

# EXPERT SYSTEM TOOL TO STUDY THE FEASIBILITY OF DIAGNOSING LUNG CANCER DISEASE BY NON-CLINICAL METHOD

**K.Balachandran,**  
Professor, MCA,  
Christ College, Bangalore-560 029,  
E-mail id. Balachander63@gmail.com.

**Dr.R.Anitha,**  
Director, MCA,  
K. S. Rangasamy College of Technology,  
Tiruchengode - 637 215.  
Email id : [aniraniraj@rediffmail.com](mailto:aniraniraj@rediffmail.com)

## Abstract

Cancer, the dreaded disease is one of the dominant cause of sufferings and death in modern world. Cancer is due to the uncontrolled proliferation of the body's cells resulting in an abnormal growth or disruption of the body's auto-regulation. Lung cancer is the leading cause of cancer deaths in both women and men in the United States and throughout the world. Its cure rate and prognosis depends mainly on the early detection and diagnosis of the disease. Manifestation of Lung cancer in the body of the patient reveals through early symptoms in most of the cases. This study is aim to find out the feasibility of developing an expert system for Lung cancer detection by systematic study of the risk factors and the symptoms. Prediction based on this system can be integrated and correlated with other clinical approaches like biopsy, mammogram, X-ray etc.,. From the supervised learning, it can be made to adopt unsupervised learning, once sufficient data is populated.

## Introduction

Expert System, a type of computer application program that makes decisions or solves problems in a particular field, such as finance or medicine, by using knowledge and analytical rules defined by experts in the field. Human experts solve problems by using a combination of factual knowledge and reasoning ability. In an expert system, these two essentials are contained in two separate but related components, a knowledge base and an inference engine. The knowledge base provides specific facts and rules about the subject, and the inference engine provides the reasoning ability that enables the expert system to form conclusions.

Cancerous (malignant) cells can develop from any tissue within any organ. As cancerous cells grow and multiply, they form a mass of cancerous tissue – called a tumor- that invades and destroys normal adjacent tissues. The term 'tumor' refers to an abnormal growth or mass; tumors can be cancerous or non cancerous. The risk of developing cancer increases as we age, so age along with gender, race and personal and family medical history, is risk factors for cancer. Other risk factors are largely related to lifestyle choices, while certain infections, occupational exposures and some environmental factors can also be related to developing cancer. On the basis of strength of evidence in the scientific literature, we distinguish

below between known and possible risk factors by site<sup>[2]</sup>.

## Lung cancer symptoms

Non-clinical symptoms (warning signs) are some of the generic indicators of the Lung cancer diseases, though these warning signs may be common among other diseases also.

Some of the symptoms of Lung cancer are: new cough or change in an existing cough, hemoptysis (flecks of blood in the sputum when coughing), unexplained weight loss, unexplained persistent fatigue, or unexplained deep aches or pains, coughing up a large amount of blood, sudden shortness of breath, sudden weakness, sudden vision problems, or persistent chest pain.

## Lung Cancer Risk factors

Cigarette smoking is the most important cause of lung cancer. Research as far back as the 1950s clearly established this relationship. Cigarette smoke contains more than 4,000 chemicals, many of which have been identified as causing cancer. A person who smokes more than one pack of cigarettes per day has a risk of developing lung cancer 20-25 times greater than someone who has never smoked. Once a person quits smoking, his or her risk for lung cancer gradually decreases. About 15 years after quitting, the risk for lung cancer decreases to the level of someone who never smoked.

Cigar and pipe smoking increases the risk of lung cancer but not as much as smoking cigarettes. About 90% of lung cancers arise due to tobacco use. The risk of developing lung cancer is related to the following factors: the number of cigarettes smoked, the age at which a person started smoking and how long a person has smoked (or had smoked before quitting).

Other causes of lung cancer include the following: Passive smoking, or secondhand smoke, presents another risk for lung cancer. An estimated 3,000 lung cancer deaths occur each year in the U.S. that are attributable to passive smoking.

Air pollution from motor vehicles, factories, and other sources probably increase the risk for lung cancer, and many experts believe that prolonged exposure to polluted air is similar to prolonged exposure to passive smoking in terms of risk for developing lung cancer.

Asbestos exposure increases the risk of lung cancer by nine times. A combination of asbestos

exposure and cigarette smoking raises the risk to as much as 50 times.

Lung diseases, such as tuberculosis (TB) and chronic obstructive pulmonary disease (COPD), also create a risk for lung cancer. A person with COPD has a four to six times greater risk of lung cancer even when the effect of cigarette smoking is excluded.

Radon exposure poses another risk. Radon is a by-product of naturally occurring radium, which is a product of uranium. Radon is present in indoor and outdoor air. The risk for lung cancer increases with significant long-term exposure to radon, although no one knows the exact risk. An estimated 12% of lung cancer deaths are attributable to radon gas, or 15,000 to 22,000 lung cancer-related deaths annually in the U.S. Radon gas is the second leading cause of lung cancer in the U.S. As with asbestos exposure; smoking greatly increases the risk of lung cancer with radon exposure. Certain occupations where exposure to arsenic, chromium, nickel, aromatic hydrocarbons, and ethers occurs may increase the risk of lung cancer. A person who has had lung cancer is more likely to develop a second lung cancer than the average person is to develop a first lung cancer. Screening examinations, tests, or procedures are usually not diagnostic of cancer but instead indicate that a cancer may be present. The Diagnosis is then made following a workup that includes a biopsy and pathological confirmation.

This study is to find the feasibility of developing a system for the diagnosis of the Lung cancer disease based on the symptoms and risk factors. Confirmation of Lung cancer is generally done diagnosed by following techniques.

Microscopic, X-ray, CT Scan, Bronchoscopy, Mediastinoscopy : Thoracotomy

### Methodology

Symptoms and risk factors generally aid in the diagnosing the disease. Sometimes the symptoms and risk factors may mislead the general diagnosis or sometimes the cancer disease even in the advanced stage may not give enough symptoms. Leaving aside these extreme cases this study is trying to probe the diagnosis of the cancer diseases by the symptoms and risk factors associated with different types of cancers.

Sequence of operations are:

1. Obtaining personal information of the patient like name, age, sex
2. Getting symptoms & risk factors of the patient
3. Classifying the possibility of Lung cancer disease based on the symptoms and risk factors from the knowledge base( table 1 & 2)
4. Assigning weightage factors for various symptoms and risk factors (Ref. Table 1)
5. Finding the cumulative score of these influencing factors.

$$\left[ \frac{\sum_i W_i S_i}{\sum_i W_i + \sum_j W_j R_j} / \sum_j W_j \right] * 100$$

where  $W_i$  and  $W_j$  are the weightage factors of  $i^{\text{th}}$  and  $j^{\text{th}}$  symptom and risk factor respectively and where

$S_i$  and  $R_j$  are  $i^{\text{th}}$  and  $j^{\text{th}}$  symptom and risk factor respectively

6. Based on the step 5 result putting the patient for further specific cancer disease type
7. Tabulating the result and generating inference
8. Guiding the patient regarding the future measures
9. Adding the profile of the patient to the knowledge base after validation by the expert

Table 1:

Lung			
Symptoms	Yes / No	Weighted Score	Knowledge Base Field
Do you have Cough with expectoration	1	40	C34SYM01
Do you Experience Chest pain	1	20	C34SYM02
Do you Experience Loss of weight	2	0	C34SYM03
Do you Experience fatigue?	2	0	C34SYM04
Do you Experience Fever	1	5	C34SYM05
Do you Experience Hoarseness of voice	2	0	C34SYM06
Do you Experience Puffiness of face	2	0	C34SYM07
Do you Experience Loss of appetite	2	0	C34SYM08
Do you Experience Nausea and vomiting	2	0	C34SYM09
<b>Risk factors</b>			
Smoking	2	8	C34RIS01
High dose of ionizing radiation	1	3	C34RIS02
Occupational exposure to mustard gas, chloromethyl ether, inorganic arsenic, chromium, nickel, vinyl chloride, radon asbestos	1	3	C34RIS03
Radon exposure	1	1	C34RIS04
Air pollution	1	1	C34RIS05
Insufficient consumption of fruits & vegetables	2	0	C34RIS06
<b>Sum:</b>		81	<u>REPORT</u>
<b>Probable</b>		69.82759	

**Table 2:**  
**Some of the factors are further expanded like smoking**

Risk Factors	User response	Calculated/ assigned weightage factor	Period in years	Average Quantity	Factor
Smoking (chain smoker, occasional, passive smoking, rare, None)	2	8			720
Cigarette			10	5	600
Beedi			5	2	120
Hukka					0
Tobacco chewing					0
Alcohol (frequent, occasional, rare, None)	1	7			720
Country liquor- Toddy			10	5	600
Beer			5	2	120
Hot drinks					0
Weight (obese, overweight, normal, underweight)	1	5			
Insufficient consumption of fruits & vegetables	2	0			

Weightage parameters are added and termed as *ChanceSum* parameter. This chancesum is normalized for a maximum score of 100 (if all the parameters are high the maximum score is taken as 100). Based on the *ChanceSum* score, possibility of the person having cancer disease is categorized into any one of the four possibilities. Viz. 1) Most probable (if chancesum > 80), 2) probable (if chancesum > 60 and <=80), 3) likely but further probing is required (if chancesum > 35 and <=60) and 4) unlikely (if chancesum <=35).

This software is intended for use by the general medical practitioners especially the Primary Health center medical practitioners, the registration personnel in the oncology specialty hospitals and users who don't have access to expert medical advise.

To assign the weightage factor initially all the available records of the Lung cancer patients can be stored in a common database and by linear correlation weightage factors for various factors can be assigned. With the advancement of technology and knowledge about the particular disease's manifestation these factors maybe modified, new factors can be

incorporated and some of the factors may be deleted as and when need arises.

In case of positive result on any type of cancers patient may be to guided for proper medical care from the concerned specialists.

Early detection of the cancer disease is crucial in diagnosing and treating the patient. The cure, metastasis (spreading of cancer disease to completely new location), recurrence (relapse), Remission (absence of all evidence of a cancer after treatment) and survival rate (Percentage of people who survive for a given period of time after treatment all directly attributable to the phase of detection of the cancer disease. Hence it is very essential that common man who has some symptoms and risk factors are better to under go medical examination by a specialist at the earliest.

#### Discussion and conclusion

This study is trying to analyse the various factors that give indications of the disease and give proper suggestion. Timely taking a remedial course of action helps the patient to overcome the disease or at least to minimize the impact of the disease. As specialized diagnostic facilities are not available except in major cities this study aims to help those patients who can identify their symptoms and risk factors and try to get information so that they can go for specialized medical diagnosis at the earliest.

Caution to be taken as some of the cancer diseases even in the advanced stage does not give disease symptoms externally. So purely based on this software result (which may predict unlikely) one cannot neglect medical attention. Equally some of the symptoms and risk factors are common for other types of chronic diseases also. Hence other chronic diseases presence has to be ruled out before analyzing the result from this study.

#### References

1. Microsoft® Encarta® Encyclopedia. © 2001 Microsoft Corporation
2. Mark Beers, Andrew J.Fletcher, Robert Porter, Michael Berkwits & Justin L.Kaplan[2002], *The Merck Manual of Medical Information Second Home Edition*
3. Eugene Braunwald, Anthony S.Fauci, Dennis L.Kasper, Stephen L.Hauser, Dal L. Longo & J.Larry Jameson[2001], *Harrison's Principles of Internal Medicine Volume I (15<sup>th</sup> Edition) (part six: Oncology & Hematology) Page 491-652.*
4. Whyatt and Spiegelhalter, Edward Feigenbaum, Pamela McCorduck, & H. Penny Nii [2001], *Medical Expert system*
5. Indian council of Medical Research's [2005] *National Cancer Registry Programme: Two year report of the Population based Cancer Registries 1999-2000 Incidence and Distribution of cancer*
6. James O. Prochaska Ph.D. Wayne F. Velicer Ph.D., Colleen Redding Ph.D., Joseph S. Rossi Ph.D., Michael Goldstein M.D., Judith DePue Ed.D., Geoffrey W. Greene Ph.D., R.D, L.D.N.d, Susan R. Rossi Ph.D., R.N.,

Xiaowu Sun Ph.D., Joseph L. Fava Ph.D., Robert Laforge Ph.D., William Rakowski Ph.D. and Brett A. Plummer Ph.D. [1998] *Stage-based expert systems to guide a population of primary care patients to quit smoking, eat healthier, prevent skin cancer, and receive regular mammograms*

### Websites

- W-1. National Cancer Institute website:  
[www.cancer.com](http://www.cancer.com) New Cancer statistics review 1975-2005 - U.S cancer statistics report
- W-2. Nikki AI [1991] in web site  
[http://www.wtec.org/loyola/kb/c3\\_s2.htm](http://www.wtec.org/loyola/kb/c3_s2.htm)
- W-3. World Health Organization's report May 1990  
[www.who.int/classifications/icd/en](http://www.who.int/classifications/icd/en)
- W-4. [www.aaai.org](http://www.aaai.org)
- W-5. [www.cee.hw.ac.uk](http://www.cee.hw.ac.uk)
- W-6. [www.healthandenergy.com](http://www.healthandenergy.com)
- W-7. [www.nsf.gov](http://www.nsf.gov)
- W-8. [www.preventcancer.com/alerts/alert3.html](http://www.preventcancer.com/alerts/alert3.html)
- W-9. [www.radonc.washington.edu](http://www.radonc.washington.edu)
- W-10. [www://rex.nci.nih.gov/](http://www://rex.nci.nih.gov/)
- W-11. <http://www.sparc.org.nz/Low-Awareness-that-Physical-Activity-Reduces-Risk-of-Cancer>
- W-12. <http://www.heapro.oxfordjournals.org/cgi/content/abstract/16/4/305>
- W-13. <http://www.cancer.gov>