ORGAN FAILURE PREDICTION BASED ON CLINICAL ADVERSE EVENTS: A CLUSTER MODEL APPROACH

Álvaro Silva[1], João Pereira[2], Manuel Santos[3], Lopes Gomes[4], José Neves[5]

[1]Serviço de Cuidados Intensivos, Hospital Geral de Santo António,
Porto, Portugal
a.moreirasilva@mail.telepac.pt
[2]Escola Superior de Educação de Viana do Castelo
Viana do Castelo, Portugal
jpereira@ese.ipvc.pt
[3]Departamento de Sistemas de Informação, Universidade do Minho,
Guimarães, Portugal
mfs@dsi.uminho.pt
[4]Clínica Medica I, Instituto de Ciências Biomédicas Abel Salazar,
Porto, Portugal
Cardiologia.hgsa@mail.telepac.pt
[5]Departamento de Informática, Universidade do Minho,
Braga, Portugal
jneves@di.uminho.pt

Abstract

Medical prognosis has played an increasing role in health, namely in the critical care medicine. These factors have induced the medical community to take a more active interest in developing models for mortality prediction based on Artificial Intelligence (AI) techniques [1], that make possible the doctors pro-active action. In this context, the existence of large Databases (DB) containing Intensive Care Units (ICU) clinical information, motivate and enable the application of Data Mining (DM) techniques, in a Knowledge Discovery Database process (KDD), to induce prediction models of organ failure in a much more efficient way than other approaches (e.g., Logistic Regression)[2].

In this study we used a clustering approach as a strategy to predict the organs failures, based on the occurrence of Clinical Adverse Events (CAE) during the patient's stay in the ICU. It follows other studies made in this area (Santos e Silva [3,4]), where DM techniques were used to predict hospital mortality.

The clinical database used was created from part of the database originated from the EURICUS II [5] study, after being subjected to processes elaborated with SPSS and SPSS Clementine tools[6].

Various models were created by applying the C5 Algorithm [7] in order to generate decision rules for easy interpretation, of each organ, for each day of stay, and for each cluster defined. This approach enables a better accuracy on some models, a homogeneous segmentation of patient's (for further medical analysis) and a tracking model of the patients' evolution during his stay. In such a way it is possible to identify the cluster where he fits and, consequently, to make use of the adjusted model to predict the status of a certain organ on the following day.

Key Words

Data Mining, Knowledge Discovery from Databases, Artificial Neural Networks, Organ Failure, Mortality Predicting Models based on Intermediate Outcomes, Intensive Care Units.

1. Introduction

Decision Support Systems (DSS) are present on different areas of activity where various solutions for different problems origins are needed. Health care is not an exception. So we shouldn´t be surprised that realities like hospital ICUs take advantage of these DSS.

The Sequential Organ Failure Assessment (SOFA) [8,9], score the degree of failure of six organs. This score is calculated on a daily basis.

Considering the admission data and other variables taken on first admission day, as well as CAEs taken during the patient's stay in ICU, it is possible to predict the failure of each organ for the day following the last day of collecting data (time series). This is the assumption this study is going to validate.

This document shows a study on the influence of the CAEs to the dysfunction or failure of human organs occurring during the patients' stay at the ICU.

To generate predicting models for each organ, and for each day, we followed a cluster model approach. This allowed the conception of a clustering framework, representing the patient evolution along his stay. The paper is organized as follows: in the second and third chapters present the clinical data and some definitions about events and critical events; the fourth and fifth chapters introduce the process of data preparation, transformation and model generation; the last two chapters, before the eighth that concludes the article, are dedicated to the results presenting the achieved accuracies and the framework to organize the models into clusters.

2. Clinical Data

In this study a database was created based on EURICUS II, a study made in 42 ICUs on 9 UE countries, between 1997 and 1999. For a period of 10 months every admission to the ICU was included. This database includes the characteristics that define case-mix [10], namely the Age, the Type of Admission (unscheduled surgery, scheduled surgery and medical), the Admission Source (Operating Bloc, Recovery Room, Emergency Room, Infirmary, other ICU, other Hospital, other sources), Diagnosis, Gravity Index defined by SAPS II [11], SOFA of each Organ System (Respiratory, Coagulation, Liver, Cardiovascular, Central Nervous and Renal), Mortality in the ICU and in the Hospital; Number of CAEs for each of the parameters monitored continuously, Length of Stay and Admission Day.
In this study, from the 5355 patients admitted to the ICUs only 4425 (82.63%) stayed for two or more days, 3105 (57.98%) stayed three or more days and 2329 (43.49%) four days or over. For the data concerning the fifth day of stay, only 1845 (34.35%) patients were considered.

3. Clinical Adverse Events

We considered as Events (Ev) or Critical Events (CrEv) the occurrences of values out of the established limits for four physiologic variables that are monitored continuously. These four variables are Heart Rate (HR), Systolic Blood Pressure (BP), Oxygen Saturation (SaO2) and Urine Output (Diur). A group of clinical specialists determined the intervals considered normal for each one of these parameters. In the Table 1 it is presented the information related with the intervals considered as normal, as well as the situations in that they are considered as out of the limits.

Table 1 – Clinical Adverse Events.

Events

Event	Suggested Range	Continuosly out-of-range	Intermittently out-of-range
BP(mmHg)	**90-180**	≥ 10’	≥ 10’ in 30’
SaO2(%)	**≥ 90**	≥ 10’	≥ 10’ in 30’
HR(bpm)	**60-120**	≥ 10’	≥ 10’ in 30’
Diur(ml/hour)	**≥ 30**	≥ 1 hour	

Critical Events

Critical Event	Suggested Range	Continuosly out-of-range	Intermittently out-of-range	Event Anytime
BP(mmHg)	**90-180**	≥ 60’	≥ 60’ in 120’	BP < 60
SaO2(%)	**≥ 90**	≥ 60’	≥ 60’ in 120’	SpO_2 < 80
HR(bpm)	**60-120**	≥ 60’	≥ 60’ in 120’	HR < 30 or HR> 180
Diur(ml/hour)	**≥ 30**	≥ 2 hours		≤ 10

Adverse events were defined as binary variables, whose values correspond to one of two situations, in that the variable is within or not of the established limits (if yes, by how long). We considered as an Event when the value of the analyzed parameter maintains out of the limits, for a same or a period superior of continuous time to 10 min. (1 h. in the case of Diur.) and less then 60 min. (2 h in the case of Diur).
It is still considered an Event when, in a discontinuous way, values are verified out of the limits, but that are inferior to 10 min. and in a period of time of 30 min. maximum, since the sum of those is equal or greater to 10 min.
The definition of Critical Event is similar to the Event, but with different values. The times of 10 min. referred in the definition of Events, should be replaced by 1 hour, the 30 min. for 2 hours and 1 hour for 2 hours respectively.
A Critical Event can also be defined in some special situations, when the value of the analyzed parameter places among certain values that we could see in the Table 1.
We only can consider a new event, after a recovery period of 30 min. or more for BP, SaO2 and HR, and of 2 hours or more for Diur, with values inside of the intervals. In Critical Events, it should be considered a period greater than 2 hours for Diur and greater than 60 minutes for the remaining ones.

4. Data preparation

In this work we followed the Crisp-DM Methodology [12] and in the data preparation phase has been necessary to treat the wrong or omitted data. Besides, not all the variables were considered to generate the prediction models, as it is the case of the age, once it is already considered within SAPSII score.
The Table 2 shows the variables that were considered in this study and their description.

Table 2 - Variables Description.

Variable	Description	
ID	Pacient number	*
Respirat	Respiration System	**
Coagulat	Coagulation System	**
Liver	Liver System	**
Cardiova	Cardiovascular System	**
Cns	Central Nervous System	**
Renal	Renal System	**

Variable	Description	
Nrbpevnt	Number of BP Events/day	
Nrbpcriv	Number of BP Critical Events/day	
Nrofhrev	Number of HR Events/day	
Nrofhrcr	Number of HR Critical Events/day	
Nrofo2ev	Number of O2 Events/day	
Nrofo2cr	Number of O2 Critical Events/day	
Nrofurev	Number of Diur Events/day	
Nrofurcr	Number of Diur Critical Events/day	
Admfrom	Admission From	***
Admtype	Admission Type	***
SapsiiII	Simplified Acute Physiology Score	***
Diagn01	Diagnostic	***

* Not considered into the prediction models.
** Dependent variables.
*** Variables just considered in the first day.

Once we intend to predict an organ failure in a certain day, based in the data of previous days, it was necessary to transform the database structure, in order to capture a temporary sequence of the variables (time series). The Tables 3 and 4 represent these two states.

Table 3 – Database before the transformation.

ID	AdmFrom	AdmType	...	Nrofo2cr	Nrofurev	Nrofurcr
03	1	1		0	0	0
03	1	1		0	1	1
03	1	1		0	1	1
...	...	...	...	...	...	...

Table 4 – Database transformated.

ID	Adm From	Adm Type	...	Nrof o2c1	Nrof ure1	Nrof urc1	...	Nrof o2c3	Nrof ure3	Nrof urc3
03	1	1		0	0	0		0	1	1

There are variables that are only related to the first (AdmFrom, AdmType, SAPSII and Diagn01) and are not considered to the remaining days.
For instance, if we considered the prediction of respiratory organ failure in the third day of stay at the ICU, we can verify that the SOFAs obtained for the first and second days are not considered. However the number of Ev and CrEv obtained for these two days are considered for the prediction in cause.
We just considered a temporary horizon of five days, because, in medical terms, the fifth day of stay in a ICU is considered a critical point in terms of the evolution of the patient's clinical state. The first day was not considered for prediction effects, once the goal is to predict organ failure based on the data collected in the previous days. In any way, the data collected in the first day is used as input for all the other ones.

5. Models Definitions and Construction

Making use of SPSS Clementine tool, we submitted the database to a Kohonen Network [13], to segment it in three distinct groups. Later, it was fallen back upon the C5 algorithm in way to generate a model of decision rules to understand each one of those clusters. This way, we obtained 3 models for each one of the dependent variables, and for each one of the days of stay in the ICU, totalling 18 models for each day.
The necessary procedures for mining and generation of models for each one of the clusters, can be visualized in the Figure 1, that presents a stream of Clementine, elaborated for predicting central nervous system failure, in the fourth day of stay.

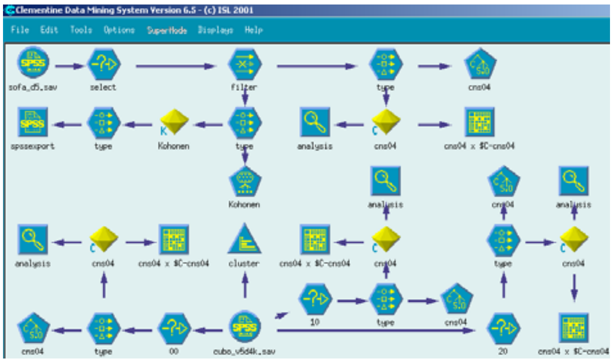

Figure 1 – Clementine stream example.

After having selected the variables that will be used in the generation of models, a Kohonen Network was applied to the database, in order to generate two additional variables, that are two coordinates which are assigned at each record, identifying the cluster that it belongs. These coordinates make possible the partition of the patients into three clusters and, later, applying the C5 algorithm to each one of these, in way to generate the respective decision rules.
The validation of those models was made through the cross validation method (10 fold) and, at the end, we could see the achieved results through the application of the models generated in each one of the clusters using a 2x2 matrix, which compares the results obtained with the real ones.

6. Achieved Accuracies

The fact of segmenting the database, enabled that in some clusters the achieved accuracy has been very high, as we can verify in the Table 5 (reached accuracies and standard errors, for the predicting models of the fifth day of stay).

Table 5 – Accuracies obtained for the 5th day of stay.

		Cluster 1	Cluster 2	Cluster 3
Respiratory System	Accuracy	74.20	75.00	75.80
	Std.Error	0.70	4.80	0.90
	N.Cases	1132	64	649
Coagulation System	Accuracy	92.60	98.60	94.50
	Std.Error	0.40	1.40	0.20
	N.Cases	649	65	1131
Liver System	Accuracy	97.90	98.60	97.20
	Std.Error	0.10	1.40	0.20
	N.Cases	1132	63	650
Cardiovascular System	Accuracy	80.60	89.80	81.50
	Std.Error	0.50	3.80	0.50
	N.Cases	1132	64	649
Central Nervous System	Accuracy	82.20	91.00	79.80
	Std.Error	0.50	2.50	1.00
	N.Cases	1132	64	649
Renal System	Accuracy	89.10	92.10	91.10
	Std.Error	0.50	3.60	0.50
	N.Cases	1134	62	649

7. Clustering Framework

The organ failure prediction models are denoted by $m(d_j, o_k, c_i,)$, they were generated for each of the 6 different organs o_k *where* $k \in$ *{Renal, Central Nervous, Cardiovascular, Liver, Coagulation, Respiratoty}*, for each of the first 5 days of the stay d_i, *where* $j \in \{2,3,4,5\}$ and were organized in three distinct clusters c_i, *where* $i \in \{0,1,2\}$ (the cases in each of these clusters are closer).
These models are organized in a cube framework that makes possible the graphical presentation of the patient course along the stay in ICU as we can see in the Figure 2. For a given patient we have a prediction model for each one of the 6 organs (o_k) indexed to the day of stay (d_j) and the correspondent cluster (c_i).

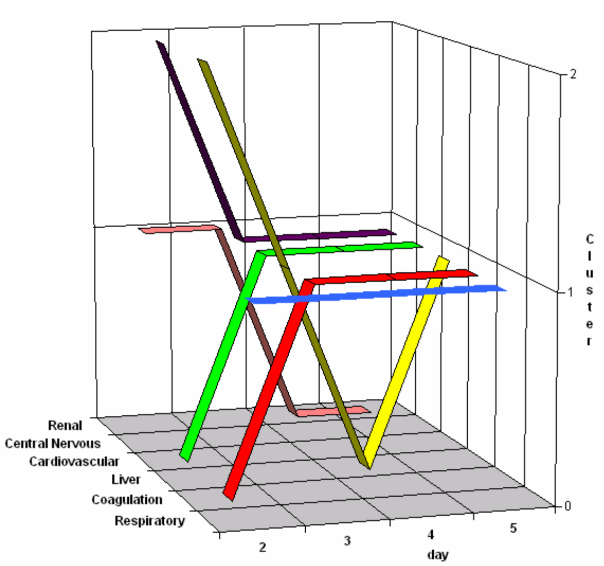

Figure 2 – Clustering Framework.

Be noticed that, in the same day, the correspondent models of a particular patient may belong to different clusters.
In the example presented in the Figure 2, the prediction models for day 2 are given by:

m(2, Renal, 1);
m(2, Central Nervous, 2);
m(2, Cardiovascular, 0);
m(2, Liver, 2);
m(2, Coagulation, 0);
m(2, Respiratory, 1).

For the same patient, for the fourth day, the models are given by:

m(4, Renal, 0);
m(4, Central Nervous, 1);
m(4, Cardiovascular, 1);
m(4, Liver, 0);
m(4, Coagulation, 1);
m(4, Respiratory, 1).

As we can see, the prediction model for the *Central Nervous* system of this patient changed from the cluster 2 to the first cluster.
The database segmentation criteria, for the third day of stay related with the *Renal* system failure, can be visualized under the form of rules as following:

Rules for 0:
Rule #1 for 0:
if admfrom1 > 1
and admtype1 > 2
then -> 0 (1764, 0.999)
Rules for 1:
Rule #1 for 1:
if admfrom1 > 1
and admtype1 > 1
and admtype1 <= 2
then -> 1 (98, 0.99)
Rule #2 for 1:
if admfrom1 <= 1
and admtype1 > 2
then -> 1 (12, 0.929)
Rules for 2:
Rule #1 for 2:
if admtype1 <= 1
then -> 2 (774, 0.999)
Rule #2 for 2:
if admfrom1 <= 1
and admtype1 <= 2
then -> 2 (922, 0.999)
Default : -> 0

In this case, the variables that determined the classification in three clusters, were the Admission Type and Admission From. The values presented between

parentheses stand for the support level and the confidence level, respectively.
As we can see, there is only one rule that respect to cluster 0, and two rules for each one of the clusters 2 and 3. If the admission from is medical, and the admission type is not the Operating Bloc or Recovery Room, the patient will be in the cluster 0. This rule was applied to 1764 cases.
The first day of stay was not considered, once it doesn't make sense to predict organ failure for this day, because the only data we have was collected in the same day.
It can be said for instance that, a patient with SAPSII 37, with the type of admission scheduled surgery, coming from the infirmary, with diagnosis pos-operative and with an event of O2 in the first day; it's framed in the cluster 1.
Each of the 72 models referred in the framework correspond to rules generated by the C5 algorithm. Consider for example the rules that predict the *Central Nervous* system for the fifth day:

```
Rules for 0:
    Rule #1 for 0:
        if sapsii1 <= 54
        then -> 0
    Rule #2 for 0:
        if nrofurev1 <= 5
        and nrofhrev2 > 2
        and nrofhrev4 <= 3
        then -> 0
Rules for 1:
    Rule #1 for 1:
        if sapsii1 > 54
        and nrbpevnt3 <= 5
        and nrofhrev4 > 3
        then -> 1
    Rule #2 for 1:
        if admfrom1 <= 3
        and sapsii1 > 54
        and diagn01<= 0
        and nrofurev4 > 2
        then -> 1
    Rule #3 for 1:
        if nrofurev1 > 1
        and sapsii1 > 54
        and diagn01 > 0
        and nrbpevnt2 <= 0
        and nrofhrcr2 <= 0
        then -> 1
          ……
          ……
    Rule #11 for 1:
        if nrbpevnt1 <= 0
        and admfrom1 <= 3
        and sapsii1 > 54
        and nrofurcr2 <= 1
        and nrbpcriv3 <= 0
        then -> 1
  Default : -> 0
```

8. Conclusions and Further Work

As demonstrated in this study, the Clinical Adverse Events (CAE) that happens during the stay period in the ICU are important factors in the process of organ failure prediction. These CAEs, casemix and SAPSII were the base of the variables considered for organ failure prediction, and the interesting results obtained, are result of all this, as well of the clustering approach adopted.
Besides, this study presented a clustering framework, with the purpose of identifying and applying the model generated for the cluster in which a patient frames to, according to his characteristics. The majority of the models revealed high accuracies; witch is very useful for decision support.
Further work will include the optimization of each generated model and the deep medical characterization of the patients of each model (in the cube).

References

[1] Hanson William C., Marshall Bryan E.: Artificial Intelligence applications in the intensive care unit, *Critical Care Medicine, Volume 29, Number 2*, 2001, 427:435.

[2] Gilles Clermont, Derek C. Angus, Stephen M. DiRusso, et al: Predicting hospital mortality for patients in the intensive care unit: A comparison of artificial neural networks with logistic regression models, *Critical Care Medicine, Volume 29, Number 2*, 2001, 291:296.

[3] Santos Manuel, Silva Álvaro, Abelha António, et al: Predicting Outcome in Critically Ill Patients using Artificial Intelligence Models, *Artificial Intelligence and Applications*, Málaga, Spain, 2002, 47-51.

[4] Santos Manuel, Neves José, Abelha António, et al: Augmented Data Mining over Clinical Databases: Using Learning Classifier Systems, *Fourth International Conference on Enterprise Information Systems*, Ciudad Real, Spain, 2002, 512-516.

[5] http://www.frice.nl, Foundation for Research on Intensive Care in Europe, visited in March 2003.

[6] http://www.spss.com, SPSS Inc., visited in March 2003.

[7] http://www.rulequest.com, Quinlan J.R. C5.0 Data Mining Tool, visited in March 2003.

[8] Vincent J-L, De Mendonça A, Cantraine F. et al., Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study, *Critical Care Medicine, Volume 26,* 1998, 1793-1800.

[9] Moreno R., Vincent J-L., Matos R. et al: The use of maximum SOFA score to quantify organ dysfunction/failure in intensive care. Results of a prospective, multicentre study, 1999, *Intensive Care Med 25*, 686-696.

[10] Fetter RB, Shin Y, Freeman JL, et al. Case-mix definition by diagnosis-related groups, 1980, *Medical Care 18*, 1-53.

[11] Le Gall JR, Lemeshow S, Saulnier F., A new simplified acute physiology score (SAPS II) base on an European/North American multicenter study, *JAMA* 1993:270:2957-2963.

[12] Chapman, P., et al., CRISP 1.0 Process and User Guide. 2000, CRISP–DM Consortium (http://www.crisp-dm.org).

[13] T. Kohonen, "Self--Organizing Maps," Springer-Verlag, 1995.