Estimation of Treatment Outcome in elderly ICU Patients using Clustering Techniques

Project supervisor:
Dr. Ameen Abu-Hanna

Project Machine Learning
October 2007

Written by:
Maarten Buiter
Mark de Groot
Evelyn Lai

Contents

1. Introduction

1.1 Elderly patients and the intensive care unit

Worldwide, life expectancy of people is increasing with every decade. Healthcare plays a large role in this process, as it has been able to prevent, cure and manage many diseases that used to lower average life expectancies. The downside of this is that although many people live longer in good quality, the people who do fall ill at an advanced age present with a complicated spectrum of co-morbidity and functional disabilities, which makes caring for their health a very delicate matter.[1,2] This is especially true for Intensive Care treatment, where the increasing number of elderly patients demands large amounts of capacity.

1.2 Problem

Patients from this elderly population have a high mortality risk when admitted to an Intensive Care unit (ICU). For some of these patients starting a high intensity treatment will not lead to a positive outcome and the patient will not survive his hospital stay, despite the intensive treatment.

1.3 Our goal

It would be beneficial, from both an economic and a humane point of view, to be able to identify the patients for whom treatment is not going to make a difference. De Rooij et al. focussed on developing a prognostic model for very old ICU patients (80 years or older) that could reliably identify patients with very high mortality risk. They obtained their data from the Dutch National Intensive Care Evaluation (NICE) database and based their research on the following four reasons:

1) *Focus attention on groups of patients for whom current treatments may be insufficient; this in itself could lead to an improvement in care.*
2) *For some medical studies, enrolment of high-risk patients in clinical trials may provide the highest likelihood for finding a positive effect or facilitate investigating treatments with serious adverse side effects which are acceptable only if other treatments are not effective.*
3) *Identification of high-risk subgroups may be used for case-mix correction when comparing the outcomes of very elderly patients in different comparing the outcomes of very elderly patients in different ICUs.*
4) *It may be used for providing optimal information to patients, their relatives, and caregivers. Very elderly patients do not necessarily prefer intensive care treatment over palliative care that aims at comfort and pain relief.*[1]

These four reasons show that it is important to get a clear understanding of the risk factors in elderly patients, and how they are related to a patient's response to an intensive care treatment. In order to come to this understanding, we used methods from the field of Artificial Intelligence (AI) to analyse the original data of the Rooij at al study. With these methods, we built a

predictive model that can be used to estimate, based on information available within the first 24 hours after admission to the ICU, what the outcome of the intensive treatment would be.

1.4 Research questions

We limited ourselves to answering the following research questions:

1) *Can the clustering techniques taken from Artificial Intelligence techniques be used to describe our patient population in a sensible manner, and which technique produces the best results?*
2) *What does the model we created tell us about our patient population?*

2. Methods

2.1 Data Exploration

We used parts of the original dataset used by de Rooij et al. for our research. All patients in this set were aged 80 years or older. This dataset was modified for use with our clustering methods. The original data is collected from several Dutch Intensive Care facilities, and were assembled by the NICE Foundation. We consulted the NICE data dictionary for a description of the variables in the dataset.[1]

We used Microsoft Excel (versions 2000, 2003 and 2007) and the Statistical Package for the Social Sciences (SPSS) 14 and 15 for our preliminary data description. The Waikato Environment for Knowledge Analysis (Weka) 3.4.11, developed by the University of Waikato, was used to perform the clustering. Literature on clustering itself, and the COBWEB clustering technique in particular, was consulted to better understand the fundamental concepts of the clustering algorithms.

2.2 Clustering

The two algorithms used to cluster the data with were “SimpleKMeans” (K-Means clustering) and “Cobweb” (hierarchical conceptual clustering), both unsupervised techniques. We have chosen to focus on predicting mortality. We identified six attributes in the dataset that had a strong affiliation with the outcome variable. This outcome variable (*died*) indicated whether a patient had survived his stay at the ICU. We identified the entropy of the entire dataset, and used Weka’s “InfoGainAttributeEval” function to determine which attributes had the strongest predictive value for the outcome variable.

We compared the quality of the clusters produced by both algorithms based on the amount of incorrectly clustered instances, as this is a clear measure that can be used in both clustering techniques. Each algorithm’s parameters have been chosen such that the algorithm would produce three clusters, with the lowest amount of incorrectly clustered instances.

3. Results

3.1 Initial Data Exploration

The NICE dataset that we used in our research contained 500 instances of patients who were admitted to intensive care units. Each instance consisted of 59 variables (two nominal, 20 binary and 37 continuous). We explored the whole dataset and made table baseline characteristics containing the means/median of all the variables. The complete table can be found in the appendix, table 1 is a part of it.

Table 1[1] describes the attributes that we identified important for the clustering process. A complete description of the dataset can be found in appendix A. In Table 1, *GCS15* stands for the Glasgow Coma Scale, *heartrate.min* for the minimum heartrate, *urine.24* for the amount of urine produced in the first 24 hours, *los.icu* for the number of days the patient was treated at the ICU and *creat.max* for the maximum blood creatinine level. *AGR* represents the admittance reason where the number 0 is a scheduled surgery, the number 1 a medical reason and the number 2 an emergency surgery. In the supplied data the GCS15 variable which had a missing value were substituted with the value '15'.

Table 1 - Baseline Characteristics

Variable	Survivors (n = 340)	Non-survivors (n = 160)	Missing – n (%)
Aggregated Scores			
GCS15 - (median)	15,0	15,0	
Heartrate			
heartrate.min - mean (sd)	71,3 (17,1)	69,0 (31,3)	77 (15,4)
Various			
urine.24 - mean (sd)	3,3 (1,9)	2,0 (1,8)	94 (18,8)
los.icu - mean (sd)	2,7 (4,6)	4,4 (7,9)	2 (0,4)
Blood values			
creat.max - mean (sd)	121,1 (83,9)	178,7 (123,8)	92 (15,6)
AGR (n)	1 (83); 2 (31); 3 (46)	1 (70); 2 (56); 3 (214)	0 (0)

[1] See 'Table 9 – Baseline Characteristics' in the appendix for detailed information on all attributes

3.2 Information Gain Attribute Evaluation

We first explored the dataset as described earlier. Table 1 shows the differences between the survivors and the non-survivors. There are a couple of reasons why an attribute may not be selected:

- There are too many missing values
- There is (almost) no difference between survivors and non-survivors
- The attribute was important afterwards when the clusters were formed
- The information gain of the attribute was not high enough

The first two reasons have something to do with the fourth reason, when the first two reasons apply for a specified attribute, than the information gain (of that attribute) will not be high enough to be one of the first of six attributes which will be selected for clustering. When the information gain of an attribute is high it can differentiate more between cases than when an attribute has a low information gain. For clustering it is important to select the attributes which have the highest information gain. These attributes can discriminate the most between the cases that are present in the data. The third reason is special as some attributes have a high information gain, but consist of other attributes. Like the *nice.saps.score* and *nice.saps.prob*. It is more desirable to use these kinds of attributes later on when the clusters are formed, to determine if there are differences between each cluster.

Table 2. Ranked attributes

0.10837335	59	GCS15
0.08637301	58	AGR
0.06905964	31	urine.24
0.05049076	57	los.icu
0.05039664	20	heartrate.min
0.04933517	38	creat.max
0.04893298	37	creat.min

Out of the "InfoGainAttributeEval" that we have done with Weka, we obtained a ranking list (table 2), we decided on 6 attributes out of the top 10 of the list as our main focus on our research; *GCS15*, *AGR*, *urine.24*, *los.icu*, *creat.max* and *heartrate.min*.

3.3 Data Entropy

The amount of people, who died in the hospital or not			Probability and the first part of the calculation	
Attribute		n	p	$\log_2 p$
Died	0	340	340/500= 0,68	-0,55639
	1	160	160/500= 0,32	-1,64386
	Total	500		

$$-P_{+} \log_2 P_{+} = 0{,}32 * -1{,}64386 = -0{,}526033981$$

$$-P_{-} \log_2 P_{-} = -0{,}68 * -0{,}55639 = 0{,}378347477$$

$$Entropy = -P_{+} \log_2 P_{+} + -P_{-} \log_2 P_{-} = 0{,}904381$$

$$GCS15\ InfoGain = 0{,}108373$$

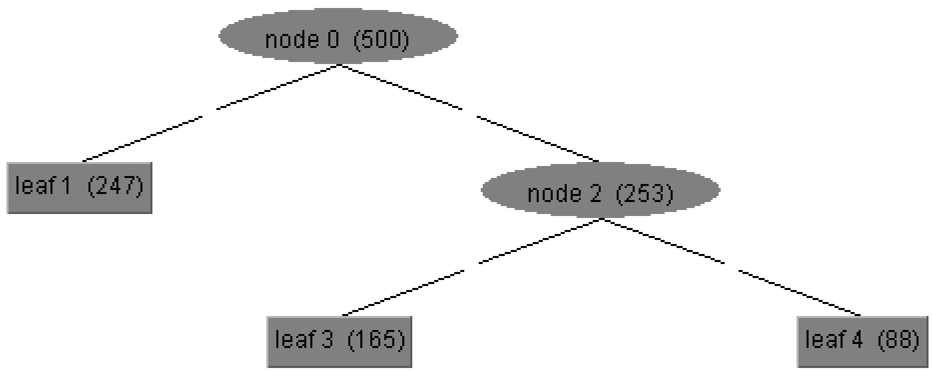
$$\frac{Informationgain}{Entropy} = \frac{0{,}108373}{0{,}904381} = 0{,}119831$$

The entropy of the whole dataset is 0,904381. The entropy is zero when a dataset is perfectly homogeneous and the entropy is one when a dataset is perfectly inhomogeneous. Our dataset is (nearly) perfectly inhomogeneous. We would like to cluster our dataset and therefore create three as much as possible homogeneous clusters. Therefore we selected the six attributes which have the highest information gain to lower the entropy of the dataset. GCS15 has the highest information gain and the ratio is 0,12. The higher the gain ratio, the more likely the subdivision into regions is useful for classification and clustering. We do not think this is much, but the information gain ratio of GCS15 is the highest of all. Therefore this attribute is useful for clustering (the remaining five out of six selected attributes are useful for the same reason).[II]

3.4 Cluster description

We obtained the Acuity and Cutoff with trial and error and decided this is the best cutoff and acuity; Cutoff = 0,12 and Acuity=1. With these parameters, Cobweb produced three clusters. Other values for Cutoff or Acuity would produce either more or fewer clusters, or create clusters which had a higher amount of incorrectly clustered instances. Fig 1 is a visual representation of the clustering produced by Cobweb. It consists of a root-node, one sub-node, and three leaves.

[II] http://www.cs.ualberta.ca/~aixplore/learning/DecisionTrees/InterArticle/1-DecisionTree.html



Tree from the cutoff and acuity mentioned above for all attributes accept for died

3.4.1 The Cobweb Clusters[III,IV]

Survival (nom)
When we look at the clusters – and look at the most interesting attributes – we can see that when a patient is assigned to cluster A the patient has a high probability to survive, while a patient assigned to cluster B results in a probability lower than the overall chance to survive. Cluster C compared to the overall population has almost the same probability to survive.

AGR (nom)
We have used this attribute to cluster on and patients in cluster A had an elective/planned surgery. While cluster B consists of all three reasons and cluster C consists of both medical reason for admission and a speedy/urgent surgery. Cluster B is populated the most with patients who have a medical admission reason and in cluster C the assignments are mostly patients which have an elective/planned surgery as a reason.

The other nominal attributes (nom)
Overall the patients in cluster A have less co-morbidities than in cluster C (cluster C has mostly the same probabilities as the overall population). Patients in cluster B are more likely – than the overall population – to have co-morbidities.

Creat.max (num)
The value of this attribute is – overall – the highest for patients in cluster B, while patients in cluster A have the lowest value and patients in cluster C are not that far from standard (the overall population).

[III] See 'Table 3 – cluster_Cobweb' and 'Table 6 – cluster_Cobweb' in the appendix for detailed information
[IV] Nom = nominal, num = numeric

Los.icu (num)

The length of stay of patients is – overall – the longest for patients clustered in cluster C, for patients in cluster B the length of stay is slightly longer than standard (the overall population) and patients in cluster A have left the ICU (very) fast, their length of stay is the shortest among all clusters.

Nice.saps.score (num)

The SAPS-score for patients in cluster B is the highest, while for patients in cluster A the score is the lowest and for patients in cluster C the score is almost the same as the overall population.

3.4.2 K-Means (s100)[V]

First of all cluster 2 is (very) small, 34 patients out of 500 are assigned to this cluster (almost 7%). A single patient has much more influence in this cluster than a patient assigned to cluster 3 which consists of 316 patients. Second, we will compare the K-Means clustering and Cobweb clustering by describing the (main) differences here.

Survival (nom)

The probability to survive in the K-Means clusters differ more from each other than the Cobweb clusters do. Patients in cluster 2 have a probability of 95% to pass away, while in cluster 3 patients have a probability of 82% to survive.

AGR (nom)

The AGR is more distributed over the three clusters than in the Cobweb clusters, still each cluster has a main reason for admission.

The other nominal attributes (nom)

Patients with co-morbidities are mainly scattered over clusters 1 and 2. Cluster 3 consists mainly of patients without co-morbidities.

Creat.max (num)

Cluster 2 consists of patient with the highest creatinine level followed by patients in cluster 1 and cluster 3 which consist of patients with the lowest creatinine level (compared with the overall population).

Los.icu (num)

The length of stay is more or less the same in cluster 2 and 3; the difference is only one day in favor of cluster 3. Patients in cluster 1 stay (much) longer than patients in the other two clusters.

Nice.saps.score (num)

The SAPS-score of patients in cluster 2 is (very) high, while cluster 1 the score is a bit higher than the overall population and the score of cluster 3 is a bit lower than the overall population.

[V] See 'Table 4 – cluster_KMeans_s100' and 'Table 7 – cluster_KMeans_s100' in the appendix for detailed information

4. Conclusion and discussion

4.1 Our findings

K-Means and Cobweb clustering produced clusters from our data, and both techniques can be used to predict mortality. However, the two clustering methods gave us two completely different sets of clusters. They produced clusters in which 34% for K-Means and 39,4% for Cobweb of the instances was incorrectly clustered. With this in mind, it seems that K-Means produces better clusters but Cobweb produces clusters in which the instances have been distributed more equally whereas K-Means produces clusters with very many and very few instances.

The models built by the clustering techniques describe the patient population based on several characteristics, in relation to the outcome of their treatment (survival or non-survival). The most important thing we can learn from the Cobweb clusters is that patients who suffer from acute renal insufficiency and chronicle renal insufficiency are more likely to not survive their ICU stay. Many of the other attributes, like co-morbidity, do not seem to make a large difference.

4.2 Recommendation

It appears that the clustering techniques are suitable for predicting outcome of an ICU treatment in elderly patients. These predictions can be used to inform patients and their family about what they can expect from a treatment. Nonetheless, we do not recommend using the results of our research to decide on whether or not a patient should receive treatment or not. Therefore we recommend using a larger dataset and to do more research into the differences produced by the K-Means and Cobweb clustering techniques.

4.3 About the clustering techniques

We have used K-Means and Cobweb clustering techniques to group together several patient instances, where the average mortality in each cluster was used to predict mortality in new patients assigned to a cluster. However, some remarks can be made about our methods, and about our results.

4.3.1 Chosen parameters do not guarantee global minimum in error-rate function

During the clustering process, the parameters for each clustering algorithm (seed for K-Means, and Cutoff and Acuity for Cobweb) were chosen so that the process would produce three clusters, with the lowest amount of incorrectly clustered instances. This produced clusters in K-Means and Cobweb with an error-rate of respectively 34,0% and 39,4%. Yet this result depends to a large extend on these parameters; a slight variation in the parameters produces large variations in the amount of clusters and the amount of incorrectly clustered instances. There is little evidence that the values chosen are the best values possible. Small changes in the values cause the clustering to produce less than or more than three clusters, with higher error-rates, but it cannot be guaranteed that the values only represent a global minimum in the error-rate function. Other choices could lead to better clusters, but better values have not been found.

4.3.2 High error-rates do not provide accurate predictive information

With both types of clusters there are large differences in the survival-rates between clusters and the overall a-priory change of survival. Nevertheless, the error-rates of both types of clusters are very high and seem to make no significant difference with the base chances of survival that a patient has when he is admitted to the ICU. It can be doubted if the result of the clustering process can have any major clinical relevance, because of the low predictive value of the clusters.

4.3.3 Unidentified clinical relevance may influence model validity

One can ask questions as to how, if at all, these clusters can be used in clinical practice. A new patient can be assigned to one of the clusters, which tells something about the possible outcome of the treatment. Several actions can be chosen based on this information, of which termination of treatment because of low expected outcome is the most extreme. Whatever action is chosen, it will change the outcome, thereby making the difference between model and reality larger with every case. Building a model that will be used in clinical situations will have to include consideration about how many cases can be dealt with by the model before it has to be rebuilt. We did not consider this in our study.

4.3.4 Real-life performance of model is unknown

All 500 instances in the dataset were used as training data, and no data was used to test the model for correctly classifying new instances. Due to this omission we cannot tell how well our model will perform in reality, which should be part of any model-building study.

4.3.5 Cobweb clustering variables are not independent

The variables on which the Cobweb clustering was based on (*GSC15, los.icu, urine.24, heartrate.min, creat.max* and *AGR*) are not completely independent of one another. In Cobweb cluster B, many patients suffer from acute renal failure. This cluster has a low survival rate (21,2%) and acute renal failure is associated with impaired urine production and elevated blood levels of creatinine. This way, cluster B has mainly become a cluster for people with acute renal failure, who do not have a good prognosis. It would have been better is there had been no dependency between the variables used in the clustering.

4.3.6 Not all variables used in the clustering are available in the first 24 hours

We used *los.icu* as one of the six variables upon which we had Cobweb create the clusters. This is in conflict with our intention to create a model that would predict mortality based on parameters available within the first 24 hours at the ICU, as *los.icu* represents the total amount of days the patient stayed at the ICU. To determine the influence *los.icu* had on the clustering, we made new clusters without *los.icu* (and included *creat.min* as a new sixth attribute) and compared these clusters with the original clusters.

With los.icu		Without los.icu	
Cluster A	247 (49%)	Cluster A	256 (51%)
Cluster B	165 (33%)	Cluster B	165 (33%)
Cluster C	88 (18%)	Cluster C	79 (16%)

2% (9 patients out of 500) of the patients shifted from cluster C to cluster A. We concluded that this was not large enough a change to redo our calculations. An added advantage of this is that our clustering can also be used to predict the length of stay.

4.3.7 Cobweb vs. K-Means

K-Means has some weaknesses, as the name says it can only work with numeric data from which it is possible to calculate the mean.[3] For some attributes it is not possible to calculate the mean (categorical/nominal data). For the clustering process we used the attribute AGR, this attribute is a nominal attribute; K-Means is unable to cope with this attribute.

For this assignment it was necessary to come up with three clusters, this is an advantage of K-Means, but that is not normally the case for unsupervised learning. So in case of unsupervised learning (you do not know much or anything about the data) it is not preferable to use K-Means.

Cobweb uses probability for assigning objects to clusters.[4] It can create a new class for an object or place it in an existing class. Later on (to certain limits) Cobweb can combine two classes into a single class (merging) or divide a class into several classes (splitting). This is not possible with K-Means and it is sometimes required.

There are some limitations for Cobweb; the assumption that the attributes are independent of each other is often too strong, because correlation may exist. As in our dataset (some attributes are dependent on another attribute) and it is said that Cobweb cannot handle large datasets, because of expensive probability distributions.

To conclude we are in favor of Cobweb, because we do not know exactly what the dependencies are and our dataset is not a very large one and Cobweb can later on split and merge classes to make clusters better. We also use AGR for clustering purposes and K-Means cannot handle nominal attributes so the only option left is Cobweb.

Reference List

(1) de Rooij SE, Abu-Hanna A, Levi M, de JE. Identification of high-risk subgroups in very elderly intensive care unit patients. Crit Care 2007; 11(2):R33.

(2) de Rooij SE, Abu-Hanna A, Levi M, de JE. Factors that predict outcome of intensive care treatment in very elderly patients: a review. Crit Care 2005; 9(4):R307-R314.

(3) Ian H.Witten EF. Implementations: Real Machine Learning Schemes; Clustering. Datamining, practical machine, learning tools and techniques with java, implementations. Morgan Kaufmann; 2000. 210-221.

(4) Jude W.Shawlik TGD. Knowledge Acquisition Via Incremental Conceptual Clustering. Readings in Machine Learning. Morgan Kaufmann; 1990. 267-283.

Appendix

Table 3 - cluster_Cobweb

		Cluster (n=247)			clusterB (n=185)			clusterC (n=88)			Overall (a-priory)
		Count	Row N %	Column N %	Count	Row N %	Column N %	Count	Row N %	Column N %	Row N % of total
died	***0***	***210***	***61,76%***	***85,02%***	***72***	***21,18%***	***43,64%***	***58***	***17,06%***	***65,91%***	***68,00%***
	1	***37***	***23,13%***	***14,98%***	***93***	***58,13%***	***56,36%***	***30***	***18,75%***	***34,09%***	***32,00%***
AGR	**1**	**0**	**0,00%**	**0,00%**	**153**	**100,00%**	**92,73%**	**0**	**0,00%**	**0,00%**	**30,60%**
	2	**0**	**0,00%**	**0,00%**	**9**	**10,34%**	**5,45%**	**78**	**89,66%**	**88,64%**	**17,40%**
	3	**247**	**95,00%**	**100,00%**	**3**	**1,15%**	**1,82%**	**10**	**3,85%**	**11,36%**	**52,00%**
ac.ren.fail	***0***	***240***	***52,75%***	***97,17%***	***135***	***29,67%***	***81,82%***	***80***	***17,58%***	***90,91%***	***91,00%***
	1	***7***	***15,56%***	***2,83%***	***30***	***66,67%***	***18,18%***	***8***	***17,78%***	***9,09%***	***9,00%***
cardio.vasc.insuf	0	155	46,41%	62,75%	116	34,73%	70,30%	63	18,86%	71,59%	66,80%
	1	92	55,42%	37,25%	49	29,52%	29,70%	25	15,06%	28,41%	33,20%
chr.renal.insuf	***0***	***239***	***50,96%***	***96,76%***	***145***	***30,92%***	***87,88%***	***85***	***18,12%***	***96,59%***	***93,80%***
	1	***8***	***25,81%***	***3,24%***	***20***	***64,52%***	***12,12%***	***3***	***9,68%***	***3,41%***	***6,20%***
confirm.infection	0	241	52,97%	97,57%	133	29,23%	80,61%	81	17,80%	92,05%	91,00%
	1	6	13,33%	2,43%	32	71,11%	19,39%	7	15,56%	7,95%	9,00%
cpr	0	246	52,23%	99,60%	139	29,51%	84,24%	86	18,26%	97,73%	94,20%
	1	1	3,45%	0,40%	26	89,66%	15,76%	2	6,90%	2,27%	5,80%
cva	0	239	49,59%	96,76%	160	33,20%	96,97%	83	17,22%	94,32%	96,40%
	1	8	44,44%	3,24%	5	27,78%	3,03%	5	27,78%	5,68%	3,60%
dysrhytmia	0	219	51,77%	88,66%	128	30,26%	77,58%	76	17,97%	86,36%	84,60%
	1	28	36,36%	11,34%	37	48,05%	22,42%	12	15,58%	13,64%	15,40%
gastro.bleed	0	244	51,15%	98,79%	149	31,24%	90,30%	84	17,61%	95,45%	95,40%
	1	3	13,04%	1,21%	16	69,57%	9,70%	4	17,39%	4,55%	4,60%
hem.malign	0	247	49,50%	100,00%	164	32,87%	99,39%	88	17,64%	100,00%	99,80%
	1	0	0,00%	0,00%	1	100,00%	0,61%	0	0,00%	0,00%	0,20%
imm.insuf	0	244	49,39%	98,79%	162	32,79%	98,18%	88	17,81%	100,00%	98,80%
	1	3	50,00%	1,21%	3	50,00%	1,82%	0	0,00%	0,00%	1,20%
intracran.mass	0	245	49,60%	99,19%	164	33,20%	99,39%	85	17,21%	96,59%	98,80%
	1	2	33,33%	0,81%	1	16,67%	0,61%	3	50,00%	3,41%	1,20%
mech.ventil.24	0	59	46,09%	23,89%	54	42,19%	32,73%	15	11,72%	17,05%	25,60%
	1	188	50,54%	76,11%	111	29,84%	67,27%	73	19,62%	82,95%	74,40%
neoplasm	0	245	49,70%	99,19%	162	32,86%	98,18%	86	17,44%	97,73%	98,60%
	1	2	28,57%	0,81%	3	42,86%	1,82%	2	28,57%	2,27%	1,40%
resp.insuf	0	239	52,64%	96,76%	134	29,52%	81,21%	81	17,84%	92,05%	90,80%
	1	8	17,39%	3,24%	31	67,39%	18,79%	7	15,22%	7,95%	9,20%
vasdrug	0	72	48,98%	29,15%	50	34,01%	30,30%	25	17,01%	28,41%	29,40%
	1	175	49,58%	70,85%	115	32,58%	69,70%	63	17,85%	71,59%	70,60%

Legenda **Incorrectly clustered instances :** 197 39,40%

colored columns are percentages characteristic for the cluster, bold variable is one of the six selected variables for clustering, bold italic variable is a variable of interest, e.g. died, saps-score

Table 4 - cluster_KMeans_s100

		cluster1 (n=150)			cluster2 (n=34)			cluster3 (n=316)			Overall (a-priory)
		Count	Row N %	Column N %	Count	Row N %	Column N %	Count	Row N %	Column N %	Row N % of total
died	***0***	***79***	***23,24%***	***52,67%***	***2***	***0,59%***	***5,88%***	***259***	***76,18%***	***81,96%***	***68,00%***
	1	***71***	***44,38%***	***47,33%***	***32***	***20,00%***	***94,12%***	***57***	***35,63%***	***18,04%***	***32,00%***
AGR	**1**	**130**	**84,97%**	**86,67%**	**23**	**15,03%**	**67,65%**	**0**	**0,00%**	**0,00%**	**30,60%**
	2	**20**	**22,99%**	**13,33%**	**9**	**10,34%**	**26,47%**	**58**	**66,67%**	**18,35%**	**17,40%**
	3	**0**	**0,00%**	**0,00%**	**2**	**0,77%**	**5,88%**	**258**	**99,23%**	**81,65%**	**52,00%**
ac.ren.fail	***0***	***127***	***27,91%***	***84,67%***	***21***	***4,62%***	***61,76%***	***307***	***67,47%***	***97,15%***	***91,00%***
	1	***23***	***51,11%***	***15,33%***	***13***	***28,89%***	***38,24%***	***9***	***20,00%***	***2,85%***	***9,00%***
cardio.vasc.insuf	0	108	32,34%	72,00%	21	6,29%	61,76%	205	61,38%	64,87%	66,80%
	1	42	25,30%	28,00%	13	7,83%	38,24%	111	66,87%	35,13%	33,20%
chr.renal.insuf	***0***	***134***	***28,57%***	***89,33%***	***29***	***6,18%***	***85,29%***	***306***	***65,25%***	***96,84%***	***93,80%***
	1	***16***	***51,61%***	***10,67%***	***5***	***16,13%***	***14,71%***	***10***	***32,26%***	***3,16%***	***6,20%***
confirm.infection	0	118	25,93%	78,67%	30	6,59%	88,24%	307	67,47%	97,15%	91,00%
	1	32	71,11%	21,33%	4	8,89%	11,76%	9	20,00%	2,85%	9,00%
cpr	0	134	28,45%	89,33%	23	4,88%	67,65%	314	66,67%	99,37%	94,20%
	1	16	55,17%	10,67%	11	37,93%	32,35%	2	6,90%	0,63%	5,80%
cva	0	147	30,50%	98,00%	32	6,64%	94,12%	303	62,86%	95,89%	96,40%
	1	3	16,67%	2,00%	2	11,11%	5,88%	13	72,22%	4,11%	3,60%
dysrhytmia	0	120	28,37%	80,00%	22	5,20%	64,71%	281	66,43%	88,92%	84,60%
	1	30	38,96%	20,00%	12	15,58%	35,29%	35	45,45%	11,08%	15,40%
gastro.bleed	0	135	28,30%	90,00%	31	6,50%	91,18%	311	65,20%	98,42%	95,40%
	1	15	65,22%	10,00%	3	13,04%	8,82%	5	21,74%	1,58%	4,60%
hem.malign	0	149	29,86%	99,33%	34	6,81%	100,00%	316	63,33%	100,00%	99,80%
	1	1	100,00%	0,67%	0	0,00%	0,00%	0	0,00%	0,00%	0,20%
imm.insuf	0	147	29,76%	98,00%	34	6,88%	100,00%	313	63,36%	99,05%	98,80%
	1	3	50,00%	2,00%	0	0,00%	0,00%	3	50,00%	0,95%	1,20%
intracran.mass	0	149	30,16%	99,33%	34	6,88%	100,00%	311	62,96%	98,42%	98,80%
	1	1	16,67%	0,67%	0	0,00%	0,00%	5	83,33%	1,58%	1,20%
mech.ventil.24	0	53	41,41%	35,33%	3	2,34%	8,82%	72	56,25%	22,78%	25,60%
	1	97	26,08%	64,67%	31	8,33%	91,18%	244	65,59%	77,22%	74,40%
neoplasm	0	146	29,61%	97,33%	34	6,90%	100,00%	313	63,49%	99,05%	98,60%
	1	4	57,14%	2,67%	0	0,00%	0,00%	3	42,86%	0,95%	1,40%
resp.insuf	0	122	26,87%	81,33%	28	6,17%	82,35%	304	66,96%	96,20%	90,80%
	1	28	60,87%	18,67%	6	13,04%	17,65%	12	26,09%	3,80%	9,20%
vasdrug	0	51	34,69%	34,00%	3	2,04%	8,82%	93	63,27%	29,43%	29,40%
	1	99	28,05%	66,00%	31	8,78%	91,18%	223	63,17%	70,57%	70,60%

Incorrectly clustered instances : 170 34,00%

		Table 5 - cluster_KMeans_s186									
		cluster1 (n=260)			cluster2 (n=87)			cluster3 (n=153)			Overall (a-priory)
		Count	Row N %	Column N %	Count	Row N %	Column N %	Count	Row N %	Column N %	Row N % of total
died	***0***	***214***	***62,94%***	***82,31%***	***56***	***16,47%***	***64,37%***	***70***	***20,59%***	***45,75%***	***68,00%***
	1	***46***	***28,75%***	***17,69%***	***31***	***19,38%***	***35,63%***	***83***	***51,88%***	***54,25%***	***32,00%***
AGR	**1**	**0**	**0,00%**	**0,00%**	**0**	**0,00%**	**0,00%**	**153**	**100,00%**	**100,00%**	**30,60%**
	2	**0**	**0,00%**	**0,00%**	**87**	**100,00%**	**100,00%**	**0**	**0,00%**	**0,00%**	**17,40%**
	3	**260**	**100,00%**	**100,00%**	**0**	**0,00%**	**0,00%**	**0**	**0,00%**	**0,00%**	**52,00%**
ac.ren.fail	***0***	***251***	***55,16%***	***96,54%***	***79***	***17,36%***	***90,80%***	***125***	***27,47%***	***81,70%***	***91,00%***
	1	***9***	***20,00%***	***3,46%***	***8***	***17,78%***	***9,20%***	***28***	***62,22%***	***18,30%***	***9,00%***
cardio.vasc.insuf	0	163	48,80%	62,69%	65	19,46%	74,71%	106	31,74%	69,28%	66,80%
	1	97	58,43%	37,31%	22	13,25%	25,29%	47	28,31%	30,72%	33,20%
chr.renal.insuf	***0***	***250***	***53,30%***	***96,15%***	***86***	***18,34%***	***98,85%***	***133***	***28,36%***	***86,93%***	***93,80%***
	1	***10***	***32,26%***	***3,85%***	***1***	***3,23%***	***1,15%***	***20***	***64,52%***	***13,07%***	***6,20%***
confirm.infection	0	254	55,82%	97,69%	80	17,58%	91,95%	121	26,59%	79,08%	91,00%
	1	6	13,33%	2,31%	7	15,56%	8,05%	32	71,11%	20,92%	9,00%
cpr	0	259	54,99%	99,62%	84	17,83%	96,55%	128	27,18%	83,66%	94,20%
	1	1	3,45%	0,38%	3	10,34%	3,45%	25	86,21%	16,34%	5,80%
cva	0	252	52,28%	96,92%	82	17,01%	94,25%	148	30,71%	96,73%	96,40%
	1	8	44,44%	3,08%	5	27,78%	5,75%	5	27,78%	3,27%	3,60%
dysrhytmia	0	230	54,37%	88,46%	76	17,97%	87,36%	117	27,66%	76,47%	84,60%
	1	30	38,96%	11,54%	11	14,29%	12,64%	36	46,75%	23,53%	15,40%
gastro.bleed	0	257	53,88%	98,85%	83	17,40%	95,40%	137	28,72%	89,54%	95,40%
	1	3	13,04%	1,15%	4	17,39%	4,60%	16	69,57%	10,46%	4,60%
hem.malign	0	260	52,10%	100,00%	87	17,43%	100,00%	152	30,46%	99,35%	99,80%
	1	0	0,00%	0,00%	0	0,00%	0,00%	1	100,00%	0,65%	0,20%
imm.insuf	0	257	52,02%	98,85%	87	17,61%	100,00%	150	30,36%	98,04%	98,80%
	1	3	50,00%	1,15%	0	0,00%	0,00%	3	50,00%	1,96%	1,20%
intracran.mass	0	258	52,23%	99,23%	84	17,00%	96,55%	152	30,77%	99,35%	98,80%
	1	2	33,33%	0,77%	3	50,00%	3,45%	1	16,67%	0,65%	1,20%
mech.ventil.24	0	62	48,44%	23,85%	14	10,94%	16,09%	52	40,63%	33,99%	25,60%
	1	198	53,23%	76,15%	73	19,62%	83,91%	101	27,15%	66,01%	74,40%
neoplasm	0	258	52,33%	99,23%	85	17,24%	97,70%	150	30,43%	98,04%	98,60%
	1	2	28,57%	0,77%	2	28,57%	2,30%	3	42,86%	1,96%	1,40%
resp.insuf	0	251	55,29%	96,54%	80	17,62%	91,95%	123	27,09%	80,39%	90,80%
	1	9	19,57%	3,46%	7	15,22%	8,05%	30	65,22%	19,61%	9,20%
vasdrug	0	74	50,34%	28,46%	24	16,33%	27,59%	49	33,33%	32,03%	29,40%
	1	186	52,69%	71,54%	63	17,85%	72,41%	104	29,46%	67,97%	70,60%

Incorrectly clustered instances : 203 40,60%

	Table 6 - cluster_Cobweb											
	clusterA (n=247)			clusterB (n=165)			clusterC (n=88)			allClusters (n=500)		
	Mean	Median	Std. Deviation	Mean	Median	Std. Deviation	Mean	Median	Std. Deviation	Mean	Median	Std. Deviation
age	82,55	82,00	2,875	83,83	83,00	3,943	84,08	83,00	3,471	83,24	82,00	3,430
album.max	27,68	27,00	4,921	29,61	30,00	5,772	28,87	27,00	5,841	28,70	28,00	5,511
album.min	25,29	25,00	5,217	25,65	26,00	6,382	24,98	25,00	6,590	25,37	25,00	5,948
bicarb.max	23,03	22,90	2,811	23,39	23,30	5,132	23,42	23,75	3,769	23,21	23,00	3,876
bicarb.min	20,56	20,40	2,820	19,51	19,55	5,657	19,57	19,90	4,118	20,04	20,10	4,205
bili	26,23	16,50	40,500	16,99	12,00	13,399	22,58	18,50	20,259	20,64	14,00	25,465
creat.max	**109,30**	**99,00**	**49,228**	**188,71**	**146,00**	**146,804**	**122,49**	**102,00**	**54,037**	**137,66**	**107,50**	**100,315**
creat.min	96,56	90,00	41,105	161,27	126,00	114,304	107,62	95,00	46,934	119,66	98,00	79,571
GCS15	**14,92**	**15,00**	**0,508**	**12,47**	**15,00**	**4,408**	**14,85**	**15,00**	**0,578**	**14,10**	**15,00**	**2,809**
hb.max	6,76	6,65	0,824	7,06	6,80	1,047	7,08	7,00	0,828	6,92	6,75	0,914
hb.min	5,62	5,80	0,838	6,05	5,75	1,024	5,79	5,45	0,768	5,80	5,75	0,914
heartrate.max	102,75	100,00	19,929	123,39	119,50	32,482	107,43	101,00	23,434	110,41	105,00	26,928
heartrate.min	**72,39**	**72,00**	**16,374**	**66,59**	**70,00**	**30,058**	**73,75**	**74,00**	**18,987**	**70,71**	**72,00**	**22,383**
ht.max	12,680	0,330	32,912	12,722	0,350	32,970	14,578	0,350	35,079	13,021	0,340	33,252
ht.min	0,120	0,260	0,420	0,152	0,285	0,412	0,097	0,260	0,449	0,127	0,270	0,422
los.icu	**1,64**	**0,96**	**1,516**	**4,45**	**2,05**	**7,864**	**5,46**	**1,00**	**7,896**	**3,23**	**1,01**	**5,901**
meanbl.max	95,03	92,00	17,887	92,73	93,00	22,235	95,24	92,50	18,942	94,31	93,00	19,578
meanbl.min	60,60	60,00	12,539	50,99	55,00	22,843	64,00	64,00	16,400	58,04	60,00	17,924
nice.saps.prob	0,147739	0,116838	0,111	0,513069	0,507019	0,326	0,321288	0,266086	0,208	0,290482	0,181019	0,268
nice.saps.score	***31,74***	***31,00***	***8,217***	***53,87***	***51,00***	***22,104***	***42,74***	***41,00***	***11,292***	***40,98***	***36,00***	***17,715***
pao2.24	88,97	85,00	32,711	82,40	76,00	34,178	89,87	85,00	28,860	86,93	82,00	32,664
ph.min	7,37	7,38	0,080	7,33	7,34	0,108	7,38	7,39	0,087	7,36	7,37	0,094
potas.max	4,45	4,40	0,406	4,59	4,45	0,793	4,48	4,45	0,616	4,50	4,40	0,597
potas.min	3,76	3,70	0,428	3,78	3,70	0,742	3,66	3,60	0,584	3,75	3,70	0,576
ptt	18,61	17,10	8,574	24,03	17,00	17,267	19,34	18,00	5,717	20,66	17,10	12,246
resprate.max	19,76	19,00	5,776	24,95	23,00	8,977	19,62	18,00	5,547	21,46	20,00	7,377
resprate.min	11,47	11,00	3,757	12,99	13,00	5,341	12,32	12,00	3,829	12,12	12,00	4,402
sodium.max	141,08	141,00	4,259	140,55	141,00	5,942	140,11	140,00	5,413	140,74	141,00	5,072
sodium.min	138,30	138,00	3,891	137,28	138,00	5,337	136,11	136,00	4,620	137,59	138,00	4,604
syst.max	148,91	146,00	27,637	140,57	140,00	34,113	144,77	140,00	26,032	145,45	144,00	29,814
syst.min	92,27	90,00	21,465	77,79	84,00	34,481	96,53	95,00	27,535	88,31	90,00	28,354
temp.max	37,35	37,30	0,619	37,44	37,40	1,068	37,33	37,20	0,665	37,38	37,40	0,799
temp.min	35,62	35,60	0,952	35,92	36,20	1,667	35,90	35,90	1,012	35,77	35,80	1,246
throm.min	150,45	142,00	69,952	183,52	173,50	93,863	148,77	141,00	75,200	160,81	150,00	80,619
urea	11,00	8,60	10,618	19,28	16,80	10,836	13,41	10,60	14,033	14,44	10,80	11,958
urine.24	**3,09**	**3,00**	**1,537**	**2,34**	**2,00**	**1,735**	**3,23**	**2,80**	**3,083**	**2,87**	**2,70**	**1,991**
wbc.max	12,98	11,20	9,963	15,73	13,40	8,093	12,63	12,25	6,125	13,92	12,30	8,814
wbc.min	11,66	10,35	9,483	12,35	11,00	6,398	9,73	9,00	5,371	11,57	10,20	7,903

Legenda

colored columns are the carateristic values for the cluster, bold variable is one of the six selected variables for clustering, bold italic variable is a variable of interest, e.g. died, saps-score

	Table 7 - cluster_KMeans_s100											
	cluster1 (n=150)			cluster2 (n=34)			cluster3 (n=316)			allClusters (n=500)		
	Mean	Median	Std. Deviation	Mean	Median	Std. Deviation	Mean	Median	Std. Deviation	Mean	Median	Std. Deviation
age	83,67	83,00	3,757	84,62	83,00	4,465	82,89	82,00	3,074	83,24	82,00	3,430
album.max	29,78	30,00	5,474	27,64	30,00	6,422	28,10	27,00	5,285	28,70	28,00	5,511
album.min	25,61	25,00	6,662	23,20	25,00	5,694	25,52	25,00	5,439	25,37	25,00	5,948
bicarb.max	23,53	23,30	4,959	21,96	22,00	4,948	23,22	23,00	3,028	23,21	23,00	3,876
bicarb.min	19,65	19,80	5,557	17,14	17,70	5,319	20,61	20,50	2,933	20,04	20,10	4,205
bili	17,52	12,00	13,514	17,50	14,00	12,665	25,21	16,50	36,330	20,64	14,00	25,465
creat.max	**178,75**	**135,50**	**142,230**	**224,56**	**204,00**	**131,333**	**109,70**	**98,00**	**48,337**	**137,66**	**107,50**	**100,315**
creat.min	151,91	120,00	108,614	184,50	171,00	114,320	97,32	90,00	40,320	119,66	98,00	79,571
GCS15	**14,60**	**15,00**	**1,123**	**4,44**	**3,00**	**1,988**	**14,90**	**15,00**	**0,809**	**14,10**	**15,00**	**2,809**
hb.max	7,084	6,800	0,920	6,975	6,350	1,303	6,820	6,700	0,820	6,92	6,75	0,914
hb.min	5,98	5,70	0,974	6,08	5,85	1,057	5,65	5,80	0,830	5,80	5,75	0,914
heartrate.max	120,45	115,00	32,492	124,53	121,00	30,865	103,77	100,00	20,631	110,41	105,00	26,928
heartrate.min	**71,63**	**70,50**	**24,044**	**50,41**	**57,50**	**39,183**	**72,89**	**72,00**	**16,706**	**70,71**	**72,00**	**22,383**
ht.max	13,23	0,35	33,543	9,13	0,36	28,689	13,36	0,33	33,651	13,021	0,340	33,252
ht.min	0,14	0,27	0,418	0,17	0,29	0,372	0,11	0,26	0,430	0,127	0,270	0,422
los.icu	**5,76**	**2,38**	**8,913**	**2,92**	**1,64**	**3,821**	**2,07**	**0,96**	**3,429**	**3,23**	**1,01**	**5,901**
meanbl.max	93,89	93,00	19,795	86,59	89,00	28,440	95,45	93,00	17,983	94,31	93,00	19,578
meanbl.min	55,44	58,00	19,679	39,15	45,00	27,084	61,65	60,00	13,455	58,04	60,00	17,924
nice.saps.prob	0,415477	0,391925	0,267	0,922861	0,969298	0,094	0,173366	0,128047	0,133	0,290482	0,181019	0,268
nice.saps.score	***46,75***	***46,50***	***15,001***	***85,26***	***85,50***	***15,932***	***33,47***	***32,00***	***8,947***	***40,98***	***36,00***	***17,715***
pao2.24	81,50	77,00	24,698	85,24	74,00	53,306	89,75	85,50	32,328	86,93	82,00	32,664
ph.min	7,34	7,34	0,100	7,29	7,30	0,128	7,38	7,38	0,079	7,36	7,37	0,094
potas.max	4,57	4,40	0,755	4,84	4,70	0,920	4,43	4,40	0,431	4,50	4,40	0,597
potas.min	3,75	3,70	0,694	3,92	3,80	0,922	3,73	3,70	0,448	3,75	3,70	0,576
ptt	23,57	17,00	15,838	25,13	17,35	18,571	18,46	17,10	7,656	20,66	17,10	12,246
resprate.max	25,36	24,00	8,595	20,73	17,00	8,715	19,68	19,00	5,687	21,46	20,00	7,377
resprate.min	13,30	13,00	4,920	11,76	12,00	5,820	11,60	12,00	3,803	12,12	12,00	4,402
sodium.max	140,03	141,00	5,792	142,73	142,50	5,901	140,86	141,00	4,528	140,74	141,00	5,072
sodium.min	136,57	137,00	5,173	139,19	138,00	5,023	137,89	138,00	4,161	137,59	138,00	4,604
syst.max	143,05	141,00	29,410	131,09	140,00	43,197	148,45	145,00	27,270	145,45	144,00	29,814
syst.min	85,76	90,00	29,413	55,59	70,00	37,510	93,77	91,00	23,068	88,31	90,00	28,354
temp.max	37,47	37,45	0,916	37,26	37,00	1,425	37,35	37,30	0,615	37,38	37,40	0,799
temp.min	36,01	36,20	1,619	35,30	35,40	1,531	35,71	35,70	0,950	35,77	35,80	1,246
throm.min	182,03	170,00	99,696	161,83	152,50	67,263	151,01	143,00	69,994	160,81	150,00	80,619
urea	18,40	16,05	10,624	20,25	17,10	9,903	11,45	8,70	12,029	14,44	10,80	11,958
urine.24	**2,41**	**2,50**	**1,373**	**1,82**	**1,10**	**2,432**	**3,23**	**3,00**	**2,083**	**2,87**	**2,70**	**1,991**
wbc.max	15,81	13,80	8,095	14,67	13,10	8,039	12,78	11,30	9,130	13,92	12,30	8,814
wbc.min	12,08	10,30	6,868	12,17	11,30	5,472	11,22	10,00	8,648	11,57	10,20	7,903

Table 8 - cluster_KMeans_s186

	cluster1 (n=260)			cluster2 (n=87)			cluster3 (n=153)			allClusters (n=500)		
	Mean	Median	Standard Deviation	Mean	Median	Standard Deviation	Mean	Median	Standard Deviation	Mean	Median	Std. Deviation
age	82,55	82,00	2,853	84,44	83,00	3,943	83,74	83,00	3,748	83,24	82,00	3,430
album.max	27,69	27,00	4,831	28,54	26,50	6,481	29,88	31,00	5,490	28,70	28,00	5,511
album.min	25,06	25,00	5,254	24,87	25,00	6,871	26,03	26,00	6,205	25,37	25,00	5,948
bicarb.max	23,02	22,80	2,815	23,65	24,00	3,950	23,28	23,30	5,178	23,21	23,00	3,876
bicarb.min	20,42	20,20	2,874	19,96	20,15	4,198	19,46	19,70	5,757	20,04	20,10	4,205
bili	25,62	16,50	39,630	23,68	20,00	20,219	16,62	12,00	13,212	20,64	14,00	25,465
creat.max	**110,86**	**99,00**	**49,976**	**123,16**	**102,00**	**56,186**	**190,79**	**146,00**	**150,287**	**137,66**	**107,50**	**100,315**
creat.min	97,58	90,00	41,732	109,22	94,00	50,620	163,22	127,00	116,581	119,66	98,00	79,571
GCS15	**14,79**	**15,00**	**1,369**	**13,72**	**15,00**	**3,402**	**13,14**	**15,00**	**3,788**	**14,10**	**15,00**	**2,809**
hb.max	6,75	6,60	0,787	7,39	7,65	1,041	6,98	6,80	0,991	6,92	6,75	0,914
hb.min	5,61	5,75	0,815	6,09	5,90	0,973	6,00	5,70	0,995	5,80	5,75	0,914
heartrate.max	103,00	100,00	19,845	107,28	104,50	23,651	124,84	120,00	32,984	110,41	105,00	26,928
heartrate.min	**72,37**	**72,00**	**17,163**	**70,07**	**70,00**	**21,621**	**68,24**	**71,00**	**29,495**	**70,71**	**72,00**	**22,383**
ht.max	12,881	0,330	33,137	14,752	0,350	35,255	12,301	0,350	32,490	13,021	0,340	33,252
ht.min	0,117	0,260	0,422	0,098	0,260	0,453	0,159	0,285	0,404	0,127	0,270	0,422
los.icu	**2,23**	**0,97**	**3,670**	**4,01**	**0,94**	**6,394**	**4,49**	**2,10**	**8,042**	**3,23**	**1,01**	**5,901**
meanbl.max	94,51	90,00	17,806	96,32	95,00	19,378	92,82	93,00	22,469	94,31	93,00	19,578
meanbl.min	60,24	60,00	12,629	63,69	65,00	18,170	51,00	55,00	22,902	58,04	60,00	17,924
nice.saps.prob	0,161101	0,128047	0,137	0,378596	0,305597	0,266	0,484706	0,449105	0,319	0,290482	0,181019	0,268
nice.saps.score	***32,54***	***32,00***	***9,420***	***46,91***	***43,00***	***17,069***	***51,94***	***49,00***	***21,244***	***40,98***	***36,00***	***17,715***
pao2.24	88,17	83,50	32,590	93,76	86,00	39,682	80,96	76,00	27,170	86,93	82,00	32,664
ph.min	7,37	7,38	0,080	7,38	7,39	0,095	7,33	7,34	0,108	7,36	7,37	0,094
potas.max	4,46	4,40	0,413	4,48	4,50	0,605	4,58	4,40	0,814	4,50	4,40	0,597
potas.min	3,77	3,70	0,437	3,63	3,60	0,594	3,78	3,70	0,743	3,75	3,70	0,576
ptt	18,70	17,10	8,388	19,03	18,00	5,696	24,45	17,00	17,734	20,66	17,10	12,246
resprate.max	19,71	19,00	5,872	19,43	18,00	5,071	25,59	24,00	8,987	21,46	20,00	7,377
resprate.min	11,51	12,00	3,684	12,36	12,00	4,008	13,02	13,00	5,474	12,12	12,00	4,402
sodium.max	141,28	141,00	4,483	140,40	140,00	5,509	140,02	140,00	5,661	140,74	141,00	5,072
sodium.min	138,48	138,00	4,005	136,28	136,00	4,885	136,80	137,00	5,079	137,59	138,00	4,604
syst.max	147,98	145,00	27,560	145,34	150,00	27,419	141,10	140,00	34,335	145,45	144,00	29,814
syst.min	91,43	90,00	21,632	95,49	95,00	30,334	78,64	84,00	34,560	88,31	90,00	28,354
temp.max	37,33	37,30	0,615	37,32	37,35	0,815	37,50	37,45	1,035	37,38	37,40	0,799
temp.min	35,63	35,60	0,935	35,77	36,00	1,224	36,01	36,20	1,642	35,77	35,80	1,246
throm.min	148,23	140,00	69,556	152,60	147,00	73,980	187,66	180,00	95,268	160,81	150,00	80,619
urea	11,15	8,70	10,472	14,09	10,80	14,452	19,34	17,00	10,895	14,44	10,80	11,958
urine.24	**3,10**	**3,00**	**1,977**	**3,07**	**2,80**	**2,419**	**2,36**	**2,10**	**1,622**	**2,87**	**2,70**	**1,991**
wbc.max	12,94	11,20	9,855	12,36	11,75	5,921	16,12	14,00	8,164	13,92	12,30	8,814
wbc.min	11,58	10,35	9,320	9,62	9,00	5,325	12,59	11,00	6,495	11,57	10,20	7,903

Table 9 - Baseline Characteristics			
Variable	**Survivors (n = 340)**	**Non-survivors (n = 160)**	**Missing (%)**
Male sex - n (%)	160 (47,1)	69 (43,1)	
Age - year (median)	82,0	83,0	
Re-admission - n (%)	22 (6,5)	15 (9,4)	
Comorbidity			
ac.ren.fail - n (%)	14 (4,1)	31 (19,4)	
dysrhythmia - n (%)	45 (13,2)	32 (20,0)	
cva - n (%)	9 (2,6)	9 (5,6)	
gastro.bleed - n (%)	16 (4,7)	7 (4,4)	
intracran.mass - n (%)	5 (1,5)	1 (0,6)	
chr.renal.insuf - n (%)	18 (5,3)	13 (8,1)	
chron.dialysis - n (%)	0 (0)	0 (0)	
neoplasm - n (%)	4 (1,2)	3 (1,9)	
aids - n (%)	0 (0)	0 (0)	
hem.malign - n (%)	1 (0,3)	0 (0)	
cirrhosis - n (%)	0 (0)	0 (0)	
cardio.vasc.insuf - n (%)	125 (36,8)	41 (25,6)	
resp.insuf - n (%)	25 (7,4)	21 (13,1)	
imm.insuf - n (%)	4 (1,2)	2 (1,3)	
confirm.infection - n (%)	23 (6,8)	22 (13,8)	
Life support			
cpr - n (%)	8 (2,4)	21 (13,1)	
mech.ventil.24 - n (%)	252 (74,1)	120 (75,0)	
Aggregated Scores			
saps.score - mean (sd)	34,4 (12,8)	45,3 (22,5)	239 (47,8)
GCS15 - (median)	15,0	15,0	
Heartrate			
Minimum - mean (sd)	71,3 (17,1)	69,0 (31,3)	77 (15,4)
Maximum - mean (sd)	105,9 (22,7)	120,6 (32,6)	77 (15,4)
Respiratory rate			
Minimum - mean (sd)	11,3 (3,8)	13,0 (5,5)	82 (16,4)
Maximum - mean (sd)	20,8 (6,4)	23,0 (9,0)	81 (16,2)
Blood pressure			
syst.min - mean (sd)	92,5 (22,6)	78,7 (36,8)	83 (16,6)
syst.max - mean (sd)	147,8 (26,6)	140,0 (35,9)	84 (16,8)
meanbl.min - mean (sd)	60,6 (14,2)	51,9 (23,4)	77 (15,4)
meanbl.max - mean (sd)	95,5 (17,8)	91,7 (23,1)	79 (15,8)

Temperature				
	temp.min - mean (sd)	35,7 (1,2)	35,8 (1,3)	89 (17,8)
	temp.max - mean (sd)	37,4 (0,6)	37,4 (1,1)	90 (18,0)
Various				
	ptt - mean (sd)	19,4 (10,8)	23,9 (15,0)	182 (36,4)
	urine.24 - mean (sd)	3,3 (1,9)	2,0 (1,8)	94 (18,8)
	vasdrug - n (%)	242 (48,4)	111 (22,2)	201 (40,2)
	paO2 - mean (sd)	85,0 (27,8)	81,0 (31,6)	94 (18,8)
	los - mean (sd)	2,7 (4,6)	4,4 (7,9)	2 (0,4)
Blood values				
	ph.min - mean (sd)	7,4 (0,07)	7,3 (0.1)	100 (20)
	wbc.min - mean (sd)	11,3 (8,4)	12,3 (6,7)	105 (21)
	wbc.max - mean (sd)	13,4 (9,1)	15,0 (8,1)	135 (27)
	creat.min - mean (sd)	106 (56,7)	153,7 (111,9)	78 (15,6)
	creat.max - mean (sd)	121,1 (83,9)	178,7 (123,8)	92 (15,6)
	potas.min - mean (sd)	3,7 (0,4)	3,8 (0,7)	75 (13,0)
	potas.max - mean (sd)	4,4 (0,5)	4,6 (0,8)	66 (13,2)
	sodium.min - mean (sd)	137,6 (18,3)	137,5 (5,3)	71 (14,2)
	sodium.max - mean (sd)	140,7 (21,9)	140,7 (35,5)	72 (14,4)
	bicarb.min - mean (sd)	20,8 (3,6)	18,3 (5,0)	112 (22,4)
	bicarb.max - mean (sd)	23,6 (3,5)	22,1 (4,5)	113 (22,6)
	urea - mean (sd)	13,5 (12,8)	16,6 (9,4)	142 (28,4)
	bili - mean (sd)	22,1 (31,7)	18,7 (13,2)	362 (72,4)
	ht.min - mean (sd)	0,17 (0,37)	0,04 (0,51)	13 (2,6)
	ht.max - mean (sd)	9,9 (29,5)	19,6 (39,5)	13 (2,6)
	hb.min - mean (sd)	5,8 (0,9)	5,9 (0,9)	396 (79,2)
	hb.max - mean (sd)	6,9 (0,9)	7,0 (1,0)	396 (79,2)
	album.min - mean (sd)	25,9 (31,7)	25,4 (6,4)	177 (35,4)
	album.max - mean (sd)	28,9 (5,0)	28,4 (40,0)	216 (43,2)

Table 2. *The means from the 59 variables for patients, who died in the hospital (non-survivors) and the ones who didn't (survivors). The 6 attributes, which we have chosen are in bold and cursive.*