Hierarchal Fuzzy Logic Medical Database with Decision Algorithm for Metabolic Syndrome Diagnosis

I. Salgado [(1)], A. Rodríguez [(1)], J. Chairez [(1)], A. Zúñiga [(2)] and P. Santillan [(2)]

[(1)] Unidad Profesional Interdisciplinaria de Biotecnología del IPN.

[(2)] Instituto de Ciencias Médicas y Nutrición "Salvador Zubirán"

jchairez@ctrl.cinvestav.mx

Abstract. Probably, the most complicated aspect in medical care is an efficient and opportune diagnostic from the Medic. Few years ago this branch of medical attention was depending only on physician's experience and his knowledge, who were dedicated to the patients care. However, this practice has been changed since Data Base (DB) related and intelligent systems (like experts systems) have been used. In this document a Related DB for Nutriology Clinician Department of Instituto Nacional de Ciencias Médicas y Nutrición (INCMNSZ), is shown, whose objective is to classify the patient's information for a better clinical state visualization to be gotten. An appropriate designation for an operational risk scale that indicates the attention degree that must be applied in sick persons will be determinate by the mixed application of fuzzy logic and relational database. The information system response is based in a technique called Fussy Logic, which is applied in medical variables analysis in order to find the existence of Metabolic Syndrome in the patient, and in a direct test about their life quality, alimentation habits and Clinic History. The connection between the DB and Fussy System was created with Matlab and the ODBC Data Transfer System.

1 Introduction

The clinical record is the most important element in the medical treatment that can be applied to a patient by a physician, due the information included into it. Roughly speaking, the data set contained in the corresponding patient file, is constructed by the answers (given by a specific patient) to specific question designed by an expert medical group. The questions are related to the patient's health conditions, the sickness events over the entire life, a short version of the patient's family medical record, lifestyle, some relevant accidents, surgery interventions, allergies, vaccination history and any current medicine intake by the patient [1].

These groups of data are one of the most important aspects for the medic to take decisions in a correct way, diagnose efficiently the illness' patient and propose a solution to him. In this paper the CH is a key to follow the patients who are candidates to suffer *Metabolic Syndrome*. The metabolic Syndrome is defined by the OMS as an altered regulation of glucose or diabetes (it means an insulin unresponse

© A. Gelbukh, Á. Kuri (Eds.)
Advances in Artificial Intelligence and Applications
Research in Computer Science 32, 2007, pp. 430–441

Received 20/06/07
Accepted 31/08/07
Final version 22/10/07

that is defined as glucose make out under the last quartile levels for population in the study), besides 2 or more of the following components:

- High blood pressure.
- A triglyceride level above 150 mg/dl
- A High density lipoprotein level (HDL)
- Obesity or high IMC.

The exact cause of metabolic syndrome is unknown. Most researchers believe it is caused by a combination of the genetic makeup and lifestyle choices, including the types of eaten food and the level of physical activity. If the metabolic syndrome is diagnose, the body suffers a series of biochemical changes. Over time, these changes lead to the development of one or more associated medical conditions. The sequence begins when insulin, a hormone excreted from your pancreas, loses its ability to make your body's cells absorb glucose from the blood-your body uses glucose for energy. When this happens, glucose levels remain high after you eat. Your pancreas, sensing a high glucose level in your blood, continues to excrete insulin. Loss of insulin production may be genetic or secondary to high fat levels with fatty deposits in the pancreas [2]. Until today, the SM's diagnose is given for the Nutriology medic's experience. Actually an automatic or semiautomatic system, that can give us an appropriate and exactly diagnose, doesn't exist because each factor of SM has a dissimilar relative weigh and affect in a different way the metabolic patients state, which can't be so solved whit classic programming methods.

The Fuzzy Logic (FL) can associate uncertainly o eventually variables whit a mathematic function, which is called Membership Function, obtaining an associate grade inside a range between 0 and 1; that is one of the differences whit the classical binary theories, because the only can give two different values to the variables. The FL is conformed for four steps: Fuzzification, rules base, inference mechanism and defuzzification [3]. The FL is used in a lot of applications, like an uncertainly models controller, expert systems and decision systems whit partially defined variables. This is making whit an input's review, based in the *modus-ponens* structure: *condition-action.* An example applied to this work is the following:

If ... the patient has a high glucose level... so the patient has diabetes

These kinds of sentences are going to use to determinate SM and its related illness.

2 Methodology

The Methodology is described for the following steps:

2.1 Design of the Nutriology Department's Data Base

Microsoft Access was used to design de Data Base because it has an easy form to development the elements (tables, forms, queries, etc) of it. MatLab and the OBDC Data System were the tools to export and import data from the Data Base to the fuzzy system, which were done in Matlab. The DB has eight areas: *I.* Diagnose, *II.* Clinician

Information, *III.* Life Style, *IV.* Poverty, *V.* Habits, *VI.* Metabolic Syndrome, *VII.* ECD and Comorbility, *VIII.* Nutritional Evaluation.

2.2 Data export and import

Matlab has a different kind of tools to link it whit external software, in other words matlab can import data from access to make an analysis of the DB. "*Visual Query Builder*" is software to take the information inside the DB tables and show it in a matrix form in MatLab.
The program code to import de information from the DB is the as follows:

```
1. logintimeout(10);
2. conn = database('bd2', '', '');
3. ping(conn);
4. curs = exec(conn, 'select * from prueba');
5. setdbprefs('DataReturnFormat','numeric');
6. curs = fetch(curs);
7. z = curs.Data;
```

Description of the code Line1. Specify the time connection to Matlab before send an error message. Line 2 and 3. "database" is the function to make the link between the DB and Matlab, the function contains the DB's name, user name and password The line 3 specifies the connection state. Line 4, 5, 6 and 7. The fields in a table are chosen whit the function shown in line 4. The Data Format Return are given by the function setdbprefs.
The program code to export information from Matlab to the DB is the following,:

```
1. setdbprefs('DataReturnFormat','cellarray');
2. d='MALA';
3. exdata={d};
4. colnames(i) = {'calvida'};
5. insert(conn, Prueba, Colnames, exdata)
```

Line 1. Specify the format to return the data. Line 2 and 3. These lines show the data thal wll be export to access. Line 4. Shows the name of the DB field. Line5. To export the information to access is used the insert function which inserts a new line into the DB.

2.3 Pre- Diagnostic fussy System Development

A specialist medic and a lot of bibliography sources were consulted to make a correct diagnose, in special the National Cardiology journal was read and the diagnose algorithm was taken from this journal. This algorithm is shown in the picture 1 and describes the steps to follow to find evidence of SM. The diagnose standards are explained in the figure 1. The fuzzy system only analyzed the part in bold of the algorithm in figure 1, the variables to get involved used the parameters shows in table 1. The figure 1 shows the diagnostic algorithm. The variables included in the MS description and its corresponding range is depicted in Table 1. The parameters defining the discourse universe are shown in the same table.

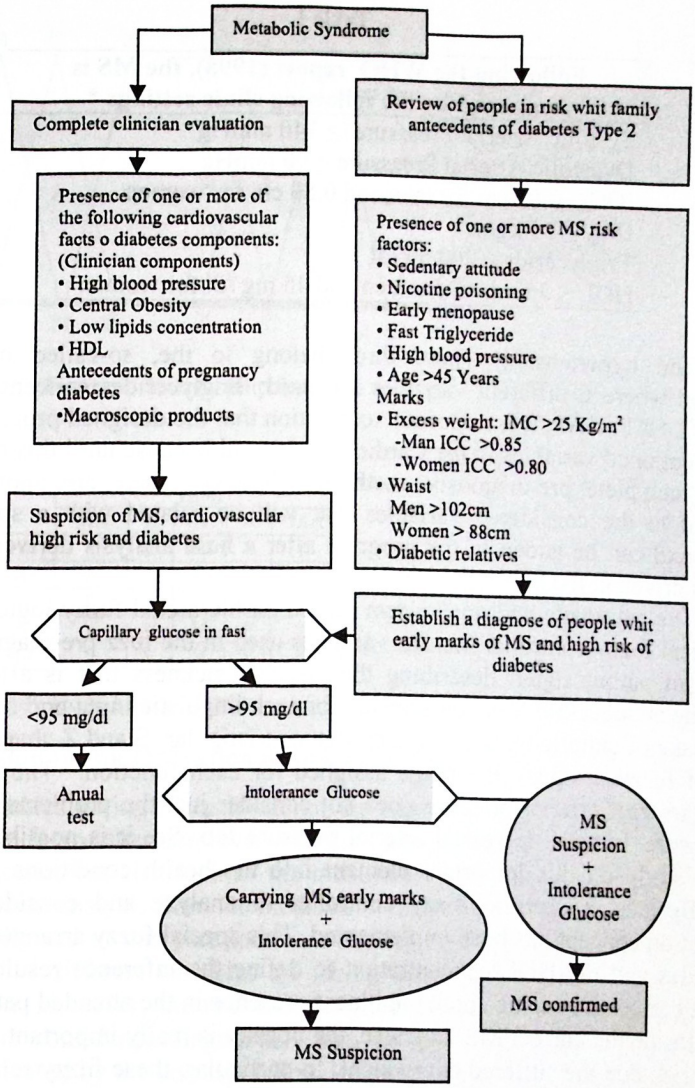

Figure 1. Diagnose algorithm by

These are diagnostic criterion: It is considered that an abnormal fasting condition and with high glucose concentration, glucose intolerant or affected by 2-type diabetes mellitus had MS if and only if there is any possible association with both components enlisted above.

FUZZY ALGORITHM DEVELOPEMENT

There were designed 3 different fuzzy pre-diagnostic computer programs. Each one of them is dedicated to one of the following clinical conditions: low lipids concentration,

Table 1

Following the WHO report (1998), the MS is componed by the following clinic settings *
Systolic Arterial Preassure ≥ 140 mmHg
Dyastolic Arterial Preassure ≥ 90 mmHg
ICC ≥ 0.85 cm for men and 0.80 cm for women
IMC ≥ 25 kg/m^2
Triglycerids ≥ 200 mg/dl
HDL < 35 mg /dl for men and 45 mg /dl for women

obesity and hypertension. These three belong to the, so-called overall clinic evaluation where 6 different variables are used: Triglycerides concentration, HDL, IMC, ICC, SAP, DAP. It is important to mention that the designed program don't use all the mentioned variables in the Cardiology Journal because until this moment there is not a complete pre-diagnostic method. However, there are another illnesses associated to the considered variables that will be related with a specific health problem and can be added to the program after a hard analysis derived by medical protocols driven at INCMNSZ.

The automatic decision system (based on hierarchal fuzzy logic) is based on the classical fuzzy controller, i.e. the variables used in the fuzz pre-diagnostic system includes an output signal describing the possible sickness that is affecting to the patient, the membership function selected for each linguistic input and output variable (for the sake of simplicity, they were chosen as triangular, S and Z shapes as usual in mostly of fuzzy designs), the range assigned for each function. The fuzzy system designed to treat arterial pressure does not consider just the problems associated to high pressure, but the decreased arterial pressure too. So, it is possible to define a complete study on this important element into the health conditions given for the patient. Besides, a specific fuzzy controller to analyze and consider low lipids concentration concept has been implemented. This special fuzzy arrangement uses the triglycerides and the HDL concentration to define the inference result, and then to provide a correct diagnostic about the illness presence in the attended patient.

Inside the correct MS diagnose, the obesity is really important to specify the patient's risk due the suffered overweight. In particular, these fuzzy relationships are based on two simple concepts: a) the relation or the, so-called, hip-waist index and b) the corporal mass index. These both concepts are so useful to define in a unique form the obesity condition on the patient. An example of the fuzzy system process to diagnose one of the three illnesses mentioned (***Low Lipid Concentration***) is shown by the following membership functions (Figure 2) and its corresponding knowledge fuzzy matrix (Table 2).

The fuzzy matrix to compare both clinician aspects gives us de rules to diagnose the illness

The membership functions (Input and Output) and the matrix were suggested by the nutrition specialist at hospital after a detail revision of the article in the Cardiology Journal. The defuzification was made by the centroid method (equation 2).

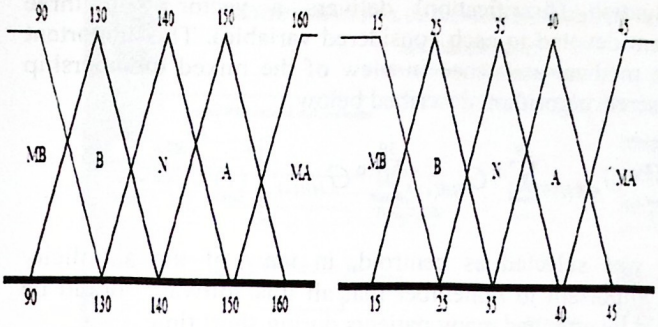

Figure 2. Membership functions for triglyceride and HDL

Low Lipid Concentration		HDL				
		MB	B	N	A	MA
Triglycerides	MB	P	P	N'	N	N
	B	P	P	N	N	N
	N	P	N	N	N	N
	A	V	V	V	P	P
	MA	V	V	V	V	P

Table 2. Knowledge Fuzzy Matrix

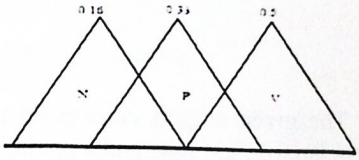

Figure 3. Output Membership Functions

2.4 Designation for an Operational Risk Scale Fuzzy System

The main variables taking place to assign the risky condition for the patient and to determinate the future attention level with the corresponding self-care are linked whit the MS diagnostic, but also include the meal habits, the carcinogenic tumors and a complete clinician analysis from the medicals to the patients. The next table illustrates the method followed in this part and the variables ranges are also depicted there. In this paper, just 4 different variables are taking into account to define the necessary medic attention. However, the real DB designed considers 39 different elements to produce a more confidence diagnostic. Once the variables are extracted of the DB, they are fuzzificated by the corresponding designed membership functions (Figure 4). The "risk" functions, for example, are classified in 1) minimum, 2) reasonable and 3) high patient risk. In general, this fuzzy structure is applied in all others variables.

Variable	Minimum Risk	Reasonable risk	High Risk
Glucose	<90	<150	>300
Triglycerides	<150	<300	>500
Uric Acid	<6	>8	<3
Albumin	>3.5	>3.4	<2.9

Table 5.MS Variables

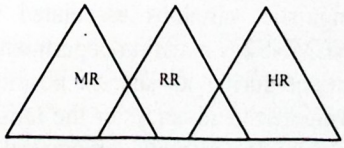

Figure 4. Input Membership fuctions

The ultimate evaluation (fuzzification) delivers a vector with three components (each one of them devoted to each considered variable). This important element, in fact, defines the medical assistance in view of the mixed membership function describing by the discrete algorithm, described below:

$$^{\circ}G = \left[\sum_{i=1}^{39} {}^{\circ}G_{MR(i)}, \sum_{i=1}^{39} {}^{\circ}G_{RR(i)}, \sum_{i=1}^{39} {}^{\circ}G_{HR(i)} \right] \quad (1)$$

The fuzzification method was selected as centroid, in view of the simplicity demanded for this job. It is important to remember that all this software should be working 24 hours and it should be attended many patients during short time.

$$y_q^{crisp} = \frac{\sum_{i=1}^{R} b_i^q \int_{Y_q} \mu_{B_q^i}(y_q) dy_q}{\sum_{i=1}^{R} \int_{Y_q} \mu_{B_q^i}(y_q) dy_q} \quad (2)$$

The given outputs correspond to the attention model: a) First time visiting the physician, b) First time and a predefined consequent attention and 3) Hard Nutritional Support (Figure 5).

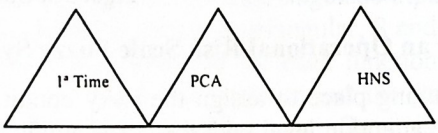

Figure 5. Membership function output

3 Results

At this moment, the DB is receiving the patient's information, however, many designs related with it have been developed: for example the visual interfaces and the corresponding formularies. The final DB design was completed using the commercial software ACCESS and using the ODBC transferred method. Additionally, this database contains some special request that allows to assign numerical values for each linguistic variables associated with the original MS scheme proposed by the INCMNSZ's nutrition department (Fig. 6). Besides, the DB was build to attend each patient during its last six hospital appointments; this fact permits the physician to evaluate the advances or the falling backs in the patient's health. This aspect is novel because the current commercial databases just give a general view on the patient response but they do not suggest any kind of treatment and, obviously, many of them are not provided by an artificial intelligence algorithm to define the corresponding diagnostic.

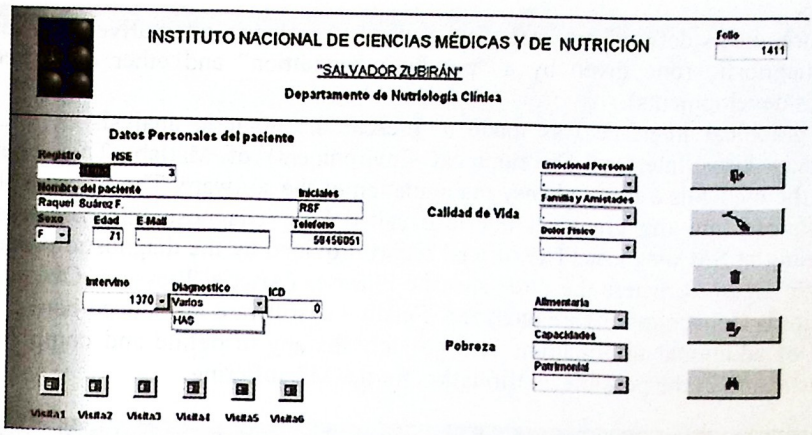

Figure 6. Lifestyle and general data for the treated patient.

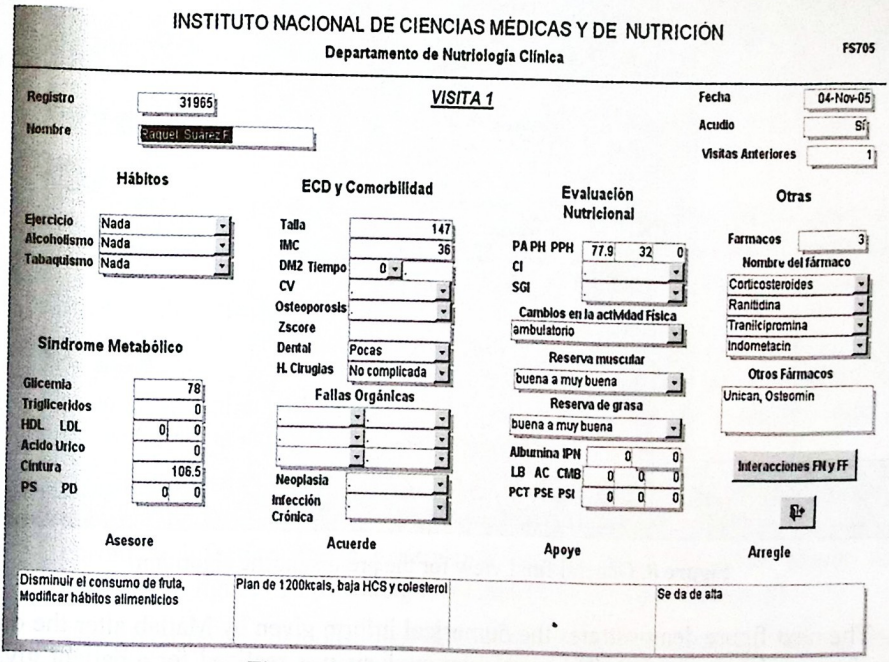

Figure 7. Clinical variables for each patient.

The figure 6 shows the DB view corresponding to the general data for each patient as well as those related with the lifestyle carried out by him or her. The next figure 7 demonstrates the frequently questions supplied by the medic to the patient and his/her clinical studies that were suggested before. Of course, these elements are used as the input data for the decision systems generated by the fuzzy arrangement.

The DB was linked with MATLAB in order to analyze the stored data. It is important to note the numerical representation considered at this point because this

procedure allows defining any statistical method to derive an alternative form for the same diagnostic (one given by a "possibility condition" and other described by statistics developments).

A graphical interface was made to present the numeric analysis by GUIDE (Graphical User Interface Development Environment) of Matlab. This interface allows the medicals a fast and easy manipulation of the software. The figure 8 shows the diagnose into the graphical interface called DIS (Diagnose Interface System) concerning at SM diagnose. The official report provided by the diagnostic process is shown in figure 8, where the three specific illnesses (Arterial Pressure, Obesity and Low Lipids Concentration) are analyzed. Finally the software introduced here gives a couple of additional information: 1) a possible therapy to define and complete the diagnostic and 2) the possible confirmation for the MS suffering.

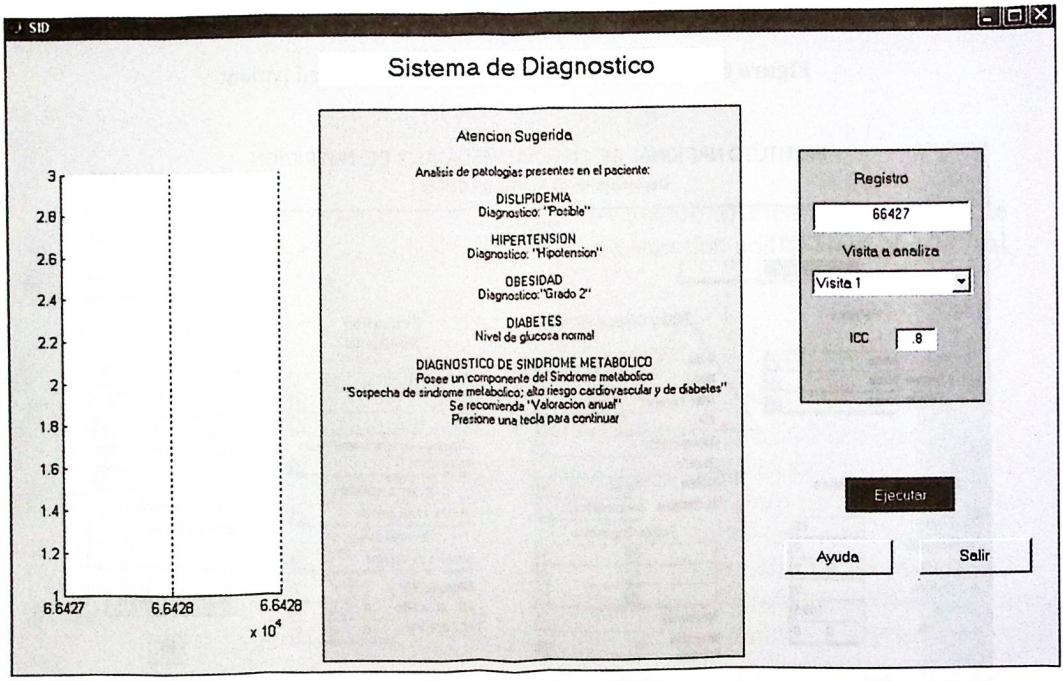

Figure 8. General final view for the pre-diagnostic algorithm.

The next figure demonstrates the numerical inform given by Matlab alter the data were imported from the DB. This particular analysis was realized for a patient given in the previous diagnosis analysis (Figure 8).

- Sex: Female. This patient is affected by the MS illness and by the type II Diabetes Mellitus (bad glucose regulation), high HDL, uric acid outside of the normal ranges and obesity (classified as level 1).

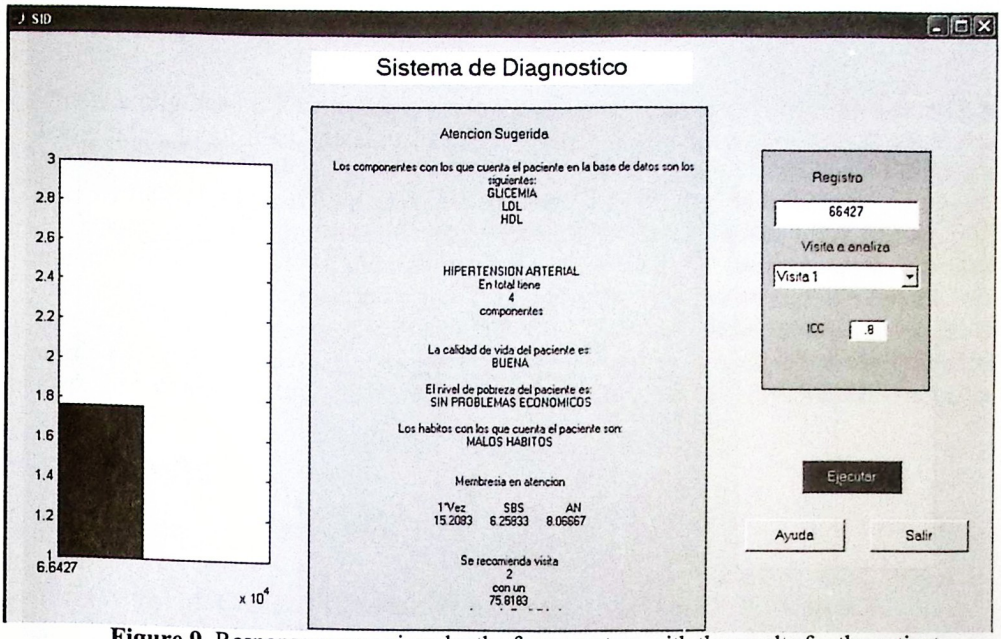

Figure 9. Response screen given by the fuzzy system with the results for the patient number 1. It should be noted the suggested therapy included at the bottom of the specific report.

The output in the graphic represents the defuzzification result, using the designed membership functions, the linguistic variables selected and the overall fuzzy inference model. The final value (1.7581) indicates the patient requires a new session to be attended with a medium urgency (possibility) determined by a 75.18 %. If the number in the response is 3, indicates the patient should be keep at the hospital to be observed by the medics and 1 with 0 % indicates the patient can be released for the treatment. To prove the workability of the suggested method, a second case is considered. The clinical values for this patient are:

- Sex: Female. She is a patient in death danger whit a lot of complications in her metabolic state. There is some evidence related with tumor activity, chronicle infection and a generalized organic fail. The given results are depicted in figures 10 (SM diagnose) and 11 (risk scale allocation).

The given results demonstrated this patient should be attended "before" the results derived before (patient 1), because the possibility or urgency to treat this patient is 90.91 %. At the same time, the results suggest a Risky Nutritional Support.

The analysis derived by the MS pre-diagnostiv system allows stating a well defined methodology to provide a range or degree of attendance and to define repeatability concepts to treat the MS illness, due the numerical algorithm introduced in this study.

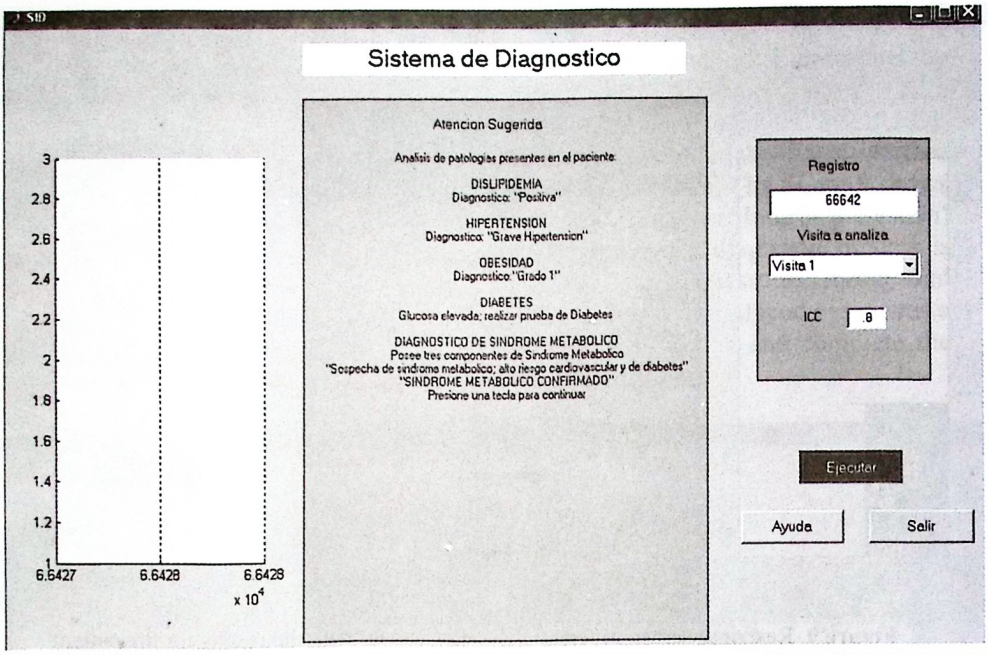

Figure 10. Results derived by the second patient's data from the diagnose algorithm.

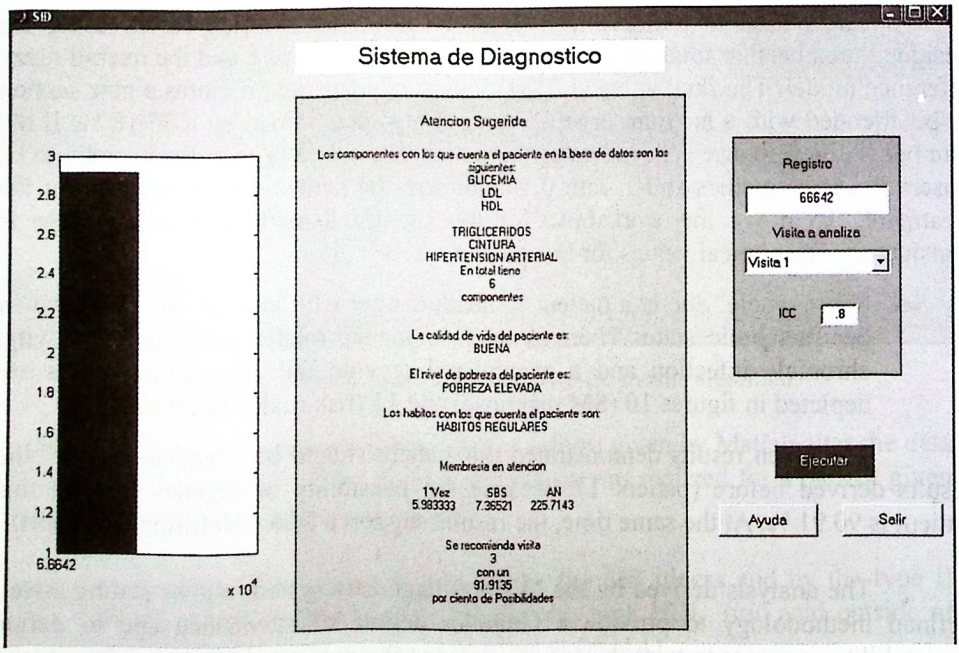

Figure 11. Results given by the risk analysis fuzzy system

4 Conclusions

The friendly design for the DB gives to the medic an easy way to acquire the enough knowledge about the clinical history, the lifestyle and some general aspect about the patient that the physician is attended. Besides, the complete elements considered in the DB fields allow defining a better manner to study the patient sufferings.

The linking process between the ODBC system interface (ACCES) and the numerical software (Matlab), give us the possibility to analyze (with advance statistics tools) and to treat the stored data using the well known decision capabilities of the hierarchical fuzzy logic method (a novel approach in the database field). This approach is really useful because many of the elements considered at the inference step are based on the medic opinions (they are with the INCMNSZ' nutrition department).

References

1. Castro, M. Gámez. "Historia Clínica", Reporte Clínico del HNCMNSZ
2. R. Porto. "Síndrome Metabólico. Enfoque Actual" , Revista Mexicana de Cardiología, 2001.
3. I. Chairez "Control inteligente aplicado a incubadoras neonatales", Tesis de Licienciatura, UPIBI, 2003.
4. C. Pérez, "Access 2003" , 2002, Mc. Graw Hill.
5. L. Zadeh, :A: " Fuzzy sets. " 1965 Inf. Cont. 8:338-353
6. H. Zimmermann. –J. " Description and optimization of fuzzy systems" Int. J. Gen. Syst. 2:209-215.
7. S. Gómez "Sistemas Difusos Jerárquicos"
8. L. Groop, M. Orho-Melander. The dysmetabolic Syndrome. J Intern Med 2001; 250
9. A. Zimmet "Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications"
10. Isselbacher, Braunwald "Principios de medicina Interna", Mc Graw-Hill.