The DTP Model: Integration of intelligent techniques for the decision support in Healthcare Assistance

David Riaño [1], John A. Bohada [2], Tatjana Welzer [3]
[1]Department d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Tarragona, Spain. drianyo@etse.urv.es
[2]PhD Student at the Universitat Politècnica de Catalunya, Barcelona, Spain. jbohada@lsi.upc.es
[3]Institut of Informatics, Unviersity of Maribor, Maribor, Slovenia. welzer@uni-mb.si

Abstract
The paper introduces a new model, called the DTM model, that uses Artificial Intelligence techniques to obtain health-care knowledge that can be applied in a combined way to support the decision making in some relevant medical activities as diagnosis, treatment selection, and prognosis. The DTP model applies inductive learning techniques to hospital data and obtains action rules, clinical guidelines and belief networks. These knowledge structures are respectively exploited by an inference engine, a case-based reasoner, and a probability propagation system to automatically propose a Diagnostic-Treatment-Prognostic (DTP) sequence that the new patients should follow. The model and its implementation has been tested with data concerning cardiopathologies of the patients assisted in the Hospital Joan XXIII in Tarragona.

Keywords: Data Mining and Knowledge Extraction, Machine Learning, Case based Reasoning, Bayesian Networks, Healthcare Assistance, Clinical Guidelines.

1 Introduction

Physicians that attend a patient which is affected by one or several diseases must decide about the best possible heath-care attention that could improve the patient's health. This attention begins with the application of diagnostic techniques, continues with the selection and adaptation of a concrete treatment, and finishes with the accomplishment of a prognostic that measures, in advance, the evolution of the patient with respect to the treatment.

The *diagnosis* process is based on the different tests, analysis, samples, clinical events, etc., that help the physician to identify a main disease (semiology) [1]. These tests are also useful in the process of *selecting a concrete treatment*, and also to adapt it to the patient so that the maximum improvement could be achieved. The way that the future improvement of the patient is predicted is called the *prognosis process*.

This sequence is related to the following aspects:

Elicitation: physicians confront its experience with medical literature, physical examinations and medical files, in order to obtain one or several hypotheses that explain the causes that are producing the ailments.

Verification: physicians use additional medical tests to increase their confidence in the final diagnosis. The results are used to adjust the initial hypotheses and to reach a definitive conclusion about the disease that the patient has (diagnosis).

Recommendation: physicians compare the feasible risks and the advantages of ordering additional tests or making therapeutic interventions to develop alternative therapies for the patient (treatment selection). These therapies are discussed with the patient, and his or her preferences can be taken into account (treatment adaptation) whenever the risk do not increase (prognosis). After that, the arranged treatment is applied.

A medical treatment, or therapy, is described as the temporal distribution of drug prescriptions and the application of medical procedures. The ability to propose a good treatment is something that grows with the experience of the physician as he or she (a) attends similar cases, (b) is aware of similar clinical episodes in the hospital record file, (c) knows about the advances in the treatment of the disease, and (d) uses helpful information computer systems [2]. Moreover, these elements have some influence in the capability of the physician to predict the future clinical evolution of the patient state and the probable results (cure, stabilization, *exitus*, etc.). There are several factors that conditionate these predictions [3]: the etiologic diagnosis, the patient's general clinical state, the stage of the disease, the treatment followed, etc. [4].

Along the last decades, multiple computer-based tools have been proposed with the purpose of helping to improve the processes in healthcare assistance. These tools have allowed physicians to have other options in the health-care decision making processes of diagnosis, treatment definition and prognosis. So, the development of Knowledge-Based Systems [5] has played a very important role as decision support elements in healthcare. As far as the evolutive perspective is concerned, the most relevant features of these systems are the *reasoning process* that they use [6] (e.g.

associative, causal, uncertain, rule-based, method-based, case-based, etc.), the *knowledge representation model* (e.g. trees, rules, frames, networks, ontologies, etc.), and the *approach model* ranging from quantitative and probabilistic (e.g. k-Nearest Neighbour, Instance-Based Learning, Neural Networks, Bayesian Belief Networks, Regression, etc.) to qualitative and symbolic (e.g. Decision Trees, Production and Association Rules, Inductive Logic Programming, Genetic Algorithms, etc.). A short sample of some of the oldest well-known Artificial Intelligent systems that were applied in health-care are MYCIN [7], QMR [8], DxPLAIN [9], CASNET [10], CONSIDER [11], ILIAD [12].

Another problem treated in health-care refers to the management of the great *amount of information* that are generated with the clinical practice. In order to come to grips with this increasing saturation of medical data, some Artificial Intelligence technologies have been developed directed to the knowledge discovery (KDD) [14] and the intelligent analysis of data (IDA) [15]. Although these technologies are narrowly related, they are different from the point of view of the scope of application [16][17]. Broadly speaking, these technologies provide reliable automatic methods to fight the increasing nature of the information by discovering principles, mechanisms and causes that are implicitly contained in the health-care data. If these elements are organized in knowledge structures, they can be used to support medical decision making.

Nevertheless, and in spite of the efforts that have been done in the development of approaches that involve different techniques to the support to the medical decision making, we observe that all the known works that combine diagnosis, treatment and prognosis [19][20][31][3] use either a single approach (decision-theoretic planning, case-based reasoning, etc.), or are not based on explicit knowledge structures, or the knowledge is supplied by the physicians and not by a learning process.

Unlike these approaches, we propose the modelling of a system of support to the decision making called DTP that it allows to the integration of the activities of diagnosis, treatment and prognosis applied to the medical domain. This system will be based on the execution of two fundamental processes: the application of techniques of KDD for the extraction of essential knowledge in the processes of diagnosis, treatment and prognosis in health-care and the knowledge exploitation in aid to physicians in the decision making.

For a better understanding of this process, the modelling of system DTP is detailed in section 2. A case study is supplied in section 3 and the conclusions in section 4.

2 The DTP model

The *Diagnostic-Treatment-Prognostic model*, or DTP, retakes aspects followed in the *domino model* [19][20], for the accomplishment of diagnosis and treatment. Additionally one becomes jumbled predictive characteristics (figure 1) which allow integrating all the aspects of the healthcare assistance, objective of model DTP.

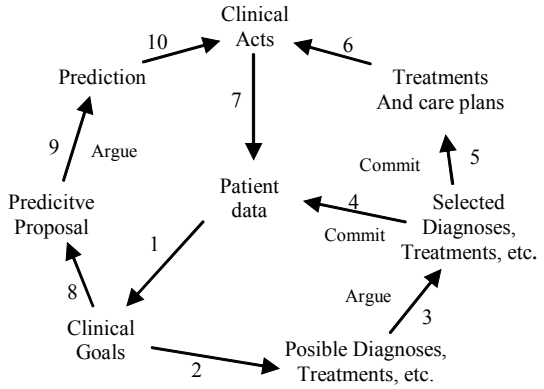

Figure 1. The DTP model

After having developed the processes of diagnosis and followed treatment actions clinical, united with data of patients, the prognosis process begins with the exposition of clinical objective directed to determine predictive factors that they allow to advance future facts with respect to the evolution of the disease (arrow 8), the prediction becomes and their results are evaluated (arrow 9), helping to decide the actions pertinent clinical for favourer the health of the patient (arrow 10).

In the modelling of DTP system, the accumulation of data will be used to capture the essence of the diagnosis processes, therapeutic procedures, and prognosis factors, followed by the unit that provide these data. Its base will be in the clinical data generated in the consultations given to the patients and the data already stored of healthcare assistance. In order to fulfil the essence of a system of support to the decision making, our model allows to integrate Inductive Learning techniques [21][24] (production rules), case-based reasoning [22] and Bayesian Network [23], for the diagnosis, treatment and prognosis processes respectively.

2.1 Diagnosis

The left side column in Figure 2 describes the process by which a data base is transformed into a knowledge base that is used to diagnose the new patients. The starting point is a supervised data matrix that relates the description of a set of patients with the diagnosed disease. All the patients are described according to the same predefined medical features (e.g. age, sex, blood pressure, previous heart attack Y/N, etc.).

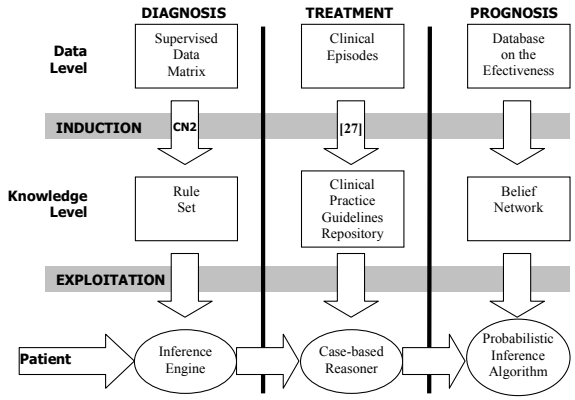

Figure 2. Modular description of the DTP model

An inductive learning algorithm is used to transform these data into a set of production rules that represent *diagnosis knowledge*. Although the DTP model is compatible with many of such algorithms [24], only the CN2 inductive procedure [25] is considered here. Once the diagnosis knowledge is generated in the form that Figure 3 describes, it can be used to predict the disease of the new patients. This knowledge exploitation stage is based in an inference engine that applies the rules to the description of the new patients, inducing a final diagnosis for them.

```
IF chest-pain-type = asymptomatic
 AND resting-blood-pressure >107
 AND ST-depression-induced > 2.45
THEN heart-disease = presence [29 0]
ELSE
IF 111 < resting-blood-pressure < 150
 AND max-heart-rate > 174.5
 AND ST-depression-induced < 3.65
THEN heart-disease = absence [0 27]
ELSE
IF sex = female
 AND chest-pain-type ≠ asymptomatic
 AND 105.5 < max-heart-rate < 173.5
 AND number-of-major-vessels < 1.5
THEN heart-disease = absence [0 46]
ELSE
IF 243.5 < serum-cholesterol < 320.0
 AND max-heart-rate >106.5
 AND exercise-induced-angina = yes
THEN heart-disease = presence [28 0]
ELSE ...
```

Figure 3. CN2 ordered rules for heart diseases

Some metrics, as accuracy, can be calculated in order to conclude about the reliability of the diagnosis process.

2.2 Treatment

The second column in Figure 2 shows the artificial intelligent process to deal with the problem of choosing and adapting a treatment to a particular diagnosed patient. As treatments are temporal structures, we propose that the generation of *treatment knowledge* must be based on temporal data describing the stages of therapies of the patients. A *clinical episode* [26] is defined as the description of the medical actions that a patient receives for a particular disease between the admission and the discharge times. Each episode can contain one or more *description lines* that represent physician decisions like drug prescriptions (change medication, increase or decrease dose or frequency) or orders for clinical procedures (blood analysis, X ray, etc.). Here, each description line contain information about the minimum basic data set (MBDS): patient id, age, sex, residence, primary diagnosis, secondary diagnoses, primary procedure, secondary procedures, doctor id, hospital service, number of days in the hospital service, and discharge reason; extended with drug information.

```
Procedure MakeGuideline (S,R)
Let S be the set of the episodes understudy.
Let R be the node where the guideline must
be rooted.
If S=0, stop.
Selected P the procedures which are in all
the episodes of S.
If P≠0
    Remove P from all the episodes in S.
    Make an Action A with the procedures in
    P, and root it to R.
    For all the empty episodes,
    Generate a state for each discharge
    reason.
    Root them to A.
    Remove empty episodes from S.
    Call makeGuideline (S,A).
  else
    B=BestCondition(S).
    Make a Decision D with the condition B,
    and root it to R.
    Let S+ be the episodes in S fulfilling B.
    Let S- the rest of the episodes in S.
    Call MakeGuideline (S+,D.yes).
    Call MakeGuideline (S-,D.no).
```

Figure 4. Algorithm for guideline generation

The algorithm displayed in Figure 4 is used to transform a data base with clinical episodes into a case base where each case is a clinical practice guideline that generalizes the treatment followed for the patients diagnosed with a particular disease [27]. The clinical practice guidelines are defined following the *PROforma* [30] style and can contain action nodes (describing medical orders concerning medication and clinical procedures), decision points (representing conditions that allow forks in the treatment), and state points (indicating the discharge reasons). Figure 5 shows a clinical practice guideline representing the treatment of Acute Idiopathic Pericarditis.

Different guidelines can be obtained for the same disease if the algorithm of Figure 4 is applied to the episodes that come from different hospitals or those belonging to different years. Therefore, the *case base* or guideline repository is designed to contain one ore more

clinical practice guidelines for each target disease or diagnosis.

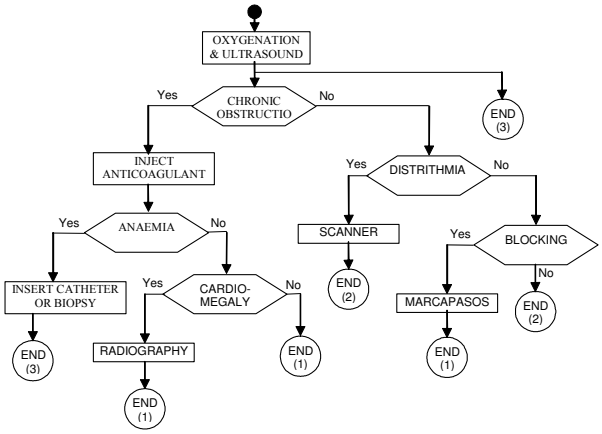

Figure 5. Acute Idiopathic Pericarditis Guideline

The *treatment knowledge*, represented by the case base, is exploited with a *case-based reasoning* (CBR) process that defines the *retrieval function* (f_r) in terms of the minimal distance between the new diagnosed patient p (the one that we want to prescribe a treatment for) and the patients x that were used to make each guideline. The way this function is defined is represented by the next equations.

$$f_r(p,C) = \arg\min_{P \in C} \{\arg\min_{x \in P} \{dist(p,x)\}\}$$

$$dist(p,x) = dif(p.sex, x.sex) + \frac{|p.age - x.age|}{max.age} + \frac{p.procs \cap x.procs}{p.procs \cup x.procs} + \frac{\sum_{m \in p.meds \cap x.meds} m.dosi \cdot m.freq}{\sum_{m \in p.meds \cup x.meds} m.dosi \cdot m.freq}$$

The *adaptation function* is a process by which the guideline removes the decision points and the banches that evaluate to false for this patient. For example, an anaemic patient with a chronic obstruction of his aerial ways will obtain the treatment of Figure 6, if he is diagnosed with Acute Idiopathic Pericarditis.

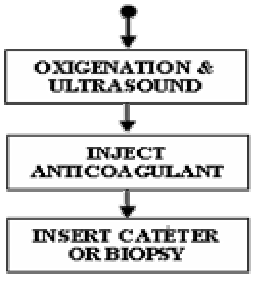

Figure 6. Guideline adapted to a particular patient

2.3 Prognosis

The right column in Figure 2 shows the prognosis process [2][28]. It begins with a database that contains all the data about the historical effectiveness of the treatments that are stored in the guideline repository. Effectiveness is measured in terms of *relative mortality* (the ratio of the number of deaths for a particular diagnosis when a treatment is followed in one year divided by the total amount of patients following that treatment during the same year) and *relative morbidity* (the ratio of the number of cases for which there is a prevalecence of a disease following a treatment divided by the number of people which have received the treatment).

The data about the effectiveness of the guidelines is used to calculate the conditional probabilities in a Bayesian Belief Network as the one that Figure 7 shows. The network has nodes corresponding to discrete variables that represent patient data, diagnosis, treatments, mortality, and morbidity. These nodes are connected following the logics of the DTP model: all the patient data are connected to all the diagnosis (e.g. *P(Acute Lung Edema | sex=female)*), the diagnoses are only connected to the particular treatments (e.g. *P(guideline_CCI2 | Congestive Cardiac Insufficiency)*), and all the treatments to all the prognosis variables (e.g. *P(morbidity=low | guideline_CCI2)*).

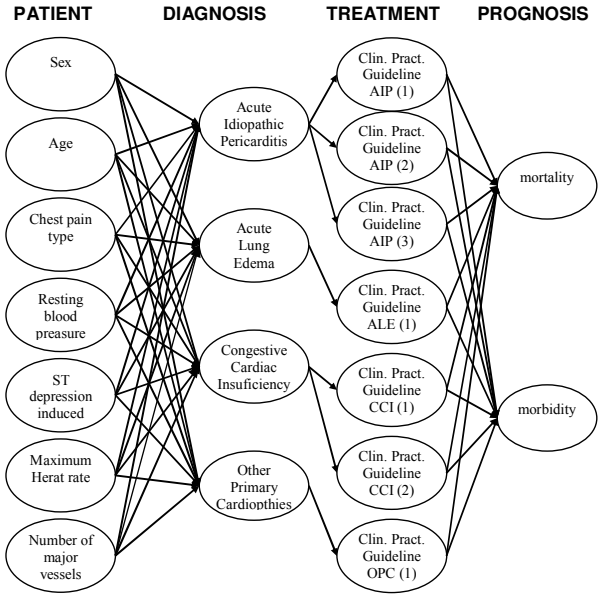

Figure 7. Layered BN for heart disease prognosis

The Bayesian Network obtained has multiple prognostic utilities according to the sort of conditional probability that the Probabilistic Inference Algorithm calculates.

$$P(mortality = high \mid age = adult)$$

$$P(morbidity = medium \mid diagnosis = CCI)$$

$$P(morbidity = low \mid treatment = guidelineCCI(1))$$

It is possible to calculate the prognosis value of mortality or morbility conditioned to one or several descriptions of the patients (e.g. age, sex, symptoms, signs, etc.), or conditioned to a particular diagnosis, or conditioned to a concrete treatment or guideline.

3 A Case Study: Cardiopathologies

Cardiac diseases are one of the greatest causes of mortality nowadays in the world. Defining healthy nutritional, physical, psychological, and spiritual habits, together with prevention, informative campaigns, early diagnoses, and suitable treatments are some of the ways to fight these diseases. Here, the DTP model has been applied to study some of these diseases so that the tasks of diagnosis, treatment prescription, and prognosis can be supervised with a knowledge-based decision support system.

Two data sources have been used: the four "Heart Diseases Databases" available at the Irvine University Repository [29] (cleveland.dat, hungarian.dat, long-beach-va.dat, and switzerland.dat), and the database of the Hospital Joan XXIII, Tarragona (Spain). The first four databases describe the "angiographic disease status" of 303+294+123+200 (= 290) patients according to the patient's age, sex, chest pain type, resting blood pressure, serum cholestoral, fasting blood sugar, electrocardiographic results, maximum heart rate, induced angina, ST depression, ST slope of the peak, number of major vessels colored by fluoroscopy, and thal.

The last database contains the episodes of all the patients admitted at the *Hospital Joan XXIII* in the years 2000, 2001, and 2002. The number of patients and mean stage in days are 13 and 10.5 for "Acute Idiopathic Pericarditis", 103 and 9 for "Acute Lung Edema", 64 and 11 for "Congestive Cardiac Insufficiency", and 21 and 8 for "Other Primary Cardiopathies".

		TRAINING			
		Cleveland	Hungarian	Long Beach	Switzerland
TESTING	Cleveland	100,0	72,4	67,5	80,5
	Hungarian	68,3	100,0	68,5	51,2
	Long Bch	63,7	64,6	100,0	89,4
	Switzerland	70,0	76,9	75,0	100,0

Table 1. Results of the diagnosis process

The cross fold application of the CN2 algorithm on the Cleveland, Hungarian, Long Beach, and Switzerland databases shows the accuracies of Table 1. We can conclude that Long Beach rules performed the best for Switzerland patients (89,4%), followed by the Cleveland rules with Switzerland patients (80,5). We could think that Switzerland patients are easily diagnosed heart disease cases, but this is not the case since they obtain the lowest accuracy with the Hungarian rules (51,2%, which is equivalent to diagnose by throwing a coin). If we consider 75%, the minimum acceptable result, none of the rule sets can assure this accuracy for all the data sets. Finally, Switzerland rules are the ones that obtain the mean best results.

The clinical practice guidelines obtained for "Acute Idiopathic Pericarditis", "Acute Lung Edema", "Congestive Cardiac Insufficiency", and "Other Primary Cardiopathies" with the data of the patients attended in the year 2000 are depicted in Figure 5, Figure 8, Figure 9, and Figure 10, respectively.

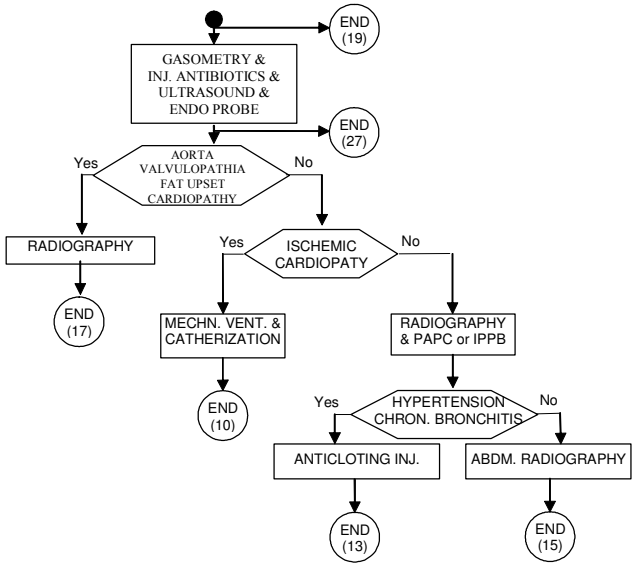

Figure 8. Acute Lung Edema Guideline

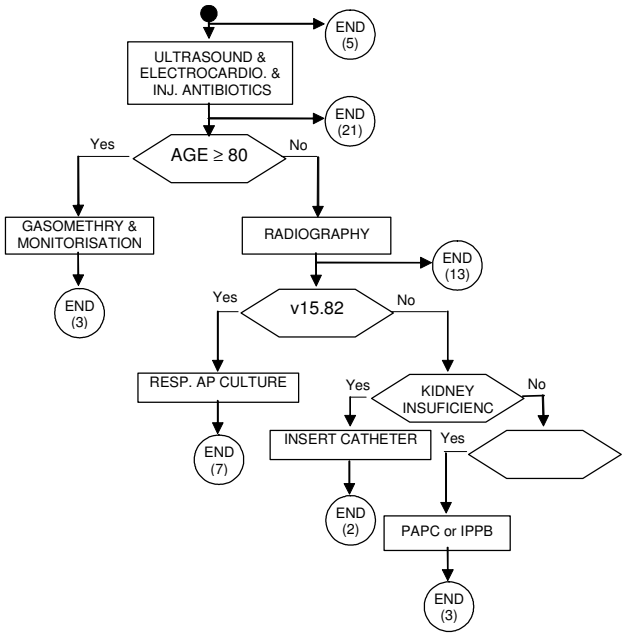

Figure 9. Congestive Cardiac Insufficiency Guideline

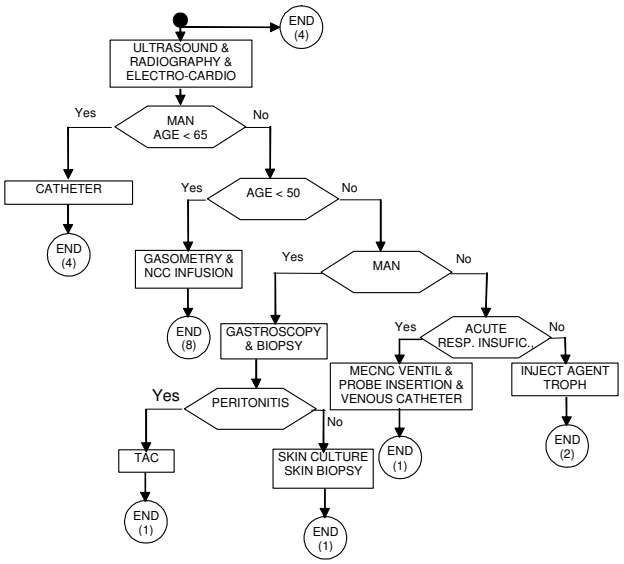

Figure 10. Other Primary Cardiopathies Guideline

Finally, the prognosis process has been tested for the diagnosis variables of *sex* (considering men and women separately) and *age* (discretizing between young, medium-aged and old people). The conditional probabilities in Table 2 represent the concepts of "probability of diagnose a particular disease conditioned to the fact that the patient has some feature (e.g. is a man, is young, etc.)". The probabilities in Table 2 are very small since they are calculated over the total amount of patients in the hospital and not only on those with heart disease. This is a requirement of the DTP model, which is intended to be applied to patients which are not diagnosed.

Whenever there have been patients of each of the four considered diseases in the years 2000, 2001, and 2002, a separate clinical guideline has been created and incorporated in the case base. The proportion of patients in each year has been used to calculate the conditional probabilities in Table 3. These percentages refer to the idea of the "probability of being treated following a concrete clinical guideline conditioned to the fact that has one of the four heart diseases under study (i.e. Acute Idiopathic Pericarditis or AIP, Acute Lung Edema or ALE, Congestive Cardiac Insufficiency or CCI, and Other Primary Cardiopathies or Other)".

	AIP	ALE	CCI	Other
M	0,07%	0,51%	0,32%	0,14%
W	0,07%	0,57%	0,36%	0,08%
<17	0,00%	0,00%	0,00%	0,00%
17-60	0,02%	0,17%	0,12%	0,08%
>60	0,16%	1,28%	0,78%	0,20%

Table 2. Conditional probability at the diagnostic level

Finally, Table 4 contains the percentages that describe the "probability of a patient to *cure*, *die* (mortality), or *remain ill* (morbidity) after applying one of the clinical guidelines generated with the patients in the year 2000".

From the former tables we can conclude that there are not clear differences in the diagnosis between men and women (e.g. $P(ALE \mid M) \cong P(ALE \mid W)$), there are not young people affected with heart diseases (i.e. $P(AIP \mid <17)=P(ALE|<17)=P(CCI|<17)=P(Other|<17)=0,0\%$) and the majority of affected persons are older than 60, there are not AIP cases in the years 2001 and 2002 (i.e. $P(2001 \mid AIP)=P(2002 \mid AIP)=0$), patients treated according to the AIP guideline of the year 2000 do not cure completely and either die (i.e. $P(cure \mid AIP2000)=P(mortality \mid AIP2000)=0,0\%$), and very often all four heart diseases remain chronic after following any of the proposed treatments (i.e. $P(morbidity \mid AIP2000)=P(morbidity \mid ALE2000)= P(morbidity \mid CCI2000)= P(morbidity \mid Other2000)> 60,0\%$).

	2000	2001	2002
AIP	100,00%	0,00%	0,00%
ALE	34,79%	27,70%	37,50%
CCI	19,39%	29,09%	51,51%
Other	33,87%	24,19%	41,93%

Table 3. Conditional probability at the treatment level

	cure	mortality	morbidity
AIP 2000	0,00%	0,00%	100,00%
ALE 2000	30,35%	5,35%	64,28%
CCI 2000	13,33%	0,00%	86,66%
Other 2000	7,69%	3,84%	88,46%

Table 4. Conditional probability at the prognosis level for the guidelines related to the patients in 2000

4 Conclusions

A new model has been presented that integrates the clinical activities of diagnosis, treatment selection, and prognosis. Several Artificial Intelligence techniques has been proposed for the automatic construction of several sorts of knowledge that are used to perform those clinical activities in an integrated mode by a computer system to support clinical decision making.

References

1. Kukar M. Transductive reliability estimation for medical diagnosis. *Artificial Intelligence in medicine*. 2002 (to appear).
2. Lucas J., Abu-Hanna A. Editorial. Prognostic Models in Medicine: IA and statistical Approaches. Methods of information in medicine. Schattauer GmbH, 2001.

3. Colloc J., Bouzidi L. A Case Based Reasoning Decision Support System for use in Medicine. Upgrade II(1): 30-35, 2001.
4. Fok S. C., Ng E. Y. K., Thimm G. Developing Case-Based Reasoning for Discovery of Breast Cancer. In *Journal of Mechanics in Medicine and Biology*, 2003 (in press).
5. Kalogeropolulos D, et al. Towards knowledge-based systems in clinical practice: Development of an integrated clinical information and knowledge management support system. *Computer Methods and Programs in Biomedicine* 72(2003): 65-80.
6. Miller R. A. Medical Diagnosis Decision Support System – Past, Present, and future: A threaded Bibliography and Brief Commentary. *Journal of the American Medical Informatics Association: JAMIA*. 8-27. 1994.
7. Shortliffe E. H. Computer-Based Medical Consultations: MYCIN. *Artificial Intelligence Series*. New York: Elsevier Computer Science Library, 183-185, 1976.
8. Berman L., Millar R. A. Problem area formation as an element of computer aided diagnosis: a comparison of two strategies within Quick Medical Reference (QMR). *Methods Inf. Med*. 30(1991): 90-5.
9. Barnett G. O., Cimino J. J., Hupp J. A., Hoffer E. P. DXplain. An evolving diagnostic decision-support system. *JAMA* 258(1): 67-74, 1987.
10. Weiss S. M., Kulikowski C. A., Amarel S., Safir A. A model-based method for computer-aided medical decision making. *Artificial Intelligence* 11 (1978): 145–172.
11. Broering NC. Corn M, Ayers WR, Mistry P. implementing RECONSIDER, a diagnostic prompting computer system, at the Georgetown University Medical Center. *Bull Med Libr Assoc*. 76(1988): 155-8.
12. Warner H. R., Haug P., Bouhaddou 0., et al. ILIAD as an expert consultant to teach differential diagnosis. In: *Proceedings of the Twelfth Annual Symposium on Computer Applications in Medical Care*. New York: IEEE Computer Society Press 1987; 371-6.
13. Hasman A., Safran C., Takeda H. Quality of health care: informatics foundations. *YearBook of Medical Informatics*. 2001: 143-151.
14. Fayyad U., Storloz P. Data Mining and KDD: Promise and challenges. *Future Generation Computer System* 13(1997): 99-115.
15. Deogun J., Raghavan V., Sarkar A., and Sever H., Data Mining: Trends in Research and Development. In *Rough Sets and Data Mining: Analysis for Imprecise Data*, Boston, MA, 9-45, 1996.
16. Lavrac N., Keravnou E., and Zupan B. Intelligent data Analysis in Medicine, in A. Kent et al., eds., *Encyclopedia of Computer Science and Technology*, 42 (2000): 113-157.
17. Lavrac, N. Data Mining in medicine: selected Techniques and Applications. *Artificial Intelligence in Medicine*, 16(1): 3-23, 1999.
18. Michie D., Spiegelhalter D. J., Taylor C. C. Eds, Machine Learning, Neural and Statistical Classification, Ellis Horwood, 1994.
19. Das K. A., Fox J. A flexible architecture for autonomous agents. *Revised version is being considered by JETAI*, 1995.
20. Das K. A., Fox J., Krause P. A Unified Framework for Hypothetical and Practical reasoning (1): Theoretical Foundations. In *International Conference on Formal and Applied Practical Reasoning*. LNAI: 58-72, Springer-Verlag, 1996.
21. Golding Andrew, Rosembloom Paul. Improving accuracy by combining rule-based and case-based reasoning. *Artificial Intelligence* 87(1996) 215-254.
22. Golobardes Elzabet, et al. Computer aided diagnosis with case-based reasoning and genetic algorithms. *Knowledge-Based System*, 15(2002): 45-52.
23. Fernández G. S. Redes Bayesians temporales: aplicaciones médicas e industriales. Tesis doctoral. Universidad Complutense de Madrid. 2002.
24. Riaño D. On the process of making descriptive rules, *LNAI* 1624 : 182-197, 1999.
25. Clark P. and Niblett T. The CN2 induction algorithm. *Machine Learning* 3 (1989): 261–283.
26. Riaño D, Prado P. The Analysis of Hospital Episodes, In *Medical Data Analysis*, LNCS 2199: 231-37, 2001.
27. Riaño D. Guideline Composition from Minimum Basic Data Set. The 16th *IEEE Symposium on Computer-Based Medical Systems*, 231-235, NY, 2003.
28. Lucas Peter, Abu-Hanna A. Prognostic methods in medicine. *Artificial Intelligence in Medicine* 15 (1999) 105-119.
29. Blake, C.L. and Merz, C.J. (1998). UCI Repository of machine learning databases [http://www.ics.uci.edu/~mlearn/MLRepository.html] Irvine, CA: University of California, Department of Information and Computer Science.
30. Fox J., Johns N., Rahmanzadeh A. Disseminating Medical Knowledge: The PROforma Approach. *Artificial Intelligence in Medicine* 14: 157-181, 1998.
31. Peek N.B. Decision-Theoretic Planning of Clinical Patient Management. PhD dissertation, 2000.