University of Minho
School of Engineering

Carlos Filipe da Silva Portela

Pervasive Intelligent Decision Support in Critical Health Care

PhD in
Information Systems and Technologies

Supervised by
Professor Doctor Manuel Filipe Vieira Torres dos Santos

Co-supervised by
Professor Doctor Álvaro Moreira Silva

October 2013

Declaration

Name: Carlos Filipe da Silva Portela

Email: cfp@dsi.uminho.pt

Phone: +351 967562540

Citizen Card: 12718892

Thesis Title:

Pervasive Intelligent Decision Support in Critical Health Care

Supervisors:

Professor Doctor Manuel Filipe Vieira Torres dos Santos

Professor Doctor Álvaro Moreira Silva

Conclusion Year: 2013

PhD Knowledge area:

Information System and Technologies

University of Minho, 15/10/2013

Signature:

Filipe Portela

Acknowledgments

In this unique moment, I would like to thank all the people who have always been on my side and contributed to the success of this work.

First of all, I want to thank for the commitment, dedication and wisdom of my supervisor Professor Manuel Filipe Santos, and my co-supervisor Doctor Álvaro Silva as well as the interventions from Professor José Machado and Professor António Abelha. They were both the key in the progress and success of this work.

I would also to thank the contributions from the physician and head of the ICU department, Dr. Fernando Rua, and the chief nurse José António and all the doctors and nurses who helped me along the way. Your commitment and professionalism allowed to establish contact with the real environment and to receive feedback so to improve the processes.

I would also like to thank the kindness and hospitality offered by CHP. In the persons of Dr. César Quintas and Dra. Manuel Salazar, I thank to all Information System Service professionals for their commitment and professionalism.

I wish here to offer a word of appreciation to all PhD colleagues, the faculty, and master co-supervised students and staff of the Department of Information Systems that coexisted with me throughout the process.

I also leave a word of thanks to all those who have crossed with me in one way or another, motivated me and helped me to overcome the difficulties of the route.

I would also like to express my thanks to FCT which funded the project through the PhD grant: SFRH/BD/70156/2010 and) and for the financial support through the contract PTDC/EIA/72819/ 2006 (INTCare).

Here, I let a special word of thanks to my family, particularly parents, brother, grandmother and godparents. Their presence, patience and words of motivation have been remarkable and fundamental to keep the track.

Finally, I let an appreciation to a special person, who has accompanied me throughout these long four years. It all started in 2009, when I met her in the political campaign and by that time I began my PhD. So far I have no words to express how much her presence and support were crucial for me and this project. She was the cornerstone of the whole construction that held firm until the end.

Ana Sampaio Portela, thank you for always having been on my side and for helping me to complete this process.

Without your cooperation, this PhD would not have come to fruition,

For all much and much thanks.

Eternally grateful,

Filipe Portela

Abstract

Intensive Care Units (ICU) are recognized as being critical environments, due to the fact that patients admitted to these units typically find themselves in situations of organ failure or serious health conditions. ICU professionals (doctors and nurses) dedicate most of their time taking care for the patients, relegating to a second plan all documentation tasks. Tasks such as recording vital signs, treatment planning and calculation of indicators, are only performed when patients are in a stable clinical condition. These records can occur with a lag of several hours. Since this is a critical environment, the Process of Decision Making (PDM) has to be fast, objective and effective. Any error or delay in the implementation of a particular decision may result in the loss of a human life. Aiming to minimize the human effort in bureaucratic processes and improve the PDM, dematerialization of information is required, eliminating paper-based recording and promoting an automatic registration of electronic and real-time data of patients. These data can then be used as a complement to the PDM, e.g. in Decision Support Systems that use Data Mining (DM) models.

At the same time it is important for PDM to overcome barriers of time and space, making the platforms as universal as possible, accessible anywhere and anytime, regardless of the devices used. In this sense, it has been observed a proliferation of pervasive systems in healthcare. These systems are focused on providing healthcare to anyone, anytime and anywhere by removing restrictions of time and place, increasing both the coverage and quality of health care. This approach is mainly based on information that is stored and available online.

With the aim of supporting the PDM a set of tests were carried out using static DM models making use of data that had been collected and entered manually in Euricus database. Preliminary results of these tests showed that it was possible to predict organ failure and outcome of a patient using DM techniques considering a set of physiological and clinical variables as input. High rates of sensitivity were achieved: Cardiovascular - 93.4%; Respiratory - 96.2%; Renal - 98.1%; Liver - 98.3%; hematologic - 97.5%; and Outcome and 98.3%. Upon completion of this study a challenge emerged: how to achieve the same results but in a dynamic way and in real time? A research question has been postulated as: "To what extent, Intelligent Decision Support Systems (IDSS) may be appropriate for critical clinical settings in a pervasive way? ".

Research work included:

1. To percept what challenges a universal approach brings to IDSS, in the context of critical environments;
2. To understand how pervasive approaches can be adapted to critical environments;
3. To develop and test predictive models for pervasive approaches in health care.

The main results achieved in this work made possible:

1. To prove the adequacy of pervasive approach in critical environments;
2. To design a new architecture that includes the information requirements for a pervasive approach, able to automate the process of knowledge discovery in databases;
3. To develop models to support pervasive intelligent decision able to act automatically and in real time. To induce DM ensembles in real time able to adapt autonomously in order to achieve predefined quality thresholds (total error < = 40 %, sensitivity > = 85 % and accuracy > = 60 %).

Main contributions of this work include new knowledge to help overcoming the requirements of a pervasive approach in critical environments. Some barriers inherent to information systems, like the acquisition and processing of data in real time and the induction of adaptive ensembles in real time using DM, have been broken. The dissemination of results is done via devices located anywhere and anytime.

Resumo

As Unidades de Cuidados Intensivos (UCIs) são conhecidas por serem ambientes críticos, uma vez que os doentes admitidos nestas unidades encontram-se, tipicamente, em situações de falência orgânica ou em graves condições de saúde.

Os profissionais das UCIs (médicos e enfermeiros) dedicam a maioria do seu tempo no cuidado aos doentes, relegando para segundo plano todas as tarefas relacionadas com documentação. Tarefas como o registo dos sinais vitais, o planeamento do tratamento e o cálculo de indicadores são apenas realizados quando os doentes se encontram numa situação clínica estável. Devido a esta situação, estes registos podem ocorrer com um atraso de várias horas. Dado que este é um ambiente crítico, o Processo de Tomada de Decisão (PTD) tem de ser rápido, objetivo e eficaz. Qualquer erro ou atraso na implementação de uma determinada decisão pode resultar na perda de uma vida humana. Com o intuito de minimizar os esforços humanos em processos burocráticos e de otimizar o PTD, é necessário proceder à desmaterialização da informação, eliminando o registo em papel, e promover o registo automático e eletrónico dos dados dos doentes obtidos em tempo real. Estes dados podem, assim, ser usados com um complemento ao PTD, ou seja, podem ser usados em Sistemas de Apoio à Decisão que utilizem modelos de *Data Mining* (DM).

Ao mesmo tempo, é imperativo para o PTD superar barreiras ao nível de tempo e espaço, desenvolvendo plataformas tão universais quanto possíveis, acessíveis em qualquer lugar e a qualquer hora, independentemente dos dispositivos usados. Nesse sentido, tem-se verificado uma proliferação dos sistemas *pervasive* na saúde. Estes sistemas focam-se na prestação de cuidados de saúde a qualquer pessoa, a qualquer altura e em qualquer lugar através da eliminação das restrições ao nível do tempo e espaço, aumentando a cobertura e a qualidade na área da saúde. Esta abordagem é, principalmente, baseada em informações que estão armazenadas disponíveis *online*.

Com o objetivo de suportar o PTD, foi realizado um conjunto de testes com modelos de DM estáticos, recorrendo a dados recolhidos e introduzidos manualmente na base de dados “Euricus”. Os resultados preliminares destes testes mostraram que era possível prever a falência orgânica ou a alta hospitalar de um doente, através de técnicas de DM utilizando como valores de entrada um conjunto de variáveis clínicas e fisiológicas. Nos testes efetuados, foram obtidos elevados níveis de sensibilidade: cardiovascular - 93.4%; respiratório - 96.2%; renal - 98.1%; hepático - 98.3%; hematológico - 97.5%; e alta hospitalar - 98.3%. Com a finalização deste estudo, observou-se o

aparecimento de um novo desafio: como alcançar os mesmos resultados mas em modo dinâmico e em tempo real? Uma questão de investigação foi postulada: "Em que medida os Sistemas de Apoio à Decisão Inteligentes (SADIs) podem ser adequados às configurações clínicas críticas num modo *pervasive*?". Face ao exposto, o trabalho de investigação inclui os seguintes pontos:

1. Perceber quais os desafios que uma abordagem universal traz para os SADIs, no contexto dos ambientes críticos;
2. Compreender como as abordagens *pervasive* podem ser adaptadas aos ambientes críticos;
3. Desenvolver e testar modelos de previsão para abordagens *pervasive* na área da saúde.

Os principais resultados alcançados neste trabalho tornaram possível:

1. Provar a adequação da abordagem *pervasive* em ambientes críticos;
2. Conceber uma nova arquitetura que inclui os requisitos de informação para uma abordagem *pervasive*, capaz de automatizar o processo de descoberta de conhecimento em base de dados;
3. Desenvolver modelos de suporte à decisão inteligente e *pervasive*, capazes de atuar automaticamente e em tempo real. Induzir *ensembles* DM em tempo real, capazes de se adaptarem de forma autónoma, com o intuito de alcançar as medidas de qualidade pré-definidas (erro total <= 40 %, sensibilidade> = 85 % e acuidade> = 60 %).

As principais contribuições deste trabalho incluem novos conhecimentos para ajudar a ultrapassar as exigências de uma abordagem *pervasive* em ambientes críticos. Algumas barreiras inerentes aos sistemas de informação, como a aquisição e o processamento de dados em tempo real e a indução de *ensembles* adaptativos em tempo real utilizando DM, foram transpostas. A divulgação dos resultados é feita através de dispositivos localizados, em qualquer lugar e a qualquer hora.

Contents

List of Abbreviations

ABI	Adaptive Business Intelligence
ACE	Accumulated Critical Events
AI	Artificial Intelligence
ANN	Artificial Neuronal Networks
ArtDia	Blood Pressure – Diastolic
ArtSys	Blood Pressure – Systolic
BI	Business Intelligence
BP	Blood Pressure
CDSS	Clinical Decision Support System
CE	Critical Event
CEnv	Critical Environment
CHP	Centro Hospitalar do Porto
DB	Data Base
DDS	Data-Driven System
DM	Data Mining
DMP	Decision Making Process
DSM	Data Stream Mining
DSM	Data Stream Mining
DSS	Decision Support System
DT	Decision Trees
DW	Data Warehouse
ENR	Electronic Nursing Record
ETL	Extract transforming and Loading
ExS	Expert System
HL7	Health Level Seven
HSA	Hospital Santo António
ICU	Intensive Care Unit
IDSS	Intelligent Decision Support System
IM	Intensive Medicine
IS	Information Systems
KDD	Knowledge Discovery in Database
LR	Linear regression

MAS	Multi-agent System
Max	Maximum
Min	Minimum
NB	Naïve Byes
PaCO2	Partial pressure of CO2
PC	Pervasive Computing
PH	Pervasive Health Care
PID	Patient Identification
PIDSS	Pervasive Intelligent Decision Support System
PMML	Predictive Model Mark-up Language
RBS	Rules Based System
SAPS	Simplified Acute Physiology Score
SOFA	Sequential Organ Failure Assessment
SPO2	Saturation of Oxygen
SVM	Support Vector Machine
TAM	Technology Acceptance Model
TISS	Therapeutic intervention scoring system
STAFF	ICU Physicians and nurses
TEMP	Temperature
ubiDSS	Ubiquitous DSS
UDM	Ubiquity Data Mining
UR	Urine Output
MEWS	Modified Early Warning Score
VSBM	Vital Signs Bedside monitors

List of Figures

List of Tables

CHAPTER I – Introduction

1. Scope

This work explores the use of Intelligent Decision Support Systems in Intensive Medicine (IM). A pervasive approach was considered in order to deal with dynamic aspects of the environment (Kwon, Yoo, & Suh, 2005).

IM allows for recovering patient in end-of-life or in organ failure conditions. The recovering phase depends on the decisions that are taken in Intensive Care Units (ICU). These decisions can influence more the patient outcome than applying of newer intervention. Decision Making Process (DMP) in ICU is complex and involves distinct stakeholders and delicate situations.

The development of this type of systems encompasses a process of Knowledge Discovery from Databases (KDD). KDD process enables the selection and transformation of the data collected in order to be interpreted as new knowledge. After that, new knowledge can be used by the physicians to support their decisions.

Intelligent Decision Support Systems (IDSS) make use of prediction and/or decision models induced by a KDD process to create new knowledge and support DMP, adapting them continuously.

To overcome time and geographic barriers, IDSS should be pervasive and universal being accessible anywhere and anytime. Decisions can then be taken "remotely". As a proof of concept, a prototype of a pervasive IDSS has been implemented in the ICU of Centro Hospitalar do Porto (CHP), in Oporto in the north of Portugal.

These changes allowed to obtain new important knowledge to the decision making process such as critical events, medical scores and the prediction of organ failure and patient outcome. Once the artefacts have been generated, they were evaluated through the use of questionnaires in order to prove that it is possible to implement an IDSS in a Critical Environment, such as Intensive Medicine, using pervasive features. TAM approach has been used in order to understand how technology was accepted by its users.

All achieved results were published in journals, book chapters and important conferences in the PhD area.

This thesis is based on an article format and consists of an introduction, an overview of the research methodologies used, a critical overview of the state of the art and a selection of the most significant articles, which represents the work done and the results achieved during the work.

The first chapter introduces the theme, presents the motivation, the goals, expected results and the scientific contributions. The second chapter makes a description of the methodologies used and presents their organic structure. The third chapter introduces the research project and presents the research question as well as a description of the research plan. Fourth chapter explores the state of the art. Then, the chapter V is the key-part of the document. In this chapter it is presented the main contributions and a list of the published articles. The published articles are presented by sub-sections: Pervasive Environment, Data Overview, Results of Decision Support, and Assessment of the Results. Firstly, in the sub-chapter I it is presented the pervasive environment and associated features in order to design a Pervasive Intelligent Decision Support System (PIDSS). The studies carried out and published in the articles of this sub-section contributed to a better understanding of the environment and helped to define the information needs of the Intensive Care Units. At the same time, it was also possible to define a measure to evaluate whether or not an environment is prepared to receive a Pervasive IDSS.

The following sub-section II presents the evolution of the Information System architecture, i.e. the way of the data is collected, processed and stored automatically and in real-time. The articles included in this sub-section present some contributions in terms of the Knowledge Discovery in Database (KDD) process and also explain how some common errors of ICU, such as noise values (acquisition of bad values) and missing data (data without patient identification), can be solved. These articles also present the transforming process in order to prepare the data for the main goals of the IDSS.

The sub-section III presents the main results attained in this work, that can contributes to the decision support process:

- Critical Events tracking System;
- ICU scoring system;
- Predictive system (Ensemble Data Mining).

Includes a collection of published articles explaining what, when and how the new knowledge was obtained. Real-world experimentation has been used in order to prove the concepts and artefacts explored / prototyped.

The last sub-section presents the results of another strand. This chapter contains a group of articles that discuss the results previously presented. All results obtained were assessed in terms of data and system quality and user acceptance. Quality measures were defined and the Technology Acceptance Model (TAM) was used to perceive the level of technology acceptance by ICU professionals using four constructs. Is also identified the knowledge provided for the Decision Making Process and a way for monitoring the system data quality.

The last two chapters conclude the document. In chapter VI is presented an overview of all the achieved results and it is made a final remark on the work developed and published. Chapter VII makes depicts of possible future research work.

1.1. Intelligent Decision Support Systems

Information Systems can be defined as systems which collect process, store, analyse and disseminate information to a specific purpose (E. Turban, McLean, & Wetherbe, 2001). They are planned and developed focusing on the organization goals and implemented with the objective to support the business regardless of the environment in which they operate. Since 1980 there are many organizations which are betting in the Expert System (ExS) to support the resolution of complex and specialized problems. With the evolution of the concept, the decision makers require proactive IDSS that use contextual data (Kwon, et al., 2005) in real-time.

DSS are proliferating in many areas of human activity that involve the resolution of complex problems, such as Intensive Care Medicine (IM). The systems using decision algorithms comprehend a mix of Rule-Based System (RBS) and learning techniques. Some of these systems attempt to imitate the process of human decision, whereas others provide predictions based in Data Mining and knowledge to help the decision process. In the case of the health care, the most common systems are designed to help the physicians make a diagnostic and decide about a treatment.

Knowledge engineering should use simple techniques of Artificial Intelligence (AI) to simulate the actions of human experts (Andrade, 1999). This approach is stimulating researchers on the study of human decision processes, because they have discovered many systemic and robust flaws in human reasoning (Kahneman & Tversky, 1982). It is argued here that such an understanding may help identify strategies to improve human performance (Lindgaard, Pyper, Frize, & Walker, 2009).

Intelligent Decision Support Systems (IDSS) are DSS able to online adapt to environmental changes (Gago, Silva, & Santos, 2007). This kind of system must have the ability to learn and adapt to new rules maintaining its reliability and consistency. The development of an IDSS typically involves the adoption of AI techniques. This type of system allows for the exploration of other approaches and concepts such as automatic data acquisition and automatic data transformation, online-learning, real-time and pervasiveness.

1.2. Intensive Medicine

Intensive medicine has as main goals to diagnose, monitor and care patients in weak conditions and with serious diseases and also recover them to their previous health condition, improving their life quality (Suter et al., 1994). The Clinical practices are focused in the support of patient with Multisystem Organ Failure (Coombs, 2003). This type of medicine is mostly used in the Intensive Care Units (ICUs) for life supporting and continuous monitoring of organic systems.

The first computers in ICU arose around 1960 (Clemmer, 2004) and were focused in the monitoring of cardiovascular and respiratory patients. Then, the Vital Signs (VS) of patients were incorporated into the monitors to show signs trends of electrocardiogram (ECG), blood pressure, central venous pressure, cardiac output, among others (Clemmer, 2004). At a later stage, it was realized that it was possible to store large amounts of data relating to the signs that were monitored at the bedside. The initial idea was then to assist physicians in the process of care at the bedside of patients (Clemmer, 2004). Over time, was possible detect some gaps in the ICU systems. The systems didn't know what it was necessary to the decision process - what the physicians wanted and what the physicians really need.

In addition, it was necessary to determine the current state of a patient; for this purpose, a set of several indexes had to be developed for assessing the severity of a disease. These indexes were intended to analyse and quantify the severity level related to the patient. For these indexes, it was taken into account the patient's physiological data and some parameters measured. To process automatically and to embed these indexes into an IDSS is a very difficult and slow mission, mainly because necessary data aren't collected or they are recorded into a paper format (Mador & Shaw, 2009).

According to Vankayala and Dasgupta (Dasgupta & Vankayala, 2007), the real-time intelligent systems are receiving increasing attention from the scientific community. Indeed, the development and maintenance of these systems is a challenging task and present a critical issue: how to ensure that the systems can actually work with almost zero latency. This represents an enormous difficulty in critical areas. In this case, due to the complex condition of critically ill patients and the huge amount of data, it is difficult for physicians to decide on the best procedure. The human factor can lead to make mistakes in the decision process, because there is not enough time to completely analyse the situations, often due to circumstances of stress. Moreover, it is not possible to analyse and store all data continuously (Pereira et al., 2007).

A quick interpretation of physiological data, which are monitored in intensive care, is critical to assess accurately the patient's condition. Data analysis enables the decision making support through the use of predictive and decision models. The use of artificial intelligence techniques can contribute to achieve these tasks, but their development requires correct patient data (K. Morik, 2003; Ying, Silvers, & Randolph, 2007). The quality of data is another barrier to the implementation of IDSS in IM. In this project, quality procedures were implemented in order to validate all the values collected and assess the results achieved.

2. Motivation

Health information systems are a research area facing interesting challenges. In particular, pervasive IDSS represent a promise due to recent progresses attained on predictive models. Interesting results have been obtained in research project INTCare (Gago et al., 2006; Gago, et al., 2007; M. F. Santos, 2006). Nevertheless, some limitations should be addressed, namely:

- To prove that the use of pervasive information technologies can significantly reduce the number of medical errors and thus save lives;
- To prove the importance of such systems to the Decision Support Process.
- To reduce considerably the costs of health (U. Varshney, 2009).

In this point, it is essential to encourage the users to participate in the process, motivating the hospital staff (doctors and nurses) to use this type of technology. These technologies should allow them to work with greater efficiency and effectiveness. They can also help in the process of decision-making, supplying information about the real condition of the patient. Complementarily, the help should be available in the exact moment of the problem occurrence, supporting decisions on the best treatment for a patient and selecting the most adequate procedure. Figure 1 summarizes the main factors that contributed to choose the main research areas.

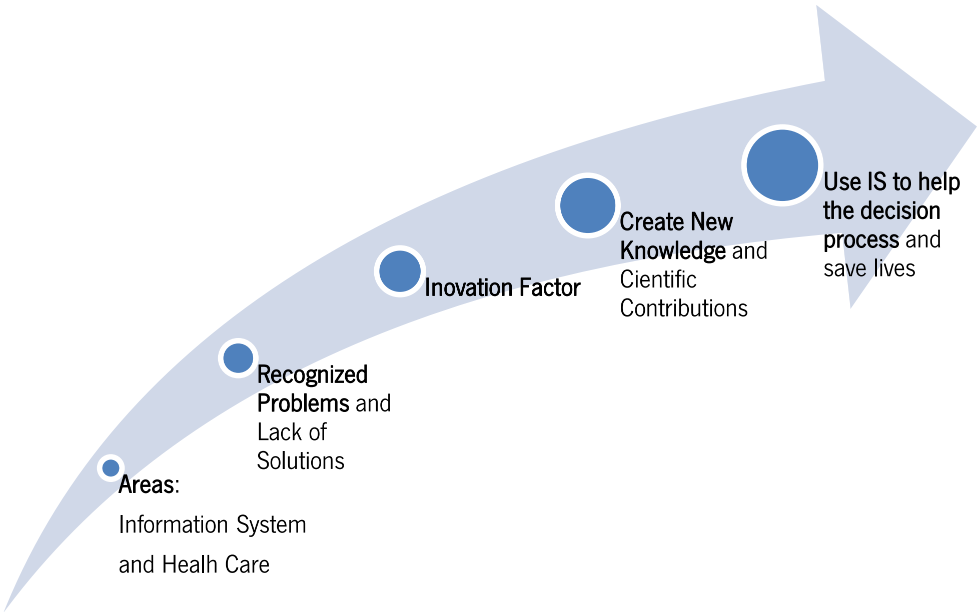


Figure 1 – Motivation factors to the development of the thesis.

3. Goals and expected results

The goals of this work can be postulated as:

1. **To understand the implications of a pervasive approach to Intelligent Decision Support Systems (IDSS) in the context of critical environments**
    i. To understand the impact of the pervasiveness in the construction of IDSS;
    ii. To analyse what are the changes relatively to the traditional model;
    iii. To study how IDSS should be adapted to operate in critical environments and in a pervasive way;
    iv. To study the impact of this approach on each one of the tasks inherent to IDSS.
2. **To adopt pervasive approaches in critical environments**
    i. To characterize Critical Environments (CEnv);
    ii. To understand and interpret the problem associated with pervasive approach in CEnv;
    iii. To analyse the impact and implications of the approach at level of:
        a. Environment;
        b. Predictive and Decision Models;
        c. Decision factors / criteria;
        d. Real-time;
        e. Technologies;
        f. People;
        g. Services.
    iv. To study the impact of the actions of prepare the environment to be pervasive;
    v. To study the processes inherent to ICU toward to its automation and dematerialization;
3. **To develop and test pervasive models in Intensive Care**
    i. To analyse and define the decision-making process;
    ii. To create predictive models to help in decision, prognosis, diagnosis and treatment;
    iii. To prototype an IDSS;
    iv. To test the models developed;
    v. To include pervasive capabilities in the IDSS;
    vi. To assess the results.

4. Scientific contributions

Nowadays the use of Decision Support Systems in Medicine is a reality, however, despite of data mining models for predicting already in use, the decision-making process is "stuck" because is not adaptive. At this time, taking into account the changes that have occurred at the level of Information Systems, communications and demands of critical environments, as well as the emergence of ubiquitous and pervasive environments, there is a set of challenges for which the critical areas of medicine are not yet prepared. In this sense, it was generated scientific knowledge to two different areas: Health and Information Systems.

4.1. Health Area

In this area it is expected the introduction of new results at level of the healthcare information systems and new knowledge that can contribute to better understanding of the real condition of the patient. As main goal of this new knowledge it is the possibility to improve the clinical condition of a patient based on the prediction, decision and in the implementation of improvements in the environment and in the process of decision making of the Intensive Care Units.

4.2. Information Systems Area

In this work one of the objectives is to introduce new knowledge which allow to help in the resolution of problems related to ubiquity and pervasiveness in critical environments. This work promotes a set of studies on these themes. As consequence of this work it is possible to identify what are the main information systems limitations and what path to follow when someone is faced with such problems.

At level of Information System in real-time, this work overcome a set of barriers allowing an automatic data acquisition, an automatic data processing and an induction of ensembles data mining in real-time. With this knowledge it is possible presenting always the best result in the right moment, being the results disseminated through the ubiquitous devices anywhere and anytime. This work was based in a major investigation with scientific rigor and presents a set of rules, models and changes that need to be made on this type of environments in order to support the real-time decision. The process was implemented in a real environment where now everyone who have access privileges can "enter" and consult crucial information to the Decision Making Process anywhere and anytime.

CHAPTER II – Research Methodologies

1. Description of Methodologies

This research project was guided by quantitative methodologies. In an initial step, it was used an exploratory approach in order to understand the needs and set the variables to be used. At this stage, a number of studies and tests were carried out. After obtaining the results, these were tested through a confirmatory stage, where the main objective was to prove the validity of the models, architectures or artefacts found after the exploratory phase.

Throughout this work some methodologies were used in a crossed way: Design Research, Case Studies, Field Experiments and Proof of Concept. Complementarily, to endure the research process some techniques were used, namely: Literature Review, Questionnaires and Techniques of Data Collection and Modelling.

The definition of methodologies allows for developing a research design of the project. Research design distinguishes which information might be useful and which might be potentially harmful (Mitchell & Jolley, 2012).

1.1. Design Research

Design Research (DR) is the main methodology of this work, i.e., DR drives the entire project, because it is fundamental in the development of effective solutions. DR is foundational to create products, services, and systems that respond to human needs (Lee, 2012). According to Lunenfeld (Lunenfeld & Laurel, 2003), research for design is the hardest to characterize, as its purpose is to create objects and systems that display the results of the research and prove its worth.

In DR, the methods and data collected diverge from the multiple data sources. Ethnographic approaches for participant interaction can clarify complex human needs, behaviours, and perspectives (Lee, 2012). These types of approaches are fundamental to understand the variables and the needs of critical environments such as Intensive Medicine. One of the main goals of DR is to understand and improve the design processes and working practices. This represents more than developing a specific knowledge domain in a professional field because it also encloses the environment and their stakeholders.

Table 1 presents an overview of the DR characteristics mapped with the work developed in this work. In the related work column are presented the tasks developed regarding the design research features.

Table 1 – Design Research Characteristics in the Work (adapted by (Lee, 2012))

	Design Research	*Related Work*
Main Goal	To generate value (often utility for the end User)	Improve the Decision making process through deploying a Pervasive Intelligent Decision Support System (PIDSS)
Artefact	Define the artefacts in study	A set of new knowledge presented as a Pervasive Intelligent System to Support Decision Making Process (DMP)
Process	Cumulative gathering of human experiences and artefacts that is then synthetically processed	Collect data from ICU data sources and consider the opinions of the specialists: nurses and physicians
Primary Processing Tool	Empathy	Motivate the ICU professionals to participate in the process of changing
Enables Practitioners to	Develop a solution that meets identified needs	Develop a pervasive solution to help ICU staff in the DMP

DR can be depicted as a cycle (Figure 2) initiated by a designer that find a problem and foresee an opportunity to research and then design a solution to be used by the users. Finally, the users give some feedback and the designer can use this feedback to improve the system, creating then a continuous cycle.

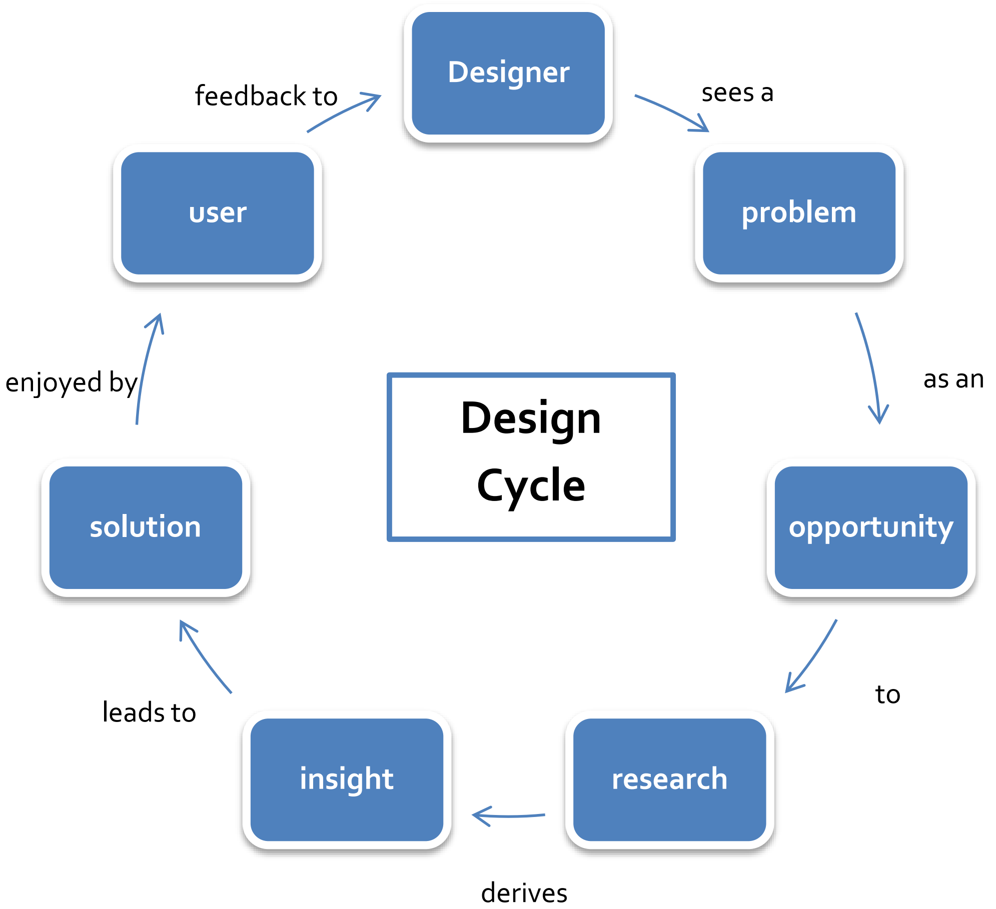


Figure 2 – Design Cycle (adapted from (Lee, 2012)).

Analysing the context of the problem, DR encompasses: the knowledge flow, the process steps and a logical formalism for, when it is facing a problem, find a solution and present a conclusion for the problem. Figure 3 represents the stages of a normal DR cycle. Both cycles presented in figure 2 and 3 were followed in this project and are presented in the articles in an implicit way.

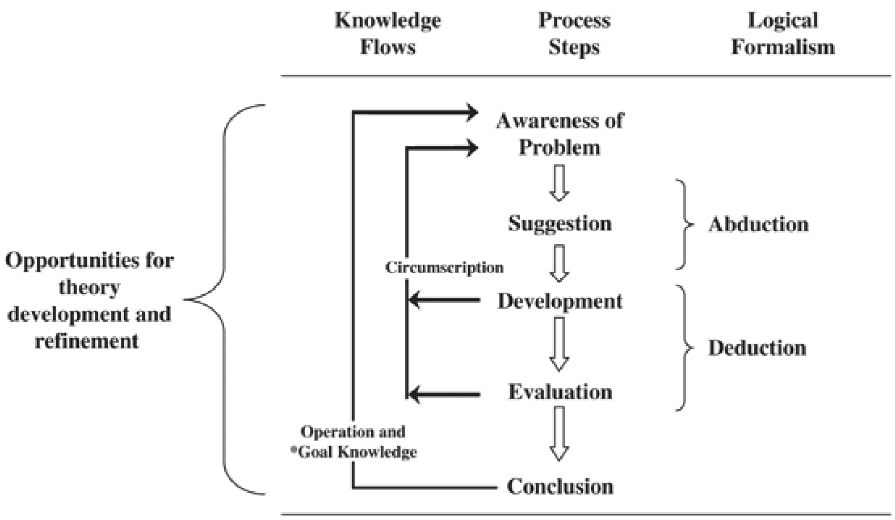


Figure 3 – Reasoning in the design research cycle (adapted from (Kuechler & Vaishnavi, 2008).

The research project can also be recognized as a project based in the action research methodology because whenever some new artefacts / suggestions were found or defined, they were immediately evaluated by the ICU staff, in order to validate the veracity and viability of the information and ideas obtained / generated. However, due to the high number of interactions, these steps aren't documented.

The results should be assessed in terms of concept proofing (proof of Concept) and user acceptance (Technology Acceptance Model).

1.2. Case Study

Case study aims to answer specific research questions and has different points of view, proofs, solutions and setting cases. The information obtained is collected and analysed to obtain, whenever possible, the best answer to the research question. In this type of methodologies, often the users are not satisfied with one answer, because they are always seeking for more studies (Gillham, 2000).

Normally, the case study is based on looking for similar cases and on the interpretation of issues and solutions adopted or, uses these cases to obtain data that are essential to solve the problem that arises within the investigation.

In this work was performed the analysis of the results already achieved by INTCare project, other similar studies, pervasive systems, information systems architectures, as well as studies and publications on the computation of medical scores, critical events, organ failure and patient outcome.

1.3. Field Experiments

This type of methodology consists of an experimenting with something on the spot where it is being tested, in this case, in a real environment (field). All results, ideas, models and architectures provided / developed were tested in the ICU of CHP before the use and the publication.

1.4. Proof of Concept

The Proof Of Concept (POC), as its name implies, aims to prove a concept, a theory, a practice. Usually, it is used in the computer sciences where the end result of a project is a new artefact. A proof of concept is also conducted in order to demonstrate the feasibility of the proposed scenarios (Bendavid, Wamba, & Lefebvre, 2006).

During this process, were used the methods that allow us to show that the artefact is efficient and fits the requirements. At the end, two questionnaires were introduced to assess the system developed, the changes introduced in the legacy systems and the new knowledge created for decision making process. The system was firstly assessed in terms of technical and functional characteristics (M. F. S. Filipe Portela, José Machado, António Abelha, José Neves, Álvaro Silva, Fernando Rua, 2012). Then, the results and user behaviour were evaluated through the use of Technology Acceptance Model III (TAM) (Chooprayoon & Fung, 2010). Four constructs were considered: Perceived Usefulness, Perceived Ease of Use, Behavioral Intention and Use Behaviour.

1.5. Questionnaires

Nowadays, it is easy to draw up questionnaires due to technological developments. A lot of artefacts are available allowing for information-gathering from mobile phones, internet, email and direct interpellation. The surveys are known to have a set of features that allows us to determine the veracity of the information collected, such as the sampling frame that contains the error and universe of respondents.

The main feature of the questionnaires is the data collection through pre-defined parameters, being this information approved by the interviewer, i.e. there is consent of both parties to the interview (Lavrakas, 2008). Samples taken should be comprehensive and representative within a given universe.

During this work two types of questionnaires were conducted: semi-structured and structured. The first one was used at the first phase to find the most suitable variables to understand the ICU staff opinions about the system and what should be improved. The second type was used to validate the system / results achieved in order to understand the User Technology Acceptance.

1.6. Data Acquisition Techniques

Such techniques helped in the process of Knowledge Discovery from Databases, as well as allowed for collecting data from a number of sources (manual or automatic): database, acquisition systems, information sources and other systems. In this case a set of agents was implemented to help the data acquisition process and a set of procedures was defined to process the data automatically.

1.7. Modelling Techniques

Modelling techniques were used to develop models for predicting the targets for the data previously collected. Prediction models were induced using Oracle Data Mining (ODM) ensemble data mining and three different techniques: Naïve Byes (NB), Decision Trees (DT) and Support Vector Machine (SVM).

1.8. Literature review

As a way to support the whole process, it was carried out a literature review. The scientific community sees in the literature review a way to find solutions to their problems. This method consists of reading several articles in order to get an idea on what has been done in the area and how the problems have been resolved. This technique was fundamental to understand the problems and contributed for the design of solutions for DMP.

In this project it was adopted the Knowledge Discovery from Database (KDD) process to obtain the new knowledge. KDD process goes through five phases: Selection, Pre-Processing, Processing, Data Mining and Interpretation (Fayyad, Piatetsky-Shapiro, & Smyth, 1996). Figure 4 presents the KDD process and the results obtained at the end of each task. The process starts from the raw data collected from the sources (databases) and culminates with the achievement of new knowledge.

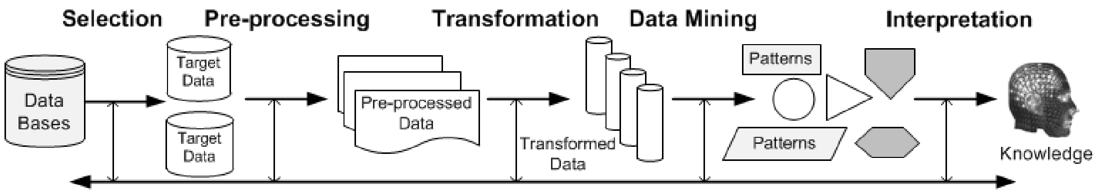


Figure 4 – KDD Process.

The Extract, Transform and Load (ETL) process involves the three initial phases and consists on the extraction of data from the sources, transformation of these data and loading the final data into the data warehouse, i.e. preparation of the data to be used by data mining algorithms.

At same time, it was used the Cross Industry Standard Process for Data Mining (CRISP-DM) methodology to guide the Data Mining project (Wirth & Hipp, 2000). CRISP-DM is composed of six phases: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation and Deployment. For each phase, a set of sub-tasks were performed. As shown in figure 5, it is possible to revisit preceding phases, i.e., jump among the phases according to the results obtained or requirements defined.

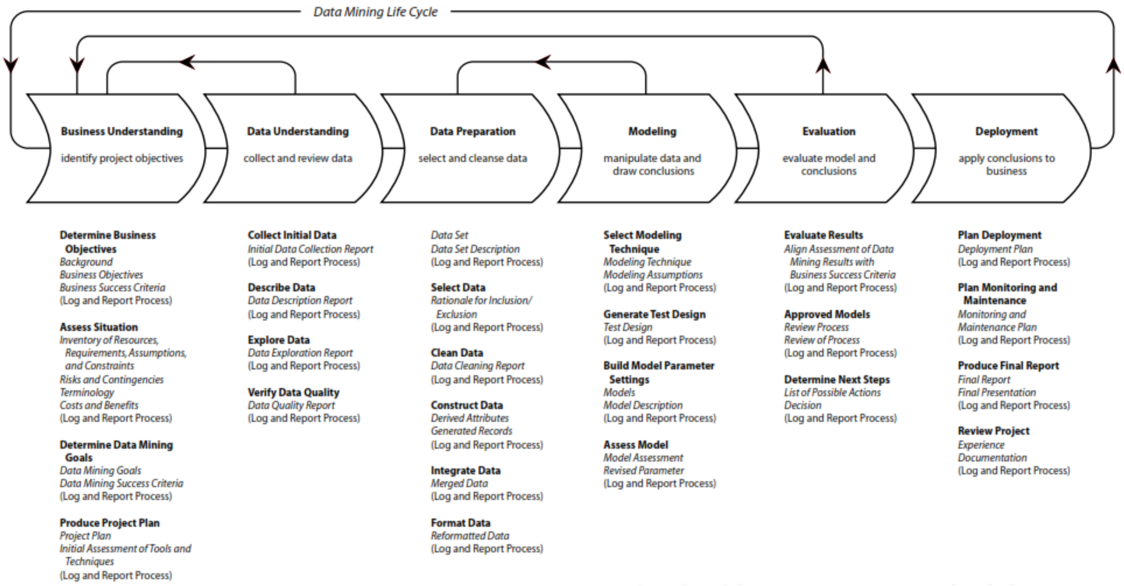


Figure 5 – CRISP-DM: Data Mining Life Cycle (adapted by (Wirth & Hipp, 2000)).

2. Characterization of Methodologies

Figure 6 shows a characterization of methodologies / techniques used by type (science or technology) and by local (laboratory or field).

As it can be observed, the research was conducted in the field (University of Minho / Intensive Care Unit, Centro Hospitalar do Porto) and encompasses research of both types: science and technology.

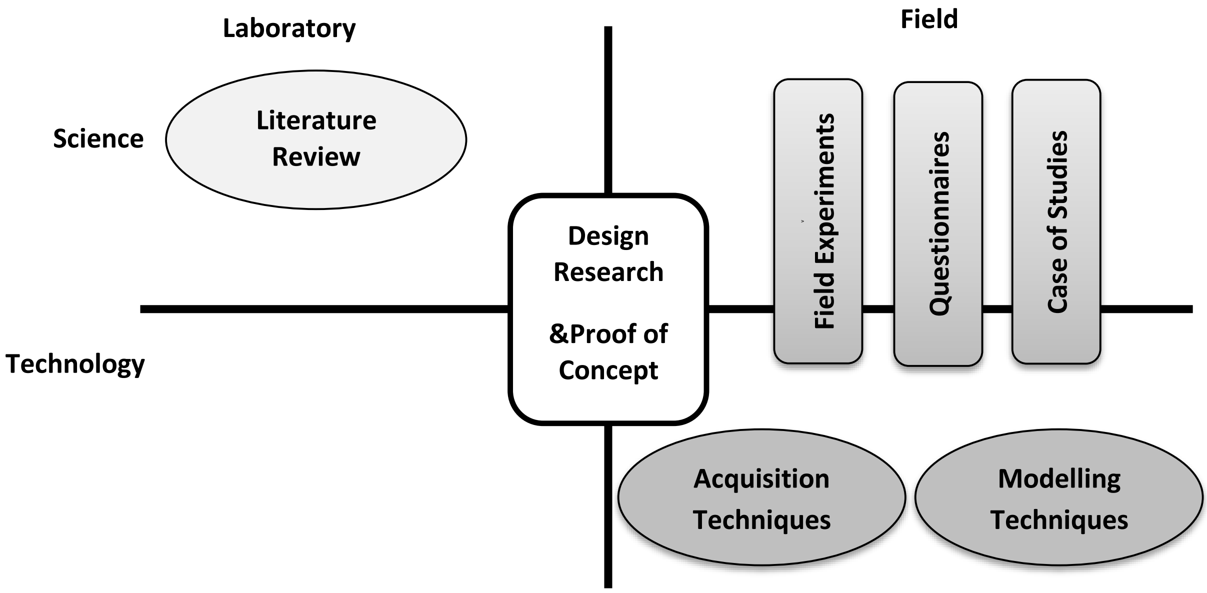


Figure 6 – Matrix of the Research Methodologies and Techniques.

Figure 7 presents the type of research conducted, the approaches taken, the data collecting and data analysis techniques used. Two different types of research were used: in the first two phases (general approach and data collection techniques) it is used the exploratory research and, then, the confirmatory research culminates in the proof of concept.

The design research is present in all the phases.

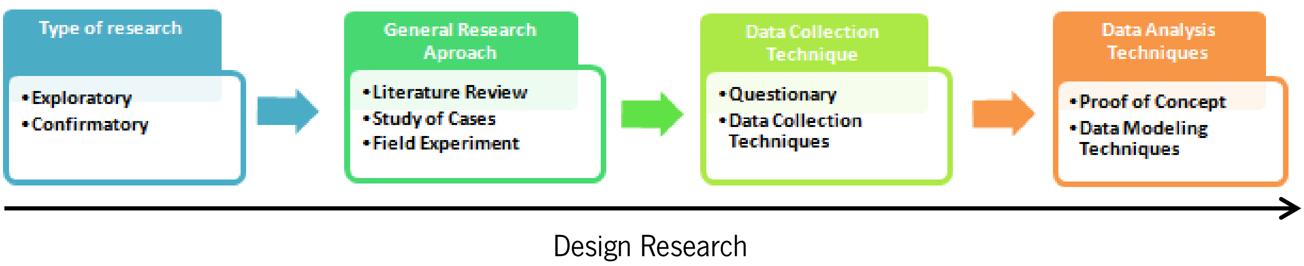


Figure 7 – Research Process.

The Literature Review and Study of Cases methodologies were fundamental to understand and define the problem and design a solution. After design a solution, some field experiments were carried out in order to validate the Information System Architecture. Consequently, a set of data collection techniques and questionnaires was used to acquire the data. Using the data collected were induced ensemble data mining models and designed some decision models in order to obtain critical events and scores.

Design Research is transversal to all phases because the main objective was to prove the concept through a system implementation and analysis of technology acceptance by the ICU users.

At the end, it was possible to confirm / validate the results obtained during the whole the project (exploratory phase), and assess the Pervasive Intelligent Decision Support in Critical Health Care by validating an artefact: a new approach to support the Decision Making Process.

CHAPTER III – Research Project

1. Introduction

This project was designed and guided by the research methodologies and techniques mentioned in the previous chapter. A research question was defined taking into account the state of the art and the requirements encountered during the INTCare project.

At the first stage and in order to answer this question, a set of objectives were identified and expected results. To achieve the proposed objectives, a research plan was developed. This research plan includes the activities and tasks to be performed and the estimated time for the completion of each one as well as the tasks to be undertaken in each semester. This plan was detailed and discussed in the thesis proposal.

For each activity, it was written a detailed description. Aiming to a better understanding of the research project, a diagram was drawn. Figure 5 presents a global overview and summary of the research process.

2. Research Question

Given the complexity of the environment and the requirements for an intelligent decision support system that integrates pervasive features and a set of requirements previously analysed, the following research question was formulated:

"To what extent, Intelligent Decision Support Systems may be suitable for critic's clinical environments in a pervasive way? "

3. Research Process

Figure 8 concentrates in a single diagram the research question, the objectives, the keywords inherent to the state of the art, the methodologies followed and the results achieved. In accordance to the objectives, a review of literature has been carried out in areas such as: real-time, critical environments, pervasiveness and health. A set of methodologies and techniques were used to support further work and to achieve the expected results. Subsequent tasks involved: environmental characterization, definition of the variables in use, studies on the implementation methods and the impact of these implementations. Complementarily, two techniques were used: data acquisition and data modelling. These techniques implicated, among others: to analyse data sources, to define new IS architecture, to induce data mining models. Finally, a methodology was used to prove the concept. A set of predefined tasks allowed us to validate some artefacts such as data acquisition system, medical scores, critical events, prediction models and the IDSS as a whole. All the methodologies were applied having in consideration the design research context.

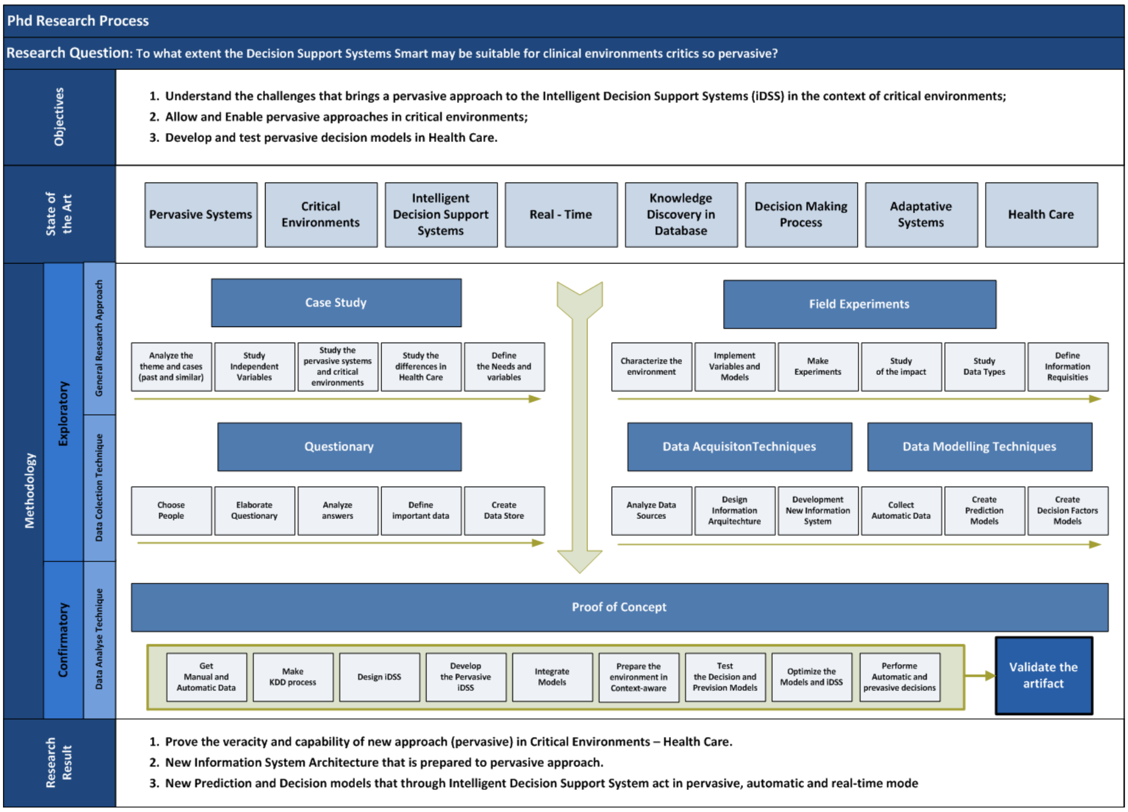


Figure 8 – Research Project.

CHAPTER IV – State Of the Art

1. Introduction

This chapter attempts to establish a position on Pervasive IDSS and Critical Health Environments. Gives a small explanation of the concepts and an overview of the most recent contributions made by the scientific community, as well as what issues are still opened.

The following scientific and technological databases were explored:

- ✓ "BioMed Central";
- ✓ "Inter Science";
- ✓ "Isi Web of Knowledge";
- ✓ "Science Direct";
- ✓ "Google Scholar";
- ✓ "PubMed";
- ✓ "SpringerLink";
- ✓ "Scopus".

The search focused on the following concepts:

Adaptive Business Intelligence, Adaptive Systems, Artificial Intelligence, Business Intelligence, Clinical Decision Support System, Critical Environments, Critical Events, Data Acquisition, Data Mining, Decision-Making, Decision Support System, Ensembles, Health Care, Intelligent Systems, Intensive Care, Intensive Medicine, INTCare, Knowledge Discovery, Medicine, Medical Scores, Online-Learning, Pervasive Health Care, Pervasive Systems, Real-Time and Ubiquitous System.

Among the various terms defined, some combinations were performed to filter information and find the best publications. Other languages have been also considered in order to extent the searching scope (e.g. Portuguese).

The selection of articles was based on the number of citations, the interest of information in the article, reputation associated to the editors/authors as well as the novelty. The articles associated to the INTCare project were those that more contributed for the state of the art.

The state of the art is presented in seven separate parts. Subchapter "Critical Health Environments" characterizes critical environments, explaining what are the main concerns and goals of Intensive Care Medicine. Subchapter "Decision Making Process" presents the most recent developments of the Decision-Making Process (DMP) in medicine. Next subchapter complements the previous one, introducing the concept of "Decision Support System" (DSS). The following chapter addresses the development of "Intelligent Decision Support Systems" (IDSS) and Clinical Decision Support Systems (CDSS). Subchapter "Pervasive Health Care" presents an outlook on pervasive health and prospective challenges. Subchapter "Ubiquity and Pervasiveness" presents the requirements to assure a pervasive environment. Finally, the chapter is summarized eliciting the results achieved in the project INTCare.

2. Critical Health Environments

Critical environments in health are characterized for being a place where the decision process is fundamental to the success of an operation. In this type of places failures are not allowed because they can result in a life lost. Critical environments demand for decisions in real-time, which require efficient, secure and ubiquitous processes. This kind of environments includes, for instance, intensive care units, operation rooms, emergency rooms, decision rooms and services where decision making can jeopardize a number of factors due to the urgency of action, and delicacy and type of the variables used in those situations. Typically, they deal with critical variables which are in constant change.

Intensive Medicine (IM) is recognized as a critical area. IM appeared in the century XX and, according to Silva (Álvaro Silva, Cortez, Santos, Gomes, & Neves, 2008), it can be defined as a multidisciplinary field of Medical Sciences that specifically addresses the prevention, diagnosis and treatment of acute potentially reversible situations in patients with failure of one or more vital functions.

Intensive Care Units (ICUs) are recognized as a critical environment because they are concerned with these types of patients and they focus their efforts on the resuscitation of patients who are in terminally ill. At the same time they are focused in the treatment of patients who are vulnerable to dysfunction, benefiting the preventive care for each system dysfunction in accordance to the principles of restoring of the normal physiology (Hall, et al., 2005).

In ICUs the act of treating a patient requires continuous monitoring of the clinical condition by the nursing staff. A nursing sheet is commonly used to record all therapeutics, serums and treatments administered to the patient as well as the fluid balance. The activities performed in this type of environment are complex because the various organ systems can be affected at the same time (Apostolakos & Papadakos, 2001). In ICUs therapeutic goals are established and achieved through the formulation and testing of clinical hypotheses. These therapeutic goals should always to show the relationship between the intensity of treatment and the benefit to the patient. The use of new treatments requires the execution of rigorous clinical tests, whereas traditional therapies require the clarifying of the objectives and the adverse effects on the patient (Hall, et al., 2005).

The treatments that are applied daily in intensive care medicine can be of two types: treatments that lead to cure the patient or palliative treatment, which has as main objective the alleviation of the suffering of the patient.

According to Hall (2005), the decision-making process contributes more to the clinical outcome of a patient than to innovative interventions or technologies. As it can be proved with this assertion, the decision-making process is important for critical care medicine.

This process may be more effective when supported by IDSS technology with a high degree of accuracy and high level of pervasiveness. In this sense, it is necessary to continue the technological development that began in medicine and that has proliferated especially in intensive care medicine.

3. Decision Making Process

The Decision Making Process (DMP) is a key point in critical areas such as critical care medicine, in both diagnosis and assessment of prognosis as well in the treatment of the patient (Á. Silva, 2007). According to some studies, medical error is the eighth leading cause of death in industrialized countries (Kohn, Corrigan, & Donaldson, 2000). The development of an IDSS that helps in the decision process, giving the right information at the right moment by the use of models to predict the organ failure and patient outcome, can reduce the error.

In this sense, it is urgent to conceive IDSS that, using data mining models, are able to predict an event before its occurrence. At the same time they can help physicians deciding on what to do in order to prevent this event. In 2009 only one IDSS, using an offline approach and data mining models, was able to predict organ failure and outcome in intensive care. This system resulted from the INTCare research project (Gago, et al., 2006; Santos M.F. & J., 2009b; M. F. Santos, Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., Neves, J., 2009a; M. F. V.-B. Santos, Marta. Portela, Filipe. Silva, Álvaro. Rua, Fernando, 2010). INTCare was tested in the ICU of Centro Hospitalar do Porto (CHP).

Decision-making process can be approached using various models: Simon, political process, the container and the McGrath (Marreiros, 2007). One of those models was proposed by one of the pioneers of Artificial Intelligence and Nobel Prize in Economics (Marreiros, 2007). Herbert Simon suggested he's own model, bounded rationality, where he says that people act rationally on the basis of knowledge and insights they get (Simon, 1960). This model is one of the most used in the scientific community. Simon's model is based on three stages: intelligence, design and choice. A few years later, this model has changed and a new phase has been added – the implementation – by Sprague Jr. and Carlson (Sprague Jr & Carlson, 1982). This model does not include negotiation processes, so it aims to find a solution, and the solution is generally accepted by all members of a group (Marreiros, 2007). Although not including a negotiation process, at appropriate situations, all stakeholders can give their contribution to the process of making the final decision, without ever jeopardizing the idea initially set.

More recently, in 2005 and according to Efraim Turban (Efraim Turban, Aronson, & Liang, 2005), DMP was defined as a five phases process, so long as the traditional process must include a post-deployment. Starting from this definition, the process can be divided into four initial phases (Intelligence, Design, Choice and Implementation) more a fifth phase called monitoring. These phases can facilitate the interpretation and application of models in a real context. All the phases of DMP are fundamental to the development of an IDSS.

Addressing the interpretation of each phase defined by Turban (Efraim Turban, Sharda, & Delen, 2010), the first phase - "intelligence" - allows to study the environment in a continuous or intermittent way and identifies problems and opportunities. An adaptive perspective also allows for evaluating the results of the implementation phase (monitoring). The main tasks are:

- ✓ Identification of the problem;
- ✓ Ranking the problem;
- ✓ Decomposition of the problem;
- ✓ Define the owner of the problem.

The second stage – "design" – consists of finding, developing and analysing possible courses of action. This phase includes the perception of the problem and experimenting of the solutions in order to evaluate their feasibility (Efraim Turban, et al., 2010). This phase is the most important due to its decisive factor in understanding the problem and determining solutions. This phase includes the following main tasks:

- ✓ Explore the principle of choice;
- ✓ Create normative and descriptive models;
- ✓ Introduce the sub-optimization;
- ✓ Design of alternatives;
- ✓ Measure the viability of solutions;
- ✓ Define risk assessment;
- ✓ Develop scenarios for the solutions;
- ✓ Analyse the errors presented in DMP.

The third stage – "Choose" – is a critical stage because is where, after defining the problem and developing solutions, the optimal solution is chosen. In this sense, it is necessary to seek the most appropriate course of action (Efraim Turban, et al., 2010). This phase includes activities such as looking for reviews and recommendations for an appropriate solution for the model.

The fourth phase – "Implementation" – is the final stage of the process itself, and is the stage at which the solution is implemented. The solution can be as simple as a change in the behaviour, or the introduction of major changes both inside (inter-partnership) and outside (enter-partnership) of the organization, in the people, environment or technicians.

In 2010 Turban (Efraim Turban, et al., 2010) proposed a fifth phase: "Monitoring". In this sense the phase "intelligence" is also used as a way to monitor, control and evaluate the solutions. Thus the decision-making process becomes an iterative process. INTCare adopted the decision process in accordance to the four phases-based model of Turban (Gago, et al., 2006).

Currently, in the most of ICUs, physicians analyse the available clinical data and, relying on their experience, they decide whether or not to intercede. The information usually comes from bedside monitors or from observations recorded periodically. This information is used to find out which action should be taken and what path should be followed. However, sometimes the information provided is not enough, is not in the most adequate format or not arrive in the right time for the decision.

Actually, INTCare is the unique system in the medical environment that seeks to correct the gaps found in ICU and oriented by DMP. Main gaps in ICU include the insufficient provision of information in real-time and the out-of-time response to the problems.

This system is being deployed and aims to provide information in a high level of confidence. Supplying relevant information about the clinical condition of patients, physicians can improve the DMP. The information is available in the electronic format (online) and in real-time (Filipe Portela et al., 2010) and the results are disseminated through the ICU devices.

Out of the medical field there are some important IDSS projects such is the case of systems of intangible investment plans (Kivijarvi & Tuominen, 1999) and money laundering (Shijia Gao & Dongming Xu, 2009). In these projects it is possible to observe that the DSS can benefit from the application of the concepts and methodologies from Business Intelligence (BI) in conjunction with the creation of decision models. Despite of the lack of knowledge observed in the real-time systems described above, the IDSS already represents an improvement for the systems. Being the IDSS based on techniques of DM it don't only provide a solution, but also can integrate decision models automatically and in real-time.

Daily, ICU physicians have to make important decisions to the future of the patient, elaborating diagnosis and planning treatments. In the 70s, the decision on the treatment of a patient was performed based only on the evaluation of the patient's symptoms, discarding the differences obtained by the medical diagnosis (Coombs, 2003). Nowadays, decisions take into account patients' symptoms and medical diagnoses.

The toughest decisions are those that are related to the life support so, in this case, it is fundamental a correct exam of the more sensitive issues, because a small mistake can result in worsening of the patient's life. Discontinuation of a treatment is a controversial issue in ICUs, since it is now possible to extend life for long periods of time with no hope of recovery (McMillen, 2008). These types of decisions are an example of decisions that can be taken in ICUs and require a complete knowledge base to help the staff to perform the best choice.

As it has been described, one of the problems in the DMP is the time to decide. According to the nurses, the concept of time has emerged as an important factor in the decision. In tasks such as to withdraw or change of the treatment, the right time is not only important for the patient and their family, but also for the nurses at professional level (McMillen, 2008). At this point the presence of an intelligent system increases the capacity and promptness of the action. In addition, sometimes the limitation of resources in intensive care units forces the staff to make decisions assuring that the treatments are only applied to people who can benefit from them (Gago, et al., 2006).

Therefore and due to lack of some information, critical decisions such as discontinuation of life-sustaining treatments and non-prescription orders of resuscitation are considered futile (Gago, et al., 2006). Proving this idea, it is one of the maxims of the Intensive Medicine: first it is the patient care and only then the patient documentation (Baggs et al., 2007; Saarinen & Aho, 2005; Tang, Mazabob, Weavind, Thomas, & Johnson, 2006). This idea is correct and is the most assertive. However, it blocks any type of development of an IDSS, unless the process could be computerized and automated. This is the main challenge to introduce IDSS technology and, consequently, change the DMP. To fortify this fact, it is important to supply physicians with a set of complementary information that may help them deciding, regardless of where and who provides the information.

The quality of this information can determine what to do with a patient or a specific treatment. For modelling the process of decision-making it can therefore be concluded that the information and opinions of nurses should never be set aside and they should be complementary to the doctors. Only then it is possible to develop an intelligent system that is capable of supporting the most critical moments in the decision. The aim is to create a universal DMP based on a pervasive IDSS and able to assist physicians enabling them to act in a ubiquitous and continuous way.

In order to help the decision process physicians use information provided by some medical scores measured manually and adverse / critical events. The most used scores are SOFA, SAPS II and Glasgow.

Sequential Organ Failure Assessment score (SOFA) is used to describe the severity of the organ failures (0 to 4 (worst level)). The score is based on six different results, one for each organ failure – respiratory, cardiovascular, hepatic, coagulation, renal and neurological systems (J. L. Vincent et al., 1996) – resulting in a global score to all organs.

Simplified Acute Physiology Score (SAPS) is a severity of disease classification system to predict mortality using logistic regression (Le Gall, Lemeshow, & Saulnier, 1993; Metnitz et al., 2005). In ICU, the most widely used score is SAPS II (Le Gall, et al., 1993), however was recently introduced SAPS III (Metnitz, et al., 2005; Moreno et al., 2005).

Glasgow Coma Scale (GSC) is a neurological scale that aims to understand the conscious state of a person in three types of response – eye (4 levels), verbal (5 levels) and motor (6 levels) (Jones, 1979) – where 3 indicates deep unconsciousness and 15 indicates totally awake.

Until the beginning of this work there were no systems to collect and process data in real-time in order to obtain the respective ratios automatically. Normally, the ICU staff use pre-defined sheets to fill manually and to calculate the scores. The introduction of IDSS implies that all calculations should be done in an automatic way and that the results should be provided online and in real-time.

Finally, there are the critical events. These events are defined in the medical community as a quantity of time where the patient presents values out of the normal range to some variables. Nevertheless, they are not yet used in the ICU due to the difficulty in obtaining the results manually. The introduction of an intelligent systems can use the automatic data acquisition to track the patient critical events in real-time.

4. Decision Support Systems

Former Decision Support Systems (DSS) appeared in the 1970s (Vahidov, 2002) and were based in decision-making models. These systems analyse a large set of variables which allow for answering a particular question. Defined as "a computer-based system that aids the process of decision making" (Finlay, 1994), they are also "interactive, flexible, and adaptable computer-based information systems, especially developed for supporting the solution of a non-structured management problem for improved decision making. They utilize data, provide an easy-to-use interface, and allow for the decision maker's own insights" (Efraim Turban, et al., 2005). These kind of systems can be grouped in two types (Arnott & Pervan, 2004): the first one involves the rule-based system, and the second one usually uses Artificial Intelligence techniques (IA) – neuronal networks, genetic algorithms and fuzzy logic (Efraim Turban, et al., 2005).

DSS are a particular sort of information systems and are focused in the improving of the support of the decision-making process (Arnott & Pervan, 2004).

4.1. Intelligent Decision Support Systems

Intelligent Decision Support Systems (IDSS) for health are mainly based in the Clinical Decision Support Systems (CDSS). CDSS are characterized as computer systems that have an impact on the patient particularly in the moment when the decisions are taken (Berner & La Lande, 2007). CDSS have the following main objectives: creating alerts and reminders, assistance to the diagnosis, treatment planning, decision support on prescriptions, information retrieval and image recognition and interpretation (Coiera, 2003). Figure 9 shows a typical architecture of a CDSS.

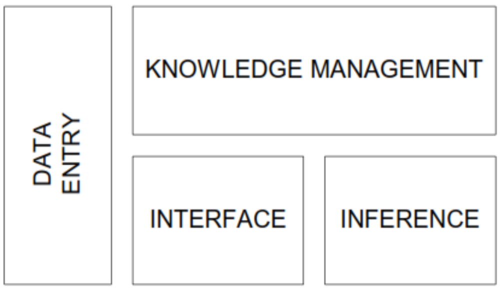


Figure 9 – CDSS architecture (Gago & Santos, 2009).

In the case of INTCare, data input are extremely important for the correct operation of the system. These data come from the Vital Signs Monitors and Nursing Records (M. F. Santos, Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., Neves, J., 2009). The Knowledge Management component retains the prediction models and also includes the evaluation of the predictions performed. The inference component responds to user requests made by the interface (P. Gago & M. F. Santos, 2009). The development of an information model for IDSS requires a particular set of characteristics for the system. It should be developed with the purpose of filling the gaps previously identified, which are essential for the good execution of the system.

4.2. Characteristics

In the work developed by Filipe Portela (Filipe Portela, 2009) during the Master of Engineering and Management of Information Systems dedicated to the theme "Decision Support Systems in Intensive Care Medicine based on Knowledge Discovery in Database", some features were identified as being essential for the development of an IDSS:

a) **Online-Learning:** The system must have the ability to act online, i.e., the models must be induced using online data;
b) **Real-Time:** The system must be able to act in real-time. The process of acquisition, transformation and storage of the data must occur immediately after the data is available. Moreover, the decisions should be taken at a time when certain event occurs;
c) **Adaptability:** The system needs to have the ability to automatically optimize the models considering the most recent data. This information is normally obtained after evaluating stored results;
d) **Prediction Models:** The success of an IDSS depends on the accuracy of the data mining models, i.e. the degree of reliability of these models should be high. These models make possible the prediction of events and consequently prevent certain medical complications for patients
e) **Decision Models:** The achievement of the best results is strongly dependent on the decision models developed. These models are based on several factors such as the decision and differentiation that are applied in the prediction models. Then, the models can help the physicians to choose what is the best decision during the DMP;
f) **Intelligent Agents:** This type of agents helps the system to work through autonomous actions that perform essential tasks. These tasks support some of the INTCare system modules: Data Acquisition, Knowledge Management, Inference and Interface. The flexibility and effectiveness of such systems depend on the agents and the interactions among them. An agent is an intelligent program that acts autonomously on behalf of the user. Multi-agent Systems (MAS) are composed by a number of agents that communicate with each other

and work together, toward a common goals, with a degree of reactivity and either reasoning (Wooldridge, 1999).

4.3. IDSSs in the Medicine

The conception of DSS that operate in real-time and in critical environments is a constant challenge, since they typically use a set of critical variables that are varying constantly. In this case, there are no margins to fail, i.e., the system only should be deployed in the real environment after its proper functioning and after the execution of a high number of tests. In this point, it is central to the proper functioning of the IDSS the communication between the various systems already present in the service (interoperability). Systems such as data acquisition, electronic health record, drugs circuit and laboratory system must be connected through the interoperability of systems and operate together in order to be useful to the IDSS.

The use of Artificial Intelligence (AI) algorithms (e.g. Support Vector Machines (SVM), Neural Networks, Fuzzy Logic, Ensembles) and intelligent agents can contribute for a faster defining the problems to be studied and then find, in the right time, the best solutions for them. The use of agents to support the decision-making process is important in the medical industry. They allow physicians and nurses for processing and collecting the information quickly and in several ways (prepared according the goal – monitoring vital signs, present the lab results, control the fluid balance, etc.) in order to assist them in decision-making about diagnoses and treatments (Foster, McGregor, & El-Masri, 2005).

In critical care medicine, there are some IDSS with different goals that help decision-making such as the Guardian (Larsson & Hayes-Roth, 1998). Through an independent agent and cooperation among various algorithms, Guardian produces diagnosis and treatment plans for action in real-time. Although this system is acting in real-time, the knowledge used to produce the results were before stored (offline and manual) in the database. Analysing a study about systems with temporal abstraction modules published in 2007 (Stacey & McGregor, 2007), is possible to find some systems that use data in real-time (streaming data), while others use data from offline databases. However, none of them follows a structured DMP based in the four stages of a decision-making process (intelligence, design, choice and implementation) (Efraim Turban, et al., 2005).

Among the systems found, the most notable is an intelligent system for the management of ventilation and oxygenation (Belal, Taktak, Nevill, & Spencer, 2005). This system uses real-time monitored data, the laboratory results, the variations of the variables SaO2, and tcpO2 tcpCO and the existing events to define the limits for them. Whenever these limits are exceeded, it sends alerts. Among the events, it is defined for example the minutes that the system must wait after changes in each variable or after any aspiration procedure or after a change in the patient clinical condition. Only after this timeout, the system makes recommendations and sends alerts. In addition to these constraints, the user can also set extra configurations such as launching diagnostics and changing the mode of ventilation, which influences the manner how alerts are sent.

Another available system, which is based on the use of streaming data, proposes a new approach for the recognition of cardiac arrhythmias in electrocardiograms (ECG) (Carrault, Cordier, Quiniou, & Wang, 2003). ECG signals reflect the changes in tension occurring in the body surface when the heart muscle is triggered by electrical events (Carrault, et al., 2003). The purpose of the system is to characterize the arrhythmias from a set of representative elements of ECG tests. The classification rule of an arrhythmia is based on twelve derivations present in classic ECG (Carrault, et al., 2003). The derivations are usually collected by the cardiologist. In this case, the derivations start being monitored by the system, which together with pre-defined rules can recognize the presence of an arrhythmia. The algorithms of the system are ready to process the data changes and to produce new sets of rules based in the information available that can be provided by other sensors. This system takes advantage in terms of signal processing due to continuous monitoring.

On the basis of the information provided, is possible to conclude that both systems are intelligent because they have the capability to automatically analyse the values collected, i.e., they collect the data in real-time and use the values obtained to decide, in real-time, what to do. However, since both the intelligence are based on rules, they only indicate a possible solution to the problem, not addressing the effects that this may have in the future nor presenting other ways to solve the problem.

The oxygenation system should take into account other variables. Decision models, commonly used to send alerts and then give directions to the doctor, should be present in a pervasive system with an intuitive interface. In the case of arrhythmias, the system should be able to recognize an arrhythmia and allow for the determination of the origin of the arrhythmia.

Despite the progress achieved, none of the systems presented can operate in real-time and ubiquitously. Because of this, it is essential to provide medical information through ubiquitous devices (U. Varshney, 2007a), avoiding many medical errors that occur due to the lack of accurate information on the location and the time at which is required, as for example misdiagnosis (Kohn, et al., 2000; Makeham M, 2002).

The lack of accurate information can lead to decision errors in 50% of the cases (Bergs, Rutten, Tadros, Krijnen, & Schipper, 2005). Ubiquitous electronic medical records enhances the analysis by authorized users, anytime and anywhere (U. Varshney, 2009).

Time granularity and retrieval of information play an important role in the models to predict automatically and in real-time diagnostics, prognostics, and treatments (Augusto, 2005). Pervasive systems should be focused in the creation of new and important knowledge for the decision process and make it available in the right place at the right time.

5. Pervasive HealthCare

"Pervasive HealthCare" (PH), according to Varshney (U. Varshney, 2009), can be defined as "conceptual system of providing healthcare to anyone, at any time, and anywhere by removing restraints of time and location while increasing both the coverage and the quality of healthcare". This approach is based on information that is stored and available online (Mikkonen, Vyrynen, Ikonen, & Heikkila, 2002). However, although the PH has the potential to reduce costs, improve service quality and facilitate the treatment to the patient, it also faces many technical and administrative obstacles (Upkar Varshney, 2003), such as resistance to change and significant changes in technology and systems.

Figure 10 presents a possible scenario of operations in a context of "Pervasive healthcare." In this illustration it can be observed that mobile devices are central to the work because they use network access of the hospital to communicate with local services, check the patient's vital signs, collect information and implement decisions by using telemedicine.

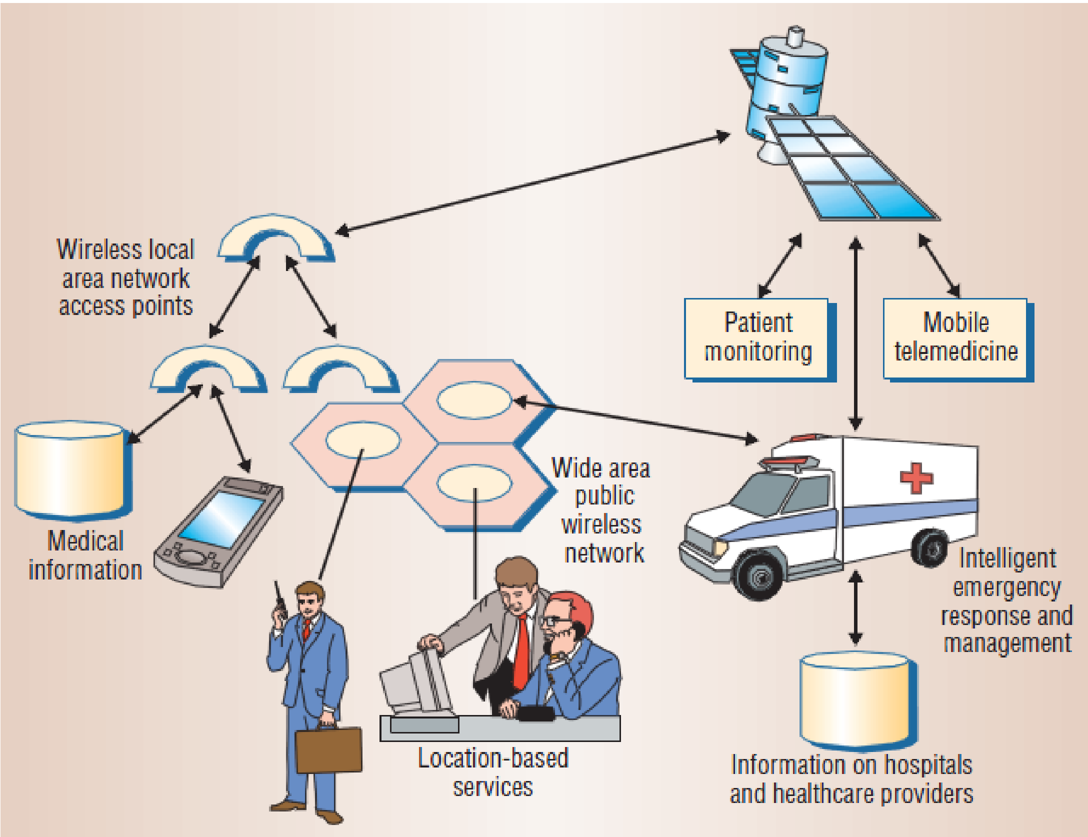


Figure 10 – Pervasive Healthcare Scenario (Upkar Varshney, 2003).

The main difficulty to implement PH is that the information is not always available when it is needed. Often, this happens because there is a lot of information on paper. This limitation prevents physicians take the best decision for the patient. As a solution, the indispensable information may be available electronically and complemented by predictive models that can help the physician making the best decision in real-time.

The existence of electronic medical records in an ubiquitous way can change the paradigm making the information available to be analysed by authorized persons, anytime and anywhere (U. Varshney, 2009). The information returned by the operational systems should be stored and accessible from a website and from portable devices (Mikkonen, et al., 2002). As an example of applications that can be developed for this type of environments, we have: universal systems for monitoring clinical data, intelligent emergency management, universal access to clinical data and mobile and ubiquitous telemedicine (U. Varshney, 2007a).

A pervasive IDSS can fill most of these gaps and help in the coordination of several activities that may be important to the clinical condition of the patients, as the accurate diagnosis and the realization of appropriate procedures (Scicluna, Murray, Xiao, & Mackenzie, 2008). The system can be based in models (with real data) of prediction and decision and accessed anywhere. Agent paradigm may also be an important contribution in this area, considering multiple agents that work together in order to find answers to the problems (Foster, et al., 2005; Machado, 2006).

6. Ubiquity and Pervasiveness

6.1. Ubiquity

The technology of ubiquitous computing aims to improve the living and working environments including home, office, city developing advanced devices, network infrastructure and operating systems. Researchers focus on the development of so-called "smart environments", enabling the appearance of ubiquitous technology environments (Kwon, et al., 2005).

Ubiquitous computing is viewed as a third way of computing (Weiser, 1991), which has two views: (1) reduce the need for users in concentrating when they are interacting with computing devices and (2) provide the computing power anytime and anywhere (Lupiana, O'Driscoll, & Mtenzi, 2009). The goal is to give to these users the possibility to use computers and, in an unconscious way, support their daily activities (Streitz & Nixon, 2005). The ubiquitous computing environment is emerging as a primary controller to change the environment that the makers of the tasks operate, and to reshape the way of we do business, creating a new era (Kwon, et al., 2005). Figure 11 shows the characteristics of a ubiquitous DSS (ubiDSS).

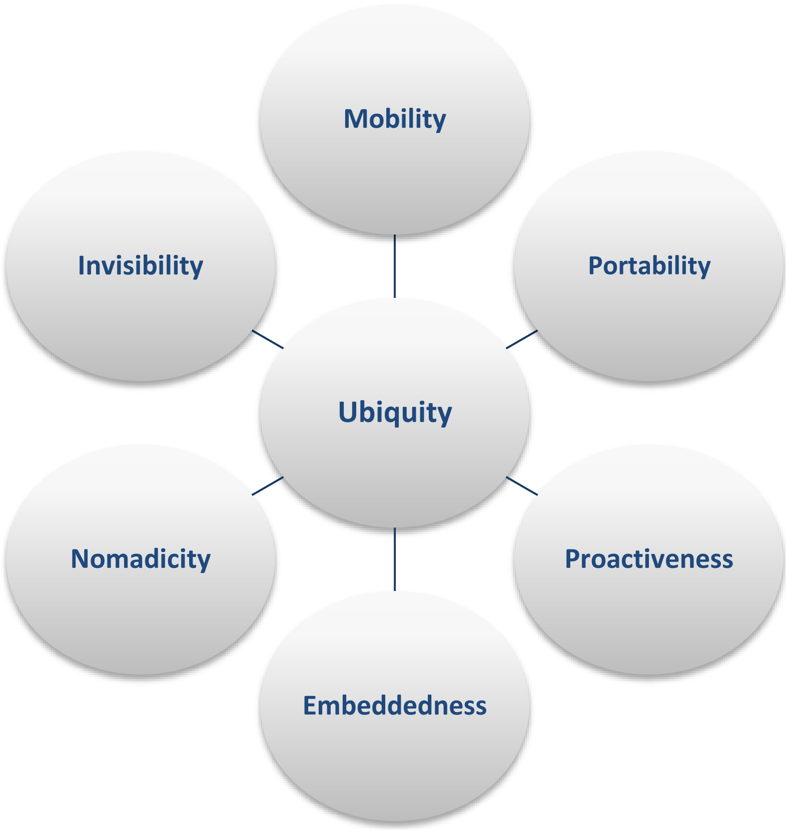


Figure 11 – Capabilities enabled by the ubiquity in ubiDSS. (Kwon, et al., 2005).

6.2. Ubiquitous Data Mining

Ubiquitous Data Mining (DMU) is mainly based on Data Mining (DM) and may be defined as the process of selection, exploration and modelling of large amounts of data to discover unknown patterns or relationships that provide a clear and useful data analysis (Giudici, 2003) using ubiquity characteristics. The abundance of mobile devices, such as PDAs and cell phones, along with the progress in wireless communications, made possible the achievement of Mobile Data Mining, which is the analysis and mining of data streams from portable devices significantly more viable (Horovitz, Gaber, & Krishnaswamy, 2005). Consequently, considerable efforts have focused on research and development in this area. Alternative approaches include the execution of data mining in a web server and the diffusion of the results in ubiquitous systems. Mobile devices can consult and execute the models stored in the server.

DMU is a process flow that uses DM ubiquitous devices in an environment of limited resources (Horovitz, et al., 2005), i.e., the analysis of data streams to discover useful knowledge on moving parts, and inserted into ubiquitous devices (Gaber, Krishnaswamy, & Zaslavsky, 2004). It is an emerging focus on DM, encouraged and supported by the steady growth of computational abilities of mobile devices. This results in systems that are capable of performing intelligent data analysis and critical moments on continuous streams of data, thereby simplifying the pervasive paradigm "anywhere and anytime" (Horovitz, et al., 2005).

Data Stream Mining (DSM) consists in the extracting of knowledge (e.g. patterns, clusters, forecasts) from information flows through the application of techniques such as clustering, classification, frequency count and time series analysis (Goñi et al., 2009). The main functions performed by these monitoring systems in medicine are: (1) acquisition of data from biological sensors, (2) the data analysis to detect certain anomalies, and if they are detected, then (3) generation of alarm notifications to the physicians so they can make the right decision (Goñi, et al., 2009).

6.3. Pervasiveness

According to Saha (Saha & Mukherjee, 2003), “Most computing systems and devices today cannot sense their environments and therefore cannot make timely, context-sensitive decisions”. However, pervasive computing requires systems and devices that recognize the context. Mobile computing addresses location and mobility-management issues but in a reactive context, responding to discrete events. Pervasive computing is more complex because it is proactive. Intelligent environments are a prerequisite to pervasive computing. Perception, or context-awareness, is an intrinsic characteristic of intelligent environments. Implementing perception introduces significant complications: location monitoring, uncertainty modelling, real-time information processing, and merging data from multiple and possibly disagreeing sensors. The information that defines context awareness must be accurate; otherwise, it can confuse or interfere on the user experience.

Pervasive computing is about making our lives simpler through digital environments that are sensitive, adaptive, and responsive to human needs. Far more than mobile computing, this technology will fundamentally change the nature of computing, allowing most objects we encounter in daily life to be “aware,” interacting with users in both the physical and virtual worlds. While research challenges remain in all areas of pervasive computing, all the basic component technologies exist today.

Satyanarayanan (Satyanarayanan, 2002) characterizes the universal computation (pervasive) as an evolutionary step resulting from previous two steps: first distributed computing and then mobile computing (Figure 12). According to the author, research involves not only issues associated with the communication between the environments and their interaction with users, but also issues related to the supporting mobility of the users.

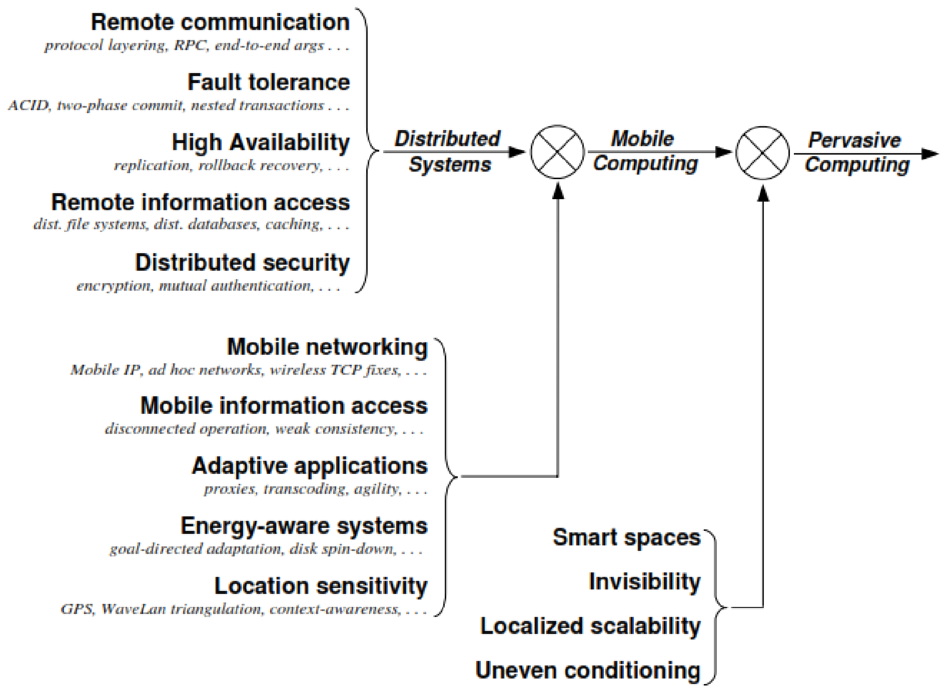


Figure 12 – Pervasive computing taxonomy (Satyanarayanan, 2001, 2002).

This type of computing has evolved with the emergence of the Internet, a valuable new tool essential to the global information system, not only being widely adopted by organizations, but also by people. Nowadays, the internet is easily accessible in all developed countries, schools, public organizations, at home, inside or outside of building. This new reality promotes the introduction of pervasive computing.

The appearance and ease of access networks and wireless communication systems contributed to its dissemination. The incorporation of computing devices, objects or monitoring sites provides a glimpse of a world real, physical and enhanced with information and computing resources that can be used to facilitate human life in its various tasks (personal or social) or to improve business or organizational processes (Banavar et al., 2000).

Banavar (Banavar, et al., 2000) further believes that pervasive computing can be about the following: (i) how people view and use mobile computing devices to perform tasks, (ii) how applications are developed and implemented in order to support these tasks, (iii) how the environment is reinforced by the emergence and ubiquity of new information and features.

In an economic context applied to health, according to Varshney (Upkar Varshney, 2003), the Pervasive computing can improve productivity of healthcare professionals and facilitate the provision of a wide range of medical services. The increasing use of portable devices and deploying wireless solutions will accelerate the development of such services, especially in areas where there are good infrastructures and also where the infrastructures are impractical, improving them. Faster and more accurate communication results in substantial gains that could be used to expand basic health care for all and reduce long-term costs.

For a system to be pervasive, it needs to achieve some features (Henricksen, Indulska, & Rakotonirainy, 2001; Orwat, Graefe, & Faulwasser, 2008; Saha & Mukherjee, 2003):

Scalability - Future pervasive computing environments will likely face a proliferation of users, applications, networked devices, and their interactions on a scale never experienced before. As environmental smart ness grows, so will the number of devices connected to the environment and the intensity of human-machine interactions;

Heterogeneity - Conversion from one domain to another is integral to computing and communication. Assuming that uniform and compatible implementations of smart environments are not achievable, pervasive computing must find ways to mask this heterogeneity—or uneven conditioning as it has been called— from users;

Integration - Though pervasive computing components are already deployed in many environments, integrating them into a single platform is still a research problem;

Invisibility - A system that requires minimal human intervention offers a reasonable approximation of invisibility;

Context awareness – Once a pervasive computing system can perceives the current context, it must have the means of using its perceptions effectively. Richer interactions with users will require a deeper understanding of the physical space. Smartness involves accurate sensing (input) followed by intelligent control or action (output) between two worlds, namely, machine and human.

Verifying that such systems are beneficial to those who use them, now it is necessary to prove that the use of real-time data in Intensive Medicine can influence positively the decision-making process. Making the DMP more efficient, it is possible to use models to support real-time decision and to take action in time in order to solve the problem independently the local where physicians are.

In this work it is important to reinforce the idea that the system is also working with pervasive information, i.e., information are collected, transformed and made available automatically and in real-time.

7. INTCare

The INTCare project (FCT PTDC/EIA/72819/2006) seeks to develop an Intelligent Decision Support System (IDSS) for Intensive Medicine (IM) (M. F. Santos, Cortez, Gago, Silva, & Rua, 2006). The aim of INTCare is to create predictive models to support physicians during the process of decision making, taking into account information about the probability associated to mortality and organ failure.

7.1. INTCare System

Figure 13 presents an overview of INTCare System. The system is composed by four subsystems: Data Entry, Knowledge Management, Inference and Interface. The system uses autonomous intelligent agents (M. F. Santos, et al., 2006) to perform their tasks, without the need of human intervention. These agents are responsible by a set of activities such as data acquisition, models' processing and data mining.

These data are collected from Vital Signs Monitors (MSV) and from the nursing records. The data are stored in a local database. Then the Data Mining models use these data to predict future scenarios interacting with the physicians through an interface (M. F. Santos, et al., 2006). The data are then applied to the prediction of organ failure and treatment outcome. Until 2009 only the data entry subsystem was explored.

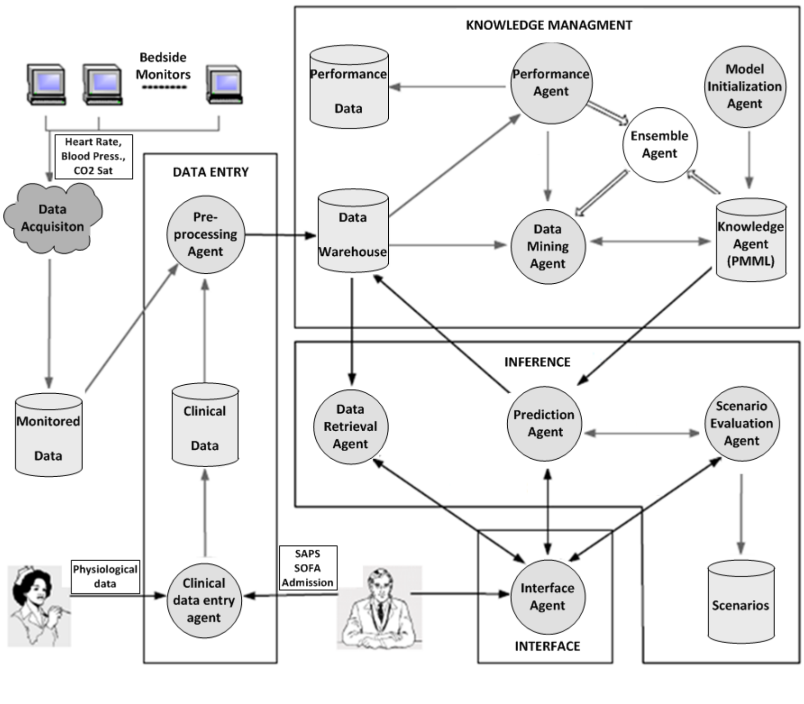


Figure 13 – INTCare System Architecture.

7.2. Previous achieved results

Previous results include:

- The design of the whole architecture of the data collection and information system (M. F. Santos, Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., Neves, J., 2009), (Santos M. & Abelha A., 2009; Santos M.F. & J., 2009);
- The introduction of new methods for data collection (electronic nursing sheet) that privileges the dematerialization of information, putting as much data as possible in electronic format (Filipe Portela, Vilas-Boas, Santos, & Fernando, 2010; Santos M.F. & J., 2009b).

Preliminary results were obtained using offline learning (Vilas-Boas, Santos, Portela, Silva, & Rua, 2010a, 2010b; Villas Boas et al., 2010). For the prediction of each organ system as well as the outcome, three different scenarios were explored considering transformed variables:

M1 = (Hour, Case Mix, Accumulated Critical Events (ACE)

M2 = (Hour, Case Mix, ACE, Ratios)

M3 = (Hour, Case Mix, ACE, SOFA)

These variables are represented by the following sets:

Case Mix = {Age, type of admission, source of admission}

Critical Events = {CE Blood Pressure (BP), CE Oxygen Saturation (OS), EC Heart Rate (HR), EC Urine Output (UO)}

Ratios = {ECA ratios of BP / length of stay to date (LSD), ACE OS / LSD, ACE HR / LSD, ACE UO / LSD }

SOFA = {Cardiovascular, Respiratory, Renal, Hepatic, Coagulation, Neurological}

For certain models and variables, several models were tested according to some given DM techniques (Artificial Neural Networks (ANN), decision trees (DT), Linear Regression (LR)). Table 2 resumes the best results obtained for each organ system and outcome in terms of sensitivity.

Table 2 – Top Results for the organ failure and Outcome (Vilas-Boas, et al., 2010a)

System	Technique	Scenario	Sensibility (%)
Cardiovascular	ANN	M3	93.4
Respiratory	ANN	M2	96.2
Renal	ANN	M3	98.1
Coagulation	ANN	M2	97.5
Hepatic	ANN	M3	98.3
Outcome	ANN	M1	98.3

These results prove that would be possible to obtain better results than the previously presented (Á. Silva, 2007), however still using offline approach.

At this point we can conclude that there is too much to research, evolve and innovate. In this way, is essential to automate the data acquisition and transforming tasks, induce the models online and in real-time, optimize the results and make the results available through pervasive devices.

8. Conclusion

This chapter presented the state of the art of Intelligent Decision Support System in the Critical Health Care. A review of literature was made and a set of contents addressed. With this chapter was possible observe that the systems which support decision-making process and problem solving have proliferated and evolved over the last decades. They started from applications based on spread sheets (first expert systems (ExS)) to Intelligent Decision Support Systems (IDSS) incorporating data mining and optimization models with statistical and Artificial Intelligence features (Michalewicz, Schmidt, Michalewicz, & Chiriac, 2005). This may be seen through the developments that have occurred in the Business Intelligence concept and the transition that took place between the ExS and the IDSS

Innovation in the fields of wireless data communications, mobile devices and biosensor technology enables the development of new types of systems, which provide assistance anywhere and anytime (Goñi, et al., 2009) and remote access to prediction models such as Ubiquitous Data Mining (UDM).

To develop pervasive IDSS, a high number of improvements should be introduced. These improvements must take into account the model and architecture and the information sources. INTCare is an IDSS operating in real-time, whose main purpose is to improve patient health care, allowing physicians acting in a proactive way in the best interest of patients (M. F. Santos, et al., 2006).

CHAPTER V – Contributions

This chapter introduces the main contributions of the work. First, a brief summary of the most important results is presented. Then, the chapter includes a list of the most significant published articles.

1. Main Contributions

In the next lines is a description of the major contributions is presented organised in turn of the objectives previously defined.

1. **To understand the implications of a pervasive approach to Intelligent Decision Support Systems (IDSS) in the context of critical environments**

Introducing of a pervasive perspective to the decision making process. The doctors can remotely, through a secure connection, consult the patient data or make a decision. This consult may have two purposes and can be done in real-time through the web platforms:

- ENR - shows all clinical data validated about the patient, allow monitoring patient condition;
- INTCare - helps to make the decision based on some clinical predictions about the patient suggested by the Data Mining Engine.

The doctor's decision making process can start after the patient data is consulted and can be made in a single fashion or discussed in a collaborative platform.

INTCare system is designed to address know issues of the ICU setting, such as noisy, high dimensional numerical time series data in real-time (K. Morik, et al.), as well as the data acquisition in real-time, storage, integration and rapid availability of all clinical information. The pervasive approach brought new features to the system enabling the remote access. However is necessary to reformulate the ICU environment taking into account some questions like privacy, security, connection and other concerns

2. **To allow and enable pervasive approaches in critical environments**

Was proved, that it is possible create a Pervasive and Intelligent Critical Health Care Environment for data acquisition and data consulting with the maximum security and reliability for the ICU, their professionals, patients and applications. Almost all tasks can be intelligent and performed autonomously and in real-time. The proposed environment allows for a complete availability of data in electronic format wherever we need at the time that we want.

A Data engineering process for an IDSS in Intensive Medicine based in the processing and transformation of data collected in real-time was developed.

3. **To develop and test pervasive models in Intensive Care**

In this point two main contributions are noteworthy. The first one is related to the possibility of handling data stream collected directly from the devices and systems present in the ICU. A pervasive data acquisition component has been implemented. Online and real-time acquisition, processing and transforming of data is now possible. This increases the number of data provided electronically and in real-time.

The second contribution refers to the ensemble approach adopted in order to adapt the predictive models automatically and in real-time. This work corroborated the initial hypothesis that it would be possible to induce data mining models automatically and in real-time using data streaming and online-learning in critical areas as is intensive medicine.

Furthermore was presented a new approach for tracking critical events. This represents an innovation in the way of the data are collected and used to support the decision making process. It is now possible to acquire, process, and present knowledge automatically and in real-time using online-learning, anywhere and anytime.
Aiming at promoting proactive actions with the patient this pervasive intelligent system is able to evaluate the number and duration of critical events patients.

At the same level a scoring system was proposed. This system processes automatically the scores and adapt the results according to the new values collected, generating new knowledge. The main gains in using this approach can be summarized as:

- The data acquisition, the scores calculation and the results are made in real-time;
- All values are considered - no missing values;
- The data is displayed in a new way – real-time charts to compare trends;
- Less human intervention in the scores calculation – less errors;
- The scores are available anywhere and anytime;
- Help decision making process through a continuous scores monitoring - a real-time calculation of the scores according the most recent patient results, allow a quick and better comprehension of the patient condition.

Summarizing, using the data provided by the system it is possible: to present better information to support the decision process and create new knowledge. It is also possible to calculate ICU Scores: SAPS, SOFA, MEWS and TISS; calculate Critical Events (BP, Temperature, SPO2 and Urine Output) and predict organ failure or patient outcome with high level of confidence. All of these changes allow for having a bigger control of the data and a high number of data available automatically and in real-time. With this system the professionals have a complete control of the data (e.g. they can change the values whenever the values aren't correct).
Finally, results were accessed making use of the Technology Acceptance Model (TAM 3) jointly with the Delphi's method in order to evaluate the acceptance by the users, their perceptions and the impact on the INTCare's system usage behaviour.

2. Published articles

In the next section a list of published articles is presented:

A1 – A Pervasive Approach to a Real-Time Intelligent Decision Support System in Intensive Medicine

A2 – Enabling A Pervasive Approach for Intelligent Decision Support in Critical Health Care

A3 – Enabling Real-Time Intelligent Decision Support in Intensive Care

A4 – Towards Pervasive and Intelligent Decision Support in Intensive Medicine – A Data Stream Mining Approach

A5 – Pervasive Real-Time Intelligent System for Tracking Critical Events in Intensive Care Patients

A6 – Intelligent Data Acquisition and Scoring System for Intensive Medicine

A7 – Intelligent Decision Support in Intensive Care – Towards Technology Acceptance

A8 – Real-Time Decision Support in Intensive Medicine - An Intelligent Approach for Monitoring Data Quality

A9 – Pervasive Intelligent Decision Support System - Technology Acceptance in Intensive Care Units

To facilitate the reading each one of the articles is preceded by an introductory summary containing the objectives covered and the results attained.

I – Pervasive Environment

ARTICLE A1

Objectives

To understand the implications of a pervasive approach to Intelligent Decision Support Systems (IDSS) in the context of critical environments

i. To understand the impact of the pervasiveness in the construction of IDSS;
ii. To analyse what are the changes relatively to the traditional model;
iii. To study how IDSS should be adapted to operate in critical environments and in a pervasive way;
iv. To study the impact of this approach on each one of the tasks inherent to IDSS.

Results

The development of an ENR allows for the integration of all necessary information regarding the patients' condition to be collected and integrated in just one application, which is a great gain in time and performance for the medical staff operating in the ICU. In addition to the patients' vital signs, data regarding their LR, procedures, medication, is also available by the time it is generated.

Moreover, the INTCare system is designed to address know issues of the ICU setting, such as noisy, high dimensional numerical time series data in real-time (K. Morik, et al.), as well as the data acquisition in real-time, storage, integration and rapid availability of all clinical information.

The pervasive approach brought new features to the system enabling the remote access. However is necessary to reformulate the ICU environment taking into account some questions like privacy, security, connection and other concerns (Black et al.; Cook & Das, 2005; O'Donoghue, Herbert, & Sammon, 2008; U. Varshney, 2009)

A PERVASIVE APPROACH TO A REAL-TIME INTELLIGENT DECISION SUPPORT SYSTEM IN INTENSIVE MEDICINE

Filipe Portela, Manuel Filipe Santos and Marta Vilas-Boas

Departamento de Sistemas de Informação, Universidade do Minho
Campus de Azurém 4800-058 Guimarães, Portugal
{cfp, mfs mvb}@dsi.uminho.pt

Abstract. The decision on the most appropriate procedure to provide to the patients the best healthcare possible is a critical and complex task in Intensive Care Units (ICU). Clinical Decision Support Systems (CDSS) should deal with huge amounts of data and online monitoring, analysing numerous parameters and providing outputs in a short real-time. Although the advances attained in this area of knowledge new challenges should be taken into account in future CDSS developments, principally in ICUs environments. The next generation of CDSS will be pervasive and ubiquitous providing the doctors with the appropriate services and information in order to support decisions regardless the time or the local where they are. Consequently new requirements arise namely the privacy of data and the security in data access. This paper will present a pervasive perspective of the decision making process in the context of INTCare system, an intelligent decision support system for intensive medicine. Three scenarios are explored using data mining models continuously assessed and optimized. Some preliminary results are depicted and discussed.

Keywords: Real-time, Pervasive, Remotely access, Knowledge discovery in databases, Intensive care, INTCare, Intelligent decision support systems.

1. Introduction

Intensive care units (ICU) are a particular environment where a great amount of data related to the patients' condition is daily produced and collected. Physiological variables such as heart rate, blood pressure, temperature, ventilation and brain activity are constantly monitored on-line (Mahmoud, 2003). Due to the complex condition of critical patients and the huge amount of data, it can be hard for physicians to decide about the best procedure to provide them the best health care possible. The human factor can lead to errors in the decision making process; frequently, there is not enough time to analyse the situation because of stressful circumstances; furthermore, it is not possible to continuously analyse and memorize all the data (Pereira, et al., 2007) .

Care of the critically ill patients requires fast acquisition, registering and availability of data (Gardner, Hawley, East, Oniki, & Young, 1991). Accordingly, rapid interpretation of physiological time-series data and accurate assessment of patient state is crucial to patient monitoring in critical care. The data analysis allows supporting decision making through prediction and decision models. Algorithms that use Artificial Intelligence (AI) techniques have the potential to help achieve these tasks, but their development requires well- annotated patient data (K. Morik, 2003; Ying, et al., 2007).

We are deploying a real-time and situated intelligent decision support system, called INTCare1, whose main goal is to improve the health care, allowing the physicians to take a pro-active attitude in the patients' best interest (Gago, et al., 2006; M. F. Santos, et al., 2006).

INTCare is capable of predicting organ failure probability, the outcome of the patient for the next hour, as well as the best suited treatment to apply. To achieve this, it includes models induced by means of Data Mining (DM) techniques (M. F. Santos, et al., 2006), (Gago & Santos, 2008; Gago, et al., 2007; Á Silva, Cortez, Santos, Gomes, & Neves, 2004; Á Silva, Pereira, Santos, Gomes, & Neves, 2003).

Further improvements include the adjustment of the system to new requirements in order to make it pervasive and ubiquitous (Kwon, et al., 2005). This allows the system to be universal, i.e. can be used anywhere and anytime, eliminating any sort of barrier be it time or place.

This paper is organized as follows. Section 2 presents some background related to Intelligent Decision Support Systems (IDSS), Knowledge Discovery in Databases (KDD), Intensive Medicine and the Pervasive Computing. In next sections the INTCare system is presented, focusing on its features (section 3), the information architecture (section 4), pervasive approach (section 5) and the latest DM models developed (section 6). Section 7 and 8 conclude this paper, presenting a discussion, a conclusion and pointing to future work.

2. Background

2.1. Intelligent Decision Support Systems

According to Turban (Efraim Turban, et al., 2005), a Decision Support System (DSS) is an interactive, flexible and adaptable information system, developed to support a problem solution and to improve the decision making. These systems usually use AI techniques and are based on prediction and decision models that analyse a vast amount of variables to answer a question.

The decision making process can be divided in five phases: Intelligence, design, choice, implementation and monitoring (Efraim Turban, et al., 2005). Usually it is used in the development of rule based DSS (Arnott & Pervan, 2004). However, these DSS are not adaptable to the environment in which they operate. To address this fault, Michalewicz (Michalewicz, Schmidt, Michalewicz, & Chiriac, 2007) introduced the concept of Adaptive Business Intelligence (ABI). The main difference between this and a regular DSS is that it includes optimization that enables adaptability. An ABI system can be defined as "the discipline of using prediction and optimization techniques to build self-learning decision systems. ABI systems include elements of data mining, predictive modelling, forecasting, optimization, and adaptability, and are used to make better decisions." (Michalewicz, et al., 2007).

As it is known, predictive models' performance tends to degrade over time, so it is advantageous to include model re-evaluating on a regular basis so as to identify loss of accuracy (Gago & Santos, 2008) and enable their optimization.

There is a particular type of DSS, the real-time DSS. Ideally, the later includes adaptive behaviour, supporting the decision making in real-time.

To achieve real-time DSS, there is a need for a continuous data monitoring and acquisition systems. It should also be able to update the models in real-time without human intervention (M. F. Santos, et al., 2006). In medicine, most systems only use data monitoring to support its activities, without predictive behaviour and with poor integration with other clinical information.

2.2. Knowledge Discovery from Databases

KDD is one of the approaches used in Business Intelligence (BI). According to Negash (Negash & Gray), BI systems combine data gathering, data storage, and knowledge management with analytical tools to present complex and competitive information to planners and decision makers. KDD is an interactive and nontrivial process of extracting implicit and previously unknown and potentially useful and understandable information from data (Frawley, Piatetsky-Shapiro, & Matheus, 1992).

The KDD process is divided in 5 steps: Selection, pre-processing, transformation, data mining and interpretation/evaluation (Fayyad, et al., 1996). This process starts with raw data and ends with knowledge.

The automation of the knowledge acquisition process is desirable and it is achieved by using methods of several areas of expertise, like machine learning (Gago, et al.). The knowledge acquisition takes advantage of KDD techniques, simplifying the process of decision support (Gago & Santos).

Knowledge discovery is a priority, constantly demanding for new, better suited efforts. Systems or tools capable of dealing with the steadily growing amount of data presented by information system, are in order (Lourenco & Belo, 2003).

2.3. Intensive Medicine

Intensive medicine can be defined as a multidisciplinary field of the medical sciences that deals with prevention, diagnosis and treatment of acute situations potentially reversible, in patients with failure of one or more vital functions (A. Silva, 2007).

These can be grouped into six organic systems: Liver, respiratory, cardiovascular, coagulation, central nervous and renal (Hall, et al., 2005).

ICU are hospital services whose main goal is to provide health care to patients in critical situations and whose survival depends on the intensive care (Ramon et al., 2007), (Rao & T., 2003).

In the ICU, the patients' vital signs are continuously monitored and their vital functions can be supported by medication or mechanical devices, until the patient is able to do it autonomously (Ramon, et al., 2007).

Clinical intervention is based on the degree of severity scores like the SOFA (Sequential Organ Failure Assessment) score, that allow the evaluation of the patient's condition according to a predefined set of values (J. L. Vincent, et al., 1996).

The assessment of these severity scores are based on several medical data acquired from bedside monitors, lab results and clinical records.

2.4. Real-Time

A system that aims to support decision making must analyse many parameters and output in short real-time and consider online monitoring (K. Morik, Brockhausen, & Joachims). It is known that in the ICU setting, there is a huge amount of noisy, high dimensional numerical time series data describing patients. Consequently, such systems must go beyond classical medical knowledge acquisition, since they have to handle with high dimensional data in real-time.

Data acquisition in real-time implies the need for a system responsible for collecting the relevant data to the DSS. This process can be divided in two phases: monitoring and acquisition and storage.

Initially, the required data (variables) for the project is identified for further being monitored by sensors or other technology. Subsequently, data is acquired and stored in DB.

This is a critical phase, for technical, human and environment factors are involved and may condition the quality of the data acquired by a gateway, for example, and its storage on a DB. Usually, the monitoring is continuous and there are a small percentage of failures. Although they may occur, they are relatively easy to correct. The biggest problem occurs in the communication between the monitoring system and the storage system.

In conclusion, monitoring in real-time is relatively easy; usually, problems arise in the data storage process.

2.5. Pervasive Computing

Satyanarayanan (Satyanarayanan, 2002) characterizes the pervasive computing (ubiquitous) as an evolutionary step resulting from the harmonization of the fields of distributed computing and mobile computing (Figure I.1). This involves not only issues related with saturated environments of communication and user interaction, but also in the support of user's mobility.

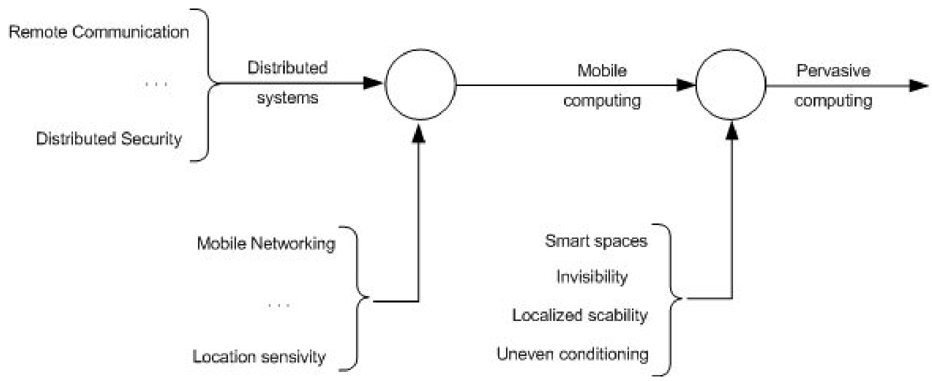


Figure I.1 – Pervasive Computing (Satyanarayanan, 2002).

This type of computing has evolved with the emergence of the Internet. Nowadays, the Internet is easily accessible in all developed countries, schools, public organizations, at home, inside or outside of buildings.

The appearance and ease of access to wireless networks and communication systems contributed to its dissemination. The incorporation of computing devices, objects or sites for monitoring, allows a glimpse of real, physical, and enhanced with information and computing resources that can be used to facilitate human life in its various tasks (personal or social) or to improve business or organizational processes (Banavar, et al., 2000). Complementarily, Banavar (Banavar, et al., 2000) also considers that pervasive computation can be:

i. the way people think about and use mobile computing devices to perform tasks;
ii. how applications are developed and implemented as form of support for these tasks;
iii. How the environment is strengthened by the emergence and ubiquity of new information and features.

Almost all the concepts mentioned above can be migrated to pervasive health care. According to Varshney (Upkar Varshney, 2003), Pervasive HealthCare (PH) can be defined as "health for everyone, anytime and anywhere by removing restrictions such as location and time, increasing both coverage and quality of healthcare". Such approach is essentially based on information that is stored and available online (Mikkonen, et al., 2002).

One of the targets of the system is to achieve the three points described above and to deliver medical information remotely and online through ubiquitous devices (Kohn, et al., 2000; Makeham M, 2002).

Such approach contributes for the mitigation of medical errors caused by the lack of correct information at the place and time that it is required e.g., misdiagnosis. Decision errors may reach 50% of the total errors (Bergs, et al., 2005).

Ubiquity in electronic medical records enhances the analysis of the information by authorized users, anytime and anyplace (U. Varshney, 2009) promoting a real-time operation.

The granularity of time, the representation and retrieval of information play an important role in the mode as the diagnosis, prognosis and treatments are performed (Augusto, 2005).

3. The INTCare system

INTCare is an IDSS for intensive medicine that is being developed in the ICU of the Hospital Santo António (HSA) in Porto, Portugal. It makes use of intelligent agents (Abelha et al.; M. F. Santos M.F.; Manuel Filipe. Santos et al., 2011a) that are capable of autonomous actions in order to meet its goals (Gago, et al., 2006), (Jennings, 2000).

To allow a pervasive access to the IDSS some also is essential take into account some requirements and concerns like privacy and security of the users, patients and system.

3.1. System Features

In order to model the information for KDD processing, the system attends some requirements:

a) **Online Learning** - The system acts online, i.e., the DM models are induced using online data in opposition of an offline approach, where the data is gathered and processed afterwards;
b) **Real-Time** - The system actuates in real-time, for the data acquisition and storing is made immediately after the events take place to allow that decisions are taken whenever an event occurs;
c) **Adaptability** - The system has the ability to, automatically, optimize the models with new data when needed. This information is obtained from their evaluation results;
d) **Data mining models** - The success of IDSS depends, among others, on the acuity of the DM models, i.e., the prediction models must be reliable. These models make it possible to predict events and avert some clinical complications to the patients;
e) **Decision models** - The achievement of the best solutions depend heavily on the decision models created. Those are based in factors like differentiation and decision that are applied on prediction models and can help the doctors to choose the better solution on the decision making process;
f) **Optimization** – The DM models are optimized over time. With this, their algorithms are in continuous training so that increasingly accurate and reliable solutions are returned, improving the models acuity;
g) **Intelligent agents** - This type of agents makes the system work through autonomous actions that execute some essential tasks. Those tasks support some modules of the system: Data acquisition, data entry, knowledge management, inference and interface. The flexibility and efficiency of this kind of system emerges from the intelligent agents and their interaction (Gago, et al., 2006).
h) **Accuracy**: The data available in the IDSS need to be accurate and reliable. The system need to have an autonomous mechanism to a pre-validation of the data. The final validation will be always done by a Human, normally by the nurse staff. This operation should be done on the ENR, moments after collection. With this, the user is sure that the data he can see online is guaranteed true.

i) **Safety**: All patient data should be safely stored in the database. The data security has to be ensured the access should be restricted. This is the one of the most critical aspects in this type of approach.

j) **Pervasive / Ubiquitous:** The system need to be prepared to work in ubiquitous devices like notebooks, PDAs and mobile phones. The internet plays an important role making the system available for users in anyplace. The ICU access policy should be available.

k) **Privacy**: There are two types of privacy: i) related to the patient and; ii) related to the health care professional. The patient identification should be always hidden to the people out of hospital. On the other hand the pieces of information recorded on this environment need to be identified and associated to one user, in order to find out responsibilities. Both types of identifications should be protected and masked.

l) **Secure Access from Exterior**: The hospital access point has to be protected from exterior connections and encrypted. A Virtual Private Network (VPN) with appropriate access protocols is a good option. Only people who have access to the ICU can see the information and operate, locally or remotely, with the IDSS. This system should implement a secure policy access and be prepared to work in a protected environment.

m) **User Policy**: The IDSS should include an inside (ICU environment) and an outside (remote connections) access policy, e.g. where and who can be consult or edit the data.

In order to accomplish the features presented above, some requirements should be considered:

- Fault tolerance capacities;
- Processing to remove null and noisy data;
- Automatic detection and processing of null patient identification;
- Automatic validation of the data taking into account the ranges (min, max) of each variable;
- Continuous data acquisition process;
- Time restrictions for the data acquisition and storage;
- Online learning mode;
- Digital data archive in order to promote the dematerialization of paper based processes (e.g., nursing records);
- Database extension to accommodate the data structures;
- Correct usage of the equipment that collects the vital signs.

4. Information Architecture

Patient management is supported by complex information systems, which brings the need for integration of the various types and sources of data (Fonseca, Ribeiro, & Granja, 2009).
In order to follow the requirements enumerated above, an information model was drawn, regarding the data acquisition module which includes three types of information sources:

- Bedside Monitors (BM);
- Lab Results (LR); and
- Electronic Nursing Records (ENR).

All sources can produce information to the system and that information can be used to develop predicting models in Intensive Care (knowledge). The development of an automated information system for ICU has to be in harmony with the whole information system and activities within the unit and the hospital (Fonseca, et al., 2009).

The first type of sources relates to data acquisition from BM. This acquisition is in real-time, the data is received by a gateway, and it is stored on a DB table by an agent. Automatic acquisition eliminates transcription errors, improves the quality of records and allows the assembly of large electronic archives of vital sign data (Fonseca, et al., 2009)

The second type of sources (LR) is the one that contains the less frequent observations, because the patient normally does this type of clinical analysis once or twice per day, except in extraordinary situations. With this method we can collect the data related with some clinical analysis, such as: number of blood platelets, creatinine, bilirubin, SOFA scores, and partial pressure of oxygen in arterial blood and fraction of inspired oxygen.

4.1. INTCare sub-systems functionality

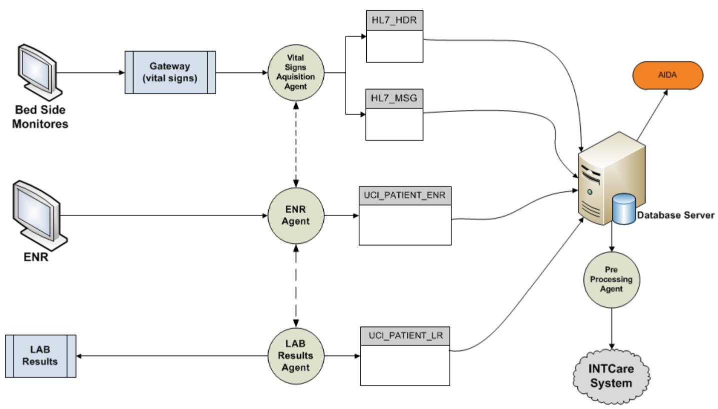


Figure I.2 – The INTCare Data Acquisition Subsystem.

The INTCare System (Gago, et al., 2006; M. F. Santos, et al., 2006) is divided into five subsystems: data acquisition, data entry, knowledge management, inference and interface. Figure I.2 shows a model that is a part of INTCare system and represents an evolution of two subsystems: data acquisition and data entry. This subsystem is responsible for all activities of data acquisition and data store and will gather all required data into a data warehouse (Santos M.F. & J., 2009b; M. F. Santos, et al., 2009; M. F. Santos, Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., Neves, J., 2009a). The evolution of this architecture is prominent. Formerly, most of the data was registered in paper format, and it was necessary to manually put it in electronic format, i.e., the information was rarely stored in computers, except the information from the BM, which was automatically collected and stored in electronic format.

The new architecture (Santos M.F. & J., 2009b; M. F. Santos, et al., 2009; M. F. Santos, Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., Neves, J., 2009a) contemplates the data acquisition from three sources and, regarding the information input, it is done either automatically (BM, LR) or automatically and manually (nursing records). The adjustment made to the system was the addition of one more data source and the creation of two more agents that enable storing information in the database (DB).

This modification is in course and it is the most important, because it makes possible the data acquisition in electronic and automatic mode for all data sources through multi-agent system. Whit this change, we will have all the necessary information in electronic format for the DM models and the decision support process, addressing the timing requirements of critical tasks.

4.2. How these subsystems work:

The first type of data sources is the BM, which collects the patients' vital signs (VS). The gateway is connected to the monitors, reads the information and stores it on a DB through the data acquisition agent. This agent splits a HL7 (Hooda, Dogdu, & Sunderraman) message in two, one with the header information and another with patient data.

The second source is the ENR (Santos M.F. & J., 2009b). It was developed with the objective of registering electronically the paper-based nursing records. With the ENR, the medical and nursing staff can register various types of data, like confirming if some therapeutic was performed or not, and they may consult all the present and past data about the patients. The last type of data sources is the LR, which is controlled by the clinical analyses agent that automatically stores all the LB from the patients.

All the data is stored in one DB and it can be accessed by the medical staff through a computer. The integrated data will be used by the INTCare system to create prevision and decision models.

The DM agent belongs to the sub-system knowledge management and it is in charge of retrieving the required data to feed the DM models and to train new models whenever their performance becomes unsatisfactory.

5. Pervasive Approach

One of the most important limitations in the ICU environment is related to the information availability which prevents the doctors to make the best decision for the patient. This is caused by the lack of integration of information in electronic format (a lot of records are still made in paper format).

Overcoming this limitation by means of a pervasive approach gives the doctors the possibility to review the data and act in the right time, i.e. before patient's clinical condition worsens, otherwise the doctor can only treat the patient when he is physically in the service.

In this context electronic medical records with detailed information about the patient may be analysed by anyone authorized, anytime and anywhere (U. Varshney, 2009).

The information that returns from the applications is stored and made accessible from one single site. This implies some modifications to enable the access from small portable devices [29] (as defined in chapter 3).

Figure I.3 shows a pervasive perspective of the decision making process. The doctors can remotely, through a secure connection, consult the patient data or make a decision. This consult may have two purposes and can be done in real-time through the web platforms:

- ENR - shows all clinical data validated about the patient,
- INTCare - helps to make the decision based on some clinical predictions about the patient suggested by the Data Mining Engine.

The doctor's decision making process can start after the patient data is consulted or at the start, and can be made in a single fashion or discussed in a collaborative platform (Miranda, Abelha, Santos, Machado, & Neves, 2009; Villas Boas, et al., 2010).

After the decision be taken, this can be performed remotely, by a directly configuration in an ICU System or by giving some clinical instruction to the nurses. At the end, the performed decision will be used to optimize and adapt the decision and prevision models.

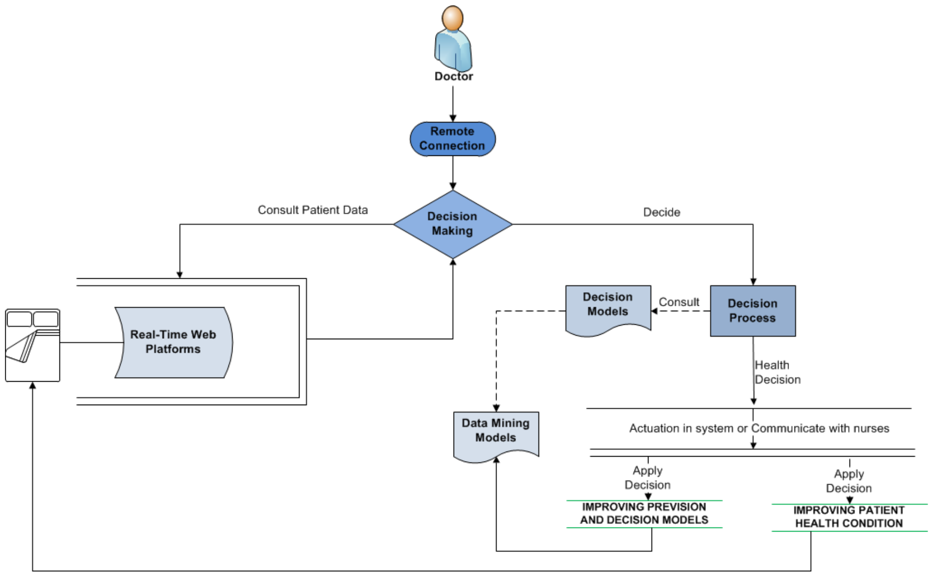


Figure I.3 – Decision Support - Pervasive Approach.

6. Data Mining Models

6.1. Data description

The data used to generate the DM models originates from three distributed and heterogeneous sources: LR, BM and paper-based nursing records, presented and explained previously. Additionally, variables containing the case mix (information that remains unchanged during the patient's stay - age, admission from, admission type) were also considered. It was also included some calculated variables: Critical Events (CE), SOFA scores and a set of ratios relating the previous variables to the patients' length of stay.

The data was gathered in the ICU of HSA and it was collected in the first five days of stay of thirty two patients. The construction of the dataset was not automatically, the data from the LB and nursing records was manually registered, for the new adjustments of the system regarding the data acquisition and data entry were not developed at the time the models were generated.

6.2. Features selection

For the prediction of the dysfunction/failure of each organic system and outcome, three scenarios were explored regarding the inclusion of the variables mentioned above – M1, M2, and M3 – where,

M1 = {Hour, Case Mix, CE}
M2 = {Hour, Case Mix, CE, Ratios}
M3 = {Hour, Case Mix, CE, SOFA}.

For each model, the techniques applied were Artificial Neural Networks, Decision Trees, Regression and Ensemble methods.

6.3. Results

Table I.1 shows the best results achieved for cardiovascular, respiratory, renal, liver, coagulation and neurological systems and outcome in terms of sensibility (i.e. percentage of failure and death correctly classified as such) as well as the scenario that produced the best results.

The models were developed for hourly prediction with the intent to make predictions as fast as possible, in the patients best interest.

Table I.1 – Sensibility of the DM models by system and outcome

System	*Sensibility (%)*	*Scenario*
Cardiovascular	93,4	M3
Respiratory	96,2	M2
Renal	98,1	M3
Coagulation	97,5	M2

7. Discussion

In this paper we presented the INTCare system, which is an IDSS for intensive medicine. It relies on the KDD process and AI algorithms to apply DM techniques for predicting outcomes that might support the course of action of doctors' decision.

Relying on intelligent agents, the system in divided into five sub-systems (data acquisition, data entry, knowledge management, inference and interface) that guarantee its functionality.

Since its beginning, INTCare has evolved towards using real-time and online clinical data so that the predictions can be as accurate and as soon as possible. As an IDSS, INTCare uses continuous data monitoring and acquisition systems that make possible for all information being available at the right time. This allows doctors to have a proactive attitude in patients' care.

The development of an ENR allows the integration of all necessary information regarding the patients' condition to be collected and integrated in just one application, which is a great gain in time and performance for the medical staff operating in the ICU. In addition to the patients' vital signs, data regarding their LR, procedures, medication, is also available by the time it is generated.

Moreover, the INTCare system is designed to address know issues of the ICU setting, such as noisy, high dimensional numerical time series data in real-time (K. Morik, et al.), as well as the data acquisition in real-time, storage, integration and rapid availability of all clinical information. The pervasive approach brought new features to the system enabling the remote access. However is necessary to reformulate the ICU environment taking into account some questions like privacy, security, connection and other concerns (Black et al.; Cook & Das, 2005; O'Donoghue, Herbert, & Sammon, 2008; U. Varshney, 2009).

8. Conclusions and future work

The main concern in ICU is to avoid or reverse organ failure, in order to preserve the patients' lives. The INTCare system is being developed for hourly prediction of the patients' clinical condition, i.e. the prediction of dysfunction/failure of the organ systems (cardiovascular, respiratory, renal, coagulation and liver systems) and outcome.

We believe that, with this fine grained prediction, it will be possible for the healthcare professionals to have a timely intervention and a proactive attitude, regardless the space where they are, so that worst complications for the patients may be avoided. A pervasive approach will have a strong impact. Context awareness is an important issue to adapt the applications to the current situation.

Further work will encompass the test of the DM models generated so far, with online and real-time data from the ICU of HSA, in order to guarantee their accuracy or, in case their performance decays, to optimize them. The models presented used data manually entered and the next step is to use them with the new adjustments of the system, i.e., online and in real-time. Prediction, optimization and adaptability are features that make INTCare an ABI system, whose maid goal it to allow the medical staff to make better decisions, at the right time and place, improving quality in health care. At the same time the pervasive approach will be tested and evaluated remotely.

The integration with the various data sources and with the rest information systems of the hospital has been supported by the development of an ENR and further related work include its test in the ICU and subsequently, its optimization.

Acknowledgment

The authors would like to thank FCT (Foundation of Science and Technology, Portugal) for the financial support through the contract PTDC/EIA/72819/2006.

References

Abelha, A., Machado, J., Santos, M., Allegro, S., Rua, F., Paiva, M., et al. *Agency for Integration, Diffusion and Archive of Medical Information*.

Arnott, D., & Pervan, G. (2004, Jul). *A critical analysis of decision support systems research.* Paper presented at the Conference on Decision Support Systems, Prato, ITALY.

Augusto, J. C. (2005). Temporal reasoning for decision support in medicine. *Artificial Intelligence in Medicine, 33*(1), 1-24.

Banavar, G., Beck, J., Gluzberg, E., Munson, J., Sussman, J., & Zukowski, D. (2000, 2000). *Challenges: an application model for pervasive computing*.

Bergs, E. A. G., Rutten, F. L. P. A., Tadros, T., Krijnen, P., & Schipper, I. B. (2005). Communication during trauma resuscitation: do we know what is happening? *Injury, 36*(8), 905-911.

Black, J. P., Segmuller, W., Cohen, N., Leiba, B., Misra, A., Ebling, M. R., et al. (2004). *Pervasive Computing in Health Care: Smart Spaces and Enterprise Information Systems*.

Cook, D. J., & Das, S. K. (2005). *Smart environments: technologies, protocols, and applications*: Wiley-Interscience.

Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery: an overview.

Fonseca, T., Ribeiro, C., & Granja, C. (2009). Vital Signs in Intensive Care: Automatic Acquisition and Consolidation into Electronic Patient Records. *Journal of Medical Systems, 33*(1), 47-57.

Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992). Knowledge Discovery in Databases: An Overview. *AI Magazine, 13*(3), 57-70.

Gago, P., & Santos, M. F. (2008). *Towards an Intelligent Decision Support System for Intensive Care Units*. Paper presented at the 18th European Conference on Artificial Intelligence.

Gago, P., Santos, M. F., Silva, Á., Cortez, P., Neves, J., & Gomes, L. (2006). INTCare: a knowledge discovery based intelligent decision support system for intensive care medicine. *Journal of Decision Systems*.

Gago, P., Silva, A., & Santos, M. F. (2007, Dec 03-07). *Adaptive decision support for intensive care.* Paper presented at the 13th Portuguese Conference on Artificial Intelligence, Guimaraes, PORTUGAL.

Gardner, R. M., Hawley, W. L., East, T. D., Oniki, T. A., & Young, H. F. W. (1991). Real-time data acquisition: recommendations for the Medical Information Bus (MIB). *International journal of clinical monitoring and computing, 8*(4), 251-258.

Hall, J. B., Schmidt, G. A., & Wood, L. D. H. (2005). *Principles of Critical Care*: McGraw-Hill's AccessMedicine.

Hooda, J. S., Dogdu, E., & Sunderraman, R. (2004). *Health Level-7 compliant clinical patient records system*.

Jennings, N. R. (2000). On agent-based software engineering. *Artificial Intelligence, 117*(2), 277-296.

Kohn, L. T., Corrigan, J., & Donaldson, M. S. (2000). *To Err Is Human: Building a Safer Health System*: National Academy Press.

Kwon, O., Yoo, K., & Suh, E. (2005). UbiDSS: a proactive intelligent decision support system as an expert system deploying ubiquitous computing technologies. *Expert Systems with Applications, 28*(1), 149-161.

Lourenco, A., & Belo, O. (2003). Promoting agent-based knowledge discovery in medical intensive care units. *WSEAS Transactions on Computers, 2*(2), 403-408.

M. F. Santos M.F., V.-B. M., Machado and A. A. J., Neves J., Silva A., Rua F., Salazar M., Quintas C., Cabral A.F. Intelligent Decision Support in Intensive Care Units - Nursing Information Requirements.

Mahmoud, M. (2003). Real-time data acquisition system for monitoring patients in intensive care unit. (Vol. 5090, pp. 320-326). Multisensor, Multisource Information Fusion: Architectures, Algorithms and Applications.

Makeham M, D. S., County M, Kidd M. (2002). An international taxonomy for errors in general practice: a pilot study. *The Medical Journal of Australia, 177*(2), 68-72.

Michalewicz, Z., Schmidt, M., Michalewicz, M., & Chiriac, C. (2007). *Adaptive Business Intelligence*: Springer.

Mikkonen, M., Vayrynen, S., Ikonen, V., & Heikkila, M. O. (2002). User and concept studies as tools in developing mobile communication services for the elderly. *Personal and Ubiquitous Computing, 6*(2), 113-124.

Miranda, M., Abelha, A., Santos, M., Machado, J., & Neves, J. (2009). A group decision support system for staging of cancer. *Electronic Healthcare, 0001*, 114-121.

Morik, K. (2003). Data analysis and knowledge validation in intensive care monitoring.

Morik, K., Brockhausen, P., & Joachims, T. (1999). *Combining statistical learning with a knowledge-based approach-a case study in intensive care monitoring*.

Negash, S., & Gray, P. Business intelligence. *Handbook on Decision Support Systems 2*, 175-193.

O'Donoghue, J., Herbert, J., & Sammon, D. (2008). Patient sensors: A data quality perspective. In S. Helal, S. Mitra, J. Wong, C. Chang & M. Mokhtari (Eds.), *Smart Homes and Health Telematics* (Vol. 5120, pp. 54-61). Berlin: Springer-Verlag Berlin.

Pereira, M., Curra, A., Rivas, R., Pereira, J., Banos, G., Teueiro, J., et al. (2007). Computer aided monitoring system of intensive care unit patients. *WSEAS Transactions on Information Science and Applications, 4*(1), 78-84.

Ramon, J., Fierens, D., Güiza, F., Meyfroidt, G., Blockeel, H., Bruynooghe, M., et al. (2007). Mining data from intensive care patients. *Advanced Engineering Informatics, 21*(3), 243-256.

Rao, S. M., & T., S. (2003). Organization of intensive care unit and predicting outcome of critical illness. *Indian J. Anaesth, 47 (5)*, 328-337.

Santos M.F., P. F., Vilas-Boas M., Machado, & J., A. A., Neves J., Silva A., Rua F., Salazar M., Quintas C., Cabral A.F.,. (2009). Intelligent Decision Support in Intensive Care Units - Nursing Information Requirements. In Springer (Ed.), *WSEAS Transactions on INFORMATICS*.

Santos, M. F., Cortez, P., Gago, P., Silva, Á., & Rua, F. (2006). *Intelligent decision support in Intensive Care Medicine.* Paper presented at the 2nd International Conference on Knowledge Engineering and Decision Support, Lisbon, Portugal.

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., & Neves, J. (2011). *INTCARE - Multi-agent approach for real-time Intelligent Decision Support in Intensive Medicine*. Paper presented at the 3rd International Conference on Agents and Artificial Intelligence (*ICAART*).

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., Neves, J., et al. (2009). Information Modeling for Real-Time Decision Support in Intensive Medicine. In S. Y. Chen & Q. Li (Eds.), *Proceedings of the 8th Wseas International Conference on Applied Computer and Applied Computational Science - Applied Computer and Applied Computational Science* (pp. 360-365). Athens: World Scientific and Engineering Acad and Soc.

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., Neves, J. (2009). *Information Architecture for Intelligent Decision Support in Intensive Medicine*. Paper presented at the 8th WSEAS International Conference on APPLIED COMPUTER & APPLIED COMPUTATIONAL SCIENCE (ACACOS '09).

Satyanarayanan, M. (2002). Pervasive computing: vision and challenges. *Personal Communications, IEEE, 8*(4), 10-17.

Silva, A. (2007). Modelos de Inteligência Artificial na análise da monitorização de eventos clínicos adversos, Disfunção/Falência de órgãos e prognóstico do doente critico.: Tese de doutoramento, ciências médicas, universidade do Porto.

Silva, Á., Cortez, P., Santos, M. F., Gomes, L., & Neves, J. (2004). *Multiple organ failure diagnosis using adverse events and neural networks.* Paper presented at the 6 th International Conference on Enterprise Information Systems.

Silva, Á., Pereira, J., Santos, M., Gomes, L., & Neves, J. (2003). *Organ failure prediction based on clinical adverse events: a cluster model approach*, 3th International Conference on Artificial Intelligence and Applications.

Turban, E., Aronson, J. E., & Liang, T.-P. (2005). *Decision Support Systems and Intelligent Systems* (7 ed.): Prentice Hall.

Varshney, U. (2003). Pervasive Healthcare. *Computer, 36*(12), 138-140.

Varshney, U. (2009). *Pervasive Healthcare Computing: EMR/EHR, Wireless and Health Monitoring*: Springer-Verlag New York Inc.

Villas Boas, M., Gago, P., Portela, F., Rua, F., Silva, Á., & Santos, M. F. (2010). *Distributed and real-time Data Mining in the Intensive Care Unit*. Paper presented at the 19th European Conference on Artificial Intelligence - ECAI 2010.

Vincent, J. L., Moreno, R., Takala, J., Willatts, S., De Mendonca, A., Bruining, H., et al. (1996). The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive care medicine, 22*(7), 707-710.

Ying, Z., Silvers, C. T., & Randolph, A. G. (2007). *Real-Time Evaluation of Patient Monitoring Algorithms for Critical Care at the Bedside.* Paper presented at the Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE.

ARTICLE A2

Objectives

To adopt pervasive approaches in critical environments

i. To characterize Critical Environments (CEnv);
ii. To understand and interpret the problem associated with pervasive approach in CEnv;
iii. To analyse the impact and implications of the approach
iv. To study the impact of the actions of prepare the environment to be pervasive;

The main objective of this paper is to present the requirements that should be addressed in order to bring pervasiveness to an IDSS and an evaluation framework in order to guarantee the results.

Results

In order to set up the environment to an IDSS like INTCare, some features of the system (Filipe Portela, Santos, et al., 2010) need to be assured. The main features for this environment was defined.

The features defined and evaluated in this work made possible to transform the system in order to be more secure, robust, easily accessible and intelligent. Is expected the system will improve the patient outcome in the future due to some new facilities like the data availability in online, real-time and in electronic format. With the ICU pervasive access recast, is possible to accede to knowledge portions that can support the decision making process, anytime, anywhere. Beyond the performance of INTCare, the evaluation of the success of the proposed environment depends on the results obtained after a sample of 100 surveys answers that will be done to all people that directly interacted with this new ICU environment and can answer (patient, doctors, nurses, other professionals). Based on the results obtained for each category and according to table 2 we will decide on what to do in the future. The results also can give us indications about if the new environment is or not prepared to implement a pervasive IDSS like is INTCare.

ENABLING A PERVASIVE APPROACH FOR INTELLIGENT DECISION SUPPORT IN CRITICAL HEALTH CARE

Filipe Portela[1], Manuel Filipe Santos[2], Álvaro Silva[3], José Machado[4], António Abelha[4]

[1,2] Centro Algoritmi, Departamento de Sistemas de Informação, Universidade do Minho, Portugal
[3] Serviço Cuidados Intensivos, CHP – Hospital Santo António, Porto, Portugal,
[4] Departamento de Informática, Universidade do Minho, Portugal
[1, 2]{cfp, mfs}@dsi.uminho.pt; [3]moreirasilva@clix.pt; [4]{jmac, abelha}@di.uminho.pt

Abstract. The creation of a pervasive and intelligent environment makes possible the remote work with good results in a great range of applications. However, the critical health care is one of the most difficult areas to implement it. In particular Intensive Care Units represent a new challenge for this field, bringing new requirements and demanding for new features that should be satisfied if we want to succeed. This paper presents a framework to evaluate future developments in order to efficiently adapt an Intelligent Decision Support System to a pervasive approach in the area of critical health (INTCare research project).

Keywords. Intensive Care; Pervasive Environments; Critical Health Care; Intelligent Environment; Real-Time; Online; Remote Connection

1. Introduction

INTCare (Gago, et al., 2006; Gago, et al., 2007; Manuel Filipe. Santos, et al., 2011a; M. F. Santos, et al., 2009) is an on-going research project involving the Intensive Care Unit of the Porto Hospital Centre whose objective is to implement an Intelligent Decision Support System (IDSS) to predict the dysfunction or failure of six organic systems and the patient outcome in order to help doctors deciding on the better treatments or procedures for the patient. The good results obtained so far (Filipe Portela, Vilas-Boas, et al., 2010; Vilas-Boas, et al., 2010a; Villas Boas, et al., 2010) motivated the transformation of this system into a pervasive IDSS. A framework to evaluate the efficacy of the new characteristics has been developed as a way to attest their feasibility. A Critical Environment has special characteristics and needs, such as fast, efficient, secure and ubiquitous operations in real-time. At the moment, we choose the ICU to make our tests. ICUs are considered a critical environment because they have some complex health care situations (Bricon-Souf & Newman, 2007). The activities occurring in it are sometimes adverse, dangerous and tiring. and the various organ systems of the patient may be affected at the same time (Apostolakos & Papadakos, 2001) what represents a challenge for the systems (Bricon-Souf & Newman, 2007) that operate in this environment. The ICU main objectives are diagnose, monitor and treat patients with serious illnesses and recover them for their health previous state and quality of life (Suter, et al., 1994). With the introduction of Intelligent Environments (IE) this type of monitoring could be done remotely. These operations could be supported by the new technologies based on Pervasive Healthcare Computing that allow the execution of remote tasks (access and control).

Pervasive HealthCare can be defined as "healthcare to anyone, anytime, and anywhere by removing locational, time and other restraints while increasing both the coverage and the quality of healthcare" (U. Varshney, 2007a).

A pervasive medical environment is designed to allow smart interactions by mobile devices with the patient sensor and data servers. Is imperatively necessary help the doctors make the best decision and take a pro-active attitude in the patients' best interest (Gago, et al., 2006; M. F. Santos, et al., 2006). For this is important that medical staff have all important and essential information, in real-time, online, and in electronic format. The best solution is modify the environment paradigm, making it intelligent, where all information will be available at the right time and the right place, eliminating any kind of barrier either be it time or place. This is very important because the communications between health care professionals represents a large part of their activity (Coiera, 2000). This development aims the remote access to the health data and future predictions of patient conditions made available by the INTCare system, guarantying the maximum quality, security and privacy.

The main objective of this paper is to present the requirements that should be addressed in order to bring pervasiveness to an IDSS and an evaluation framework in order to guarantee the results. This paper is divided into eight chapters. The first one introduces the necessities of the new environment and the motivation of the work, the second gives an idea about what is the INTCare system. The third one presents the environment evolution, comparing the past (As was) and the present (As is). The forth chapter presents the main features necessary to enable a pervasive environment. In order to evaluate the implementations, the chapter five proposes a framework to evaluate the environment. At the end, some discussion about the significations of each result will be done, including conclusions and future work.

2. Background

Like Roy and Das (Das & Roy, 2008) told: "The essence of pervasive computing lies in the creation of smart environments saturated with computing and communication capabilities, yet gracefully integrated with human users (inhabitants)". Nowadays, we can verify that pervasive and ubiquitous computing and ambient intelligence are concepts evolving in a growing number of applications in health care and the influence of it increase every days (Orwat, et al., 2008). Some similarities can be found in other study related to the developing of pervasive computing (Black, et al.): there will be an evolution from isolated “smart spaces” to more integrated Hospital environments, which can be accessed remotely. The necessities are global and the medical staffs "require" information about their patient available through the world and, if possible, with some patient condition previsions to help in the decision-making process. An intelligent and pervasive environment requires that a large number of tasks could be done automatically and the information should be accessed electronically. however the Health Care providers are registering, manually, the patient characteristics into the computer system , hindering automation of some tasks like the retrieval, registration and display of information (Villas Boas, et al., 2010).

INTCare is IDSS that is prepared to work in critical environments and are being modified to also be pervasive. This IDSS is implemented in terms of a Multi Agent System in order to process some tasks automatically and to ensure their success. The flexibility of such approach makes the incorporation of new agents to integrate the IDSS in the Pervasive Intelligent Environment an easy task. INTCare is capable of predicting organ failure probability and the outcome of the patient for the next hour, as well as the best suited treatment to apply(Filipe Portela, Santos, et al., 2010) .

To achieve this, it includes models induced by means of Data Mining techniques (M. F. Santos, et al., 2006), (Gago & Santos, 2008; Gago, et al., 2007; Á Silva, et al., 2004; Á Silva, et al., 2003). This system has so particular Features (Filipe Portela, Santos, et al., 2010) that allow operate in pervasive environment: Online Learning, Real-Time, adaptability, Data Mining and Decision Models, Optimization and Intelligent Agents. Now some pervasive and ubiquitous characteristics are being added.

3. Pervasive Environment

The objective is to create a “smart environment” able to operate autonomously, prepared to acquire and apply knowledge about its users and their surroundings, and adapt to the users’ behaviour or preferences with the ultimate goal to improve their experience in those environment (Cook & Das, 2005), facilitating the mode how they work and improve the patient condition. To do this, we had to understand as was the environment and define the necessary modifications to make it pervasive (As is). Accomplishing these changes we are extending the existing environment in order to turn it prepared to an automated information system for ICU taking into account the harmonisation with the whole information system and activities within the unit and the hospital (Kalli et al., 1993).

3.1. As was

The project started in 2008 and during this time some important steps were given: transforming a paper based and manual process information system to an automatic and electronic one. The data was collected in offline and in an irregularly mode making the analysis of patient data and the search of past information a very hard and time-consuming work. The existing information systems were only used information consultation and not to register patient data. For example the bedside monitors only showed patient vital signs (VS) values, these weren't stored in any place. In figure II.1 is visible how Medical Staff have been working in the ICU. Normally they analyse the patient condition, looking for the VS values from the bed side monitors and, every one hour, they wrote the results in the Paper Based Nursing Record (PBNR). After that, the PBNR values were made available to be analysed and stored manually into tables.

The Lab Results constitute another important data source being collected, in average, one or two times a day. The data was only available to be consulted in PDF format only 2 hour after the measuring. To store the results into a database the medical staff should read the documents and insert, manually, the values into the tables. Only after this process the patient information will be available in the database (DB) and in electronic format. This process was very inefficient and was the origin of many errors.

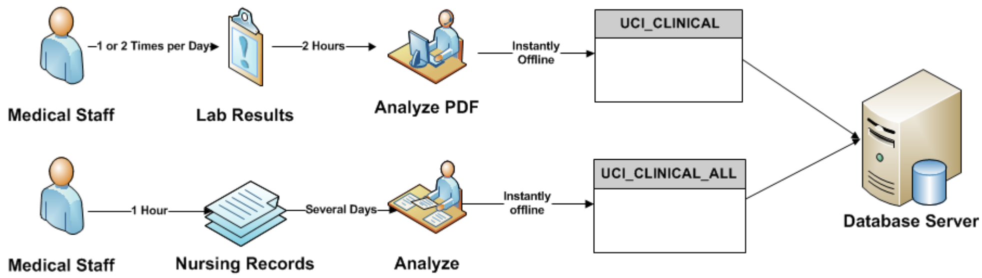


Figure II. 1 – ICU Past Environment.

3.2. As is

After some interactions with ICU staff and some analyses of the environment, we defined the necessities and requirements for the ICU. We had to change several processes and reformulate the information system (IS) architecture. Along the last three years we introduced a lot of modifications into the IS architecture (Filipe Portela, Santos, et al., 2010; M.F. Santos, F. Portela, et al., 2009b). In particular some intelligent agents were designed to do some tasks automatically and replacing some manual operations. “Intelligent agents with their properties of autonomy, reactivity, and proactivity are well suited for dynamic, ill-structured, and complex environments” (S. Gao & D. Xu, 2009), such as an ICU. The system actually working in the ICU environment can perform all tasks in online, real-time and electronic mode. The agents are used to perform automatic tasks like collecting and storing patient data. The intelligent agents collect data from four data sources: Electronic Health Record (EHR), Lab Results, Vital Signs and drugs system. All data is received and send, automatically and at the same time, to the Electronic Nursing Record (ENR).

ENR is a system which was developed with the objective to receive all medical data and put it available to Doctors and Nurses (responsible to consult and validate the data in an hourly-based mode). It is a mechanism to integrate and subsequently access patient data. The digital nature of an ENR allows data contained within it to be searched and retrieved (M. Santos et al., 2009). When some data errors appear, they should be corrected putting the exact values on ENR. On the other hand, we have some data that are still collected manually; this data is registered by the doctors or nurses and is concerned to three information's types: procedures, adverse events and some nursing records like fluid balances.

These types of operations normally are not programmed and are done by the nurses near the patient. The use of ENR improves the data quality and reduces the number of unidentified patients in the monitoring system (Filipe Portela, Vilas-Boas, et al., 2010). When the data is valid, it should be stored into the database. With this process we can ensure that the data that are available in the environment are real and are correctly associated to the patients.

The figure II.2 outlines the actual reality of ICU. Like the figure shows the environment has changed, and now only some tasks are manual. Be noticed that the data validation kept manual because only the humans can see if the data associated to some patient are correct or not (due to the semantics associated to clinical interpretation).

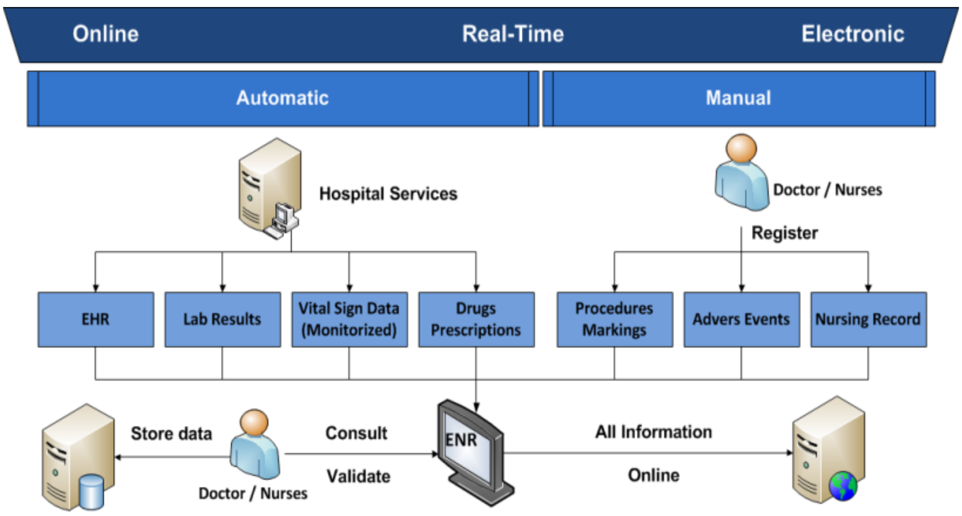


Figure II.2 – UCI Current Environment.

4. Environment Features

In order to set up the environment to an IDSS like INTCare, some features of the system (Filipe Portela, Santos, et al., 2010) need to be assured. The main features for this environment are:

a) **Real-time**: All patient data have to be collected in real-time for that some patient sensors have to be added to ICU. Is necessary creates some control tasks in the Environment to ensure that the manual data are inserted in database, as immediately as possible, after the events occurs.

b) **Electronic Mode**: In this environment all data have to be available in electronic mode for this the nursing staffs has to verify if all data have electronic access and if not, they have to register the manual data in ENR.

c) **Online**: The information of each patient need to be available online, i.e. all data have to be accessed through the ENR independently on the environment type.

d) **Autonomous**: It has to have as many as possible automatic tasks. Almost all of the tasks in the environment can be performed automatically by Intelligent Agents, however is very important that data validations continue to be manual operations.

e) **Safety**: All patient data presented on database and servers have to be safe and protected. The data security has to be ensured and nobody without access can consult the data. This is the one of the most critical aspect in these environments.

f) **Reliable**: This aspect has two types. The first, the nurse staff is responsible to validate the data, in ENR, moments after collection. The second, the systems that will operate in these environments only can see the data that it looked for. With this, the user is sure that the data he can see online is guaranteed true.

g) **Accurate**: All operations have to be approved before someone does something. The tasks have to be good defined and precise, i.e. the operation have to be valid and never can put the patient in risk. With this is possible avoid any phishing.

h) **Privacy**: There are two types of privacy: the patient and the health care professional. PID always has to be hidden to people out of hospital. On the other hand all tasks done on this environment need to be identified and associated to one person. With this we can know who did the operations and blame for something that happened.

i) **Adaptive**: The environment has to be the capability of adapt to the change of some variables, always ensure the proper functioning of the same. Sometimes is necessary optimizing some tasks and the system have to be prepared for that. If this is possible this type of operations should be performed automatically.

j) **Secure Access**: The hospital access point has to be protected and encrypted. A virtual private network with access protocols is essential because this type of operations have to be secure. Only people who have access to the ICU can see the information and operate, local or remotely, in this environment. Is extremely dangerous if foreign people accede to the hospital and see what they can't see

k) **Context awareness**: is concerned with the situation a device or user is in, and with adapting applications to the current situation (Das & Roy, 2008). All users need to be focused on the environment and to know the importance of the success of the operations because there can't be any type of negligence.

l) **Risk List and Contingence Plan**: This is a critical area where anything can fail. Is necessary to predict some problems, define the risk impact, the happen probability and prepare some contingence task.

If the aspects safety (e), privacy (h), and secure access (j) aren't always guaranteed is better not do nothing, because any type of operations can put in risk the patients, the hospital and all people that operate there. The other important aspect is the Effective data consistency since it provides the foundation to ensure that medical practitioners receive up-to-date a correct data on time, whenever they need (O'Donoghue, et al., 2008). This type of environments have to be concerned with the safety of users, reduction of cost of maintaining the environment, optimization of resources task automation or the promotion of an intelligent independent living environment for health care services and wellness management [10]. An important characteristic for the pervasive environment lies in the autonomous and pro-active interaction of smart devices used for determining users' important contexts, realization of tasks, consults of vital sign and patient information, which ensure the success of distance operations. Our environment aims to support web applications for different platforms in order to allow users to access the application using one portable device (laptop, mobile, PDA).

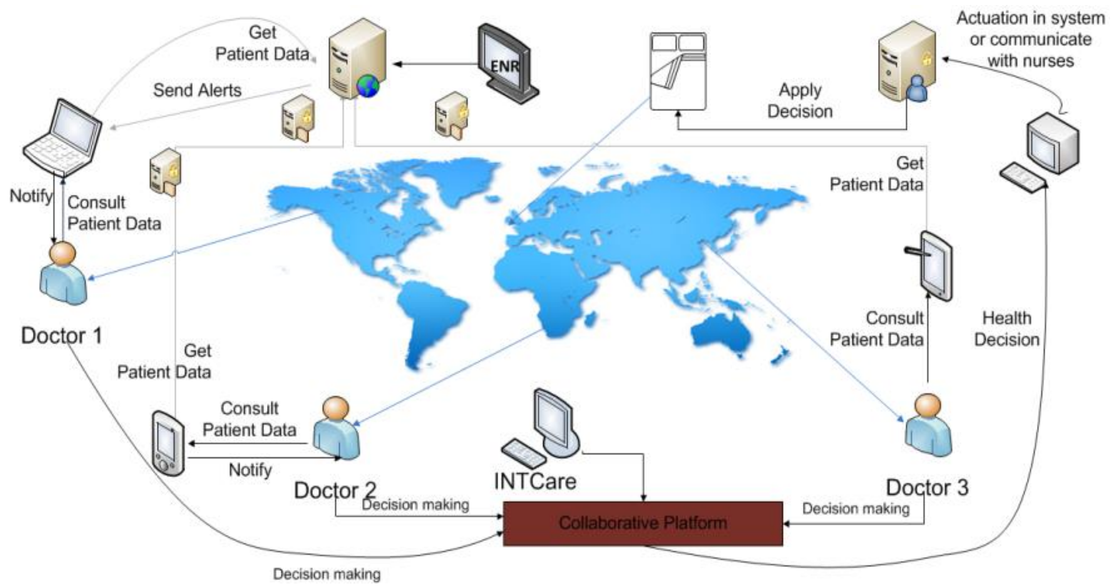


Figure II.3 – Hospital Pervasive Access.

Now, at any time, the user (ICU professional) can consult patient's information by some device regardless of where he is. Like was presented before (Villas Boas, et al., 2010) exist many possible communications in the Hospital throughout the world. However is necessary to make some modification in the architecture with the objective to enable a secure connection. The figure II.3 illustrates the alterations introduced, the doctors only can connect to the hospital through online applications in this case the interface will be the ENR. ENR gives the possibility to put, online, all data about the patient. This data will be validating by the doctors / nurses that are present in the ICU and near the patient. The connection is protected and only users with access privileges can connect to the hospital and see the data. This group will be limited and if someone needs access to the system, have to request access to the team in charge of the ICU, with this option we can control who is accessing to the system remotely.

Out of the ICU the doctors only can consult the data; the validation is not possible because they are so far the patient. The patient is in ICU on Porto, doctors make a secure connection to the hospital, regardless of where they are, the time at which they are accessing and the device they are using, they can consult and analyse patient data. If they see something wrong like bad lab results, bad prescriptions, bad previsions or anything else, they can use a collaborative platform and talk with other specialists about what is the best decision / treatment to apply to the patient. In this task, INTCare can be a good help because it gives the best decisions based in some previsions and earlier treatments to particular patient problems.

After the decision, this can be communicated to someone on ICU to accomplish it, may be given an instruction to the system to make an autonomous operation, can be made some treatments changes, or can be prescribed new drugs through drug system. All decision and their results will be stored in database with the objective to improve the decision; this requires an adaptation of the predictive and decision system.

5. A Framework to Evaluate Environment Quality

An important issue is the evaluation of the proposed approach in order to determine its efficiency. The evaluation should incorporate two strands: technical and human. The data from the remote connections will be used to support the technical evaluation. A framework will be defined in order to facilitate the users giving their opinion about their interaction with the environment. The results obtained will be submitted to a panel of experts from several areas (e.g. Information System, Communication, and Health). The system will start working for tests very soon; we are making the latest implementations and modifications in the ICU. This framework considers a set of measures and a scale of possible answers. The scale ranges from one (R1 – the lowest level) to five (R5 – the highest level). Each measure is then associated to particular range, e.g., data access has three possible options R [1-3], and the Reliability System has five possible answers, R [1-5].

5.1. Surveys

To the human evaluation we consider some online questionnaires that will be concerned with the connection (A), the data (B), and the environment and user satisfaction (C). The table II.1 shows the measures and the evaluation scale of each.

Table II.1 – Environment Evaluation Measure

Measure	ID	Evaluation Scale				
		R1	R2	R3	R4	R5
Connection Secure	A_1	No	Yes			
Connection Velocity	A_2	Too Slow	Slow	Normal	Fast	Very Fast
Data Access	A_3	Many Delay	Some Delay	Just In Time		
Reliability System	A_4	None	Little	Some	A Lot	All
Privacy	A_5	No	Yes			
Data Accuracy	B_1	50% or less	51% to 75%	76% to 99%	All	
Electronic Data	B_2	50% or less	51% to 75%	76% to 99%	All	
Online Data	B_3	50% or less	51% to 75%	76% to 99%	All	
Access	C_1	Difficult	Normal	Easy		
Adaptive Environment	C_2	No	Yes			
Interface	C_3	Very Bad	Bad	Good	Very Good	Excellent
Global Satisfaction	C_4	Very Bad	Bad	Good	Very Good	Excellent

5.2. Results Analysis

Understand the environment, realize if it is really intelligent and verify if is working as expected, is the base of success of a pervasive environment. The questionnaires will be filled by people that interacted with the environment inside and outside of ICU. The table I.1 configures the base of the evaluation framework. In this table a scale of measures was defined ranging from R1 (min) to R5 (max). However, for some measures the max score are R2 (A1, A5, C2), R3 (A3, C1) or R4 (B1, B2, B3). For each question, the user will choose an answer according to the possibilities (table 1). Based on the answer one score will be assigned. The final score will be based in the weighted evaluation of each set: connection, Data and Environment and User Satisfaction. The weight of the measure (a percentage value) was defined taking into account the importance of each environment attribute for the ICU. The calculation of each metric is equal to (for all the set):

$$\sum_{k=0}^{n} \frac{K_\chi \text{ x } R_\chi}{M_\chi}$$

Where,

n: is the number of metrics in the group;
K: is the weight associated to this measure;
R: stands for the value of R score for the question chosen by user (e.g. R2 → 2);
M: is the maximum of the score (K x Max R Value to this measure).
χ: corresponds to the question id.

Individually, the metrics are defined by:

Connection	$\chi = \frac{0{,}25A_1 + 0{,}1A_2 + 0{,}3A_3 + 0{,}15A_4 + 0{,}2A_5}{3{,}05}$	**(1)**
Data	$\alpha = \frac{0{,}4B_1 + 0{,}3B_2 + 0{,}3B_3}{4}$	**(2)**
Environment /User Satisfaction	$\beta = \frac{0{,}2X_1 + 0{,}3X_2 + 0{,}2X_3 + 0{,}3X_4}{3{,}7}$	**(3)**
Global Evaluation Environment	$\Omega = \frac{0{,}3\chi + 0{,}4\alpha + 0{,}3\beta}{3}$	**(4)**

According to the results obtained we can make some modifications or improve some aspects. Table II.2 establishes the significance for each metric (1-4) in terms of a classification value ranging from Very Bad to Very Good. These percentages were agreed by the ICU and Information System professionals.

Table II.2 – Results Evaluation

Metric	*Very Bad*	*Bad*	*Acceptable*	*Good*	*Very Good*
1	< = 70%	< = 75%	< = 85%	< = 90%	> 90%
2	< = 75%	< = 85%	< = 90%	< = 99%	= 100%
3	< = 75%	< = 80%	< = 85%	< = 90%	> 90%
4	< = 80%	< = 85%	< = 90%	< = 95%	> 95%

6. Discussion

The features defined and evaluated in this work made possible to transform the system in order to be more secure, robust, easily accessible and intelligent. Is expected the system will improve the patient outcome in the future due to some new facilities like the data availability in online, real-time and in electronic format. With the ICU pervasive access recast, is possible to accede to knowledge portions that can support the decision making process, anytime, anywhere. Beyond the performance of INTCare, the evaluation of the success of the proposed environment depends on the results obtained after a sample of 100 surveys answers that will be done to all people that directly interacted with this new ICU environment and can answer (patient, doctors, nurses, other professionals). Based on the results obtained for each category and according to table 2 we will decide on what to do in the future. The results also can give us indications about if the new environment is or not prepared to implement a remote IDSS like is INTCare.

7. Conclusion

The right technologies in a smart environment can get some important objective: "help health care professionals to manage their tasks while increasing the quality of patient care" (Bricon-Souf & Newman, 2007). Like we can prove, is possible create a Pervasive and Intelligent Critical Health Care Environment for data acquisition and data consult with the max security and reliability for the ICU, their professionals, patients and applications. Almost all tasks can be intelligent and performed autonomous and in real-time. The proposed environment allows for a total availability of data in electronic format wherever we need at the time that we want. However there are some operations that need to be performed manually like is the mode how data is validated. We can change the mode of data is validated. The data can be pre-validated automatically but the final verification has to be done by ICU professionals. We can improve the form how data is validated, but this process always needs to be human. Nothing can fail because patients' lives are at risk. The impact of this on the society will be big, because the doctors can see the patient condition and treat him remotely. New systems have to be created and others need to be modified, because this environment will allow pervasive computing and new features can be added to the actual systems. The principal difference between this model and the previous is in the form of the data is collected, available and how it can be accessed. With the implementation of new environment and the guarantee of the success of operations at a distance, some new treatments can be performed and some lives can be saved.

8. Future work

Now we are analysing the new environment and arrange some meetings with the objective to develop a Risk List and a contingency plan. The objective is to define actions that can avoid problems if something fails on the system. In the future we hope to have the entire ICU adapted to a pervasive and intelligent approach. The prevision and decision models will be improved and prepared to the pervasive environment.

Acknowledgment

The authors would like to thank FCT (Foundation of Science and Technology, Portugal) for the financial support through the contract PTDC/EIA/72819/2006. The work of Filipe Portela was supported by the grant SFRH/BD/70156/2010 from FCT.

References

Apostolakos, M. J., & Papadakos, P. J. (2001). *The Intensive Care Manual*: McGraw-Hill Professional.

Black, J. P., Segmuller, W., Cohen, N., Leiba, B., Misra, A., Ebling, M. R., et al. (2004). *Pervasive Computing in Health Care: Smart Spaces and Enterprise Information Systems*.

Bricon-Souf, N., & Newman, C. R. (2007). Context awareness in health care: A review. [10.1016/j.ijmedinf.2006.01.003]. *International Journal of Medical Informatics, 76*(1), 2-12.

Coiera, E. (2000). When conversation is better than computation. *Journal of the American Medical Informatics Association, 7*(3), 277.

Cook, D. J., & Das, S. K. (2005). *Smart environments: technologies, protocols, and applications*: Wiley-Interscience.

Das, S. K., & Roy, N. (2008). Learning, Prediction and Mediation of Context Uncertainty in Smart Pervasive Environments. In R. Meersman, Z. Tari & P. Herrero (Eds.), *On the Move to Meaningful Internet Systems: Otm 2008 Workshops* (Vol. 5333, pp. 820-829). Berlin: Springer-Verlag Berlin.

Gago, P., & Santos, M. F. (2008). *Towards an Intelligent Decision Support System for Intensive Care Units*. Paper presented at the 18th European Conference on Artificial Intelligence.

Gago, P., Santos, M. F., Silva, Á., Cortez, P., Neves, J., & Gomes, L. (2006). INTCare: a knowledge discovery based intelligent decision support system for intensive care medicine. *Journal of Decision Systems*.

Gago, P., Silva, A., & Santos, M. F. (2007, Dec 03-07). *Adaptive decision support for intensive care.* Paper presented at the 13th Portuguese Conference on Artificial Intelligence, Guimaraes, PORTUGAL.

Gao, S., & Xu, D. (2009). Conceptual modeling and development of an intelligent agent-assisted decision support system for anti-money laundering. *Expert Systems with Applications, 36*(2), 1493-1504.

Kalli, S., Ambroso, C., Gregory, R., Heikelä, A., Ilomäki, A., Leaning, M., et al. (1993). Inform: Conceptual modelling of intensive care information systems. *Journal of Clinical Monitoring and Computing, 9*(2), 85-94.

O'Donoghue, J., Herbert, J., & Sammon, D. (2008). Patient sensors: A data quality perspective. In S. Helal, S. Mitra, J. Wong, C. Chang & M. Mokhtari (Eds.), *Smart Homes and Health Telematics* (Vol. 5120, pp. 54-61). Berlin: Springer-Verlag Berlin.

Orwat, C., Graefe, A., & Faulwasser, T. (2008). Towards pervasive computing in health care - A literature review. [10.1186/1472-6947-8-26]. *BMC Medical Informatics and Decision Making, 8*(1), 26.

Portela, F., Santos, M., Vilas-Boas, M., Rua, F., Silva, Á., & Neves, J. (2010). *Real-time Intelligent decision support in intensive medicine*. Paper presented at the *KMIS* 2010- International Conference on Knowledge Management and Information Sharing.

Portela, F., Vilas-Boas, M., Santos, M. F., & Fernando, R. (2010). *Improvements in data quality for decision support in Intensive Care*. Paper presented at the eHealth 2010 - 3rd International ICST Conference on Electronic Healthcare for the 21st century.

Santos, M., Portela, F., Boas, M., Machado, J., Abelha, A., Neves, J., et al. (2009). *Intelligent Decision Support in Intensive Care*
Units Nursing Information Requirements. Paper presented at the APPLIED INFORMATICS AND COMMUNICATIONS (AIC '09).

Santos, M. F., Cortez, P., Gago, P., Silva, Á., & Rua, F. (2006). *Intelligent decision support in Intensive Care Medicine.* Paper presented at the 2nd International Conference on Knowledge Engineering and Decision Support, Lisbon, Portugal.

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., & Neves, J. (2009). Information Architecture for Intelligent Decision Support in Intensive Medicine. *8th WSEAS International Conference on APPLIED COMPUTER & APPLIED COMPUTATIONAL SCIENCE (ACACOS '09), 8*(5), 810-819.

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., & Neves, J. (2011). *INTCARE - Multi-agent approach for real-time Intelligent Decision Support in Intensive Medicine*. Paper presented at the 3rd International Conference on Agents and Artificial Intelligence (*ICAART*).

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., Neves, J., et al. (2009). Information Modeling for Real-Time Decision Support in Intensive Medicine. In S. Y. Chen & Q. Li (Eds.), *Proceedings of the 8th Wseas International Conference on Applied Computer and Applied Computational Science - Applied Computer and Applied Computational Science* (pp. 360-365). Athens: World Scientific and Engineering Acad and Soc.

Silva, Á., Cortez, P., Santos, M. F., Gomes, L., & Neves, J. (2004). *Multiple organ failure diagnosis using adverse events and neural networks.* Paper presented at the 6 th International Conference on Enterprise Information Systems.

Silva, Á., Pereira, J., Santos, M., Gomes, L., & Neves, J. (2003). *Organ failure prediction based on clinical adverse events: a cluster model approach*, 3th International Conference on Artificial Intelligence and Applications.

Suter, P., Armaganidis, A., Beaufils, F., Bonfill, X., Burchardi, H., Cook, D., et al. (1994). Predicting outcome in ICU patients. *Intensive Care Medicine, 20*(5), 390-397.

Varshney, U. (2007). Pervasive healthcare and wireless health monitoring. *Mobile Networks and Applications, 12*(2), 113-127.

Vilas-Boas, M., Santos, M. F., Portela, F., Silva, Á., & Rua, F. (2010). *Hourly prediction of organ failure and outcome in intensive care based on data mining techniques*. Paper presented at the 12th International Conference on Enterprise Information Systems.

Villas Boas, M., Gago, P., Portela, F., Rua, F., Silva, Á., & Santos, M. F. (2010). *Distributed and real-time Data Mining in the Intensive Care Unit*. Paper presented at the 19th European Conference on Artificial Intelligence - ECAI 2010.

II – Data Overview

ARTICLE A3

Objectives

To adopt pervasive approaches in critical environments

i. To study the processes inherent to ICU toward to its automation and dematerialization;

The Extract, Transform and Load (ETL) phase is one of the most critical as many of the data collected needs to be validate and treated, with the objective to present data with a high level of quality and thus be able to get the best prediction models. To automate the data acquisition process and the ETL phase the development of some agents were necessary.

Results

The results obtained after the KDD process enforcement, were essentially related with the data quality. The new forms defined and implemented to collect, store and validate the data increase significantly the number of real values. The same evolution was verified in the number of data collected with values out of the range defined by ICU. The processing and data process was divided into some individual tasks, this tasks use the data collected in real-time and according the ranges of ICU, accept or not the values collected. The development of intelligent agents and the integration of a Multi agent system is a good choice when intended make some process autonomous. The modification presented increases the IDSS performance giving guarantees of execution success in order to support a decision process in critical areas as is Intensive Medicine.

This paper presented a data engineering process for an IDSS in Intensive Medicine with base in the processing and transformation of data collected in real-time. The results achieved are very important to improve the data mining models, acquire new knowledge and increase the sensibility and accuracy values inherent to the prevision models. Concluding, the INTCare data processing phase is now well defined and is prepared to run autonomously by the agents in charge to process and transform the data collected which makes possible the creation of real-time models.

ENABLING REAL-TIME INTELLIGENT DECISION SUPPORT IN INTENSIVE CARE

Filipe Portela[1], Manuel Filipe Santos[1], Pedro Gago[2], Álvaro Silva[3], Fernando Rua[3], António Abelha[4], José Machado[4] and José Neves[4]

[1] Centro Algoritmi, Universidade do Minho, Guimarães, Portugal
{cfp,mfs}@dsi.umnho.pt

[2]ESTG, Instituto Politécnico de Leiria, Leiria, Portugal
pedro.gago@ipleiria.pt

[3]Centro Hospitalar do Porto, Porto, Portugal
moreirasilva@clix.pt fernandorua.sci@hgsa.min-saude.pt

[4]CCTC, Universidade do Minho, Braga, Portugal
{jmac, abelha, jneves}@di.uminho.pt

Abstract Medical devices in ICU allow for both continuous monitoring of patients and data collection. Nevertheless, the amount of data to be considered is such that it is difficult for doctors to extract all the useful knowledge. In order to help uncover some of that knowledge we have built an IDSS based in the agent's paradigm and using data mining techniques to build prediction models. With the intention of collecting as much data as possible the data acquisition process was automated. Furthermore, given the paramount importance of data quality for data mining a data quality agent responsible for detecting the errors in the data was devised. Indeed, data acquisition in the ICU is error prone as, for instance, sensors may be displaced as patients move. The aim of this paper is to present the overall KDD process implemented, presenting in detail the data transformations that were done and the benefits achieved.

Keywords: Real-Time, Intelligent Decision Support, Data Engineering, Data Mining, Intensive Medicine, Agents, KDD.

1. Introduction

Intensive Medicine is a critical health area where lots of data are continuously collected. Most of it comes from electronic medical devices present in the Intensive Care Units (ICU) and are associated with some patient. However, this automated data collection is usually done without any type of data treatment or selection and because of that many of these values are null or considered out of range. Moreover, the volume of data is such that if makes it difficult for the medical staff to interpret this data. In addition, they can't extract useful knowledge and quickly understand which are the true values. We can argue that this delay and the associated indecision about the values collected can keep doctors from reaching lifesaving conclusions that could have been reached if adequate decision support tools were present. With the objective to help the doctors in their decision making process and give to them the real values about the patient, is necessary develop an Intelligent Decision Support System (IDSS) that operates in a real-time and is able to present the data with a very fine granularity. The success of this system depends on its ability to fulfil all the tasks associated with the Knowledge Discovery in Databases (KDD) process.

The Extract, Transform and Load (ETL) phase is one of the most critical as many of the data collected needs to be validate and treated, with the objective to present data with a high level of quality and thus be able to get the best prediction models. To automate the data acquisition process and the ETL phase the development of some agents were necessary.

This paper is divided in five chapters, the first and second introduces the theme and all concepts related. The third chapter is the base of this paper and presents all phases of KDD that were completed on the ICU, some processing and transforming tasks performed, and some data mining result obtained. The last chapters present a little analysis of the results obtained with the completion of KDD, the conclusions of those and some future work to this area.

2. Background

2.1. Intensive Medicine

Intensive Medicine (IM) is a "Multidisciplinary area of Medical Sciences that specifically addresses the prevention, diagnosis and treatment of acute potentially reversible in patients with failure of one or more vital functions" (Á. Silva, 2007). This is a particular environment were anyone and anything can fail. The patients admitted to Intensive Care Units are in serious conditions. The critical care medicine allows the recovery of the patients in terminally ill or in a state of organ failure.

The quickly patient recovery depends largely, the patient data quality and decisions taken in the ICU. The main objectives are to diagnose, monitor and treat patients with serious illnesses and recover them to the quality of life they had prior to being admitted into the ICU (Suter, et al., 1994). In the ICU physiological patient variables such as heart rate, blood pressure, temperature, ventilation and brain activity are constantly monitored on-line (Mahmoud, 2003) and others like administered drugs and fluid balance are registered Hourly. Normally for the doctors is very difficult to interpret this data quickly and in useful time.

2.2. Decision Making Process

The process of decision making is a key point in critical areas such as critical care medicine, or assessment of prognosis and diagnosis or treatment of the patient (Á. Silva, 2007), since, according to some studies, the medical error can be the eighth leading cause of death in industrialized countries (Kohn, et al., 2000). Normally this process is complex, need to be done quickly and deals with a large number of variables that are in a constant modification.

Currently in the ICUs, doctors analyses the clinical data and relying on its experience they decide whether or not to intervene. The information usually comes from the monitors that are placed beside the bed, or notes that are taken periodically. This information is used to find out what action to take, which should be the course of this and the treatments to apply.

However, we can verify that the information provided sometimes isn’t enough, isn't displayed in correct way or it arrives only after a decision has been made.

The decision making process is a continuous process that encompasses four phases: Intelligence, Decision, Choice and Implementation (Efraim Turban, et al., 2010).

After the implementation, the results obtained must be monitored and should be used to reformulate the models and these will be used with new sources on the consequently intelligence phase.

Touching the phases of the decision making process and analysing the area of critical care medicine there is none IDSS that follow these parameters. The INTCare system seeks to correct the gap in the ICU (Gago, et al., 2006; Filipe Portela, Santos, et al., 2010), in the provision of information and rapid response to a problem.

2.3. Real-Time

While never easy, the development of real-time systems is particularly demanding in critical areas where every second counts and nothing can fail. Such settings make it very difficult and even dangerous to develop this type of systems. Real-Time is a particularly feature of this systems and involves a lot of others characteristics like online learning, data security, flawless processing and autonomous process. In this cases any process that we try modify or create is a new and entire challenge for the human resources, technologies and the environment (Scicluna, et al., 2008).

The data acquisition process is one of the most critical phases because technical, human and environment factors are involved and may condition the quality of the data acquired and the success of each task (Filipe Portela, Santos, et al., 2010). However, in order to have an IDSS that can operate in real-time all tasks need to be developed in real-time mode, the data needs to be ready when they are necessary and we have to automate the highest number of tasks possible (Katharina Morik, Imboff, Brockhausen, Joachims, & Gather, 2000) .

2.4. Knowledge Discovery Process

Before you get the knowledge, the KDD process goes through five phases: Selection, Pre-Processing, Processing, Data Mining and Interpretation (Fayyad, et al., 1996). Figure III.1 presents the KDD process and the results obtained at the end of each task. The process starts with the raw data obtained in the sources (databases) and culminates with the obtaining of new knowledge.

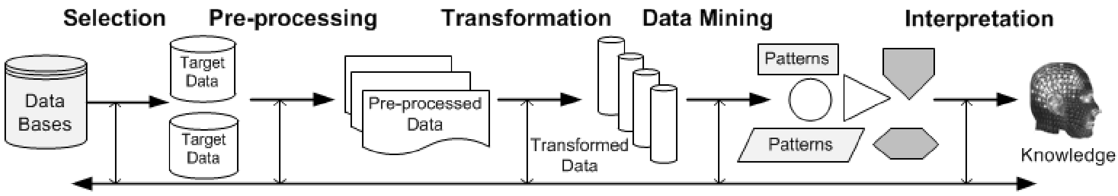


Figure III.1 – KDD Process.

The Extract, transform and load (ETL) process involves the three initial phases and consists in the extraction of data from the sources, transformation of these data and loading the final data to the data warehouse, i.e. prepare the data to be used by data mining algorithms.

2.5. INTCare

INTCare is a research project which main goal is to develop an Intelligent Decision Support System for Intensive Medicine that is capable to predict the patient outcome and the organ failure (Gago, et al., 2006; Gago, et al., 2007) in real-time (Vilas-Boas, et al., 2010a). The system is being developed and tested in the ICU of the Hospital Santo António (HSA) in Porto, Portugal.

The INTCare system is divided into four subsystems: data acquisition, knowledge management, inference and interface (Filipe Portela, Santos, et al., 2010) and uses intelligent agents (Manuel Filipe. Santos, et al., 2011a) who are responsible for a range of activities, such as the automation of data collection and updating of the predictive models in real-time, without the need for human intervention.

2.6. Multi-agent System

An agent is an intelligent program that acts autonomously on behalf of the user." (Manuel Filipe. Santos, et al., 2011a) The intelligent agents used by INTCare are capable to performs autonomous actions in order to meet its goals (Gago, et al., 2006), (Jennings, 2000). With the objective to support the KDD process a multi-agent system was integrated in INTCare. This agents are responsible for perform, automatically, all tasks of data processing and data transformation.

3. ICU KDD Process

Data Engineering is a continuous and adaptive process. We started this process two years ago when after some requirements analyses for INTCare and consequent necessities of ICU we defined the reformulation of the information system architecture.

The base of the entire process was: the results obtained in the past (offline) (Á. Silva, 2007; Álvaro Silva, et al., 2008), the necessity of innovation, systems integration, dematerialization of processes, turn all information electronic and make it available online and in real-time eliminating the high number of errors present in the patients records. The first challenge in the ICU, consisted of the change of record method and in the obtaining of some access privileges to some data sources. In 2009 80% of registers were made on paper, weren't available electronically (e.g. PDF format) or could not be accessed.

Nowadays, all data that are needed to create the models and obtain new knowledge, are computerized and are available online and in real-time, except to some blood gas results that is collected by ICU nurses and is not present on the normal lab results.

Table III.1 presents the evolutions that took place in the ICU in the last 2 years, i.e., an analysis of the data that we now can monitor (consult, edit, validate) and document (save) electronically. Most of the changes were achieved with the development and introduction in the ICU, of the ENR (Filipe Portela, Vilas-Boas, et al., 2010; M.F. Santos, F. Portela, M. Vilas-Boas, J. Machado, A. Abelha, J. Neves, et al., 2009) and with the access to Electronic Health Process by AIDA (Abelha et al., 2003).

Table III.1 – Monitoring and Documentation of data sources

Data Sources / Variables	2009		2011	
	Monitor	Document	Monitor	Document
Nursing Record	X	PDF	√	√
Vital Signs	√	X	√	√
Drug System	√	Paper	√	√
Patient EHR	X	PDF / Paper	√	√
Fluid balance	√	Paper	√	√
Procedures	√	√	√	√
Lab Results	X	PDF	√	√
Patient Events	X	X	√	√
Ventilations	√	Paper	√	√
Patient Scores	√	Paper	√	√
ICU Scores	√	Paper	√	√
ICU Blood Gas	X	X	X	X

Figure III.2 shows the ICU KDD process. The database is populated with data from seven major sources. The data will be selected to the data warehouse to be processes or transformed, depending of the use goal to each variable. After this task the data will be available to be presented by the ENR and prepared for the creation of Data Mining Models. Finally all models will be evaluated and the knowledge obtained will be presented in INTCare.

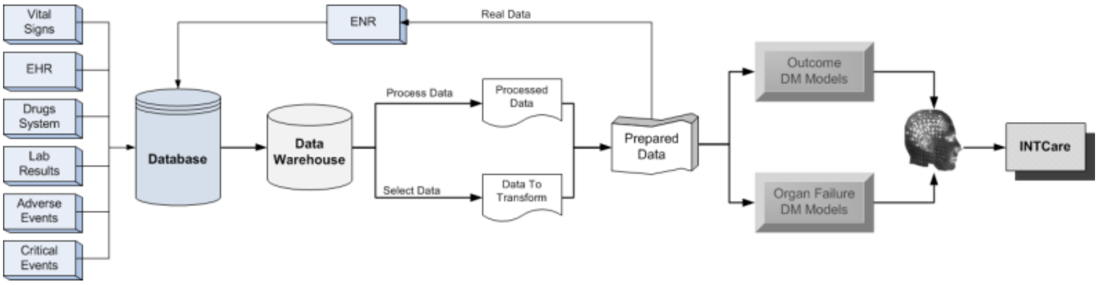


Figure III.2 – ICU Knowledge Discovery in Database Process.

3.1. Selection

After the modifications we can obtain a lot of information regularly (up to once per minute) to be pre-processed and transformed. The ICU Database is composed by various tables that store all data collected by the various systems. The database is compose by tables that contain the raw data collected by gateway and tables that store the data that were processed and transformed (real data). The data are provided from a lot of systems that are integrated by a single application: the Electronic Nursing Record (ENR) that presents all data collected. The ENR gives to the nurses and the doctors the possibility to consult, edit and validate a lot of values and results related with the patients. The Database integrates the following tables:

ICU_HL7	$\subseteq$	{Vital Signs}
ICU_HL7_T	$\subseteq$	{Vital Signs auto validated (real values)}
ICU_PARAM	$\subseteq$	{ICU Limits (max, min) values}
ICU_LR	$\subseteq$	{All Lab Results}
ICU_DRUGS	$\subseteq$	{All Patient Drugs administrated}
ICU_ENR	$\subseteq$	{Data validated and provided from ENR}
ICU_ADEVENTS	$\subseteq$	{ICU Adverse Patient Events}
ICU_CEVENTS	$\subseteq$	{ICU Critical Patient Events}
ICU_PATIENT	$\subseteq$	{HER, Patient Information (admission, outcome, age}
ICU_DATABASE	$\cup$	{ICU_HL7, ICU_HL7_T, ICU_PARAM, ICU_DRUGS, ICU_LR, UCI_ENR, ICU_ADEVENTS, UCI_CEVENTS, UCI_PATIENT}

The data necessary to create the models and the Real-Time IDSS will be selected from these tables.

3.2. Pre-Processing

All data will be processed and transformed even if it is not necessary for the creation of the data mining models currently in use. This option makes possible, in the future, the introduction of new variables on the models (adaptive features).

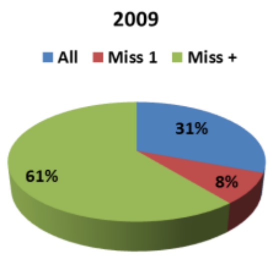
All data collected that have null values and values out of the range will be treated. This operation will be two phases: the auto processing and the human validation. The auto processing is based in the achieving of automatic task, by intelligent agents (Manuel Filipe. Santos, et al., 2011a), which will run after receipt some values. The validation of values by the nurses is needed, because they are the only that are near the patient bed and can validate the veracity of the values. This validation will be required to specific variables that due the critical area that we are can have dubious values. This phase will be autonomous and will be done by the pre-processing agent and is concerned with the problems:

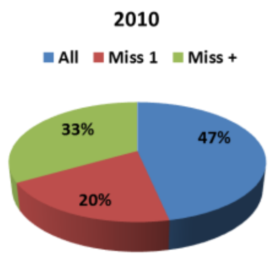
a) No patient ID;
b) Null Values; and
c) Values out of range.

3.2.1. No patient ID

The process of collecting data from the gateway without patient identification (PID) is something normal in an ICU, because the PID is manually inserted by the nurses and isn't usually considered important as patient treatment can proceed even if no PID is registered in the system. Nonetheless, this is a very important field for automated systems and as such we needed to have it filled. Our first option was to motivate the nursing staff to insert the PID and we noticed some benefits to the system. However the results weren't as expected. We witnessed a decrease in the number of missing PID but this value continued very high. In 2011 we opted for another solution: auto-update hl7 message.

The next 3 figures (3, 4, and 5) present the evolution (2009 - 2011) of the collect of data of patient with: all days, the PID failing one day or failing more than one day.





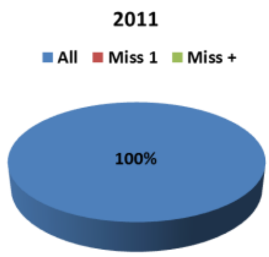


Figure III.3 – PID in 2009. **Figure III.4 –** PID in 2010. **Figure III.5 –** PID in 2011.

Ω – Patient ID collected
β – Patient ID on EHR
γ – Bed ID
δ – ICU ID
θ – Date of discharge.

if Ω = null then
get β in (γ, δ, θ)
set $\Omega = \beta$

A modified version of this formula was used to recover the vital signs data that were collected in the past and stored into the database without patient identification. A SQL function was developed to process each row in the table. The date and bed number early stored are compared with the admission and discharge date and the bed number stored in the EHR on that date (message date).

For each row, it will try to find the correct PID on the EHR and, if it found it will substitute, in the PID column of the row without ID of the patient (i.e. PID = null), with the number of process of the patient (PID) that was admitted in the ICU in this date.

Ω – Patient ID collected
β – Patient ID on EHR
γ – Bed ID
δ – ICU ID
¥ – message date
– Date of discharge.
φ –Date of Admission

if Ω = null then
get β in (γ, δ)
where φ >= ¥ <= θ
set $\Omega = \beta$

3.2.2. Null Values

The null values occur when some sensor was disconnected from the patient and continues collecting the data. The solution found for this problem is simple. The function (2) will delete the row with null values:

Ω – database row γ – result

β – variable id

if β is null or γ is null

delete Ω

3.2.3. Values out of range

This is a critical task and is based on a set of analyses that will determine if a value is possible or not in the ICU environment, i.e. if the value collected from a patient is or not true. If the value collected be within the range previously defined this will be validated e inserted on the table with the real values, otherwise this value will be inserted in a different table. The equation (3) shows those operations:

Ω – Database Row δ – min_value

β – variable_id θ – max_value

γ – result

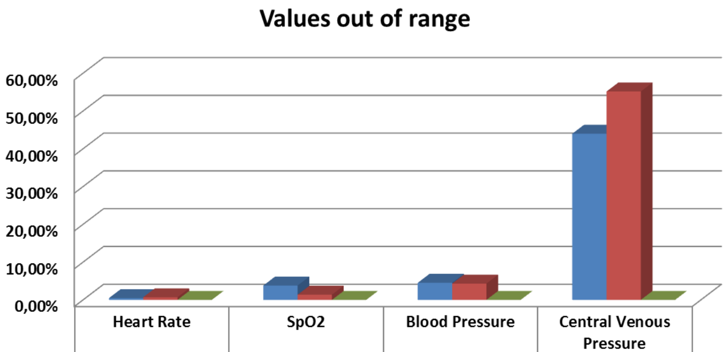
if γ >= δ in (β) and γ <= θ in (β)

insert Ω on table_in

else

insert Ω on table_out

Now, like shows figure III.6, ICU vital signs tables don't have any bad values stored because, if the gateway receive a value out of the range defined by ICU, the vital signs agent ignore that hl7 message and didn't store this value in database.



	Heart Rate	SpO2	Blood Pressure	Central Venous Pressure
2009	0,45%	3,79%	4,51%	43,96%
2010	0,69%	1,36%	4,29%	55,14%
2011	0,00%	0,00%	0,00%	0,00%

Figure III.6 – Example of comparison of values out of range.

The manual validation of values will be done in hourly in the ENR for: vital signs, drugs system and fluid balance. The values that will appear in the ENR will be the data that were automatic validate and are present in real data table. With this option is possible, if necessary, correct the bad values or insert values out of defined range if that happen with any patient.

3.3. Transformation

Data collect as shown in Figure III.2 will be divided in two parts, one that only needs to be processed and validated and other that must undergo some transformations. These transformations are necessary in order to calculate some variables necessary for the development of the DM models (Manuel Filipe Santos & Portela, 2011):

a) Critical Events;
b) Accumulated Critical Events;
c) SOFA;
d) SPAPS;
e) MEWS.

3.4. Critical Events

The calculation of critical events (4) is done with the base of the maximum values defined for ICU, the limits of a normal value and the time for an event be critical (Álvaro Silva, et al., 2008). First is verified if the value is normal or critical, next, will calculated the time of the event and stored in the critical events table, an identification of the event.

Ω – Database Row
β – variable_id
Ω – event_id (id, β, ¥, PID, Ͼ, Ͽ)
β – variable_id
γ – result
δ – min_value
θ – max_value
ρ – min_normal_value
φ – max_normal_value

δ – min_value
θ – max_value
Ɯ – value_type
Ͼ – event_date_start
Ͽ – event_date_finish
¥ – event_time
§ – critical_event_table
Δ – time_to_event_be_critic
Θ – acumulate_time_to_critic

if γ >= δ in (β) and γ <= θ in (β) and (γ <= ρ in (β) or γ <= φ in (β))
set Ɯ = 1 else set Ɯ = 0
¥ = sum (Ͽ - Ͼ)
If ¥ >= Δ or sum (¥) >= Θ
insert Ω in §

3.5. Scores

For all scores the method of calculation (5) is similar. The number allocated will be according the score punctuation table (Gardner-Thorpe, Love, Wrightson, Walsh, & Keeling, 2006; Le Gall, et al., 1993; J. Vincent et al., 1998). Using the points associated the final score will be calculated and will the result will be inserted in the scores table.

Ω – final score $\quad\quad$ θ – final_score_category

β – variable_id $\quad\quad$ § – score_table

γ – result $\quad\quad$ φ – score_id (PID, Ω, θ)

δ – score_result_punctuation

$$\Omega = \text{sum}(\ \delta(\gamma, \beta))$$
$$\Omega \Leftrightarrow \theta$$
$$\text{insert } \varphi \text{ in } \S$$

The process presented before are defined in database. To obtain all data some other functions, procedures or triggers were associated to the ICU database tables.

To perform these tasks, the multi-agent system for KDD process will be used, with the configurations defined the processes will start automatically trough some hourly schedule or are activated after some database operation (insert or update). The results obtained after each task will be stored in database according the necessary data and structures defined for the DM models.

3.6. Data Mining

Being of utmost importance to INTCare's performance the knowledge management module provides the functionalities that allow the system to learn from the existing data. This subsystem is composed by three agents: Data Mining, Performance and Ensemble. Together they implement the Data Mining step of the KDD process.

The creation of predictive models is the responsibility of the Data Mining agent. Given a prediction objective, this agent retrieves the relevant data from the database and prepares it to be used by the chosen data mining algorithm. Next it creates a predictive model and stores it in the Knowledge Base.

Nevertheless, we can only have an autonomous system if it is capable of knowing what is going on and can evaluate itself in order to continue to operate with good performance levels. A Performance agent does exactly that by continuously sifting through the data that is being collected and updating a set of statistics that allow it to populate a Performance Database (Gago, et al., 2006; Filipe Portela, Santos, et al., 2010).

It analyses the new data that was stored in data warehouse and verifies the performance of the prediction models through DM.

If the collected statistics show that the performance has fallen below a predefined parameter action must be taken in order to try to correct that situation (Gago, et al., 2006; Filipe Portela, Santos, et al., 2010).

One of the actions that can be triggered by the Performance agent is the replacement of a prediction model by another one built using the most recent available data. However, this procedure is frowned upon by doctors as it forces learning to happen in a sequence of "jumps" thus preventing a smooth learning curve for the system. Also, it is know that ensemble systems lead to better predictive results (as long as some conditions are met) (Dietterich, 2000). Hence, an Ensemble agent was created to enhance predictive performance by combining several models in order to produce models with better results.

3.7. Evaluation

Some work has been done in evaluating the effectiveness of INTCare. In one of the experiments we divided the available data into two mutually exclusive datasets. Models were created using the first dataset. Those models were then evaluated using the second dataset. That work showed that allowing for dynamic changes in the ensemble of predictive models gave way to better results (P. Gago & M. Santos, 2009) (Gago, et al., 2006; Filipe Portela, Santos, et al., 2010). This can be seen in Figure III.7 where we show the increase in the area under the ROC curve (Zweig & Campbell, 1993) when we used this approach versus a traditional (static) ensemble.

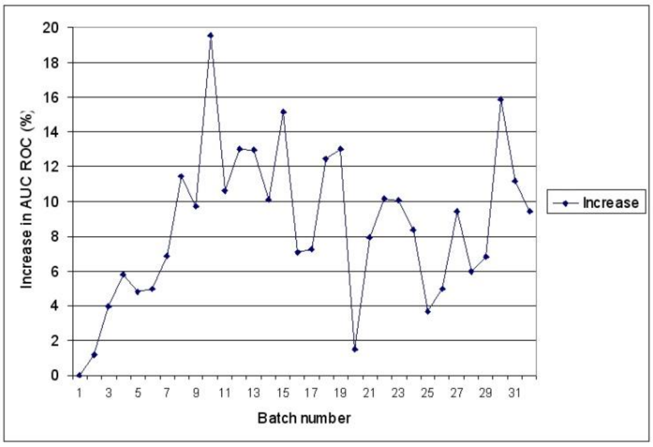


Figure III.7 – Increase in AUC ROC (%).

4. Discussion

The results obtained after the KDD process enforcement, were essentially related with the data quality. The new forms defined and implemented to collect, store and validate the data increase significantly the number of real values.

A major problem that normally was reported by the ICU staff was related with the lack of patient ID in the records collected automatically. Like we can verify in the chart 3 the success of our modifications and implementations, reaching a fantastic number of 0% of data collected without PID.

The same evolution was verified in the number of data collected with values out of the range defined by ICU. The biggest difference was verified in the Central Venous Pressure where the values down from 55% to 0%. Other variable that wasn't represented in the example (chart 4) but also was a common problem in ICU and now is resolved, was the body temperature. This variable sometimes had the sensor disconnected from the patient and collected the environment temperature, now this result is ignored by the agents.

The processing and data process was divided into some individual tasks, this tasks use the data collected in real-time and according the ranges of ICU, accept or not the values collected. On the other hand the data needed to for transformation and preparation is available to the automatic calculation of UCI Scores in order to integrate the data mining models, studied before.

The development of intelligent agents and the integration of a Multi agent system is a good choice when intended make some process autonomous. Improve the DM engine and obtain the better result as possible, having all information always updated, privileged models adaptive and in real-time and online processing.

The modification presented increases the IDSS performance giving guarantees of execution success in order to support a decision process in critical areas as is Intensive Medicine.

5. Conclusion and Future Work

This paper presented a data engineering process for an IDSS in Intensive Medicine with base in the processing and transformation of data collected in real-time.

The results achieved are very important to improve the data mining models, acquire new knowledge and increase the sensibility and accuracy values inherent to the prevision models. Concluding, the INTCare data processing phase is now well defined and is prepared to run autonomously by the agents in charge to process and transform the data collected which makes possible the creation of real-time models. However, if some new variables or ICU configurations appear, this process can be adapted in order to obtain the best data quality possible.

In the future work, new models will be created using the data collected in real-time and treated using the KDD process presented in this paper. Also these models will be tested in a pervasive approach that are being developed and allow accessing by the ICU professionals to the Intelligent Decision Support System anywhere, anytime, with privileges access and security restrictions.

Acknowledgements

The authors would like to thank FCT (Foundation of Science and Technology, Portugal) for the financial support through the contract PTDC/EIA/72819/2006. The work of Filipe Portela was supported by the grant SFRH/BD/70156/2010 from FCT.

References

Abelha, A., Machado, M., Santos, M., Sollari, A., Rua, F., Paiva, M., et al. (2003). Agency for Archive, Integration and Diffusion of Medical Information. *Proceeding of AIA*.

Dietterich, T. (2000). Ensemble methods in machine learning. *Multiple classifier systems*, 1-15.

Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery: an overview.

Gago, P., & Santos, M. (2009). Evaluating Hybrid Ensembles for Intelligent Decision Support for Intensive Care. *Applications of Supervised and Unsupervised Ensemble Methods*, 251-265.

Gago, P., Santos, M. F., Silva, Á., Cortez, P., Neves, J., & Gomes, L. (2006). INTCare: a knowledge discovery based intelligent decision support system for intensive care medicine. *Journal of Decision Systems*.

Gago, P., Silva, A., & Santos, M. F. (2007, Dec 03-07). *Adaptive decision support for intensive care.* Paper presented at the 13th Portuguese Conference on Artificial Intelligence, Guimaraes, PORTUGAL.

Gardner-Thorpe, J., Love, N., Wrightson, J., Walsh, S., & Keeling, N. (2006). The value of Modified Early Warning Score (MEWS) in surgical in-patients: a prospective observational study. *Annals of The Royal College of Surgeons of England, 88*(6), 571.

Jennings, N. R. (2000). On agent-based software engineering. *Artificial Intelligence, 117*(2), 277-296.

Kohn, L. T., Corrigan, J., & Donaldson, M. S. (2000). *To Err Is Human: Building a Safer Health System*: National Academy Press.

Le Gall, J. R., Lemeshow, S., & Saulnier, F. (1993). A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA, 270*(24), 2957-2963.

Mahmoud, M. (2003). Real-time data acquisition system for monitoring patients in intensive care unit. (Vol. 5090, pp. 320-326). Multisensor, Multisource Information Fusion: Architectures, Algorithms and Applications.

Morik, K., Imboff, M., Brockhausen, P., Joachims, T., & Gather, U. (2000). Knowledge discovery and knowledge validation in intensive care. *Artificial Intelligence in Medicine, 19*(3), 225-249.

Portela, F., Santos, M., Vilas-Boas, M., Rua, F., Silva, Á., & Neves, J. (2010). *Real-time Intelligent decision support in intensive medicine*. Paper presented at the *KMIS* 2010- International Conference on Knowledge Management and Information Sharing.

Portela, F., Vilas-Boas, M., Santos, M. F., & Fernando, R. (2010). *Improvements in data quality for decision support in Intensive Care*. Paper presented at the eHealth 2010 - 3rd International ICST Conference on Electronic Healthcare for the 21st century.

Santos, M. F., & Portela, F. (2011). *Enabling Ubiquitous Data Mining in Intensive Care - Features selection and data pre-processing*. Paper presented at the 13th International Conference on Enterprise Information Systems.

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., & Neves, J. (2011). *INTCARE - Multi-agent approach for real-time Intelligent Decision Support in Intensive Medicine*. Paper presented at the 3rd International Conference on Agents and Artificial Intelligence (*ICAART*).

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., Neves, J., et al. (2009). Intelligent Decision Support in Intensive Care Units - Nursing Information Requirements. In Springer (Ed.), *WSEAS Transactions on INFORMATICS*.

Scicluna, P., Murray, A., Xiao, Y., & Mackenzie, C. F. (2008). Challenges to Real-Time Decision Support in Health Care. *Agency for Healthcare Research and Quality*.

Silva, Á. (2007). *Modelos de Inteligência Artificial na análise da monitorização de eventos clínicos adversos, Disfunção/Falência de órgãos e prognóstico do doente critico.* Universidade do Porto, Porto.

Silva, Á., Cortez, P., Santos, M. F., Gomes, L., & Neves, J. (2008). Rating organ failure via adverse events using data mining in the intensive care unit. *Artificial Intelligence in Medicine, 43*(3), 179-193.

Suter, P., Armaganidis, A., Beaufils, F., Bonfill, X., Burchardi, H., Cook, D., et al. (1994). Predicting outcome in ICU patients. *Intensive Care Medicine, 20*(5), 390-397.

Turban, E., Sharda, R., & Delen, D. (2010). *Decision Support and Business Intelligence Systems* (9th Edition ed.): Prentice Hall.

Vilas-Boas, M., Santos, M. F., Portela, F., Silva, Á., & Rua, F. (2010). *Hourly prediction of organ failure and outcome in intensive care based on data mining techniques*. Paper presented at the 12th International Conference on Enterprise Information Systems.

Vincent, J., Mendonça, A., Cantraine, F., Moreno, R., Takala, J., Suter, P., et al. (1998). 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, 26*, 1793-1800.

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem, 39*(4), 561-577.

III – Results of Decision Support

ARTICLE A4

Objectives

To develop and test pervasive models in Intensive Care

i. To analyse and define the decision-making process;
ii. To create predictive models to help in decision, prognosis, diagnosis and treatment;
iii. To prototype an IDSS;
iv. To test the models developed;
v. To include pervasive capabilities in the IDSS;
vi. To assess the results.

An agenda has been defined in order to triumph over defined barriers, namely:

- ✓ **Millstone 1** - To provide INTCare with a pervasive component able to perform online data acquisition and data processing automatically and in real-time;
- ✓ **Millstone 2** - To explore DM approaches to induce / adapt the models automatically and in real-time, in order to ensure the best performance.

Results

During the project, a continuous action research process was applied. An ensemble approach has been adopted and tested. The ensemble is well suited for online-learning and streaming data. To induce data mining models automatically, autonomously and in real-time, using online- learning in critical environments as is Intensive Medicine, the ensemble approach has been explored. From the six targets considered in this study, three satisfy completely the quality measures defined: outcome, cardiovascular and coagulation. The results are in line with the medical knowledge (Kaplow & Hardin, 2007; J. Vincent et al., 1998; Vincent, et al., 1996). The best models that meet the quality measures are executed using the data of the patient admitted in ICU. Then the results, the probabilities associated to the targets, are shown through INTCare system. This system is deployed by a hospital server and can be assessed anywhere and anytime through laptops, mobile devices or situated displays.

This paper presents the most recent advances achieved in the context of the INTCare project, an IDSS to predict organ failure and outcome of the ICU patients. The results are mainly related to the KDD process. KDD process is currently executed in real-time using online learning and streamed data. This has been possible due to the modifications introduced in the environment and in the information system of the ICU. Intelligent agents are in charge for data acquisition, data processing and data transformation, in order to: induce data mining models; assess the models; and prepare the results to be accessible in the INTCare console.

TOWARDS PERVASIVE AND INTELLIGENT DECISION SUPPORT IN INTENSIVE MEDICINE – A DATA STREAM MINING APPROACH

Portela, F. [1], Santos, M.F. [1], Silva, A. [2]

[1]Centro Algoritmi, University of Minho, Guimarães, Portugal

[2]Serviço de Cuidados Intensivos, CHP Porto – Hospital Santo António, Porto, Portugal

Abstract. Extracting knowledge from multiple and distributed data streams is one of the most significant challenges that should be faced in areas like the intensive medicine. This work presents the latest progresses in the context of the INTCare project. INTCare brings a new approach to intensive care units endowing them with an intelligent decision support system for predicting organ failure and the outcome of the patients. INTCare uses data collected and processed in real-time and combines adaptive data mining models with online-learning. This paper depicts the data stream acquisition component and explores the ensemble approach implemented in order to make possible the data stream mining. A pervasive system was developed to present the results anywhere and anytime. This study considered the data collected from 129 patients, during 105 days, to predict the outcome and the probability of dysfunction for five organic systems: renal, hepatic, respiratory, cardiovascular and coagulation. A quality measure was introduced in order to evaluate the ensemble performance. The results demonstrated that the approach is very suited to predict the outcome and the dysfunction probability of the cardiovascular and coagulation systems.

Keywords: INTCare; Ensembles; Data Mining; Pervasive; Real-time; Online-Learning; Streaming Data; Knowledge Discovery from Databases, Intensive Care;

1. Introduction

Nowadays it is very common to find systems for collecting and for showing data in a user-friendly way. The data collected by these systems are often used to present some results or facts but not to predict goals or to create new knowledge in order to help the decision makers. This gap is notorious in critical and specific areas such as intensive medicine. Devices in Intensive Care Units (ICU) produce a large amount of data for both continuous monitoring of patients and data collecting.

These data can be streamed into databases or data warehouses and then used to extract knowledge. Streaming data in ICU is recognized to be a process where the data is acquired automatically in a continuous way and stored into a database. Alternatively, the data can be streamed into a data warehouse after a transformation process. This approach is normally associated with the data provided from bedside monitors (e.g., vital signs, ventilation parameters). In ICU there are other types of sources whose data are not being acquired continuously. Data like laboratory results (e.g. blood exams), therapeutics, fluid balance and proceedings can be stored at the moment where the results are obtained. All of these data can be acquired and processed in real-time.

Data Stream Mining (DSM) – the process of applying data mining techniques in stream data - using the amount of data available in the ICU is an interesting approach to obtain knowledge useful for the decision making. Guided by such knowledge, Intelligent Decision Support Systems (IDSS) can help the doctors to take a pro-active attitude. INTCare (Gago et al., 2006) is a pervasive IDSS based on intelligent agents (Manuel Filipe Santos et al. 2011a) and DSM models to predict the organ failure and outcome of patients anywhere and anytime.

Early results (Álvaro Silva, Paulo Cortez, Manuel Filipe Santos, Lopes Gomes, & José Neves, 2008) demonstrated that is possible to predict health targets using data collected in an offline mode. These results led to a new challenge: how to develop an IDSS to predict the outcome and organ failure of patients in real-time and in an online mode?

At the ICU, professionals need to decide promptly and accurately. The patient's direct care is always the first priority, meaning that the data recording / storing is in second place (normally, this task is done when the doctors/nurses have free-time). The main barrier to overcome these limitations was changing the data nature and the way that they are collected. More than 85% of required data to induce DM models were in a paper and/or in a closed format:

1) The fluid balance has been recorded in paper based nursing records;
2) The vital signs were available in the bedside monitors but they didn't communicate with the hospital systems. The values were translated manually to the nursing paper sheet (paper) in a hourly base mode;
3) The laboratory results were only available in a closed digital document format;
4) The scores and critical events associated to the patients weren't recorded.

An agenda has been defined in order to triumph over those barriers, namely:

- ✓ **Millstone 1** - To provide INTCare with a pervasive component able to perform online data acquisition and data processing automatically and in real-time;
- ✓ **Millstone 2** - To explore DM approaches to induce / adapt the models automatically and in real-time, in order to ensure the best performance.

During the project, a continuous action research process was applied. An ensemble approach has been adopted and tested. The ensemble is well suited for online-learning and streaming data. 126

different scenarios are maintained and evaluated using Support Vector Machine, Decision Trees and Naïve Bayes in order to predict the failure of coagulation, renal, hepatic, cardiovascular, and respiratory system and the patient outcome. Finally, the knowledge reached by the ensemble is disseminated through the ICU platforms using an online and pervasive interface. The research was carried out in the ICU of Centro Hospitalar do Porto, Portugal.

Remaining sections of the paper include: a description of the materials and methods used in the research work in the section 2; an introduction to the background knowledge and state of the art of the concepts involved in the work, in section 3. Sections 4 to 8 present the tasks inherent to the process of knowledge discovery from databases. All the tasks are described recurring to the agent abstraction. The experimental work realized is described in the section 9. Section 10 discusses the results attained and identifies the limitations of the study. Finally, section 11 concludes the work and points future directions.

2. Materials and Methods

The working methodology is a problem-oriented, based on the instrumentalist research, contributing to make human intervention in real world environments more efficient. In order to achieve the goals, three methodologies have been followed: case study, action research, and implementation. To the development of this research project some methodologies and technologies have been considered: Knowledge Discovery from Databases (KDD) (Fayyad, Piatetsky-Shapiro, & Smyth, 1996; Frawley, Piatetsky-Shapiro, & Matheus, 1992), the decision making process (Shim et al., 2002; Simon, 1960), intelligent agents (Manuel Filipe Santos, et al., 2011a; Michael Wooldridge, 1999; M. Wooldridge & Jennings, 1995) and Data Mining (DM). In particular, Support Vector Machines, Decision Trees, Naïve Bayes and ensembles (Fayyad, et al., 1996; João Gama, Santos, Filipe, Marques, & Cortez, 2010; Han & Kamber, 2006; Kantardzic, 2011; Tamayo et al., 2005; Wirth & Hipp, 2000; Wu et al., 2008) have been used in DM tasks.

The experimental environment includes the ICU of Centro Hospitalar do Porto (CHP). The ICU possesses now 12 beds, each one of them equipped with a complete data acquisition system capable of monitoring vital signs from the patients. All patients do recurrently a set of laboratory exams and are constantly medicated. Their conditions are continuously evaluated making use of scores and critical events (M. F. S. Filipe Portela, José Machado, Álvaro Silva, Fernando Rua, António Abelha, 2012; P. G. Filipe Portela, Manuel Filipe Santos, Álvaro Silva, Fernando Rua, 2012a). All patients admitted in the ICU have also an electronic health record containing the admission and discharge information. These records are maintained permanently by the nursing staff. The data used in this research flows from multiple sources, e.g., Vital Signs Monitors, Electronic Nursing Records (ENR), Lab Information System, and Electronic Health Records (EHR) (Portela, Santos, Silva, Machado, & Abelha, 2011; Portela, Santos, & Vilas-Boas, 2012; Manuel Filipe Santos & Portela, 2011). The interoperability between INTCare and the other systems is assured by the AIDA platform (Abelha et al., 2003; Duarte, Portela, Santos, António, & José, 2011). All the experiments were conducted using the Oracle Data Mining (ODM) technology (Tamayo, et al., 2005).

3. Background

3.1. INTCare

INTCare is an IDSS engineered to help the doctors in their decision making process. The main objective of this system is to avoid the gap between the data used in the ICU and the data that the doctors normally need to support their decisions. Using streaming data (Portela et al., 2010) and data transformation processes in real-time, INTCare predicts the organ failure and patient outcome (Gago, et al., 2006; Manuel Filipe Santos et al., 2011b) for the next 24 hours, during the first five days of stay.

This system is composed by several semi-autonomous agents in charge to: automate the data collection; processing and transforming the data; and predicting the results in real-time. These tasks don't require human intervention. Conceptually, the INTCare system can be decomposed in four subsystems (Figure IV.1): Data Acquisition, Knowledge Management, Inference and Interface (Portela, et al., 2010).

Formally, INTCare is defined as a tuple (Villas Boas et al., 2010):

$$\Xi \equiv < C_{INTCare}, \Delta_{INTCare}, a_{gat}, a_{vsa}, a_{enr}, a_{,lr}\, a_{pp}, a_{cde}, a_{dm,}\, a_{pf,}\, a_{mi,}\, a_{dr,}\, a_{pd,}\, a_{sc,}\, a_{int,}\, a_{ic} >$$

$C_{INTCare}$ is the context and corresponds to a logical theory represented as a triple Lg, Ax, D, where *Lg* stands for an extension to the language of programming logic, *Ax* is a set of axioms over *Lg*, and *D* is a set of inference rules;

$\Delta_{INTCare}$ is the set of bridge rules defining the interaction among the systems' components (the agents);

a_{gat} *to* a_{ic} are the system's agents.

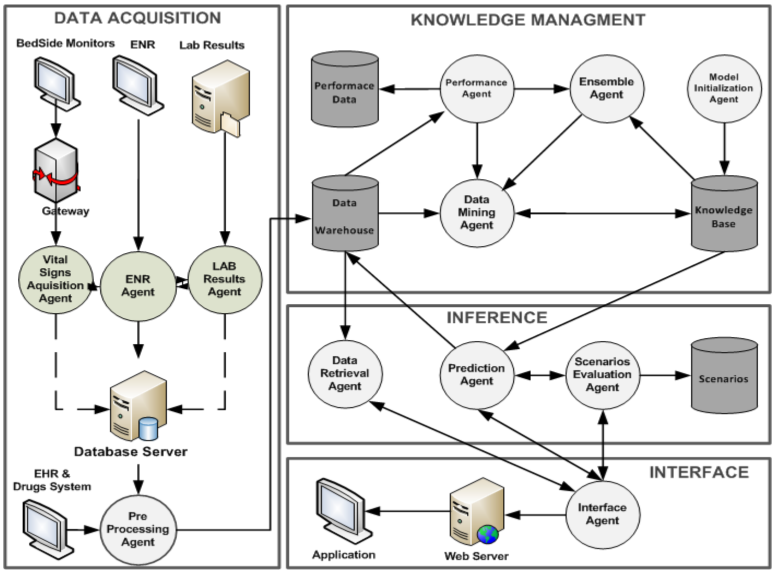


Figure IV.1 – INTCare Architecture.

The features of the INTCare System were defined taking into account the environment, the information needs and the requirements for Data Mining. These features are important for developing a comprehensive system as possible. Several concepts and technologies have been explored, namely (Manuel Filipe Santos & Portela, 2011):

a) Online Learning;
b) Real-Time;
c) Adaptability;
d) Streaming data;
e) Data mining models;
f) Decision models;
g) Optimization;
h) Intelligent agents;
i) Pervasiveness;
j) Accuracy;
k) Safety;
l) Privacy;
m) Secure Access from Exterior;
n) User Policy.

At the same time, two goals related with data mining process were defined as fundamental to the INTCare success:

a) To ensure a correct execution of the real-time process (data acquisition, data transformation and ensemble inducing);
b) To ensure a constant and continuous evaluation of Data Mining models.

INTCare is an agent based system were intelligent agents carry out their tasks autonomously (Manuel Filipe. Santos, et al., 2011a). As depicted in figure IV.1 all the subsystems are intended in terms of agents. The agents are also responsible for the communication between each subsystem. The intelligent agents used are able to act without the human interaction, in order to meet the system goals (Gago, et al., 2006).

Data acquisition agents are used in the KDD process to obtain data from ICU data sources. A pre-processing agent is used for validating the values and for storing them into the data warehouse. Then, Knowledge Management agents are used to induce DM models from the data warehouse and to store them into a knowledge base. Finally, in the inference subsystem, all the scenarios generated by DM are compared and the best result for each target is made available by INTCare. The agent technology allows for developing of an automated system that performs all the tasks in real-time.

3.2. Previous Results

Previous work using data collected manually achieved promising results for organ failure prediction (Silva, 2007; Álvaro Silva, et al., 2008). In this work hourly values were considered for patient biometrics and adverse events. Two data mining methods have been used: artificial neural networks and multinomial logistic regression. Those methods provided a better discrimination and good calibration. Results can be condensed as: 74.64% and 69% for Area under the Curve (AUC), and Brier scores of 0.09, 0.18 and 0.16 for the failure, dysfunction and normal conditions, respectively. Former studies showed that adverse events which are taken from bedside monitors are important to intermediate outcomes and can contribute to a timely recognition of organ dysfunction and failure during ICU stay. The results obtained with those DM methods showed that it is possible to create decision support systems using biometrics and adverse events data.

The deployment of an intelligent system using data collected in real-time in a critical environment is very difficult and needs to be performed carefully. This happens due to the complexity inherent to the environment and to the difficulty in obtaining automatically some fundamental data, such as critical events, fluid balance and laboratory results.

3.3. Decision Support in Intensive Medicine

According to Álvaro Silva (Silva, 2007) Intensive Medicine (IM) is a Multidisciplinary area of Medical Sciences that specifically addresses the prevention, diagnosis and treatment of acute potentially reversible in patients with failure of one or more vital functions, and has unique characteristics and specificities. Decision making is a complex task due the environment uncertainties.

IM is a critical area where anything can fail, because the intensive care professionals are dealing with human lives in very weak conditions and in a serious life-risk situation. Intensive Care Units (ICU) are a critical environment where complex health care situations are a constant and physicians use IM to treat their patients (Bricon-Souf & Newman, 2007). The main goal of an ICU is to avoid or reverse organ failure process by adopting a timely intervention, i.e., diagnose, monitor and treat patients with serious illnesses and recover them to the quality of life they had prior to being admitted into the ICU (Suter et al., 1994).

Decisions are commonly taken based on physiological variables such as heart rate, blood pressure, temperature, ventilation and brain activity that are constantly monitored online (Mahmoud, 2003). Other variables registered hourly are also considered like administered drugs, laboratory results and fluid balance. Normally, it is very difficult to interpret this data quickly and in useful time. This situation leads to a decision making process not supported by the data collected in real-time and represents a waste of helpful information.

Actually, all data is collected automatically excepting only a few cases that require a human observation. Although there is an automatic validation, in some cases, the manual validation is required because the data can only be verified by humans (due to the semantics associated to clinical interpretation) (Filipe Portela, Manuel Filipe Santos, et al., 2011). IDSS can improve patient safety and reduce the medication-related costs, since it introduces automation at the time of ordering, a key process in health care (Kuperman et al., 2007).

3.4. Data Stream Mining

According to Gaber (Gaber, Zaslavsky, & Krishnaswamy, 2005), Data Stream Mining (DSM) is concerned with extracting knowledge structures represented in models and patterns in non-stopping streams of information.

The main problem of developing systems using DSM is the environment and the issues related with streaming and distributed data (Gaber, et al., 2005). These barriers should be overcome in order to develop a pervasive system using ubiquitous data mining. (J. Gama & Gaber, 2007; João Gama, et al., 2010). The data analysis is the most complex phase. During this phase special attention should be taken to some factors as is computing, communication, storage, and humans (J. Gama & Gaber, 2007). Managing DSM is very difficult due to data scalability (H. Wang, Fan, Yu, & Han, 2005). DSM methods are constantly changing by real-time systems that generate enormous amounts of data at unprecedented rates (H. Wang, et al., 2005).

3.5. Ensemble Approach

The use of ensembles has to do with the principle established that the sensitivity can usually be improved by using an assortment of predictive models instead of a single model (Dietterich, 2000). The ensemble-learning methodology consists in two sequential phases: the training and the testing phase (Kantardzic, 2011). In the training phase, several different predictive models are generated from the training set. In the test phase, the ensemble is executed and aggregates the outputs for each predictive model (Kantardzic, 2011). In this work the ensemble is used to compare the models and to define which one is the best for a specific target. The different types of models can contribute to a lower number of coincident failures thus increasing the ensemble performance (W. Wang, Partridge, & Etherington, 2001).

3.5.1. Ensemble methodology

In this project was followed the Stacked Generalization Methodology (SGM) (Ting & Witten, 2011; Wolpert, 1992). The learning procedure SGM is divided in four steps (Kantardzic, 2011):

1. Split the training set into two disjoint sets;
2. Train several base learners on the first set;
3. Test the base learners on the second set;
4. Using the predictions made in the stage 3 as the inputs and the correct responses as the outputs, train a higher level learner.

To split the dataset stratification technique was used. For each target a different dataset was considered with the same distribution on the target classes (0, 1).

3.5.2. Ensemble Assessment

To assess and compare the ensemble models, confusion matrix and Receiver Operating Characteristic (ROC) curve (Bradley, 1997) have been used. In this work sensitivity, accuracy and total error measures were considered. Sensitivity is the most significant measure to decide on models in critical health care. Doctors need to know with a high precision how bad the patient can be. For that reason the models to estimate the patient outcome and organ failure should be as sensible as possible. The intension is to avoid situations of no-return and no-recover of the patient condition. In order to reduce the number of false positives with a higher sensibility and a lower specificity a quality measure was introduced. This measure combines the sensibility, the accuracy and the total error and it is presented in section 7.4

3.5.3. Data Mining Techniques

This work explored different models using three techniques: Decision Trees (DT), Support Vector Machines (SVM) and Naïve Bayes (NB) (Wu, et al., 2008).

3.5.3.1. Decision Trees

Decision Trees (DT) based models produce rules in the form of IF predictive information THEN target (Concepts, 2005). In this work the CART algorithm (Breiman, Friedman, Olshen, & Stone, 1984) and *Gini coefficient* were adopted. According to Steinberg (Steinberg & Colla, 1997) CART is a decision tree learning technique that produces DT as a binary recursive partitioning procedure capable of processing continuous and nominal attributes as targets and predictors.

3.5.3.2. Support Vector Machines

Support Vector Machines (SVM) can be defined as a two-class learning task where the aim is to find the best classification function to distinguish between members of the two classes in the training data (Wu, et al., 2008). SVM records the input data into a kernel space and builds a linear model there (Concepts, 2005). SVM is based in the Vapnik-Chervonenkis algorithm (Cherkassky & Mulier, 1999). According to Milenova (Milenova, Yarmus, & Campos, 2005), Eq. 1 is a linear equation in the space of attributes $\vartheta j = k(xj, xi)$. The kernel function K, can be linear or non-linear. If K is a linear kernel, Eq. 1 reduces to a linear equation in the space of the original attributes x. If K is a non-linear kernel, Eq. 1 defines a linear equation on a new set of attributes. The number of new attributes in the kernel-induced space can be as many as the number of rows in the training data.

The output of an SVM binary classification model using (Milenova, et al., 2005) is given by :

$$f_i = b + \sum_{j=1}^{m} a_j \, y_j \, K(x_j, x_i) \qquad (1)$$

Where,

f_i (also called margin) is the distance of each point (data record) to the decision hyperplane defined by setting $f_i = 0$;

b is the intercept;

α is the Lagrangian multiplier for the j^{th} training data record $\mathbf{x}_j$;

y_j is the corresponding target value (± 1).

3.5.3.3. Naïve Bayes

Naïve Bayes (NB) algorithm is based on conditional probabilities and uses Bayes' theorem (Laplace, 1820) to derive the probability of a prediction from the underlying evidence (Concepts, 2005). Bayes' Theorem finds the probability of an event occurring given the probability of another event that has already occurred (Concepts, 2005).

Bayes' Theorem: Prob(B given A) = Prob(A and B)/Prob(A)

3.6. Pervasive HealthCare

Pervasive HealthCare (PH) aims to mitigate some of the flaws mentioned as: missing data, to many records in paper format, information available in the wrong time (delay), information available only in one place, and others. According to Varshney (Varshney, 2009) PH can be defined as "conceptual system of providing healthcare to anyone, at any time, and anywhere by removing restraints of time and location while increasing both the coverage and the quality of healthcare". This approach is based on information that is stored and available online (Mikkonen, Va¨ yrynen, Ikonen, & Heikkila, 2002).

Satyanarayanan (Satyanarayanan, 2002) characterizes the pervasive computing as an evolutionary step resulting from previous two steps: first distributed computing and then mobile computing. As main contributions we can point out the smart spaces, invisibility, localized scalability and uneven conditioning. According to Saha (Saha & Mukherjee, 2003), "Most computing systems and devices today cannot sense their environments and therefore cannot make timely, context-sensitive decisions. Pervasive computing, however, requires systems and devices that perceive context. Mobile computing addresses location- and mobility-management issues but in a reactive context—responding to discrete events. Pervasive computing is more complex because it is proactive. Intelligent environments are a prerequisite to pervasive computing. Perception, or context-awareness, is an intrinsic characteristic of intelligent environments. Implementing perception introduces significant complications: location monitoring, uncertainty modelling, real-time information processing, and merging data from multiple and possibly disagreeing sensors. The information that defines context awareness must be accurate; otherwise, it can confuse or intrude on the user experience.

Pervasive computing is about making our lives simpler through digital environments that are sensitive, adaptive, and responsive to human needs. Far more than mobile computing, this technology will fundamentally change the nature of computing, allowing most objects we encounter in daily life to be "aware," interacting with users in both the physical and virtual worlds. While research challenges remain in all areas of pervasive computing, all the basic component technologies exist today." Research in PH involves not only issues associated with the communication/interaction between the environments and the users but also the mobile support of the users.

INTCare is a PH system because using streaming data can automatically process the values in the purpose of a target and make the information (knowledge reached) available anywhere and anytime. At the same time the system can reduces the human efforts, increasing the time to patient care. INTCare provides important knowledge to the decision making process and also is considered a pervasive IDSS because attends the main pervasive features (Saha & Mukherjee, 2003):

Scalability - Pervasive computing environments face a proliferation of users, applications, networked devices, and their interactions on a scale never experienced before. As environmental smart grows, so the number of devices connected to the ICU environment and the intensity of human-machine interactions are higher;

Heterogeneity - Conversion from one domain to another is integral to computing and communication. Assuming that uniform and compatible implementations of smart environments are not achievable, pervasive computing must and ways to mask this heterogeneity—or uneven conditioning as it has been called— from users. This system processes the data from several data sources and presents the information to heterogeneous users and environments;

Integration - Though pervasive computing components are already deployed in many environments, integrating them into a single platform can be a research problem. INTCare surpassed this problem and integrates a set of platforms and patient data in a single platform called Electronic Nursing Record;

Invisibility - A system that requires minimal human intervention offers a reasonable approximation of invisibility. INTCare is an example of invisibility because many of the operational tasks are provided by intelligent agents;

Context awareness (CA) - Once a pervasive computing system can perceive the current context it must have the means of using its perceptions effectively. Richer interactions with users will require a deeper understanding of the physical space. Smartness involves accurate sensing (input) followed by intelligent control or action (output) between two worlds, namely, machine and human. The CA are very important to the success of INTCare, being the system adaptable, it can easily adapt to environment and users changes or change to other context.

4. Streaming Data Sources

Figure IV.2 shows the data acquisition architecture (Filipe Portela et al., 2011; Portela, et al., 2012). Data collected by the agents are stored into tables and then they are processed and sent to the database server. This data are now available to be used by INTCare. The sources and load frequency of each one of the variables in use are presented in table IV.1. They are provided by five data systems:

a) Laboratory (LR);
b) Pharmacy System (PHS);
c) Electronic Nursing Records (ENR);
d) Electronic Health Records (EHR);
e) Vital Signs Monitor (VS).

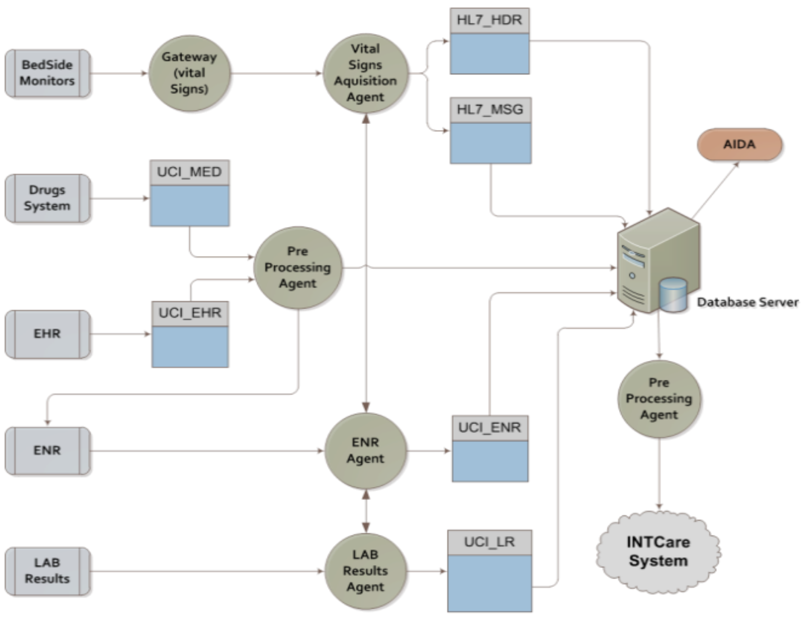


Figure IV.2 – Data acquisition architecture.

The most part of the data (vital signs, laboratory results and therapeutics) are collected in real-time by a streaming process. At the same time they are stored into the database. During this process a set of agent procedures are executed in order to construct a data warehouse and prepare data for data mining tasks.

This operation is executed in a real-time streaming cycle due to the continuity and sequence of the tasks. As presented in figure IV.1, INTCare system comprises a set of cycles operating in real-time (e.g. the cycles between data warehouse and knowledge base). These tasks process data streams with hard real-time constraints on heterogeneous systems (Verner, Schuster, & Silberstein, 2011). The system collects real-time data streams from a variety of external sources into the data warehouse.

The objective of those cycle is line with the Real-Time Stream Warehouse (Golab, Johnson, & Shkapenyuk, 2009): to keep all the tables and views up-to-date as new data arrive over time.

Table IV.1 presents the frequency and the source of each variable extracted from the streamed data into INTCare System.

Table IV.1 – Frequency and data source of the variables considered

Input		*Variable*	*Source*	*Frequency*
Case Mix		Age	EHR	Once
		Admission Type	EHR	Once
		Admission From	EHR	Once
SOFA	Cardiovascular	Blood Pressure	VS	Minute
		Dopamine, dobutamine, noradrenaline, adrenaline	PHS	All day
	Respiratory	PaO_2/FiO_2	LR	All day
	Renal	Creatinine	LR	All day
	Liver	Bilirubin	LR	All day
	Coagulation	Blood platelets	LR	All day
ACE		Hear Rate	VS / ENR	Minute
		Blood Pressure	VS / ENR	Minute
		Diuresis	ENR	Hour
		SPO2	VS / ENR	Minute
Ratios		All ratios	VS / ENR	All day
Outcome		Outcome	HER	All day

4.1. Data Selection

One of the organic systems (neurologic) was disregarded since Glasgow Coma Score is rarely registered. The variables in use only are not registered in different sources, but also they have different frequencies as shown in table IV.1. The base of the data selection for the DSM process are the following: the variables discovered in the previous work (Case Mix and Critical Events), SOFA (0 or 1), and the ratios introduced now. Data is presented in database (DBT) and is used to obtain DM attributes:

a) Case MIX;
b) Sofa;
c) Accumulated Critical Events (ACE);
d) Ratios;
e) Outcome;

The next lines will explain these attributes and their natural domains.

4.1.1. Case mix

Case Mix (CM) is composed by a set of variables presented in the patient Electronic Health Record. These variables (age, admission type and admission from) are obtained at patient admission and are automatically transformed according the DM attributes.

4.1.2. Sequential Organ Failure Assessment

Sequential Organ Failure Assessment (SOFA) is commonly used in ICU to score the degree of dysfunction/failure of six organic systems – Cardiovascular, Respiratory, Renal, Liver, Coagulation and Neurological (Vincent et al., 1996). SOFA is set in a scale from 0 (normality) to 4 (failure) for each organic system and uses the worst values occurred along the day. In opposition to the original SOFA, where the result is only calculated at the end of the day, INTCare is in a constant scoring computation throughout the day (M. F. S. Filipe Portela, José Machado, Álvaro Silva, Fernando Rua, António Abelha, 2012). SOFA score is now collected and computed in a continuous mode and uses the worst values occurred per hour. For the DM models it is used a binary variable for each organ, where 0 describes normality and 1 describes dysfunction/ failure.

4.1.3. Accumulated Critical events

Critical Events (CE) were developed due to the recognize incapability of SOFA to predict precisely the patient outcome (J. L. Vincent et al., 1998). CE were defined by a panel of experts (Á Silva, P.

Cortez, M. F. Santos, L. Gomes, & J. Neves, 2008) and they are related to four physiological variables – Blood Pressure (BP), Heart Rate (HR), Urine Output (UR) and Oxygen Saturation (SPO2). CE values are obtained by analysing the duration of values out of the normal range to a specific variable. They are calculated hourly and, subsequently, a new variable is derived: Accumulated Critical Events (ACE).

Accumulated Critical Events (ACE) reflects the patients' clinical evolution/severity of illness. These variables calculate the accumulated value by hour. In the DM models, ACE is the number of critical events verified in a patient during their admission. For example, the ACE of SPo2 represents the sum of values by hour of Oxygen Saturation. If two critical events occur in first hour and other two in the second hour the ACE of the second hour is set to four.

4.1.4. Ratios

Ratios were added in consequence of the introduction of ACE. These ratios give the possibility to determinate the number of ACE by hour (R1) and a correspondence between the number of ACE and the maximum number of events verified in the past, grouped by category and patient (R2).

4.1.5. Outcome

Outcome is a binary variable and it is used to identify patient hospital discharge results (alive or died). In order to prepare these attributes to the DSM models, it is necessary to proceed automatically some pre-processing and transformation phases.

5. Data Pre-Processing

This phase is responsible for solving some data acquisition problems (P. G. Filipe Portela, Manuel Filipe Santos, Álvaro Silva, Fernando Rua, 2012a). All tasks are performed by pre-processing agent. It is also responsible to clear the data set, validating the data received, deleting null values received and making a correct patient identification (Filipe Portela, Pedro Gago, et al., 2011). This process is applied to the values received from three data sources: bedside monitors, electronic nursing records and laboratory.

This agent is executed whenever a new value is collected:

a) Validation of all data collected
    i. Validate / classify the values according to the range defined in the ICU (table 2) (F. Portela et al., 2011);

ii. Validate the patient identification (PID) according to the EHR;

b) Preparation of the data mining input table (DMIT) – creation of the model input structure for each patient;

c) Validation of DSM input data – clean null and wrong values to ensure that the dataset is real.

d) Process the variables with all-day frequency.

Table IV.2 presents the range of the values that normally can be accepted in ICU (ICUMIN and ICU MAX). This table also presents a second range, which contains some abnormal values that can be verified ('*Anytime Min*' and '*Anytime Max*') and sometimes can be considered valid. In the automatic pre-processing phase '*Anytime Min*' and '*Anytime Max*' are used. These limits were defined in accordance to the common results verified in ICU.

Table IV.2 – ICU Vital Signs Range

Vital Sign	*Unit*	*ICU Min*	*ICU Max*	*Anytime Min*	*Anytime Max*
Blood Pressure (BP)	bpm	50	180	0	300
SPO2	%	80	100	40	100
Temperature (Temp)	°C	34	42	30	45
Respiratory Rate (RR)	%	0	40	0	40
Heart Rate (HR)	*mmHg*	40	160	0	250

6. Data Transforming

In order to develop the predictive models some of the input attributes need to be transformed. The values of the variables are transformed into discrete and normalized values. All transforming tasks are performed automatically and in real-time by the INTCare intelligent agents. These agents enclose a set of predefined procedures, functions and triggers. All of them are adaptable and can be automatically modified after some environment change. The data are prepared and inserted into DMIT. In this phase the variables are transformed according to the attributes:

1) CM Variables;
2) SOFA variables;
3) Accumulated Critical Events (ACE);
4) Ratios.

All transforming tasks are based in the values defined in the table 3. Table 3 presents for each one the variables the list of values to be considered according to a range of values (*Min*, *Max*). For the real numbers (values $\in$ {| R_{0+}}), some ranges were introduced according to the normality and importance / significance of the value to ICU. For the ACE, R1 and R2 the transformed value is calculated according to limits (*Min* and *Max*) defined in the table 6.

6.1. Case Mix and SOFA

The first transformation process is a simple analytic task, where the values collected are transformed according to some rules (if then else). This process is applied to the variables presented in table IV.3. Regarding to the age parameter the patient age is classified by a level from 1 to 4. For Case Mix attributes only one the values is used for each case. For instance, the admission type only can be urgent (u) or programmed (p). As another example, if a patient is 35 years old and has an urgent admission from surgical, the transformed values are 1, u, 1 respectively for each ID column.

In the case of SOFA the values are collected in real-time and in a continuous way, being the data automatically streamed into the database. Data mining models are based on the worst SOFA value received by each hour / variable. The value collected is transformed, according to DM attribute, to 0 (normal) and to 1 (failure) (M. F. S. Filipe Portela, José Machado, Álvaro Silva, Fernando Rua, António Abelha, 2012). Finally, for outcome attribute the default value is 0, except when patient dies all the values of outcome attribute are updated to 1.

Table IV.3 – Variables transforming values dies

	BP(mmHg)	*SpO2 (%)*	*HR(bpm)*	*UR(ml/h)*
Normal range	90 - 180	>= 90	60 - 120	>= 30
Critical event a	>= 1h	>= 1h	>= 1h	>= 2h
Critical event b	< 60	<80	<30 V> 180	<= 10

6.2. Critical Events and Ratios

The second phase of transformation uses CE. Firstly the value is categorized as normal (0) or critical (1) according to the Table IV.4. If it is within the range, the value is classified as normal; otherwise, it is classified as critical. Then the typology of the event is determined: critical event type a, or b. The typology (a) corresponds to the extent of an event (out of the normal range) in order to be considered critical. The typology (b) refers to values that are critical independently the duration of the event.

Table IV.4 – The protocol for the out of range physiologic measurements (adapted from Álvaro (Álvaro Silva, et al., 2008))

ID		*Variable*	*Min*	*Max*	Value
Age		Age	18	46	1
			47	65	2
			66	75	3
			76	130	4
Admission Type		Urgent	-	-	u
		Programmed	-	-	p
Admission From		Chirurgic	-	-	1
		Observation	-	-	2
		Emergency	-	-	3
		Nursing Room	-	-	4
		Other ICU	-	-	5
		Other Hospital	-	-	6
		Other Situation	-	-	7
SOFA	Cardiovascular	BP (mean)	0	70	1
		Dopamine	0,01	-	1
		Dobutamine	0,01	-	1
		Epi / Norepi	0,01	-	1
	Renal	Creatinine	1.2	-	1
	Respiratory	Po2/Fio2	0	400	1
	Hepatic	Bilirubin	1.2	-	1
	Coagulation	Platelets	0	150	1
ACE			0	+∞	SET
R1			0	1	SET
R2			0	1	SET
Outcome		Live	-	-	0
		Dead	-	-	1

a Defined when continuously out of range.

b Defined anytime.

Data mining agent is responsible to prepare DMIT. During its execution, some tasks are performed (P. G. Filipe Portela, Manuel Filipe Santos, Álvaro Silva, Fernando Rua, 2012b):

a) To identify value importance;
b) To identify critical event typology;
c) To calculate ACE;
d) To calculate ratios;
e) To delete null and wrong values.

In previous experiments it was used the real values obtained without transformation in DM process. Further experiments considered the *Bin TopN* approach, where N it is the number of bins (Concepts, 2005). In this case the DM attributes contain only 8 different values. The 7 most frequently values have bins of their own. The remaining values are thrown into a bin called "Other". In more recent experiments a discretization technique has been considered. The values were grouped and categorized in accordance to an interval (Min and Max). Using this technique the sets were defined according to the respective average (R1) or higher value (R2) of the data collected. The categories defined represent the significance / importance of each value to the patient in the ICU. These ranges are flexible and are updated automatically according to the values collected in the ICU. This approach adjusts the sets as defined in the table 5 making them adaptive and regulated by the patient results and the streamed data.

The ranges were created using a 7-point-scale adapted by Clinical Global Impression - Severity scale (CGI-S) (Guy, 1976). The goal of CGI-S is to allow the clinician to rate the severity of illness (W. Guy Modified From: Rush J, 2000).

Table IV.5 presents the rules to create the ranges. In the case of R1 (elapsed time) the average of values collected are used. In the case of R2 (maximum number of ACE) a percentage of the maximum value obtained in the range is used. This value is updated when a superior value appears. When this happens, the maximum value is swapped with the new value. The criterion used to define the percentages concentrates the most part of patient values within a scale between 0 and 5. More severe cases are assigned to the levels 6 and 7. The sets to be used by the DM models are defined according the CGI-S severity level.

Table IV.5 – Discretization rules

SET	*R1*		*R2*		*Definition*
	Average		Maximum		
	>	<=	>	<=	
0	-	0%	-	0%	Inexistence
1	0%	25%	0%	10%	Normal condition
2	25%	50%	10%	25%	Borderline condition
3	50%	100%	25%	50%	Mild condition
4	100%	150%	50%	75%	Moderate condition
5	150%	200%	75%	90%	Marked condition
6	200%	300%	90%	100%	Severe condition
7	300%	1000%	100%	200%	Extreme condition

Using Table IV.5 some ranges were introduced according to the importance /significance of the value for the ICU.

Table IV.6 presents the rules defined to discretize each continuous value (values ϵ{|R0+}). The different sets are identified at the top of the table. The left column identifies the variable. In the centre of the table are defined the ranges for each set. R2Min and R2Max are used by R2. R2 uses the maximum number of ACE occurred in ICU for each variable until the current moment. R2 is categorized according to the percentage of the values collected, i.e., at level 2 corresponds to 15% of the cases (values collected between 10% and 25% of the Maximum).

For example, if the ACE maximum of O2 is 4 they are categorized at the fourth level. The same is verified to patients who present values of R2 between 0.1 and 0.25. All attributes of R1 have the same limits. The attributes of R1 the set is determined by the rows (R1 BP Min to R1 TOT Max). In the case of urine output was difficult to define the range due to the existence of many null values. For all attributes of R2, the ranges of the set are equal, i.e. the first level (1) is between 0.00 (0%) and 0.10 (10%) for all R2 cases (BP, O2, HR and Total).

Finally, all ACE attributes are grouped in accordance to their importance and the number of occurrences. These values were defined by ICU experts, but can be modified in the future.

Table IV.6 – Discretization sets of Data Mining Input

SET		*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*
R1 BP	Min	-0,1	0,000	0,010	0,021	0,041	0,062	0,082	0,123
	Max	0,000	0,010	0,021	0,041	0,062	0,082	0,123	2,000
R1 O2	Min	-0,1	0,000	0,018	0,036	0,072	0,108	0,144	0,216
	Max	0,000	0,018	0,036	0,072	0,108	0,144	0,216	2,000
R1 HR	Min	-0,1	0,000	0,004	0,008	0,015	0,023	0,030	0,045
	Max	0,000	0,004	0,008	0,015	0,023	0,030	0,045	2,000
R1 TOT	Min	-0,1	0,000	0,020	0,041	0,081	0,122	0,162	0,243
	Max	0,000	0,020	0,041	0,081	0,122	0,162	0,243	2,000
R2	Min	-0,1	0,000	0,100	0,250	0,500	0,750	0,900	1,000
	Max	0	0,100	0,250	0,500	0,750	0,900	1,000	2,000
ACE	Min	-0,1	0	3	5	8	10	12	15
	Max	0	3	5	8	10	12	15	50

At level of Data Mining (after transformation tasks being accomplished), the values are stored into the data warehouse. Then, all data are validated and prepared to be used by the DM models and inserted into DM input table (DMIT).

7. Data Mining

In this phase, data mining models are induced using the data processed and transformed in preceding phases. The DM engine uses DMIT content to train and test the models. To compute the probabilities of the targets, the data presented in DMPT is used.

INTCare uses the models presenting the better results for each target. The results obtained are then compared and evaluated according to the quality measure. This task is performed by the ensemble. The results obtained determinate whether it is or not necessary to induce new models.

This process is done by the performance agent. Finally, after the selection of the best model for each target it is necessary to calculate the probabilities associated to organ failure and to death, i.e., the results associated to the class 1 (target = 1). The objective is to predict if the organ failure will occur or if the patient will die in the next 24 hours. Because of this, the best model needs to be extremely sensible to the class 1. The prediction agent is used to obtain the probabilities defined above. The DM tasks are performed by the following agents:

Data Mining (a_{dm}) agent is the most important of the Knowledge Management subsystem. It is responsible to:

a) perform all transforming tasks;
b) create and test the models in real-time using online data stored in the data warehouse;
c) store the induced models into the Knowledge Base;
d) ensure that all data essential to the application of DM algorithms are currently and are properly validated;
e) answer Performance Agent (*apf*) requests.

Ensemble (a_{ens}) agent enhances predictive performance by combining several models in order to produce models with better results, i.e., with higher sensitivity and accuracy (Dietterich, 2000). This agent maintains an ensemble grouping the models induced by target. Three DM techniques are considered for this purpose: DT, SVM and NB.

Performance (a_{pf}) agent is responsible to ensure the quality and performance of the models. It is constantly looking for new data in order to compute assessment parameters (Gago, et al., 2006). It analyzes the new data available and verifies the performance of the predictive models. The metrics for evaluating the classification models include accuracy, sensitivity, specificity and total error. According to some metrics defined (subsection 7.4) it activates the Data Mining agent. If the metric result of a model decreases to a level lower than a predefined threshold, a new model is required to substitute the first one (Gago, et al., 2006). This agent is in a continuous interaction with the ensemble and data mining agents.

Prediction (*apd*) agent answers user questions by applying the adequate models stored in the Knowledge Base. This agent is also responsible to determine in an hourly base the probability (%) associated to the organ failure and to the death.

Data Retrieval (*a_{dr}*) agent retrieves, from the data warehouse, the information requested by the interface agent (Gago, et al., 2006).

Model Initialization (*a_{mi}*) agent is used to initiate the models presented in Knowledge base. This agent is executed whenever some new knowledge is required by the *data mining agent*. The models results also are sent to the *prediction agent*.

7.1. Ensemble

The ensemble is organized in terms of six independent components. Each one of these components is dedicated to a different target. Seven different scenarios are considered (*s*) and three distinct DM techniques (*z*): Decision Trees (DT), Support Vector Machine (SVM) and Naïve Bayes (NB).

The ensemble can be defined as a three-dimensional matrix *M* composed by *s=7* scenarios (*s1* to *s7*) x *t=6* targets (t1 to t7) x z=3 techniques (z1 to z3). Each element of *M* corresponds to a particular model and can be defined as:

$$M_{s,t,z} = \begin{Bmatrix} s = 1 \ldots 7 \\ t = 1 \ldots 6 \\ z = 1 \ldots 3 \end{Bmatrix}$$

Where,

s:
1 = {CASE MIX}
2 = {CASE MIX, ACE, R}
3 = {CASE MIX, ACE, R1}
4 = {CASE MIX, ACE, SOFA}
5 = {CASE MIX, ACE, SOFA, R}
6 = {CASE MIX, ACE, SOFA, R2}
7 = {CASE MIX, ACE, SOFA, R1}

t:
1 = Respiratory
2 = Cardiovascular
3 = Coagulation
4 = Renal
5 = Hepatic
6 = Outcome

z:
1 = Support Vector Machine
2 = Decision Trees
3 = Naïve Bayes

Each model is induced automatically and in real-time using streamed data. The data mining engine uses the data present in DMIT. This table contains tuples of the type:

$DMIT = <$ *pid, date, hour*, v_{ace_bp}, v_{aceBP_time}, v_{aceBP_max}, v_{ace_hr}, v_{acehr_time}, v_{acehr_max}, v_{ace_spo2}, $v_{acespo2\_time}$, $v_{acespo2\_max}$, v_{ace_ur}, v_{aceur_time}, v_{aceur_max}, v_{total_ace}, $v_{totalace\_time}$, $v_{totalace\_max}$, v_{age}, v_{adminF}, v_{adminT}, v_{sofa_cardio}, v_{sofa_resp}, v_{sofa_renal}, v_{sofa_coag}, v_{sofa_hepa}, v_{sofa_neuro}, $v_{outcome}$ $>$

Where,

pid is the patient identification;

date is the date of the values;

hour is the number of hours elapsed since the patient admission;

v_{ace_bp} ... $v_{outcome}$ are the values obtained for each patient and date.

The next representation illustrates an instantiation for the model M6 which corresponds to SVM and *outcome* target:

$DMIT_{svmOutcomePidDate} = <$ *outcome, pid, date, hour*, v_{age}, v_{adminF}, v_{adminT}, v_{sofa_cardio}, v_{sofa_resp}, v_{sofa_renal}, v_{sofa_coag}, v_{sofa_hepa}, v_{sofa_neuro}, v_{aceBP_max}, v_{acehr_max}, $v_{acespo2\_max}$, v_{aceur_max}, v_{total_ace}, $v_{totalace\_time}$, $v_{totalace\_max}$ $v_{outcome}$ $>$

To obtain a DM model the data stored in the input table are loaded. Real values (ACE and ratios) are categorized using the sorting method. The other values are maintained as they are. *DMIT* contains 32 columns (ACE, Ratios, SOFA, Case Mix and Outcome) and they are used to induce seven different configurations of attributes for each target (renal, hepatic, coagulation, cardiovascular, respiratory and outcome) and for each DM technique (DT, NB, and SVM). The models are induced in real-time using online-learning by the DM agent. This agent runs whenever a request is sent or when the performance of the models decreases. Figure IV.3 presents an overview on how the ensemble works and how the predictive models are created:

- **Predictive Models** – 126 models are induced combining seven scenarios (S1 to S7), six targets and three different techniques (SVM, DT and NB);
- **Ensemble** – the models are assessed in terms of the sensitivity, accuracy, total error and specificity. The best model for each target (t) is then selected

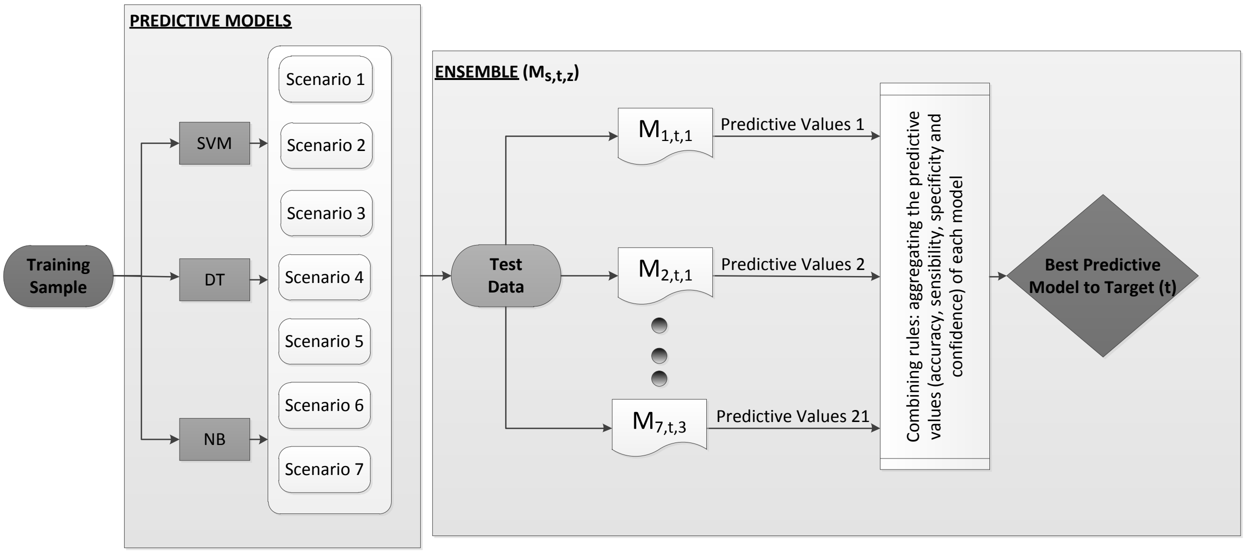


Figure IV.3 – Ensemble main cycle and predictive models induction (target T).

In order to select the best predictive model for each target, a set of tasks are performed:

1. Create the confusion matrix for each scenario;
2. Obtain the assessment measures;
3. Apply the quality measure;
4. Determine the confidence rate for each prediction.

The above tasks will be explained in more detail in the following lines.

7.2. Confusion Matrix

In the first step, a Confusion Matrix (CMX) is automatically created for each model $M_{s,t,z}$ using the data stored in DMIT. In this way the following measures are obtained: the number of True Positive (*TP*), False Positive (*FP*), True Negatives (*TN*) and False Negatives (*FN*) presented in the dataset. The CMX is organized in terms of the expected values (*REAL*) and the predicted values (*PREDICTION*). Table IV.7 presents a typical confusion matrix, where 0 and 1 are the classes. CMX contains the number of occurrences of TN, FN, FP and TP obtained by a model during a test. The values obtained by the confusion matrix are stored in the database. Each Model has a different matrix table (DMMT)

Table IV.7 – Confusion Matrix

		PREDICTION	
		0	1
REAL	0	TN	FP
	1	FN	TP

7.3. Assessment measures

In this task a set of complementary measures are calculated based on the values of *TP, FP, FN* and *FP*: sensitivity, specificity, accuracy, negative values (*NPV*), positive values (*PPV*), positive error (*PError*), negative error (*Nerror*) and total error (Terror). For a model M the measures are calculated using the following expressions:

$$Sensitivity(M) = \frac{TP(M)}{TP(M) + FN(M)}$$

$$Specificity(M) = \frac{TN(M)}{TN(M) + FP(M)}$$

$$Accuracy(M) = \frac{TP(M) + TN(M)}{TP(M) + TN(M) + FP(M) + FN(M)}$$

$$PPV(M) = \frac{TP(M)}{TP(M) + FP(M)}$$

$$NPV(M) = \frac{TN(M)}{TN(M) + FN(M)}$$

$$NERROR(M) = \frac{FP(M)}{FP(M) + TN(M)}$$

$$PERROR(M) = \frac{FN(M)}{TP(M) + FN(M)}$$

$$TERROR(M) = \frac{FP(M) + FN(M)}{TP(M) + TN(M) + FP\ (M) + FN(M)}$$

The measures are then stored in the table DMMT:

$DMMT = < cm_{tp}, cm_{tn}, cm_{fp}, cm_{fn}, m_{sensitivity}, m_{specificity}, m_{accuracy}, m_{ppv}, m_{npv}, m_{perror}, m_{nerror}, m_{terror} >$

Where,

cm_{tp} ... cm_{fn} represent the number of true positives, true negative, false positive and false negatives obtained by the confusion matrix for each model;

$m_{sensitivity}$ ... m_{terror} are the results obtained by the model for each measure.

7.4. Apply the quality measure

The main purpose of the ensemble is to select the most suited model from a group of candidates. In order to evaluate the models, a quality measure was defined. This measure is based in the results obtained by the models in terms of *sensitivity*, *accuracy* and *total error*. The selected models are used by the pervasive system only if they satisfy the following conditions:

i. *Total Error* <= 40%
ii. *Sensitivity* >= 85%
iii. *Accuracy* >= 60%

These thresholds have been defined in order to assure a minimum level of quality in models. The measure was defined in accordance to ICU doctors and can be adjustable when necessary. The ensemble executes a set of models for each target (t1... t6) and separates the models into two sets: one group ($BESTMODEL_{t,k}$) includes the models that achieve the thresholds (*i, ii and iii*); and another set containing the remaining models:

$BESTMODEL_{t,k} = < ModelID_{stz}, Accuracy_{stz}, Sensitivity_{stz}, Terror_{stz}, Specificity_{stz} >$

$OTHERMODEL_{t,k} = < ModelID_{stz}, Accuracy_{stz}, Sensitivity_{stz}, Terror_{stz}, Specificity_{stz} >$

Where,

ModelID is the identification of the $model_{stz}$;

t is the target;

s is the scenario;

z is the technique;

k is the index of the array.

To unravel models with the same results the *sensitivity, accuracy, TError and specificity* should be considered in that order.

7.5. Determine the confidence rate

The confidence rate is the probability of the target result to be equal to 1 (TP). For each model the scenarios evaluation agent sends a message to the prediction agent with the identification of the best model and the indication to run the model for each target. The confidence rate is determined for every prediction. The results obtained for each target are stored into an array containing the models identification and respective results. For example, if the target is the *cardiovascular* system an entry will be created in the set *BESTMODELt,k* containing the models which achieved the thresholds.

At the first level (*k=0*) is identified the name of the model which presents the better result to the target t. At the end, the prediction probability for each target / hour is disseminated through the INTCare system. If, for a given target, none of the models matches the quality measures no prediction is presented.

8. Experimental work

A number of experiments has been carried out in order to test the ensemble performance. Online and real-time data collected from the ICU has been considered. The data used to generate the DM models were gathered in the ICU of CHP - HSA during the period between February and June 2012 and covering the first five days of stay of 129 patients.

8.1. Experimental setting

Three DM techniques were considered: Decision Trees (DT), Support Vector Machine (SVM) and Naive Bayes (NB). Table IV.8 presents the configurations considered for each one of the techniques. For each parameter is indicated whether the used value is a default value (DEFAULT) or an user-specified value (INPUT).

Table IV.8 – Configuration of DM Techniques

Technique	*Setting Name*	*Setting Value*	*Setting Type*
DT	Minrec Node	10	nput
	Max Depth	7	nput
	Minpct Split	.1	nput
	Impurity Metric	Gini	Input
	Minrec Split	20	nput
	Minpct Node	.05	nput
	Prep Auto	On	nput
NB	Pairwise Threshold	0	nput
	Singleton Threshold	0	nput
SVM	Conv tolerance	.001	nput
	Active learning	Enable	nput
	Kernel function	Linear	Default
	Complexity factor	0.142831	Default
	Prep auto	On	nput

8.2. Training and Testing Datasets

In order to evaluate the 126 models induced, an automatic and real-time test phase was performed using the data streamed and transformed. The original dataset (DMIT) was divided into two data sets using the holdout sampling method. 70% of the data were considered for training and 30% for testing (stratified by the target).

Cross Validation (CV) method is considered one of the most accurate methods to sample the data and evaluate the models. Nevertheless, this option wasn't available in the version of ODM used in the study Because of this reason the CV method can't be used, however a new approach was explored using the ensemble. Ten runs (10 folds) were executed with different data sets. The data related to a patient are maintained as an indivisible piece. For each target a separate dataset has been considered. The data corresponds to:

- Period in analysis: 105 days;
- Number of patients: 129;
- Time Frame Considered: the number of inpatient days from 1 to *n* (n<=5 only first five days are considered in the maximum).

Invalid records were not considered based on the following criteria:

- Exclusion criterion I: Patient with intermittent data, i.e., the collecting system failed at least one hour in a continuous way;
- Exclusion criterion II: Existence of null values.

Figure IV.4 presents the distribution of the classes (in percentage) for each one of the targets. For example, in the dataset related to the outcome, 28.98 % of the rows have an outcome equal to 1. This doesn't means that 28,98% of the patients died, because the values are related to the rows (not to patients). In this experiment, only two targets, the cardiovascular (62,70%) and the respiratory (67,48%) present a high level of organ failure (result = 1). For all the other targets, the number of positive cases (final result = 1) is substantially lower and less than the cases of without organ failure.

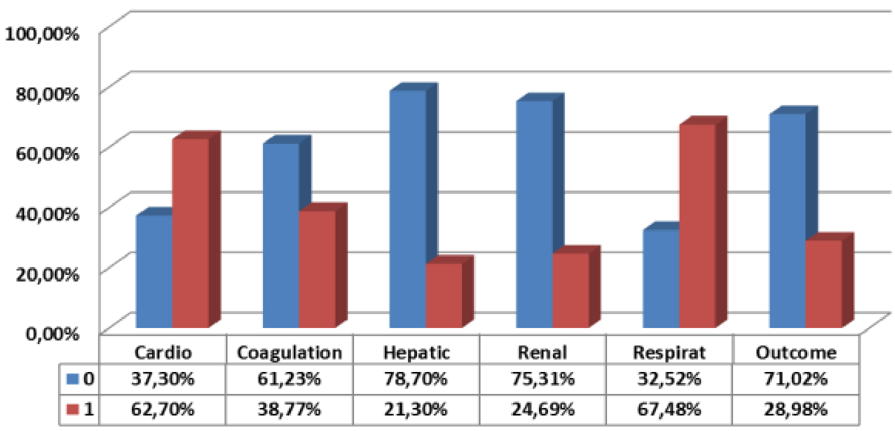


Figure IV.4 – Distribution of the values for the target (%).

8.3. Results

To evaluate the ensemble three measures were considered: *Sensitivity, Accuracy,* and *Terror*. For each measure the average and the standard deviation of 10 runs was taken. Table A1 (appendix) presents the top 3 results in terms of sensitivity obtained for each target and run (R1 to R10). The best values for each run are identified in bold and the others are underlined. From this table is possible to emphasize some interesting results. For cardiovascular, coagulation and renal systems the best model along the 10 executions are always the same, respectively M6,2,2, M4,3,3 and M4,3,4 . This situation reduces the ensemble to a singular model. For the targets of outcome (M1,6,1; M4,6,1) and hepatic (M4,5,3; M5,5,1; M6,5,3) two / three models compete to be elected in the ensemble. In general, the competition for the top 3 is animated.

Taking as an example the respiratory system, along the 10 runs, it can be observed that 9 of the 21 models were considered as one of the three best models. The use of ensemble helps to choose the best model in the cases where more than one model present good results like is the outcome, the hepatic and the respiratory system. In the table A1 the lines identified with the title *ensemble* present the best result attained in terms of the sensitivity for each run and for each target. Table IV.9 presents the performance of the ensemble for each target. The values correspond to the average of the measures obtained during ten runs. For each one the values are associated the standard deviation. Respiratory, hepatic and renal systems don't meet the quality measure established and have been excluded.

Table IV.9 – Ensemble Results

Target	Comply the quality measure	Sensitivity	Accuracy	Specificity	Terror
Cardiovascular	YES	**97,95** ± 0,31	**76,81** ± 2,35	**41,81** ± 5,75	**23,19** ± 2,35
Coagulation	YES	**91,20** ± 3,57	**65,69** ± 3,83	**49,61** ± 6,15	**34,31** ± 3,84
Hepatic	NO	69,24 ± 9,41	**82,89** ± 2,57	**87,34** ± 3,22	**17,10** ± 2,57
Outcome	YES	**99,77** ± 0,33	**63,58** ± 3,11	49,58 ± 4,90	**36,42** ± 3,11
Renal	NO	77,17 ± 12,41	43,08 ± 4,66	43,08 ± 4,66	49,09 ± 5,39
Respiratory	NO	67,11 ± 5,67	**63,86** ± 4,27	**60,39** ± 6,75	**36,14** ± 4,27

The graphs of the figures IV.5 to IV.10 present the Receiving Operator Curve (ROC) (Bradley, 1997; Zweig & Campbell, 1993) for each one of the targets . These graphs show two ROC curves one for the best model and another with the ensemble results. For example, in Figure IV.9 the ROC curve of the best cardiovascular model (M6,2,2) are is to the ensemble curve. Figure IV.11 presents the ROC for the ensemble.

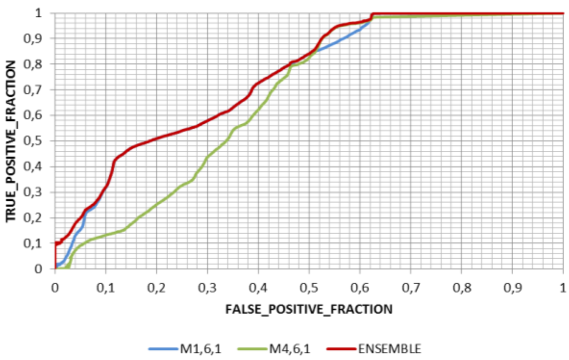


Figure IV.5– Outcome Target.

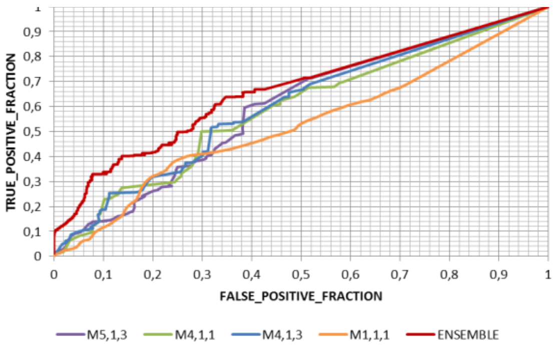


Figure IV.6–Respiratory Target.

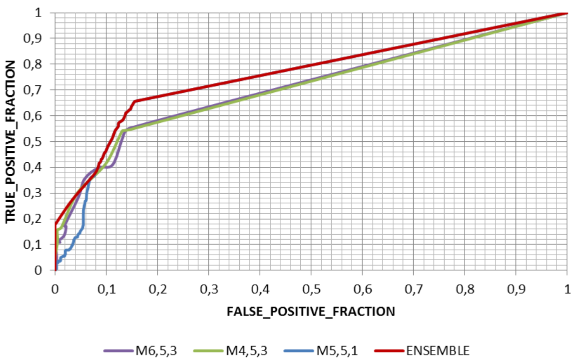


Figure IV.7 –Hepatic Target.

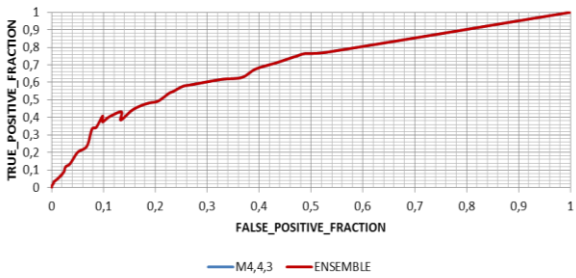


Figure IV.8 –Renal Target.

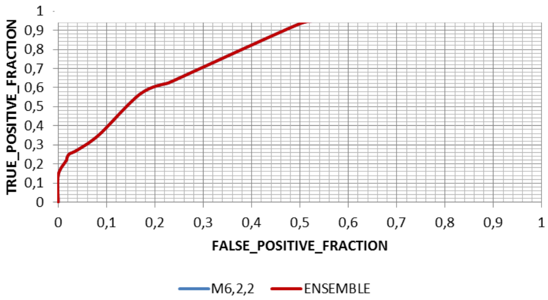


Figure IV.9 – Cardiovascular target.

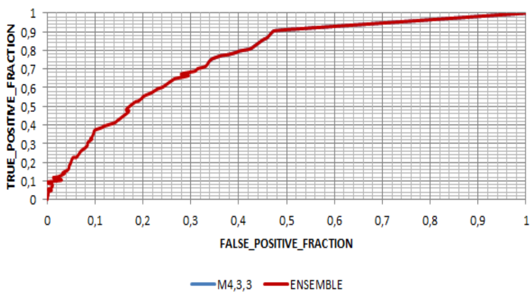


Figure IV.10 – Coagulation Target.

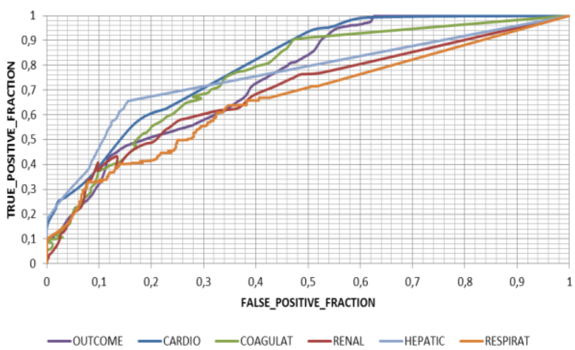


Figure IV.11–ROC for all targets.

9. Discussion

To induce data mining models automatically, autonomously and in real-time, using online- learning in critical environments as is Intensive Medicine, the ensemble approach has been explored.

From the six targets considered in this study, three satisfy completely the quality measures defined in the table IV.9: outcome, cardiovascular and coagulation. For the outcome the results obtained are very good (the sensitivity is equal to 99.77% ± 0.33). For some runs the sensitivity attained 100%. For the cardiovascular system the results obtained are slightly lower (97.95 ± 0.31). The cardiovascular target gathered better results for sensitivity (98.25%). Although presenting interesting results for various runs, the renal system denoted an instable behavior. For instance, for the 2nd, 3rd and 10th round, the values obtained for sensitivity have been: 91.15; 89.93; 95.30. Hepatic organic system was also discarded. This happened because the prediction models associated to this system are more specific than sensible. These models are better predicting the class 0 (specificity = 87.34 ± 3.22). At the same time the predictions for the hepatic system are the most accurate (accuracy = 82.89 ± 2.57 and total error = 17.10 ± 2.57). However, the medical requirements dictated the exclusion of this system. The hepatic ROC curve in the graph of the figure IV.11 presents the better specificity.

The results are in line with the medical knowledge (Kaplow & Hardin, 2007; J. Vincent et al., 1998; Vincent, et al., 1996). Renal and hepatic systems are directly or indirectly associated to the urine output attribute. The occurrence of numerous dubious values compromises the results obtained.

The respiratory system obtained fair results simultaneously for sensitivity, accuracy and specificity (around 60%). This happens because the variables available and regularly monitored in the ICU are not enough to categorically understand if this system failing. A deeper analysis demonstrated the necessity of considering further variables, e.g. from the ventilator systems (Eslami, de Keizer, Abu-Hanna, de Jonge, & Schultz, 2009; Hoo, 2009; Tehrani, 2008a, 2008b)

The ROC graphs for coagulation, cardiovascular and renal targets show that the ensemble curve is equal to the curves of the better models. In the other cases (outcome, hepatic and respiratory) are presented the models which obtained the better result during at least one of the runs (consult table A1), being the ensemble a combination of the better models, i.e., result of ensemble is always better than a single model.

For example, in the case of the hepatic system three models present the better result for sensitivity along the 10 runs: M6,5,3 (1 time); M4,5,3 (1 time) and M5,5,1 (8 times). In figure 10 is presented the curve for these three models and the curve for the ensemble. In figures 7, 8 and 10 the ensemble (red curve) presented the better results for sensitivity, i.e., the curve is closer to 1 in the true positive axis. Figure IV.11 presents the best results of the ensemble for the six independent components. In a global observation the ensemble is very efficient predicting the class 1 for outcome, cardiovascular and coagulation.

The low level of accuracy verified is due to the dynamic characteristics of the ICU environment. In ICU there are a high number of false positives (FP) and false negatives (FN). The patient condition changes over the time and, sometimes these changes are caused by the clinical procedures performed by the nurses and doctors. Often the patients recover due to the treatments performed, augmenting the FP or FN occurrences. However, is a medical imposition that the models shouldn't present a total percentage of error higher than 40%. The sensitivity can´t be lower than 85% because the objective is not to give a correct answer but to predict what will happen to the patient in the next hour. The minimum accuracy of the models is limited to 60%. This value avoids bad predictions or false positives. The quality measure introduced tries to decrease these limits.

The best models that meet the quality measures are executed using the data of the patient admitted in ICU. Then the results, the probabilities associated to the targets, are shown through INTCare system. This system is deployed by a hospital server and can be assessed anywhere and anytime through laptops, mobile devices or situated displays.

10. Conclusions and Future Work

This paper presents the most recent advances achieved in the context of the INTCare project, an IDSS to predict organ failure and outcome of the ICU patients. The results are mainly related to the KDD process. KDD process is currently executed in real-time using online learning and streamed data. This has been possible due to the modifications introduced in the environment and in the information system of the ICU. Intelligent agents are in charge for data acquisition, data processing and data transformation, in order to: induce data mining models; assess the models; and prepare the results to be accessible in the INTCare console.

Two main contributions are noteworthy. The first one is related to the possibility of handling data stream collected directly from the devices and systems present in the ICU. A pervasive data acquisition component has been implemented. Online and real-time acquisition, processing and transforming of data is now possible. This increases the number of data provided electronically and in real-time.

The second contribution refers to the ensemble approach adopted in order to adapt the predictive models automatically and in real-time. This work corroborated the initial hypothesis that it would be possible to induce data mining models automatically and in real-time using data streaming and online-learning in critical areas as is intensive medicine. However, some limitations remain in data acquisition. Some attributes still need to be recorded manually (e.g. Glasgow Score and Urine Output).

Experiments have been carried out in order to compare single model usage with an ensemble. The results obtained show that, in general, the ensemble performs better. The results were evaluated in terms of three metrics: sensitivity, accuracy and total error. Additionally, a quality measure has been established (sensitivity >= 85%; accuracy >= 60% and total error <=40%). Despite the good results attained, only three of the six initial targets fulfilled the quality measure: outcome, cardiovascular and coagulation. All of the three targets presented a high level of sensitivity (> 90%), and an acceptable level of total error and accuracy. This means that the doctors can use the predictions to avoid non return situations in those organic systems.

The advances attained in this work enable the implementation of a pervasive IDSS in ICU. The probabilities associated to organ failure and patient death for the next 24 hours are now available anywhere, anytime and in real-time (Mahmoud, 2003). The knowledge provided by the ensemble is also disseminated through the ICU platform and mobile devices. Despite its pervasiveness, the system doesn't interfere with the patient condition. The human efforts towards the information recording are minimized, providing more time to patient care and giving important contributions into the decision making process.

This work brought to the light some open issues that should be investigated in the near future. The first one is related to the quantity and diversity of data available. More cases and more diverse situations should be considered in the future in order to refine the models. In particular a deeper analysis should be done on the renal and hepatic systems. Ventilator variables will be used to understand if they improve the performance of the respiratory models.

Further work will involve also the tuning of the quality measure introduced in this work and will explore other methods to induce, validate and testing the ensemble.

Acknowledgments

This work is supported by FEDER through Operational Program for Competitiveness Factors – COMPETE and by national funds though FCT – Fundação para a Ciência e Tecnologia in the scope of the project: FCOMP-01-0124-FEDER-022674.

The authors would like to thank FCT (Foundation of Science and Technology, Portugal) for the financial support through the contract PTDC/EIA/72819/ 2006 (INTCare) and PTDC/EEI-SII/1302/2012 (INTCare II). The work of Filipe Portela was supported by the grant SFRH/BD/70156/2010 from FCT.

The authors would like to express their gratitude to the nursing and medical staff of the ICU of CHP and in particular to Dr. Fernando Rua for their valuable contributions to this project.

References

Abelha, A., Machado, M., Santos, M., Sollari, A., Rua, F., Paiva, M., et al. (2003). Agency for Archive, Integration and Diffusion of Medical Information. *Proceeding of AIA*.

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern recognition, 30*(7), 1145-1159.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and regression. *Wadsworth, Belmont, CA*.

Bricon-Souf, N., & Newman, C. R. (2007). Context awareness in health care: A review. [10.1016/j.ijmedinf.2006.01.003]. *International Journal of Medical Informatics, 76*(1), 2-12.

Cherkassky, V., & Mulier, F. (1999). Vapnik-Chervonenkis(VC) learning theory and its applications. *IEEE Transactions on Neural Networks, 10*(5), 985-987.

Concepts, O. D. M. (2005). 11g Release 1 (11.1). *Oracle Corp, 2007*.

Dietterich, T. (2000). Ensemble methods in machine learning. *Multiple classifier systems*, 1-15.

Duarte, J., Portela, F., Santos, M. F., António, A., & José, M. (2011). *Electronic Health Record in Dermatology Service.* Paper presented at the *HCist 2011* – International Workshop on Health and Social Care Information Systems and Technologies, Algarve, Portugal.

Eslami, S., de Keizer, N. F., Abu-Hanna, A., de Jonge, E., & Schultz, M. J. (2009). Effect of a clinical decision support system on adherence to a lower tidal volume mechanical ventilation strategy. *Journal of Critical Care, 24*(4), 523-529.

Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery: an overview.

Filipe Portela, M. F. S., José Machado, Álvaro Silva, Fernando Rua, António Abelha. (2012). *Intelligent Data Acquisition and Scoring System for Intensive Medicine.* Paper presented at the 3rd International Conference on Information Technology in Bio- and Medical Informatics, Viena, Austria.

Filipe Portela, P. G., Manuel Filipe Santos, Álvaro Silva, Fernando Rua. (2012a). *Intelligent and Real-time Data Acquisition and Evaluation to Determine Critical Events in Intensive Medicine.* Paper presented at the *HCist 2012* - International Conference on Health and Social Care Information Systems and Technologies, Portugal.

Filipe Portela, P. G., Manuel Filipe Santos, Álvaro Silva, Fernando Rua. (2012b). *Intelligent Data Acquisition and Calculation System for Intensive Medicine Critical Events.* Paper presented at the *HCist 2012* - International Conference on Health and Social Care Information Systems and Technologies, Portugal.

Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992). Knowledge Discovery in Databases: An Overview. *AI Magazine, 13*(3), 57-70.

Gaber, M. M., Zaslavsky, A., & Krishnaswamy, S. (2005). Mining data streams: a review. *ACM Sigmod Record, 34*(2), 18-26.

Gago, P., Santos, M. F., Silva, Á., Cortez, P., Neves, J., & Gomes, L. (2006). INTCare: a knowledge discovery based intelligent decision support system for intensive care medicine. *Journal of Decision Systems*.

Gama, J., & Gaber, M. M. (2007). *Learning from data streams: processing techniques in sensor networks*: Springer-Verlag New York Inc.

Gama, J., Santos, Filipe, M., Marques, N., & Cortez, P. (2010). Ubiquitous Data Mining (UDM). Retrieved 07-01-2010, 2010, from http://www.dei.isep.ipp.pt/~cgf/ECAI10/ws.html

Golab, L., Johnson, T., & Shkapenyuk, V. (2009, 2009). *Scheduling updates in a real-time stream warehouse*.

Guy, W. (1976). *ECDEU assessment manual for psychopharmacology*: Rockville, Md.

Han, J., & Kamber, M. (2006). *Data mining: concepts and techniques*: Morgan Kaufmann.

Hoo, G. W. S. (2009). Barotrauma and Mechanical Ventilation. 24. Retrieved from http://emedicine.medscape.com/article/296625-print

Kantardzic, M. (2011). *Data mining: concepts, models, methods, and algorithms*: Wiley-IEEE Press.

Kaplow, R., & Hardin, S. R. (2007). *Critical care nursing: synergy for optimal outcomes*: Jones & Bartlett Learning.

Kuperman, G. J., Bobb, A., Payne, T. H., Avery, A. J., Gandhi, T. K., Burns, G., et al. (2007). Medication-related clinical decision support in computerized provider order entry systems: a review. *Journal of the American Medical Informatics Association, 14*(1), 29-40.

Laplace, P. S. (1820). *Théorie analytique des probabilités*: Courcier.

Mahmoud, M. (2003). Real-time data acquisition system for monitoring patients in intensive care unit. (Vol. 5090, pp. 320-326). Multisensor, Multisource Information Fusion: Architectures, Algorithms and Applications.

Mikkonen, M., Va¨ yrynen, S., Ikonen, V., & Heikkila, M. O. (2002). User and concept studies as tools in developing mobile communication services for the elderly. *Personal and Ubiquitous Computing, 6*(2), 113-124.

Milenova, B. L., Yarmus, J. S., & Campos, M. M. (2005, 2005). *SVM in oracle database 10g: removing the barriers to widespread adoption of support vector machines*.

Portela, F., Gago, P., Santos, M. F., Silva, A., Rua, F., Machado, J., et al. (2011). *Knowledge Discovery for Pervasive and Real-Time Intelligent Decision Support in Intensive Care Medicine*. Paper presented at the *KMIS* 2011- International Conference on Knowledge Management and Information Sharing.

Portela, F., Santos, M., Vilas-Boas, M., Rua, F., Silva, Á., & Neves, J. (2010). *Real-time Intelligent decision support in intensive medicine*. Paper presented at the *KMIS* 2010- International Conference on Knowledge Management and Information Sharing.

Portela, F., Santos, M. F., Gago, P., Silva, Á., Rua, F., Abelha, A., et al. (2011). Enabling real-time intelligent decision support in intensive care.

Portela, F., Santos, M. F., Silva, Á., Machado, J., & Abelha, A. (2011). *Enabling a Pervasive approach for Intelligent Decision Support in Critical Health Care*. Paper presented at the *HCist 2011* – International Workshop on Health and Social Care Information Systems and Technologies.

Portela, F., Santos, M. F., & Vilas-Boas, M. (2012). A Pervasive Approach to a Real-Time Intelligent Decision Support System in Intensive Medicine. In Springer (Ed.), *Communications in Computer and Information Science* (Vol. 0272, pp. 14).

Saha, D., & Mukherjee, A. (2003). Pervasive computing: a paradigm for the 21st century. [10.1109/MC.2003.1185214]. *IEEE Computer, 36*(3), 25-31.

Santos, M. F., & Portela, F. (2011). *Enabling Ubiquitous Data Mining in Intensive Care - Features selection and data pre-processing*. Paper presented at the 13th International Conference on Enterprise Information Systems.

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., & Neves, J. (2011a). *INTCARE - Multi-agent approach for real-time Intelligent Decision Support in Intensive Medicine.* Paper presented at the 3rd International Conference on Agents and Artificial Intelligence (*ICAART*), Rome, Italy.

Satyanarayanan, M. (2002). Pervasive computing: vision and challenges. *Personal Communications, IEEE, 8*(4), 10-17.

Shim, J. P., Warkentin, M., Courtney, J. F., Power, D. J., Sharda, R., & Carlsson, C. (2002). Past, present, and future of decision support technology. *Decision support systems, 33*(2), 111-126.

Silva, Á. (2007). *Modelos de Inteligência Artificial na análise da monitorização de eventos clínicos adversos, Disfunção/Falência de órgãos e prognóstico do doente critico.* Universidade do Porto, Porto.

Silva, Á., Cortez, P., Santos, M. F., Gomes, L., & Neves, J. (2008). Rating organ failure via adverse events using data mining in the intensive care unit. *Artificial Intelligence in Medicine, 43*(3), 179-193.

Silva, Á., Cortez, P., Santos, M. F., Gomes, L., & Neves, J. (2008). Rating organ failure via adverse events using data mining in the intensive care unit. *Arificial Intelligence in Medicine, 43*(3), 179-193.

Simon, H. A. (1960). The New Science of Management Decision Making. *New York-Evanston*.

Steinberg, D., & Colla, P. (1997). CART: classification and regression trees. *San Diego, CA*.

Suter, P., Armaganidis, A., Beaufils, F., Bonfill, X., Burchardi, H., Cook, D., et al. (1994). Predicting outcome in ICU patients. *Intensive Care Medicine, 20*(5), 390-397.

Tamayo, P., Berger, C., Campos, M., Yarmus, J., Milenova, B., Mozes, A., et al. (2005). Oracle Data Mining. *Data Mining and Knowledge Discovery Handbook*, 1315-1329.

Tehrani, F. T. (2008a). Automatic control of mechanical ventilation. Part 1: Theory and history of the technology. *Journal of clinical monitoring and computing, 22*(6), 409-415.

Tehrani, F. T. (2008b). Automatic control of mechanical ventilation. Part 2: the existing techniques and future trends. *Journal of clinical monitoring and computing, 22*(6), 417-424.

Ting, K. M., & Witten, I. H. (2011). Issues in stacked generalization. *Arxiv preprint arXiv:1105.5466*.

Varshney, U. (2009). *Pervasive Healthcare Computing: EMR/EHR, Wireless and Health Monitoring*: Springer-Verlag New York Inc.

Verner, U., Schuster, A., & Silberstein, M. (2011, 2011). *Processing data streams with hard real-time constraints on heterogeneous systems*.

Villas Boas, M., Gago, P., Portela, F., Rua, F., Silva, Á., & Santos, M. F. (2010). *Distributed and real-time Data Mining in the Intensive Care Unit*. Paper presented at the 19th European Conference on Artificial Intelligence - ECAI 2010.

Vincent, J., Mendonça, A., Cantraine, F., Moreno, R., Takala, J., Suter, P., et al. (1998). 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, 26*, 1793-1800.

Vincent, J. L., de Mendonca, A., Cantraine, F., Moreno, R., Takala, J., Suter, P. M., et al. (1998). 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, 26*(11), 1793.

Vincent, J. L., Moreno, R., Takala, J., Willatts, S., De Mendonca, A., Bruining, H., et al. (1996). The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive care medicine, 22*(7), 707-710.

W. Guy Modified From: Rush J, e. a. (2000). Clinical Global Impressions (CGI) Scale. *Psychiatric Measures, APA*.

Wang, H., Fan, W., Yu, P. S., & Han, J. (2005, 2003). *Mining concept-drifting data streams using ensemble classifiers.* Paper presented at the ACM SIGKDD international conference.

Wang, W., Partridge, D., & Etherington, J. (2001, 2001). *Hybrid ensembles and coincident-failure diversity*.

Wirth, R., & Hipp, J. (2000). *CRISP-DM: Towards a standard process model for data mining.* Paper presented at the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining.

Wolpert, D. H. (1992). Stacked generalization*. *Neural networks, 5*(2), 241-259.

Wooldridge, M. (1999). Intelligent agents *Multiagent systems: a modern approach to distributed artificial intelligence* (pp. 27-77): MIT Press.

Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: Theory and practice. *Knowledge engineering review, 10*(2), 115-152.

Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., et al. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems, 14*(1), 1-37.

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem, 39*(4), 561-577.

Appendix

Table A1. 10 runs results by model

Target	Model	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	AVG
CARDIOVASCULAR	$M_{6,2,2}$	0,9825	0,9792	0,9823	0,9778	0,9825	0,9810	0,9745	0,9802	0,9806	0,9740	0,9795
	$M_{1,2,2}$	0,9778	0,9750	0,9823	0,9778	0,9759	0,9802	0,9731	0,9802	0,9753	0,9726	0,9770
	$M_{5,2,2}$	0,9778	0,9750	0,9823	0,9778	0,9759	0,9802	0,9731	0,9802	0,9753	0,9726	0,9770
	Ensemble	0,9825	0,9792	0,9823	0,9778	0,9825	0,9810	0,9745	0,9802	0,9806	0,9740	0,9795
COAGULATION	$M_{4,3,3}$	0,9355	0,8954	0,9115	0,9168	0,9405	0,9534	0,9229	0,9192	0,9008	0,8237	0,9120
	$M_{6,3,3}$	0,9275	0,8901	0,9069	0,9038	0,9290	0,9458	0,9194	0,9061	0,8947	0,7878	0,9011
	$M_{5,3,3}$	0,8672	0,8183	0,8542	0,7916	0,8641	0,7855	0,8962	0,8603	0,7931	0,7954	0,8326
	$M_{4,3,1}$	0,8679	0,7656	0,7908	0,8122	0,8084	0,8168	0,7998	0,7366	0,8084	0,7809	0,7987
	Ensemble	0,9355	0,8954	0,9115	0,9168	0,9405	0,9534	0,9229	0,9192	0,9008	0,8237	0,9120
HEPATIC	$M_{6,5,3}$	0,5792	0,4806	0,6000	0,5611	0,6222	0,4681	0,6806	0,6889	0,5028	0,6840	0,6017
	$M_{4,5,3}$	0,6096	0,5069	0,6667	0,5764	0,6218	0,4542	0,7278	0,6889	0,4903	0,6684	0,5867
	$M_{5,5,1}$	0,5431	0,5323	0,8465	0,6389	0,3597	0,6625	0,7087	0,7542	0,7792	0,7458	0,6571
	$M_{5,5,3}$	0,5833	0,4958	0,5569	0,6297	0,5250	0,6386	0,3347	0,4583	0,3225	0,6097	0,5345
	Ensemble	0,6096	0,5323	0,8465	0,6389	0,6278	0,6625	0,7278	0,7542	0,7792	0,7458	0,6924
OUTCOME	$M_{1,6,1}$	1,0000	0,9990	1,0000	1,0000	0,9969	1,0000	0,9969	0,9765	0,9918	0,9949	0,9956
	$M_{4,6,1}$	0,9275	1,0000	1,0000	1,0000	0,9254	0,9980	1,0000	0,9928	0,9908	0,9081	0,9743
	$M_{6,6,1}$	0,9162	0,9521	0,9816	0,9734	0,9162	0,9642	0,9837	0,9070	0,9285	0,8529	0,9376
	$M_{5,6,1}$	0,8764	0,9510	0,9571	0,9683	0,8989	0,8979	0,9856	0,9530	0,9070	0,8447	0,9240
	$M_{3,6,1}$	0,8917	0,9425	0,9734	0,8917	0,9540	0,8713	0,8760	0,9459	0,7508	0,8458	0,8943
	$M_{2,6,1}$	0,8274	0,9223	0,8693	0,8856	0,9397	0,8621	0,8172	0,9142	0,8029	0,8662	0,8707
	Ensemble	1,0000	1,0000	1,0000	1,0000	0,9969	1,0000	1,0000	0,9928	0,9918	0,9949	0,9977
RENAL	$M_{4,4,3}$	0,7352	0,9115	0,8993	0,6602	0,7240	0,7003	0,8501	0,5755	0,7074	0,9530	0,77,17
	$M_{6,4,3}$	0,7338	0,7842	0,8200	0,6601	0,6794	0,6595	0,7794	0,4887	0,7026	0,7806	0,7088
	$M_{5,4,3}$	0,6667	0,7074	0,7058	0,6559	0,6578	0,6451	0,7086	0,4477	0,6415	0,7626	0,6599
	Ensemble	0,7352	0,9115	0,8993	0,6602	0,7240	0,7003	0,8501	0,5755	0,7074	0,9530	0,77,17
RESPIRATORY	$M_{1,1,1}$	0,6295	0,6220	0,7241	0,7464	0,7418	0,6576	0,7356	0,6269	0,6616	0,7041	0,6850
	$M_{6,1,3}$	0,6383	0,6437	0,5825	0,6541	0,7374	0,6502	0,7764	0,6659	0,5752	0,6800	0,6604
	$M_{4,1,3}$	0,6475	0,6451	0,5755	0,6523	0,7435	0,6383	0,7698	0,6589	0,5572	0,6993	0,6588
	$M_{5,1,3}$	0,6230	0,6427	0,6104	0,6677	0,7199	0,6326	0,7448	0,6694	0,5879	0,6725	0,6571
	$M_{4,1,1}$	0,6068	0,6586	0,5885	0,5791	0,7339	0,5879	0,7229	0,6103	0,5861	0,6703	0,6344
	$M_{5,1,1}$	0,5783	0,6316	0,5834	0,5686	0,7242	0,5616	0,7067	0,5607	0,5620	0,6348	0,6112
	$M_{6,1,1}$	0,5300	0,6480	0,5564	0,6015	0,7155	0,5335	0,7080	0,4660	0,6028	0,6905	0,6052
	$M_{2,1,1}$	0,4936	0,5572	0,5582	0,5783	0,5831	0,4761	0,5822	0,4336	0,6120	0,6567	0,5531
	$M_{3,1,1}$	0,4437	0,5848	0,5694	0,5787	0,6068	0,4713	0,5804	0,4266	0,6151	0,6138	0,5491
	Ensemble	0,6475	0,6586	0,7241	0,7464	0,7435	0,6576	0,7764	0,6694	0,6616	0,7041	0,6989

ARTICLE A5

Objectives

To develop and test pervasive models in Intensive Care

i. To prototype an IDSS;
ii. To test the models developed;
iii. To include pervasive capabilities in the IDSS;

To a better comprehension of the patient's condition it is very important to understand how critical the values associated to some patient are. Until now this type of values weren't considered. In order to overcome this problem a Critical Events tracking system was developed and deployed in the Electronic Nursing Record (ENR).

Results

Each patient has an ENR workstation at the bedside. The ENR contains information regarding several variables: vital signs, ventilation, medical scales, fluid balances. Simultaneously the data provided by them they are used to automatically calculate the number and the duration of an event.

Having in account the settings of a real environment it was necessary to define a set of procedures to automatically compute critical events (CE) for five variables: Urine Output, Blood Pressure, Heart Rate, Respiratory and Temperature. Furthermore an intuitive and user-friendly interface was designed. This interface is accessible anywhere and anytime and the number and duration of CE are disseminated through situated ICU devices (touch monitors), one for each patient bed.

This new approach represents a set of innovations in the way of the data are collected and used to support the decision making process. It is now possible to acquire, process, and present knowledge automatically and in real-time using online-learning, anywhere and anytime.

Aiming at promoting proactive actions with the patient this Pervasive Intelligent System is able to determinate and evaluates the number and duration of critical events patients. The doctors in ICU of Centro Hospital do Porto now have some new information that helps them to have a better understanding of the patient's condition.

The doctors can see, for each patient and in real time the number of events by hour and the time in which the patient was in a critical event. In addition, they possess a traffic light system (green, yellow, red) to alert / show the present situation for each event.

PERVASIVE REAL-TIME INTELLIGENT SYSTEM FOR TRACKING CRITICAL EVENTS IN INTENSIVE CARE PATIENTS

Filipe Portela [1], Pedro Gago[2], Manuel Filipe Santos[1], José Machado[3], António Abelha[3], Álvaro Silva[4], Fernando Rua[4]

[1] Centro Algoritmi, Universidade do Minho, Guimarães, Portugal
cfp@dsi.uminho.pt, mfs@dsi.uminho.pt

[2] Instituto Politécnico de Leiria, Leiria, Portugal,
pedro.gago@.ipleiria.pt, fpinto@.ipleiria.pt

[3] Centro de Ciências e Tecnologias de Computação, Universidade do Minho, Braga, Portugal

[4] Serviço de Cuidados Intensivos, Centro Hospitalar do Porto, Hospital Santo António, Porto, Portugal
moreirasilva@clix.pt, fernandorua.sci@hgsa.min-saude.pt

Abstract. Nowadays it is fundamental in critical areas as is Intensive Medicine to have intelligent systems that are able to support the decision making process (DMP) giving important information in the right moment. Some of the biggest problems faced by such systems are related both to the number and the different types of data sources present in Intensive Care Units (ICU). Even though in such a setting the values for some variables are easy to collect, data collection is still performed manually for some others. In order to help the DMP in ICU, a Pervasive Intelligent Decision Support System, called INTCare was deployed in the ICU of Centro Hospitalar do Porto in Portugal. This system changed the way of the information is collected and presented. Taking advantage of the change of the environment and the data acquisition system, a system for critical events tracking was developed as the use of information regarding critical events to support decision making in Intensive Care Units is considered very useful.
The tracking system was deployed in a particular module of INTCare – Electronic Nursing Record (ENR) and it is accessible anywhere and anytime. The system allows for the calculation of the critical events regarding five variables that are usually monitored in an ICU. Moreover, this system is composed by a grid that shows the events by type and duration, a warning system to alert the doctors and intuitive graphics that allow them to follow the patient evolution. User acceptance was measured through a questionnaire designed in accordance with the Technology Acceptance Methodology (TAM). This paper presents the tracking system, its interface and the results achieved with TAM.

Keywords: Critical Events; Adverse Events; Intensive Care Units; INTCare; Real-Time Data Processing; Pervasive Systems; Technology Acceptance Methodology; Tracking Systems; Decision Making Process

1. Introduction

Support for the decision making process in Intensive Care Units is still far from being enough. Even though ICUs are flooded by a wealth of situated devices that give much information about the patient it is sometimes still very difficult to make use of the collected information due the high number of data sources. Patients have sensors connected to bedside monitors for monitoring a set of data as vital signs (eg. blood pressure, oxygen saturation, heart rate, temperature) ventilation (eg. ventilation type, PEEP) and others (Urine Output, Fluid Balance). Regularly they make a set of laboratory exams and other types of procedures. Some of that data is usually collected manually and even when it is registered in digital format that often done using proprietary applications that limit future data access. Some modifications were done in the environment and in the information system architecture in order to both change the way how the data are collected and to make these data accessible so as to be available to support the decision making process. These modifications in the environment were performed having in account the INTCare system features, i.e., the environment was prepared to allow applications such as INTCare to work (Filipe Portela et al., 2011; Portela, Santos, & Vilas-Boas, 2012).

INTCare is a pervasive intelligent decision support system developed in the Intensive Care Unit of Hospital Santo António, Centro Hospitalar do Porto in Portugal. This system uses an ensemble of classifiers in order to perform online-learning to predict both organ failures in the next 24 hours and the patient's hospital outcome (dead or alive).

Instead of a more traditional approach where a static dataset is used, INTCare is able to work with a number of data streams (Mador & Shaw, 2009). Indeed, for each monitored variable in the ICU one has a data stream and, to further complicate matters, it is often the case that data from different sources are gathered at different time intervals. Currently most patient data are automatically acquired and processed in real-time. Data thus collected and stored constitutes an important knowledge base that will be used to support the decision making process (DMP).

INTCare represents an important contribution to DMP. This system helps the doctors to have a better comprehension of patient condition. In this step the information provided by the IDSS are crucial. Indeed, given the huge number of variables to be considered and the rate at with they vary; automatic data processing is of paramount importance to highlight value changes that are potentially medically relevant. To a better comprehension of the patient's condition it is very important to understand how critical the values associated to some patient are. Until now this type of values weren't considered. In order to overcome this problem a Critical Events tracking system was developed and deployed in the Electronic Nursing Record (ENR).

Each patient has an ENR workstation at the bedside. The ENR contains information regarding several variables: vital signs, ventilation, medical scales, fluid balances. Simultaneously the data provided by them they are used to automatically calculate the number and the duration of an event. Having in account the settings of a real environment it was necessary to define a set of procedures to automatically compute critical events (CE) for five variables: Urine Output, Blood Pressure, Heart Rate, Respiratory and Temperature.

Furthermore an intuitive and user-friendly interface was designed. This interface is accessible anywhere and anytime and the number and duration of CE are disseminated through situated ICU devices (touch monitors), one for each patient bed. During this work the CE system was also assessed in terms of usability and technology acceptance (Chooprayoon & Fung, 2010). Using the Technology Acceptance Methodology (TAM) (Chooprayoon & Fung, 2010) and a structured questionnaire, the INTCare system (Filipe Portela, Jorge Aguiar, Manuel Filipe Santos, Álvaro Silva, & Fernado Rua, 2013) was evaluated in four aspects: perceived usefulness (PU), perceived ease of use (PEOU), behavioural intention (BI) and use behaviour (UB). As part of INTCare the CE system also was evaluated. The results obtained show that the users are satisfied with the implementation of the critical events tracking system.

The paper is divided in seven sections, the first (this one) introduces the paper and the second section addresses the problem (INTCare, Intensive Care, Critical Events, Pervasive Health Cate and Technology Acceptance Methodology). The third and fourth sections, present the data acquisition process, the data analysis and the pervasive system. Section five introduces the CE tracking system and the sixth section presents the results obtained at level of interface and technology acceptance. Finally some concluding considerations are made and future work presented.

2. Background

2.1. INTCare

The development of a system for Critical Events tracking in Intensive Medicine is part of a research project called INTCare (Gago et al., 2006). The algorithms in use in INTCare's prediction module – ensemble data mining, require information regarding the critical events that have occurred during the stay of the patient in the ICU. In order to be able to use INTCare with real-time data one must have some means of extracting the CE also in real-time.

INTCare includes an automatic data acquisition module to acquire data in real-time and an Electronic Nursing Record (ENR) to collect some data that requires manual observation (Filipe Portela, Pedro Gago, et al., 2011) as is urine output. ENR maintains the patients hourly clinical records and other information relevant to the decision making process.

The data used by the INTCare system are collected from several different platforms and include the patient's vital signs, medical procedures, therapeutic plans, lab results, medical scores, information regarding ventilation and others. The ENR is a touch screen platform and it is deployed in a workstation near each ICU bed. Medical and nursing staff may use them to monitor the patient, i.e., record, validate and visualize all clinical information about each patient. INTCare (F. Portela et al., 2011; Portela, et al., 2012) is implemented in the Intensive Care Unit of Hospital Santo António, Centro Hospitalar do Porto.

2.2. Intensive Care

Intensive care is a critical area of medicine, where the patients are in too weak conditions and/or in serious life-risk (Bricon-Souf & Newman, 2007). Intensive Care units are the place where this type of medicine is applied. Usually, during the stay of these patients is possible to verify a set of adverse events. These events can influence the future outcome and can occur several times a day (Rothschild et al., 2005). The ability to calculate the events automatically and in real-time is an important support to the decision making process. In addition, the development of an Intelligent Decision Support System, like INTCare, which uses these variables to predict the patient condition, can help the doctors to maintain a pro-active action that will benefit the patients.

2.3. Critical Events

Studies done in the past reported that the most common adverse errors were due to wrong mechanical or human performance (Kaur, Pawar, Kohli, & Mishra, 2008). However, there are other issues that are difficult to analyze (eg. patient clinical events). This happens because it is very difficult to quantify the number of clinical errors due to a lack of automatic data acquisition systems in the ICUs. Normally, these results are collected by some alerts provided from bedside monitors (Keegan, Gajic, & Afessa, 2011). This paper will explain an approach to obtain the number and the duration of clinical adverse events for five variables (Heart Rate, Blood Pressure, Oxygen Saturation, Diuresis and Temperature).

The data needed to determine adverse events were continually recorded and processed by an electronic application. To understand if an event is critic or not, two main criteria were used (Silva, Cortez, Santos, Gomes, & Neves, 2008):

- occurrence and duration should be registered by physiological changes;
- related physiological variables should be routinely registered at regular intervals.

An event is considered critical when out of range values occur for a longer time or when the values are extremely out of range (Silva, et al., 2008).

In this project we used two different definitions: critical values and critical events. Critical values are values that are out of a normal range. Critical event are defined as labels to signal that a variable had critical values for more than the admissible time span, as defined in Table V.1. Also, a critical event may signal that the critical value was so out of range that it is considered serious regardless of the duration of that observation. For example, a critical event happens whenever the patient's heart rate stays above 120 bpm for more than 1 hour. Also, a critical event happens every time the heart rate drops below 30 bpm or rises above 180 bpm.

Table V.1 – The protocol for the out of range physiologic measurements (adapted from (Álvaro Silva, et al., 2008))

	BP	*SpO2*	*HR*	*UR*
Normal range	90—180 mmHg	>= 90%	60—120 bpm	>= 30 ml/h
Critical event [a]	>= 1h	>= 1h	>= 1h	>= 2h
Critical event [b]	< 60 mmHg	<80%	<30 bpm V >180 bpm	<= 10 ml/h

[a] Defined when continuously out of range.

[b] Defined anytime.

2.4. Pervasive health care

Pervasive health care derives from the concept of pervasive computing. Pervasive computing is characterized by Satyanarayanan (Satyanarayanan, 2002) as an evolutionary step resulting from two other steps: first distributed computing and then mobile computing. According to the own research it involves not only issues associated with the communication between the environments and their interaction with users, but also in supporting mobility of the users.

In this context Varshney (Varshney, 2009) defined pervasive health care as "conceptual system of providing healthcare to anyone, at any time, and anywhere by removing restraints of time and location while increasing both the coverage and the quality of healthcare".

The information returned by the applications should be stored and accessible from a website and includes some modifications to be able to access from small portable devices (Mikkonen, Va¨ yrynen, Ikonen, & Heikkila, 2002). As an example of applications that can be developed for this type of environments we have: universal systems of monitoring of clinical data, intelligent emergency management, universal access to clinical data and mobile and ubiquitous telemedicine (Varshney, 2007).

INTCare is categorized as a pervasive system because it is accessible anywhere and anytime through portable devices with access to the hospital network.

2.5. Technology Acceptance Methodology

Technology Acceptance Methodology(TAM) "is adapted from the Theory of Reasoned Action (TRA) model which describes human behaviours in a specific situation" (Fishbein & Ajzen, 1975). The main goal of TAM is study the effects of external variables towards people's internal beliefs, attitudes, and intentions (Chooprayoon & Fung, 2010).

TAM is supported by the use of questionnaires. In this case a questionnaire was prepared by a coordination team, composed by professionals of ICU and Information System, and sent to a set of participants (a group of experts from the ICU nurses team).

The questionnaire was prepared taking into account the constructs of TAM (Venkatesh & Bala, 2008; Venkatesh & Davis, 2000): perceived usefulness (PU), perceived ease of use (PEOU), behavioural intention (BI) and use behaviour (UB).

3. Data acquisition Process

The data acquisition process is the first step of the tracking system. In order to develop an automatic system the platform must be able to automatically gather the values for all variables continuously and in real-time. As previously reported, initially none of the variables was automatically acquired or/and in an electronic mode. Every hour the values were manually registered in a paper based nursing record. This reality lead the research team to implement some changes in the environment and in the data acquisition system (Filipe Portela, Pedro Gago, et al., 2011; F. Portela, et al., 2011; Portela, Santos, Silva, Machado, & Abelha, 2011). Currently, it is possible acquire automatically, with the exception of the volume of urine output, the entire set of patient data associated with the critical events. In the case of urine output the value is manually entered by the nurses in the ENR.

For the critical events system, the following data sources are being used:

- Bed Side Monitor - vital signs (VS);
- Electronic Nursing Record (ENR) – Hourly values (Diuresis);

To perform the tasks associated to the data acquisition process, a set of intelligent agents has been implemented (Santos et al., 2011) to automatically acquire and process the data from the different ICU data sources.

Firstly, the data are extracted from the ICU data sources by the agents; during the extraction some automatic procedures are executed. Then, all process of interpretation of values is executed by the pre-processing agent and finally all data are loaded into a data warehouse that stores all critical values to be used by the Critical Events System.

Figure V.1 summarizes the process of the data acquisition system which has been developed in order to collect data from bedside monitors and ENR. These data are either automatic collected from the bedside monitors through the gateway agent or manually validated / inserted in the ENR by the ICU nurses. Afterwards, all the data is processed in order to obtain the critical events and the CE system is executed.

Afterwards the results attained are disseminated through the ENR. At same time the results are available to the INTCare prediction agent so that it may obtain the target probabilities. These results are available to consult by the doctors anywhere and anytime using a mobile device.

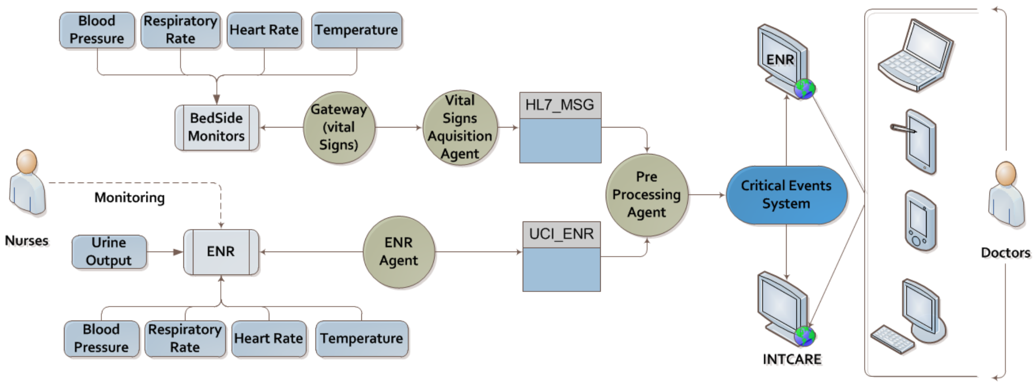


Figure V.1– ICU Critical Events – Data acquisition.

Nowadays it is important to concentrate the efforts on detecting the critical events that are deemed most important by the doctors we are working with. Automated data acquisition without human intervention is already in place for most of them even though values can be manually inserted / validated by ICU staff. However, the urine output measurement isn't automated at the ICU and the respective values have to be manually entered by the nursing staff.

Table V.2, shows which are the variables collected in real-time, the correspondent data source and if they are acquired in an automatic or manual way. In this case, only one of the variables requires human intervention, all the other are collected automatically and can be monitored manually.

Table V.2 – Critical variables data source and acquisition type

Variable	Data Source	Acquisition
Blood Pressure (BP)	BM	Automatic
Temperature (TEMP)	BM	Automatic
Hearth Rate (HR)	BM	Automatic
Oxygen Saturation (O2)	BM	Automatic
Urine Output (UR)	ENR	Manual (hour)

4. Data Analysis

Data pre-processing is executed by the pre-processing agent. This agent is responsible for the automatic execution of all the tasks required to determine the number and the duration of critical events. This process is central to the entire system.

A first pass through the data allows for validation and for detection of values that may be part of CE. The data quality is fundamental to obtain the correct number of critical events. In this case, all values collected are being used, i.e., all values are necessary to be able to do a continuous calculation of events.

In order to streamline the process it is necessary to have an automatic validation procedure. This procedure is activated whenever a new value arrives. Each value is checked to see if it is inside the valid range of values for the respective variable and then it is flagged as potentially critical or non-critical. This task is executed according to the values defined in Table V.3 (Filipe Portela, Pedro Gago, Manuel Filipe Santos, Álvaro Silva, & Rua, 2012). For example, a value of 10 for the urine output will be discarded as it is out of the valid range. A value of 190 (anytime) for the same variable will be flagged as critical and 70 as potentially critical (it will be confirmed as critical if the values stay out of range for at least 1 hour).

Table V.3 – Data Ranges

Ev_Id	Descr	Min EC	Max EC	Min Val	Max Val	Min Any	Max Any
3510	TEMP	36	38	34	45	35	40
1011	BP	90	180	0	300	60	
3000	O2	90	100	0	100	80	
2009	HR	60	120	0	300	30	180
DIU	UR	30	1000	0	1000	10	

The procedure referred is applied only to the variables that are to be used in the calculation of the critical events. This procedure is executed through a trigger (1) which is activated before the value is inserted in the table containing HL7 data. For each value collected:

```
BEGIN
    IF CATEGORY IN (3000, 2009, 1011, 3510, 'DIU') THEN
      IF VALUE >= MIN_VALUE_ICU AND VALUE <= MAX_VALUE_ICU THEN
        IF (VALUE   <= MIN_ANYTIME OR VALUE >= MAX_ANYTIME) THEN
          SET CRITIC TO 2
          SET VALID TO 1
        ELSIF VALUE <= MIN_EC OR VALUE >= MAX_EC THEN
          SET CRITIC TO 1
          SET VALID TO 1
        ELSE
          SET CRITIC TO 0
          SET VALID TO 1
        ENDIF
```

```
        ENDIF
      ENDIF
  END
```

Immediately after this procedure is executed another trigger, which is responsible to calculate the time of these events, is activated.

5. Critical Events Tracking System

The tracking system is executed in real-time using automatic data acquisition and data processing tasks. After collecting the critical values it is necessary to check if they originated a critical event. The rules for the calculation are in Table V.1. In order to help ascertain if the event type collected is or is not the same which was before collected, a flag in the table will be used. The flag φ will identify the state of the event type collected, i.e., the flag value present in the table allows the system to know if there is some event in the table that has the same type and if that event is open (1) or not (0). An “open” event is still in progress while a “closed” one has already finished.

When a value is collected and, immediately after it is validated and inserted in the HL7 table, the trigger will verify if there is some “open” record for this event category. If it doesn’t exist, a new row will be created. If there is a record for the same event type, nothing it is done, otherwise the event finish date will be defined according the collected date of the obtained value. For each event type (0, 1, 2) some operations will be done according to the respective event state (open or not). Finally and after the event start and event finish date is filled, a procedure, which calculates the time of each event is executed. For each patient, event category and value collected (PEVID) and, after the values are correctly inserted in database, a trigger is executed to verify if the event is closed and if it is the same type in order to calculate the event duration and then classify the event as critic or normal. After the value is correctly identified and inserted into the database, with total time filled, another trigger will be executed. This other trigger is designed to verify if an event is critic or not. To identify if a set of collected results is or not a critical event it is necessary to analyse the table which contains all values collected with the respective time interval. By computing the length of that time interval, the procedure identifies which set of values are critical:

```
  BEGIN
      IF OPEN = 0 THEN
        SET DATE_START TO SYSDATE
        SET DATE_FINISH TO NULL
        SET OPEN TO 1
        SET CRITIC TO EVTYPE
      ENDIF
      IF OPEN = 1 THEN
        IF CRITIC <> EVTYPE THEN
```

```
          SET DATE_FINISH TO SYDATE
          SET OPEN TO 0
          SET CRITIC TO EVTYPE
          SET TOTAL_TIME TO DATE_START - DATE_FINISH
        ELSE
          NULL
        ENDIF
      ENDIF
  END
```

After the value is correctly identified and inserted into the database, with total time filled, another trigger will be executed. This other trigger is designed to verify if an event is critic or not.

To identify if a set of collected results is or not a critical event it is necessary to analyse the table which contains all values collected with the respective time interval. By computing the length of that time interval, the procedure will identify which set of values are critical:

```
  BEGIN
    READ ROW
    IF EV_TYPE = 2 THEN
      SET CRITICALEVENT TO 1
    ELSIF EV_TYPE = 1
      IF CATEGORY = 'DIU'
        IF TOTAL_TIME >= 7200 THEN
          SET CRITICALEVENT TO 1
        ELSE
          SET CRITICALEVENT TO 0
        ENDIF
      ELSIF TOTAL_TIME >= 3600 THEN
        SET CRITICALEVENT TO 1
      ELSE
        SET CRITICALEVENT TO 0
      ENDIF
    ENDIF
  END
```

6. Results

Our system is continually collecting and processing data and is continually providing us updated information regarding the number and duration of critical events for each patient. It is a two-step procedure: in the first step the value being analysed is flagged as critical or not critical and in a second step the systems calculates the Accumulated Critical Events (ACE) – to reflect the patients' clinical evolution/severity of illness by hour. The values obtained will be used to create some ratios (Filipe Portela, et al., 2012). The procedures presented next are executed at regular intervals and have the objective to understand the number of events by type and hour, and their sum. The procedure is used to count the number of events and to sum the time; the values are grouped by type and hour.

In order to illustrate the functionality of the system, Figure V.2 presents the distribution of the values collected during the last six months. The values are grouped by variable, event type and hour. In the graph is possible to see the percentage of occurrence for three of the five variables (blood pressure (BP), SPo2, Heart Rate (HR) by hour (0-23) and event type (1 or 2). Analysing the figure 2 it is possible to retain some important conclusions (Filipe Portela, et al., 2012). The critical events of BP (59,85%) and HR (85,91%) are strongly associated to the events type 1 (less critical). In the case of SPo2 the critical events are mostly of the type 2 (68,22%). The vast majority of the events (almost 20%) occur between 9 and 11 am. This is most evident for the HR (critical event type 1) and SpO2 (critical event type 2). For the Blood Pressure, the 15th hour is the most critical (6,64% = 3,91% (1) + 2,73% (2) of the critical events occurred during this hour). This sort of analysis can help the doctors understand when each the event type normally occurs. The information obtained about the critical events combined with other variables can be used to alert the doctors in order to avoid critical events.

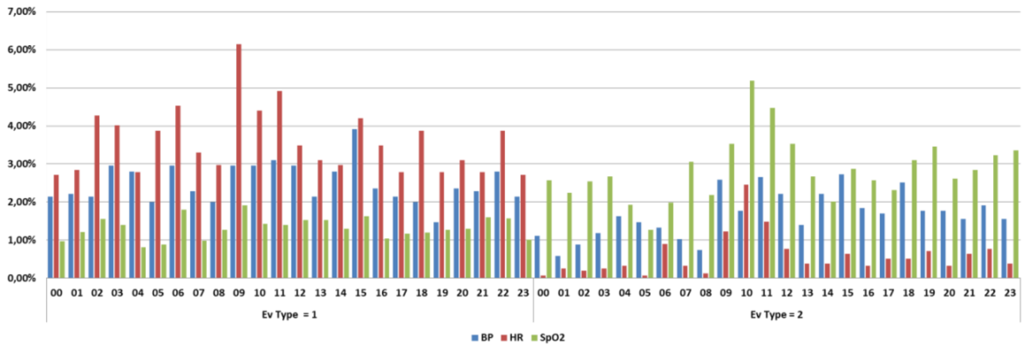


Figure V.2– ICU Critical Events – Data acquisition distribution by hour.

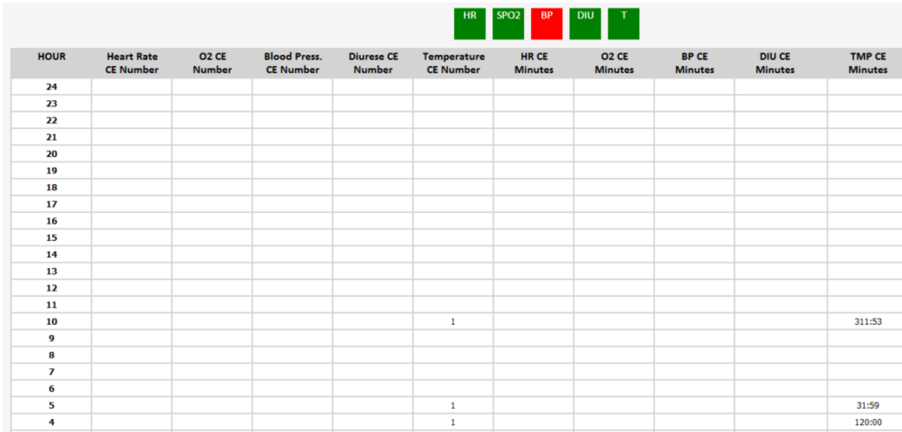
6.1. Tracking System - Interface

The results obtained by the tracking system are presented inside the Electronic Nursing Record application. The results are presented in two different ways: a grid (table) and a chart. The table header is composed by 13 columns: the number, the duration of each CE and the total of events. The table also has 24 lines, one for each hour of the day.

The system fills the grid according to the patient values, i.e., if an event is critical or not and their duration.

Figure V.3 presents an overview of the CE grid. For example, in the tenth hour of the day the patient finished a Temperature critical event with the duration of 311 minutes.

This way of visualizing events also has a warning system to alert if current values are out of range, i.e., critical values. It is represented by a colour system to alert to the patient condition. For each variable, if some event is “open”, the event type (1 or 2) is checked. Then, for the event type = 1, it calculates the time between the system date and the starting date. If the time in minutes is between 10 and 20 the label will be yellow. If this event is open more than 20 minutes the label will be red. In the cases of event type = 2 the label will be always red whatever the time it is started. In the other cases the label will be green.



HOUR	Heart Rate CE Number	O2 CE Number	Blood Press. CE Number	Diurese CE Number	Temperature CE Number	HR CE Minutes	O2 CE Minutes	BP CE Minutes	DIU CE Minutes	TMP CE Minutes
24										
23										
22										
21										
20										
19										
18										
17										
16										
15										
14										
13										
12										
11										
10					1					311:53
9										
8										
7										
6										
5					1					31:59
4					1					120:00

Figure V.3 - Critical Events grid.

The warning system is composed of five boxes, one for each category (z) and has three different colours. The box colour is refreshed every minute and is defined according to the event type and duration. The following procedure shows the rule used:

```
IF EVENT(Z).TYPE = 2 THEN
          EVENT(Z).BOXCOLOR = RED
    ELSE IF EVENT(Z).TYPE = 1 THEN
          IF EVENT_DURATION >= 30 THEN
                EVENT(Z).BOXCOLOR = RED
          ELSEIF EVENT_DURATION >= 10 THEN
                EVENT(Z).BOXCOLOR = YELLOW
          ELSE
                EVENT(Z).BOXCOLOR = GREEN
          END IF
    ELSEIF EVENT(Z).TYPE = 0 THEN
          EVENT(Z).BOXCOLOR = GREEN
    END IF
```

The charts present a new way of tracking the Critical Events. The user (doctor / nurse) can anywhere and anytime consult the evolution of a patient with regard to Critical Events. Users may view this information by minute, hour and day. When minutes are selected, it presents the evolution of the values in last 25 minutes using to the effect a continuous line graph. Figure V.4 shows an overview of a chart for the minutes. In this figure it is possible to observe that the patient is having a critical event (SpO2) from 46 last minutes.

The second and third types of analyses are similar. In figure V.5 is possible to observe that the information by hour aspect is presented in a different manner than information by minute. In this chart we use a bar chart that accumulates the minutes of the critical events during a day. Figure VI.5 shows a set of events (duration and type) verified from a patient in the last 12 hours. For example in the sixth hour the patient had a critical event (type 2) to the temperature during 94 minutes. The third view (day) presents the sum of critical events by day. All the graphs are grouped by category (BP, SpO2, Temperature, Urine Output, and Heart Rate) and event type (1 or 2).

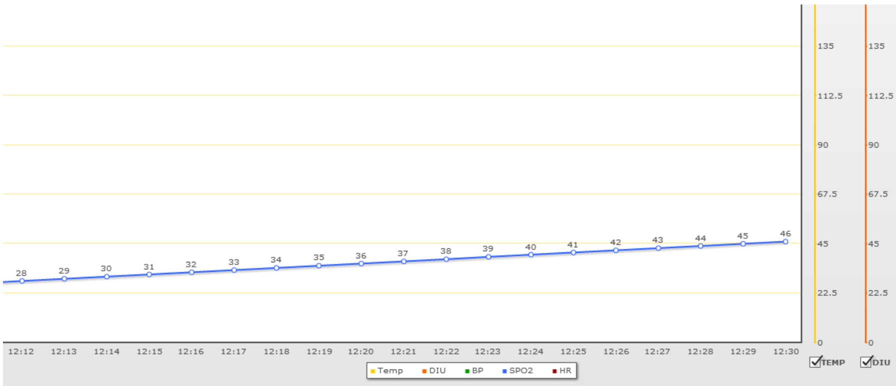


Figure V.4 - Critical Events Chart by minute.

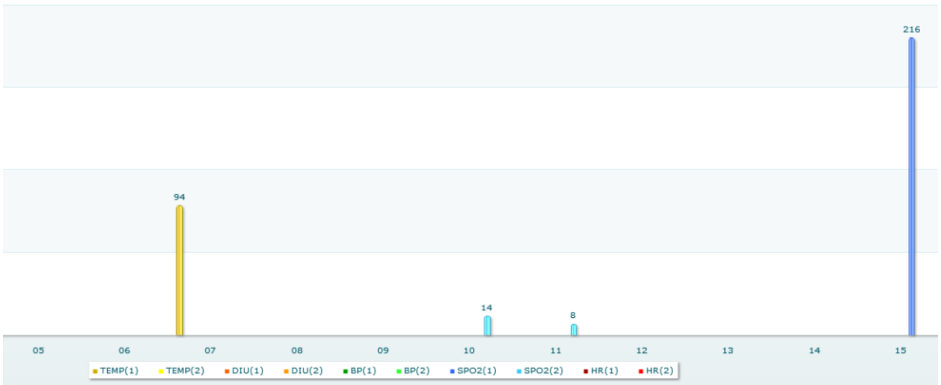


Figure V.5 - Critical Events chart by hour.

6.2. Tracking system TAM assessment

The Critical Events tracking system was evaluated in questionnaire of INTCare system. Even though the questionnaire has 91 questions only six are concerned with CE System. This questionnaire was answered by 13 nurses (33% of ICU nursing staff). The questions were filled using Likert Scale (Johns, 2010). The considered levels were the following:

1) Not satisfies/in complete disagreement (< 20% of cases);
2) Satisfies a bit/in some level of disagreement (20-40%);
3) Satisfies/under some level of agreement (40-60%);

4) Satisfies a lot/strongly agreement (60-80%);
5) Satisfies completely/full agreement (> 80%).

Table V.4 presents the question and respective construct of TAM: perceived usefulness (PU), perceived ease of use (PEOU), behavioural intention (BI) and use behaviour (UB).

Table V.4 – Answer VS TAM Construct

#	*Answer*	*PU*	*PEOU*	*BI*	*UB*
1	Potentiates a proactive action to the professionals	X	X	X	X
2	Allows the tasks to be performed with greater precision	X	X	-	-
3	Allow an easy operation through touch interface at the head of the beds	X	X	X	X
4	Critical Events - Utility of the System	X	-	X	X
5	Critical Events - Utility of warning system	X	X	-	X
6	Critical Events System - Interface	-	X	-	-

Having in account the answers received the results were grouped. Table V.5 presents the average (avg), standard deviation (Stdev), minimum (min), maximum (max) and mode of the values collected by answer and TAM construct.

Table V.5 – Analysis of the results collected

#	*AVG*	*Stdev*	*MIN*	*MAX*	*MODE*
1	3,30769231	0,60569291	2	4	3
2	3,15384615	0,8634594	2	5	3
3	3,07692308	0,91664425	2	5	3
4	3,61538462	0,48650426	3	4	4
5	3,61538462	0,48650426	3	4	4
6	3,61538462	0,62492603	3	5	4

	AVG	*Stdev*	*Min*	*Max*	*Mode*
PU	3,35385	0,715169	2	5	3
PEOU	3,28846	0,753061	2	5	3
IC	3,19231	0,727952	2	5	3
CMPU	3,19231	0,686676	2	5	4

Analysing Table V.5 it is possible to observe that in general the results are satisfactory. All questions present positive results and the level of accordance by question / user it is high. The questions associated in particular to the CE system were which present better results, having an average of 3,65 and a mode of 4.

At level of the constructs it is obvious the importance that the CE tracking system has to the decision making process, being the perceived usefulness the better construct. In all the cases the answers varies from two and five. This value signifies that nobody is totally in disagreement with the system.

7. Conclusions & Future Work

This new approach represents a set of innovations in the way of the data are collected and used to support the decision making process. It is now possible to acquire, process, and present knowledge automatically and in real-time using online-learning, anywhere and anytime.

Aiming at promoting proactive actions with the patient this Pervasive Intelligent System is able to determinate and evaluates the number and duration of critical events patients. The doctors in ICU of Centro Hospital do Porto now have some new information that helps them to have a better understanding of the patient's condition.

The doctors can see, for each patient and in real-time the number of events by hour and the time in which the patient was in a critical event. In addition, they possess a traffic light system (green, yellow, red) to alert / show the present situation for each event.

The system presents its results in three different ways: the first one presents all events in a based the second shows the evolution of the events in a chart and the third presents results from the prediction models for organ failure and final outcome and also shows the number of critical events. In order to assess the tracking system, a questionnaire based in the TAM methodology was performed. The results attained were very satisfactory and motivate futures developments in this area. The tracking system interface is the aspect which reaps greater acceptance by the ICU users. At the same time they recognize the importance of this type of system to help the decision.

Due to the success of this implementation, in the future, further adverse events types will be explored and added to the system. A set of new data mining models will be addressed in order to explore other targets.

Acknowledgements

This work is supported by FEDER through Operational Program for Competitiveness Factors – COMPETE and by national funds though FCT – Fundação para a Ciência e Tecnologia in the scope of the project: FCOMP-01-0124-FEDER-022674.

The authors would like to thank FCT (Foundation of Science and Technology, Portugal) for the financial support through the contract PTDC/EIA/72819/ 2006 (INTCare) and PTDC/EEI-SII/1302/2012 (INTCare II). The work of Filipe Portela was supported by the grant SFRH/BD/70156/2010 from FCT.

References

Bricon-Souf, N., & Newman, C. R. (2007). Context awareness in health care: A review. [10.1016/j.ijmedinf.2006.01.003]. International Journal of Medical Informatics, 76(1), 2-12.

Chooprayoon, V., & Fung, C. C. (2010). TECTAM: An Approach to Study Technology Acceptance Model (TAM) in Gaining Knowledge on the Adoption and Use of E-Commerce/E-Business Technology among Small and Medium Enterprises in Thailand.

Filipe Portela, Jorge Aguiar, Manuel Filipe Santos, Álvaro Silva, & Fernado Rua. (2013). Pervasive Intelligent Decision Support System - Technology Acceptance in Intensive Care Units. In Springer (Ed.), Advances in Intelligent Systems and Computing: Springer.

Filipe Portela, Pedro Gago, Manuel Filipe Santos, Álvaro Silva, & Rua, F. (2012). Intelligent and Real-time Data Acquisition and Evaluation to Determine Critical Events in Intensive Medicine. Paper presented at the HCist'2012 - International Conference on Health and Social Care Information Systems and Technologies, Portugal.

Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention and behaviour: An introduction to theory and research: Addison-Wesley.

Gago, P., Santos, M. F., Silva, Á., Cortez, P., Neves, J., & Gomes, L. (2006). INTCare: a knowledge discovery based intelligent decision support system for intensive care medicine. Journal of Decision Systems.

Hooda, J. S., Dogdu, E., & Sunderraman, R. (2004, 2004). Health Level-7 compliant clinical patient records system.

Johns, R. (2010). Likert Items and Scales. Survey Question Bank: Methods Fact Sheet, 1.

Kaur, M., Pawar, M., Kohli, J. K., & Mishra, S. (2008). Critical events in intensive care unit. Indian Journal of Critical Care Medicine: Peer-reviewed, Official Publication of Indian Society of Critical Care Medicine, 12(1), 28.

Keegan, M. T., Gajic, O., & Afessa, B. (2011). Severity of illness scoring systems in the intensive care unit. Critical care medicine, 39(1), 163.

Mador, R. L., & Shaw, N. T. (2009). The impact of a Critical Care Information System (CCIS) on time spent charting and in direct patient care by staff in the ICU: a review of the literature. International Journal of Medical Informatics, 78(7), 435-445.

Mikkonen, M., Va¨ yrynen, S., Ikonen, V., & Heikkila, M. O. (2002). User and concept studies as tools in developing mobile communication services for the elderly. Personal and Ubiquitous Computing, 6(2), 113-124.

Portela, F., Gago, P., Santos, M. F., Silva, A., Rua, F., Machado, J., et al. (2011). Knowledge Discovery for Pervasive and Real-Time Intelligent Decision Support in Intensive Care Medicine. Paper presented at the KMIS 2011- International Conference on Knowledge Management and Information Sharing.

Portela, F., Santos, M. F., Gago, P., Silva, Á., Rua, F., Abelha, A., et al. (2011). Enabling real-time intelligent decision support in intensive care.

Portela, F., Santos, M. F., Silva, Á., Machado, J., & Abelha, A. (2011). Enabling a Pervasive approach for Intelligent Decision Support in Critical Health Care. Paper presented at the HCist 2011 – International Workshop on Health and Social Care Information Systems and Technologies.

Portela, F., Santos, M. F., & Vilas-Boas, M. (2012). A Pervasive Approach to a Real-Time Intelligent Decision Support System in Intensive Medicine. In Springer (Ed.), Communications in Computer and Information Science (Vol. 0272, pp. 14).

Rothschild, J. M., Landrigan, C. P., Cronin, J. W., Kaushal, R., Lockley, S. W., Burdick, E., et al. (2005). The Critical Care Safety Study: The incidence and nature of adverse events and serious medical errors in intensive care*. Critical care medicine, 33(8), 1694.

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., & Neves, J. (2011). INTCARE - Multi-agent approach for real-time Intelligent Decision Support in Intensive Medicine. Paper presented at the 3rd International Conference on Agents and Artificial Intelligence (ICAART), Rome, Italy.

Satyanarayanan, M. (2002). Pervasive computing: vision and challenges. Personal Communications, IEEE, 8(4), 10-17.

Silva, Á., Cortez, P., Santos, M. F., Gomes, L., & Neves, J. (2008). Rating organ failure via adverse events using data mining in the intensive care unit. Artificial Intelligence in Medicine, 43(3), 179-193.

Varshney, U. (2007). Pervasive healthcare and wireless health monitoring. Mobile Networks and Applications, 12(2), 113-127.

Varshney, U. (2009). Pervasive Healthcare Computing: EMR/EHR, Wireless and Health Monitoring: Springer-Verlag New York Inc.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. Decision Sciences, 39(2), 273-315.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management science, 186-204.

ARTICLE A6

Objectives

To develop and test pervasive models in Intensive Care

i. To prototype an IDSS;
ii. To test the models developed;
iii. To include pervasive capabilities in the IDSS;

Intensive-care medicine score systems serve to quantify the severity of diseases and to characterize patient groups on the basis of objective criteria (Keegan, et al., 2011). The scores describe patient severity by adding up points. As main objectives, they: assess the prognosis; establish the amount of treatment required; provide information on the prognosis of patients; indicate the efficacy of therapeutic interventions; and serve for the stratification in clinical studies and workload (Schusterschitz & Joannidis, 2007), helping doctors in the decision process. In the ICU a high number of scoring systems are used that can be grouped in: diseases, patient and universally mortality prediction. The work developed was only focused on the scores that have variables that are used by Decision Support System, i.e., SAPS, SOFA and Glasgow. Those limitations dictated the development of a completely new scoring system to collect the data and to score the measures in real-time. A lot of research has been done to understand how it would be possible to acquire and prepare the necessary data. An intelligent agent based approach has been followed to perform crucial tasks such as the automatic data acquisition and the online scoring.

Results

This paper presents an intelligent scoring system to feed DM models as a way to predict the organ failure and the outcome of the ICU patients. A new concept for scores visualization was introduced, based in an hourly and continuous observation of the scores' results. The results are presented both in a grid of values and in a chart format. An intelligent scoring system has been presented to support the decisions taken in the ICU environment and, at the same time, improve the patient results. This approach makes it possibility to provide a set of scores calculated / updated in real time. The scoring system proposed processes automatically the scores and adapt the results according to the new values collected, generating new knowledge.

INTELLIGENT DATA ACQUISITION AND SCORING SYSTEM FOR INTENSIVE MEDICINE

Filipe Portela[1], Manuel Filipe Santos, José Machado[2], Álvaro Silva[3], Fernando Rua[3], António Abelha[2]

[1] Centro Algoritmi, University of Minho, Guimarães, Portugal
{cfp, mfs}@dsi.uminho.pt
[2] CCTC, University of Minho, Braga, Portugal
{jmac, abelha}@di.uminho.pt
[3] Serviço de Cuidados Intensivos, Centro Hospitalar do Porto, Hospital Santo António, Portugal
moreirasilva@clix.pt
fernandorua.sci@hgsa.min-saude.pt

Abstract. In a critical area as is Intensive Medicine, the existence of systems to support the clinical decision is mandatory. These systems should ensure a set of data to evaluate medical scores like is SAPS, SOFA and GLASGOW. The value of these scores gives the doctors the ability to understand the real condition of the patient and provides a mean to improve their decisions in order to choose the best therapy for the patient. Unfortunately, almost all of the required data to obtain these scores are recorded on paper and rarely are stored electronically. Doctors recognize this as an important limitation in the Intensive Care Units. This paper presents an intelligent system to obtain the data, calculate the scores and disseminate the results in an online, automatic, continuous and pervasive way. The major features of the system are detailed and discussed. A preliminary assessment of the system is also provided.

Keywords: Intensive Medicine Scores, Real-time, Pervasive System, Scoring System, SAPS, SOFA, GLASGOW

1. Introduction

During the last three decades, several physiological-based prognostic models have emerged (Keegan, et al., 2011). In 1980, the severity scoring systems were introduced in the Intensive Care Units (ICUs) (Schusterschitz & Joannidis, 2007). Intensive-care medicine score systems serve to quantify the severity of diseases and to characterize patient groups on the basis of objective criteria (Keegan, et al., 2011). The scores describe patient severity by adding up points. As main objectives, they: assess the prognosis; establish the amount of treatment required; provide information on the prognosis of patients; indicate the efficacy of therapeutic interventions; and serve for the stratification in clinical studies and workload (Schusterschitz & Joannidis, 2007), helping doctors in the decision process. In the ICU a high number of scoring systems are used that can be grouped in: diseases, patient and universally mortality prediction.

This work has been developed in the context of INTCare research project (Gago, et al., 2006), an intelligent decision support system that makes use of online learning to predict the organ failure and the outcome of the ICU patients. The work developed was only focused on the scores that have variables that are used by Decision Support System, i.e., SAPS, SOFA and Glasgow. In the context of INTCare, a new set of data had to be acquired in order to induce Data Mining (DM) models able to predict the organ failure and the outcome of the patients. These scores rarely were stored in an electronic way. As a normal procedure, the nurses / doctors calculate these scores by consulting the patient's bed side monitors, interpreting the lab results and reading admission documents in an offline mode. The existence of multiple data sources for each organic system (Strand & Flaatten, 2008) difficult the evaluation of the variables by the human and, consequently, can delay / interfere negatively in the decision making process.

For the development of INTCare system this situation was unaffordable. The inexistence of all data in an electronic mode and in real-time made impossible the construction of prediction models. Those limitations dictated the development of a completely new scoring system to collect the data and to score the measures in real-time. A lot of research has been done to understand how it would be possible to acquire and prepare the necessary data. An intelligent agent based approach has been followed to perform crucial tasks such as the automatic data acquisition and the online scoring. The main goal of the project was the concentration of the decision support tasks in a single platform.

According to Keegan (Keegan, et al., 2011) there are several limitations inherent to the ICU prognostic models: i) errors in data collection and data entry, flaws in development and validation of the models; ii) variables used to make predictions may not be easily measured; and some laboratory values may not be routinely obtained. The lack of standardization in obtaining values leads to missing data and to the fault of important prognostic variables, compromising the performance of the models (Metnitz, et al., 2005). The empirical experience accumulated in this field of acting allowed the development of a system able to overcome the limitations described above. This system has been deployed in the ICU of the Centro Hospitalar do Porto (CHP), Porto, Portugal.

This paper presents an intelligent scoring system to feed DM models as a way to predict the organ failure and the outcome of the ICU patients. Chapter one introduces the problematic and the subject, the second chapter makes an overview of the situation and explains the main concepts. Chapter three present the data manipulation process. Chapters four and five will present the scoring system and the results achieved. The preliminary results of a study on the technological acceptance are presented. Finally, some conclusions will be done about the work and the future work.

2. Background

It is common sense that in the ICU a set of clinical measures should be used regularly. Many of the scoring processes are executed manually and require a high level of human intervention. Currently there are systems able to collect all data produced during the day and execute some scoring. An example of this is the system presented by Shabot (Shabot, 2005) where "Three measures of severity of illness are automatically calculated for each adult ICU patient on admission and again daily". When the scores are calculated manually the value considered are not the worst among the measured values but the worst value registered by someone. This configures a major limitation. For a more accurate decision it is desirable to develop a system able to calculate the scores automatically according to the data captured from the patients along the day in real-time.

2.1. Intensive Care Scores

Scores are integrated in the diagnosis-related groups (Brenck, Hartmann, Mogk, & Junger, 2008) and can be used, for example, to predict the outcome (J. L. Vincent & Bruzzi de Carvalho, 2010). In the ICU of CHP, the most used scores are: SOFA, SAPS II and Glasgow.

Sepsis-related Organ Failure Assessment (SOFA) is used to daily score, as objectively as possible, the degree of organ dysfunction/failure of a patient (J. L. Vincent, et al., 1996). SOFA considers the worst value occurred along the day for the calculation of a score of 0 (no failure) to 4 (severe failure) for each organ: respiratory, renal, cardiovascular, neurologic, coagulation and hepatic (Strand & Flaatten, 2008).

Simplified Acute Physiology Score II (SAPS II) is an evolution of SAPS and provides an estimation of the risk of death without having to specify a primary diagnosis. SAPS II scores are converted into a probability of hospital mortality (Le Gall, et al., 1993) making use of logistic regression analysis (Strand & Flaatten, 2008). It uses twelve physiologic variables more age, admission type and the presence of metastatic or haematological cancer or AIDS (Strand & Flaatten, 2008).

More recently SAPS III has been developed. SAPS III includes a set of new variables, prefacing a total of 20 variables (Metnitz, et al., 2005): socio-demographics, chronic conditions, diagnostic information, physiological derangement at ICU admission, number and severity of organ dysfunctions, length of ICU and hospital stay, and vital status at ICU and hospital discharge. SAPS III is a model that was developed to assess severity of illness and to predict vital status at hospital discharge based on ICU admission data (Moreno, et al., 2005).

Glasgow Coma Score (GCS) (Jones, 1979) describes the patient's level of consciousness. GCS is scored between 3 and 15, where 3 is the worst value, and 15 the best. This score can't be automatically calculated because it requires human observation. It can be calculated several times along the day and is composed by three parameters: best eye response, best verbal response and best motor response.

2.2. INTCare System

INTCare (Gago, et al., 2006) is a research project whose main goal is to develop an Intelligent Decision Support System to, automatically and in real-time, predict the organ failure and patient outcome by means of data mining models (Vilas-Boas, et al., 2010a). INTCare system is divided in fours subsystem: data acquisition, knowledge management, inference and interface (Filipe Portela, Gago, et al., 2011). The DM models are fed from several data sources including ICU scores, in this case the SOFA values and some data from SAPS. The results described in this paper were obtained through the INTCare system. INTCare has been implemented in terms of intelligent agents able to react and / or reasoning that work together to obtain the ICU scores autonomously (Manuel Filipe. Santos, et al., 2011; Wooldridge, 1999).

3. Data Extraction, Transformation and Loading

Extraction, Transformation and Loading (ETL) is the process responsible for the extraction of data from several sources, their cleansing, customization and insertion into a data warehouse (Vassiliadis, Simitsis, & Skiadopoulos, 2002). For scoring purposes, the following systems and data were considered:

- Bedside Monitors - vital signs (VS);
- Electronic Health Record (EHR) - patient admission values;
- Electronic Nursing Record (ENR) – Hourly values (Fluid balance, Glasgow);
- Drugs System (DS) – therapeutic plan;
- Laboratory Results (LR) – Blood and blood gas exam results;

A set of agents of the INTCare data acquisition subsystem (Filipe Portela, Gago, et al., 2011) are in charge of the tasks associated to ETL (Manuel Filipe. Santos, et al., 2011). Figure VI.1 illustrates the overall ETL process. The data acquisition system is composed by a set of agents in charge of collecting data from several data sources automatically (bedside monitors, drugs system, EHR, laboratory and ENR). After this, the data is corrected and stored in order to be used by the scoring system. Finally, the data collected is used to calculate the ICU measures. This process is explained in detail in the next lines making use of the agent's paradigm.

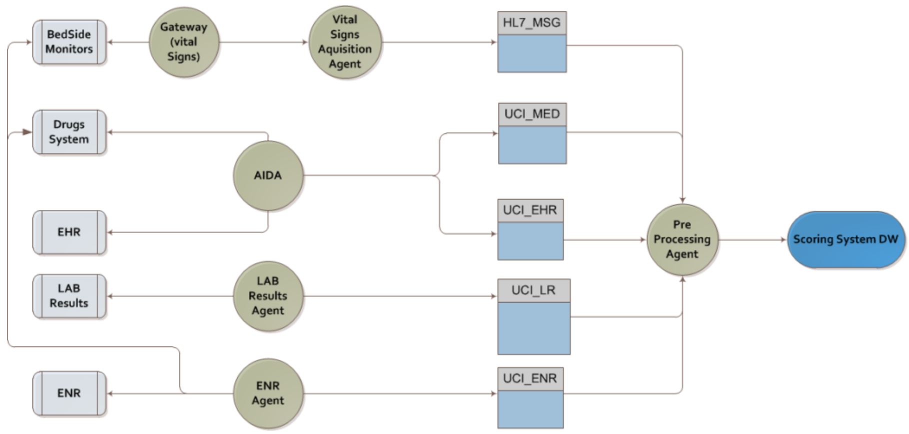


Figure VI.1 – ETL Process.

3.1. Extraction Process

Each score uses a set of different variables however, some of them are common. Table VI.1 presents the variables collected in real-time. For each variable indicates its correspondent data source and the way it is acquired (automatically or manually).

Table VI.1 – Scores Variables data source and acquisition type

Variable	Score	Data Source	Acquisition
Acute infection	SAPS III	EHR	Manual
Admission Glasgow Score	Glasgow	EHR	Automatic
Age	SAPS II	EHR	Automatic
Anatomical surgery site	SAPS III	EHR	Manual
Bicarbonate	SAPS II	LR	Automatic
Bilirubin	SOFA, SAPS II, SAPS III	LR	Automatic
Chronic diseases	SAPS II	EHR	Automatic
Co-Morbidities	SAPS III	EHR	Automatic
Creatinine	SOFA, SAPS II, SAPS III	LR	Automatic
Eye response	Glasgow	ENR	Manual (hour)
Glasgow coma scale	SAPS II, SAPS III	EHR	Automatic
Glasgow coma scale	SOFA	ENR	Manual (hour)
Heart rate	SAPS II, SAPS III	VS | ENR	Automatic
Hydrogen ion	SAPS III	LR	Automatic
Intra-hospital location	SAPS III	EHR	Automatic
Length of stay	SAPS III	EHR	Automatic
Leukocytes	SAPS III	LR	Automatic
Mean Arterial Pressure	SOFA	VS | ENR	Automatic
Mechanical ventilation	SAPS II	ENR	Automatic

Motor response	Glasgow	ENR	Manual (hour)
PaO_2/FiO_2	SOFA, SAPS II, SAPS III	LR	Automatic
Platelets	SOFA, SAPS III	LR	Automatic
Potassium	SAPS II	LR	Automatic
Reason(s) for admission	SAPS III	EHR	Automatic
Serum Urea or BUN	SAPS II	LR	Automatic
Sodium	SAPS II	LR	Automatic
Surgical status	SAPS III	EHR	Automatic
Systolic BP	SAPS II, SAPS III	VS | ENR	Automatic
Temperature	SAPS II, SAPS III	VS | ENR	Automatic
Type of admission	SAPS II, SAPS III	EHR	Automatic
Urine Output	SAPS II, SAPS III	ENR	Manual (hour)
Use of major therapeutic	SAPS III	EHR	Manual
Vasopressors	SOFA	DS | ENR	Automatic
Verbal response	Glasgow	ENR	Manual (hour)
White Blood Cell	SAPS II	LR	Automatic

Next, each agent responsible to collect data from the data sources will be explained in terms of its actions and variables instantiated.

Gateway (a_{gat}) operates in real-time and is responsible for capturing the vital signs data from bedside monitors. These data are packed into Health Level Seven (HL7) (Hooda, Dogdu, & Sunderraman, 2004) messages and sent to the Vital Signs Acquisition Agent. Gateway collects information, in average, once at each eight minutes and restarts at each hour to ensure that no communications failures compromise the system.

a_{gat} { Blood Pressure (systolic, mean), temperature, heart rate}

Vital Signs Acquisition (a_{vsa}) don't collects any data, only is used to, in real-time, extract information blocks, splitting the HL7 message and storing them into the database.

ENR Agent (a_{enr}) is associated to the Electronic Nursing Record. It was developed to allow the introduction of ICU information in an electronic mode, being responsible to capture the clinical data from the doctors and nurses (ICU Staff) (Gago, et al., 2006).

a_{enr} { Urine Output, Temperature, Eye Response, Glasgow Coma Scale, Mechanical Ventilation, Motor Response, Verbal Response, Blood Pressure (Systolic, Mean), Temperature, Heart Rate}

LR (a_{lr}) is responsible for capturing the clinical data from the lab results, i.e. blood and blood gas exams. The ENR agent requests to (a_{lr}), every five minutes, new results from the laboratory. This agent verifies if there are new results from a patient and, if so, it stores them into the database.

a_{lr}{ Bicarbonate, Bilirubin, Creatinine, Hydrogen ion, Leukocytes, PaO2, FiO2, Platelets, Potassium, Serum Urea, BUN, Sodium, Vasopressors, White Blood Cell}

AIDA (a_{ada}) is an agency to archive and to disseminate medical exams and results, earlier implemented at the hospital(Abelha, et al., 2003). In this case, it supplies the patient admission data.

a_{ada} {Acute infection, Admission Glasgow Score (Hospital and Service), Age, Anatomical surgery site, Chronic diseases, Co-Morbidities, Intra-hospital location, Length of stay, Reason(s) for admission, Surgical status, Type of admission, Use of major therapeutic}

Pre-Processing (a_{pp}) agent is responsible for the correct linking of all the values in order to create a valid medical record for each patient (Gago, et al., 2006). It is in charge of solving some data acquisition problems (M.F. Santos, F. Portela, et al., 2009a) and prepare the data to the scoring system.

3.2. Transforming Process

This process is the most important for the scoring system. In this stage all data collected will be processed, i.e., validated and transformed according to the scoring scales. The process is totally ensured by the pre-processing agent and requires a correct acquisition of the data. The vital signs data, after been collected, need to be validated, due to the errors that normally occur in ICU. All data collected will be automatically validated (F. Portela et al., 2011b) according to the ranges defined in ICU (Table VI.2) Those values are also subject of a manual validation after the automatic validation. If some values are out of range, they can be corrected by the nurses. The data validation process is quite simple: every time a value is collected by the agent a trigger is executed (1).

Table VI.2 – ICU Data ranges

Vital Sign	*Min*	*Max*
Blood Pressure (BP)	0	300
Temperature (Temp)	35	45
Respiratory Rate (RR)	0	40
Heart Rate (HR)	0	250

```
If value is null
delete row;
If value >= min and value <= max
move row_data to table "real_data"
set valid_data = true
```
(1)

After having all data validated, a set of procedures / functions or triggers are executed to prepare the data to be used in the scores.

The next process is a typical process of transformation and cataloguing of the variables according to the scores table. Table VI.3 is an example of a scoring table that is stored in the database to assist in the transforming process. This table contains the identification of the score, the variable measured, the possible values, the minimum and maximum values allowed for this variable and the points associated to each possible value. For each value collected a new analysis is done. The value will be verified and catalogued according to its importance / significance. The respective score result will be assigned after querying the database. In the case of numeric variables, the value collected is evaluated according to the min and max defined to each possibility. For example, in the case of bilirubin, three different types of measures are used, one for SAPSII, another one for SAPSIII and a last one for SOFA. For each case, the scale (min and max) varies and, according to the score in study the respective point will be associated to the value collected. If the variable collected is based in text type the scoring process will be based on the expression (definition). For example, in the case of Glasgow, the eyes will be evaluated taking into account the expression / reaction (absent to pain, to speech, spontaneous) stored in the database and the correct point (1, 2, 3, 4) will be assigned, according to the expression collected.

Table VI.3 – Scores table (example)

Score ID	*Variable ID*	*Definition*	*Min*	*Max*	*Point*
GLASGOW	Eye	Absent			1
GLASGOW	Eye	To pain			2
GLASGOW	Eye	To speech			3
GLASGOW	Eye	Spontaneous			4
SAPSII	Admission Type	Schedule Chirurgic			0
SAPSII	Admission Type	No chirurgic			6
SAPSII	Admission Type	Urgency Chirurgic			8
SAPSII	AGE	< 40	0	39	0
SAPSII	AGE	[40 ; 59]	40	59	7
SAPSII	AGE	[60 ; 69]	60	69	12
SAPSII	AGE	[70 ; 74]	70	74	15
SAPSII	AGE	[75 ; 79]	75	79	16
SAPSII	AGE	> 80	80	120	18
SAPSII	Bilirubin	<4	0	3,9	0
SAPSII	Bilirubin	4 a 5,9	4	5,9	4
SAPSII	Bilirubin	>= 6	6	9999	8
SAPSIII	Admission Time	<14	0	13	0
SAPSIII	Admission Time	[14 ; 27]	14	27	6
SAPSIII	Admission Time	>=28	28	9999	7
SAPSIII	Bilirubin	<2	0	1,9	0
SAPSIII	Bilirubin	[2 ; 5,9]	2	5,9	4

SAPSIII	Bilirubin	>= 6	6	9999	8
SOFA	Bilirubin	< 1,2	0	1,1	0
SOFA	Bilirubin	1,2-1,9	1,2	1,9	1
SOFA	Bilirubin	2,0-5,9	2	5,9	2
SOFA	Bilirubin	6,0-11,9	6	11,9	3
SOFA	Bilirubin	>=12	11,9	9999	4
SOFA	Creatinine	< 1,2 or < 110	0	1,1	0
SOFA	Creatinine	1,2-1,9 or 110-170	1,2	1,9	1
SOFA	Creatinine	2,0-3,4 or 171-299	2	3,4	2
SOFA	Creatinine	3,5-4,9 or 300-440	3,5	4,9	3

The main difference among SOFA, GLASGOW and SAPS scoring formula is the time / date used. In the case of SAPS, only is used the worst value collected along the first 24 hours. In the other cases, the score is obtained at the end of each day and the worst values of the day are used.

The next formulas (2-4) present some examples of the data processing. Each variable has different configurations according to the score id. An example of each case mentioned before (number and text) is presented.

δ – value collected
Ω – min score value
β – max score value
γ – scoring point
ρ – score id
Θ – data source variable id
Ɯ – score variable id
§ - variable definition

Δ – date (value collected)
φ – date (today)
μ – admission day + 1
θ – patient id
¥ – data source table
Є – score final table
Э – scores table

In case of number (example):
For each δ
Insert into Є (
select γ, φ, θ, Ɯ, ρ from ¥, Э
where δ >= Ω or δ <= β and Δ = φ and ρ = 'SOFA' and Ɯ = Θ); (2)
Select max(γ), θ, φ, Ɯ from Є where ρ = 'SOFA' and Ɯ = Θ
group by
θ, φ, Ɯ

In case of text (example):
For each δ
Insert into Є (
select γ, φ, θ, Ɯ, ρ from ¥, Э
where δ = Э and Δ = φ and ρ = 'GLASGOW' and Ɯ = Θ); (3)
Select max(γ), θ, φ, Ɯ from Є where ρ = 'GLASGOW' and Ɯ = Θ
group by
θ, φ, Ɯ

In case of SAPS (text example):
For each δ
Insert into Ͼ (
select γ, φ, θ, Ɯ, ρ from ¥, Ͽ (4)
where δ = Ͽ and Δ = μ and ρ = 'SAPSII' and Ɯ = Θ);
Select max(γ), θ, φ, Ɯ from Ͼ where ρ = 'SAPSII' and Ɯ = Θ
group by
θ, φ, Ɯ

In some cases it is necessary to do transformations before the scoring point is associated as is the case of the variable pao2 / fio2. In this case, the values are normally obtained separately and only then the calculation measure is applied: Pao2 / (Fio2 / 100). Another example is the SOFA cardiovascular, where two different types of measure are used: Mean Blood Pressure and Vasopressors. For this case, all variable possibilities will be evaluated and, in the final, only the worst score will be considered. SAPS II and SAPS III also have similar situations; in all the cases the same method is applied.

3.3. Loading Process

This is the last ETL process. Here, all transformed data are loaded into the data warehouse (DW). Then, all of those data are prepared to be interpreted by the scoring system, allowing for the calculation of the final scores of each measure: SOFA, SAPS II, SAPS III, and GLASGOW.

4. Scoring System

The Scoring System (SS) is integrated in the Electronic Nursing Record (ENR). The ICU staff can consult the results through this application. This application is also used for registering some values that require a human observation like is Glasgow and some SAPS parameters. The next figure (Figure VI.2) presents a print screen of a score's view for SAPS II. The other three scores' views are similar. Being this a touch-screen interface, the user can permute among the SCORES sliding the views.

The idea is to give to the ICU staff the better way for consulting the real condition of the patients whenever they want. It is possible to verify the results obtained for each parameter and also the parameters not evaluated until the moment. In the case of the SAPS, the result presented is the worst of the values obtained during the firsts 24 hours. In the other cases (SOFA and Glasgow) the result presented is the worst of the results along the day captured until the actual moment.

The user interface is simple and easy to understand. All the variables of the scale are disposed in a page where, at the left side, are the measure variables names (age, temperature, blood pressure) and at the right the worst absolute value collected for that variable. In the middle of the page are disposed all of the possible results for each score variable in agreement with the score scale. For example, in the first line of the SAPS II score (figure VI.2) a label for the age is positioned in the left side, in the middle are presented the six possibilities for the age (table VI.3) and at right the patient age. When a result is obtained the box with the correspondent result will be automatically highlighted in green (eg. 82 years = last box green). With this option the user can quickly understand what are the worst values for the patient and for a specific variable.

Figure VI.2 – SAPS II Score System.

To correct wrong values or add the values in fault is simple. Using the touch screen, the user only needs to click in the correct result and automatically the result is filled or actualized. Although the SS has an automatic saving system the final results require a manual validation. This option ensures that only correct values are stored and they are correctly associated to the patient. The results appear in the system moments after they are collected. Whenever a different result is recorded, the scores are refreshed according to the new data.

Taking into accounting the Table VI.2, it is possible to verify that many of the values are collected automatically (eg. values provided by the laboratory and the bedside monitors). However, in the case of GLASGOW and SAPS III, there is a set of variables that needs to be inserted manually (e.g.

urine output, Glasgow). Operation (5) corresponds to the calculation performed to obtain the final results by day and by score.

ρ – score id
Δ – date (value collected)
Θ – data source variable id
φ – date (today)
Ꝡ – score variable id
Ͼ – score final table
θ – patient id

For each score and day
finalScore(ρ, φ) = 0
For i= 1 to count(Ꝡ)
Score(i) = (Select max (γ) (5)
from Ͼ where ρ = 'SCORE' and Ꝡ = 'VariableID(i)' and Δ = φ
group by
θ, φ, Ꝡ)
finalScore(ρ, φ) = finalScore(ρ, φ) + Score(i)
next;

The final score is calculated and presented only when all variables are filled. In the case of the SAPS, the results obtained are also used to predict the death of the patient. SAPS scores are calculated making use of the formulas developed by the field researchers (Le Gall, et al., 1993; Metnitz, et al., 2005) and uses the SAPS II or SAPSS III final results. Each SAPS score uses a different formula to obtain the Predicted Death Rate (PDR). PDR is obtained using the equation 6:

PDR = e(Logit)/(1+e(Logit)) (6)

SAPS II Logit formula is:

Logit = -7,7631+0,0737*(SAPS II)+0,9971*ln((SAPSII)+1) (7)

In case of SAPSIII the formula is:

Logit = −32.6659 +ln(SAPS III +20.5958) ×7.3068 (8)

5. Results

A new concept for scores visualization was introduced, based in an hourly and continuous observation of the scores' results. The results are presented both in a grid of values and in a chart format. This is only possible due to a continuous and real-time execution of the data acquisition and data processing. The data processing performed for each hour is the same as that performed for the day (equation 5). The difference is in the value assigned to the φ parameter. If no data is collected for a particular hour, the last result obtained will be considered. This occurs often, principally in the cases where the laboratory results are used (e.g. Creatinine, bilirubin, pao2/fio2). The exams are normally performed in average two/three times a day. When a significant number of values exist for the same hour, as is the case of the vital signs, only the worst value of the hour, correctly collected, will be considered (9). In the hourly approach, the scoring system will be executed using the values obtained until that point.

ρ – score id	Δ – date (value collected)
Θ – data source variable id	φ – date (today)
Щ – score variable id	ƚ – date (hour of today)
Щn – score variable id (number)	𝔵 – hour (value collected)
θ – patient id	Ͼ – score final table
γ – scoring point	

```
for each score variable
select Щ, max(γ), φ, ƚ, 𝔵 from Ͼ
where Δ = φ and θ = 'patient id' and
ρ = 'SCORE'
and Щ = 'Score Variable'
group by Щ, ƚ, 𝔵
order by ƚ asc;
then →

hNow = 1
For kj = hNow To 𝔵
If φ is not null and kj = 𝔵 then
valueHour(kj, Щn) = max(γ)                    (9)
ElseIf kj > 0 Then
valueHour(kj, Щn) = valueHour(kj - 1, Щn)
End If
End If
Next
hNow = 𝔵 + 1
next;
```

A different chart is associated to each score. Each chart is composed by the variables that make up the score and the results obtained through the process presented before (9). In the case of SOFA, the chart has seven variables (Neurologic, Respiratory, Coagulation, Liver, Hepatic, Cardiovascular and Total). In figure VI.3 it is possible observe the interface of the system developed. For a better comprehension of the values, the scales considered are the same and range between 0 and 24. Only the Sofa total can have values within this range, the other variables (neuro, coag, resp, hepatic, renal, cardio) range from 0 to 4.

These charts are included in the ENR platform and can be consulted in two different ways: hourly (along the 24 hours) and by day (since the admission day). The user can also consult all variables in simultaneous (default) or can isolate a particular variable for a fine-grained study. In the same platform, the ICU staff can record, validate and consult the scores.

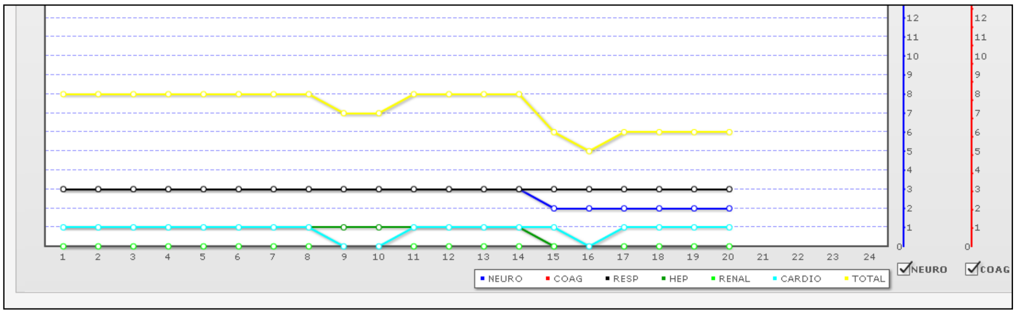


Figure VI.3 – SOFA score chart.

Since its inception in last March, the scoring system was used in more than 20 patients. In order to understand the level of technological acceptance of the system, a questionnaire with ten questions has been conceived. From the ten questions, five are dedicated to the scoring system. 15% of the ICU professionals have been asked to respond to the questionnaire. The evaluation scale was defined around 5 different levels of agreement:

1) In complete disagreement;
2) Disagrees;
3) Agrees;
4) Satisfactory agreement;
5) Complete agreement.

The questions were grouped into two different groups: the functional aspects and the technical aspects. The five questions concerning the scoring system are:

1. An efficient consulting of information for nursing decision support is allowed?
2. An efficient consulting of information for medical decision support is allowed?
3. A proactive performance of the professionals is enhanced by the system?
4. Is the access to the information, in terms of speed and availability, adequate to the needs?
5. Is the access to the system easy and secure?

The Table VI.4 presents the results obtained, in terms of the percentage of the answers for each question and each evaluation level.

Table VI.4 – Questionnaire result (%)

Query	1	2	3	4	5	Query	1	2	3	4	5
Functional characteristics						Technical characteristics					
1	0,00	0,00	0,00	16,67	83,33	3	0,00	0,00	0,00	66,67	33,33
2	0,00	0,00	0,00	0,00	100,00	4	0,00	0,00	0,00	33,33	66,67
						5	0,00	0,00	0,00	50,00	50,00

The results obtained showed that the ICU professionals are very comfortable with the new system. In general, the questions related with decision making support were scored to 4 or 5 points. This gives a good motivation to continue the work improving the entire system.

6. Conclusion and Future Work

An intelligent scoring system has been presented to support the decisions taken in the ICU environment and, at the same time, improve the patient results. This approach makes it possibility to provide a set of scores calculated / updated in real-time. The scoring system proposed processes automatically the scores and adapt the results according to the new values collected, generating new knowledge. The main gains in using this approach can be summarized as:

- The data acquisition, the scores calculation and the results are made in real-time;
- All values are considered - no missing values;
- The data is displayed in a new way – real-time charts to compare trends;
- Less human intervention in the scores calculation – less errors;
- The scores are available anywhere and anytime;
- Help decision making process through a continuous scores monitoring - a real-time calculation of the scores according the most recent patient results, allow a quick and better comprehension of the patient condition.

Further work includes the study of other scores (e.g. MEWS) and their impact in the DM models.

Acknowledgments

The authors would like to thank FCT (Foundation of Science and Technology, Portugal) for the financial support through the contract PTDC/EIA/72819/ 2006. The work of Filipe Portela was supported by the grant SFRH/BD/70156/2010 from FCT.

References

Abelha, A., Machado, M., Santos, M., Sollari, A., Rua, F., Paiva, M., et al. (2003). Agency for Archive, Integration and Diffusion of Medical Information. *Proceeding of AIA*.

Brenck, F., Hartmann, B., Mogk, M., & Junger, A. (2008). [Scoring systems for daily assessment in intensive care medicine. Overview, current possibilities and demands on new developments]. *Der Anaesthesist, 57*(2), 189.

Gago, P., Santos, M. F., Silva, Á., Cortez, P., Neves, J., & Gomes, L. (2006). INTCare: a knowledge discovery based intelligent decision support system for intensive care medicine. *Journal of Decision Systems*.

Hooda, J. S., Dogdu, E., & Sunderraman, R. (2004, 2004). *Health Level-7 compliant clinical patient records system*.

Jones, C. (1979). Glasgow coma scale. *AJN The American Journal of Nursing, 79*(9), 1551.

Keegan, M. T., Gajic, O., & Afessa, B. (2011). Severity of illness scoring systems in the intensive care unit. *Critical care medicine, 39*(1), 163.

Le Gall, J. R., Lemeshow, S., & Saulnier, F. (1993). A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA, 270*(24), 2957-2963.

Metnitz, P. G. H., Moreno, R. P., Almeida, E., Jordan, B., Bauer, P., Campos, R. A., et al. (2005). SAPS 3—From evaluation of the patient to evaluation of the intensive care unit. Part 1: Objectives, methods and cohort description. *Intensive care medicine, 31*(10), 1336-1344.

Moreno, R. P., Metnitz, P. G. H., Almeida, E., Jordan, B., Bauer, P., Campos, R. A., et al. (2005). SAPS 3—From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission. *Intensive care medicine, 31*(10), 1345-1355.

Portela, F., Gago, P., Santos, M. F., Silva, A., Rua, F., Machado, J., et al. (2011). *Knowledge Discovery for Pervasive and Real-Time Intelligent Decision Support in Intensive Care Medicine*. Paper presented at the *KMIS* 2011-International Conference on Knowledge Management and Information Sharing.

Portela, F., Santos, M. F., Gago, P., Silva, Á., Rua, F., Abelha, A., et al. (2011). *Enabling real-time intelligent decision support in intensive care.* Paper presented at the 25th European Simulation and Modelling Conference- ESM'2011, Guimarães, Portugal.

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., & Neves, J. (2009). Information Architecture for Intelligent Decision Support in Intensive Medicine. *8th International Conference on Applied Computer and Applied Computational Science (ACACOS '09), 8*(5), 810-819.

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., & Neves, J. (2011). *INTCARE - Multi-agent approach for real-time Intelligent Decision Support in Intensive Medicine.* Paper presented at the 3rd International Conference on Agents and Artificial Intelligence (*ICAART*), Rome, Italy.

Schusterschitz, N., & Joannidis, M. (2007). Predictive capacity of severity scoring systems in the ICU. *Contributions to nephrology, 156*(R), 92.

Shabot, M. M. (2005, 2005). *Automated data acquisition and scoring for JCAHO ICU core measures*.

Strand, K., & Flaatten, H. (2008). Severity scoring in the ICU: a review. *Acta Anaesthesiologica Scandinavica, 52*(4), 467-478.

Vassiliadis, P., Simitsis, A., & Skiadopoulos, S. (2002, 2002). *Conceptual modeling for ETL processes*.

Vilas-Boas, M., Santos, M. F., Portela, F., Silva, Á., & Rua, F. (2010). *Hourly prediction of organ failure and outcome in intensive care based on data mining techniques*. Paper presented at the 12th International Conference on Enterprise Information Systems.

Vincent, J. L., & Bruzzi de Carvalho, F. (2010). Severity of illness. *Semin Respir Crit Care Med, 31*(1), 031-038.

Vincent, J. L., Moreno, R., Takala, J., Willatts, S., De Mendonca, A., Bruining, H., et al. (1996). The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive care medicine, 22*(7), 707-710.

Wooldridge, M. (1999). Intelligent agents *Multiagent systems: a modern approach to distributed artificial intelligence* (pp. 27-77): MIT Press.

IV– Results Assessment

ARTICLE A7

Objectives

To develop and test pervasive models in Intensive Care;

i. To assess the results.

This paper will present the first questionnaires done after conclude the test phase, with the objective to understand the importance of the system to ICU and their professionals.

Results

The results obtained encourage the continuation the development and optimization of the solution and also to deep evaluation of all features, using TAM. After the new features were incorporated in the INTCare system, the benefits of them for ICU were notorious. The first benefit is related to the number of different information platforms supporting the ICU processes. The ICU professionals normally had to access to six different platforms – Bedside Monitors (BM), Nursing Record Paper Based (NRPB), AIDA, Medical Platform (Mplat), Nurses Platform (Nplat) and Pharmacy, to consult essential information to decision / job. All of the improvements were assessed in terms of their functionality and usability. A restrict and specialized number of users were asked. During this time two doctor and six nurses (~15 % of ICU professionals) explored the platform and answered the questionnaires. All of them were familiarized with the ICU Clinical systems, being a daily user of them. The use of a reduced number of professionals was a strategy, because during the test phase of the system is too important not demotivate the professionals with some possible failures. This objective, is easily achieved with a performing of lot of test near the ICU professionals which are more trained for direct patient care and understand the risks / complexity of this type of systems.

The users answered to a questionnaire in order to assess the technology. With the objective to give a global idea about the usefulness and ease of use of the system in superficial way, an easy and quick questionnaire was done. The study carried out on the technology acceptance (Table 3) corroborated the options made during the development of INTCare system. The system is now in use in the ICU of Centro Hospitalar do Porto, Porto, Portugal.

INTELLIGENT DECISION SUPPORT IN INTENSIVE CARE – TOWARDS TECHNOLOGY ACCEPTANCE

Filipe Portela[1], Manuel Filipe Santos[1], José Machado[2], António Abelha[2], José Neves[2],
Álvaro Silva[3], Fernando Rua[3]

[1]Algoritmi Center, University of Minho, Campus de Azurém, Guimarães, Portugal
E-mail: {cfp, mfs}@dsi.uminho.pt
[2]CCTC, University of Minho, Braga, Portugal
E-mail: {jmac, abelha, jneves}@di.uminho.pt
[3]Serviço de Cuidados Intensivos, Centro Hospitalar do Porto, Hospital Santo António, Porto, Portugal
moreirasilva@clix.pt , fernandorua.sci@hgsa.min-saude.pt

ABSTRACT Decision support technology acceptance is a critical factor in the success of the adoption this type of systems by the users. INTCARE is an intelligent decision support system for intensive care medicine. The main purpose of this system is to help the doctors and nurses making decisions more proactively based on the prediction of the organ failure and the outcome of the patients. To assure the adoption of INTCARE by the doctors and by the nurses, several requirements had taken into account: process dematerialization (information is now in electronic format); interoperability among the systems (the AIDA platform was used to interoperate with other information systems); on-line data acquisition and real-time processing (a set of software agents has been developed to accomplish these tasks). A technology acceptance model has been followed in the Intensive Care Unit (ICU) of Centro Hospitalar do Porto in order to assure the most perfect alignment between the functional and technical characteristics of INTCARE and the user expectations. Results showed that the ICU staff is permeable to the system. In general more than 90 % of the answers are scored with 4 or 5 points which gives a good motivation to continue the work.

KEYWORDS: INTCARE, Intelligent Decision Support Systems, Real-Time, Interoperability, Technology Acceptance, Intensive Care Medicine.

1. Introduction

The INTCare project, started in 2009 with the purpose to develop an Intelligent Decision Support system to predict the organ failure and patient outcome in intensive care medicine. Early results obtained with the data from EURICOS project, in an offline learning approach, motivated further developments (Álvaro Silva, et al., 2008). However, a huge challenge arisen to adapt the system to a real-time and online environment. The main challenge was related to the distributed nature of the information sources and to the paper based processes still in use in the Intensive Care Unit (ICU). In order to overcome these issues some requirements were defined.

The requirements defined to turn INTCare more suitable for the ICU environment, may be summarized as:

1. To implement an online data acquisition component;

2. To make available the lab results data in open format;
3. Allow for open access to prescriptions, interventions and therapeutics attitudes;
4. To dematerialize the nursing record;
5. To integrate the main systems used in ICU in a single platform;

To achieve these objectives a new system and platform was developed. This platform was based in a concept of having a multi data source in one application with the goal to reduce the time the doctors or nurses spent on documentation and consequently increase the time of patient care (Häyrinen, Saranto, & Nykänen, 2008; Vandijck, Labeau, & Blot, 2008). This new solution can integrate all data sources that are essential to work in the best interest of the patients. Like Brailer (Brailer, 2005) said without interoperability, Electronic Medical Record (EMR) adoption will further strengthen the information silos that exist in today's paper-based medical files, resulting in even greater proprietary control over health information and, with it, control over patients themselves. In addition, providing patient data electronically, online and in real-time, the information will be available anywhere and anytime.

Finally, decision support technologies like the INTCare system should be more framed in the user environment, assuring the user acceptance (Chooprayoon & Fung, 2010).

The Technology Acceptance Model (TAM) (Venkatesh & Bala, 2008; Venkatesh & Davis, 2000) is been applied to assure the adoption of INTCare by the ICU users (doctors and nurses). This paper will present the first questionnaires done after conclude the test phase, with the objective to understand the importance of the system to ICU and their professionals. The results obtained encourage the continuation the development and optimization of the solution and also to deep evaluation of all features, using TAM (Chooprayoon & Fung, 2010).

The following sections of the paper will be dedicated to: present the background knowledge and related work in section 2; section 3 introduces the departure point of the INTCare system; section 4 is focused in the improvements implemented; section 5 is dedicated to the electronic nursing record, an important component in the prosecution of the objectives; section 6 presents the results obtained so far in terms of the degree of interoperability, dematerialization of processes and the technology acceptance; and, finally, section 7 concludes the paper criticizing and points future directions.

2. Background and Related Work

2.1. INTCare

INTCare (Gago, et al., 2006) is a research project which has as main goal to develop an Intelligent Decision Support System (IDSS) that can, in real-time and in an online learning mode (Filipe Portela, Santos, et al., 2010), predict the outcome and organ system failure in a pervasive approach (Filipe Portela, Santos, et al., 2011). INTCare is in use in the Intensive Care Unit (ICU) of the Centro Hospitalar do Porto, Porto, Portugal. The main motivation to the developing of a new information system were the number of data in paper and information silos, which exist in ICU and, difficulties to obtain data to INTCare Data Mining models. These data were defined based in the results obtained in the past (Álvaro Silva, et al., 2008).

To the INTCare a set of agents were developed with aim to automatize some important tasks, avoiding the manual effort. The use of intelligent agents is fundamental to an implementation of a real-time and online information system. The agents are integrated in a Multi-Agent System (MAS), they communicate with each other and work together towards common goals, with a degree of reactivity and / or reasoning (Wooldridge, 1999).

2.2. Intensive Care

Normally, to Intensive Care Units (ICU) go patients in seriously weak conditions, where the organ failure is / can be a reality. This environment is considered critical because have some complex health care situations and the risk life is obvious. ICU is a data-rich environment (Lapinsky, Holt, Hallett, Abdolell, & Adhikari, 2008) and every hour a great amount of clinical data are generated and stored in paper format. This process difficult the data dissemination and data consult. Typically the information used is provided from a lot of different sources that can be electronic or manual. Other problem is the delay of data, i.e., the time difference between data acquisition and data availability.

At the beginning in this ICU the data was collected in offline and in an irregularly mode making the analysis of patient data and the search of past information a very hard and time-consuming work. The existing information systems were only used for information consultation and not to register patient data. For instance, the bedside monitors were only used to visualize the patient Vital Signs (VS), these weren't stored in a database or used by the information systems (Filipe Portela, Santos, et al., 2011). To develop a pervasive health care (PHC) system to ICU is important to have present the general requirements defined by Varshney(U. Varshney, 2007b).

2.3. AIDA

AIDA(Duarte, Portela, Santos, António, & José, 2011) is an Agency for Integration, diffusion and Archive of Medical Information. This platform is implemented in CHP and according Abelha (Abelha, Machado, Alves, & Neves, 2004), provides intelligent electronic workers, here called pro-active agents, and is in charge of tasks such as communicating with the heterogeneous systems, sending and receiving information, managing and saving the information and answering to information requests, with the necessary resources to their correct and in time accomplishment. Although AIDA was implemented in ICU, still there is a set of clinical applications which can't communicate between them (e.g. platform of nurses and medics). In these cases, normally, the information will be lost in the time. In other case, is too difficult accede to data in a quick and efficient way. The introduction of a new information system which, can integrate, using systems interoperability, all ICU data sources, can be the best solution.

2.4. Technology Acceptance Model

To understand if some technology is or not adequate to the environment and if the users are happy with the solution, is extremely necessary evaluate the success of the application. One of the models most used in this area is TAM. Since it development, in 1989, it suffered a set of improvements and changes of content. "TAM is adapted from the Theory of Reasoned Action (TRA) model which describes human behaviours in a specific situation" (Fishbein & Ajzen, 1975). The main purpose of TAM is to present an approach to study the effects of external variables towards people's internal beliefs, attitudes, and intentions (Chooprayoon & Fung, 2010), understanding, the ease of use (PEOU) and usefulness (PU) of technology. The goals of TAM can be achieved, for example, by using methodologies based in questionnaires. TAM (Venkatesh & Bala, 2008; Venkatesh & Davis, 2000) was the guide of the questions done and will be used, in the future, to do a most deep study of the application.

3. Global Overview

To develop INTCare a deep study, about the environment, were done. The necessities of innovation and features that need to be implemented were defined (Filipe Portela, Santos, et al., 2011). In this context was possible verify that:

Vital signs monitors were used to visualize the values;

- The lab results were available in a closed format;
- The prescriptions were registered in paper by the nurses;
- The patient data, ventilation, fluid balance and medical scores are registered and calculated manually and in paper;
- Some registries weren't done regularly;
- Many sources containing vital data were distributed by distinct silos.

Figures VII.1 depicts an overview of the data sources that daily are used by nurses and doctors in ICU. Besides the high number of data sources and the necessity to use different platforms to consult / store the information, the nursing record were still made in paper and can't be used by other platforms. Orange shapes represent the data sources, associated to the information that normally are requested. The other shapes represent this information. For example, in case of AIDA, ICU normally request information about Electronic Health Record (EHR), patient and diagnosis tests.

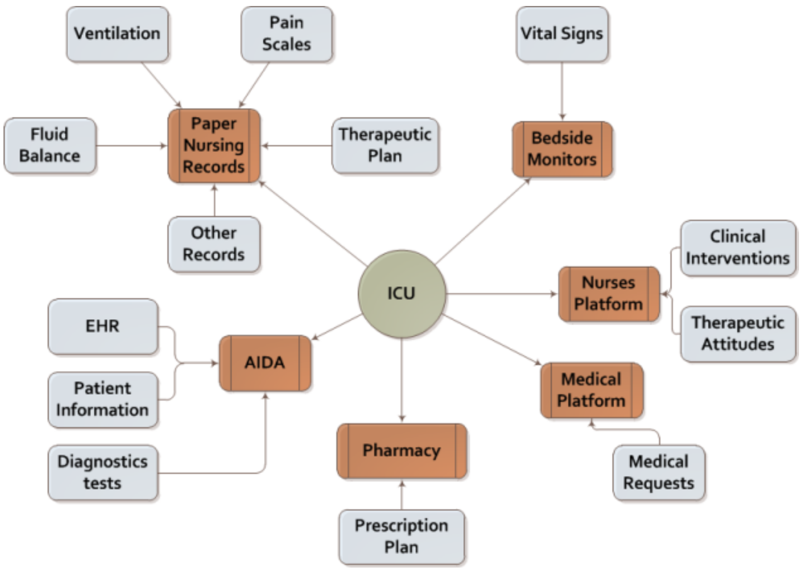


Figure VII.1 – ICU Data Sources.

4. Improvements Introduced

In order to resolve the limitations reported before, a set of modifications were introduced. These improvements will be briefly explained in the next sub-sections.

4.1. Data acquisition system

The first concern was about the vital signs, the possibility of use the values that were in patient bedside monitor and store them into a database. To resolve this, a gateway was implemented; it is connected to the vital signs monitors, reads the patient information and stores it on a database through the data acquisition agent. Due to the high number of data provided by the gateway, the interval of time defined for sampling was 8 minutes. After the process is concluded the data will be displayed in the ENR platform. During this process some important problems were resolved. The problem of patient identification and data out of range or badly collected were overcome. In both cases, the introduction of new tasks to the agent was the solution. The agent can verified who patient are in the bed (admitted in EHR) and can compare the values collected with the max and min normally allowed in the ICU (F. Portela, et al., 2011a).

After all values are correctly identified and validated, they will be, on moment, available in the Electronic Nursing Record platform (online). Now, instead be necessary see the monitors to consult the patient vital signs and then store them manually in a paper, the ICU professional only needs to confirm those values in the platform.

4.2. Lab results

Second, and considering that laboratory are an important source for ICU, an effort were made to have the lab results in an open format, making the results available for ICU immediately after the patient exams be concluded. This change gives the possibility to have the results in a comparative format during the patient stay in ICU.

The Lab Results are not under the nurse’s control in the ICU, because they are ordered by the doctors and executed by the labs (Filipe Portela, Gago, et al., 2011). Normally these analyses are done 1 to 3 times-a-day, however they can be done more times. Those analyses have different types, are executed by different services and at different hours. The integration and grouping of obtained results are done by ENR. Lab Results agent is responsible to receive and store the results in a table. All data acquired are available to be consulted in the ENR. Laboratory releases the results of fowling services: Microbiology, Haematology, Urgency exams, Serology, Immunology, Clinical Chemistry, Respiratory gas measurement

4.3. Open access to prescriptions

In this step, the objective was obtaining a partial access to data provided from the pharmaceutical information system. After some dealing was possible have an access to patient prescriptions. These prescription, were totally controlled by pharmacy, and always someone needed to consult the patient therapeutic plan had to open a too slow platform. The interaction between the pharmaceutical system and ENR is made by an agent. Periodically (10 in 10 minutes), the ENR agent sends a request to the pharmacy drugs system, and then, the data requested will be sent into a table. After the data is store, a pre-processing agent runs to prepare all data to ENR and stores them. This process involves a high number of processing tasks in real-time.

5. Electronic Nursing Record

Electronic Nursing Record (ENR) is a platform that was developed with the objective to receive all medical data and put it available to doctors and nurses in an hourly-based mode where all operations can be registered through a simple interface. The information used in the ICU is provided by different platforms. Figure VII.2 represents well the interoperability between ENR and the other data sources. Opposing the Figure VII.1 is possible verify that all data sources converge to the same platform: ENR. Now, aren't the ICU professionals who request the data but the platform which is always searching from new data from other data sources. The implementation of ENR gives also the opportunity to have more data available. Data which were not recorded electronically like is the case of ventilation, pain scales, medical scores, fluid balance, therapeutic plan, and others.

The other data sources that previously were accessed through a set of distinct application are now available in the ENR. The data provided from the platforms of nurses and doctors are integrated in ENR. The data provided by the laboratory results are now in an open format and they are displayed in a grid format organized by day and exams. The prescription plan provided by the pharmacy is divided by drugs, hour and dosages and it is put in a table with 25 columns one with the name of drug and the others with the hours (24). All the prescriptions data are displayed in a grid. Finally, the platform of AIDA and EHR also are integrated in the ENR. ENR has a special tab with the AIDA. The user of ENR can consult all the data provided by all the platforms however, only can registry and validate the data inserted manually and provided by the bedside monitors and the pharmacy.

The digital nature of the ENR turns the data contained in it searchable and retrievable (Filipe Portela, Gago, et al., 2011).

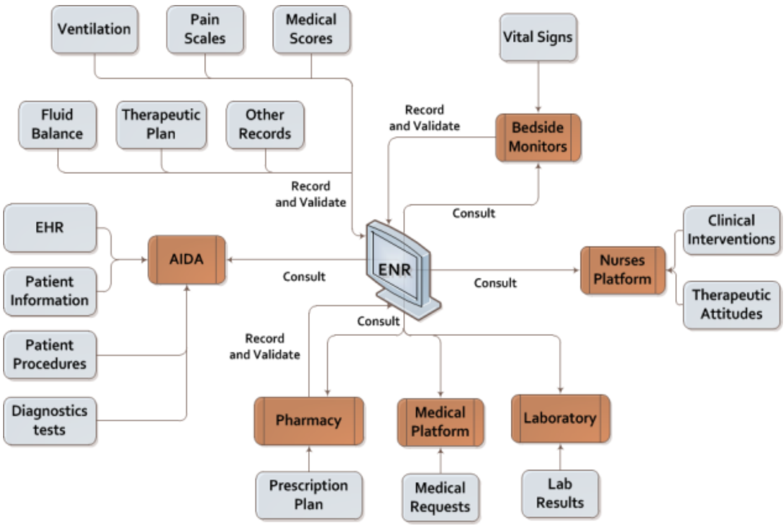


Figure VII.2 – ENR interoperability overview.

This platform is touch and web-based and is composed by different screens, grouping the data by the information provenance. Data about vital-signs and prescriptions will be available in the same grid. This grid has one column for the variable identification, 24 for each hour (1-24) and two more which is used by fluid balance to store the input and output values. In case of vital signs only the first value correctly validated of the hour will be inserted. The data obtained from prescription were extremely and carefully prepared and distributed, in a grid, by hour, dosage, category and type. The data provided from Lab Results has an own screen. This screen also is composed by a grid which facilitates the data consulting. The results collected were grouped by day, hour and type. The professionals can, in an easy and quick way, analyses the patient results today and for examples two days ago. For all cases, always some new result were available the platform will refresh and put them in the correct place.

6. Results

After the new features were incorporated in the INTCare system, the benefits of them for ICU were notorious. The first benefit is related to the number of different information platforms supporting the ICU processes. The ICU professionals normally had to access to six different platforms – Bedside Monitors (BM), Nursing Record Paper Based (NRPB), AIDA, Medical Platform (Mplat), Nurses Platform (Nplat) and Pharmacy, to consult essential information to decision / job. Table VII.1 presents the kind of pieces of information normally consulted from the six different platforms. All of them are now available in a single platform (ENR).

Table VII.1 – ICU Information Platforms

<table>
<tr><th>Information</th><th>Before</th><th>Now</th></tr>
<tr><td>Ventilation</td><td rowspan="5">NRPB</td><td rowspan="16">ENR</td></tr>
<tr><td>Pain Scales</td></tr>
<tr><td>Fluid Balance</td></tr>
<tr><td>Therapeutic Plan</td></tr>
<tr><td>Other Records</td></tr>
<tr><td>EHR</td><td rowspan="3">AIDA</td></tr>
<tr><td>Patient Information</td></tr>
<tr><td>Diagnostics tests</td></tr>
<tr><td>Vital Signs</td><td>BM</td></tr>
<tr><td>Prescription Plan</td><td>Pharmacy</td></tr>
<tr><td>Medical Requests</td><td>Mplat</td></tr>
<tr><td>Clinical Interventions</td><td>Nplat</td></tr>
<tr><td>Therapeutic Attitudes</td><td>-</td></tr>
<tr><td>Lab Results</td><td>-</td></tr>
<tr><td>Patient Procedures</td><td>-</td></tr>
<tr><td>Medical Scores</td><td>-</td></tr>
</table>

Another benefit (Table VII.2) obtained was the number of pieces of information available electronically (E), online and in real-time (ORT). The initial situation (50% of ORT) evolved to a complete ORT approach.

Table VII.2 – Information type available in ICU (comparison)

Information Type	Before		Now	
	Source	ORT	Source	ORT
Ventilation	P	X	E	√
Pain Scales	P	X	E	√
Fluid Balance	P	X	E	√
Therapeutic Plan	P	X	E	√
Other Records	P	X	E	√
HER	E	√	E	√
Patient Information	E	√	E	√
Diagnostics tests	E	√	E	√
Vital Signs	P	X	E	√
Prescription Plan	P / E	√	E	√
Medical Requests	E	√	E	√
Clinical Interventions	E	√	E	√
Therapeutic Attitudes	E	√	E	√
Lab Results	P	X	E	√
Patient Procedures	P / E	√	E	√
Medical Scores	P	X	E	√

The benefits obtained with the changes in the environment and in the information system gives the possibility to create a pervasive intelligent decision support system – INTCare. INTCare system provide three groups of information which can support the decision making process in the different parts of them. Using the streamed data (Gama & Gaber, 2007), i.e., data collected automatically and in real-time, the ENR, AIDA (Abelha, et al., 2003), and the hospital interoperability system

(Brailer, 2005; Häyrinen, et al., 2008) , is possible processing the data automatically and using online learning in order to obtain:

a) ICU Scores;
b) Critical Events;
c) Predict organ failure and patient outcome.

6.1. Improvements and System assessment

All of the improvements were assessed in terms of their functionality and usability. A restrict and specialized number of users were asked. During this time two doctor and six nurses (~15 % of ICU professionals) explored the platform and answered the questionnaires. All of them were familiarized with the ICU Clinical systems, being a daily user of them. The use of a reduced number of professionals was a strategy, because during the test phase of the system is too important not demotivate the professionals with some possible failures. This objective, is easily achieved with a performing of lot of test near the ICU professionals which are more trained for direct patient care and understand the risks / complexity of this type of systems. The users answered to a questionnaire in order to assess the technology. With the objective to give a global idea about the usefulness and ease of use of the system in superficial way, an easy and quick questionnaire was done.

1) To the questionnaire, a metric evaluation scale was defined:
2) Does not meet / in complete disagreement (<20% of cases),
3) Meets some / disagree (20 -40%),
4) Meets / agreement (40-60%),
5) Very satisfied / very agreement (60-80%),
6) Fully meet / fully agree (> 80%))

The questions were grouped in two groups:

1 Functional characteristics
- 1.2 The registration system allows efficient information?
- 1.3 The system allows for an efficient consulting of information for nursing decision support?
- 1.4 The system allows for an efficient consulting of information for medical decision support?
- 1.5 The system enhances the proactive performance of the professionals?

2 Technical characteristics
- 2.2 The system responds with an appropriate quality and speed to the needs?
- 2.3 The access to the information, in terms of speed and availability, correspond to the needs?
- 2.4 The access to the system is easy and secure?
- 2.5 The system interoperability, with other systems (e.g. EHR), suits the needs?
- 2.6 The touch interface at the bed side allows for an easy operation of the system?

The Table VII.3 presents the results obtained, in terms of the percentage of answers for each question and each score level.

Table VII.3 – Questionnaire results

Query	*1*	*2*	*3*	*4*	*5*
Functional characteristics					
1.1	0,00	0,00	16,67	0,00	83,33
1.2	0,00	0,00	0,00	16,67	83,33
1.3	0,00	0,00	0,00	0,00	100,00
1.4	0,00	0,00	0,00	83,33	16,67
Technical characteristics					
2.1	0,00	0,00	0,00	100,00	0,00
2.2	0,00	0,00	0,00	66,67	33,33
2.3	0,00	0,00	0,00	33,33	66,67
2.4	0,00	0,00	0,00	50,00	50,00
2.5	0,00	0,00	16,67	0,00	83,33

The results obtained were treated and despite according the responsibility of each inquired. For example in case of the nurses only question about the nursing care were considered as answer. In a first approach, is possible to observe that the results obtained are good and the ICU professionals are very comfortable with the new system. In general, more than 90 % of the answers are scored with 4 or 5 points which gives a good motivation to continue the work, improving this system.

Only two questions were answered with 3 points in a five-scale point. These questions are related with a functional aspect and a technical aspect.

At level of functional aspect the question which obtained 3 points was concerned with the registration system. This result happens due the possibility of user with the haste records wrong data. This is a valid answer because the clinical care is always the first concern by the nurses. In the technical aspect, the fact of touch-screen being a novelty and the users are accustomed with records in paper it is difficult to some users at first contact have a quick understanding of the system and their benefits. Resuming, it is possible observe that this system allow an efficient and quick access to the essential data which support the clinical decision process.

Now, and after this test phase finish a set of new questionnaires will be done for each TAM area and most people will be added to the process.

7. Conclusions and Future Work

The success of decision support systems depends on the technology acceptance by the users. In the intensive care area this is of the most importance and is very difficult to be achieved. Further than contribute to accurate decision models, a set of specific requirements was considered:

- A complete dematerializations of the processes;
- An online data acquisition and processing;
- A real-time decision support;
- Access to data in open-format in order to evaluate the values in terms of scores and alarms;
- Interoperate with the systems and equipment present in the ICU environment.

This paper presented all the improvements introduced in INTCare system towards to the user's acceptance fulfilling the requirements enunciated. Tables 1 and 2 resume the benefits obtained in terms of the degree of interoperability and the dematerialization attained.

The study carried out on the technology acceptance (Table 3) corroborated the options made during the development of INTCare system. The system is now in use in the ICU of Centro Hospitalar do Porto, Porto, Portugal.

In the future, an analysis of effective gains with the reduction of time recording and patient caring will be done and the impact on ICU Cost/Quality (Cassi, 2009). In parallel all features of the system will be carefully evaluated by the ICU professionals. The system will be in a continuous development and optimization according the answers obtained and recitals done.

Acknowledgements

The authors would like to thank FCT (Foundation of Science and Technology, Portugal) for the financial support through the contract PTDC/EIA/72819/ 2006. The work of Filipe Portela was supported by the grant SFRH/BD/70156/2010 from FCT.

References

Abelha, A., Machado, J., Alves, V., & Neves, J. (2004). Data warehousing through multi-agent systems in the medical arena.

Abelha, A., Machado, M., Santos, M., Sollari, A., Rua, F., Paiva, M., et al. (2003). Agency for Archive, Integration and Diffusion of Medical Information. *Proceeding of AIA*.

Brailer, D. J. (2005). Interoperability: the key to the future health care system. *Health Affairs-Millwood Va Then Bethesda MA-, 24*, 5.

Cassi, A. (2009). Despite Wide Adoption, Little Evidence of eICU Impact on Intensive Care Cost/Quality. *Health Affairs*.

Chooprayoon, V., & Fung, C. C. (2010). TECTAM: An Approach to Study Technology Acceptance Model (TAM) in Gaining Knowledge on the Adoption and Use of E-Commerce/E-Business Technology among Small and Medium Enterprises in Thailand.

Duarte, J., Portela, F., Santos, M. F., António, A., & José, M. (2011). *Electronic Health Record in Dermatology Service.* Paper presented at the *HCist 2011* – International Workshop on Health and Social Care Information Systems and Technologies, Algarve, Portugal.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: An introduction to theory and research*: Addison-Wesley.

Gago, P., Santos, M. F., Silva, Á., Cortez, P., Neves, J., & Gomes, L. (2006). INTCare: a knowledge discovery based intelligent decision support system for intensive care medicine. *Journal of Decision Systems*.

Gama, J., & Gaber, M. M. (2007). *Learning from data streams: processing techniques in sensor networks*: Springer-Verlag New York Inc.

Häyrinen, K., Saranto, K., & Nykänen, P. (2008). Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *International Journal of Medical Informatics, 77*(5), 291-304.

Lapinsky, S., Holt, D., Hallett, D., Abdolell, M., & Adhikari, N. (2008). Survey of information technology in Intensive care units in Ontario, Canada. *BMC medical informatics and decision making, 8*(1), 5.

Portela, F., Gago, P., Santos, M. F., Silva, A., Rua, F., Machado, J., et al. (2011). *Knowledge Discovery for Pervasive and Real-Time Intelligent Decision Support in Intensive Care Medicine*. Paper presented at the *KMIS* 2011- International Conference on Knowledge Management and Information Sharing.

Portela, F., Santos, M., Vilas-Boas, M., Rua, F., Silva, Á., & Neves, J. (2010). *Real-time Intelligent decision support in intensive medicine*. Paper presented at the *KMIS* 2010- International Conference on Knowledge Management and Information Sharing.

Portela, F., Santos, M. F., Gago, P., Silva, Á., Rua, F., Abelha, A., et al. (2011). Enabling real-time intelligent decision support in intensive care.

Portela, F., Santos, M. F., Silva, Á., Machado, J., & Abelha, A. (2011). *Enabling a Pervasive approach for Intelligent Decision Support in Critical Health Care*. Paper presented at the *HCist 2011* – International Workshop on Health and Social Care Information Systems and Technologies.

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., Neves, J. (2009). *Information Architecture for Intelligent Decision Support in Intensive Medicine*. Paper presented at the 8th WSEAS International Conference on Applied Computer &Amp; Applied Computational Science (ACACOS '09).

Silva, Á., Cortez, P., Santos, M. F., Gomes, L., & Neves, J. (2008). Rating organ failure via adverse events using data mining in the intensive care unit. *Artificial Intelligence in Medicine, 43*(3), 179-193.

Vandijck, D. M., Labeau, S. O., & Blot, S. I. (2008). Level of knowledge articulated by intensive care nurses and clinical decision-making. *Intensive and Critical Care Nursing, 24*(1), 6-7.

Varshney, U. (2007). Pervasive healthcare and wireless health monitoring. *Mobile Networks and Applications, 12*(2-3), 113-127.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences, 39*(2), 273-315.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, 186-204.

Wooldridge, M. (1999). Intelligent agents *Multiagent systems: a modern approach to distributed artificial intelligence* (pp. 27-77): MIT Press.

ARTICLE A8

Objectives

To develop and test pervasive models in Intensive Care;

i. To assess the results.

The main objective of this paper is to present an intelligent and integrated data quality system. This system increases the quality of information used in the decision process by the ICU professionals.

Results

The quality process is divided in a set of different continuous and integrated tasks. This process are performed automatically and in real-time. At first level are the data collected automatically by bedside monitors. Now and after talk with nurses about the importance of maintain the sensors connected, only Central Venous Pressure (CVP) (~40%) and temperature (~15%) present bad values. The quality system validates the values collected. After this process all variables don't have values out of range. Figure IX.2 presents an overview of the hours when are verified more cases of data out of range. These hours are associated to the tasks of hygienic (10h-12h) or patient visits (15h-17h).

At the second level it is the patient with correct identification and the records with predefined data. It is possible observe that in 2012 all data collected has a PID and are available electronically, online and in real-time. These changes result in an increase of the data quality.

Finally, it is assessed the quality of DM system. The data quality of the DM Models can be measured using the procedures earlier described. To a better comprehension of the quality results by the user a monitoring interface were implemented. The monitoring system is similar to a traffic light. Using this system the user can know, on moment when he is consulting the prediction, the quality level of results presented by the selected models. Each label corresponds to one of the target quality measure and it is filled with one colour (green, yellow and red).

REAL-TIME DECISION SUPPORT IN INTENSIVE MEDICINE - AN INTELLIGENT APPROACH FOR MONITORING DATA QUALITY

Filipe Portela[1], Manuel Santos[1], José Machado[2], António Abelha[2], Álvaro Silva[3] and Fernando Rua[3]

[1]Algoritmi Center, University of Minho, Azurém, Guimarães, Portugal email: {cfp, mfs}@dsi.uminho.pt
[2]CCTC, University of Minho, Braga, Portugal,
email: {jmac, abelha}@di.uminho.pt
[3]Serviço de Cuidados Intensivos, Centro Hospitalar do Porto, Hospital Santo António, Porto, Portugal,
email: moreirasilva@clix.pt, fernandorua.sci@hgsa.min-saude.pt

Abstract - Intensive Medicine is an area where big amounts of data are generated every day. The process to obtain knowledge from these data is extremely difficult and sometimes dangerous. The main obstacles of this process are the number of data collected manually and the quality of the data collected automatically. Information quality is a major constrain to the success of Intelligent Decision Support Systems (IDSS). This is the case of INTCare an IDSS which operates in real-time. Data quality needs to be ensured in a continuous way. The quality must be assured essentially in the data acquisition process and in the evaluation of the results obtained from data mining models. To automate this process a set of intelligent agents have been developed to perform a set of data quality tasks. This paper explores the data quality issues in IDSS and presents an intelligent approach for monitoring the data quality in INTCare system.

Index Terms— Intensive Care; Data Quality, Real-Time; Intelligent Decision Support; INTCare.

1. Introduction

Nowadays and due the importance which the patient data have for the decision making process, the number of data present in Intensive Care Units (ICU) increased significantly (Lapinsky, et al., 2008; Strand & Flaatten, 2008). This situation motivates the use of Intelligent Decision Support System (IDSS) to obtain knowledge. However the implementation of this type of system is difficult due to the complexity of the environment; the number of data which are regularly collected in paper format (White, Braddock, Bereknyei, & Curtis, 2007) and the quality of information / data collected (Filipe Portela, Vilas-Boas, et al., 2010). To overcome these limitations, a system called INTCare (Gago, et al., 2006; Filipe Portela, Santos, & Vilas-Boas, 2012) was developed with the objective to predict the organ failure and patient outcome for the next 24 hours. This system uses Data Mining (DM) techniques and some variables presented in ICU to predict the targets. At the beginning most of those variables were collected manually, in paper format and in an hourly and offline base mode. In order to resolve this problem an electronic data acquisition system was developed.

This new system gets values from the, patient sensors (vital signs, ventilation), electronic health record (EHR), laboratory and pharmacy system. The data collected from the sensors attached to the patients require a more carefully analysis on the data quality provided. The sensors can be easily disconnected from the patients; measuring bad values and gives a wrong idea of the patient situation. At same time many of the values haven't a correct patient identification. These situations can influence significantly the Data Mining (DM) models and the decision making process. With the objective to validate the values collected automatically, was necessary to implement some procedures. When the data is collected it needs to be processed and transformed according to DM input variables. Finally, the models are induced in real-time using online learning and some new knowledge will be available. At this point some other problems arise, like is the assessment of the quality of the results obtained by the models. All process reported before is executed automatically and in real-time. The difficult of validate each process automatically is evident. The most difficult it is the process of evaluating the DM models in a continuous and real-time way. The quality of models developed by INTCare is evaluated using an ensemble and an agent analyses the quality measures defined to each result whenever a new prediction is done. This agent presents the best data mining results / models or refreshes the models according to the evaluation of the results obtained.

The main objective of this paper is to present an intelligent and integrated data quality system. This system increases the quality of information used in the decision process by the ICU professionals. This paper is divided in seven chapters. After an introduction to the subject a set of concepts are addressed in second chapter. The third chapter presents the quality process and the agents designed. The fourth chapter evaluates the data acquisition process in terms of bedside monitored values, patient identification and process dematerialization. The DM process is evaluated in the fifth chapter. Finally, in the sixth chapter some results are analysed and in the seventh chapter some conclusions and future work are outlined.

2. Background

2.1. Intensive Medicine

Intensive medicine (IM) is a particular area of Medicine. It has specialists which apply their knowledge in Intensive Care Units (ICU). ICU is recognized as a critical environment because it has some complex health care situations and their patients are in very weak health conditions (Bricon-Souf & Newman, 2007). The professionals of ICU spend more time in the patient direct care than in data documentation (Mador & Shaw, 2009). Due to the reduced number of automatic acquisition system, this attitude can cripple a system like it is INTCare. Being the direct care, correctly always the first concern the solution is automating as possible the acquisition and data quality processes (F. Portela, et al., 2011a; Filipe Portela, Santos, & Vilas-Boas, 2012). This solution reduces the human efforts with data documentation and makes it feasible to introduce an IDSS in intensive medicine.

2.2. INTCare

INTCare is a research project and has as main objective develop a Pervasive Intelligent Decision Support System to automatic and in real-time predict the organ failure (renal, coagulation, hepatic, cardiovascular, neurologic and respiratory) and patient outcome. During the development of the project a set of change in the ICU were done. A gateway was implemented to obtain the vital signs values from bedside monitors. Some changes in protocols of obtain the data from Laboratory and pharmacy systems were done.

The objective of these changes was increase the number of data available and getting the patient results and therapeutic plans in an open format.

2.3. Data Mining variables

The idea of designing this system arose after obtaining good results using offline data. These results (Álvaro Silva, et al., 2008) were obtained during a studied using EURICUS database which showed be possible using some patient variables predict the organ failure. The variables used were: Age, Critical Events, Admission Variables, Outcome, and SOFA. During the INTCare some other variables were added to the Data Mining Input:

a) **SOFA** Cardio, Respiratory, Renal, Liver, Coagulation, neurologic;
b) **Case Mix** = {Age, Admission type, Admission from};
c) **Accumulated Critical Events** (ACE) = {ACE of Blood Pressure, ACE of Oxygen Saturation, ACE of Heart Rate, ACE of Urine Output}
d) **Ratios**;
e) **Outcome**;

To ensure the quality of the values used, which are now collected in real-time it is necessary a constant and automatic validation of those variables. After the induction of the models and having in account that INTCare also is an adaptive system is too important have the possibility for automatically and in real-time assess the quality of models and induce new models when necessary.

2.4. Oracle Data Mining

One of the most difficulty to implement INTCare was the possibility of perform all KDD tasks automatically and in real-time. Most of the solutions explored don't allow a data transformation, model induction or apply the results using always a database. After do a depth research a solution was found. Oracle Data Mining (ODM) provides a comprehensive collection of Data Mining analytics as part of the Oracle database environment which supports the development, integration and deployment of Data Mining applications (Tamayo, et al., 2005). All KDD tasks are now performed using ODM.

3. Data Quality Process

The introduction of data quality measures in the INTCare is too important to the success of the system. The data quality process is totally ensured by intelligent agents (Shijia Gao & Dongming Xu, 2009). These agents are responsible for perform automatically some tasks. INTCare system is composed by four subsystems (Filipe Portela, Santos, et al., 2010): Data acquisition, knowledge management, Inference and Interface. Figure VIII.1 makes an overview of interaction between the agents which are in each subsystem. The process has two starting points: having some new data or there is an INTCare prevision request. In the first case the validity of data collected are analysed. In the second case the Data Mining models are executed and the quality of the scenarios created is evaluated.

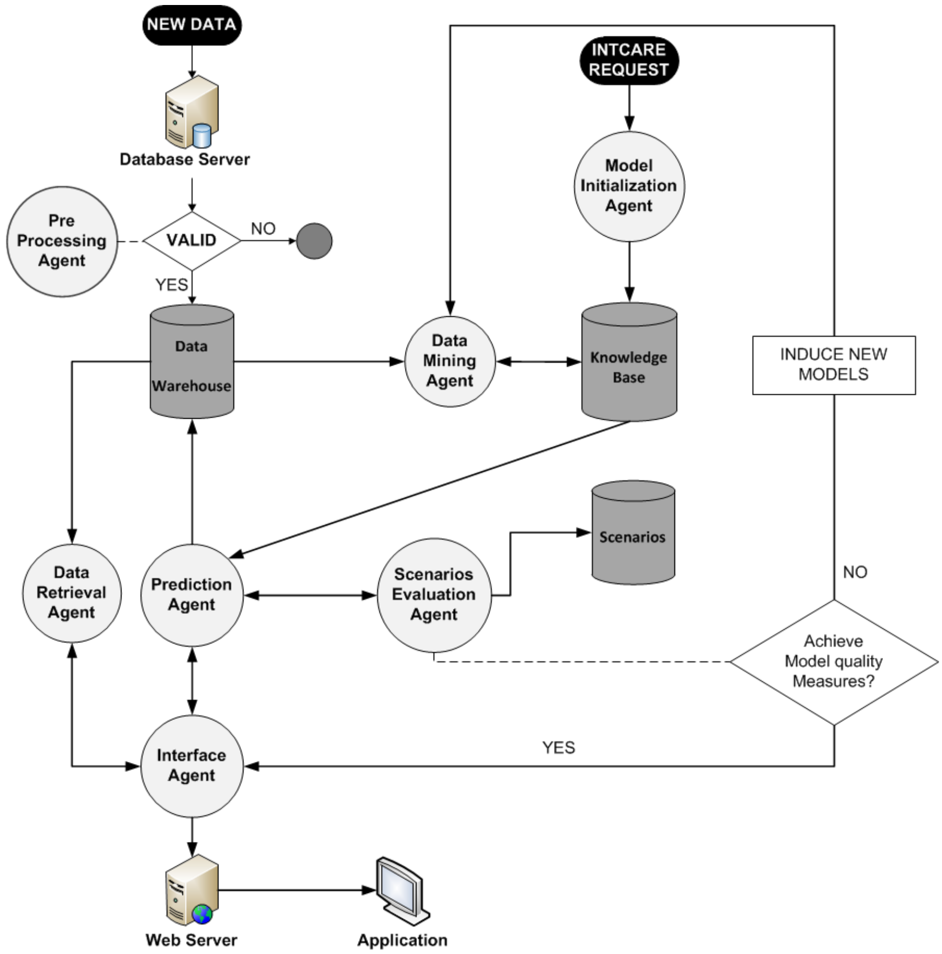


Figure VIII.1 – Overview of INTCare data quality system.

The Data quality process is ensured by three agents. However there are other four agents which their tasks are dependent from the data quality process. The agents in use or directly conditioned by this process are:

- ✓ **Pre-processing** agent is integrated in Data Acquisition sub-system and verifies all data collected automatically. For each value collected it analyses if the value is possible or not.
- ✓ **Data Mining** Agent executes the data mining engine always some new request to improve data mining models is done. By default it is executed once a day or always the INTCare system is used by first time.
- ✓ **Model Initialization** Agent is responsible by answer to the INTCare request. After receive the message the agent processes the models according the request.
- ✓ **Prediction Agent** processes the results presented in knowledge base using the data warehouse. This agent only is executed when scenarios evaluation agents gives some indication about quality of scenarios.

- ✓ **Scenarios Evaluation** Agent is the other agent which has data quality tasks. This agent has the responsibility of evaluate the results provided from prediction agents. According the quality of the models it sends a message to the data mining agent to induce new models or send a message to the prediction agent with the results of the scenario chosen.
- ✓ **Data Retrieval** Agent is responsible to ensure that the data in use by prediction agent are correct and are the most recently collected to each patient. This agent is used to cases where there are subsequent changes, as is the patient outcome – patient is discharged in live condition and days after he die, or cases where since models are executed until the results are obtained some new data arrive. This agent is also responsible to delete rows with null values or values not recognized.
- ✓ **Interface** agent is executed by indication of prediction agents and only has the task of put the results available in the INTCare interface.

The agents are integrated in decision making process and allow having data with high level of quality for automatically and in real-time attain:

a) Intensive Care Scores (Filipe Portela et al., 2012) ;
b) Patient Critical Events (P. G. Filipe Portela, Manuel Filipe Santos, Álvaro Silva, Fernando Rua, 2012);
c) Predict organ failure and patient outcome (F. P. Filipe Portela, Manuel Filipe Santos, 2012).

4. Evaluation of Data Acquisition Values

The process of validation values in real-time and in areas like is IM is too important and isn't always correct.

4.1. Bedside Monitors values

The values validated by this process are collected by the bedside monitor (BM). BM is used to collect vital signs data and through patient sensors. These sensors are connected to the patient and can be easily moved by them or the nurses during the patient care. Proving this fact they are for example the temperature and diastolic sensors. For the temperature if the sensor is connected to the patient is measuring the body temperature (~36º) otherwise is measuring room temperature (~22º). In case of diastolic can be measuring the diastolic or, if moved, systolic values. To resolve these errors an automatic procedure was implemented. This procedure is started by a trigger always some new value is inserted in database.

The validation process has two stages. First, the values collected (Spo2 and Heart Rate) are validated according the values present in table VIII.1. Second, for the Blood Pressure (BP) the validation of Mean Arterial Pressure (MAP) is done. It is defined that MAP is (1/3(systolic blood pressure) + 2/3 (diastolic blood pressure)) (Sesso et al., 2000). The procedure compares at the same time if the BP values collected (systolic and diastolic) are between the possible ranges and if the MAP of these values are upper than 40 (minimum value possible)(Aubuchon, 2011).

Table VIII.I presents the range of the values that normally can be accepted in ICU (ICUMIN and ICU MAX); it also presents a second range, which contains some abnormal values that can be verified (Amin and Amax). In the automatic pre-processing phase is been used Amin and Amax. These ranges are defined in accordance with the ICU doctors and with the patient data reality.

Table VIII.1 – Vital Signs - Range of Values

Vital Sign	ICUMin	ICUMax	AMin	AMax
Blood Pressure (BP)	50	180	0	300
SPO2	80	100	40	100
Temperature (Temp)	34	42	30	45
Respiratory Rate (RR)	0	40	0	40
Heart Rate (HR)	5	40	0	250

In this process there are other data quality procedures. At same time it validates the data inserted, i.e., verify if there isn't a negative value, in case of scales the value collected need to be between the max and min number possible. In other case if only numbers are allowed, this agent ensures that data collected are numbers. If some value doesn't achieve these requirements the respective row is deleted. All of the data can be manual validated through Electronic Nursing Record (ENR). This situation allows having values out of range when valid. If some value is out of range it is deleted automatically. In the ENR only appear the values automatic validated. The humans can delete the automatic value and put manually the correct. When the value is manually inserted the automatic validation didn't do nothing, because only the professionals can know if some value is true or not, independent it is or not out of range. The values manually inserted are considered valid and are used by the DM Models. In case of be verified some error the humans can correct this values.

4.2. Patient Identification

One of the most commons problems in ICUs is the patient identification (PID). The PID is put manually in the bedside monitors; normally the nurses are concerned with the patient and not with putting the correct identification. This situation leads to an automatic acquisition of a high set of values without identification.

In order to overcome this situation two solutions were tested. One is verify the PID through the Electronic Health Record (EHR), crossing the Bed Number and the information of the patient admitted in this bed. The procedure verifies the bed and consults in the EHR the respective PID. Then always some new value is collected the procedure update the data automatically putting the correct PID. The second solution still is in tests and uses RFID technologies (Aguilar, van der Putten, & Kirrane, 2006; Chowdhury & Khosla, 2007) to identify the patient. When the patient comes to ICU a tag is put in the arm. Then the nurses associate in the EHR: the PID, the patient bed and TAG ID. All beds have one antenna to read the tag and identify the patient. When the patient is in the bed, the RFID system reads the TAG and identify the patient (Rui Rodrigues, 2012). When the patient goes out, the RFID can't detect any patient and stops the vital signs registry. Both processes use intelligent agents to perform the tasks described above.

4.3. Process Dematerialisation

Other change was in the record process through the processes dematerialization. Now, instead of the professional records the patient values manually in the paper, they have an electronically, online and touch-screen platform: Electronic Nursing Record (ENR) to registry the clinical data.

This platform doesn’t allow text writing. All the data possible are predefined or number. This situation increases the data quality because all of them record the data in the same way, i.e., the text records aren’t allowed because they have to choose a predefined value or a number. The ENR is not only used to register but also to validate and confirm the values that is automatically collected. ENR make available all data than before were recorded manually in the paper.

5. Evaluation of Data Mining Models

In order to ensure the quality of the models developed some evaluation procedures were designed. Like presented in figure 1 always some request is done to the Model Initialization Agent the knowledge base is used to predict the targets. For each scenario a set of evaluations are done. At the end of comparing all models, a DM ensemble is induced. INTCare System uses the models which present the better sensibility and also have a very good level of accuracy and total error.

First and after the DM input table is prepared, a procedure is executed to delete all rows with null values in any of the columns used by the models. This procedure cleans the nulls automatically and prepares the data to induce the models. Then, the DM engine is executed only after all values are filled. Until them the user only can consult individual values because the new knowledge isn't available. This solution avoids incorrect read of the previsions, because a bad prevision can be fatal to the patient. Always some new data is obtained the models are refreshed in order to present the best models according the real data in the moment.

The ensemble is organized in terms of six independent components (target) considers seven different scenario and applies three distinct DM techniques: Decision Trees (DT), Support Vector Machine (SVM) and Naïve Byes (NB). The ensemble process is divided in two steps:

- *Predictive Models* – 126 models are induced combining seven scenarios (S1 to S7), six targets and three different techniques (SVM, DT and NB);
- *Ensemble* – all the models are assessed in terms of the sensibility, accuracy and total error.

The models are induced in real-time and using online-learning by the DM agent. This agent runs whenever a request is sent or when the performance of the models decreases.

In order to choose the best predictive model for each target, a set of tasks are performed automatically and in real-time:

1. Create the confusion matrix for each scenario;
2. Obtain the assessment measures;
3. Apply the quality measure;

For each model 10 runs are performed. Then the confusion matrix (CMX) is automatically obtained for each models. Through the CMX is possible obtain:

a) Number of false positives (FP)
b) Number of true positives (TP)
c) Number of false Negatives (FN)
d) Number of true negatives (TN)

Then, a procedure is executed recursively by scenarios agents for each model (M) in order to obtain the measures:

SENSIBILITY (M) = TP (M) / (TP (M) + FN (M))
ACCURACY (M) = (TP (M) + TN (M)) / (TP (M) + TN (M) + FP (M) + FN (M)

TERROR (M) = FP (M) + FN (M) / (TP (M) + TN (M) + FN (M) + FP (M))

The main goal of the ensemble is to select the most suited model from a set of candidates. In order to assess the models quality, a measure was defined. This measure is based in the results obtained by the models during the 10 runs in terms of sensibility, accuracy and total error. The selected models are used only if they satisfy the following conditions:

- *Total Error* <= 40%
- *Sensibility* >= 85%
- *Accuracy* >= 60%

6. Results

The quality process is divided in a set of different continuous and integrated tasks. This process are performed automatically and in real-time.

At first level are the data collected automatically by bedside monitors. The Table VIII.2 presents an overview of the data collected in ICU. In 2009 most of the data were out of range (Filipe Portela, Vilas-Boas, et al., 2010). Now and after talk with nurses about the importance of maintain the sensors connected, only Central Venous Pressure (CVP) (~40%) and temperature (~15%) present bad values. The quality system validates the values collected. After this process all variables don't have values out of range. Figure VIII.2 presents an overview of the hours when are verified more cases of data out of range. These hours are associated to the tasks of hygienic (10h-12h) or patient visits (15h-17h).

Table VIII.2 – ICU Vital Signs Data Analysis

	Limit	<MIN (%)	>MAX (%)	=Limit (%)	Mean	Stdev	Mode	MIN	MAX
HR	[40;160]	0,19	0,10	99,71	100,00	35,07	89,00	0,00	250,00
SPO2	[80;100]	1,48	0,00	98,52	90,00	6,20	100,00	0,00	100,00
MAP	[30;160]	0,56	1,83	97,61	95,55	31,47	77,75	-39,66	319,96
Systolic	[50;180]	0,44	6,19	93,37	116,29	37,44	120,36	-40,00	320,00
Diastolic	[20:140]	0,31	1,46	98,23	80,00	33,96	59,57	-39,97	320,00
CVP	[0;40]	5,51	33,84	60,65	20,00	11,98	0,04	-40,00	320,00
TEMP	[34;42]	14,01	0,01	85,98	38,00	2,73	36,82	-253	158,72
RR	[5;40]	0,00	0,23	99,77	20,00	11,98	0,00	0,00	70,00

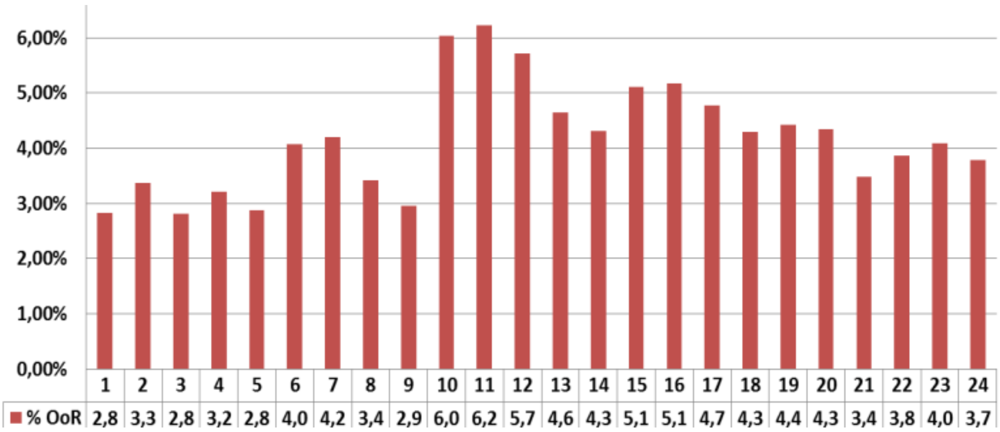


Figure VIII.2 – Overview of data out of range by hour.

At the second level it is the patient with correct identification and the records with predefined data. Figure VIII.3 present a comparison between 2009 and 2012, at level of patient identification and data format. It is possible observe that in 2012 all data collected has a PID and are available electronically, online and in real-time. These changes result in an increase of the data quality.

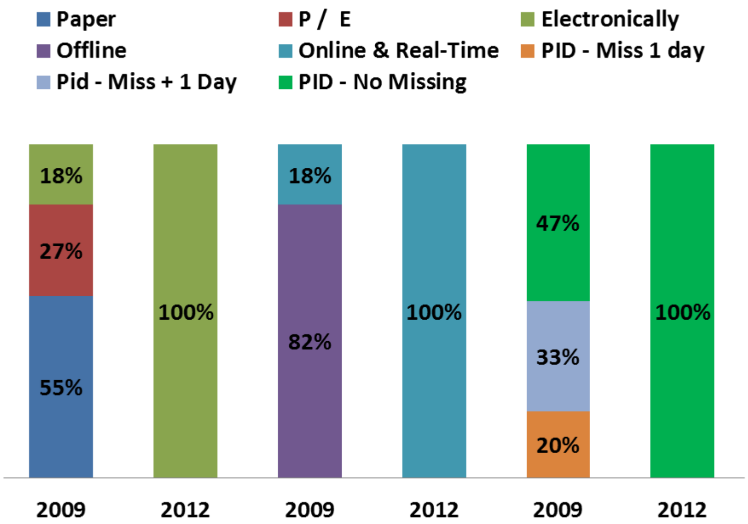


Figure VIII.3 – Overview of the patient identification and type of data in 2012.

Finally, it is assessed the quality of DM system. The data quality of the DM Models can be measured using the procedures earlier described. To a better comprehension of the quality results by the user a monitoring interface were implemented.

The monitoring system is similar to a traffic light. Using this system the user can know, on moment when he is consulting the prediction, the quality level of results presented by the selected models. Each label corresponds to one of the target quality measure and it is filled with one colour (green, yellow and red). Table VIII.3 shows the ranges, where first level is green (very good models), the second level is yellow (acceptable models) and the third level is red (excluded models).

Table VIII.3 – Evaluation Criteria

Measure	Min Level 1	Max Level 1	Min Level 2	Max Level 2	Min Level 3	Max Level 3
Accuracy	85%	100%	60%	85%	0%	60%
Sensibility	95%	100%	75%	95%	0%	75%
Toral Error	0%	40%	40%	60%	60%	100%

Table VIII.4 displays the results achieved by the ensemble having in account the quality measures. At bold it is the results that achieve the quality measure and at red it is the rejected results. Looking to the table is possible observe that only 50% of the targets present satisfactory results. The INTCare system using the quality system only shows prevision to renal, cardiovascular and respiratory system. Due the fact of the data are in constant changes this values can be quickly modified and other models present best results.

TABLE VIII.4 – Ensemble Performance

Measure	Accuracy	Sensibility	Terror
Renal	**97,95** ± 0,31	**76,81** ± 2,35	**41,81** ± 5,75
Respiratory	**91,20** ± 3,57	**65,69** ± 3,83	**49,61** ± 6,15
Coagulation	69,24 ± 9,41	**82,89** ± 2,57	**87,34** ± 3,22
Cardiovascular	**99,77** ± 0,33	**63,58** ± 3,11	49,58 ± 4,90
Hepatic	77,17 ± 12,41	43,08 ± 4,66	43,08 ± 4,66
Outcome	67,11 ± 5,67	**63,86** ± 4,27	**60,39** ± 6,75

7. Conclusions & Future Work

The introduction of the data quality component brings direct benefits to the IDSS. For a real-time system, is fundamental to assure the data quality in an autonomous and integrated way. The development of a system able to validate automatically and in real-time the data collected improves the quality of information used and the efficiency of the data mining models induced. Using the data provided by the quality system it is possible: to present better information to decision process and to create new knowledge. It is also possible to calculate ICU Scores: SAPS, SOFA, MEWS and TISS; calculate Critical Events (BP, Temperature, SPO2 and Urine Output) and predict organ failure or patient outcome with high level of confidence. All of this changes allow to have a bigger control of the data and a high number of data available automatically and in real-time. With this system the professionals have a complete control of the data (e.g. they can change the values whenever the values aren't correct).

In the future, other techniques to validate the information and RFID will be explored in order to assure the correct identification of patients. Some improvements will be done in the quality system and in the data mining models.

Acknowledgments

This work is supported by FEDER through Operational Program for Competitiveness Factors – COMPETE and by national funds though FCT – Fundação para a Ciência e Tecnologia in the scope of the project: FCOMP-01-0124-FEDER-022674. The authors would like to thank FCT (Foundation of Science and Technology, Portugal) for the financial support through the contract PTDC/EIA/72819/ 2006. The work of Filipe Portela was supported by the grant SFRH/BD/70156/2010 from FCT.

References

Aguilar, A., van der Putten, W., & Kirrane, F. (2006, 2006). Positive patient identification using RFID and wireless networks.

Aubuchon, V. (2011). Blood Pressure Chart - Normal Blood Pressure Range. Medicine Summaries Heart Health Retrieved 07-11-2012, 2012, from http://www.vaughns-1-pagers.com/medicine/blood-pressure.htm

Bricon-Souf, N., & Newman, C. R. (2007). Context awareness in health care: A review. [10.1016/j.ijmedinf.2006.01.003]. International Journal of Medical Informatics, 76(1), 2-12.

Chowdhury, B., & Khosla, R. (2007, 2007). RFID-based hospital real-time patient management system.

Filipe Portela, F. P., Manuel Filipe Santos. (2012). Data Mining Predictive Models For Pervasive Intelligent Decision Support In Intensive Care Medicine. Paper presented at the KMIS 2012 - International Conference on Knowledge Management and Information Sharing.

Filipe Portela, P. G., Manuel Filipe Santos, Álvaro Silva, Fernando Rua. (2012). Intelligent and Real-time Data Acquisition and Evaluation to Determine Critical Events in Intensive Medicine. Paper presented at the HCist'2012 - International Conference on Health and Social Care Information Systems and Technologies, Portugal.

Gago, P., Santos, M. F., Silva, Á., Cortez, P., Neves, J., & Gomes, L. (2006). INTCare: a knowledge discovery based intelligent decision support system for intensive care medicine. Journal of Decision Systems.

Gao, S., & Xu, D. (2009). Conceptual modeling and development of an intelligent agent-assisted decision support system for anti-money laundering. Expert Systems with Applications, 36(2, Part 1), 1493-1504.

Lapinsky, S., Holt, D., Hallett, D., Abdolell, M., & Adhikari, N. (2008). Survey of information technology in Intensive care units in Ontario, Canada. BMC medical informatics and decision making, 8(1), 5.

Mador, R. L., & Shaw, N. T. (2009). The impact of a Critical Care Information System (CCIS) on time spent charting and in direct patient care by staff in the ICU: a review of the literature. International Journal of Medical Informatics, 78(7), 435-445.

Portela, F., Santos, M., Vilas-Boas, M., Rua, F., Silva, Á., & Neves, J. (2010). Real-time Intelligent decision support in intensive medicine. Paper presented at the KMIS 2010- International Conference on Knowledge Management and Information Sharing.

Portela, F., Santos, M. F., Gago, P., Silva, Á., Rua, F., Abelha, A., et al. (2011). Enabling real-time intelligent decision support in intensive care.

Portela, F., Santos, M. F., Machado, J., Silva, Á., Rua, F., & Abelha, A. (2012). Intelligent Data Acquisition and Scoring System for Intensive Medicine. In Springer (Ed.), Lecture Notes in Computer Science - Information Technology in Bio- and Medical Informatics (Vol. 7451/2012, pp. 1-15). Viena, Austria.

Portela, F., Santos, M. F., & Vilas-Boas, M. (2012). A Pervasive Approach to a Real-Time Intelligent Decision Support System in Intensive Medicine. In Springer (Ed.), Communications in Computer and Information Science (Vol. 0272, pp. 14).

Portela, F., Vilas-Boas, M., Santos, M. F., & Fernando, R. (2010). Improvements in data quality for decision support in Intensive Care. Paper presented at the eHealth 2010 - 3rd International ICST Conference on Electronic Healthcare for the 21st century.

Rui Rodrigues, P. G., Miguel Miranda, Carlos Portela, Manuel Santos, José Neves, António Abelha, José Machado. (2012). Monitoring Intelligent System for the Intensive Care Unit using RFID and Multi-Agent Systems. Paper presented at the IEEE International Conference on Industrial Engineering and Engineering Management, Hong Kong.

Sesso, H. D., Stampfer, M. J., Rosner, B., Hennekens, C. H., Gaziano, J. M., Manson, J. A. E., et al. (2000). Systolic and diastolic blood pressure, pulse pressure, and mean arterial pressure as predictors of cardiovascular disease risk in men. Hypertension, 36(5), 801-807.

Silva, Á., Cortez, P., Santos, M. F., Gomes, L., & Neves, J. (2008). Rating organ failure via adverse events using data mining in the intensive care unit. Artificial Intelligence in Medicine, 43(3), 179-193.

Strand, K., & Flaatten, H. (2008). Severity scoring in the ICU: a review. Acta Anaesthesiologica Scandinavica, 52(4), 467-478.

Tamayo, P., Berger, C., Campos, M., Yarmus, J., Milenova, B., Mozes, A., et al. (2005). Oracle Data Mining. Data Mining and Knowledge Discovery Handbook, 1315-1329.

White, D. B., Braddock, C. H., Bereknyei, S., & Curtis, J. R. (2007). Toward shared decision making at the end of life in intensive care units - Opportunities for improvement. [Article]. Archives of Internal Medicine, 167(5), 461-467.

ARTICLE A9

Objectives

To develop and test pervasive models in Intensive Care;

i. To assess the results.

In order to assess the results achieved, the technology, the INTCare functionalities and their importance to ICU, a questionnaire was developed. This questionnaire is based in the Technology Acceptance Methodology (TAM) (Chooprayoon & Fung, 2010) and it is concerned to the evaluation of four aspects: perceived usefulness (PU), perceived ease of use (PEOU), behavioural intention (BI) and use behaviour (UB).

Results

After collecting answers from 14 questionnaires sent by email (35% total number of nurses in ICU) an analysis of the results was performed. First a processing was done to avoid invalid or inconsistent answers given by the participants. Then was noticed that only one participant out of the 14 nurses answered the questionnaire in an inconsistent way for the proposed questions. This situation leaded to only consider 13 surveys.

A global analysis was done in order to understand the best features, the average, the mode and the standard deviation (Stdev) for each one of constructs.

First, it should be noted that the initial objective proposed regarding to the junction of the Technology Acceptance Model (TAM 3) with the Delphi's method in order to evaluate the acceptance by users, their perceptions and the impact on the INTCare's system usage behaviour, it is totally innovative and can be considered a success. This is the first approach to assess the impact of such type of solutions in ICU environment.

In order to understand the acceptance it is possible to conclude that the ICU staff it is very comfortable with the system INTCare. They pointed the data access as the biggest problem and the utility of the information generated for the decision process as the biggest gain.

PERVASIVE INTELLIGENT DECISION SUPPORT SYSTEM - TECHNOLOGY ACCEPTANCE IN INTENSIVE CARE UNITS

Filipe Portela[1] , Jorge Aguiar[1] , Manuel Filipe Santos[1], Álvaro Silva[2], Fernado Rua[2]

[1] Algoritmi Center, University of Minho, Guimarães, Portugal

[2]Intensive Care Unit, Centro Hospitalar do Porto, Portugal

[1]cfp@dsi.uminho.pt, jorge.phena@gmail.com, mfs@dsi.uminho.pt

[2] moreirasilva@clix.pt, fernandorua.sci@hgsa.min-saude.pt

Abstract. Intensive Care Units are considered a critical environment where the decision needs to be carefully taken. The real-time recognition of the condition of the patient is important to drive the decision process efficiently. In order to help the decision process, a Pervasive Intelligent Decision Support System (PIDSS) was developed. To provide a better comprehension of the acceptance of the PIDSS it is very important to assess how the users accept the system at level of usability and their importance in the Decision Making Process. This assessment was made using the four constructs proposed by the Technology Acceptance Methodology and a questionnaire-based approach guided by the Delphi Methodology. The results obtained so far show that although the users are satisfied with the offered information recognizing its importance, they demand for a faster system.

Keywords: TAM, INTCare, Technology Acceptance, Intensive Care, Decision Support System, Pervasive, Technology Assessment

1. Introduction

Decision making in Intensive Medicine (IM) is a crucial process because deals with critical condition patients. Nothing can fail and if something wrong happens the patient can die. It is a specific area of Medicine and their knowledge it is practiced in the Intensive Care Units (ICU). ICU is recognized as a critical environment where the decision needs to be performed fast and with a high level of accuracy (Baggs et al., 2007). In the ICU the patient care is always the main concern and tasks like is patient documentation are relegated for a second plane (Mador & Shaw, 2009). The introduction of intelligent decision support system (IDSS) in the support of decision process is claimed by many of the nurses and physicians which work in ICU. This type of support can be addressed by a pervasive system which operates automatically and in real-time. This system can give to the ICU staff a better comprehension about the patient condition and at the same time predict future situations. INTCare is framed in this type of system. It is a system developed by this research team and which has as main goal the prediction of the patient organ failure and patient outcome in real-time for the next 24 hours.

With the development of the project, other types of sources were deepened and as result some new knowledge were obtained. Currently, INTCare it is considered by the ICU staff a very useful and complete platform, being composed by a set of pertinent information for the Decision Making Process (DMP). To this work, a list of requirements was defined based on the needs of ICU and the goal to make the system more suitable to the environment. They may be summarized as:

R1. To implement an online data acquisition component;

R2. To make available the laboratory results in an open format;

R3. To allow an open access to prescriptions, interventions and therapeutics;

R4. To dematerialise the nursing records;

R5. To integrate the main systems used in ICU in a single platform;

R6. Develop an automatic system to process and transforming the data.

Taking advantage of the modifications introduced (R1 to R6) it is possible to determine automatically and in real-time, using online learning:

a) ICU medical scores (Portela et al., 2012);
b) ICU Critical events (P. G. Filipe Portela, Manuel Filipe Santos, Álvaro Silva, Fernando Rua, 2012);
c) Probability of occur an organ failure probability and patient die (F. P. Filipe Portela, Manuel Filipe Santos, 2012).

In order to assess the results achieved, the technology, the INTCare functionalities and their importance to ICU, a questionnaire was developed. This questionnaire is based in the Technology Acceptance Methodology (TAM) (Chooprayoon & Fung, 2010) and it is concerned to the evaluation of four aspects: perceived usefulness (PU), perceived ease of use (PEOU), behavioural intention (BI) and use behaviour (UB). Despite of the questionnaire used be composed by a high number of questions, in this paper only are presented the TAM results associated to the decision making process, i.e., the results related to DMP.

This paper is divided in seven sections. The first and second sections introduce the work and make an overview of the concepts and work previous performed. The third section presents the improvements attained in the Intensive Care according to INTCare features. Then the fourth section introduces the PIDSS at the level of results achieved. The fifth and sixth sections are related to the

TAM; the questionnaire performed and results achieved. Finally, some remarks and future work are considered.

2. Background

During 2009 when this project started the ICU information system was composed by a set of information silos. The ICU professionals need to access to more than five hospital applications to obtain important information and to make their decisions. Now with the introduction of INTCare they have the most important information available in the Electronic Nursing Record (ENR) and they can obtain new knowledge automatically moments after the patient documentation. This situation only was possible with the modification made by INTCare project and the introduction of new knowledge into the ICU DMP.

2.1. INTCare

INTCare (Gago et al., 2006; Santos et al., 2011) is an IDSS to predict organ failure and patient outcome for the next 24 hours in real-time using online learning. This system is result of a research project.

The research and modifications done allowed to obtain new types of data in an electronically format and in real-time (Portela et al., 2011; Portela et al., 2010). The new reality and the new environment created (Portela, Santos, Silva, Machado, & Abelha, 2011) allow for obtaining new knowledge fundamental to the decision process predicting patient condition, scoring the ICU measures and tracking critical events automatically and in real-time. The data is obtained through a streaming process and the knowledge attained is disseminated in situated devices.

2.2. Decision Making Process in Intensive Care units

Making decisions in ICU is a complicated and danger process, because all tasks need to be performed quickly and accurately (Baggs, et al., 2007). The ICU professionals deal with patients in serious life-risk. The use of technologies to support this type of process is welcome (Mador & Shaw, 2009) however, normally this type of systems aren't helping, i.e., don’t present the accurate information in the right time and in the right place. These types of situations complicate the decision process.

At the same time, there is a problem associated to the patient documentation because it is always relegated to a second place. In order to overcome this situation some modifications were made in the ICU environment (Portela, Santos, et al., 2011) and in the DMP.

2.3. Technology Acceptance Methodology and Delphi Methodology

The evaluation of a technology application is crucial to comprehend its suitability in a specific environment and also to measure the satisfaction level of its users. One of the most used models in this area is the Technology Acceptance Methodology.

"TAM is adapted from the Theory of Reasoned Action (TRA) model which describes human behaviours in a specific situation" (Fishbein & Ajzen, 1975). The main purpose of TAM is to present an approach to study the effects of external variables towards people's internal beliefs, attitudes, and intentions (Chooprayoon & Fung, 2010). This model is also important because it gives an understanding about the acceptance of the decision support by the ICU staff and how can be useful in the course of their daily work.

The goals of TAM can be achieved by using methodologies based on questionnaires. As a support tool it is important to use some aspects/characteristic of the Delphi method. The principles of the Delphi method involves the use of questionnaires being one of its key features (Zackiewicz & Salles Filho, 2010) the preservation of anonymity of the participants.

A questionnaire was prepared by a coordination team, composed by professionals of ICU and Information System, and sent to a set of participants (a group of experts from the ICU nurses team). The questionnaire was prepared taking into account the constructs of TAM (Venkatesh & Bala, 2008; Venkatesh & Davis, 2000). The correlations of the answers were evaluated through the Kendall's tau (τ) coefficient. Kendall's tau is a measure of rank correlation. The values range from −1 (inversion) to +1 (perfect agreement). A value of zero indicates the absence of association

2.4. Related work - Results obtained in the first Approach

In order to make a first assessment of the technology, a quick and short questionnaire was produced (M. F. S. Filipe Portela, José Machado, António Abelha, José Neves, Álvaro Silva, Fernando Rua, 2012). The main goal was to have a first idea about the usefulness and ease of use of the system in superficial way. This questionnaire was the starting point of the second questionnaire (with tam) and it had a short scope.

The questions were divided into two groups: Functional characteristics (data registration, information access and proactive performance) and Technical characteristics (efficient consulting, response time, system security, usability, and interoperability). Finally, a last question evaluate if the system suits the needs. The questionnaire was answered using a five-scale metric: Does not meet / in complete disagreement (<20% of cases) (1) until fully meet / fully agree (> 80%) (5) (M. F. S. Filipe Portela, José Machado, António Abelha, José Neves, Álvaro Silva, Fernando Rua, 2012). In terms of results only two questions were answered with less than 4 points: one question about the registration system and other question about the understanding of the system and their benefits.

Concluding, in the first phase of assessment the users revealed to be comfortable with the system. These results motivated: i) to continue the development of the project; and ii) perform a more extensive and deep questionnaire having the objective to understand the technology acceptance by the ICU users.

3. Research Propose - Improvements introduced in ICU

The improvements made are according to the INTCare requirements (R1 to R6) defined in the introduction and can be summarized as:

3.1. Data acquisition system (R1)

The first requirement was resolved with the implementation of a gateway. The gateway is connected to the vital signs monitors, reads the patient information and stores it on a database through the data acquisition agent. This is an autonomous process and it is always in a continuous collecting process (streaming).

In this phase two problems appear: missing patient identification (PID) and the acquisition of bad values. To overcome these problems two triggers were developed. One trigger to verify on the Electronic Health Record (EHR) system the PID of the patient admitted in bed where the values are provided and other trigger to validate the values. This second process uses the range of values pre-defined by ICU. Both the procedures are executed in the moment of the values are collected.

3.2. Laboratory (R2)

Regarding to the laboratory, an effort was made to have the lab results in an open format, i.e., accessible electronically and able to be handled without restrictions. The main objective was making the results available for ICU immediately after the patient exams are concluded. This change gives the possibility to have the results in a comparative format during the patient stay in ICU. Those exams have different types and are executed by different services and at different hours (M. F. S. Filipe Portela, José Machado, António Abelha, José Neves, Álvaro Silva, Fernando Rua, 2012).

3.3. Open access to prescriptions (R3)

In this point the objective was deal with pharmacy and to study the possibility to construct an easy access to patient prescriptions. These prescriptions were totally controlled by pharmacy and whenever someone needed to consult the patient therapeutic plan had to open a too slow platform. Now, the interaction between the pharmaceutical system and ENR is made by an agent. Periodically, the ENR agent sends a request to the pharmacy drugs system and then, the requested data is sent to a database table (M. F. S. Filipe Portela, José Machado, António Abelha, José Neves, Álvaro Silva, Fernando Rua, 2012).

3.4. Electronic Nursing Record (R4 and R5)

Electronic Nursing Record (ENR) is a platform that it was developed with the objective to receive all medical data and put it available electronically and in real-time to the physicians and nurses in an hourly mode. ENR can achieve two requirements because being it electronic can dematerializes the processes and due the interoperability mode can interoperate with all of others ICU data sources. Currently the ICU staff using the ENR has more vital information about the patient in order to help to make their decisions. ENR is a touch and web-based platform and it is composed by different screens. The data is grouped by the information provenance.

3.5. Automatic data processing and transformation (R6)

After obtain all the essential data to the decision making process it was necessary introduce new features to the transformation process. The uses of intelligent agents allowed automate the whole process. Now the tasks associated to data preparation process are performed automatically and in real-time without human effort. These changes increase the speed in getting new knowledge being they useful and available in the right time, i.e., in the moment of the decision is taken.

4. Pervasive Intelligent Decision Support System

A pervasive intelligent Decision Support System (PIDSS) is recognized as a system that helps the decision making process and it is accessible anywhere and anytime. In the health care arena there are two concepts related to PIDSS as is the pervasive healthcare and the pervasive computing (Varshney, 2009). Due their pervasive features, INTCare can produce three different types of knowledge. This knowledge is available anywhere and anytime.

4.1. ICU Medical Scores

The objective of PIDSS component is to behave as an Intelligent Scoring System (ISS). The ISS (Portela, et al., 2012) is incorporated into the Electronic Nursing Record (ENR). Nowadays, the ICU professionals can record and consult the scores in real-time.

This application allows for the automatic calculation in real-time of a set of scores: simplified acute physiology score (SAPS) II (Le Gall, Lemeshow, & Saulnier, 1993), SAPSIII (Metnitz et al., 2005), Sequential Organ Failure Assessment score (SOFA) (Vincent et al., 1998), Glasgow Coma Score (GSC) (Jones, 1979), Therapeutic Intervention Scoring System (TISS-28) (Reis Miranda, de Rijk, & Schaufeli, 1996) and Modified Early Warning Score (MEWS) (Gardner-Thorpe, Love, Wrightson, Walsh, & Keeling, 2006). At the same time it is possible to analyze the patient evolution in terms of the scores through interactive graphs (in a hourly and daily base).

As mentioned before (Portela, et al., 2012): this approach makes possible to provide a set of scores calculated / updated in real-time. The ISS proposed processes automatically the scores and adapts the results according to the new values collected generating new knowledge.

The main gains in using this approach can be summarized as:

- The data acquisition, scores calculation and results are made in real-time;
- All values are considered - no missing values;
- The data is displayed in a new way – real-time charts to compare trends;
- Less human intervention in the scores calculation – less errors;
- The scores are available anywhere and anytime;
- Help decision making process through a continuous scores monitoring.

4.2. ICU Critical Events

Critical Events (CE) are very important to the development of Data Mining (DM) models. In order to develop DM models in a real setting it was necessary to define procedures to automatically compute CE for five variables: Urine Output (Diuresis), Blood Pressure, Heart Rate, Respiratory and Temperature (P. G. Filipe Portela, Manuel Filipe Santos, Álvaro Silva, Fernando Rua, 2012).

The procedure calculates according some rules the number and elapsed time of an event. Then, the value is characterized as critical or not.

As result it is possible determine a number of critical events for the patient by hour and category (P. G. Filipe Portela, Manuel Filipe Santos, Álvaro Silva, Fernando Rua, 2012). In complement it is calculated the Accumulated Critical Events (ACE) (P. G. Filipe Portela, Manuel Filipe Santos, Álvaro Silva, Fernando Rua, 2012). The CE system is composed by a grid and a system similar to a traffic light. This system is used as a way to alert about the patient condition.

The grid shows: the number of critical events by hour, the number of ACE, the time in critical event by hour and the total time in critical event. The implementation of this new approach allows to the physicians have better understanding of the patient's condition.

4.3. Ensemble Based Models

Data Mining (DM) is the centre of the PIDSS. The objective of DM system is to predict the patient organ failure (cardiovascular, hepatic, coagulation, respiratory and renal) and patient outcome for the next hour. To achieve this goal an ensemble DM was developed. To evaluate the ensemble three measures were considered: Sensibility, Accuracy and Terror. For each measure the average of 10 runs was taken.

The selected models are used only if they satisfy the following conditions (quality measure): Total Error <= 40%; sensitivity >= 85% and Accuracy >= 60%. The use of ensemble helps to choose the best model in the cases where more than one model presents good results. From the six targets, only three satisfy the quality measures defined: outcome, cardiovascular and coagulation.

The low level of results verified in the other targets it is associated to the dynamic characteristics of the environment. Currently, it is possible induce DM models in real-time using online-learning and an ensemble approach in order to adapt the predictive models automatically. The doctors can use the predictions to save lives and avoid complicated situations to the patient.

5. Technology Acceptance questionnaires

For this study it was elaborated a questionnaire based on the four constructs of TAM 3. This questionnaire was elaborated by taking into account some scientific articles that report similar processes of technological implementation and are framed in the hospital environment and the first results obtained. It means they were aggregated into several groups to represent all the aspects of TAM. The main purpose is to obtain a better understanding about the user's intentions on the use of this system in the long run as well as the functionality for them. The questionnaire is composed by 96 questions. However, in this paper only were considered the questions related to the decision process. In this questionnaire was applied the Likert Scale (Johns, 2010) to evaluate the results.

This scale was chosen because the use of short scales (scales that goes between three and four) can better constrain results into close type of answers such as a simple yes or no; and secondly, by applying a higher scale this could fall into a dispersion of results that lead the answers to inaccurate results. As a consequence the chosen scale follows a range from one to five points similar to the used in the previous work / questionnaire. It allows for giving two values for each side and at the same time finding a neutrality point (Johns, 2010).

The considered levels are the following:

1) Not satisfies/in complete disagreement (< 20% of cases);
2) Satisfies a bit/in some level of disagreement (20-40%);
3) Satisfies/under some level of agreement (40-60%);
4) Satisfies a lot/strongly agreement (60-80%);
5) Satisfies completely/full agreement (> 80%).

The level of results collected from this questionnaire vary by the fact of the participant answer in a properly manner (with consciousness) or not. The answers always depend on the goodwill of each participant by answering in a balanced way to the questions of a certain group. This hypothesis does not verify when the participant evaluates a specific characteristic as a whole and gives the same answer to the group questions. To avoid this problem it was added to the questionnaire three screening questions to understand the level of the user's consciousness (e.g. 3+2). Table IX.1 crosses the questions with the constructs: Perceived Usefulness (PU); Perceived Ease of Use (PEOU); Behavioural Intention (BI); Use Behaviour (UB).

Table IX.1 – Questions of Functional Characteristics

Functional Characteristics		*PU*	*PEOU*	*BI*	*UB*
1.1	It allows the efficient registration of the information?		X		X
1.2	It allows obtaining efficient information for decision support?		X		X
1.3	It shows the prevision of Adverse Effects in an efficacy way?		X	X	
1.4	It shows usefulness when predicts the Scores?		X	X	
1.5	It improves the proactive performance of the professionals?	X	X	X	X
1.6	It allows tasks to be performed with greater precision?	X	X		

1.7	Can help to mitigate situations of an excessive workload?	X	X	X	X
1.8	Can allow a major control of several tasks?	X	X		X
1.9	Can help to have a better decision making based in best evidences?	X	X		X
1.10	Potentiates an improvement delivery of patient's health care?	X	X		X
1.11	It allows monitoring the patient's condition?	X	X		X
1.12	It promotes automating tasks?	X	X		X
1.13	It allows answering with the appropriate information to perform the task?	X	X	X	X
Technical Characteristics		*PU*	*PEOU*	*BI*	*UB*
2.1	Can promote quality of the information?	X		X	X
2.2	Can access to information quickly?	X			X
2.3	It allows access to information in a secure way?	X	X		X
2.4	Can operate simultaneously with other hospital systems?	X	X	X	X
2.5	Can facilitate an operation by having a tactile interface beside to patient's beds?	X	X	X	X
2.6	It allows an efficient use based on the available technical support?	X	X	X	X
2.7	**Evaluate the potential of each registration presented on the Intervention's panel:**				
2.71	Can facilitate obtaining information regarding the realized interventions?		X		
2.72	Can facilitate obtaining information regarding the therapeutic attitudes?		X		
2.73	Graphic aspect?		X		
2.74	Registration of the work plan?		X		
2.75	Utility of TISS28?	X	X	X	X
2.76	Graphic aspect of TISS28?	X	X	X	
2.77	Global Evaluation of TISS28	X	X	X	X
2.8	**Evaluate the potential of each registration presented on the Score's panel:**				
2.8.1	The records made automatically, present similar values relatively to the manuscripts	X	X		
2.8.2	Utility of SOFA CHART?	X	X	X	X
2.8.3	Utility of GLASGOW CHART?	X	X	X	X
2.8.4	Graphic aspect is intuitive?	X	X	X	X
2.8.5	The graphics can help to a better understanding of the real patient's condition?	X	X	X	X
2.8.6	By using the automation registration of Scores it facilitates the registration of SAPS II?	X		X	X
2.8.7	By using the automation registration of Scores it facilitates the registration of SAPS III?	X		X	X
2.8.8	By using the automation registration of Scores it facilitates the registration of	X		X	X
2.8.9	Global Evaluation of Scores?	X	X	X	X
2.9	**Evaluate the potential of each registration presented on the Vital Sign's panel:**				
2.9.1	Utility of information?	X	X	X	
2.9.2	Utility of consulting information (hourly, daily, continuous)?	X	X	X	
2.9.3	Graphic aspect?	X	X	X	
2.9.4	MEWS – Utility of system?	X	X	X	
2.9.5	Adverse events – Utility of system?	X		X	X
2.9.6	The early warning system for Adverse Events is useful?	X	X		X
2.9.7	Graphic aspect?		X		
2.9.8	Global evaluation of the vital signs?	X	X	X	X
2.10	**It is advantageous to use this system in intensive care units?**	X		X	X

6. Results

After collecting answers from 14 questionnaires sent by email (35% total number of nurses in ICU) an analysis of the results was performed. First a processing was done to avoid invalid or inconsistent answers given by the participants. Then was noticed that only one participant out of the 14 nurses answered the questionnaire in an inconsistent way for the proposed questions. This situation leaded to only consider 13 surveys. Table IX.2 presents the technology experience of the respondents.

Table IX.2 – Level of experience in Information Technology

Question	*Answer*	*Percentage*
What is your experience in technology – How much time do you spend at the computer	Less than 2 hours/day	0%
	Between 2 to 4 hours/day	57%
	More 4 hours/day	36%
Type of User?	Full Autonomy	62%
	Rarely need technical support (less than 3 times/month)	38%
	Need regular technical support	0%
Uses computer preferably for?	Application of production staff (email, text processing, spread Sheet)	62%
	Handling/Consulting administrative information	31%
	Handling/Consulting clinical information	77%
	Handling/Consult management information	8%

6.1. Respondent Analysis

For a better perception of the answers made by each respondent, one analysis (average and mode) was carried out by the person questioned and TAM construct (figure IX.1 to IX.4). In this analysis, the persons (X axis) are represented by numbers (1 to 13).

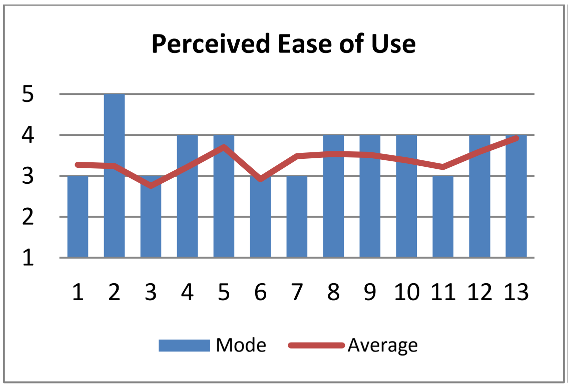


Figure IX.1 – Evaluation of PU.

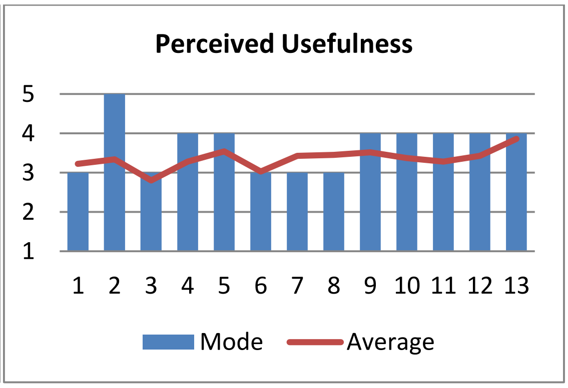


Figure IX.2 – Evaluation of PEOU.

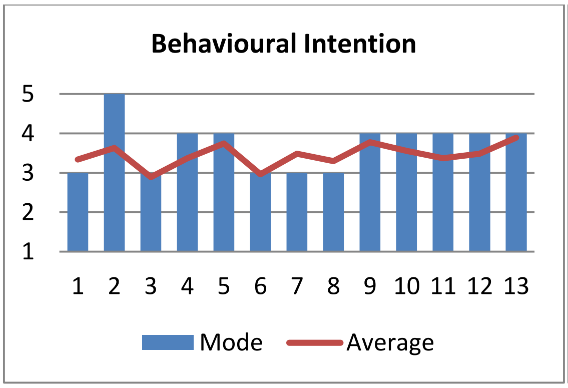


Figure IX.3 – Evaluation of BI.

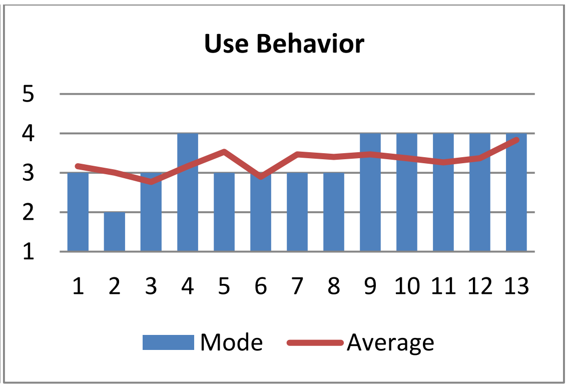


Figure IX 4 – Evaluation of UB.

At a high level of analysis, in figures IX.1, IX.2 and IX.3, is possible to observe that the second person answered most of the questions with 5 values; however the average is between 3 and 4 points. This means that this person is quite pleased with some aspects of the system and not with others. In general the evaluations are above 3 points. At same time some correlations techniques were used to understand if the users are in accordance with the answers. In a global way they are in relative accordance in some of the questions, being the overall Kendall's tau: 0,158224.

6.2. Question Analysis

In this sub-section, instead of doing an analysis by respondent, an analysis was made for each one of the question and TAM 3 construct. In the Y axis they are the possible answers of the questionnaire (1-5) and in the X axis they are the questions numbers.

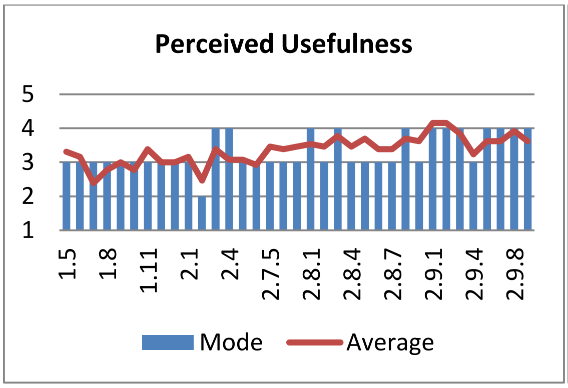


Figure IX.5 – Analysis for questions (PU).

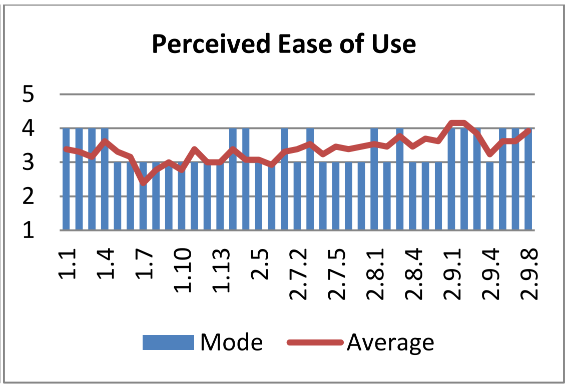


Figure IX.6 – Analysis for questions (PEOU).

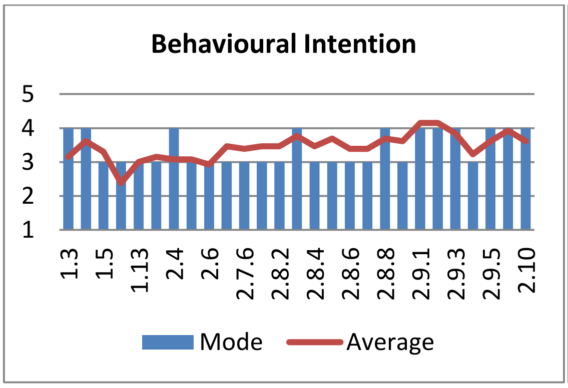


Figure IX.7 – Analysis for questions (BI).

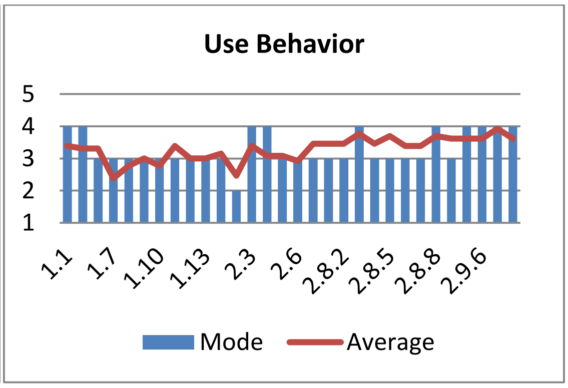


Figure IX.8 – Analysis for questions (UB).

In figure IX.5 and IX.8 it should be stressed that question 2.2 has a lowest score. On the other hand, the average of the answers for the questions related to these constructs is situated between 2 and 4 points. This result means that sometimes it is difficult to have access to the data. This happens due to hospital connectivity problems in the network. This problem represents the biggest barrier to the success of INTCare. In Figure IX.6 and IX.7, the vast majority of the answers of the questions to this constructs were stood by 3 points.

6.3. Global Analysis by question

A global analysis was done in order to understand the best features, the average, the mode and the standard deviation (Stdev) for each one of constructs. Table IX.3 offers a quick view of the answers obtained by the questionnaires. This table shows that the ICU staffs are satisfied with the system. All the constructs present positive results being the best, the Perceived Ease of Use and the worst, the Use behaviour.

Table IX.3 – Summary of mode and average for each construct and analysis overall

	PU	*PEOU*	*BI*	*UB*	*Overall*
Mode	3	3	3	3	3
Average	3,35	3,67	3,44	3,28	3,36
Stdev	0,18	0,21	0,20	0,21	0,21
Min	1	1	1	1	1
Max	5	5	5	5	5

Table IX.4 offers a quick understanding of each construct. This table presents the three best and the three worst results. In the positive side they are some information (charts) to the decision process and the importance that the graphics have in order to have a better understand of the patient condition. This characteristic has as average (avg) 4.15. In the opposite side they are some functional characteristics. It is possible to observe that there are two negative characteristics (1.7 and 2.2). Both of them are related to the system speed. Despite all that, only one question has as mode 2 points. All of the other questions have a mode between the 3 and 4 points.

Table IX.4 – The three best and worst results

	PU	*Mode*	*Avg*	*PEOU*	*Mode*	*Avg*	*BI*	*Mode*	*Avg*	*UB*	*Mode*	*Avg*
Best Results	2.9.1	4	**4,15**	2.9.1	4	**4,15**	2.9.1	4	**4,15**	2.9.8	4	3,92
	2.9.2	4	**4,15**	2.9.2	4	**4,15**	2.9.2	4	**4,15**	2.8.3	4	3,77
	2.9.8	4	3,92	2.9.8	4	3,92	2.9.8	4	3,92	2.8.5	3	3,69
Worst Results	1.7	3	2,38	1.7	3	2,38	1.7	3	2,38	1.7	3	2,38
	2.2	2	2,46	1.8	3	2,77	2.6	3	2,92	1.8	3	2,77
	1.8	3	2,77	1.10	3	2,77	1.13	3	3	1.10	3	2,77

7. Conclusions and Future Work

First, it should be noted that the initial objective proposed regarding to the junction of the Technology Acceptance Model (TAM 3) with the Delphi's method in order to evaluate the acceptance by users, their perceptions and the impact on the INTCare's system usage behaviour, it is totally innovative and can be considered a success. This is the first approach to assess the impact of such type of solutions in ICU environment.

In order to understand the acceptance it is possible to conclude that the ICU staff it is very comfortable with the system INTCare. They pointed the data access as the biggest problem and the utility of the information generated for the decision process as the biggest gain. The user acceptance was very positive (average upper than 3 points) for the four constructs assessed: Perceived Usefulness, Perceived ease of use, Behavioral Intention and Usage Behaviour.

Concluding, the ICU professionals are receptive to the INTCare system and to the new knowledge provided because it can help them in the decision making process.

The results obtained by the questionnaires allow for some concluding remarks:

- ✓ These results encourage further development and optimization of the solutions designed, as well as a deeper assessment of all the resources available;
- ✓ It is required an improvement of physical resources (e.g. memory, hardware) of the ICU.

Acknowledgements

This work is supported by FEDER through Operational Program for Competitiveness Factors – COMPETE and by national funds though FCT – Fundação para a Ciência e Tecnologia in the scope of the project: FCOMP-01-0124-FEDER-022674.

The authors would like to thank FCT (Foundation of Science and Technology, Portugal) for the financial support through the contract PTDC/EIA/72819/ 2006. The work of Filipe Portela was supported by the grant SFRH/BD/70156/2010 from FCT.

References

Baggs, J. G., Norton, S. A., Schmitt, M. H., Dombeck, M. T., Setters, C. R., & Quinn, J. R. (2007). Intensive care unit cultures and end-of-life decision making. [Article]. *Journal of Critical Care, 22*(2), 159-168.

Chooprayoon, V., & Fung, C. C. (2010). TECTAM: An Approach to Study Technology Acceptance Model (TAM) in Gaining Knowledge on the Adoption and Use of E-Commerce/E-Business Technology among Small and Medium Enterprises in Thailand.

Filipe Portela, F. P., Manuel Filipe Santos. (2012). *Data Mining Predictive Models For Pervasive Intelligent Decision Support In Intensive Care Medicine*. Paper presented at the KMIS 2012 Filipe Portela, M. F. S., José Machado, António Abelha, José Neves, Álvaro Silva, Fernando Rua. (2012). *Intelligent Decision Support in Intensive Care – Towards Technology Acceptance*. Paper presented at the 26th European Simulation and Modelling Conference - ESM'2012.

Filipe Portela, P. G., Manuel Filipe Santos, Álvaro Silva, Fernando Rua. (2012). *Intelligent and Real-time Data Acquisition and Evaluation to Determine Critical Events in Intensive Medicine.* Paper presented at the *HCist 2012* -

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: An introduction to theory and research*: Addison-Wesley.

Gago, P., Santos, M. F., Silva, Á., Cortez, P., Neves, J., & Gomes, L. (2006). INTCare: a knowledge discovery based intelligent decision support system for intensive care medicine. *Journal of Decision Systems*.

Gardner-Thorpe, J., Love, N., Wrightson, J., Walsh, S., & Keeling, N. (2006). The value of Modified Early Warning Score (MEWS) in surgical in-patients: a prospective observational study. *Annals of The Royal College of Surgeons of England, 88*(6), 571.

Johns, R. (2010). Likert Items and Scales. *Survey Question Bank: Methods Fact Sheet, 1*.

Jones, C. (1979). Glasgow coma scale. *AJN The American Journal of Nursing, 79*(9), 1551.

Le Gall, J. R., Lemeshow, S., & Saulnier, F. (1993). A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA, 270*(24), 2957-2963.

Mador, R. L., & Shaw, N. T. (2009). The impact of a Critical Care Information System (CCIS) on time spent charting and in direct patient care by staff in the ICU: a review of the literature. *International Journal of Medical Informatics, 78*(7), 435-445.

Metnitz, P. G. H., Moreno, R. P., Almeida, E., Jordan, B., Bauer, P., Campos, R. A., et al. (2005). SAPS 3—From evaluation of the patient to evaluation of the intensive care unit. Part 1: Objectives, methods and cohort description. *Intensive care medicine, 31*(10), 1336-1344.

Portela, F., Gago, P., Santos, M. F., Silva, A., Rua, F., Machado, J., et al. (2011). *Knowledge Discovery for Pervasive and Real-Time Intelligent Decision Support in Intensive Care Medicine*. Paper presented at the *KMIS* 2011- International Conference on Knowledge Management and Information Sharing.

Portela, F., Santos, M., Vilas-Boas, M., Rua, F., Silva, Á., & Neves, J. (2010). *Real-time Intelligent decision support in intensive medicine*. Paper presented at the *KMIS* 2010- International Conference on Knowledge Management and Information Sharing.

Portela, F., Santos, M. F., Machado, J., Silva, Á., Rua, F., & Abelha, A. (2012). Intelligent Data Acquisition and Scoring System for Intensive Medicine. In Springer (Ed.), *Lecture Notes in Computer Science - Information Technology in Bio- and Medical Informatics* (Vol. 7451/2012, pp. 1-15). Viena, Austria.

Portela, F., Santos, M. F., Silva, Á., Machado, J., & Abelha, A. (2011). *Enabling a Pervasive approach for Intelligent Decision Support in Critical Health Care*. Paper presented at the *HCist 2011* – International Workshop on Health and Social Care Information Systems and Technologies.

Reis Miranda, D., de Rijk, A., & Schaufeli, W. (1996). Simplified Therapeutic Intervention Scoring System: the TISS-28 items–results from a multicenter study. *Critical care medicine, 24*(1), 64.

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., & Neves, J. (2011). *INTCARE - Multi-agent approach for real-time Intelligent Decision Support in Intensive Medicine.* Paper presented at the 3rd International Conference on Agents and Artificial Intelligence (*ICAART*), Rome, Italy.

Varshney, U. (2009). *Pervasive Healthcare Computing: EMR/EHR, Wireless and Health Monitoring*: Springer-Verlag New York Inc.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences, 39*(2), 273-315.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, 186-204.

Vincent, J., Mendonça, A., Cantraine, F., Moreno, R., Takala, J., Suter, P., et al. (1998). 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, 26*, 1793-1800.

Zackiewicz, M., & Salles Filho, S. (2010). Technological Foresight–Um instrumento para política científica e tecnológica. *Parcerias estratégicas, 6*(10).

CHAPTER VI – Conclusions

1. Context

The introduction of new technologies in Intensive Care Units (ICU) increased significantly the information (knowledge) available at the moment of the decision. Nowadays is possible to process the data automatically and make use of the online learning in order to obtain tree distinct types of knowledge in real-time: Critical Events, Medical Scores, Probability of occur an organ failure or patient die. These achievements are based in streamed data (Gama & Gaber, 2007), i.e., data collected automatically and in real-time, the ENR, the AIDA(EHR) (Abelha, et al., 2003) and the hospital interoperability among systems (Brailer, 2005; Häyrinen, et al., 2008).

The implementation of this new approach allows the doctors, in ICU of Centro Hospital do Porto, to have a better understanding of the patient's condition, in the moment of the decision is made. The doctors can consult, for each patient and in real-time, the number of events, scores and predictions.

Information about events is diffused hourly. In addition, they have a kind of traffic light system (green, yellow, red) that can show the current situation of each event. At the same time, medical scores are available up-to-date, i.e., whenever some new values are collected, the scoring system is executed in order to provide the results. Complementarily, a set of predictions on the failure/dysfunction of organic and the outcome the patients are available in real time to support the medical decision.

2. Main Results

Main results have been presented in terms of a list of articles published in journals, proceedings of conferences and as book chapters. The list contains the nine most significant publications chosen from a list of 37 articles published in the context of this work. Articles were grouped according to the objectives defined (appendix b). Table 3 enumerates the most important results attained for each objective.

Table 3 – Main objectives results of the select articles

Paper	Main Objective	Main Results
A1 – A Pervasive Approach to a Real-Time Intelligent Decision Support System in Intensive Medicine	To understand the implications of a pervasive approach to Intelligent Decision Support Systems (IDSS) in the context of critical environments	New information system architecture; First overview of the pervasive approach needs and the results achieved.
A2 – Enabling a Pervasive Approach for Intelligent Decision Support in Critical Health Care	To adopt pervasive approaches in critical environments	Environment Features and a Framework to evaluate the Environment Quality
A3 – Enabling Real-Time Intelligent Decision Support in Intensive Care	To adopt pervasive approaches in critical environments	Main features and description of the KDD process. Explanation how the information can be automated and dematerialized.
A4 – Towards Pervasive and Intelligent Decision Support in Intensive Medicine – A Data Stream Mining Approach	To develop and test pervasive models in Intensive Care	Ensemble Data Mining models to predict the probability of an event occurring in the next 24 hours: Patient Organ failure and patient dead.
A5 – Pervasive Real-Time Intelligent System for Tracking Critical Events in Intensive Care Patients	To develop and test pervasive models in Intensive Care	Pervasive Intelligent Tracking system to calculate the number and the duration of the critical events: Urine Output; Blood Pressure; Spo2; Temperature; Heart Rate.
A6 – Intelligent Data Acquisition and Scoring System for Intensive Medicine	To develop and test pervasive models in Intensive Care	Pervasive Intelligent Scoring System to calculate Saps II and SAPS III; SOFA; GLASGOW; TISS 28; MEWS;
A7 – Intelligent Decision Support in Intensive Care – Towards Technology Acceptance	To develop and test pervasive models in Intensive Care	Global overview of the solution. Results of the first technology assessment.
A8 – Real-Time Decision Support in Intensive Medicine - An Intelligent Approach for Monitoring Data Quality	To develop and test pervasive models in Intensive Care	Data quality system architecture and obtained results with the introduction of the quality system
A9 – Pervasive Intelligent Decision Support System - Technology Acceptance in Intensive Care Units	To develop and test pervasive models in Intensive Care	A questionnaire to evaluate the technologies acceptance in critical environments. Results of the system assessment having in account the four constructs of TAM3

3. Final remarks

This work has been developed under the research project INTCare (Portela, F., Santos, M. F., & Vilas-Boas, M. (2012).). Being an applied research project, the results were tested and deployed in the ICU of CHP.

Morik (Katharina Morik, et al., 2000) postulated in 2000 that was not possible in a near future to collect data in real-time: "Most variables are entered by hand at the bedside. For entities such as clinical observations, nursing procedures, therapeutic measures, medications, or orders it appears very unlikely that entry of these variables can be automated in the foreseeable future". After a decade, the results of this work refute partially the Morik premises proving that data collected automatically in real-time from vital signs monitors can be used to induce accurate prediction models.

Online learning was successfully implemented to adapt in real-time an ensemble of prediction models. Intelligent agents are now responsible to maintain up-to-date models to predict the organ failure probability and outcome for the next 24 hours.

The last lines are reserved to philosophical concerns associated to the treats of technology adoption. Human are always responsible for the ultimate decisions, because nobody knows the patient better than physicians or nurses. INTCare is faded to support medical judgment on the clinical condition of the patients, never substituting the professionals in the decision making.

CHAPTER VII – Future Work

Future research will include data collected from the bedside monitors and ventilators, combined with data from the drugs system, electronic health record, electronic nursing record and lab system to extend the actual capabilities of INTCare system in order to:

1. Provide intelligent decision support on therapies and procedures by means of
    a. Data mining decision models for therapies;
    b. Data mining decision models for procedures;
    c. Models to optimize the costs and outcome;
2. Extend to other intensive care units.

CHAPTER VIII - References

Abelha, A., Machado, J., Alves, V., & Neves, J. (2004). Data warehousing through multi-agent systems in the medical arena.

Abelha, A., Machado, M., Santos, M., Sollari, A., Rua, F., Paiva, M., et al. (2003). Agency for Archive, Integration and Diffusion of Medical Information. Proceeding of AIA.

Aguilar, A., van der Putten, W., & Kirrane, F. (2006). Positive patient identification using RFID and wireless networks.

Andrade, P. J. N. (1999). Specialized computer support systems for medical diagnosis. Relationship with the Bayes' theorem and with logical diagnostic thinking. Arquivos Brasileiros de Cardiologia, 73, 545-552.

Apostolakos, M. J., & Papadakos, P. J. (2001). The Intensive Care Manual: McGraw-Hill Professional.

Arnott, D., & Pervan, G. (2004, Jul). A critical analysis of decision support systems research. Paper presented at the Conference on Decision Support Systems, Prato, ITALY.

Aubuchon, V. (2011). Blood Pressure Chart - Normal Blood Pressure Range. Medicine Summaries Heart Health Retrieved 07-11-2012, 2012, from http://www.vaughns-1-pagers.com/medicine/blood-pressure.htm

Augusto, J. C. (2005). Temporal reasoning for decision support in medicine. Artificial Intelligence in Medicine, 33(1), 1-24.

Baggs, J. G., Norton, S. A., Schmitt, M. H., Dombeck, M. T., Setters, C. R., & Quinn, J. R. (2007). Intensive care unit cultures and end-of-life decision making. [Article]. Journal of Critical Care, 22(2), 159-168.

Banavar, G., Beck, J., Gluzberg, E., Munson, J., Sussman, J., & Zukowski, D. (2000, 2000). Challenges: an application model for pervasive computing.

Belal, S. Y., Taktak, A. F. G., Nevill, A., & Spencer, A. (2005). An intelligent ventilation and oxygenation management system in neonatal intensive care using fuzzy trend template fitting. Physiological Measurement, 26(4), 555.

Bendavid, Y., Wamba, S. F., & Lefebvre, L. A. (2006, 2006). Proof of concept of an RFID-enabled supply chain in a B2B e-commerce environment.

Bergs, E. A. G., Rutten, F. L. P. A., Tadros, T., Krijnen, P., & Schipper, I. B. (2005). Communication during trauma resuscitation: do we know what is happening? Injury, 36(8), 905-911.

Berner, E. S., & La Lande, T. J. (2007). Clinical Decision Support Systems-Theory and Practice, Overview of Clinical Decision Support Systems: Spinger.

Black, J. P., Segmuller, W., Cohen, N., Leiba, B., Misra, A., Ebling, M. R., et al. (2004). Pervasive Computing in Health Care: Smart Spaces and Enterprise Information Systems.

Bradbury, H., & Reason, P. (2007). The SAGE handbook of action research: Participative inquiry and practice: Sage Publications Limited.

Brailer, D. J. (2005). Interoperability: the key to the future health care system. HEALTH AFFAIRS-MILLWOOD VA THEN BETHESDA MA-, 24, 5.

Brenck, F., Hartmann, B., Mogk, M., & Junger, A. (2008). [Scoring systems for daily assessment in intensive care medicine. Overview, current possibilities and demands on new developments]. Der Anaesthesist, 57(2), 189.

Bricon-Souf, N., & Newman, C. R. (2007). Context awareness in health care: A review. [10.1016/j.ijmedinf.2006.01.003]. International Journal of Medical Informatics, 76(1), 2-12.

Carrault, G., Cordier, M. O., Quiniou, R., & Wang, F. (2003). Temporal abstraction and inductive logic programming for arrhythmia recognition from electrocardiograms. Artificial Intelligence in Medicine, 28(3), 231-263.

Cassi, A. (2009). Despite Wide Adoption, Little Evidence of eICU Impact on Intensive Care Cost/Quality. Health Affairs.

Chooprayoon, V., & Fung, C. C. (2010). TECTAM: An Approach to Study Technology Acceptance Model (TAM) in Gaining Knowledge on the Adoption and Use of E-Commerce/E-Business Technology among Small and Medium Enterprises in Thailand.

Chowdhury, B., & Khosla, R. (2007). RFID-based hospital real-time patient management system.

Clemmer, T. P. (2004). Computers in the ICU: Where we started and where we are now. Journal of Critical Care, 19(4), 201-207.

Coiera, E. (2000). When conversation is better than computation. Journal of the American Medical Informatics Association, 7(3), 277.

Coiera, E. (2003). Guide to health informatics: Arnold.

Cook, D. J., & Das, S. K. (2005). Smart environments: technologies, protocols, and applications: Wiley-Interscience.

Coombs, M. (2003). Power and conflict in intensive care clinical decision making. Intensive & Critical Care Nursing, 19(3), 125-135.

Das, S. K., & Roy, N. (2008). Learning, Prediction and Mediation of Context Uncertainty in Smart Pervasive Environments. In R. Meersman, Z. Tari & P. Herrero (Eds.), On the Move to Meaningful Internet Systems: Otm 2008 Workshops (Vol. 5333, pp. 820-829). Berlin: Springer-Verlag Berlin.

Dasgupta, S., & Vankayala, V. K. (2007, 2007). Developing Real-time Business Intelligence Systems: The Agile Way.

Dietterich, T. (2000). Ensemble methods in machine learning. Multiple classifier systems, 1-15.

Duarte, J., Portela, F., Santos, M. F., António, A., & José, M. (2011). Electronic Health Record in Dermatology Service. Paper presented at the HCist 2011 – International Workshop on Health and Social Care Information Systems and Technologies, Algarve, Portugal.

Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery: an overview.

Filipe Portela, F. P., Manuel Filipe Santos. (2012). Data Mining Predictive Models For Pervasive Intelligent Decision Support In Intensive Care Medicine. Paper presented at the KMIS 2012 - nternational Conference on Knowledge Management and Information Sharing.

Filipe Portela, M. F. S., José Machado, António Abelha, José Neves, Álvaro Silva, Fernando Rua. (2012). Intelligent Decision Support in Intensive Care – Towards Technology Acceptance. Paper presented at the 26th European Simulation and Modelling Conference - ESM'2012.

Filipe Portela, P. G., Manuel Filipe Santos, Álvaro Silva, Fernando Rua. (2012). Intelligent and Real-time Data Acquisition and Evaluation to Determine Critical Events in Intensive Medicine. Paper presented at the HCist'2012 - International Conference on Health and Social Care Information Systems and Technologies, Portugal.

Finlay, P. N. (1994). Introducing Decision Support Systems (2nd ed.): NCC Blackwell.

Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention and behaviour: An introduction to theory and research: Addison-Wesley.

Fonseca, T., Ribeiro, C., & Granja, C. (2009). Vital Signs in Intensive Care: Automatic Acquisition and Consolidation into Electronic Patient Records. Journal of Medical Systems, 33(1), 47-57.

Foster, D., McGregor, C., & El-Masri, S. (2005, 2005). A survey of agent-based intelligent decision support systems to support clinical management and research.

Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992). Knowledge Discovery in Databases: An Overview. AI Magazine, 13(3), 57-70.

Gaber, M. M., Krishnaswamy, S., & Zaslavsky, A. (2004). Ubiquitous data stream mining.

Gago, P., & Santos, M. (2009). Evaluating Hybrid Ensembles for Intelligent Decision Support for Intensive Care. Applications of Supervised and Unsupervised Ensemble Methods, 251-265.

Gago, P., & Santos, M. F. (2008). Towards an Intelligent Decision Support System for Intensive Care Units. Paper presented at the 18th European Conference on Artificial Intelligence.

Gago, P., & Santos, M. F. (2009). Adaptive knowledge discovery for decision support in intensive care units. WSEAS Transactions on Computers, 8(7), 1103-1112.

Gago, P., Santos, M. F., Silva, Á., Cortez, P., Neves, J., & Gomes, L. (2006). INTCare: a knowledge discovery based intelligent decision support system for intensive care medicine. Journal of Decision Systems.

Gago, P., Silva, A., & Santos, M. F. (2007, Dec 03-07). Adaptive decision support for intensive care. Paper presented at the 13th Portuguese Conference on Artificial Intelligence, Guimaraes, PORTUGAL.

Gama, J., & Gaber, M. M. (2007). Learning from data streams: processing techniques in sensor networks: Springer-Verlag New York Inc.

Gao, S., & Xu, D. (2009). Conceptual modeling and development of an intelligent agent-assisted decision support system for anti-money laundering. Expert Systems with Applications, 36(2, Part 1), 1493-1504.

Gardner, R. M., Hawley, W. L., East, T. D., Oniki, T. A., & Young, H. F. W. (1991). Real-time data acquisition: recommendations for the Medical Information Bus (MIB). International journal of clinical monitoring and computing, 8(4), 251-258.

Gardner-Thorpe, J., Love, N., Wrightson, J., Walsh, S., & Keeling, N. (2006). The value of Modified Early Warning Score (MEWS) in surgical in-patients: a prospective observational study. Annals of The Royal College of Surgeons of England, 88(6), 571.

Gillham, B. (2000). Case study research methods: Continuum Intl Pub Group.

Giudici, P. (2003). Applied data mining: Statistical methods for business and industry: John Wiley & Sons, Inc.

Goñi, A., Burgos, A., Dranca, L., Rodríguez, J., Illarramendi, A., & Bermúdez, J. (2009). Architecture, cost-model and customization of real-time monitoring systems based on mobile biological sensor data-streams. Computer Methods and Programs in Biomedicine, 96(2), 141-157.

Hall, J. B., Schmidt, G. A., & Wood, L. D. H. (2005). Principles of Critical Care: McGraw-Hill's AccessMedicine.

Henricksen, K., Indulska, J., & Rakotonirainy, A. (2001, 2001). Infrastructure for pervasive computing: Challenges.

Hooda, J. S., Dogdu, E., & Sunderraman, R. (2004). Health Level-7 compliant clinical patient records system.

Horovitz, O., Gaber, M. M., & Krishnaswamy, S. (2005). Making sense of ubiquitous data streams - A fuzzy logic approach. In R. Khosla, R. J. Howlett & L. C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Pt 2, Proceedings (Vol. 3682, pp. 922-928). Berlin: Springer-Verlag Berlin.

Häyrinen, K., Saranto, K., & Nykänen, P. (2008). Definition, structure, content, use and impacts of electronic health records: A review of the research literature. International Journal of Medical Informatics, 77(5), 291-304.

Jennings, N. R. (2000). On agent-based software engineering. Artificial Intelligence, 117(2), 277-296.

Jones, C. (1979). Glasgow coma scale. AJN The American Journal of Nursing, 79(9), 1551.

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. Judgment under uncertainty: Heuristics and biases, 201.

Kalli, S., Ambroso, C., Gregory, R., Heikelä, A., Ilomäki, A., Leaning, M., et al. (1993). Inform: Conceptual modelling of intensive care information systems. Journal of Clinical Monitoring and Computing, 9(2), 85-94.

Kaur, M., Pawar, M., Kohli, J. K., & Mishra, S. (2008). Critical events in intensive care unit. Indian Journal of Critical Care Medicine: Peer-reviewed, Official Publication of Indian Society of Critical Care Medicine, 12(1), 28.

Keegan, M. T., Gajic, O., & Afessa, B. (2011). Severity of illness scoring systems in the intensive care unit. Critical care medicine, 39(1), 163.

Kivijarvi, H., & Tuominen, M. (1999). Computer based intelligence, design, choice, implementation, and control of intangible investments projects. Paper presented at the 32nd Annual Hawaii International Conference on System Sciences.

Kohn, L. T., Corrigan, J., & Donaldson, M. S. (2000). To Err Is Human: Building a Safer Health System: National Academy Press.

Kuechler, B., & Vaishnavi, V. (2008). On theory development in design science research: anatomy of a research project. European Journal of Information Systems, 17(5), 489-504.

Kwon, O., Yoo, K., & Suh, E. (2005). UbiDSS: a proactive intelligent decision support system as an expert system deploying ubiquitous computing technologies. Expert Systems with Applications, 28(1), 149-161.

Lapinsky, S., Holt, D., Hallett, D., Abdolell, M., & Adhikari, N. (2008). Survey of information technology in Intensive care units in Ontario, Canada. BMC medical informatics and decision making, 8(1), 5.

Larsson, J. E., & Hayes-Roth, B. (1998). Guardian: An Intelligent Autonomous Agent for Medical Monitoring and Diagnosis. IEEE Intelligent Systems, 58-64.

Lavrakas, P. J. (2008). Encyclopedia of survey research methods: Sage Publications, Inc.

Le Gall, J. R., Lemeshow, S., & Saulnier, F. (1993). A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. JAMA, 270(24), 2957-2963.

Lee, P. (2012). Design Research: What Is It and Why Do It? http://thereboot.org/blog/2012/02/19/design-research-what-is-it-and-why-do-it/

Lindgaard, G., Pyper, C., Frize, M., & Walker, R. (2009). Does Bayes have it? Decision Support Systems in diagnostic medicine. International Journal of Industrial Ergonomics, 39(3), 524-532.

Lourenco, A., & Belo, O. (2003). Promoting agent-based knowledge discovery in medical intensive care units. WSEAS Transactions on Computers, 2(2), 403-408.

Lunenfeld, P., & Laurel, B. (2003). Design research: Methods and perspectives: MIT press.

Lupiana, D., O'Driscoll, C., & Mtenzi, F. (2009, 28-31 July 2009). Taxonomy for ubiquitous computing environments. Paper presented at the Networked Digital Technologies, 2009. NDT '09. First International Conference on.

Machado, J. A., António. Santos, Manuel, Neves, José. (2006). Multi-agent based Problem Solving in Medical Decision Support Systems. In R. C. a. F. L. Vale Z. (Ed.), Decision Support, Knowledge and Decision Technologies. Porto: IPP, ISEP.

Mador, R. L., & Shaw, N. T. (2009). The impact of a Critical Care Information System (CCIS) on time spent charting and in direct patient care by staff in the ICU: a review of the literature. International Journal of Medical Informatics, 78(7), 435-445.

Mahmoud, M. (2003). Real-time data acquisition system for monitoring patients in intensive care unit. (Vol. 5090, pp. 320-326). Multisensor, Multisource Information Fusion: Architectures, Algorithms and Applications.

Makeham M, D. S., County M, Kidd M. (2002). An international taxonomy for errors in general practice: a pilot study. The Medical Journal of Australia, 177(2), 68-72.

Marreiros, M. G. (2007). Agentes de Apoio à Argumentação e Decisão em Grupo. Universidade do Minho.

McMillen, R. E. (2008). End of life decisions: Nurses perceptions, feelings and experiences. Intensive and Critical Care Nursing, 24(4), 251-259.

Metnitz, P. G. H., Moreno, R. P., Almeida, E., Jordan, B., Bauer, P., Campos, R. A., et al. (2005). SAPS 3—From evaluation of the patient to evaluation of the intensive care unit. Part 1: Objectives, methods and cohort description. Intensive care medicine, 31(10), 1336-1344.

Michalewicz, Z., Schmidt, M., Michalewicz, M., & Chiriac, C. (2005). Case Study: An Intelligent Decision-Support System. IEEE Intelligent Systems, 44-49.

Michalewicz, Z., Schmidt, M., Michalewicz, M., & Chiriac, C. (2007). Adaptive Business Intelligence: Springer.

Mikkonen, M., Va¨ yrynen, S., Ikonen, V., & Heikkila, M. O. (2002). User and concept studies as tools in developing mobile communication services for the elderly. Personal and Ubiquitous Computing, 6(2), 113-124.

Miranda, M., Abelha, A., Santos, M., Machado, J., & Neves, J. (2009). A group decision support system for staging of cancer. Electronic Healthcare, 0001, 114-121.

Mitchell, M. L., & Jolley, J. M. (2012). Research design explained: Wadsworth Publishing Company.

Moreno, R. P., Metnitz, P. G. H., Almeida, E., Jordan, B., Bauer, P., Campos, R. A., et al. (2005). SAPS 3—From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission. Intensive care medicine, 31(10), 1345-1355.

Morik, K. (2003). Data analysis and knowledge validation in intensive care monitoring.

Morik, K., Brockhausen, P., & Joachims, T. (1999). Combining statistical learning with a knowledge-based approach-a case study in intensive care monitoring.

Morik, K., Imboff, M., Brockhausen, P., Joachims, T., & Gather, U. (2000). Knowledge discovery and knowledge validation in intensive care. Artificial Intelligence in Medicine, 19(3), 225-249.

Negash, S., & Gray, P. Business intelligence. Handbook on Decision Support Systems 2, 175-193.

O'Donoghue, J., Herbert, J., & Sammon, D. (2008). Patient sensors: A data quality perspective. In S. Helal, S. Mitra, J. Wong, C. Chang & M. Mokhtari (Eds.), Smart Homes and Health Telematics (Vol. 5120, pp. 54-61). Berlin: Springer-Verlag Berlin.

Orwat, C., Graefe, A., & Faulwasser, T. (2008). Towards pervasive computing in health care - A literature review. [10.1186/1472-6947-8-26]. BMC Medical Informatics and Decision Making, 8(1), 26.

Pereira, M., Curra, A., Rivas, R., Pereira, J., Banos, G., Teueiro, J., et al. (2007). Computer aided monitoring system of intensive care unit patients. WSEAS Transactions on Information Science and Applications, 4(1), 78-84.

Portela, F. (2009). Sistemas de Apoio à Decisão na Medicina Intensiva Baseados na Descoberta de Conhecimento em Bases de Dados. Universidade do Minho, Guimarães.

Portela, F., Gago, P., Santos, M. F., Silva, A., Rua, F., Machado, J., et al. (2011). Knowledge Discovery for Pervasive and Real-Time Intelligent Decision Support in Intensive Care Medicine. Paper presented at the KMIS 2011-International Conference on Knowledge Management and Information Sharing.

Portela, F., Santos, M., Vilas-Boas, M., Rua, F., Silva, Á., & Neves, J. (2010). Real-time Intelligent decision support in intensive medicine. Paper presented at the KMIS 2010Portela, F., Santos, M. F., Gago, P., Silva, Á., Rua, F., Abelha, A., et al. (2011b). Enabling real-time intelligent decision support in intensive care. Paper presented at the 25th European Simulation and Modelling Conference- ESM'2011, Guimarães, Portugal.

Portela, F., Santos, M. F., Machado, J., Silva, Á., Rua, F., & Abelha, A. (2012). Intelligent Data Acquisition and Scoring System for Intensive Medicine. In Springer (Ed.), Lecture Notes in Computer Science - Information Technology in Bio- and Medical Informatics (Vol. 7451/2012, pp. 1-15). Viena, Austria.

Portela, F., Santos, M. F., Silva, Á., Machado, J., & Abelha, A. (2011). Enabling a Pervasive approach for Intelligent Decision Support in Critical Health Care. Paper presented at the HCist 2011 – International Workshop on Health and Social Care Information Systems and Technologies.

Portela, F., Santos, M. F., & Vilas-Boas, M. (2012). A Pervasive Approach to a Real-Time Intelligent Decision Support System in Intensive Medicine. In Springer (Ed.), Communications in Computer and Information Science (Vol. 0272, pp. 14).

Portela, F., Vilas-Boas, M., Santos, M. F., & Fernando, R. (2010). Improvements in data quality for decision support in Intensive Care. Paper presented at the eHealth 2010 - 3rd International ICST Conference on Electronic Healthcare for the 21st century.

Ramon, J., Fierens, D., Güiza, F., Meyfroidt, G., Blockeel, H., Bruynooghe, M., et al. (2007). Mining data from intensive care patients. Advanced Engineering Informatics, 21(3), 243-256.

Rao, S. M., & T., S. (2003). Organization of intensive care unit and predicting outcome of critical illness. Indian J. Anaesth, 47 (5), 328-337.

Reason, P., & Bradbury, H. (2006). Handbook of action research.

Rothschild, J. M., Landrigan, C. P., Cronin, J. W., Kaushal, R., Lockley, S. W., Burdick, E., et al. (2005). The Critical Care Safety Study: The incidence and nature of adverse events and serious medical errors in intensive care*. Critical care medicine, 33(8), 1694.

Rui Rodrigues, P. G., Miguel Miranda, Carlos Portela, Manuel Santos, José Neves, António Abelha, José Machado. (2012). Monitoring Intelligent System for the Intensive Care Unit using RFID and Multi-Agent Systems. Paper presented at the IEEE International Conference on Industrial Engineering and Engineering Management, Hong Kong.

Saarinen, K., & Aho, M. (2005). Does the implementation of a clinical information system decrease the time intensive care nurses spend on documentation of care? Acta anaesthesiologica scandinavica, 49(1), 62-65.

Saha, D., & Mukherjee, A. (2003). Pervasive computing: a paradigm for the 21st century. [10.1109/MC.2003.1185214]. IEEE Computer, 36(3), 25-31.

Santos, M., Portela, F., Boas, M., Machado, J., Abelha, A., Neves, J., et al. (2009). Intelligent Decision Support in Intensive Care Units Nursing Information Requirements. Paper presented at the APPLIED INFORMATICS AND COMMUNICATIONS (AIC '09).

Santos, M. F. (2006). FCT: Resumo do projecto INTCare. Retrieved 09-07-2009, 2009, from http://www.fct.mctes.pt/projectos/pub/2006/Painel_Result/vglobal_projecto.asp?idProjecto=72819&idElemConcurso=895

Santos, M. F., Cortez, P., Gago, P., Silva, Á., & Rua, F. (2006). Intelligent decision support in Intensive Care Medicine. Paper presented at the 2nd International Conference on Knowledge Engineering and Decision Support, Lisbon, Portugal.

Santos, M. F., & Portela, F. (2011). Enabling Ubiquitous Data Mining in Intensive Care - Features selection and data pre-processing. Paper presented at the 13th International Conference on Enterprise Information Systems.

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., & Neves, J. (2009a). Information Architecture for Intelligent Decision Support in Intensive Medicine. 8th International Conference on Applied Computer and Applied Computational Science (ACACOS '09), 8(5), 810-819.

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., & Neves, J. (2011a). INTCARE - Multi-agent approach for real-time Intelligent Decision Support in Intensive Medicine. Paper presented at ICAART.

Santos, M. F., Portela, F., Vilas-Boas, M., Machado, J., Abelha, A., Neves, J., et al. (2009). Information Modeling for Real-Time Decision Support in Intensive Medicine. In S. Y. Chen & Q. Li (Eds.), Proceedings of the 8th Wseas International Conference on Applied Computer and Applied Computational Science - Applied Computer and Applied Computational Science (pp. 360-365). Athens: World Scientific and Engineering Acad and Soc.

Santos, M. F. V.-B., Marta. Portela, Filipe. Silva, Álvaro. Rua, Fernando. (2010). Hourly prediction of organ failure and outcome in intensive care based on data mining techniques. Paper presented at the ICEIS - 12th International Conference on Enterprise Information Systems.

Satyanarayanan, M. (2001). Pervasive computing: vision and challenges. [10.1109/98.943998]. IEEE Personal Communications, 8(4), 10-17.

Satyanarayanan, M. (2002). Pervasive computing: vision and challenges. Personal Communications, IEEE, 8(4), 10-17.

Schusterschitz, N., & Joannidis, M. (2007). Predictive capacity of severity scoring systems in the ICU. Contributions to nephrology, 156(R), 92.

Scicluna, P., Murray, A., Xiao, Y., & Mackenzie, C. F. (2008). Challenges to Real-Time Decision Support in Health Care. Agency for Healthcare Research and Quality.

Sesso, H. D., Stampfer, M. J., Rosner, B., Hennekens, C. H., Gaziano, J. M., Manson, J. A. E., et al. (2000). Systolic and diastolic blood pressure, pulse pressure, and mean arterial pressure as predictors of cardiovascular disease risk in men. Hypertension, 36(5), 801-807.

Shabot, M. M. (2005, 2005). Automated data acquisition and scoring for JCAHO ICU core measures.

Silva, A. (2007). Modelos de Inteligência Artificial na análise da monitorização de eventos clínicos adversos, Disfunção/Falência de órgãos e prognóstico do doente critico.: Tese de doutoramento, ciências médicas, universidade do Porto.

Silva, Á., Cortez, P., Santos, M. F., Gomes, L., & Neves, J. (2004). Multiple organ failure diagnosis using adverse events and neural networks. Paper presented at the 6 th International Conference on Enterprise Information Systems.

Silva, Á., Cortez, P., Santos, M. F., Gomes, L., & Neves, J. (2008). Rating organ failure via adverse events using data mining in the intensive care unit. Artificial Intelligence in Medicine, 43(3), 179-193.

Silva, Á., Pereira, J., Santos, M., Gomes, L., & Neves, J. (2003). Organ failure prediction based on clinical adverse events: a cluster model approach, 3th International Conference on Artificial Intelligence and Applications.

Simon, H. A. (1960). The New Science of Management Decision Making. New York-Evanston.

Sprague Jr, R. H., & Carlson, E. D. (1982). Building Effective Decision Support Systems: Prentice Hall Professional Technical Reference.

Stacey, M., & McGregor, C. (2007). Temporal abstraction in intelligent clinical data analysis: A survey. Artificial Intelligence in Medicine, 39(1), 1-24.

Strand, K., & Flaatten, H. (2008). Severity scoring in the ICU: a review. Acta Anaesthesiologica Scandinavica, 52(4), 467-478.

Streitz, N., & Nixon, P. (2005). The disappearing computer. Communications of the ACM, 48(3), 32-35.

Suter, P., Armaganidis, A., Beaufils, F., Bonfill, X., Burchardi, H., Cook, D., et al. (1994). Predicting outcome in ICU patients. Intensive Care Medicine, 20(5), 390-397.

Tamayo, P., Berger, C., Campos, M., Yarmus, J., Milenova, B., Mozes, A., et al. (2005). Oracle Data Mining. Data Mining and Knowledge Discovery Handbook, 1315-1329.

Tang, Z., Mazabob, J., Weavind, L., Thomas, E., & Johnson, T. R. (2006, 2006). A time-motion study of registered nurses' workflow in intensive care unit remote monitoring.

Turban, E., Aronson, J. E., & Liang, T.-P. (2005). Decision Support Systems and Intelligent Systems (7 ed.): Prentice Hall.

Turban, E., McLean, E., & Wetherbe, J. (2001). Information Technology for Management: Transforming Business in the Digital Economy (3rd ed.): JWley & Sons Inc;.

Turban, E., Sharda, R., & Delen, D. (2010). Decision Support and Business Intelligence Systems (9th Edition ed.): Prentice Hall.

Vahidov, R. (2002). Decision Station: A Notion for a Situated DSS. Paper presented at the The Annual Hawaii International Conference on System Sciences.

Vandijck, D. M., Labeau, S. O., & Blot, S. I. (2008). Level of knowledge articulated by intensive care nurses and clinical decision-making. Intensive and Critical Care Nursing, 24(1), 6-7.

Varshney, U. (2003). Pervasive Healthcare. Computer, 36(12), 138-140.

Varshney, U. (2007a). Pervasive healthcare and wireless health monitoring. Mobile Networks and Applications, 12(2), 113-127.

Varshney, U. (2009). Pervasive Healthcare Computing: EMR/EHR, Wireless and Health Monitoring: Springer-Verlag New York Inc.

Vassiliadis, P., Simitsis, A., & Skiadopoulos, S. (2002, 2002). Conceptual modeling for ETL processes.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. Decision Sciences, 39(2), 273-315.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management science, 186-204.

Vilas-Boas, M., Santos, M. F., Portela, F., Silva, Á., & Rua, F. (2010a). Hourly prediction of organ failure and outcome in intensive care based on data mining techniques. Paper presented at the 12th International Conference on Enterprise Information Systems.

Vilas-Boas, M., Santos, M. F., Portela, F., Silva, Á., & Rua, F. (2010b). Tempo real na previsão da falência de órgãos e outcome em Medicina Intensiva. Paper presented at the CISTI'2010 (5ª Conferência Ibérica de Sistemas e Tecnologias de Informação).

Villas Boas, M., Gago, P., Portela, F., Rua, F., Silva, Á., & Santos, M. F. (2010). Distributed and real-time Data Mining in the Intensive Care Unit. Paper presented at the 19th European Conference on Artificial Intelligence - ECAI 2010.

Vincent, J., Mendonça, A., Cantraine, F., Moreno, R., Takala, J., Suter, P., et al. (1998). 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, 26, 1793-1800.

Vincent, J. L., & Bruzzi de Carvalho, F. (2010). Severity of illness. Semin Respir Crit Care Med, 31(1), 031-038.

Vincent, J. L., Moreno, R., Takala, J., Willatts, S., De Mendonca, A., Bruining, H., et al. (1996). The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. Intensive care medicine, 22(7), 707-710.

Weiser, M. (1991). The Computer for the Twenty-First Century. Scientific American, 265(3), 94-104.

White, D. B., Braddock, C. H., Bereknyei, S., & Curtis, J. R. (2007). Toward shared decision making at the end of life in intensive care units - Opportunities for improvement. [Article]. Archives of Internal Medicine, 167(5), 461-467.

Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. Paper presented at the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining.

Wooldridge, M. (1999). Intelligent agents Multiagent systems: a modern approach to distributed artificial intelligence (pp. 27-77): MIT Press.

Ying, Z., Silvers, C. T., & Randolph, A. G. (2007). Real-Time Evaluation of Patient Monitoring Algorithms for Critical Care at the Bedside. Paper presented at the Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE.

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem, 39(4), 561-577.

APPENDIX A – Glossary

Arterial blood gas (ABG) [1]**:** The sampling of the blood levels of oxygen and carbon dioxide within the arteries, as opposed to the levels of oxygen and carbon dioxide in venous blood. Typically, the acidity, or pH, of the blood is measured simultaneously in ABG sampling.

Admission, hospital [2]: The formal acceptance by a hospital or other inpatient health care facility of a patient who is to be provided with room, board, and continuous nursing service in an area of the hospital or facility where patients generally reside at least overnight.

Amine [1]**:** A chemical compound containing nitrogen. Amines are derived from ammonia. It can be norepinephrine, dopamine, Dobutamine and epinephrine.

Acute Physiology & Chronic Health Evaluation III (APACHE III) [2]: A 'third-generation' system for estimating the risk of hospital death in adult ICU patients based on physiologic assessments of most severely affected values during the first 24 hours in the ICU and subjecting the results to logistic regression modelling techniques

Arrhythmias, atrial [1]**:** Abnormal heart rhythm due to electrical disturbances in the atria (the upper chambers of the heart) or the AV node "relay station", leading to fast heart beats. Examples of atrial arrhythmias include atrial fibrillation, atrial flutter, and paroxysmal atrial tachycardia (PAT).

Electrolyte Balance [2]**:** The relative concentrations of ions in the body's extracellular and intracellular fluids, especially those produced from ionized salts;

Bilirubin [1]**:** A yellow-orange compound that is produced by the breakdown of haemoglobin from red blood cells.

[1](MedicineNet, 2012); [2](Farlex, 2009)

Blood pressure [1]**:** The blood pressure is the pressure of the blood within the arteries. It is produced primarily by the contraction of the heart muscle. Its measurement is recorded by two numbers. The first (systolic pressure) is measured after the heart contracts and is highest. The second (diastolic pressure) is measured before the heart contracts and lowest. A blood pressure cuff is used to measure the pressure. Elevation of blood pressure is called "hypertension".

BP Diastolic [1]**:** Referring to the time when the heart is in a period of relaxation and dilatation (expansion). The noun for diastolic is diastole. The diastolic pressure is specifically the minimum arterial pressure during relaxation and dilatation of the ventricles of the heart when the ventricles fill with blood. In a blood pressure reading, the diastolic pressure is typically the second number recorded. For example, with a blood pressure of 120/80 ("120 over 80"), the diastolic pressure is 80. By "80" is meant 80 mm Hg (millimetres of mercury). A diastolic murmur is a heart murmur heard during diastole, the time the heart relaxes. "Diastolic" came from the Greek diastole meaning "a drawing apart." The term has been in use since the 16th century to denote the period of relaxation of the heart muscle.

BP Systolic [1]**:** The blood pressure when the heart is contracting. It is specifically the maximum arterial pressure during contraction of the left ventricle of the heart. The time at which ventricular contraction occurs is called systole. In a blood pressure reading, the systolic pressure is typically the first number recorded. For example, with a blood pressure of 120/80 ("120 over 80"), the systolic pressure is 120. By "120" is meant 120 mm Hg (millimetres of mercury). A systolic murmur is a heart murmur heard during systole, the time the heart contracts, between the normal first and second heart sounds. "Systolic", comes from the Greek systole meaning "a drawing together or a contraction." The term has been in use since the 16th century to denote the contraction of the heart muscle.

Central venous pressure (CVP) [2]: the venous pressure as measured at the right atrium, done by means of a catheter introduced through the median cubital vein to the superior vena cava.

Clinical laboratory [2]: one for examination of materials derived from the human body for the purpose of providing information on diagnosis, prognosis, prevention, or treatment of disease.

Coagulation [2]: The entire process of blood clotting, the process of transforming a liquid into a solid, especially of the blood.

[1](MedicineNet, 2012); [2](Farlex, 2009)

Creatinine [1]**:** A chemical waste molecule that is generated from muscle metabolism. Creatinine is produced from creatine, a molecule of major importance for energy production in muscles. Approximately 2% of the body's creatine is converted to creatinine every day. Creatinine is transported through the bloodstream to the kidneys. The kidneys filter out most of the creatinine and dispose of it in the urine. Although it is a waste, creatinine serves a vital diagnostic function. Creatinine has been found to be a fairly reliable indicator of kidney function. As the kidneys become impaired the creatinine will rise. Abnormally high levels of creatinine thus warn of possible malfunction or failure of the kidneys, sometimes even before a patient reports any symptoms. It is for this reason that standard blood and urine tests routinely check the amount of creatinine in the blood. Normal levels of creatinine in the blood are approximately 0.6 to 1.2 milligrams (mg) per decilitre (dl) in adult males and 0.5 to 1.1 milligrams per decilitre in adult females. (In the metric system, a milligram is a unit of weight equal to one-thousandth of a gram, and a decilitre is a unit of volume equal to one-tenth of a litter.) Muscular young or middle-aged adults may have more creatinine in their blood than the norm for the general population. Elderly persons, on the other hand, may have less creatinine in their blood than the norm. Infants have normal levels of about 0.2 or more, depending on their muscle development. A person with only one kidney may have a normal level of about 1.8 or 1.9. Creatinine levels that reach 2.0 or more in babies and 10.0 or more in adults may indicate the need for a dialysis machine to remove wastes from the blood. Certain drugs can sometimes cause abnormally elevated creatinine levels.

Diastole [1]: The time period when the heart is in a state of relaxation and dilatation (expansion). The diastolic pressure is specifically the minimum arterial pressure during relaxation and dilatation of the ventricles of the heart. Diastole is the time when the ventricles fill with blood. In a blood pressure reading, the diastolic pressure is typically the second number recorded. For example, with a blood pressure of 120/80 ("120 over 80"), the diastolic pressure is 80. By "80" is meant 80 mm Hg (millimeters of mercury). A diastolic murmur is a heart murmur heard during diastole, the time the heart relaxes. "Diastole" came without change from the Greek diastole meaning "a drawing apart." The term has been in use since the 16th century to denote the period of relaxation of the heart muscle.

Discharge summary [2]: the clinical notes written by the discharging physician or dental professional at the time of releasing a patient from the hospital or clinic, outlining the course of treatment, the status at release, and the post discharge expectations and instructions.

[1](MedicineNet, 2012); [2](Farlex, 2009)

Diuretic [1]**:** Something that promotes the formation of urine by the kidney. All diuretics cause a person to 'lose water,' but they do so by diverse means, including inhibiting the kidney's ability to reabsorb sodium, thus enhancing the loss of sodium and consequently water in the urine (loop diuretic); enhancing the excretion of both sodium and chloride in the urine so that water is excreted with them (thiazide diuretic); or blocking the exchange of sodium for potassium, resulting in excretion of sodium and potassium but relatively little loss of potassium (potassium-sparing diuretic). Some diuretics work by yet other mechanisms, and some have other effects and uses, such as in treating hypertension. Also known as water pills. Substances in food and drinks, such as coffee, tea, and alcoholic beverages, may act as diuretics

Drainage [2]: systematic withdrawal of fluids and discharges from a wound, sore, or cavity;

Electronic health record (EHR) [1] is an evolving concept defined as a systematic collection of electronic health information about individual patients or population .It is a record in digital format that is capable of being shared across different health care settings, by being embedded in network-connected enterprise-wide information systems. Such records may include a whole range of data in comprehensive or summary form, including demographics, medical history, medication and allergies, immunization status, laboratory test results, radiology images, vital signs, personal stats like age and weight, and billing information. Its purpose can be understood as a complete record of patient encounters that allows the automation and streamlining of the workflow in health care settings and increases safety through evidence-based decision support, quality management, and outcomes reporting.

Feces / Faeces [1]**:** The excrement discharged from the 'intestines.

FiO2 [2]: is the fraction of inspired oxygen in a gas mixture. The FiO2 is expressed as a number from 0 (0%) to 1 (100%). The FiO2 of normal room air is 0.21 (21%).

Fluid Balance: The difference between the amount of water taken into the body and the amount excreted or lost. Also called as *water balance*.

[1](MedicineNet, 2012); [2](Farlex, 2009)

Gastric emptying study [1]: The most common type of gastric emptying study is a procedure that is done by nuclear medicine physicians using radioactive chemicals that measures the speed with which food empties from the stomach and enters the small intestine. Gastric emptying studies are used for evaluating patients who are having symptoms that may be due to slow and, less commonly, rapid emptying of the stomach. The symptoms of slow emptying are primarily nausea, vomiting, and abdominal fullness after eating. The symptoms of rapid emptying are diarrhoea, weakness, or light-headedness after eating

Glasgow Coma Scale (GSC) [1]: was developed to provide health-caregivers a simple way of measuring the depth of coma based upon observations of eye opening, speech, and movement. Patients in the deepest level of coma: do not respond with any body movement to pain (1-6), do not have any speech (1-5), and do not open their eyes (1-4). Those in lighter comas may offer some response, to the point they may even seem wake, yet meet the criteria of coma because they do not respond to their environment. The scale is used as part of the initial evaluation of a patient, but does not assist in making the diagnosis as to the cause of coma. Since it "scores" the level of coma, the Glasgow Coma Scale can be used as a standard method for any health-caregiver to assess change in patient status. The best use of the Glasgow Coma Scale is to allow caregivers of different clinical skills and training to consistently assess a patient over longer periods of time in order to determine whether the patient is improving, deteriorating, or remaining the same. In the initial care of a comatose patient, there may be first responders, EMTs, paramedics, emergency physicians, neurologists, neurosurgeons, and nurses evaluating the patient. The Glasgow Coma Scale allows a standard assessment that can be shared. A fully awake patient has a Glasgow Coma Score of 15. A person who has deceased has a Glasgow Coma Scale of 3 (there is no lower score).

Glucose [1]: is a simple sugar (monosaccharide) that is used to increase the level of blood glucose when the level falls too low (hypoglycaemia). Glucose is a glucose-elevating agent. Other glucose-elevating agents are dioxide (Proglycem) and glucagon. Glucose is the primary fuel used by most cells in the body to generate the energy that is needed to carry out cellular functions. When glucose levels fall to hypoglycaemic levels, cells cannot function normally, and symptoms develop such as nervousness, cool skin, headache, confusion, convulsions or coma. Ingested glucose is absorbed directly into the blood from the intestine and results in a rapid increase in the blood glucose level.

[1](MedicineNet, 2012); [2](Farlex, 2009)

Glycemic index [1]: An indicator of the ability of different types of foods that contain carbohydrate to raise the blood glucose levels within 2 hours. Foods containing carbohydrates that break down most quickly during digestion have the highest glycemic index. Also called the dietary glycemic index.

Heart rate (HR) [1]: The number of heartbeats per unit of time, usually per minute. The heart rate is based on the number of contractions of the ventricles (the lower chambers of the heart). The heart rate may be too fast (tachycardia) or too slow (bradycardia). The pulse is a bulge of an artery from waves of blood that course through the blood vessels each time the heart beats. The pulse is often taken at the wrist to estimate the heart rate.

Intensive Care / Critical Care: constant complex health care as provided in various acute life-threatening conditions such as multiple trauma, severe burns, or myocardial infarction or after certain kinds of surgery. Care is most frequently given by specially trained personnel in a unit equipped with various technologically sophisticated machines and devices for treating and monitoring the condition of the patient.

Intensive Care Unit (ICU) [2]: a hospital unit in which patients requiring close monitoring and intensive care are kept. An ICU contains highly technical and sophisticated monitoring devices and equipment and is staffed by personnel trained to deliver critical care. A large tertiary care facility usually has separate units specifically designed for the intensive care of adults, infants, children, or newborns or for other groups of patients requiring a certain kind of treatment.

Intracranial pressure (ICP) [2]: pressure of the subarachnoid fluid.

Mean arterial pressure (MAP) [2]: the average pressure within an artery over a complete cycle of one heartbeat.

Modified early warning score (MEWS) [2]: simple guide used by hospital nursing & medical staff as well as emergency medical services to quickly determine the degree of illness of a patient. It is based on data derived from four physiological readings (systolic blood pressure, heart rate, respiratory rate, body temperature) and one observation (level of consciousness, AVPU). The resulting observations are compared to a normal range (-3 to 3) to generate a single composite score.

[1](MedicineNet, 2012); [2](Farlex, 2009)

Organ failure [1]: The failure of an essential system in the body. Multiple organ failure is the failure of two or more systems, such as the cardiovascular, and renal systems, and is a common consequence of sepsis (the presence of bacteria in the bloods) and of shock (very low blood pressure).

Outcome [2]: the condition of a patient at the end of therapy or a disease process, including the degree of wellness and the need for continuing care, medication, support, counselling, or education.

Oximetry [1]: is a procedure for measuring the concentration of oxygen in the blood. The test is used in the evaluation of various medical conditions that affect the function of the heart and lungs.

Oxygen [1]: The odourless gas that is present in the air and necessary to maintain life. Oxygen may be given in a medical setting, either to reduce the volume of other gases in the blood or as a vehicle for delivering anaesthetics in gas form. It can be delivered via nasal tubes, an oxygen mask, or an oxygen tent. Patients with lung disease or damage may need to use portable oxygen devices on a temporary or permanent basis.

PaO2/FIO2 [2]: Pressure of Arterial Oxygen to Fractional Inspired Oxygen Concentration

PAO2 [2]: Partial Pressure of Oxygen in Arterial Blood

Platelet count [1]: The calculated number of platelets in a volume of blood usually expressed as platelets per cubic millimetre (cmm) of whole blood. Platelets are the smallest cell-like structures in the blood and are important for blood clotting and plugging damaged blood vessels. Platelet counts are usually done by laboratory machines that also count other blood elements such as the white and red cells. They can also be counted by use of a microscope. Normal platelet counts are in the range of 150,000 to 400,000 per microliter (or 150 - 400 x 10^9 per litter). These values many vary slightly between different laboratories.

Positive end-expiratory pressure (PEEP) [2]: a method of mechanical ventilation in which pressure is maintained to increase the volume of gas left in the lungs at the end of exhalation, reducing shunting of blood through the lungs and improving gas exchange.

[1](MedicineNet, 2012); [2](Farlex, 2009)

Invasive blood pressure[2]: is a method of measuring blood pressure internally by using a sensitive IV catheter inserted into an artery. This provides a more accurate reading of the patent's current blood pressure. This is usually used where rapid variations of blood pressure are anticipated;

Non-invasive blood pressure[2]**:** The non-invasive auscultator and oscillometric measurements are simpler and quicker than invasive measurements, require less expertise, have virtually no complications, are less unpleasant and less painful for the patient. However, non-invasive methods may yield somewhat lower accuracy and small systematic differences in numerical results. Non-invasive measurement methods are more commonly used for routine examinations and monitoring.

Plateau Pressure [2]: is the pressure applied to small airways and alveoli during positive-pressure mechanical ventilation. It is measured during an inspiratory pause on the mechanical ventilator.

Pulmonary artery wedge pressure (PAWP), **pulmonary capillary wedge pressure** (PCWP) [2]: intravascular pressure as measured by a catheter wedged into the distal pulmonary artery ; used to measure indirectly the mean left atrial pressure.

Pulse pressure [2]: the difference between systolic and diastolic pressures.

Leukocyte [2]: white cell, white blood cell; a colourless blood corpuscle capable of ameboid movement, whose chief function is to protect the body against microorganisms causing disease and which may be classified in two main groups: *granular* and *nongranular.*

Erythrocytes [2]: Red blood cells.

Globule [2]: A small spherical body, especially a drop of liquid.

Sequential Organ Failure Assessment (SOFA) [2]: were calculated for each patient's cardiovascular, renal, hematologic, respiratory, and hepatic organ systems and indicate patients who did not demonstrate organ function improvement during the first 24 hours of therapy were significantly less likely to survive. Furthermore, during the first day of therapy for severe sepsis, the change from baseline in organ dysfunction may be a better predictor of 28-day mortality than a static baseline assessment.

[1](MedicineNet, 2012); [2](Farlex, 2009)

Respiratory rate [1]: The number of breaths per minute or, more formally, the number of movements indicative of inspiration and expiration per unit time. In practice, the respiratory rate is usually determined by counting the number of times the chest rises or falls per minute. The aim of measuring respiratory rate is to determine whether the respirations are normal, abnormally fast (tachypnea), abnormally slow (bradypnea), or nonexistent (apnea).

FiO2 [2]: may be varied through the use of different Venturi masks, in combination with varying oxygen flow rates. In addition, most mechanical ventilators have controls for adjusting FiO2. An increased FiO2 is necessary in managing adequate oxygenation in patients who are critically ill due to causes such as major surgery, acute lung injury, sepsis, pneumonia, congestive heart failure, or other cardiopulmonary disease. The oxygenation to a patient on a ventilator can be manipulated by changing not only FiO2, but also the tidal volume, the respiratory rate and having a Positive end-expiratory pressure (PEEP). Generally the FiO2 is maintained at less than 60%. Higher settings can lead to oxygen toxicity. Another common misconception is that the FiO2 changes with elevation. It remains at 0.21 at all altitudes within the atmosphere. What changes is the barometric pressure of air. At altitude, therefore, the partial pressure of oxygen delivered by that 21% of oxygen is lower. The partial pressure is the driving force to oxygenate the blood and therefore a lower partial pressure makes it that much harder to get O2 delivered to the tissues that require it, resulting in hypoxia

Plasma cell [1]: A type of white blood cell that produces and secretes antibodies. A plasma cell is a fully differentiated, mature lymphocyte in the B cell lineage. As with most cell types, plasma cells can mutate to give rise to cancer. Plasma cell malignancies include plasmacytoma, multiple myeloma, Waldenstrom macroglobulinemia, and plasma cell leukaemia. Also known as plasmacytoma;

Respiratory system [1]: The organs that are involved in breathing, including the nose, throat, larynx, trachea, bronchi, and lungs. It is also known as respiratory tree.

Simplified Acute Physiology Score (SAPS II) [2]: Intensive care A 'third-generation' system for estimating in-hospital mortality in adult ICU patients, based on assessments of most severely affected values during the 1st 24 hours in the ICU and subjecting the results to logistic regression modelling techniques. It uses 17 variables.

[1](MedicineNet, 2012); [2](Farlex, 2009)

Sepsis [2]: is a potentially dangerous or life-threatening medical condition, found in association with a known or suspected infection (usually caused by but not limited to bacteria) whose signs and symptoms fulfill at least two of the following criteria of a systemic inflammatory response syndrome (SIRS)

Serum [2]: the clear portion of any liquid separated from its more solid elements.

SOFA [2]: The Sequential Organ Failure Assessment score, or just SOFA score, is used to track a patient's status during the stay in an intensive care unit (ICU). It is one of several ICU scoring systems. The SOFA score is a scoring system to determine the extent of a person's organ function or rate of failure. The score is based on six different scores, one each for the respiratory, cardiovascular, hepatic, coagulation, renal and neurological systems.

System, Neurologic [2]: evaluation of the health status of a patient with a nervous system disorder or dysfunction. The purposes of the assessment include establishing a diagnosis to guide the veterinarian in prescribing medical and surgical treatments and in planning and implementing nursing measures to help the patient cope effectively with daily living. Includes evaluation of cranial nerves, gait, mental state, muscle tone (1), postural reactions, sensory perceptivity, spinal nerves and visceral function

System, Cardiovascular [1]: The circulatory system which comprises the heart and blood vessels. The system carries nutrients and oxygen to the tissues of the body and removes carbon dioxide and other wastes from them. The cardiovascular system is a closed tubular system in which the blood is propelled by the heart. The system has two circuits, the pulmonary circuit and the systemic circuit. Each circuit has arterial, capillary, and venous components.

System, Hepatic [2]: relating to the liver

Systole [1]: The time period when the heart is contracting. The period specifically during which the left ventricle of the heart contracts. The systolic pressure is specifically the maximum arterial pressure during contraction of the left ventricle of the heart. In a blood pressure reading, the systolic pressure is typically the first number recorded. For example, with a blood pressure of 120/80 ("120 over 80"), the systolic pressure is 120. By "120" is meant 120 mm Hg (millimetres of mercury). A systolic murmur is a heart murmur heard during systole, the time the heart contracts, between the normal first and second heart sounds. "Systole" came without change from the Greek

[1](MedicineNet, 2012); [2](Farlex, 2009)

systole meaning "a drawing together or a contraction." The term has been in use since the 16th century to denote the contraction of the heart muscle.

Temperature [1]: The temperature is the specific degree of hotness or coldness of the body. It is usually measured with a thermometer.

Therapeutic [1]: Relating to therapeutics, the branch of medicine that is concerned specifically with the treatment of disease. The therapeutic dose of a drug is the amount needed to treat a disease.

TISS [2]: Therapeutic Intervention Scoring System Intensive care. It is a system used to assess intensity of care administered in the ICU, based on the number of interventions the patient needs

Vital signs [2]: Basic indicators of body function, usually meaning heartbeats per minute, breaths per minute, blood pressure, body temperature, and weight.

[1](MedicineNet, 2012); [2](Farlex, 2009)

APPENDIX B – Published Articles

Table A1 – PhD published Articles

#	Year	Type	Title	Editor	ISI	ACM	SCOPUS	IEEE	DBLP
1	2010	Book Chapter	Improvements in data quality for decision support in Intensive Care	Springer	X				
2	2010	Proceedings	Distributed and real time Data Mining in the Intensive Care Unit	IOS Press					
3	2010	Proceedings	Electronic Health Records in the Emergency Room	IEEE	X			X	
4	2010	Proceedings	Hourly Prediction of Organ Failure and Outcome in Intensive Care Based on Data Mining Techniques	SciTePress	X				X
5	2010	Proceedings	Real-Time Intelligent Decision Support in Intensive Medicine	SciTePress	X				X
6	2010	Proceedings	Real-time prediction of organ failure and outcome in intensive medicine	IEEE	X			X	
7	2011	Book Chapter	Enabling a Pervasive approach for Intelligent Decision Support in Intensive Care	Springer	X				X
8	2011	Proceedings	Enabling Real-time Intelligent Decision Support in Intensive Care	EUROSIS	X				X
9	2011	Proceedings	Enabling Ubiquitous Data Mining in Intensive Care - Features Selection and Data Pre-Processing	SciTePress	X				X
10	2011	Proceedings	INTCARE -Multi-agent Approach for Real-time Intelligent Decision Support in Intensive Medicine	SciTePress	X				X
11	2011	Proceedings	Knowledge Discovery for Pervasive and Real-Time Intelligent Decision Support In Intensive Care Medicine	SciTePress	X				X
12	2012	Book Chapter	A Pervasive Approach to a Real-Time Intelligent Decision Support System in Intensive Medicine	Springer	X				X
13	2012	Book Chapter	An Intelligent Patient Monitoring System	Springer	X				X
14	2012	Book Chapter	Grid Data Mining Strategies for Outcome Prediction in Distributed Intensive Care Units	IGI Global	X		X		X
15	2012	Book Chapter	Intelligent and Real Time Data Acquisition and Evaluation to Determine Critical Events in Intensive Medicine	Elsevier	X		X		X

#	Year	Type	Title	Editor	ISI	ACM	SCOPUS	IEEE	DBLP
16	2012	Book Chapter	Intelligent Data Acquisition and Scoring System for Intensive Medicine	Springer	X				X
17	2012	Book Chapter	Knowledge Acquisition Process for Intelligent Decision Support in Critical Health Care	IGI Global	X		X		X
18	2012	Book Chapter	Multi-agent systems for HL7 interoperability services	Elsevier	X		X		X
19	2013	Book Chapter	Intelligent Information System to Tracking Patients in Intensive Care Units	Springer	X		X		X
20	2012	Proceedings	Data Mining Predictive Models for Pervasive Intelligent Decision Support in Intensive Care Medicine	SciTePress	X				X
21	2012	Proceedings	Intelligent Decision Support in Intensive Care – Towards Technology Acceptance	EUROSIS	X				X
22	2012	Proceedings	Monitoring Intelligent System for the Intensive Care Unit using RFID and Multi-Agent Systems	IEEM Proc.				X	
23	2013	Book Chapter	Adoption of Pervasive Intelligent Information Systems in Intensive Medicine	Elsevier	X				X
24	2013	Book Chapter	Modelling Intelligent Agents to Integrate a Patient Monitoring System	Springer	X				X
25	2013	Book Chapter	Pervasive and Intelligent Decision Support in Critical Health Care Using Ensembles	Springer	X				X
26	2013	Book Chapter	Pervasive Ensemble Data Mining Models to predict Organ Failure and Patient Outcome in Intensive Medicine	Springer	X				X
27	2013	Book Chapter	Pervasive Intelligent Decision Support System - Technology Acceptance in Intensive Care Units	Springer	X		X		X
28	2013	Book Chapter	Predict Sepsis Level in Intensive Medicine - Data Mining Approach	Springer	X		X		X
29	2013	Journal	Pervasive Real-Time Intelligent System for Tracking Critical Events in Intensive Care Patients	IGI Global	X		X		X
30	2013	Journal	Real-Time Decision Support in Intensive Medicine - An intelligent approach for monitoring Data Quality	IGI Global	X		X		X
31	2013	Journal	Real-time Predictive Analytics for Sepsis Level and Therapeutic Plans in Intensive Care Medicine	IGI Global		X	X		X
32	2013	Journal	Towards Pervasive and Intelligent Decision Support in Intensive Medicine – A Data Stream Mining Approach.	Elsevier	X		X		X
33	2013	Proceedings	A Pervasive Intelligent System for Scoring MEWS and TISS-28 in Intensive Care	IGI Global	X				X
34	2013	Proceedings	An Intelligent Approach for Open Clinical Laboratory Results in Intensive Care Medicine	IEEM Proc.				X	
35	2013	Proceedings	Data Mining for Real-Time Intelligent Decision Support System in Intensive Care Medicine	SciTePress	X				X
36	2013	Proceedings	Pervasive Information Systems to Intensive Care Medicine - Technology Acceptance Model	SciTePress	X				X
37	2013	Book Chapter	Adoption of Pervasive Intelligent Information Systems in Intensive Medicine	SciTePress	X				X