluster 1 (n = 33) (%)	Cluster 2 (n = 26) (%)	Cluster 3 (n = 37) (%)	Cluster 4 (n = 26) (%)	Cluster 5 (n = 10) (%)
Gender					
Women	51.0	50	67.5	62	70
Men	49.0	50	32.5	38	30
Treatment					
HU	66.7	61.5	75.7	34.6	40
Chronic Transfusion	9.1	19.2	16.2	26.9	20
Total in treatment	75.8	80.7	91.9 ^{a,b}	61.5 ^a	60 ^b
No Treatment	24.2	19.2	8.1	38.5	40
Clinical Manifestation					
Liver abnormality	15.2 ^c	23.1	10.8 ^{d,e}	38.5 ^{c,d}	40.0 ^e
Nephropathy	39.4	46.2	48.6	34.6	50
Osteonecrosis	24.2	30.8	32.4	34.6	20.0
Acute Pain	84.8	76.9	73.0	80.8	80.0
ACS	57.6	50.0	56.8	65.4	30.0
Stroke	18.2	26.9	13.5	23.1	30.0
Thrombotic manifestations	9.1	3.8	16.2	11.5	0
Leg Ulcer	18.2	19.2	10.8	23.1	20.0
Number of clinical manifestations (mean)	2.7	2.8	2.6	3.1	2.7

Values in a row with the same superscript letters are significantly different, $p < 0.05$ (Chi-Square test). HU, Hydroxyurea; ACS, Acute Chest Syndrome.

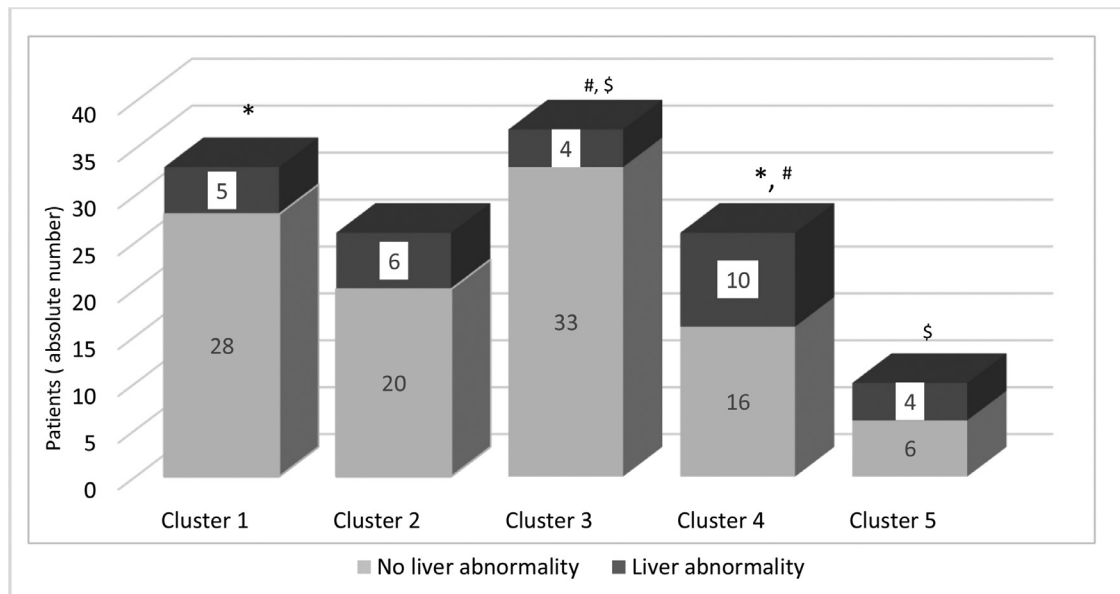


Figure 2 – Distribution of liver abnormality among the groups. Symbols represent statistical differences between groups ($p < 0.05$, Chi-Square test).

Our cohort consisted of patients with a median age of 33 years, with only three patients over 60, considered elderly in developing countries, suggesting a lower life expectancy. Moreover, most patients were receiving treatment (HU or chronic transfusion), reflecting more severe patients in our sample.

As expected, the clinical presentation was variable. The most frequent manifestation was pain, at least one episode of which occurred, with medical assistance, in 78.6% of the cases. According to the literature, pain is the main cause of emergency care and hospitalization²⁴ and could occur in 60% of patients.²⁵ In our cohort, 53.8% of patients had at least one acute chest syndrome, which corroborates the literature.^{26,27} Renal involvement was present in 40.7% of patients. Nephropathy is one of the most common chronic comorbidities and causes of death in sickle cell disease (SCD),²⁸ albuminuria can be present in up to 68% of SCD adults²⁹ and renal manifestations are frequent even in children with SCD (17-27%).³⁰ Osteonecrosis was present in 27.6% of patients, data similar to other studies.³¹

Liver disease is the result of multiple insults and pathophysiological mechanisms, such as virus infection, sickling and iron overload.³² The hepatobiliary system can be involved in 10 - 40% of SCD patients, although the concept of chronic liver disease is not clearly defined and can include iron overload, viral hepatitis, cholelithiasis and cholangiopathy.³³ Direct bilirubin ≥ 0.4 mg/dL was independently associated with mortality in an SCD cohort.¹⁵ In a previous analysis, our group obtained a conjugated bilirubin median of 0.6 mg/dL, which was higher than that found in Feld et al.^{14,15} Applying the criteria of conjugated bilirubin higher than 0.6, palpable liver on physical examination or liver abnormality in ultrasound or tomography, the liver abnormality was present in 22.8% of patients.

The leg ulcer is associated with hemolysis and can be present in thalassemia³⁴ or spherocytosis.³⁵ In our cohort, 20% of patients presented a leg ulcer (active or previous history). The frequency of leg ulcers varies in the literature and is affected by geographic distribution, occurring in 75% of SCA patients in Jamaica, but only in 8 - 10% in North America.³⁶ We found a higher number of strokes, when compared to a prevalence of 7.8% in a pediatric Jamaican study,³⁷ 8.4% in children in Nigeria³⁸ and 4% in the Cooperative Study of Sickle Cell Disease.³⁹

Finally, at least one thrombotic event occurred in 9% of the cases in our sample. There are few studies in this area and the risk of thrombotic events in sickle cell patients appears greater than in the general population. Thrombotic manifestations are probably somewhat neglected, with inadequate prophylaxis.⁴⁰ One clinical finding that was not included in the cluster analysis was priapism, occurring in at least one episode lifelong, present in 44.3% of men. This is in agreement with a previous study,⁴¹ which found a prevalence of 42%.

Considering the difficulty in classifying these patients clinically and the hypothesis that the analysis of multiple biomarkers could simultaneously offer better signatures than individual levels,⁴² we included 132 SCA patients in the hierarchical cluster analysis. Age was similar in all clusters. Variables (leukocytes count, lymphocytes, monocytes, reticulocytes, LDH and unconjugated bilirubin) were statistically different, confirming that the distribution was not by chance. Clearly, inflammation and hemolysis markers were important for the hierarchical cluster construction. The hierarchical analysis presented different hemolysis degrees, as found by Du et al.¹³ Leukocyte counts were also different among the groups. In a cohort in the Congo, an elevated baseline leukocyte count was associated with the SCA severity, suggesting that leukocytes could play a major role in the phenotypic severity of the

disease, as these cells are involved in cytokine release, chronic inflammation and infection.⁴³

Finally, to understand whether the clinical findings were associated with these clusters, they were also compared to clinical manifestations. Hepatic laboratory tests were not used in our cluster differentiation, even though the groups were different only in this clinical aspect (cluster 3 from 4 and 5 and 1 from 4). Cluster 3 (91.9%) was also different from 4 (61.5%) and 5 (60%), when comparing the percentage of patients with and without treatment, despite not being influenced by the type of treatment. The absence of treatment, high levels of inflammation and hemolysis seemed to predispose the patient to liver impairment.

The literature is not clear regarding the classification of sickle cell liver disease and this may be a bias in our study, since patients were not submitted to biopsy to demonstrate liver impairment and we do not use albumin or the prothrombin time as a routine test. Another limitation is that we could not include classical modulators of the disease in the clusters, such as fetal hemoglobin or alfa-thalassemia, because they are not routine tests. Furthermore, the small sample size (132 SS patients) may decrease the power of the statistical tests in showing important biological differences. Treatment was less frequent in clusters 4 and 5; however, we did not include the duration of treatment, which could influence the clinical presentation. Moreover, this is not a prospective study and a follow-up could be an opportunity to verify the development of comorbidities and mortality risk within the clusters. Nevertheless, we are the first study to test cluster analysis in an adult HbSS cohort, as the previous ones¹³ included only children (mean age of 15.3 years).

A widely, accessible and validated method to measure disease severity does not currently exist. It is important to detect outpatients with a worse prognosis to prevent complications and increase life expectancy. Unfortunately, SCA is still a neglected disease and affordable treatment and morbidity prevention have progressed slowly during the past years. New tools can contribute to a better understanding of this disease and its complications. Hierarchical clusters can be a tool for classifying SCA patients in an easy, accessible and important manner. It is reproducible, useful to outpatients and inexpensive.

Conclusion

Cluster analysis using 10 common routine laboratory tests detected five different subgroups. Biomarkers of hemolysis and inflammation appeared to be the factors that most influenced this grouping. When comparing the groups with clinical manifestations, only the criteria of liver abnormality showed a statistical difference.

Funding

This work was supported by the CAPES -Finance Code 001 and CNPq 162577/2015-0 and 166526/2013-4 (Authors' Fellowship) processes.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge Prof. Gianni Mara Silva dos Santos for the cluster analysis support.

REFERENCES

1. Rees DC, Williams TN, Gladwin MT. Sickle-cell disease. *Lancet*. 2010;376(9757):2018–31. [https://doi.org/10.1016/S0140-6736\(10\)61029-X](https://doi.org/10.1016/S0140-6736(10)61029-X).
2. Steinberg MH. Predicting clinical severity in sickle cell anaemia. *Br J Haematol*. 2005;129(4):465–81. <https://doi.org/10.1111/j.1365-2141.2005.05411.x>.
3. Kauf TL, Coates TD, Huazhi L, Mody-Patel N, Hartzema AG. The cost of health care for children and adults with sickle cell disease. *Am J Hematol*. 2009;84(6):323–7. <https://doi.org/10.1002/ajh.21408>.
4. Modell B, Darlison M. Global epidemiology of haemoglobin disorders and derived service indicators. *Bul World Health Organ* 86,480-7.
5. Cancado RD, Jesus JA. A doença falciforme no Brasil. *Rev Bras Hematol Hemoter*. 2007;29(3):204–6.
6. Ramalho AS, Magna LA, Paiva e Silva RB. A Portaria n° 822/01 do Ministério da Saúde e as peculiaridades das hemoglobinopatias em saúde pública no Brasil. *Cad. Saúde Pública*. 2003;19(4):1195–9.
7. Ataga KI, Stocker J. The trials and hopes for drug development in sickle cell disease. *Br J Haematol*. 2015;170(6):768–80. <https://doi.org/10.1111/bjh.13548>.
8. Murtagh F, Contreras P. Algorithms for hierarchical clustering: an overview, II. *Wiley Interdiscip Rev*. 2017;7(6). e1219. <https://doi.org/10.1002/widm.1219>
9. Pagnuco IA, Pastore JI, Abras G, Brun M, Ballarin VL. Analysis of genetic association using hierarchical clustering and cluster validation indices. *Genomics*. 2017;109(5-6):438–45. <https://doi.org/10.1016/j.ygeno.2017.06.009>.
10. Fattori A, Oliveira IM, Alves RM, Guariento ME. Cluster analysis to identify elderly people's profiles: a healthcare strategy based on frailty characteristics. *Sao Paulo Med J*. 2014;132(4):224–30. <https://doi.org/10.1590/1516-3180.2014.1324622>.
11. Williamson DJ, Burn GL, Simoncelli S, Griffié J, Peters R, Davis DM, et al. Machine learning for cluster analysis of localization microscopy data. *Nat Commun*. 2020;11(1):1493. <https://doi.org/10.1038/s41467-020-15293-x>.
12. Saraçlı S, Doğan N, Doğan İ. Comparison of hierarchical cluster analysis methods by cophenetic correlation. *J Inequal Appl*. 2013;203. <https://doi.org/10.1186/1029-242X-2013-203>.
13. Du M, Van Ness S, Gordeuk V, Nouraie SM, Nekhai S, Gladwin M, et al. Biomarker signatures of sickle cell disease severity. *Blood Cells Mol Dis*. 2018;72:1–9. <https://doi.org/10.1016/j.bcmd.2018.05.001>.
14. Seiwald MC, Vieira AC, Leonel RB, Furtado VM, Figueiredo MS, Silva AM, et al. Avaliação hepática na anemia falciforme. *Rev Bras Hematol Hemoter*. 2017;39(Supl 1):S19.
15. Feld JJ, Kato GJ, Koh C, Shields T, Hildesheim M, Kleiner DE, et al. Liver injury is associated with mortality in sickle cell disease. *Aliment Pharmacol Ther*. 2015 Oct;42(7):912–21. <https://doi.org/10.1111/apt.13347>.

16. Chou ST. Transfusion therapy for sickle cell disease: a balancing act. *Hematol Am Soc Hematol Educ Program*. 2013;2013:439–46. <https://doi.org/10.1182/asheducation-2013.1.439>. PMID: 24319217.
17. Platt OS, Brambilla DJ, Rosse WF, Milner PF, Castro O, Steinberg MH, et al. Mortality in sickle cell disease. Life expectancy and risk factors for early death. *N Engl J Med*. 1994;330(23):1639–44. <https://doi.org/10.1056/NEJM199406093302303>.
18. Platt OS. Sickle cell anemia as an inflammatory disease. *J Clin Invest*. 2000;106(3):337–8. <https://doi.org/10.1172/JCI10726>.
19. De Franceschi L, Cappellini MD, Olivieri O. Thrombosis and sickle cell disease. *Semin Thromb Hemost*. 2011;37(3):226–36. <https://doi.org/10.1055/s-0031-1273087>.
20. Coelho A, Dias A, Morais A, Nunes B, Faustino P, Lavinha J. Sickle cell disease severity scoring: a yet unsolved problem. *Eur J Haematol*. 2012;89(6):501–2. <https://doi.org/10.1111/ejh.12011>. Epub 2012 Oct 26. PMID: 22938536.
21. Diggs LM. Anatomic lesions in sickle cell disease. In: Abramson H, Bertles JF, Wethers DL, eds. *Sickle cell disease: diagnosis, management, education, and research*, St. Louis: C. V. Mosby; 1973:189–229.
22. Gaston MH, Verter JI, Woods G, Pegelow C, Kelleher J, Presbury G, et al. Prophylaxis with oral penicillin in children with sickle cell anemia. A randomized trial. *N Engl J Med*. 1986;314(25):1593–9. <https://doi.org/10.1056/NEJM198606193142501>.
23. Hamideh D, Alvarez O. Sickle cell disease related mortality in the United States (1999–2009). *Pediatr Blood Cancer*. 2013;60(9):1482–6. <https://doi.org/10.1002/pbc.24557>.
24. Shapiro BS, Benjamin LJ, Payne R, Heidrich G. Sickle cell-related pain: perceptions of medical practitioners. *J Pain Symptom Manage*. 1997;14(3):168–74. [https://doi.org/10.1016/S0885-3924\(97\)00019-5](https://doi.org/10.1016/S0885-3924(97)00019-5).
25. Sant'Ana PG, Araujo AM, Pimenta CT, Bezerra ML, Junior SP, Neto VM, et al. Clinical and laboratory profile of patients with sickle cell anemia. *Rev Bras Hematol Hemoter*. 2017;39(1):40–5. <https://doi.org/10.1016/j.bjhh.2016.09.007>.
26. Castro O, Brambilla DJ, Thorington B, Reindorf CA, Scott RB, Gillette, et al. The acute chest syndrome in sickle cell disease: incidence and risk factors. *The Cooperative Study of Sickle Cell Disease*. *Blood*. 1994;84(2):643–9.
27. Vichinsky EP, Styles LA, Colangelo LH, Wright EC, Castro O, Nickerson B. Acute chest syndrome in sickle cell disease: clinical presentation and course. *Cooperative Study of Sickle Cell Disease*. *Blood*. 1997;89(5):1787–92.
28. Hariri E, Mansour A, El Alam A, Daaboul Y, Korjian S, Aoun Bahous S. Sickle cell nephropathy: an update on pathophysiology, diagnosis, and treatment. *Int Urol Nephrol*. 2018;50(6):1075–83. <https://doi.org/10.1007/s11255-018-1803-3>.
29. Guasch A, Navarrete J, Nass K, Zayas CF. Glomerular involvement in adults with sickle cell hemoglobinopathies: Prevalence and clinical correlates of progressive renal failure. *J Am Soc Nephrol*. 2006;17(8):2228–35. <https://doi.org/10.1681/ASN.2002010084>.
30. Olaniran KO, Eneanya ND, Nigwekar SU, Vela-Parada XF, Achebe MM, et al. Sickle cell nephropathy in the pediatric population. *Blood Purif*. 2019;47(1-3):205–13. <https://doi.org/10.1159/000494581>.
31. Adesina O, Brunson A, Keegan THM, Wun T. Osteonecrosis of the femoral head in sickle cell disease: prevalence, comorbidities, and surgical outcomes in California. *Blood Adv*. 2017;1(16):1287–95. <https://doi.org/10.1182/bloodadvances.2017005256>. Published 2017 Jul 11.
32. Theocharidou E, Suddle AR. The liver in sickle cell disease. *Clin Liver Dis*. 2019;23(2):177–89. <https://doi.org/10.1016/j.cld.2018.12.002>.
33. Shah R, Tabora C, Chawla S. Acute and chronic hepatobiliary manifestations of sickle cell disease: a review. *World J Gastrointest Pathophysiol*. 2017;8(3):108–16. <https://doi.org/10.4291/wjgp.v8.i3.108>.
34. Stevens DM, Shupack JL, Javid J, Silber R. Ulcers of the leg in thalassemia. *Arch Dermatol*. 1977;113(11):1558–60.
35. Giraldi S, Abbage KT, Marinoni LP, Oliveira V, Pianowski MA, Lehmkuhl AE, et al. Leg ulcer in hereditary spherocytosis. *Pediatr Dermatol*. 2003;20(5):427–8. <https://doi.org/10.1046/j.1525-1470.2003.20512.x>.
36. Minniti CP, Eckman J, Sebastiani P, Steinberg MH, Ballas SK. Leg ulcers in sickle cell disease. *Am J Hematol*. 2010;85(10):831–3. <https://doi.org/10.1002/ajh.21838>.
37. Balkaran B, Char G, Morris JS, Thomas PW, Serjeant BE, Serjeant GR. Stroke in a cohort of patients with homozygous sickle cell disease. *J Pediatr*. 1992;120(3):360–6. [https://doi.org/10.1016/s0022-3476\(05\)80897-2](https://doi.org/10.1016/s0022-3476(05)80897-2).
38. Lagunju IA, Brown BJ. Adverse neurological outcomes in Nigerian children with sickle cell disease. *Int J Hematol*. 2012;96(6):710–8. <https://doi.org/10.1007/s12185-012-1204-9>.
39. Ohene-Frempong K, Weiner SJ, Sleeper LA, Miller ST, Embury S, Moehr JW, et al. Cerebrovascular accidents in sickle cell disease: rates and risk factors. *Blood*. 1998;91(1):288–94.
40. Noubiap JJ, Temgoua MN, Tankeu R, Tochie JN, Wonkam A, Bigna JJ. Sickle cell disease, sickle trait and the risk for venous thromboembolism: a systematic review and meta-analysis. *Thromb J*. 2018;16:27.. <https://doi.org/10.1186/s12959-018-0179-z>. Published 2018 Oct 4.
41. Emond AM, Holman R, Hayes RJ, Serjeant GR. Priapism and impotence in homozygous sickle cell disease. *Arch Intern Med*. 1980;140(11):1434–7.
42. Sebastiani P, Perls TT. Detection of significant groups in hierarchical clustering by resampling. *Front Genet*. 2016;7:144.. <https://doi.org/10.3389/fgene.2016.00144>. Published 2016 Aug 8.
43. Mikobi TM, Lukusa Tshilobo P, Aloni MN, Akilimali PZ, Mvumbi-Lelo G, Mbuyi-Muamba JM. Clinical phenotypes and the biological parameters of Congolese patients suffering from sickle cell anemia: a first report from Central Africa. *J Clin Lab Anal*. 2017;31(6):e22140. <https://doi.org/10.1002/jcla.22140>.