ORIGINAL PAPER

Design a Fuzzy Rule-based Expert System to Aid Earlier Diagnosis of Gastric Cancer

Reza Safdari[1], Hadi Kazemi Arpanahi[1], Mostafa Langarizadeh[2], Marjan Ghazisaiedi[1], Hossein Dargahi[3], Kazem Zendehdel[4]

[1]Department of Health Information Management, School of Allied Medical Sciences, Tehran University of Medical Sciences, Iran

[2]School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran

[3]School of Allied Medicine, Member of Health Information Management Research Center, Tehran University of Medical Sciences, Tehran, Iran

[4]Cancer Research Center, Cancer Institute of Iran, Tehran University of Medical Sciences, Tehran, Iran

Corresponding author: Hadi Kazemi Arpanahi. Ph.D. Candidate of Health Information Management, Department of Health Information Management, School of Allied Medical Sciences, Tehran University of Medical Sciences, Iran. E-mail: h-kazemi@razi.tums.ac.ir

ABSTRACT

Introduction: Screening and health check-up programs are most important sanitary priorities, that should be undertaken to control dangerous diseases such as gastric cancer that affected by different factors. More than 50% of gastric cancer diagnoses are made during the advanced stage. Currently, there is no systematic approach for early diagnosis of gastric cancer. **Objective:** to develop a fuzzy expert system that can identify gastric cancer risk levels in individuals. **Methods:** This system was implemented in MATLAB software, Mamdani inference technique applied to simulate reasoning of experts in the field, a total of 67 fuzzy rules extracted as a rule-base based on medical expert's opinion. **Results:** 50 case scenarios were used to evaluate the system, the information of case reports is given to the system to find risk level of each case report then obtained results were compared with expert's diagnosis. Results revealed that sensitivity was 92.1% and the specificity was 83.1%. **Conclusions:** The results show that is possible to develop a system that can identify High risk individuals for gastric cancer. The system can lead to earlier diagnosis, this may facilitate early treatment and reduce gastric cancer mortality rate.

Keywords: fuzzy logic, gastric cancer, risk analysis, preliminary detection.

1. INTRODUCTION

Today cancer is the most aggressive and the most fatal disease. Approximately, 12% of worldwide human mortality dies from cancer each year. Although lung and breast cancers are the most common cancers in the men and women, respectively, cancers of the gastrointestinal tract, including esophageal, stomach, liver, colon, and pancreas cancers, are responsible for approximately 3 million incidences and over 2 million deaths each year making them the most frequent cancers worldwide. Cancers of the Stomach are among the most frequent gastrointestinal tract cancers (1). The stomach cancer is the fourth most common cancer and second leading cause of cancer-related death in the world (2). According to Globocan 2008, gastric cancer is one of the most common cancer in Iran (3). Its incidence is particularly high in the northwest of the country. Patterns of gastric cancer in our country are similar to those reported from high-risk regions worldwide (4). "Center of Non-Communicable Disease Management of Iran Health Ministry" reports annually more than 40 thousand Iranian die due to gastric cancer (5). Northwestern province of Ardabil has been reported to have the highest incidence rate in the country with age standardized incidence rates of 49.1 and 25.4 per 100,000 in men and women, respectively (6).

Early diagnosis of gastric cancer is very crucial for reducing the mortality rate of gastric cancer but gastric cancer is asymptomatic or has non-specific signs at the early stages, patients usually diagnosed in very advanced stages as the results of this, tumor metastases to adjacent organs, the treatments are hopeless and patients die in a short time. (7) some patients even with having suspected symptoms as well as lose weight, anemia, dysphagia, the prognosis for gastric cancer remains in adequate (8).

Fuzzy logic

Fuzzy logic was first introduced by Zadeh in 1965; it is the extension of the Boolean logic that substitutes the degree of membership with crisp membership. In fuzzy sets each element has a degree of belong to membership,

Category	Variable	Relative weights	Variable Descriptions	Membership Functions
Factors are related toClinical history	Helicobacter pylori infection	0.9	Status of H pylori infection	NO=0Yes=1
	Gastric atrophic	0.8	Status atrophic gastritis	NO=0Yes=1
	Peptic ulcer	0.5	Status of peptic ulcer	NO=0Yes=1
	Gastric surgery	0.7	Prior gastric surgery	NO=0Yes=1
	Pernicious anemia	0.4		NO=0Yes=1
	Family history	0.9	History of gastric cancer in family	Non-relatives with gastric cancerSecond degree relatives with gastric cancerFirst degree relative relatives with gastric cancer
Socio-Demo-graphicFactors	Age	0.7	Age of patient in year	20-45 = young35-65 = middle- Aged55-100 = old
	Gender	0.4	Sex of patient as in male or female	Male =1Female=0
	Educational-year	0.6	Level of literacy	LiteracyElementarySecondaryHigh educational
Smoking factors	Tobacco smoking	0.6	Cigarette smoking risk habit	NO=0Yes=1
	Pack-years Smoking	0.7	Number of pack-years = (packs smoked per day) × (years as a smoker)	Never smokeEx-smokerCurrent smoker
Nutritional factors	Vegetable and fruit	0.7	Low Consumption (serves a day)	NO=0Yes=1
	High Salt intake	0.8	High consumption salt intake	NO=0Yes=1
	Content of nitrate compositions	0.8	High consumption nitrate	NO=0Yes=1

Table 1. Input variable descriptions, categorizations and Membership Functions

ranging from 0 to 1 but in Boolean logic crisp sets are used, which has any two values 0 and 1 (9).

For design a fuzzy system, initially, input-output variables are specified, then fuzzy sets and corresponding membership functions were defined. Subsequently, rule base of system formed using condition statements (if-then) based on expert's opinion. Defuzzification was implemented using appropriate defuzzification technique and fuzzy sets were converted into classical sets and numbers (10).

Fuzzy decision making mechanisms use symbolic expressions such as big, low, small, little, etc. instead of numeric expressions (11). Fuzzy logic has been identified as a substantial tool that is used to model human decision making of an expert in a computer program such that the program can mimic the expert ability to solve problem (12). Due to intrinsic imprecision and uncertainties in medical practice, the advantages provided by fuzzy logic for dealing with imprecise and uncertainty situation, makes it appropriate approach to overcome this complexes in the medical data (13). Therefore this study utilized of fuzzy logic to representation uncertainty in reasoning of medical experts to estimate gastric cancer risk.

2. AIM

The motivations of suggested system in this work are:
- To improve prognosis and preliminary measures for gastric cancer.
- To encourage appropriate referral and early diagnosis in the general populations and in high risk groups.

The aim of this study is to develop a fuzzy expert system for screening and evaluating the risk level of gastric cancer disease. Fuzzy logic was used to representation of domain knowledge for assessing gastric cancer risk in the system.

3. MATERIALS AND METHODS

Development process and architecture of the system

There are two phases in design of this system; the initial phase before the fuzzy expert system was developed that called knowledge extraction and the development phase of this proposed fuzzy expert system.

3.1. Knowledge extraction

The success of fuzzy expert system greatly depends on the proficiency of domain experts. For this purpose in this research the requisite information was gathered using interviews with seven specialist physicians in the field of gastric cancer (four oncologist and three cancer epidemiologist). Their opinions were obtained about input and output variables of the proposed model, it is very important to take into consideration all risk factors that show a substantial role in creating of gastric cancer, moreover, the degree of significance also needs to be evaluated. Thus, variables, parameters and their weighted scores were defined. Ultimately thirteen features are selected as influencing factors in leading to gastric cancer. These risk factors are considered as the input variables of the Fuzzy expert system with single output that reflect the risk status, which is further divided into " low risk ", "moderate risk", "High risk" and "very high risk" (Table 1).

3.2. Model development

The development of the system was decided upon after a thorough review of literature and conducting interview with expert physicians.

The fuzzy inference system developed in this study employs the Mamdani type fuzzy inference technique, which is made up the following steps:

a) Fuzzification of input variables

At this stage, the fuzzy sets and corresponding membership functions were defined, crisp inputs are changed to fuzzy values through the use of linguistic variables (Table 1).
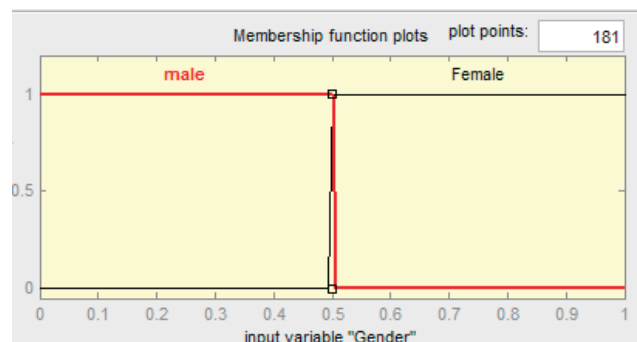

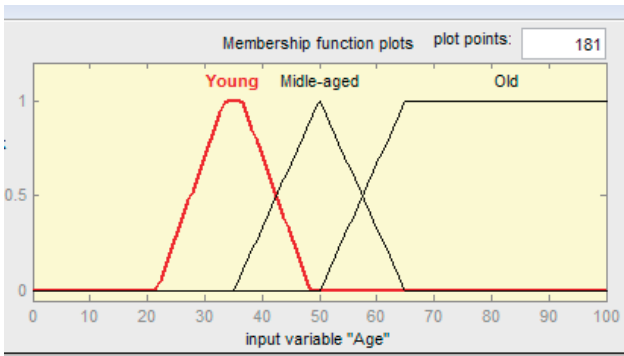Figure 1. Membership function plot for input variable Gender

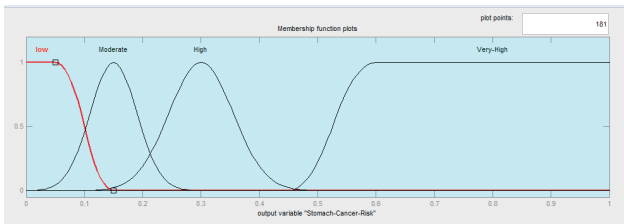**Figure 2. Membership function plot for input variable AGE**

Figure 3. Graphical representation of membership functions of Output variables for fuzzy expert system to prognosis gastric cancer

The fuzzification process defines the translation of input variables into a value between 0 and 1; this determines the degree of membership of each input variable. Figure 2-7 represents examples of the membership functions for input and output variables.

b) Fuzzy Rule Base

Rule base is the main component of a fuzzy inference system and results in such system relied on the fuzzy rule base. This system consists of a set of fuzzy rules that are formed according to input variables; together with their relevance concerning the decision process of gastric cancer risk, which presented by clinical experts and can vary somewhat from physician to physician. Since some of these risk factors stand-alone and in combination with other risk factors can effective in leading to gastric cancer. There by in order to reduce evaluating rules and even efficiency the performance of fuzzy logic system, risk factors were classified into four categories and then for each one of the groups separately Mamdani inference model was implemented. The final output of the system constitutes from accumulating the maximum of such groups (Figures 2-4).

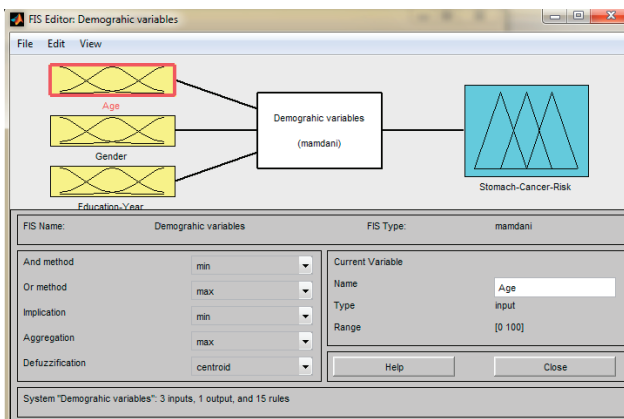Rules that constitute the rule-base were formulated with

Figure 4. Mamdani fuzzy model of demographic variables in in leading to gastric cancer
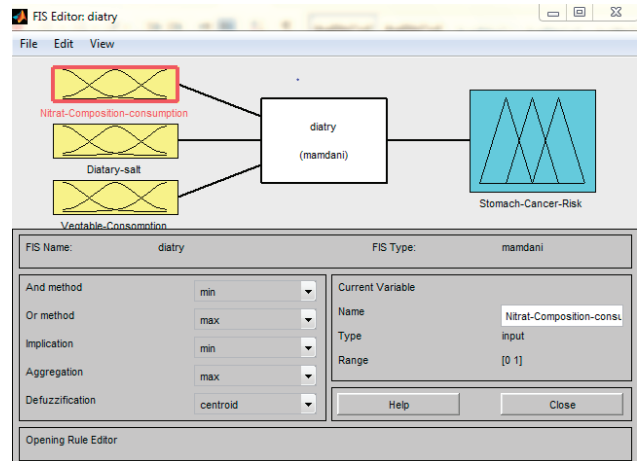

Figure 5. Mamdani fuzzy model of dietary variables in leading to gastric cancer
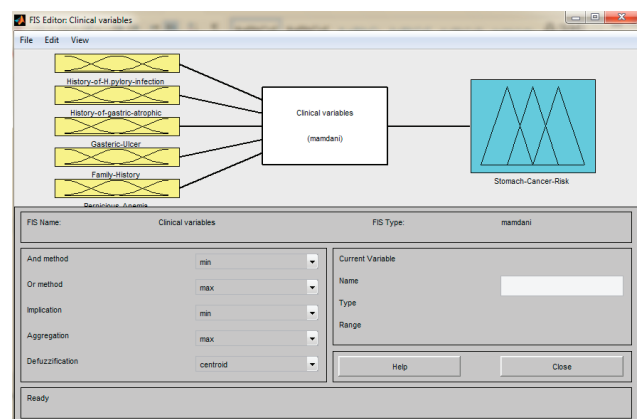

Figure 6. Mamdani fuzzy model of clinical variables in leading to gastric cancer
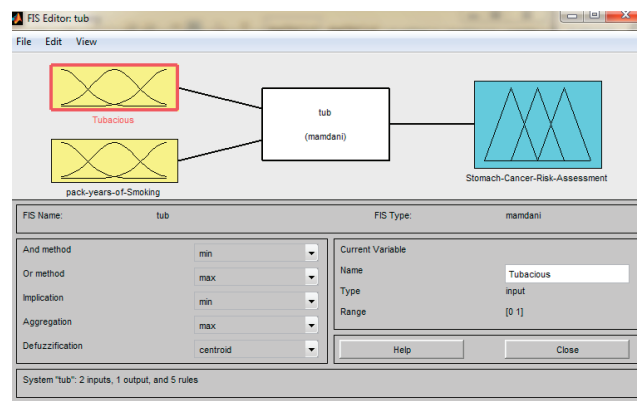

Figure 7. Mamdani fuzzy model of tobacco smoking variables in leading to gastric cancer

the consultation of medical experts. All fuzzy rules where of the form "IF x is A, and y is B then z is C". Where A, B and C are fuzzy sets. Finally, there is a total of 67 weighted fuzzy rules that successfully obtained from the knowledge acquisition are formulated in the rule base. An example of the structure of the rules in rule base: IF sex is male AND age is old AND gastric surgery is yes THEN gastric cancer risk is high.

c) Aggregation of the rule outputs

It is process of the combination of the rules. Aggregation operations on fuzzy sets are operations by which several fuzzy sets are combined in a desirable way to produce a single fuzzy set.

d) Defuzzification process

Defuzzification is the process of converting the output of inference engine into crisp values. Defuzzifier technique employed in this research is Centre of Area (COA) that is a common and useful defuzzification technique.

AND METHOD	MIN
OR METHOD	MAX
IMPLICATION	MIN
AGREGATION	MAX
Defuzzification	CENTROID

Table 2. Specifications of the fuzzy inference method used in this study

4. RESULTS

The proposed system was tested using 50 clinical case reports. In the testing phase, the information of case reports is given to the system to find risk level of each case report then obtained results were compared with expert's diagnosis. In order to evaluate agreement of the system and approved diagnosis by expert physicians, kappa test reported 87% agreement, which showed that proposed system are 87% equal to the medical expert's statement.

In addition, Sensitivity metric was calculated 92.1%, that indicates the correctly of the proposed system to detect individuals are at risk of gastric cancer. Moreover, specificity metric that indicates the correctly of the proposed system to determine individuals who are not at risk is revealed 83.1%.

The results showed that is possible to develop a fuzzy expert system that can identify high risk individuals and other risk levels for gastric cancer. The results also show that the system is user friendly that can provide a heuristic approach in the diagnosing gastric cancer at earlier stages of disease.

The suggested fuzzy expert system has been implemented in MATLAB software using the feature of fuzzy logic toolbox. Designing of user interface, membership functions and knowledge-base has been done using MATLAB. Mamdani inference method was used and finally a total 67 if-then rule was extracted according to expert's opinion as a system rule-base

5. DISCUSSION

Early detection cancer programs, focus on early diagnosis of cancer in populations with clinical symptoms and screening populations without clinical symptoms, but are at risk. The reason why gastric cancer was selected within the study is the frequency of gastric cancer patients in our country. In addition, due to poor prognosis of gastric cancer most of patients refer to hospitals in a time that leading to diagnosis disease in advanced stages. Moreover, Detection cancer greatly depends on the physician's expertise, but these experienced physicians are not available in all parts of the country. The need for the development of expert system according to the expert oncologist reasoning was highlighted to the screening and detects of gastric cancer in healthy peoples.

In consideration of this case, a fuzzy rule-based expert system has been developed. Fuzzy logic from soft computing techniques was utilized. The reason why the fuzzy logic model was used in this study is to effective drawing of a conclusion of systems where a fuzzy decision is used, depending on the uncertain linguistic information as human logic can do. Besides, fuzzy logic is conceptually easy to understand and the mathematical concepts behind fuzzy logic are very simple.

In design of this study, at first step, thorough conducting interviews with medical experts major factors that leading to gastric cancer are determined. In this step asked from medical experts to define the relative weight of each factor to leading gastric cancer. Moreover irrelevant factors or factors that until now there is not sufficient evidence supports them to be as gastric cancer risk factor are discarded based experts opinion.

Mamdani inference technique was used to stimulate reasoning of expert physicians in field of estimate gastric cancer risk. In this method, 13 risk factor used as inputs and single output that revealed the risk status.

In this study, evaluation process conducted by comparing the results obtained from by the expert system with physician diagnosis.

The results also imply that the proposed system was effective in supporting gastric cancer screening and the development of these knowledge-based gastric cancer prediction models is a step forward to a more effective screening of gastric cancer in the country. This system will be useful to the physicians for determine the risk of gastric cancer and refer susceptibility peoples to specialist settings for receiving endoscopic examination and enter to the surveillance program.

6. CONCLUSIONS

Diagnosis gastric cancer at an early stage makes significantly higher survival rate. In this paper, the use of fuzzy expert system was investigated to assess the risk level of gastric cancer. Fuzzy expert system proves to be an effective tool for diagnosis, treatment and preventive measures in medicine. Because gastric cancer is often diagnosed at an advanced stage and because symptoms are confused. The aim of the proposed system was to aid in identifying high risk individuals for gastric cancer and reduce the length of time it takes for a diagnosis is made. Our experimental evaluation reveals that this expert system, yielding good results and its efficacy has been endorsed by expert physicians in the field. We conclude that us in this approach is beneficial for early diagnosis of disease and right treatment at early stages of gastric cancer, which can improve the survival rate.

- Conflict of interest: none declared.

REFERENCES

1. Stacy CM, Matthias E, Christoph R. Gastric adenocarcinoma: epidemiology, pathologyand pathogenesis. Journal of Cancer Therapy. 2008: 6.
2. Malekzadeh R, Derakhshan MH, Malekzadeh Z. Gastric cancer in Iran: epidemiology and risk factors. Archives of Iranian medicine. 2009; 12(6): 576-83.
3. Zendehdel K, Marzban M, Nahvijou A, Jafari N. Six-fold difference in the stomach cancer mortality rate between northern and southern Iran. Archives of Iranian medicine. 2012; 15(12): 741-6.
4. Nemati A, Mahdavi R, Naghizadeh Baghi A. Case-control study of dietary pattern and other risk factors for gastric cancer. Health promotion perspectives. 2012; 2(1): 20-7.
5. Zeraati H, Amiri Z. Estimating postoperative survival of gastric

cancer patients and factors affecting it in Iran: Based on a TNM-7 Staging System. 2016; 2016: 5.

6. Ghadimi MR, Mahmoodi M, Mohammad K, Hosseini MMR. Affecting Factors on Survival of patients with gastric cancer using frailty model: PAYESH Health Monitor journal of the Iranian Institute for Health sciences Resaerch. 2011.

7. Davood Abadi AH, Sharifi H, Erfan N, Dianati M, Abdol Rahim Kashi E. An Epidemiologic and Clinical Survey on Gastric Cancer Patients Refered to Shahid Beheshti Hospital of Kashan (1994-2001). Razi Journal of Medical Sciences. 2003; 10(34): 211-20.

8. Nourinayer B, Akhondi Sh, Asadzadeh H, Ansari Sh, Behrouz N, Zali MR. Evaluation of gastric cancer related indicators in Tehran and 7 other provinces in, 1999 and 2000. Medical Sciences Journal. 2005; 15(1): 15-8.

9. Latha K. et al. Visualization of Risk in Breast Cancesr Using Fuzzy Logic in Matlab Environment. International Journal of Computational Intelligence Techniques. 2013; 4(1): 114.

10. Arslan E, Yildiz S, Albayrak Y, Koklukaya E. Rule based fuzzy logic approach for classification of fibromyalgia syndrome. Australasian physical & engineering sciences in medicine/supported by the Australasian College of Physical Scientists in Medicine and the Australasian Association of Physical Sciences in Medicine. 2016.

11. Y A, A. K. Comparison with Sugeno Model and Measurement of Cancer Risk Analysis By New Fuzzy Logic Approach. African Journal of Biotechnology. 2012; 11(4): 979-91.

12. Samuel OW, Omisore MO, Ojokoh BA. A web based decision support system driven by fuzzy logic for the diagnosis of typhoid fever. Expert Systems with Applications. 2013; 40(10): 4164-71.

13. Kunhimangalam R, Ovallath S, Joseph PK. A Novel Fuzzy Expert System for the Identification of Severity of Carpal Tunnel Syndrome. BioMed Research International. 2013; 2013: 12.