Decision trees for the severity and recurrence of acute splenic sequestration in sickle cell disease

Clara Allayous¹, Jean-Pierre Diara², Richard Emilion³, Therese Marianne-Pepin¹§

¹Laboratoire Dynamique des Protéines et Modélisation, Université des Antilles et de la Guyane, Pointe-à-Pitre, Guadeloupe, France
²Service de Pédiatrie, CHU de Pointe-à-Pitre, Guadeloupe, France
³Laboratoire MAPMO, Université d'Orleans, Orleans, France

§Corresponding author

Email addresses:

CA: callayou@univ-ag.fr
JPD: diaraky@wanadoo.fr
RE: richard.emilion@univ-orleans.fr
TMP: tmariann@univ-ag.fr
Abstract

Background

Acute splenic sequestration is one of the major clinical forms of sickle cell disease. This potentially life-threatening complication is defined by a rapid sequestration of sickle red blood cells in the spleen. Approximately 30% of young patients will develop this complication and 49% of them will have recurrences.

Methods

The decision tree method is a statistical method that can act as a medical decision tool in the case of a complex biological problem. The decision tree method is used here to determine the variables involved in the severity and recurrence of acute splenic sequestration crises.

Results

A prior decision tree shows that platelet level, variation in spleen size, patient age and hygrometry also called relative humidity define acute splenic sequestration severity. A second one shows that spleen size variation, atmospheric pressure and patient age define recurrence.

Conclusions

For the first time, biological and environmental parameters are combined to discover some interesting rules which can lead to more accurate evaluations of acute splenic sequestration severity and recurrence. Thus, this method provides a sharp diagnostic tool which could improve medical treatment for better patient care.
Background

Sickle cell disease is a monogenic hemoglobinopathy resulting from a mutation of the β-globin gene (Glu6Val). After releasing oxygen in the peripheral tissues, sickle hemoglobins interact hydrophobically and stick together to form long chains. These rigid polymers distort the red cells into a characteristic sickle shape which cause them to obstruct blood circulation, leading to tissue damage [1-3]. Acute splenic sequestration (ASS) is one of the major clinical manifestations of sickle cell disease and a potentially life-threatening complication. It is found mainly in children but only rarely in adults [4, 5]. ASS is an important cause of mortality and morbidity in HbSS children less than five years old [6, 7]. About 30% of them develop this complication and the death rate is 10 to 15%. Early detection of spleen enlargement, usually by parents using palpation techniques, reduces the mortality and morbidity rate linked to this complication [7-11].

ASS can be classified as being either severe or mild, with the latter being more common [14]. We have defined a mild ASS episode as a slight increase in spleen size with a decrease in hemoglobin of 2-3 g/dl [14]. On the other hand, severe ASSs are associated with a greater drop in hemoglobin level: less than 2 g/dl [14]. ASS is a recurrent complication in 49% of cases and recurrences have a 41% death rate, as shown by Emond et al [7, 11]. Their studies show that severity and recurrence are two major parameters in vital prognosis and patient care [15]. The aim of this paper is to provide medical doctors with some tools to assess the severity level of an ASS episode and to predict ASS recurrences in a given patient. To reach this goal, we have used a database which contains traditional biological parameters such as spleen size and patient age, but also contains environmental parameters. For the first time, the Decision Tree Method (DTM) was applied, providing some interesting decision-
making tools. As an example of our results we show that some parameters, such as platelet level and atmospheric pressure, that are not usually classified as triggering factors may play an important role in ASS occurrence.

**Methods**

**Patients, clinical event definition**

Data about HbSS patients having ASS events was collected by the Sickle Cell Center “Guy-Mérault” in Guadeloupe between 1980 and 2000. These data concern a set of 32 young patients with a total of 75 ASSs (tables 1 and 2). 16% were female and 84% were male. No deaths were observed after these ASS periods. No splenectomies were performed on these patients. Clinical, hematological, and environmental data were collected and recorded into standardized and computerized forms (tables 2 and 3). Only ASS events which required the children to be brought to the hospital pediatric department for medical care were recorded. An ASS event was defined as a sudden drop in hemoglobin level of at least 20% associated with an increase in palpable spleen size of at least 2 cm [9, 11, 12], and classified according to severity level and recurrence. Mild to moderate thrombocytopenia is often present [11, 13]. The severity and the recurrence of ASS were assessed when the patient was attended to by the medical staff.

Patients are classified as presenting severe ASS or mild ASS according to their hemoglobin level. In this study, this parameter as well as others associated with it, such as hematocrit value and red cell levels, was not used to build decision tree models (table 2).
The definition of recurrence is based on historical observations made by the Sickle Cell Center in order to distinguish between patients developing one or two fortuitous ASS events and patients developing three or more linked ASS episodes. This is an important distinction, as in some countries, patients often go through spleen surgery after one or two crises.

**Parameters**

We have taken 12 biological, environmental, and physical parameters into account. These are presented in Table 3. Biological parameters refer to the patient and to hematological values where “spleen size variation” is determined by measuring the difference between basal spleen size and its value during ASS.

Environmental parameters like atmospheric pressure and hygrometry were recorded at the meteorological unit of “Météo France”, in Le Raizet, Guadeloupe, and were expressed as average daily values. Average annual values are 77% and 1010.25 hPa respectively for hygrometry and atmospheric pressure.

**Decision trees**

A decision tree is a statistical/computer science tool that explains how a class variable depends on some predictor variables. In our case, the class variable is ASS severity or ASS recurrence, whereas the predictor variables are biological and environmental parameters.

Typically, we start with all of the individuals within a unique class called the root node of the tree. This node is then split into two child nodes according to criteria that ensure that these children are more homogenous than the root with respect to the class variable. This operation is iterated on each child node and is stopped when the node is small enough or is sufficiently homogenous. Each terminal node is called a leaf node.
A branch is a path going from the root to a leaf.

For example, in Figure 1, the root node is split into two children nodes according to variable $X$ “Platelet level” and $\alpha = 8 \times 10^4$/mm$^3$. The choice of $X$ and $\alpha$ is decided by the program as the best split among all possible splits. Further, in Figure 1, we see that one of the two child nodes is split into two nodes according to variable $X$ “Spleen size variation” and $\alpha = 7$ cm.

Hence, there is a corresponding decision rule for each branch of the decision tree, such as:

if $X_j < \alpha_j$ and $X_i \geq \alpha_i$ and $X_6 \geq \alpha_6$ then $Y = +$

with a confidence of 85%, the $X_i$’s denoting the predictor variables and $Y$ being the class variable (for two classes + and -).

In this paper, we used the well-known Classification And Regression Tree algorithm (CART) [16]. The decision tree is built in two steps, using SODAS free software. A first expansion step yields a decision tree built with a low error level for each class cut. The second pruning step yields an optimal decision tree (optimal classification). Only decision trees obtained after the pruning step are presented in this paper.

**Results**

DTM provides two models, the first one explaining ASS severity and the second ASS recurrence.

Physical, biological, and environmental parameters, listed in table 3, are used as predictor variables to build decision trees. Each branch of the tree yields an easily understandable rule explaining the class variable. Thus, we get some decision-making rules which may be useful in medical prognoses.
ASS severity

The severity DT model presented in Figure 1 suggests three rules for severe ASS events and three rules for mild ASS episodes. They depend on the four predictor variables of platelet level, spleen size variation, patient age and hygrometry.

An ASS event will be considered as severe when:

a) Platelet level is less than $8 \times 10^4$/mm$^3$, spleen size variation is less than 7 cm and hygrometry is less than 70%

b) Platelet level is less than $8 \times 10^4$/mm$^3$ and spleen size variation is more than 7 cm

c) Platelet level is more than $8 \times 10^4$/mm$^3$ and patient’s age is between 184 and 238 days (around 6 and 8 months).

On the contrary, an ASS event will be considered as mild when:

a) Platelet level is less than $8 \times 10^4$/mm$^3$, spleen size variation is less than 7 cm and hygrometry is more than 70%

b) Platelet level is more than $8 \times 10^4$/mm$^3$ and patient’s age is less than 184 days (around 6 months)

c) Platelet level is more than $8 \times 10^4$/mm$^3$ and patient’s age is more than 238 days (around 8 months).

An analysis of this decision tree showing the chances that these patients stand of developing severe or mild ASS, with a low error rate of about 4%, is presented in Table 4.

ASS recurrence

As mentioned before, a patient having at least three ASS events is classified as a recurrent one. The recurrence decision tree model is shown in Figure 2. This tree indicates that there is only one rule over the four obtained which explains ASS
recurrence. This rule relies on the three predictor variables of age, atmospheric pressure, and spleen size variation. Recurrent events will occur when spleen size variation is less than 4 cm, atmospheric pressure is more than 1013.6 hPa, and patient’s age is less than 409 days (around 14 months), during the first ASS.

On the contrary, a patient will develop one or two fortuitous ASS episodes when:

a) Spleen size variation is less than 4 cm, atmospheric pressure is less than 1013.6 hPa

b) Spleen size variation is less than 4 cm, atmospheric pressure is more than 1013.6 hPa and patient’s age is more than 409 days (around 14 months)

c) Spleen size variation is more than 4 cm.

According to our results, patients have a higher probability of developing one or two ASS events, with a low error rate of about 8%, as presented in Table 5.

When comparing the two decision trees, common predictor variables such as age and spleen size variation are found. Thus, a spleen size variation of less than 7 cm may define a severe or a mild ASS (Figure 2), whereas a spleen size variation of less than 4 cm may define a recurrent one. The lapse of time between the onset of ASS and medical attention does not seem to have any impact on recurrence.


Discussion

ASS severity and recurrence

The purpose of the present paper is to identify the factors involved in ASS severity and recurrence using the decision tree technique and to discriminate between factors triggering each phenomenon. This method is very effective for simple and fast medical decision making rules to solve complex biological problems. When a child presenting an ASS episode is brought to the hospital, the medical staff must note the level of severity of the ASS and whether or not it is a fortuitous event. These are fundamental criteria for the vital assessment of the child’s condition.

Biological and environmental variables combine to point out new decision-making rules in both events. These rules enable us to confirm classical ASS descriptors such as platelet level, spleen size variation, and age, but also define their close relationship to climatic factors such as atmospheric pressure and hygrometry. They not only determine the nature of the variables involved, but also their relative importance. “Severity DT” will make it possible to assess ASS medical severity with a low error rate (4%). Assessment of severity may influence the vital prognosis, even if this notion principally depends on the lapse between ASS beginning and medical intervention. This model offers biological predictor variables classically used by the medical staff, such as platelet level, spleen size variation and age, along with their associated values [9, 11, 12]. The second DT enables clinicians to split HbSS patients developing ASS complications into two groups. The first one concerns children having one or two fortuitous events, while the second deals with children experiencing more ASS phenomena with reinforced treatments.
These empirical medical observations are mathematically confirmed in the two
two models suggested. These observations indicate that ASS might be considered a
splenic vascular occlusion episode centered on blood cell adhesion. This hypothesis
has already been suggested by Warkentin et al [17]. These authors posit that ASS
could be an intrasplenic sickling crisis with a resulting decline in venous outflow of
blood and splenic pooling of red blood cells [17].

Moreover, the present study clearly shows a relationship between climatic
environment and ASS severity or recurrence. Thus, the primary importance of
hygrometry (severity) and atmospheric pressure (recurrence) in the two
pathophysiological processes has been stressed. Hygrometry stands out as an
important factor with a slightly lower value (70%) than the average hygrometry
observed in Guadeloupe (77%). In the same way, atmospheric pressure is linked to
the phenomenon of recurrence. Its calculated value (1013.6 hPa) is very close to the
normal value in Guadeloupe (1010.25 hPa). Overall atmospheric pressure variation is
low in Guadeloupe but is observed during some particular conditions as solstices, bad
weather conditions, and the alternation between day and night. This study was made
with daily variations, making it difficult to understand whether ASS is influenced by a
certain level of pressure or a change in atmospheric pressure.

Climatic factors, notably cold and seasonal changes, have long been known to impact
hygrometry. These climatic factors are also known to participate in HbSS vascular
occlusion, in temperate as well as tropical regions [18-22]. Our results are in
accordance with several other sickle cell disease studies showing that climatic
variables are determining parameters in vascular occlusion.
Seasonal patterns may reflect viral or other infective agents responsible for the occurrence of painful crises and are well documented [20, 21].

**Can pathophysiological models be set?**

Numerous questions can be asked using these models. Any hypotheses should be checked out with other experiments, as other information might be relevant; for instance, the number of ASS might be higher at particular times of day. ASS vascular damage can have consequences on the spleen. Some studies have suggested that HbSS patients having bacterial infections are more subject to ASS recurrence [24, 25]. This vulnerability to infection is directly connected to an early loss of spleen function [25]. In addition, splenomegaly during the first six months of life is associated with an increase in infection risks [25, 26].

The present study suggests new investigation routes and gives additional information in the pathophysiology of ASS. In the two models presented, key variables obtained are already used in case of vascular occlusion in SS disease. Platelets seem to play an important role in sickle cell disease occurrence, secreting soluble factors including thrombospondin that can contribute to vaso-occlusive complications [27, 28]. These complications are promoted by erythrocytes and sickle red cell adherence, microvascular occlusion and platelet aggregation [29-35]. In addition, neutrophils may play a role in adhesion [36, 37]. Moreover, some authors suggest that the vascular occlusion model is the result of interactions between leukocytes, erythrocytes, platelets, plasma proteins, and blood vessel walls [38]. A striking observation is that neither of the two models obtained indicate the implication of neutrophils in this complex phenomenon. Many authors suggest a link between viral infections and vascular occlusions through the adhesion of red cells and endothelium. In fact, Smolinski et al indicate that viruses may induce adhesion which could
accelerate occlusion occurrence [39]. Vascular adhesion also implies the expression of cell-surface adhesion molecules which increases in sickle cell disease, probably via a mechanical effect [15, 37, 40]. Moreover, it has been reported that HbSS children presenting infections have developed ASS. These patients present an acute drop in hemoglobin level, reticulocytopenia and sudden splenic enlargement, which is in agreement with our results [41]. These different hypotheses must be confirmed by additional studies to determine the role of the different partners involved in ASS occurrence.

Conclusions

DT methods have enabled the definition of precise, simple, and easy to use rules, almost anywhere in the world, to assess ASS severity and recurrence where biological and environmental parameters are both involved. This study mathematically confirms the predictor variables used in an empirical way by the medical staff. Nevertheless, the relationship between calculated values and ASS remains unclear. Further investigations are necessary to determine the physiological mechanisms involved and to better understand this complex phenomenon.

Thus, these two models not only confirm the classical medical factors implicated in ASS, but also confirm ASS severity even if this notion depends on the lapse of time between the onset of ASS and medical treatment. It also enables early detection of children developing multiple ASS events, leading to thorough parental prevention, specific medical treatment, and better quality of life [42, 43].

To conclude, the rules obtained need to be tested in order to confirm and refine the related biological and environmental parameters. In addition, this study could be
extended to other patients with sickle cell who develop different complications in order to determine the specific factors of specific complications. These results would be useful in establishing different sub-populations linked to specific complications, in order to predict clinical development and adapt specific treatments to each population.

**List of abbreviations**

ASS, Acute splenic sequestration

DTM, Decision tree method

**Competing interests**

The author(s) declare that they have no competing interests.

**Authors' contributions**

CA conducted the analysis, interpreted and discussed the data, and wrote the manuscript. JPD participated in the analysis and discussion of the results. TMP and RE are the primary investigators and senior authors. RE designed and conducted the decision tree method analysis and contributed to writing the manuscript. TMP designed and coordinated the study and critically reviewed and revised the manuscript. All authors read and approved the final manuscript.
Acknowledgements

This study was conducted with the financial support of “La Région Guadeloupe”, “Le Fonds Social Européen”, “Le Groupement Pharmaceutique Guadeloupéen” and “Le Soroptimist Club Riviera-Sud”.

References


23. Jones S, Duncan ER, Thomas N, Walters J, Dick MC, Height SE, Stephens AD, Thein SL, Rees DC: **Windy weather and low humidity are associated with an increased number of hospital admissions for acute pain and sickle cell disease in an urban environment with a maritime temperate climate.** *Br. J. Haematol.* 2005, **131**:530-533.


27. Richardson SG, Matthews KB, Stuart J, Geddes AM, Wilcox RM: **Serial changes in coagulation and viscosity during sickle-cell crisis.** *Br. J. Haematol.* 1979, **41**:95-103.


Figures

Figure 1 - Severity decision tree.

The different variables defined for ASS severity and their calculated values are represented as branches. Decision nodes are represented as circles. Terminal outcome nodes are represented as triangles on the right margin.

Figure 2 - Recurrence decision tree.

The different variables defined for ASS recurrence and their calculated values are represented as branches. Decision nodes are represented as circles. Terminal outcome nodes are represented as triangles on the right margin.

Tables

Table 1 - Database characteristics.

Table 2 - Database extract.

ASSN, ASS number; Atm, atmospheric pressure; P, platelet level; PN, Patient Number; L, leukocyte level; H, hygrometry; $\Delta S$, spleen size variation; $\Delta L$, leukocyte level variation; $\Delta N$, neutrophil level variation.

Table 3 - Physical, biological and environmental parameters used as predictor variables.

Table 4 - Statistics summary of « Severity Decision Tree ».

Severe ASS = +
Mild ASS = -

Error rate is indicated under the table.

Table 5 - Statistics summary of « Recurrence Decision Tree ».

3 ASS events and more = +
1 or 2 ASS events = -

Error rate is indicated under the table.
### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Patients n = 32</th>
<th>Total ASS phenomena n = 75</th>
<th>Patients with 3 ASS events and more</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girls 5 (16%)</td>
<td>6 (8%)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Boys 27 (84%)</td>
<td>69 (92%)</td>
<td>8 (26 events)</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>P N</th>
<th>ASS N</th>
<th>Birth date</th>
<th>ASS date</th>
<th>ΔS</th>
<th>ΔL</th>
<th>ΔN</th>
<th>ΔLiver size</th>
<th>Atm</th>
<th>H</th>
<th>P</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>09/09/89</td>
<td>01/20/90</td>
<td>3</td>
<td>2100</td>
<td>-261</td>
<td>0</td>
<td>1015.0</td>
<td>71</td>
<td>60000</td>
<td>8100</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>09/03/89</td>
<td>09/24/90</td>
<td>4</td>
<td>5034</td>
<td>2080</td>
<td>0</td>
<td>1011.6</td>
<td>78</td>
<td>60000</td>
<td>12600</td>
</tr>
<tr>
<td>29</td>
<td>1</td>
<td>04/19/97</td>
<td>06/01/88</td>
<td>3</td>
<td>10120</td>
<td>-3892</td>
<td>0</td>
<td>1014.7</td>
<td>78</td>
<td>60000</td>
<td>30200</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>Physical variables</th>
<th>Biological variables</th>
<th>Environmental variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASS number</td>
<td>Patient age</td>
<td>Atmospheric pressure</td>
</tr>
<tr>
<td>Patient number</td>
<td>Platelet level</td>
<td>Hygrometry</td>
</tr>
<tr>
<td>Total number of ASS episodes</td>
<td>Spleen size</td>
<td>Delta Leukocytes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Delta Neutrophils</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Delta Liver size</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Delta Spleen size</td>
</tr>
</tbody>
</table>

### Table 4

<table>
<thead>
<tr>
<th>ASS</th>
<th>Tree Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 (+)</td>
<td>9 (+), 1(-)</td>
</tr>
<tr>
<td>65 (-)</td>
<td>2 (+), 63 (-)</td>
</tr>
</tbody>
</table>

Error rate: \( \frac{1+2}{10+65} = 4\% \)
<table>
<thead>
<tr>
<th>ASS</th>
<th>Tree Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 (+)</td>
<td>6 (+), 2 (-)</td>
</tr>
<tr>
<td>18 (-)</td>
<td>0 (+), 18 (-)</td>
</tr>
</tbody>
</table>

Error rate: \( \frac{2+0}{8+18} = 8\% \)
Platelet level

Spleen size variation

Age

Hygrometry

< 8 x 10^4/mm^3

< 7 cm

< 70%

< 238 days

> 8 x 10^4/mm^3

> 238 days

> 70%

> 7 cm

< 184 days

> 184 days

Severe ASS

Mild ASS

Severe ASS

Mild ASS

Severe ASS

Mild ASS

Figure 1
Figure 2

Spleen size variation

Atmospheric pressure

< 1013.60 hPa

< 4 cm

< 409 days

1 or 2 ASS episodes

> 3 ASS episodes

Age

> 1013.60 hPa

> 4 cm

> 409 days

1 or 2 ASS episodes

1 or 2 ASS episodes