Fuzzy Rule Based Diagnostic System to Detect the Lung Cancer

Umair Ahmed
Department of Mechatronics & Control Engineering, University of Engineering & Technology Lahore, Faisalabad Campus
Faisalabad, Pakistan
umairahmed984@gmail.com

Ghulam Rasool
Department of Mechatronics & Control Engineering, University of Engineering & Technology Lahore, Faisalabad Campus
Faisalabad, Pakistan
Gr84786@gmail.com

Saqib Zafar
Department of Mechatronics & Control Engineering, University of Engineering & Technology Lahore, Faisalabad Campus
Faisalabad, Pakistan
saqibzafar704@gmail.com

Hafiz Farhan Maqbool
Department of Mechatronics & Control Engineering, University of Engineering & Technology Lahore, Faisalabad Campus
Faisalabad, Pakistan
farhan.maqbool@uet.edu.pk

***Abstract*— Lung cancer is a foreseeable cause for the death of human beings as millions of people are suffering from this disease. Therefore, there is not only need for a system which is capable of detecting and diagnosing cancer but also detects it in its earlier stages so that it can be accurately and properly controlled. A variety of lung cancer detection systems has been developed but there is still a need for improved systems capable of producing effective results. This paper presents such a system which not only encompasses improved diagnosing procedure but also capable to prescribe treatment for lung cancer. The system has nine input parameters and two output parameters. The input and output parameters explain the level or stage of cancer and suggest how to diagnose cancer, at that stage. The parameters are defined based on common signs and symptoms that are often observed while medically examining the patient. A proposed method for designing the system is Mamdani type FIS (fuzzy inference system). In addition, the major parameter which can be a key factor is the time duration which is a crucial consideration in the effective detection and control of the lung cancer at the initial stage.**

Keywords—Lung Cancer, Artificial Intelligence, Early Stage Detection, Mamdani Fuzzy Inference System (FIS)

I. Introduction

Lung cancer or tumor also named as carcinoma and pulmonary carcinoma of lungs is the uncontrolled growth of the cell in the lungs tissue. There are 2 main types of lung cancer: Non-small cell lung cancer (NSCLC) and small cell lung cancer [1]. According to the American Cancer Society (ACS), the NSCLC has 3 main types and are as follows: Adenocarcinoma (40 % cases) Squamous cell carcinoma (25-30 % cases) Large cell carcinoma (10-15 % cancer cases) [1].

- Stage 1- (60-80 %) cancer grow in the single lung, and not grown up in lymph nodes or distant organs.
- Stage 2- (30-50 %) tumor grow in nodes but not in distant organs.
- Stage 3- (stage 3a+ stage 3b) (10-15 %) tumor spread in nodes and grow in the middle of the chest in 3a cancer grown up from lymph nodes to the opposite lung and moves the neck.
- Stage 4 is diagnosed when cancer has spread throughout the body [1].

It is most important to cure the lung cancer in the early stages as in later stage it is difficult to controlled and cured. Before treatment of tumor most important thing is to detect it. Many systems have been proposed for detecting the lung cancer, but they have some major drawback as many of them focus on the desirable and fit algorithm instead of precise data which is more trustful and some of them use unrelated and unconnected symptoms. From literature survey, different diagnosing systems of lung cancer are studied and yielded following information: Every symptom has a different effect and according to symptoms membership degree changes and output of the algorithm tells the stages of cancer according to the defined level. There are three levels such as limited, extensive, and stage 4 [2].

The PCA analysis that reduces the symptoms by using the dimensionality reduction, and symptoms reduce from 56-6 and in rule database, genetic algorithm in combination with simulated annealing is used to reduce the rules. The time complexity of this system is a major issue which is very high [3]. The system includes the 19 major infections and diseases. The design of the system also checks malignant cancer, operator of the system gives input in form of the symptoms to the system, which after analyzing the symptoms returns output revealing the problems to the lungs [4]. Output design is based upon the algorithm likely mostly as [2] which provides the output determining the level of cancer categorically in 6 stages : “no cancer”, “potential `possibility”, “low stage”, “medium stage”, “high stage” and “critical stage”[5].

Different approaches are used to detect cancer such as: in breast cancer diagnostic system: the system has 96% accuracy as this mimics the opinion of an expert in the detection of risk in the patient [6]. For lung cancer detection different approaches can be used e.g. FIS using techniques as visual studio and weighted average [2]. However, the main limitation of such a technique used in previously developed systems is that this system uses many symptoms unrelated to lung cancer but mostly related to fever and pneumonia [2]. This system reduce the unrelated attributes and characteristics in the data by using the PCA techniques and genetic algorithm [3]. Another technique involves the use of the probability techniques, such that lung cancer and the sickness of the lung is revealed in ascending order [4]. The output of the system communicates the diagnoses with the

978-1-5386-7939-5/18/$31.00 ©2018 IEEE

presence of cancer along with the level of its stage using a similar algorithm technique as in [2] and [5].

Literature survey of lung cancer detection system reveals that many systems have been developed previously but they are limited in providing efficient and accurate results. Therefore, this paper will focus on the development of such a system which not only makes accurate detection of the lung cancer but also prescribe treatment of the lung cancer according to the cancer stage and patient's condition. The main difference of the proposed system over other systems is the use of time factor as a parameter to reveal the stage of cancer. Another differentiable parameter in our system is the treatment of cancer.

A system will be more effective if it has sufficient number of input symptoms. Less number of symptoms will not provide sufficient data to predict the cancer as well as the stage and treatment. Previous study reveals that it has less number of inputs as just 5 inputs and only one output [7] as compared against our system which has 9 input (symptoms) parameters and 2 output parameters, which could be a better approach to detect the lungs cancer and to predict the treatment of the cancer according to stage of cancer.

This paper also enlightens an accurate fuzzy inference system (FIS) which is practically helpful to recognize cancer by screening out the irrelevant symptoms and dealing with main and key symptoms with perspective of time span. Due to best accuracy and precision of FIS they are widely used in different medical and non-medical devices and systems. Some medical systems are breast cancer detection [6] as well as lung cancer detection systems.

II. METHODOLOGY

The proposed a model system is based on the Mamdani inference system for the detection of cancer in the lungs. 9 input and 2 output parameters were used and for these parameters, 6 parents' rules were defined. The system also proposed the treatment according to the conditions and stage of cancer. Fig. 1 shows the methodology of fuzzy rule-based design of the proposed system.

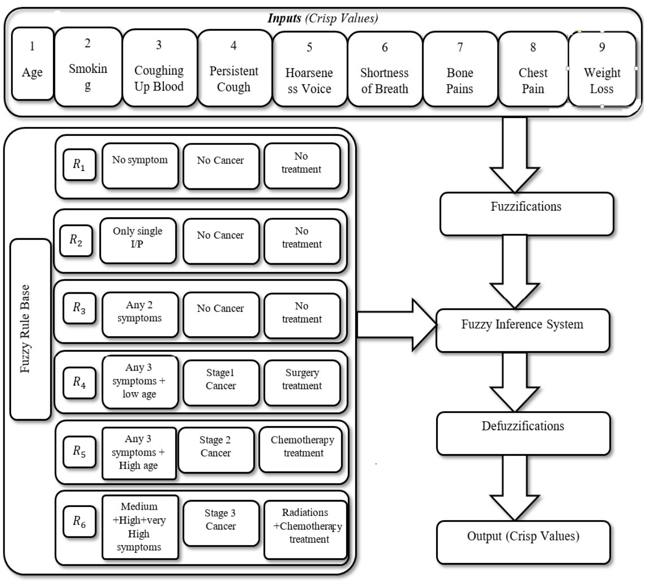

Fig. 1. Block diagram of fuzzy rule base design of the proposed system

A. Inputs

- **Age**

Age is a major factor for lung cancer as when the age exceeds from 50 years, the cancer chances increases. Thhhe range of two membership functions of age is shown in TABLE I and their graphical representation is shown in Fig.2

TABLE I. MEMBERSHIP FUNCTIONS AND RANGES FOR AGE

Logics	Membership function	Range [years]
Input-1		
Age :[0 100]		
Low danger age	Low danger age (mf1)	[0 0 55 60]
High danger age	High danger age (mf2)	[55 60 100 100]

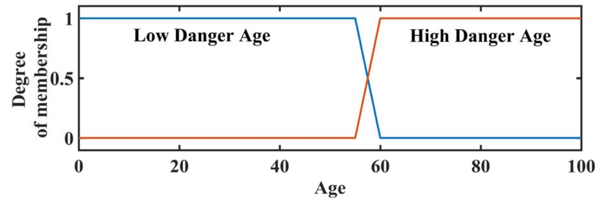

Fig. 2. Membership functions plot for age

- **Smoking**

Smoking is also an important factor or major issue for lung cancer disease such as according to a survey 90% of men and 75-80% females are died due to lung cancer caused by smoking, and 2 out of 3 patients are smokers, smoking as well as asbestos, foundry workers suffer from tumor due to gases such as carcinogens gases [8, 9]. Membership functions of smoking and their ranges are shown in TABLE II and their graphical representation is shown in Fig. 3.

TABLE II. MEMBERSHIP FUNCTIONS AND RANGES FOR SMOKING

Logics	Membership function	Range [No Yes]
Input-2		
Smoking: [0 10]		
No smoking	No (mf1)	[0 0 5 5]
Yes smoking	Yes (mf2)	[5.2 5.2 10 10]

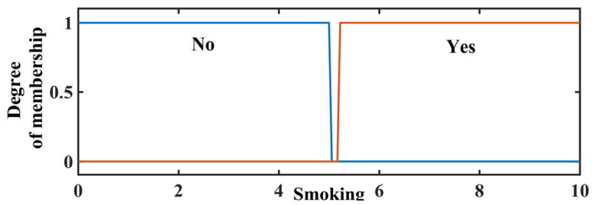

Fig. 3. Membership functions plot for smoking

- **Persistent Cough**

Cough is the result of the smoking and it can be taken as the time span unit such that if a cough is from high time span then can be a major issue toward cancer. TABLE III shows the membership functions and ranges for persistent cough also their graphical representation is shown in Fig. 4.

TABLE III. MEMBERSHIP FUNCTIONS AND RANGES FOR PERSISTENT COUGH

Logics	Membership function	Range (days)
Input-3		
Persistent Cough:[0 80]		
Low	Low (mf1)	[0 0 14 21]
Medium	Medium (mf2)	[14 31 47]
High	High (mf3)	[40 58 60]
Very High	Very High (mf4)	[58 60 80 80]

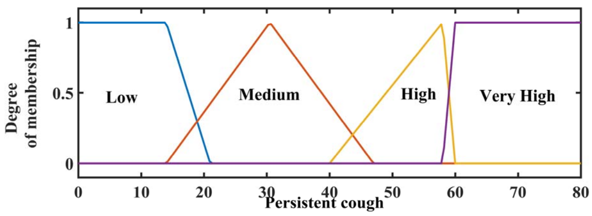

Fig. 4. Membership functions plot for persistent cough

- **Coughing Up Blood**

Blood with a cough after persistent cough is a common symptom of the lung tumor. Coughing up blood known as hemoptysis, defined as the presence of blood in the sputum (spit or phlegm) coughed up from the lungs. 7-10 % of cancer patients suffer from hemoptysis whereas 20 % experience it during lung cancer disease [10]. TABLE IV shows the membership functions and ranges for coughing Up Blood and their graphical representation is shown in Fig. 5.

TABLE IV. MEMBERSHIP FUNCTIONS AND RANGES FOR COUGHING UP BLOOD

Logics	Membership function	Range [days]
Input-4		
Coughing up Blood:[0 80]		
Low	Low (mf1)	[0 0 14 21]
Medium	Medium (mf2)	[14 30.5 47]
High	High (mf3)	[40 58 60]
Very High	Very High (mf4)	[58 60 80 80]

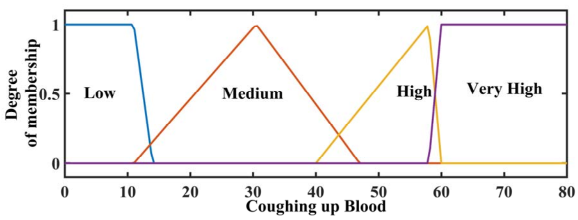

Fig. 5. Membership functions plot for coughing up blood

- **Weight Loss**

From literature, it is found out that 37% patients has huge weight loss and 63% not and most of them were females [11]. If weight loss is more than 10% with time such as 1 month then can be due to lung cancer. TABLE V shows the membership functions and ranges for weight loss and their graphical representation is shown in Fig. 6.

TABLE V. MEMBERSHIP FUNCTIONS AND RANGES FOR WEIGHT LOSS

Logics	Membership function	Range [%age]
Input-5		
Weight Loss:[0 10]		
Low	Low (mf1)	[0 0 3]
Medium	Medium (mf2)	[3 4.5 6]
High	High (mf3)	[5 6.5 8]
Very High	Very High (mf4)	[8 10 10]

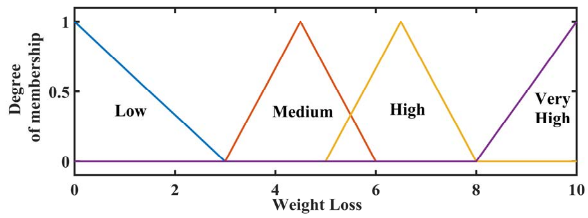

Fig. 6. Membership functions plot for weight loss

- **Shortness of Breath**

It is the most advanced case when a person suffers from lung cancer, it takes painful breathing for almost 15 to 20 days' period. In the proposed system time span is considered important for precise measurement. TABLE VI shows the membership functions and their ranges of shortness of breath and their graphical representation is shown in Fig.7.

TABLE VI. MEMBERSHIP FUNCTIONS AND RANGES FOR SHORTNESS OF BREATH

Logics	Membership function	Range [days]
Input-6		
Shortness of Breath:[0 80]		
Low	Low (mf1)	[0 0 11 14]
Medium	Medium (mf2)	[14 30.5 47]
High	High (mf3)	[30.5 58 60]
Very High	Very High (mf4)	[58 60 80 80]

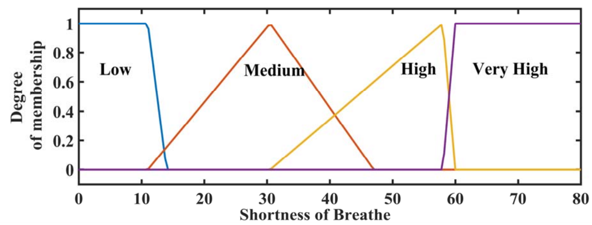

Fig. 7. Membership functions plot for shortness of breath

- **Bone Pains**

The most important parameter is bone pain. In lung cancer when pain is spread all over the body parts the pain is found mostly in backbone, knees and in waist bones . TABLE VII shows the membership functions and ranges of bone pain and their graphical representation is shown in Fig. 8.

TABLE VII. MEMBERSHIP FUNCTIONS AND RANGES FOR BONE PAIN

Logics	Membership function	Range [intensity of pain]
Input-7		
Bone Pain:[0 30]		
Low	Low (mf1)	[0 0 12 14]
Medium	Medium (mf2)	[12 14 30 30]

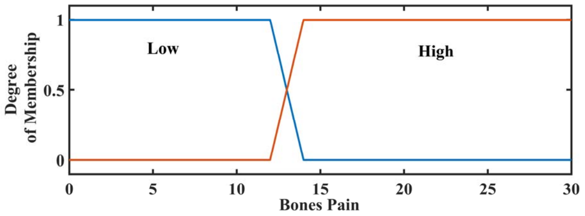

Fig. 8. Membership functions plot for bone pains

- **Hoarseness of Voice**

If a patient suffers from hoarseness of voice then the main factor is that how much pain is there and for how much time the patient has that disease, so the main factor is the time of the disease. TABLE VIII shows the membership functions and ranges of hoariness of voice and their graphical representation is shown in Fig. 9.

TABLE VIII. MEMBERSHIP FUNCTIONS AND RANGES FOR HOARSENESS OF VOICE

Logics	Membership function	Range [days]
Input-8		
Hoarness of Voice:[0 8]		
Low	Low (mf1)	[0 0 1 4]
Medium	Medium (mf2)	[3 4 5 8]
High	High (mf3)	[4.5 8 8]

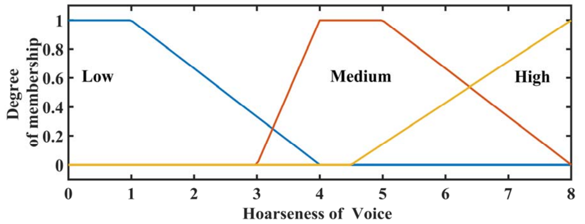

Fig. 9. Membership functions plot for hoarseness of voice

- **Chest Pain**

In case of lungs cancer, chest pain is also a considerable factor which can be under investigation if it increases the day from 10 to 20. TABLE IX shows the membership functions and ranges of chest pain and their graphical representation is shown in Fig. 10.

TABLE IX. MEMBERSHIP FUNCTIONS AND RANGES FOR CHEST PAIN

Logics	Membership function	Range [days]
Input-9		
Chest Pain:[0 80]		
Low	Low (mf1)	[0 0 7 10]
Medium	Medium (mf2)	[7 25 43]
High	High (mf3)	[40 53 55]
Very High	Very High (mf4)	[53 55 80 80]

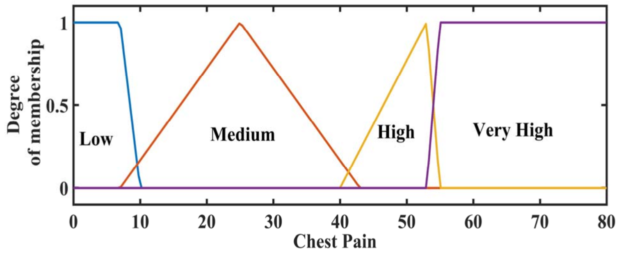

Fig. 10. Membership functions plot for chest pain

B. *Output*

- **Stage**

The proposed system classifies 4 stages of lung cancer disease i.e. "no cancer", "Lung Cancer - Step 1" Lung Cancer - Step 2" and "Lung Cancer -step 4 as shown in Fig. 11.

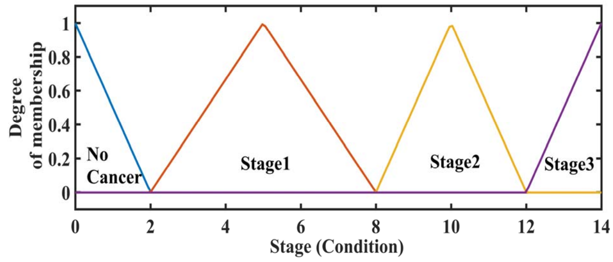

Fig. 11. Membership functions plot for stages

- **Treatment**

The proposed system also gives information about the treatment of the cancer according to condition and stage of the lung caner. TABLE X shows the membership function and ranges for treatment and their graphical representation is shown in Fig. 12.

TABLE X. MEMBERSHIP FUNCTIONS AND RANGES FOR TREATMENT

Logics	Membership function	Range
Treatments:[0 9]		
No Treatment	No Treatment (mf1)	[0 0 2]
Surgery	Surgery(mf2)	[2 3 5]
chemotherapy	chemotherapy(mf3)	[5 6 7]
Radiation	Radiation(mf4)	[7 9 9]

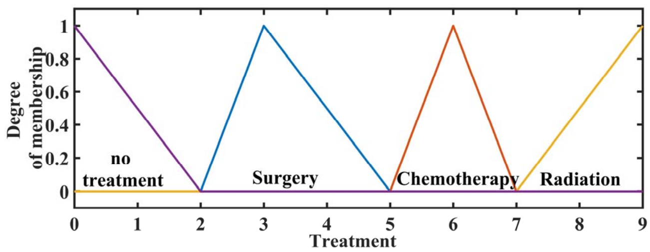

Fig. 12. Membership functions plot for treatment

III. Results and Discussion

This Intelligent Diagnostic system is implemented in MATLAB software using Fuzzy Designer toolbox. To make this system intelligent six parents' rules are defined. Each parent rule consists of sub-rules for detecting the lung cancer and suggesting the best treatment method as shown in Fig. 13. Table XI and Fig. 13 shows a sample result. All parents' rules and their results discussed as follows:

1) Parent Rule-1

Our first rule is simple i.e. if there is no symptom is present then no cancer and no treatment.

2) Parent Rule-2

This rule tells that, if there is any single symptom (input) present at a time (no more than one inputs, such as all the symptoms are at low or zero level and any single at medium) then there is no lung cancer and no treatment is required because such symptoms can be due to any other disease. Implementation of parent rule 2 is shown TABLE XI:

3) Parent Rule-3

According to this rule, if two symptoms at a time is present in a patient then there is also no cancer such that, no cancer and there is no need of treatment because these two symptoms can be due to any other disease.

4) Parent Rule-4

Parent rule 4 explains that if a person suffers from three (3) symptoms (at medium level) and with relaxed for age input, that patient may have lung cancer of initial stage. In these symptoms smoking or persistent cough is necessary. The lung cancer at initial stage can be cured by surgery.

5) Parent Rule-5

In this rule if any four symptoms are present including age (high age factor) and all inputs are at medium level then lung cancer will be at stage-2 and can be cured by chemotherapy or with combination of chemotherapy and surgery. There are 21 sub-rules defined for this parent. A sub-rule 75 is defined for all symptoms present at medium level, for this condition we defined that cancer is at stage-2 and can be cured by using chemotherapy or combined with chemotherapy and surgery.

6) Parent Rule-6

Finally, if there is present all the symptoms in the patient from medium level to high or very high level then lung cancer is at stage 3 and can be cured with radiation or combined with radiation and chemotherapy. There are 15 sub-rules for parent rule 6. These sub-rules are defined from rule# 76 to 95.

The main differences in the proposed and previous systems:

1. Newly introduced parameter of age and smoking. These parameters improve efficiency of the system such as lungs cancer can detect in such first stage.
2. Screening out the unrelated symptoms to avoid misunderstandings so that proposed system is precise to the tumor detection.
3. Factor of time; as time is an important parameter that helps prescription of treatment to the cancer appropriate to stage.
4. This system also tells the treatment of the cancer according to condition and stage of the lung caner.

TABLE XI. IMPLEMENTATION OF PARENT RULE 2

Rule #.	Age	Smoking	Persistent cough	Coughing up Blood	Chest pain	Bone pain	Hoarinessof Voice	Hoariness of breath	Weight loss	Stage	Treatment
1	H	L	L	L	L	L	L	L	L	N	N
2	L	H	L	L	L	L	L	L	L	N	N
3	L	L	H	L	L	L	L	L	L	N	N

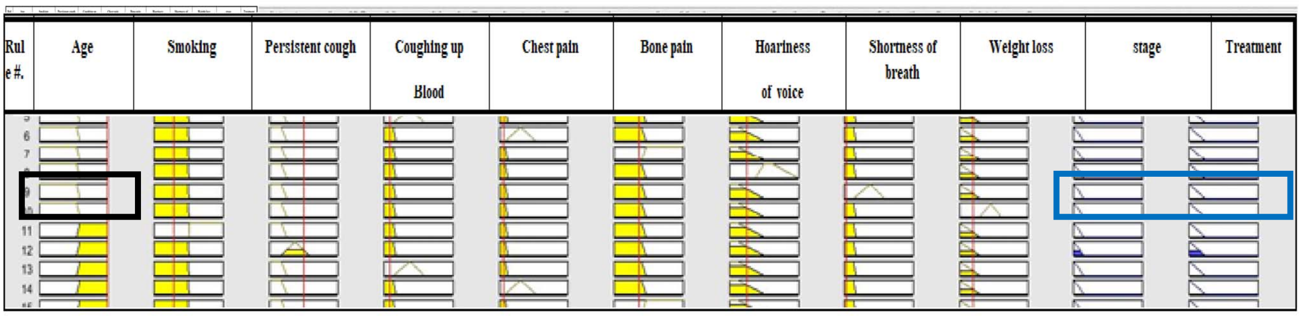

Fig.13. (Single input high: no cancer no treatment)

If a person suffers from three (3) symptoms (at medium level) and with relaxed for age input, that can be may have lung cancer which can be at initial stage. In these symptoms smoking or persistent cough is necessary. Then treatment of lung cancer can be given by surgery.

TABLE XII. MPLEMENTATION OF PARENT RULE 4

Rule #.	Age	Smoking	Persistent cough	Coughing up blood	Chest pain	Bone pain	Hoariness of voice	Shortness of breath	Weight loss	stage	Treatment
1	H	H	H	L	L	L	L	L	L	1	S
2	L	H	H	H	L	L	L	L	L	1	S
3	L	H	H	L	H	L	L	L	L	1	S

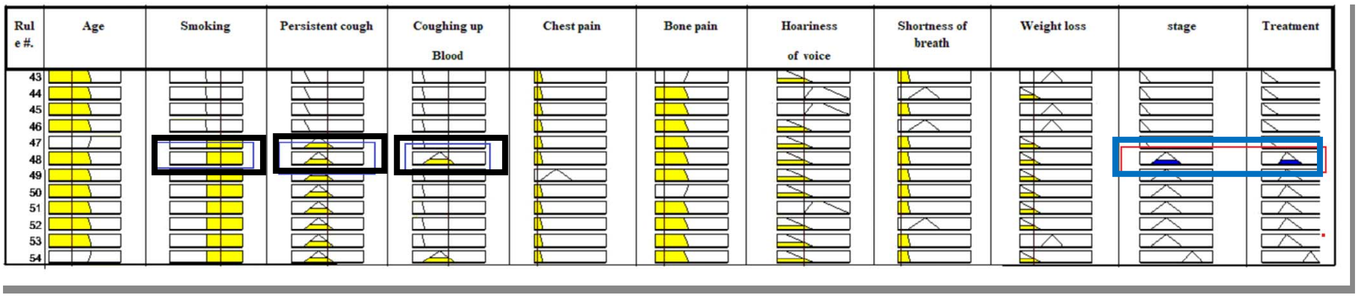

Fig.14. (Three inputs are high: cancer at stage 1 and treatment is surgery)

IV. CONCLUSIONS

Fuzzy logic designer toolbox is an efficient method for controlling a system such as, in the given proposed system, Fuzzy Rule Based Diagnostic System is developed for the detection of the lung cancer. The proposed system shows promising results and able to detect the lung cancer more accurately and more precisely. The system is intelligent which make decision on the bases of input parameters. If patient is found to be diagnosed, then system prescribes treatment such as surgery chemotherapy or radiation according to stage of the lung cancer.

V. LIMITATIONS AND FUTURE WORK

One of the limitations of the proposed system is that by the addition of more parents' rules, the processing time will increase to detect the stage of the lung cancer disease and suggest proper treatment accordingly. Future work includes the implementation of both neural networks and fuzzy rule based system as a fusion in order to increase the accuracy of the system. In addition, data will be obtained from the results of MRI and CT scan of the patient.

REFERENCES

[1] A. Wexler. (2017). *What are the stages of lung cancer?* Available: https://www.medicalnewstoday.com/articles/316198.php

[2] M. S. Durai, N. C. S. Iyengar, and A. Kannan, "Enhanced fuzzy rule based diagnostic model for lung cancer using priority values," *Int J Comput Sci Informa Technol,* vol. 2, pp. 707-710, 2011.

[3] M. Rahi and M. Firoozmand, "A model based on the fuzzy system to detect lung cancer," in *First Conference of intelligent computer systems and their applications, Tehran, Payam Noor University, Tehran*, 2012.

[4] N. H. Phuonga, H. C. Le Huu Nghia, and Y. S. Kwakb, "LDDS–A Fuzzy Rule Based Lung Diseases Diagnostic System Combining Positive and Negative Knowledge," 2003.

[5] M. S. Durai and N. C. S. Iyengar, "Effective analysis and diagnosis of lung cancer using fuzzy rules," *International Journal of Engineering Science and Technology,* vol. 2, pp. 2102-2108, 2010.

[6] R. Khezri, R. Hosseini, and M. Mazinani, "A fuzzy rule-based expert system for the prognosis of the risk of development of the breast cancer," 2014.

[7] K. Lavanya, M. S. Durai, and N. C. S. N. Iyengar, "Fuzzy rule based inference system for detection and diagnosis of lung cancer," *International Journal of Latest Trends in Computing (EISSN: 2045-5364) Volume,* vol. 2, 2011.

[8] S. S. Hecht, "Tobacco smoke carcinogens and lung cancer," *JNCI: Journal of the National Cancer Institute,* vol. 91, pp. 1194-1210, 1999.

[9] S. Markowitz, A. Morabia, and A. Miller, "Reply: asbestos, asbestosis, smoking, and lung cancer: study bias and confounding issues that complicate the interpretation of the results," *American journal of respiratory and critical care medicine,* vol. 189, pp. 116-117, 2014.

[10] Emily Downward. (January 2017.). *Symptoms – Coughing Up Blood.* Available: https://lungcancer.net/symptoms/coughing-up-blood/

[11] E. M. D. Ana Maria Rodriguez, John Friend,. (october 09, 2016). *Experience of weight loss and its burden in patients with non-small cell lung cancer: Results of an online survey.* Available: http://ascopubs.org/doi/abs/10.1200/jco.2016.34.26_suppl.71