

Predicting Pre-triage Waiting Time in a Maternity Emergency Room through Data Mining

Sónia Pereira¹, Filipe Portela^{1,2}, Manuel F. Santos¹, José Machado¹, António Abelha¹

¹Algoritmi Centre, University of Minho, Portugal

²ESEIG, Porto Polytechnique, Porto, Portugal

b7004@dps.uminho.pt; {cfp, mfs}@dsi.uminho.pt, {jmac, abelha}@di.uminho.pt

Abstract. An unsuitable patient flow as well as prolonged waiting lists in the emergency room of a maternity unit, regarding gynecology and obstetrics care, can affect the mother and child's health, leading to adverse events and consequences regarding their safety and satisfaction. Predicting the patients' waiting time in the emergency room is a means to avoid this problem. This study aims to predict the pre-triage waiting time in the emergency care of gynecology and obstetrics of Centro Materno Infantil do Norte (CMIN), the maternal and perinatal care unit of Centro Hospitalar of Oporto, situated in the north of Portugal. Data mining techniques were induced using information collected from the information systems and technologies available in CMIN. The models developed presented good results reaching accuracy and specificity values of approximately 74% and 94%, respectively. Additionally, the number of patients and triage professionals working in the emergency room, as well as some temporal variables were identified as direct enhancers to the pre-triage waiting time. The implementation of the attained knowledge in the decision support system and business intelligence platform, deployed in CMIN, leads to the optimization of the patient flow through the emergency room and improving the quality of services.

Keywords: Data mining, Classification algorithms, Gynecology and Obstetrics Care, Maternity Care, Emergency Room, Triage system, Interoperability, IDSS.

1 Introduction

Crowding and prolonged waiting times in the emergency department (ED) of a hospital unit has been considered a major problem. This problem can possibly lead to adverse events and consequences towards the patient safety as well as the quality of services [1, 2]. Addressing these concerns and therefore, optimizing patient flow through the ED can be enhanced by providing improved services such as adequate staffing, destination management and hospital bed access [2, 3]. A new approach to increase patient satisfaction in the emergency room is to estimate the waiting time [4]. The prediction of waiting times helps staff to prioritize patients and operations, as well as avoiding patients leaving the ED unanswered [5, 6]. Furthermore, it may improve morale and reduce the patient tendency to seek further opinions, as well as minimize the complaints and litigation [7]. Regarding maternity services, delivering the best emergency care is fundamental since it has the potential to affect the mother and child's health at a crucial

time. These good practices can be achieved by taking advantage of the information technologies currently available in the healthcare sector, capable of providing complete and reliable information aiding clinical and administrative decisions [8].

The emergency care of gynecology and obstetrics (GO) of Centro Materno Infantil do Norte (CMIN), the maternal and perinatal care unit of Centro Hospitalar of Oporto (CHP), owns a pre-triage system for sorting patients in emergency (URG) and consultation (ARGO) classes. The system was developed in CHP and it is a part of the Intelligent Decision Support System (IDSS) implemented in CMIN [9]. In order to improve the ED quality services by reducing crowding and upgrading the resource management, the current study induces data mining (DM) techniques using real data available in CMIN to predict the patients' waiting time at pre-triage - the time elapsed since the patient arrives to ED until she is called to go to the triage room. The required information is provided by the systems used in CMIN to collect medical data. Nursing System (SAPE) is a decision support system that allows producing and storing clinical information. In addition, the Electronic Health Record (EHR) records the patient data and their admission form. By combining the information collect by these systems in several scenarios and inducing data mining (DM) algorithms, the case study managed to achieve useful knowledge to support the CMIN's emergency care services. The best DM achieved results demonstrate the reliability of the variables provided by the information systems, reaching accuracy and specificity values of approximately 74% and 94%, respectively.

This article includes five sections in addition to the introduction. The second section presents the context and related work. The study materials and methods are described in section three. Section four states the data mining process following the Cross Industry Standard Process for Data Mining (CRISP-DM) phases. Formerly, section five contains a discussion about the obtained results, while the last section includes a set of considerations as well as possible directions for future work.

2 Background and Related Work

2.1 Context

The quality of care and patient's safety are the primary focus in clinical environments. Moreover, the patient's satisfaction is also an important outcome [10]. GO triage are trouble with reports of low satisfaction and long length of waiting times, making healthcare providers struggle with time management and budgetary constraints [11].

CMIN, as one of the four constituent hospitals of CHP, is prepared to provide neonatology, obstetrics and gynecology services and accompany the women and child conditions since the early pregnancy until the first stages of child growth. Its triage room provides outpatient care for women who require evaluation of labor, assessment of fetal wellbeing and acute obstetric issues. Studying the clinical and environmental information available at the patient's arrival and identifying its relation with the waiting times can be crucial to improve the unit services and therefore the patient's quality care and satisfaction.

2.2 Pre-triage System for Gynecology and Obstetrics Care in CMIN

In a hospital environment a vast number of triage systems are used. The most common are those with five levels of severity such as the Emergency Severity Index (ESI), the Manchester Triage System (MTS) and the Canadian Triage Acuity Scale (CTAS) [9]. These systems are limited as tools to use in Maternity Care due to their lack of flexibility. Since they were meant to be used in general emergency units, their guidelines do not have an appropriate degree of generalization to GO services [12].

Accordingly, CMIN healthcare professionals and Information System researchers developed a pre-triage system specific for GO, where the queries are focused on the particular type of patients that attend these services, i.e. pregnant women. This system is implemented in CMIN since 2010, and classifies patients according to the severity of their clinical condition, establishing clinical priorities and not diagnosis [13]. The system is inserted in the Intelligent Decision Support System (IDSS) and provides the triage result based on a set of predefined questions in the form of rules of a decision tree. The IDSS is an interactive and adaptable system, which uses artificial intelligence techniques and decision models to analyze a vast amount of variables to answer a question. The presence of the IDSS offers a better understanding of the patient's real state. In opposition to the MTS, triage is done by physicians and nurses [14].

Regarding the procedure, when a patient is admitted in the GO emergency room needing urgent care (i.e. women who arrive by ambulance or present serious conditions) she is immediately assisted, without any triage. In the remaining cases, the patients have to wait to be called to the triage room. Women in need of urgent care are classified as URG, while non-urgent cases are sent to consultation (ARGO). After the triage moment, the patients return to the waiting room where they remain until being called to the respective care [13, 15]. This study focus on the prediction of the patient's waiting time before the triage, i.e., when the patient is waiting to be called to the triage room.

2.3 Integration, Archive and Diffusion of Medical Information

As mentioned in the previous section, the current study depends of the combination of the information provided by the information systems implemented in the healthcare institution. The SAPE resulted from the Information Systems in Nursing. It records clinical episodes associated with each patient as an alternative to the traditional way of information on paper [16]. EHR stores and retrieves detailed patient information, as the admission form, helping monitoring, improving and reporting data on health care quality and safety [17]. These heterogeneous systems are only able to provide the information needed since they are connected by the Agency for Integration, Archive and Diffusion of Medical Information (AIDA), which allows the interoperability of the hospital existing systems [15]. AIDA is a platform based on the use of pro-active agents, being responsible for tasks such as communication with systems, sending and receiving information and responding to requests of information. This multi-agent system enables the standardization of clinical systems and overcomes the medical and administrative complexity of the different sources of information from the hospital, thereby allowing a suitable information management [18, 19].

2.4 Knowledge Discovery and Data Mining in Healthcare

The Knowledge Discovery in Databases (KDD) process can be described as an automatic, exploratory analysis and modeling of large data repositories, which identifies meaningful and potentially useful information [20, 21]. The KDD process is a set of five steps, beginning with the selection of the data set. The next step includes cleaning and processing the data, to make it consistent. Then, the data is transformed according to the study goal. Data mining (DM) is the core step, which results in the discovery of hidden knowledge. Finally, follows the interpretation and evaluation of the patterns attained [22]. This study follows this process to discover new knowledge, using DM classification algorithms in order to attain data models.

Data mining techniques have been used in many domains to solve classification, segmentation, association, diagnosis and prediction problems [23]. In healthcare, data mining has recently been used to reduce the number of adverse events, anticipate patient's future behavior and find solutions regarding the institution's management, providing organizations with quality services and clinical decisions at affordable costs [24, 25].

In CHP several data mining models are induced in Intensive Care Units [35] to predict patient outcome, readmissions, length of stay, organ failure, among others in real-time [16, 26-28]. Regarding the maternity care, many studies have been conducted that resorted to DM techniques to solve health services limitations. For instance, the induction of DM models to predict preterm and full term births, as it remains a complex clinical problem for families and the healthcare system [29]. In the CMIN, data mining models were already used to predict events in the Voluntary Interruption of Pregnancy [30] and to categorize GO patients through a clustering based approach [31].

3 Study Description

In order to achieve knowledge from the available data, the study follows the KDD process described in section 2.4. The DM step applies the Cross Industry Standard Process for Data Mining (CRISP-DM), a sequence of six phases well-structured and defined, which works as guidelines for data miners [32].

CRISP-DM organizes the DM process into six phases: business understanding, data understanding, data preparation, modeling, evaluation and deployment. These phases provide a path to follow while planning and implementing a DM project [33].

The DM models were induced using different classification algorithms: Decision Trees (DT), Generalized Linear Models (GLM), Support Vector Machine (SVM), Naïve Bayes (NB) and Neural Networks (NN). The selection of these techniques was based on the playability of the models and engine efficiency, given the data.

The dataset studied consists of 78186 admissions on the gynecology and obstetrics care emergency room in CMIN, comprising a period between 2010-01-06 and 2015-06-25, a total of 1850 days and 32207 women patients.

4 Data mining process

In the following section the data mining step of the KDD process is described, according to the CRISP-DM phases.

4.1 Business Understanding

The business goal of the present study is the prediction of pre-triage waiting times in the emergency room of the gynecology and obstetrics care unit in CMIN. Moreover, the study pretends to identify the relation between the environmental and personal information available at the patient's admission and the actual waiting time, in order to improve the triage process in any maternity care.

Therefore, the DM goal is to develop accurate models able to predict the pre-triage waiting time from the collected data. Acceptable models with decent statistical metrics lead to an improvement in the quality of services.

4.2 Data understanding

In this phase, the relevant variables from the two data sources were collected, forming the dataset. A total of 13 variables were considered. Some of them were temporal features: the day of the week, the part of the day, the month, the day of the month, the part of the month, the trimester, the hour and the season of the year, concerning each particular entry. In addition, the dataset is composed by the identification number (ID) of the triage professional, the number of triage professionals working (NTP) and the number of patients waiting in the room (NPW). All these variables were selected to understand how the environment influences the waiting time. Finally, it also covers two patient's characteristics: the age and the gestation weeks (in case of pregnancy).

Table 1 shows the different classes of some used variables and the percentage of their occurrence, which provides a better interpretation of the dataset. In turn, table 2 present statistical measures concerning the numerical variables studied.

Table 1. Classes and occurrences of some variables used in the dataset

<i>Variable</i>	<i>Class</i>	<i>Percentage</i>
Day of the Week	Sunday	10.14%
	Monday	18.51%
	Tuesday	14.51%
	Wednesday	15.27%
	Thursday	15.36%
	Friday	14.95%
	Saturday	11.26%
Part of the Day	Morning	44.49%
	Evening	55.26%
	Night	0.25%
Part of the Month	First Third	29.91%
	Second Third	36.32%
	Last Third	33.77%

Trimester	First Quarter	25.40%
	Second Quarter	28.22%
	Third Quarter	24.11%
	Last Quarter	22.27%
Station	Winter	28.02%
	Spring	28.92%
	Summer	24.50%
	Autumn	22.56%

Table 2. Statistical measures of the numerical variables of the dataset

<i>Variable</i>	<i>Min</i>	<i>Max</i>	<i>Avg</i>	<i>Std Dev</i>
NTP	1	21	5.24	2.85
NPW	1	24	2.18	1.48
Age	8	92	32.01	10.44
Gestation Weeks	0	42	11.39	15.48

The target variable *Waiting Time* was convened in classes, time intervals, in order to run the classification algorithms. Initially, it was simply divided in two classes: 0 being the waiting time between the minimum value and the average, and 1, the waiting time superior to the average. This approach allowed examining the viability of the study itself and confirming that the selected variables are indeed suitable to predict the pre-triage waiting time. In a second stage, the target variable was distributed in four different classes, obeying the normal distribution of the target variable. Table 3 present both distributions of the target classes. These two approaches are properly addressed through the DM process.

Table 3. Both distribution of the target variable in classes

<i>Approach</i>	<i>Class</i>	<i>Distribution</i>	<i>Percentage</i>
2 classes	0	0 – 17 minutes	64.94%
	1	18 – 340 minutes	35.06%
4 classes	0	0 – 13 minutes	51.28%
	1	14 – 17 minutes	13.66%
	2	17 – 20 minutes	7.46%
	3	21 – 340 minutes	27.60%

4.3 Data Preparation

The information provided by the information systems required some transformations to reach the dataset exposed in section 4.2.

In a pre-processing phase, some variables were attained from raw content. This phase was extremely important to create the scenarios and it required the development of several procedures. For instance, the temporal variables, such as the day of the week or the part of the month, were obtained by arranging the patients' admission date of the entrance records. Similarly, the number of patients and triage professions required creating some procedures to count the pre-triage entries meeting the required conditions, and organize the information differently. The patient age was acquired through the

transformation of the birth and episode dates. Later, some null/noise values were deleted from the dataset to ensure the information consistency, remaining 78186 pre-triage records. The entire dataset was prepared and processed through Oracle SQL Developer. The dataset was then replicated and studied with different displays of the target variable mentioned in the previous subsection.

4.4 Modeling and Evaluation

In the modeling phase, the DM models were induced using the DM techniques: GLM, SVM, DT, NB and NN. The sampling method applied was the Holdout sampling, where 30% of the data was used for testing, and the remaining entries compose the training set. Finally, different variables' scenarios were combined to identify which variables influence the pre-triage waiting time on the GO emergency room. The ten considered scenarios were:

- S0: {All variables}
- S1: {Day of the week, Part of the day, Trimester, Hour, Season}
- S2: {NTP, NPW}
- S3: {Age, Gestation Weeks, Day of the week, Part of the day, Trimester, Hour, Season, NTP, NPW}
- S4: {Month, Day of the month, Hour, ID professional, NPW}
- S5: {Age, Day of the week, Part of the day, Day of the month, Part of the month, Season, NTP}
- S6: {Month, Day of the month, Part of the month, Trimester, Season}
- S7: {Day of the week, Part of the day, Hour}
- S8: {Age, Gestation Weeks, Day of the month, Season, NTP, NPW}
- S9: {Age, Gestation Weeks, Trimester, NTP, NPW}

The DM models were induced in both target approaches. Thus, each Data Mining Model (DMM) can be identified by equation 1.

$$DMM_m = DMT_y \times A \times T_s \times S_i \quad (1)$$

DMT_y refers to the DM technique, A_i is the target approach, T_s represents the sampling method and S_i identifies the scenario. A total of 100 models were induced (10 scenarios * 5 techniques * 1 sampling method * 2 target approaches).

All models were induced using the R Studio with the configurations presented in table 4, concerning the DM classification algorithms.

Table 4. Algorithms Settings

<i>Technique</i>	<i>R algorithm</i>	<i>Setting</i>	<i>Value</i>
DT	<i>Ctree</i>	Subset	Null
		Weights	Null
		Xtrafo	Ptrafo
		Ytrafo	Ptrafo
NB	<i>naiveBayes</i>	Subset	Null
		Laplace	0
GLM	<i>Glm</i>	Subset	Null

		Weights Family	Null Binomial
SVM	<i>Ksvm</i>	Scaled Type Kernel	True C-bsvc Vanilladot
NN	<i>Nnet</i>	Size Rang Decay Max It	4 0.1 5e-04 200

The statistic metrics described in equations 2, 3 and 4 were considered to evaluate the induced models. These metrics were estimated through the results provided by the confusion matrix (CMX) of each model.

$$Sensitivity = TP / (TP + FN) \quad (2)$$

$$Specificity = TN / (TN + FP) \quad (3)$$

$$Accuracy = TP / (TP + FT + TN + FN) \quad (4)$$

CMX contains four types of results: the number of True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN). The CMX along with these metrics were obtained automatically using the package ‘caret’ in R Studio.

The best results concerning the first target approach are exposed in table 5. It contains the top 4 scenarios and the best DM techniques to each of them.

Table 5. Best statistic results in view of the first target approach (2 classes)

<i>DMT</i>	<i>Accuracy</i>	<i>Specificity</i>	<i>Sensitivity</i>	<i>DMT</i>	<i>Accuracy</i>	<i>Specificity</i>	<i>Sensitivity</i>
<i>Scenario 0</i>				<i>Scenario 2</i>			
DT	0.7135	0.9078	0.3534	DT	0.7105	0.9394	0.2859
NB	0.7038	0.9163	0.3102	GLM	0.7131	0.9237	0.3225
GLM	0.7282	0.9304	0.3323	NN	0.7106	0.9337	0.2965
<i>Scenario 3</i>				<i>Scenario 4</i>			
DT	0.7170	0.9288	0.3236	DT	0.7118	0.9270	0.3159
GLM	0.7261	0.9274	0.3383	GLM	0.7028	0.9198	0.3179
NN	0.7381	0.9366	0.3473	NN	0.7117	0.9303	0.3093

The results are satisfactory enough to prove that there is a relation between the variables studied and the pre-triage waiting time. The second target approach, which divides the target in 4 classes, was pursued. In this case, since the target is not binary, the only metric used to evaluate the models was the accuracy. Accordingly, the results are presented in table 6.

Table 6. Best accuracy results in view of the second target approach (4 classes)

<i>Scenario</i>	<i>DMT</i>	<i>Accuracy</i>
0	NB	0.6134
3	DT	0.6243
9	DT	0.6088

The best results were achieved in the first approach, reaching a specificity value of 93.94%. These high values of specificity mean that the models are highly capable predicting class 0, pre-triage waiting times shorter than the average waiting time.

In a clinical environment, the first approach will provide to the healthcare professions a larger confidence (almost 94%) predicting smaller times, the second approach represents a valuable resource to the service, since nearly 62% of the triage events fit this profile (four classes).

Overall, the DM model that attained the best outcomes was scenario 3, inducing the neural network algorithm, obtaining an accuracy of 73.81%, representing a high assertiveness. The best accuracy results provided by the second approach have statistical value, since the dataset has a four class target, but are not enough to be considered worthy clinical achievements.

4.5 Deployment

After a careful review to the results, the best DM models as well as the processed data are reported to the maternity care unit of CMIN, being implemented in the IDSS and the Business Intelligence (BI) platform already deployed in the hospital, correspondently [34]. The BI platform assists the healthcare professionals and the administrators in their decision making, regarding the services and the patients. The study will contribute to the improvement of the pre-triage system and the conditions regarding the Gynecology and Obstetrics emergency room.

5 Discussion

In the evaluation phase the study achieved satisfactory results regarding the dataset which divides the target variable *Waiting Time* in two classes (lesser than 17 minutes or greater than 18 minutes). The best 3 models are highlighted in table 7.

Table 7. Top 3 models that present the higher values of accuracy, sensitivity and specificity in view of the first target approach (2 classes)

<i>S</i>	<i>DMT</i>	<i>Accuracy</i>	<i>Specificity</i>	<i>Sensitivity</i>
0	GLM	0.7289	0.9304	0.3323
3	DT	0.7170	0.9288	0.3236
3	NN	0.7381	0.9366	0.3473

Overall, the specificity values are far higher than the sensitivity ones, allowing the gynecology and obstetrics care unit to identify upfront the patients which waiting time for pre-triage will be shorter, and subsequently, to organize the resources accordingly.

Regarding the studied variables, the scenarios 0, 2, 3, 4 and 9 achieved the best outcomes. Thus, the variables directly proportional to the pre-triage waiting time are the number of patients present in the waiting room (NPW), the number of triage professionals working at the moment (NTP) and some temporal variables – the hour of the day, the day of the month and the season of the year. This information benefits the

maternity emergency room workflow, since it can identify outstanding situations, and therefore, improve the healthcare services.

The dataset that divides the target variable into 4 groups (obeying its normal distribution) obtained good results in a statistical point of view, since it is not a binary prediction. Nevertheless, the best accuracy value achieved is 62.43%, not being good enough to be used in the maternity care decision support systems.

The analysis of the patient records from January 2010 to June 2015 also allows detecting some weaknesses in the pre-triage system implemented in CMIN. Since the system only distinguishes patients in two levels of priorities (URG and ARGO), adverse events may occur in this categorization. Besides, some triage professionals force a different output than would be expected by the pre-triage system.

6 Conclusions and Future Work

This study aimed the prediction of pre-triage waiting times in the emergency room of gynecology and obstetrics care unit. A dataset was built from real data recorded in the SAPE and EHR systems of Centro Materno Infantil do Norte (CMIN), the maternal and perinatal care unit of Centro Hospitalar of Oporto (CHP), collecting personal and environmental variables available at the patient's admission. After inducing several data mining techniques on different combinations of the dataset variables (scenarios), satisfactory results were achieved. The best DM model was acquired by inducing Neural Networks algorithm through the variables composing scenario 3, achieving an accuracy of approximately 74% and a specificity of 94%, when dividing the target variable in two classes. These results lead to quality improvements in the emergency room of the maternity care.

This study also was important to identify which variables most contributes to change the waiting time. This information is valuable to the ED. They have now a new tool able to give to them an idea of the patient waiting time basing their analysis in the data associated to the critic variables.

The Data Mining models induced and the variables identified as critical are the most important contribution of this work.

Alongside, another study is being conducted to predict the post-triage waiting time (the time between the moment of pre-triage and the doctor's admission) in CMIN. As for future work, both studies will be implemented in the IDSS and BI platform used in the maternity unit, optimizing patient flow through the ED and therefore, improving their safety and satisfaction.

Acknowledgments

This work has been supported by FCT - Fundação para a Ciência e Tecnologia within the Project Scope UID/CEC/00319/2013.

ReferencesBergs, J., Verelst, S., Gillet, J.-B., Deboutte, P., Vandoren, C., Vandijck, D.: The number of patients simultaneously present at the emergency department as an indicator of unsafe waiting times: A receiver operated curve-based evaluation. *International Emergency Nursing* 22, 185-189 (2014)

2. Hoot, N., Aronsky, D.: Systematic Review of Emergency Department Crowding: Causes, Effects and Solutions. *Annals of Emergency Medicine* 52, 126-136 (2008)
3. Stover-Baker, B., Stahlman, B., Pollack, M.: Triage nurse prediction of hospital admission. *Journal of Emergency Nursing* 38 (3), 306-310 (2012)
4. Sun, Y., Teow, K., Heng, B., Ooi, C., Tay, S.: Real-time prediction of waiting time in the emergency department, using quantile regression. *Annals of Emergency Medicine* 60 (3), 299-308 (2012)
5. Boudreaux, E., O'Hea, E.: Patient satisfaction in the emergency department: a review of the literature and implications for practice. *The Journal of emergency medicine* 26 (1), 13-26 (2004)
6. Shaikh, S., Witting, M., Winters, M., Brodeur, M., Jerrad, D.: Support for a waiting room time tracker: a survey of patients waiting in an urban ED. *The Journal of Emergency Medicine* 44 (1), 225-229 (2013)
7. Taylor, C., Bengler, J.: Patient satisfaction in emergency medicine. *The Journal of emergency medicine* 21, 528-532 (2004)
8. Khodambashi, S.: Business Process Re-Engineering Application in Healthcare in a relation to Health Information Systems. *Procedia Technology* 9, 949-957 (2013)
9. Pereira, E., Brandão, A., Salazar, M., Portela, F., Santos, M., Machado, J., Abelha, A., Braga, J.: Pre-Triage Decision Support Improvement in Maternity Care by means of Data Mining. *Integration of Data Mining in Business Intelligence Systems*, 175-192 (2014)
10. Paul, J., Jordan, R., Duty, S., Engstrom, J.: Improving Satisfaction with Care and Reducing Length of Stay in an Obstetric Triage Unit Using a Nurse-Midwife-Managed Model of Care. *Journal of Midwifery and Women's Health*, 1-7 (2013)
11. Zocco, J., Williams, M., Longobucco, D., Bernstein, B.: A systems analysis of obstetric triage. *J Perinat Neonatal Nurs* 21 (4), 315-22 (2007)
12. Murray, M., Bullard, M., Grafstein, E.: Revisions to the Canadian emergency department triage and acuity scale implementation guidelines. *Cjem*, 421-427 (2004)
13. Abelha, A., Pereira, E., Brandão, A., Portela, F., Santos, M., Machado, J., Braga, J.: Improving Quality of Services in Maternity Care Triage System. *International Journal of E-Health and Medical Communications* 6 (2), 10-26 (2015)
14. Cabral, A., Pina, C., Machado, H., Abelha, A., Salazar, M., Quintas, C., Portela, F., Machado, J., Neves, J., Santos, M.: Data Acquisition Process for an Intelligent Decision Support in Gynecology and Obstetrics Emergency Triage. *CENTERIS*, 223-232 (2011)
15. Abelha, A., Pereira, E., Brandão, A., Portela, F., Santos, M., Machado, J.: Simulating a multi-level priority triage system for Maternity Emergency. *European Simulation and Modelling Conference* (2014)
16. Portela, F., Cabral, A., Abelha, A., Salazar, M., Quintas, C., Machado, J., Santos, M.: Knowledge Acquisition Process for Intelligent Decision Support in Critical Health Care. *Healthcare Administration: Concepts, Methodologies, Tools, and Applications*, 270 (2014)
17. Abelha, A., Analide, C., Machado, J., Neves, J., Santos, M., and Novais, P.: Ambient intelligence and simulation in health care virtual scenarios. *Establishing the Foundation of Collaborative Networks*, 461-468 (2007)
18. Abelha, A., Machado, J., Santos, M., Allegro, S., Rua, F., Paiva, M., Neves, J.: Agency for Integration and Archive of Medical Information. *IASTED International Conference - Artificial Intelligence and Applications* (2002)

19. Peixoto, H., Santos, M., Abelha, A., Machado, J.: Intelligence in Interoperability with AIDA. 20th International Symposium on Methodologies for Intelligent Systems (2012)
20. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: From data mining to knowledge discovery in databases. *AI magazine* 17 (3), 37 (1996)
21. Lin, J., Gan, W., Hong, T.-P., Tseng, V.: Efficient algorithms for mining up-to-date high-utility patterns. *Advanced Engineering Informatics* (2015)
22. Maimon, O., Rokach, L.: Introduction to Knowledge Discovery and Data Mining. *Data Mining and Knowledge Discovery Handbook*, 1-17 (2005)
23. Xia, J., Xie, F., Zhang, Y., Caulfield, C.: *Artificial Intelligence and Data Mining: Algorithms and Applications*. Abstract and Applied Analysis (2013)
24. Srinivas, K., Rani, B. K., Govrdhan, A.: Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks. *International Journal on Computer Science and Engineering* 2 (2), 250-255 (2010)
25. Yoo, I., Alafaireet, P., Marinov, M., Pena-Hernandez, K., Gopidi, R., Chang, J.-F., Hua, L.: Data Mining in Healthcare and Biomedicine: A Survey of the Literature. *J Med Syst* 36, 2431–2448 (2012)
26. Braga, P., Portela, F., Santos, M., Rua, F.: Data mining models to predict patient's readmission in intensive care units. In : *ICAART - International Conference on Agents and Artificial Intelligence*, Angers, France (2014)
27. Portela, F., Santos, M., Machado, J., Abelha, A., Silva, Á.: Pervasive and Intelligent Decision Support in Critical Health Care Using Ensembles. *Information Technology in Bio-and Medical Informatics Springer Berlin Heidelberg*, 1-16 (2013)
28. Veloso, R., Portela, F., Santos, M., Sila, Á., Rua, F., Abelha, A., Machado, J.: A Clustering Approach for Predicting Readmissions in Intensive Medicine. *Procedia Technology* 16, 1307-1316 (2014)
29. Beta, J., Akolekar, R., Ventura, W., Syngelaki, A., Nicolaides, K.: Prediction of spontaneous preterm delivery from maternal factors, obstetric history and placental perfusion and function at 11–13 weeks. *Prenatal diagnosis* 31(1), 75-83 (2011)
30. Brandão, A., Pereira, E., Portela, F., Santos, M., Abelha, A., Machado, J.: Managing Voluntary Interruption of Pregnancy using Data Mining. *Procedia Technology* 16, 1297–1306 (2014)
31. Pereira, S., Portela, F., Santos, M., Abelha, A., Machado, J.: Clustering-based Approach for Categorizing Pregnant in Obstetrics and Maternity Care. In : *C3S2E*, Yokohoma, Japan, pp.1-5 (2015)
32. Shafique, U., Qaiser, H.: A Comparative Study of Data Mining Process Models (KDD, CRISP-DM and SEMMA). *International Journal of Innovation and Scientific Research* 12(1), 217-222 (2014)
33. Shearer, C.: The CRISP-DM Model: The New Blueprint for Data Mining. *Journal of Data Warehousing* 5(4), 13-22 (2000)
34. Pereira, E., Brandão, A., Portela, F., Santos, M., Machado, J., Abelha, A.: Business Intelligence in Maternity Care. In : *IDEAS - International Database Engineering & Applications Symposium*, Portugal (2014)
35. Portela, F., Santos, M.F., Machado, J., Abelha, A., Silva, Á., Rua, F.: Pervasive and intelligent decision support in Intensive Medicine—the complete picture. *Information Technology in Bio-and Medical Informatics*, pp. 87-102. Springer (2014)