

Breast Cancer Monitoring and Tracking System using Soft Computing and Expert System for Healthcare Support

Atiya Masood Rana¹, Somaya Abdullah Al Gamdi²,
Roa'a Abdulraheem bukhari³, Maryam Tami Al-Osaimi⁴

¹(Management Science, Yanbu University College, Pakistan)

^{2,3,4}(Management Science, Yanbu University College, Saudi Arabia)

Abstract: The decision process for selecting the best-suited follow-up treatment for suspected breast cancer cases are strongly dependent upon the correct diagnosis and assessment of the breast cancer risk. This study will develop an expert system with soft computing methods by using Fuzzy logic and Bayesian network. This system has five input variables and one output variable (risk status). The output variable is a value from 1 to 4; representing low risk status, intermediate risk status, high risk status and very high risk. This study will help to diagnosis and monitor breast cancer at home.

Keywords: Breast cancer diagnosis, fuzzy intelligent technique, Soft Computing, Expert system, Probabilistic reasoning, Bayesian Network.

I. Introduction

Cancer is a leading life threatening disease that cause cancer cell or lumps in a body to change and grow out of control. One of the most common cancer among women is breast cancer and it is one of the leading cause of cancer related death in women. According to research one in ten ratio of all new cancers diagnosed worldwide each year is the cancer of the female breast cancer and it is the most common cancer in women in both developing and developed countries [1]. In 2011, 39,520 death cases due to the breast cancer were recorded in USA [2]. This research is shown in Table 1.

Age	In Situ Cases	Invasive Cases	Deaths
Under 40	1,780	11,330	1,160
Under 50	14,240	50,430	5,240
50-64	23,360	81,970	11,620
65+	20,050	98,080	22,660
All ages	57,650	230,480	39,520

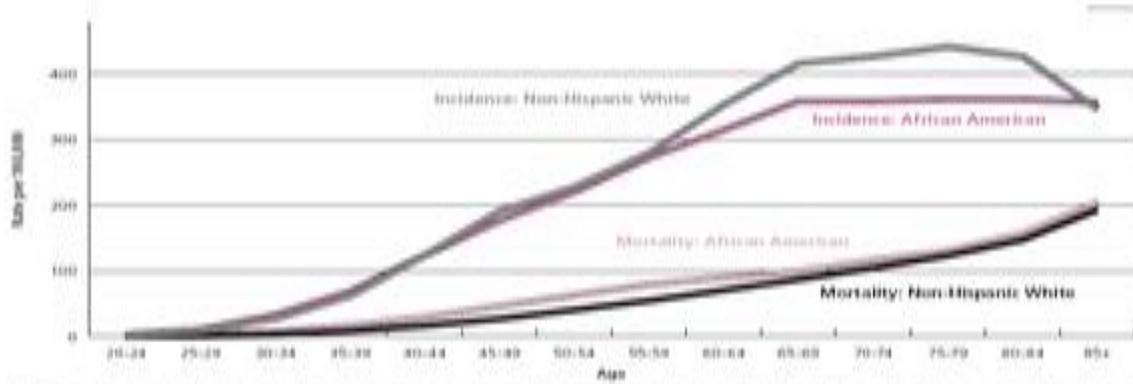
*Rounded to the nearest 10.
Source: Total estimated cases are based on 1995-2007 incidence rates from 46 states as reported by the North American Association for Central Cancer Registries. Total estimated deaths are based on data from US Mortality Data, 1969-2007, National Center for Health Statistics, Centers for Disease Control and Prevention.
American Cancer Society, Surveillance Research, 2011

Table 1: Estimated New Female Breast Cancer Cases and Deaths by Age, US, 2011

In Saudi Arabia, the number of new cases of all cancer is 2741 including about 19.9% of breast cancer in women and it is usually occurs in women at the age of 52. This ratio can be reduced, if cancer diagnosis is done in early stage. The National Cancer Institute estimates that approximately 2.6 million US women with a history of breast cancer were alive in January 2008, more than half of whom were diagnosed less than 10 years earlier. Currently, digital mammography is one of the most promising cancer control strategies in earliest stages but not all the hospitals have digital mammography and it is one of the expensive techniques for cancer control. In recent years Intelligent system that uses different Artificial Intelligence methods such as an expert system and soft computing has been used in medical field for diagnosing and treatment of many diseases. This can be true for breast cancer management. Therefore we planned to use expert system with soft computing for diagnosing the risk status of breast cancer.

II. Analytical Study

According to the American cancer society estimates that breast cancer incidence and death rates generally increase with age [2] as shown in Figure- 1. 95% of breast cancer occurred in women 40 years of age and older.



Source: Incidence: North American Association of Central Cancer Registries. Mortality: National Center for Health Statistics, Centers for Disease Control and Prevention, as provided by the Surveillance, Epidemiology, and End Results Program, National Cancer Institute. American Cancer Society, Surveillance Research, 2011

Fig-1:Age-Specific Female Breast Cancer incidence(2004-2008) and Mortality (2003-2007)Rates

Race is also an important factor of breast cancer and according to the American cancer society breast cancer incidence rates are higher in non-Hispanic white women compared to African -American women for most age groups that shows in Figure 1. However, African- American women have a higher incidence rate before 40 years of age [2]. Breast cancer is more common in African- American women. Asian, Hispanic, and Native-American women have a lower risk of developing and dying from breast cancer [3].

Breast Cancer risk is higher among whose close blood relationship , such as sister, mother or daughter has this risk . According to the research those having a relative with first degree cancer diagnosis has threat of a double risk and having 2 first-degree relatives increases her risk about 3-fold [3].

Age at first birth is one of the main reason of breast cancer . Women who have had no children or who had their first child after age 30 have a slightly higher breast cancer risk. Also age at first menarche is an important factor of breast cancer and according to the research women who have had more menstrual cycles because they started menstruating early (before age 12) and/or went through menopause later (after age 55) have a slightly higher risk of breast cancer. Late menopause increases the risk of breast cancer. Women who have undergone the menopause have a lower risk of breast cancer than pre-menopausal women of the same age and childbearing pattern [3].

Although breast cancer has many causes but in this research we consider only five factors . individual risk factors and certainty factors is shown in Table2.

Risk Factors	Breast cancer Certainty Factor
Age	
≤ 29	0.06%
30-39	0.4%
40-59	2.3%
60-70	3.8%
Race	
Mixed/Others	3%
Black	7%
Hispanic	11%
White	72%
First degree relative with breast cancer	
0	5%
1	20%
>=2	40%
Age at menarche	
12 -13	12%
< 12 or ≥15	44%
Age at first birth	
< 20	17%
20 – 29	31%
≥ 30 or None	49%

Table2: Individual risk factors and it certainty factors

III. Discussion With Proposed system

There is a strong need of monitoring and tracking that can show the severity and certainty of breast cancer . This proposed system is an Intelligent system that is based upon probabilistic reasoning and fuzzy logic . Fuzzy logic and probabilistic reasoning are techniques of soft computing and these techniques are used in inference engine of Expert system.

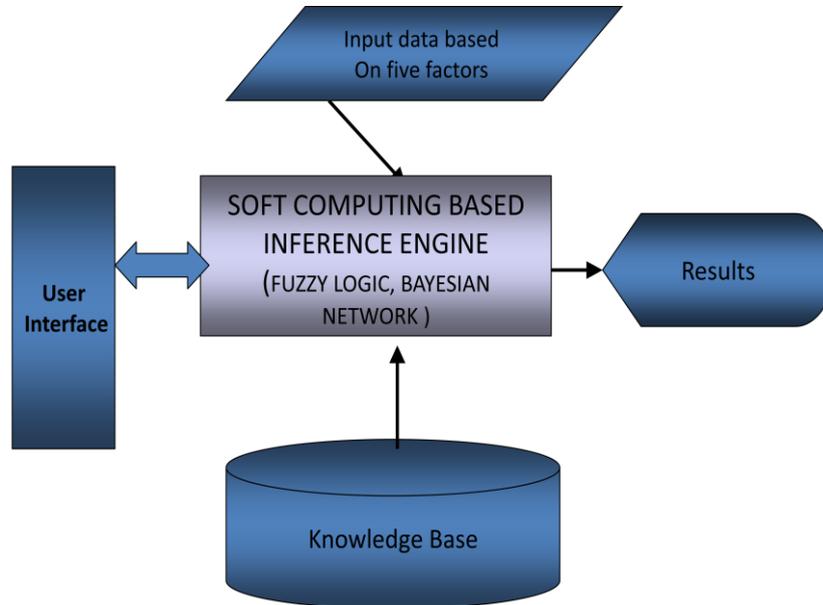


Fig- 2: Architecture of Monitoring Breast Cancer

3.1 Soft Computing

One of the most important feature of this research is diagnoses. The soft computing approach with foundation of Fuzzy Logic, Bayesian Network can be helpful for diagnosis of breast cancer. Inference engine will use five inputs and apply Bayesian and fuzzy algorithm in order to extract knowledge as output, so that in future it will help physicians and patient to diagnose breast cancer . Furthermore, this system records patient’s age ,race, gender that is a future research requirement to determine consequences based on age-group , gender-group and Race -group. Figure-2 expresses the methodological architecture of breast cancer. The Soft Computing based inference engine develops the inference using algorithms of Fuzzy Logic and Bayesian network taking the data input from user . The User Interface provides access and interaction between system and user. User could be doctor, paramedic staff, or patient itself [4].

3.2 Fuzzy Logic

Fuzzy-type uncertainty measure, expressed in verbal form, is transferred to the same types of numerical representation; the meaning is, on the one hand, a philosophical explanation, beyond the computer representation process; on the other hand, it is a simple look-up table representation attaching the 0 to 1 values to different verbal set membership measures [5].Figure-3 shows the structure of fuzzy logic system.

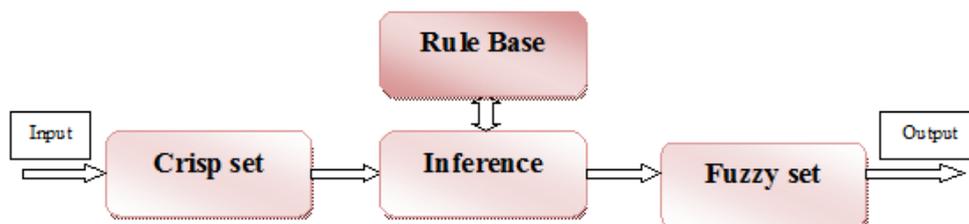


Fig 3: Architecture of Fuzzy Logic System

Fuzzy logic system has three main operations. First, fuzzification that transformed crisp sets to fuzzy sets, second Rule based inference. The fuzzy output needs to be interfaced to the crisp domain by the defuzzifier by using suitable membership functions and it is called defuzzification. In this research we also use these operation and it shows in Table 4-9.

3.3 Bayesian Network

A classic probabilistic-related model is the Bayesian; it concerns certain classes of events within the realm of the total observation. The Bayesian model supposes well-defined classes, reliable sampling methods, and a relevant amount of data for each class. The condition of *independence* belongs also to the model hypothesis. The Bayesian model establishes the well-known relation between a priori and a posteriori estimations, i.e., probability of an effect, if we know the probability of a certain cause, and vice versa, the probability of a cause based on the known probability of an effect[5]. In our research we determine conditional probability of breast cancer, where age, race, first degree relative with breast cancer, Age at menarche and age at first birth are known probabilities.

IV. Design of the system

As mentioned above that this research focused on five factors and It shows in Table 3.

Age	Less than or equal 29	Between 30 And 39	Between 40 And 59	Between 60 And 70
Race	Mixed	Black	Hispanic	White
Family History	0	one or more		
Age at menarche	Between 12 and 13 years old	Less than 12 years old	More than or equal 15 years old	
Age at first birth	Less than 20 years old	between 20 and 29 years old	More than or equal 30 years old	None

Table3: Individual risk factors and it Ranges

In our research first we describe input variables with membership function and certainty factors. In second step, we introduce the output variable with its membership functions and it is shows in figure4.

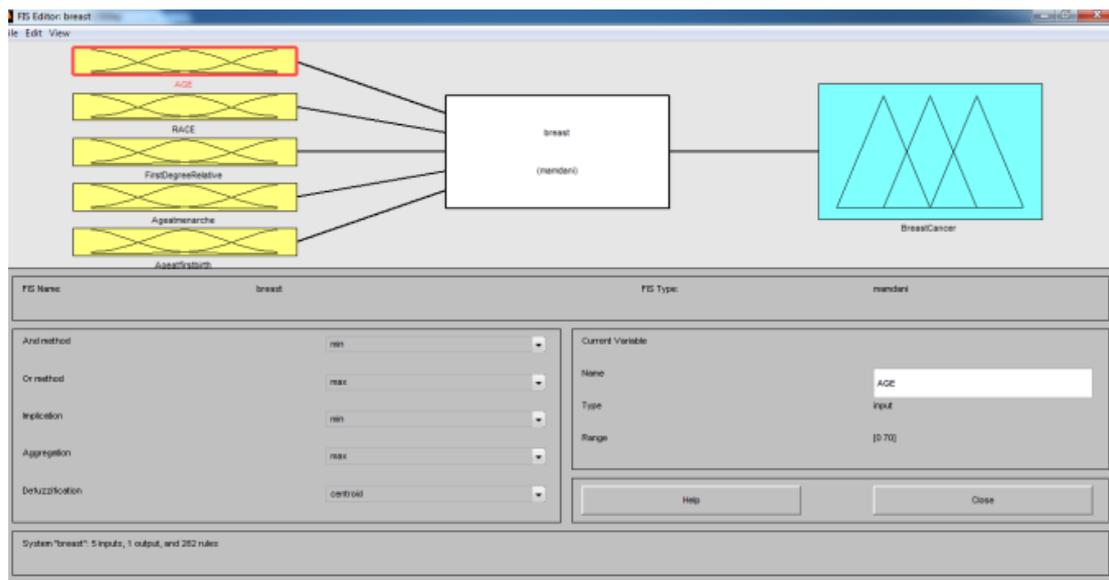


Fig- 4: Input and Output variables

Age input variable has four fuzzy sets "Low, Medium, High and Very High" with individual probabilities.

Risk Factors	Fuzzy Set	Breast cancer Risk
Age		
≤ 29	Low	0.06%
30-39	Medium	0.4%
40-59	High	2.3%
60-70	Very High	3.8%

Table 4:Fuzzy Sets Of Age

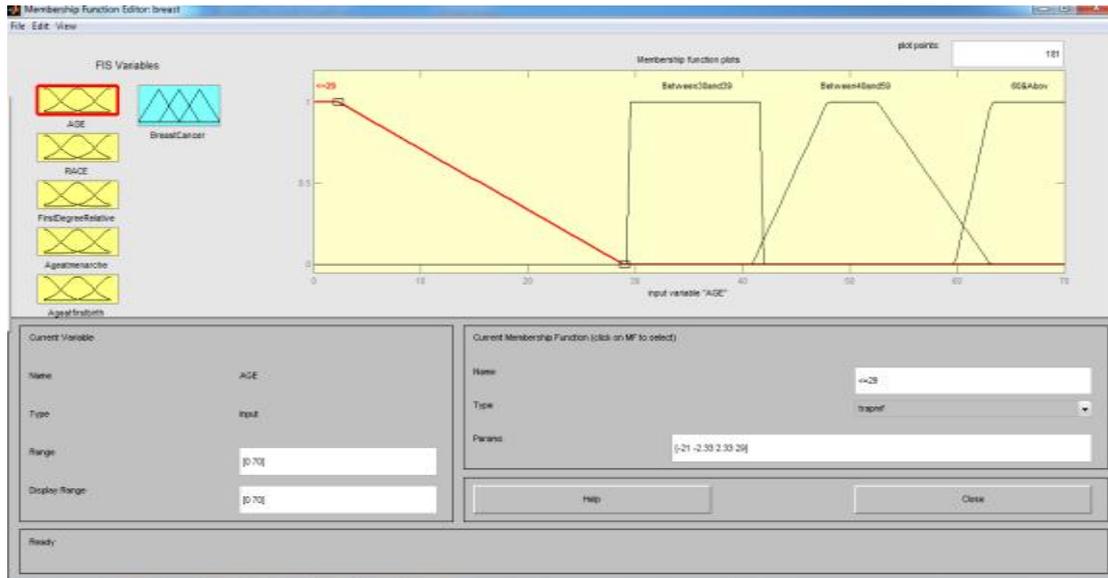


Fig-5: Membership Functions for Age

Race input variable has four fuzzy sets "Low, Medium, High and Very High" with individual probabilities .

Race		
Risk Factors	Fuzzy Set	Breast cancer Risk
Mixed/Others	Low	3%
Black	Medium	7%
Hispanic	High	11%
White	Very High	72%

Table 5:Fuzzy Sets Of Race

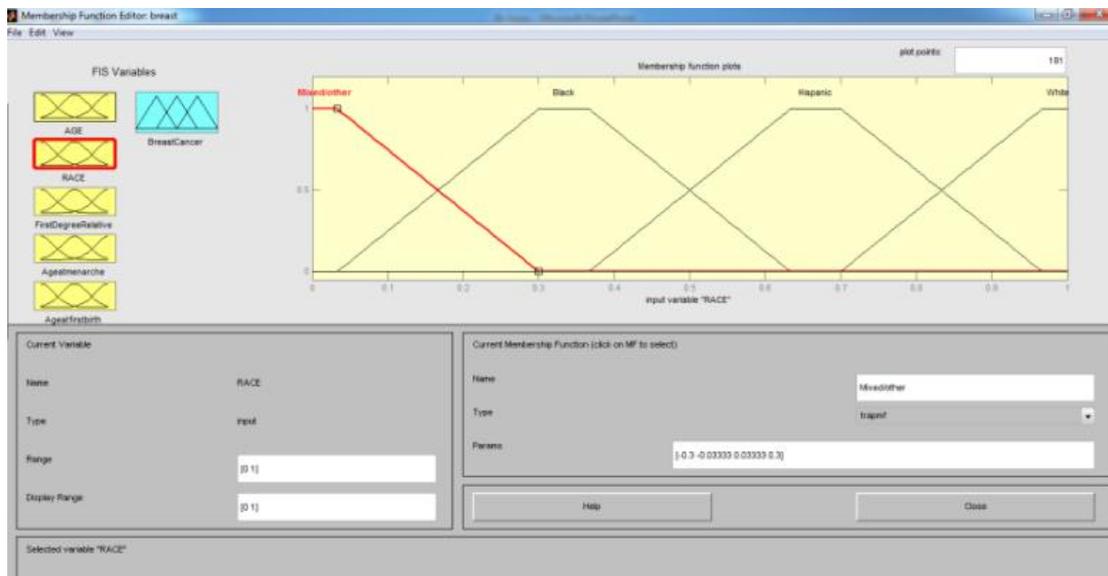


Fig- 6: Membership Functions for Race

First degree relative with breast cancer input variable has three fuzzy sets "Low, Medium and High " with individual probabilities

First degree relative with breast cancer		
Risk Factors	Fuzzy Set	Breast cancer Risk
0	Low	5%
1	Medium	20%
>=2	High	40%

Table 6:Fuzzy Sets Of First Degree Relative with breast cancer

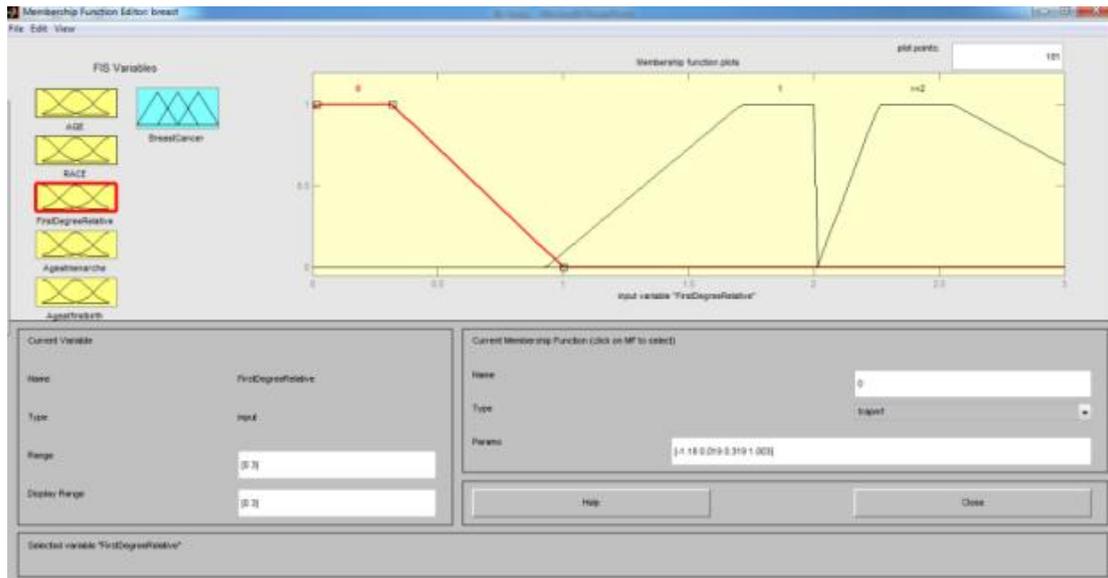


Fig-7 :Membership Functions for First degree relative with breast cancer

Age at menarche input variable has two fuzzy sets "Low, and very High " with individual probabilities .

Age at menarche		
Risk Factors	Fuzzy Set	Breast cancer Risk
12 -13	low	12%
< 12 or ≥15	Very High	44%

Table7:Fuzzy Sets Of Age at Menarche

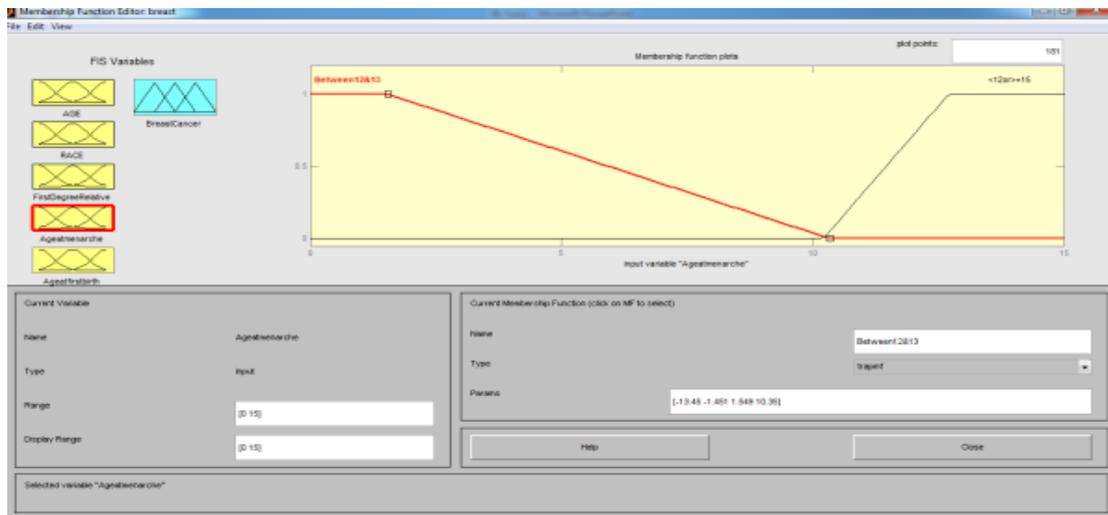


Fig- 8 :Membership Functions for Age at Menarche

Age at First Birth input variable has three fuzzy sets "Low, Medium and High " with individual probabilities

Age at first birth		
Risk Factors	Fuzzy Set	Breast cancer Risk
< 20	Low	17%
20 – 29	Medium	31%
≥ 30 or None	High	49%

Table8:Fuzzy Sets Of Age at First Birth

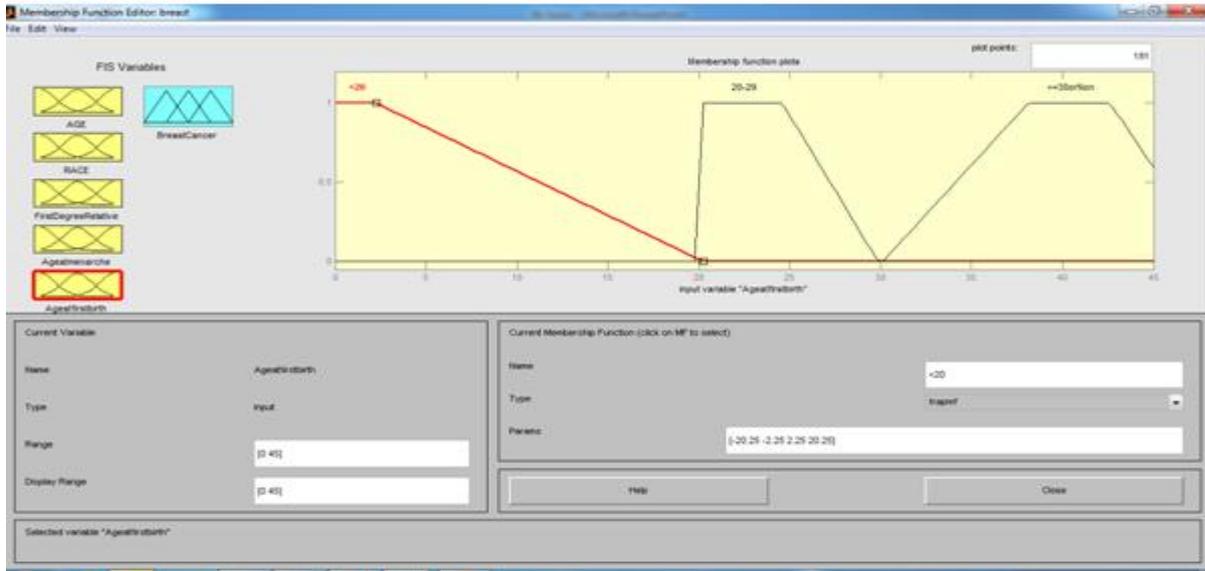


Fig-9 :Membership Functions for Age at First Birth

The "goal" of the system is to identify risk status of breast cancer. The output variable is a value from 0 to 4; representing Low Risk status, Medium Risk status High Risk status and very High Risk status.

Output		
Risk Status	Range	Fuzzy Set
	0-1	Low
	1-2	Medium
	1-3	High
	2-4	Very High

Table 9 FUZZY Sets Of Output Variable Risk Status

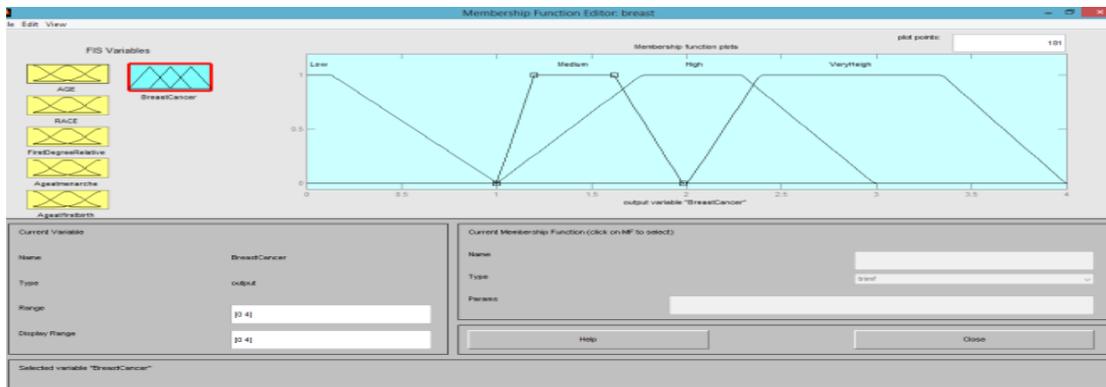


Fig-10 :Membership Functions for Risk Status

In this system, inference engine also performs probabilistic reasoning so called belief networks or bayesian Networks. Bayesian network gets probabilities about individual factor from probability tables that stored in the knowledge base and calculate the probability of breast cancer. In this system posterior probability that is breast cancer probability depends on five factors that consider prior probability and hard to compute manually. Therefore we use BayesiaLab for belief network .It shows in Figure 12

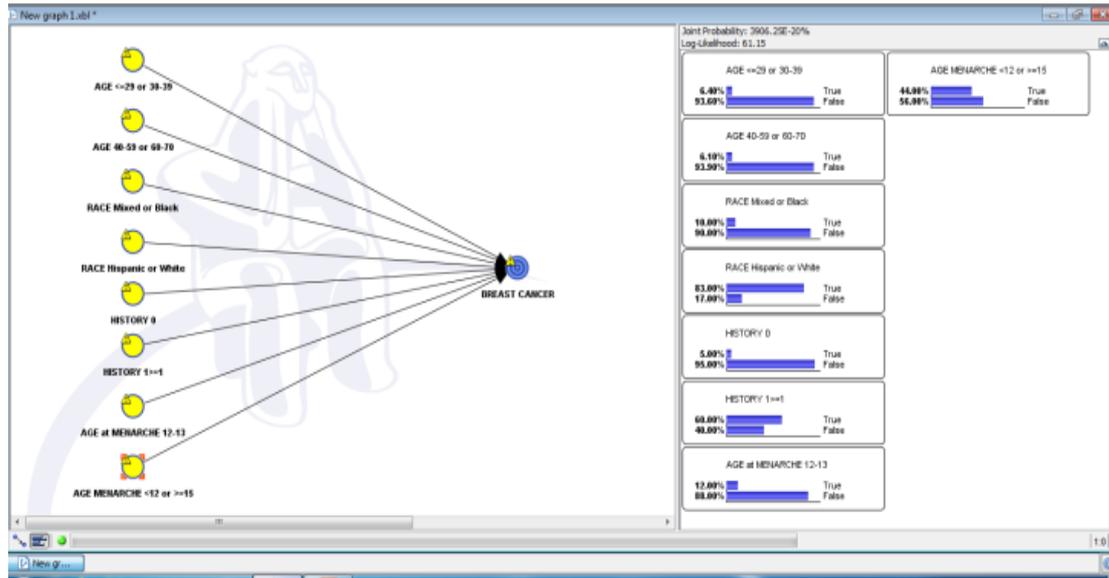


Fig-12 :Bayesian Network for Risk Status

V. System testing and discussion

System has been tested by experts and rules are defined by knowledge engineer . Some of the rules are shown in Table 9 and Figure.13 .

1.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Mixed) and (History = 0) and (12 ≤ Age at menarche ≤ 13) and (Age at first birth < 20 or 20 ≤ Age at first birth ≤ 29)) THEN the certainty factor is LOW
2.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Mixed) and (History = 0) and (12 ≤ Age at menarche ≤ 13) and (Age at first birth > 30 or Non)) THEN the certainty factor is LOW
3.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Mixed) and (History = 0) and (Age at menarche < 12 or ≥ 15) and (Age at first birth < 20 or 20 ≤ Age at first birth ≤ 29)) THEN the certainty factor is LOW
4.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Mixed) and (History = 0) and (Age at menarche < 12 or ≥ 15) and (Age at first birth > 30 or Non)) THEN the certainty factor is LOW
5.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Mixed) and (History ≥ 1) and (12 ≤ Age at menarche ≤ 13) and (Age at first birth < 20 or 20 ≤ Age at first birth ≤ 29)) THEN the certainty factor is MEDIUM
6.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Mixed) and (History ≥ 1) and (12 ≤ Age at menarche ≤ 13) and (Age at first birth > 30 or Non)) THEN the certainty factor is HIGH
7.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Mixed) and (History ≥ 1) and (Age at menarche < 12 or ≥ 15) and (Age at first birth < 20 or 20 ≤ Age at first birth ≤ 29)) THEN the certainty factor is MEDIUM
8.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Mixed) and (History ≥ 1) and (Age at menarche < 12 or ≥ 15) and (Age at first birth > 30 or Non)) THEN the certainty factor is HIGH
9.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Black) and (History = 0) and (12 ≤ Age at menarche ≤ 13) and (Age at first birth < 20 or 20 ≤ Age at first birth ≤ 29)) THEN the certainty factor is LOW
10.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Black) and (History = 0) and (12 ≤ Age at menarche ≤ 13) and (Age at first birth > 30 or Non)) THEN the certainty factor is LOW
11.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Black) and (History = 0) and (Age at menarche < 12 or ≥ 15) and (Age at first birth < 20 or 20 ≤ Age at first birth ≤ 29)) THEN the certainty factor is LOW
12.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Black) and (History = 0) and (Age at menarche < 12 or ≥ 15) and (Age at first birth > 30 or Non)) THEN the certainty factor is LOW
13.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Black) and (History ≥ 1) and (12 ≤ Age at menarche ≤ 13) and (Age at first birth < 20 or 20 ≤ Age at first birth ≤ 29)) THEN the certainty factor is MEDIUM
14.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Black) and (History ≥ 1) and (12 ≤ Age at menarche ≤ 13) and (Age at first birth > 30 or Non)) THEN the certainty factor is HIGH
15.	IF ((age ≤ 29 or 30 ≤ age ≤ 39) and (Race = Black) and (History ≥ 1) and (12 ≤ Age at menarche < 12 or ≥ 15) and (Age at first birth < 20 or 20 ≤ Age at first birth ≤ 29)) THEN the certainty factor is MEDIUM

Table 9:Rule Base Of the System

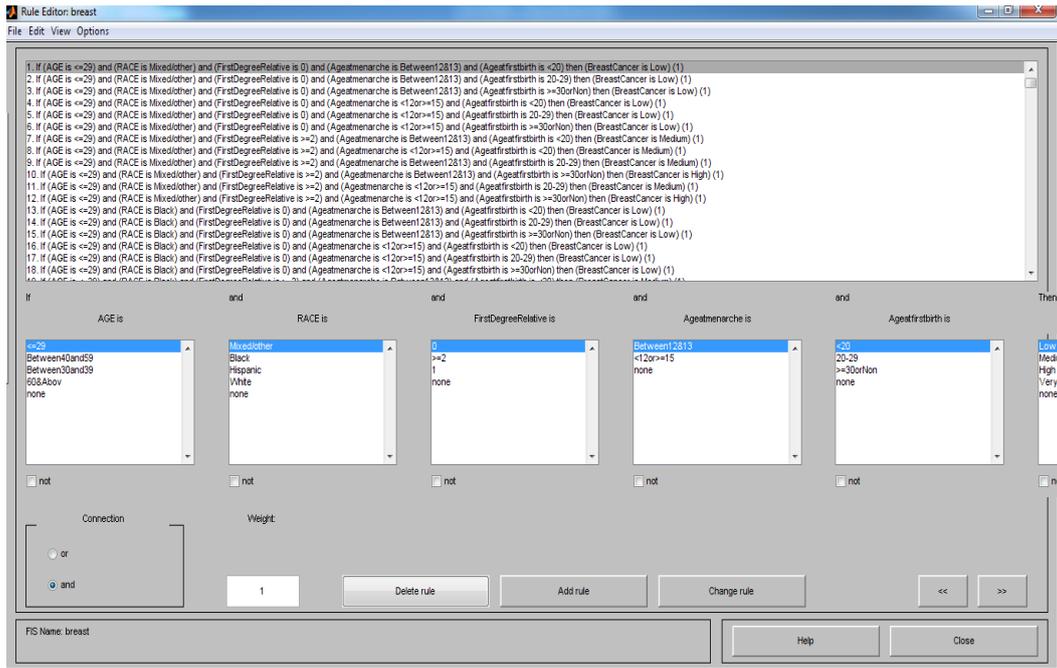


Fig-13:Rule Base Of the System

Figure 14- 17 shown surface viewer of some fields as follow

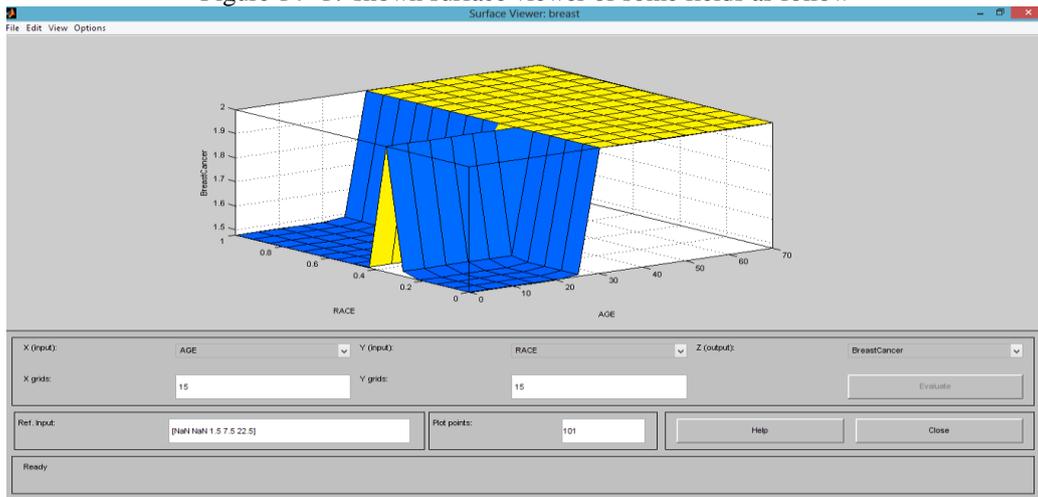


Fig- 14. Surface Viewer of Age and Race

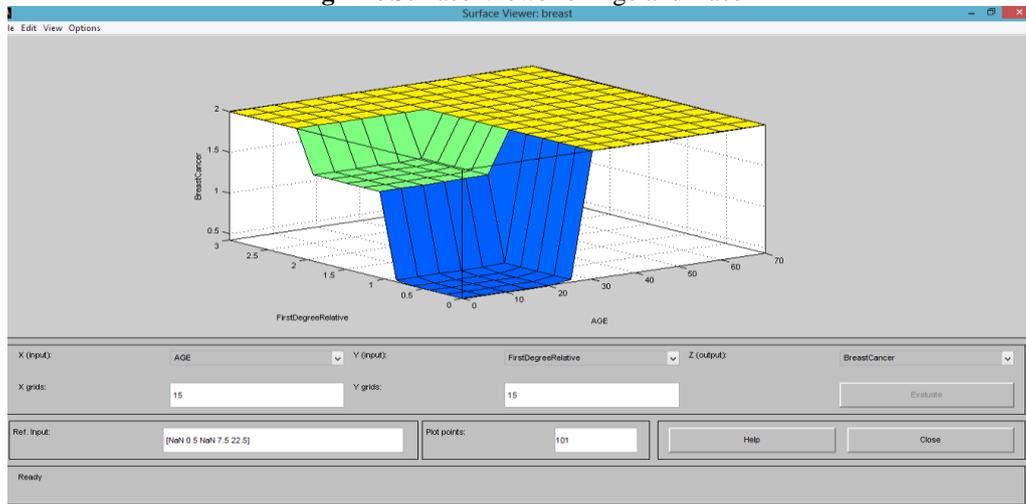


Fig-15. Surface Viewer of Age and First Degree Relative

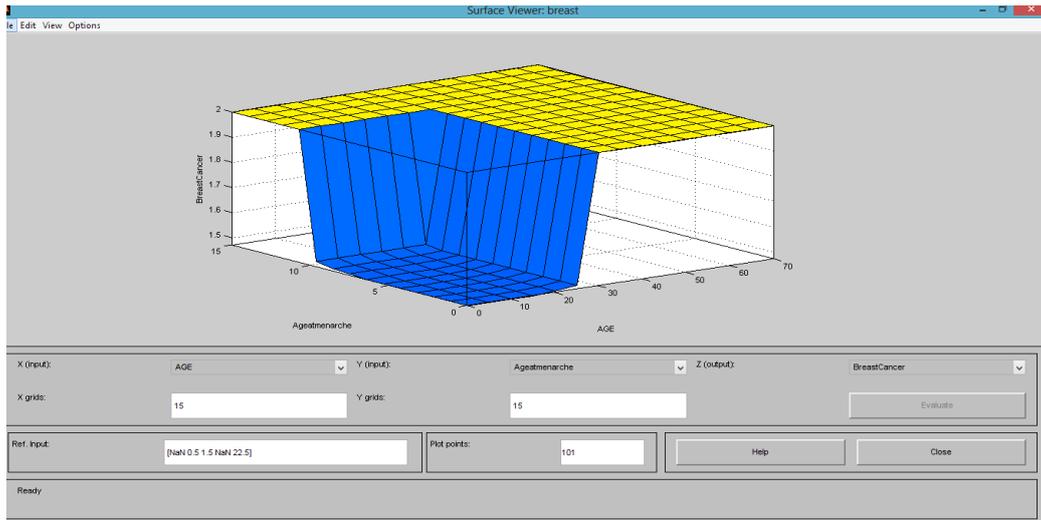


Fig-16. Surface Viewer of Age and Age at Menarche

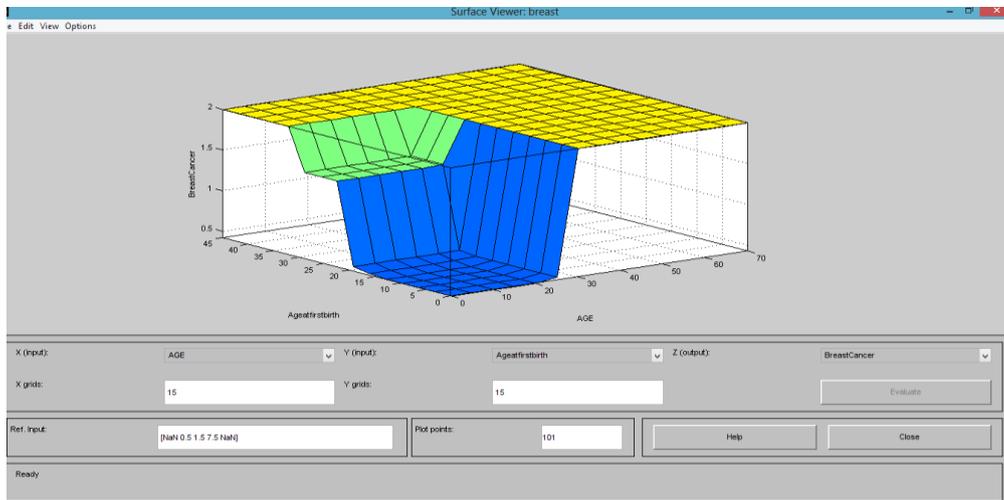


Figure 17. Surface Viewer of Age and Age at First Birth

After design of an Expert system, system has already implemented as it shown in the figure 18



Figure 18. Breast Cancer System

VI. Conclusion

This paper describes design of expert system with soft computing techniques for identification of breast cancer risk status, which can be used by specialized doctors and patient for cancer treatment. The system design is based on membership functions, input variables, output variables , rule base and probabilistic reasoning. This research also shows fuzzy set by using MatLab and belief network by using BayesiaLab .This System has been tested by expert doctors . It is well said that “Health is Wealth” and this system helps patient in monitoring and diagnosis of breast cancer at home.

Acknowledgements

We acknowledge Dr.GhadderAbdeen in Aldamamhospital and Dr. KhaledAbulkhair Associate Consultant ,Oncology in IMC Jeddah for their effort in this research.

References

- [1]. Freddie Bray, P. M. (2004). The changing global patterns of female breast cancer incidence and mortality. *BioMedCentral* , 229-239.
- [2]. Carol DeSantis, M. ., (2011). Breast Cancer Facts & Figures . American Cancer Societ .
- [3]. risk factors for breast cancer. (2014, 01 31). Retrieved from Breast Cancer: <http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-risk-factors>.
- [4]. HussainSaleem&et. al., (2010), “Traceability Management Framework for Patient Data in Healthcare Environment”, In Proc. IEEE Intl. Conf. on Computer Science & Information Technology, Chengdu, China, pp.264-8
- [5]. Masood, A. (2011). MEDICAL EXPERT SYSTEM Theory, Implementation and Application: Knowledge Discovery from Medical Data Using a Bayesian Network and Fuzzy Logic . LAP LAMBERT.