International Journal of
Science and Engineering Investigations vol. 10, issue 108, January 2021
Received on January 10, 2021 ISSN: 2251-8843

Development of Lung Cancer Awareness Framework in Nigeria: Decision Support System and Knowledge Management (KM)

Waliu Olalekan Apena
Department of Computer Engineering, the Federal University of Technology, Akure, Nigeria
Biomedical Computing and Engineering Technologies, Applied Research Group, Coventry University, Coventry, U.K.
(woapena@futa.edu.ng, apenaw@coventry.ac.uk)

***Abstract*-**This study developed a robust framework to create a repetitive lung cancer awareness among the risk group as a result of environmental devastation and lifestyle. The study considered passive and non-passive smoker as potential risk group. Factors such as exposure to harmful gases from burning of wood, fossil fuel and tobacco smoking in Nigeria pose major risk. Nigeria healthcare sector lack organizational structure to address current challenges facing cancer diagnosis procedure as a result of inadequate clinical data base and height of awareness. The communication gap between biomedical professional has hindered prognosis due to inefficient use of information technology and communication technologies (ICTs). This proven to be a complex state of cancer awareness and healthcare efficacy. The study conducted sample trial (150 respondents); outcome revealed less awareness for both passive and non-passive smokers. The research structure data management using python to support healthcare information management system. The study considered knowledge gap associated with this study to adopt decision support system in the conceptual view of knowledge management (KM) to produces a robust framework for Nigeria cancer biomedical professionals. The designed framework integrates artificial intelligence (AI) features extraction for screening/detection; to be deployed towards diagnosis and prognosis of lung cancer and enhancement of clinical decision making. The framework considered clinical knowledge sharing among the biomedical professional to ease access to secured clinical information and patients' data sharing. The framework can be available to non-governmental NGOs and other stakeholders to create a sustainable action plan in the awareness lung cancer diagnosis and prognosis activities; and promote knowledge gap sharing.

Keywords- *Lung Cancer, Awareness, Framework, Nigeria, Knowledge Management*

I. INTRODUCTION

Due to the abundance of wood as a source of fuel, many Nigerian households especially in the rural part of the country, use wood as fuel for domestic needs, while other households use wood stoves as additional heating sources. Wood is a renewable resource, unlike fossil fuels such as oil, coal and gas, which are non-renewable. Unfortunately, smoke from burning of these sources of fuel can be a major source of air pollution capable of impacting public health and the environment negatively. By-products from wood burning include appreciable number of chemical compounds that are harmful and potentially carcinogenic apart from the fact that they also can cause a number of serious respiratory and cardiovascular health problems. They include fine particulates, nitrogen oxides, sulphur oxides, carbon monoxide, volatile organic compounds, dioxins, and furans. Those at greatest health risk from wood smoke include infants, children, pregnant women, the elderly, and those suffering from allergies, asthma, bronchitis, emphysema, pneumonia other heart or lung [1].

From science, smoke that evolves from the burning of wood, fossil fuel and fireplaces contain a complex mixture of gases and particles. The size of particles has been discovered to have a direct link to their potential for causing health challenges. In [2], small particles below 10 micrometers in diameter pose the greatest problems, because they are found to get deep into the lungs, and some may even get into the bloodstream. Among these particles are "fine particles," which are 2.5 micrometers in diameter and smaller. These fine particles can affect both the lungs and heart. Emissions from burning of wood at homes also contain sulfuric oxides, nitrogen oxides, carbon monoxide and potentially carcinogenic compounds including polycyclic aromatic hydrocarbons, benzene, formaldehyde and dioxins. Some of these pollutants are known to cause cancer but their effects on human health via exposure to wood smoke have not been extensively studied [2].

The EPA conducted a review of the national ambient air quality standards (NAAQS) for carbon monoxide (CO). From the review, some standards were established for some pollutants that are most dangerous to public health and wellbeing. They are released into the air from various or multiple mobile or stationary sources. The NAAQS was meant to be based on air quality criteria, which were to accurately reflect the latest scientific knowledge useful in indicating the kind and extent of identifiable effects on public health or welfare that may be expected from the presence of the pollutant in ambient air [3].

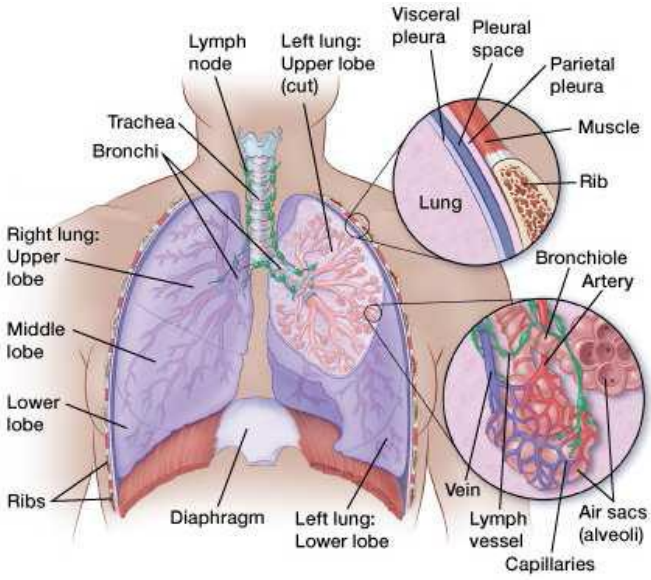

Figure 1. The Human Lung [4]

A. *The Human Anatomy*

Cell is the fundamental unit that make up the entire living organism system. Normally, body cells grow and divide into new cells after which they die naturally according to formation rate. At the early age of a person's life, normal cells divide faster to allow the person to grow. After the person becomes an adult, most cells divide only to replace worn-out or dying cells or to repair injuries. Cancer begins when cells in a part of the body start to grow out of control. There are many kinds of cancer, but they all start because of out-of-control growth of abnormal cells. Cancer cell growth is different from normal cell growth. Instead of dying, cancer cells continue to grow and form new, abnormal cells. Cancer cells can also invade (grow into) other tissues, something that normal cells cannot do. Growing out of control and invading other tissues is what makes a cell a cancer cell. Cells become cancer cells because of damage to DNA. DNA is in every cell and directs all its actions. In a normal cell, when DNA gets damaged the cell either repairs the damage or the cell dies. In cancer cells, the damaged DNA is not repaired, but the cell doesn't die like it should. Instead, this cell goes on making new cells that the body does not need. These new cells will all have the same damaged DNA as the first cell does. People can inherit damaged DNA, but most DNA damage is caused by mistakes that happen while the normal cell is reproducing or by something in our environment.

Sometimes the cause of the DNA damage is something obvious, like cigarette smoking. But often no clear cause is found. In most cases the cancer cells form a tumor. Some cancers, like leukemia, rarely form tumors. Instead, these cancer cells involve the blood and blood-forming organs and circulate through other tissues where they grow [4].

B. *Lung Cancer*

Lung cancer is a cancer that starts in the lungs. To understand lung cancer, it helps to know about the normal structure and function of the lungs. The lungs have two sponge-like organs located in the chest. The right lung is divided into three sections, called *lobes*. The left lung has two lobes. The left lung is smaller because the heart takes up more room on that side of the body.

During breathing, air goes in via the mouth or nose and enters into the lungs via the *trachea* also known as *windpipe*. The *trachea* divides into tubes called the *bronchi*, which enter the lungs and divide into smaller *bronchi*. These divide to form smaller branches called *bronchioles*. At the end of the bronchioles are tiny air sacs known as *alveoli*. Many tiny blood vessels run through the alveoli. They absorb oxygen from the inhaled air into the bloodstream and pass carbon dioxide from the body into the *alveoli*. This is expelled from the body when exhaled. The main function of the lung is to take in oxygen and takeout carbon dioxide. Surrounded by the lungs is a thin lining layer called *pleural*. The *pleura* protects the lungs and helps them slide back and forth against the chest wall as they expand and contract during breathing. Below the lungs is a *diaphragm* which separates the chest from the abdomen. The *diaphragm* moves up and down during breathing, forcing air in and out of the lungs.

C. *Prevalence of Lung Cancer*

Cancer is the leading cause of death worldwide and accounted for 7.6 million deaths in 2008 with primary lung cancer responsible for 1.37 million deaths alone. There are an estimated 226,160 new cases of lung cancer in 2012 in the United States with approximately 160,340 deaths. The lack of sensitive screening tests for early detection of lung cancer plus ineffective treatment for locally advanced and metastatic disease is responsible for the high mortality rate and the dismal overall five-year survival rate. With the extensive effort for tobacco awareness education, advancements of imaging and combined treatment modalities, the 5-year survival rate of lung cancer has improved marginally from 12% in 1977 to 16% in 2007.

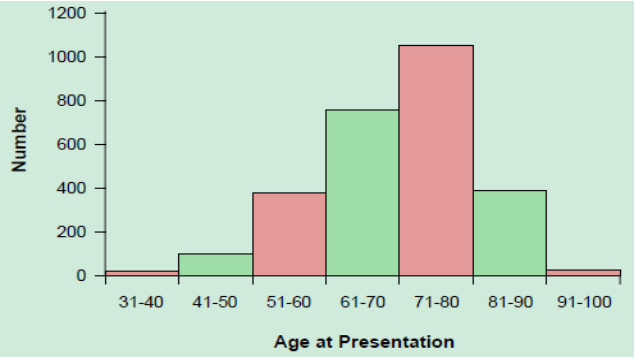

Figure 2. A Graph to Show That the Risk of Lung Cancer Increases with Age (Doncaster Lung Cancer Database, 2010)

However, if lung cancer is detected at an early stage, complete resection of the tumour may produce a 5-year survival approaching 67%. Therefore, early detection of lung cancer by sensitive screening tests could be an important strategy to improve the prognosis of lung cancer. The results of primary lung cancer screening trials using imaging modalities such as chest X-RAYs combined with sputum sampling, CT scans and LDCT will be reviewed. In addition, the role of serum biomarkers and biological modalities, circulating DNA and RNA, exosomal microRNA and circulating tumour cells (CTCs) and thermogram and nanopore sensor technologies will be discussed [5].

D. Population (Age) At Risk

In a high-risk population, there is evidence to suggest there would be about 3 undiscovered lung cancers for every 100 investigations ordered even if patients are asymptomatic and these patients will generally do well. However, screening of asymptomatic subjects has so far not been shown to reduce overall mortality, so the present emphasis should be on identifying early symptomatic disease [6].

The following table shows the percentage of children that are exposed to passive smoking in some selected nations of the world. Nigeria was found to have about 34% of children exposed to passive smoking. This figure would increase on yearly bases if the rate of active smoking increases due to non-awareness and education of its devastating health and social effects.

E. Causes Of Lung Cancer

Research on cancer biology has increased over the few decades, scientists have come to understand that the classification of cancers used in the past – grouping all cancers that arise in one organ, such as the lung, as one kind of cancer with multiple subtypes – is no longer appropriate. There are three relatively common forms of lung cancer. Of these, SCLC arises from neuroendocrine cells; squamous cell lung cancer from the squamous epithelium in the large, central airways of the bronchial system; and adenocarcinoma of the lung from the pneumocytes in the lung periphery. It is recognized that these are three distinct kinds of cancer arising from three distinct cell lineages. These tumour types exhibit mostly different mutational profiles and have different clinical histories and treatment options; and often different scientific investigators. In addition, the amount of research progress has varied for these three cancers. As a consequence, each of the three major types of cancer that originate in the lung present very different problems and require different solutions since there are combined factors that contribute to the effect [7].

F. Symptoms and Syndromes of Lung Cancer

During a diagnosis, it is possible to notice some certain syndrome. A syndrome is a collection of symptoms originating from same cause, usually observed in a group of patients having a common disease to categorize that kind of disease. Symptoms and syndromes are usually indicated to give likely things the doctors might need to know, so as to help give prognosis to a diagnostic test, and to give an idea of the range of symptoms that lung cancers may cause. From the following lists, categories of symptoms differ from one another. It is therefore recommended to discuss any symptoms with the doctor [8].

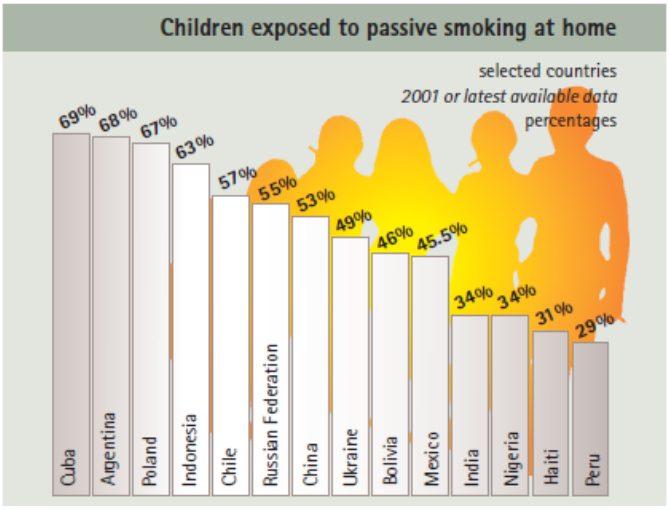

Figure 3. (who.int/tobacco/en/atlas10.pdf)

II. THEORETICAL REFERENCE

A. Literature review

Fine particulate matter, the very small particles that make up smoke and soot, may be the most dangerous component of wood smoke pollution. The most harmful particles are those ten microns or less in diameter (a human hair is approximately 70 microns in diameter). These particles can easily be inhaled deep into the lungs, collecting in the tiny air sacs (called alveoli) where oxygen enters the blood, causing breathing difficulties and sometimes permanent lung damage. Inhalation of fine particulate matter can increase cardiovascular problems, irritate lungs and eyes, trigger headaches and allergic reactions, and worsen respiratory diseases such as asthma, emphysema, and bronchitis, which could result in premature deaths. Pollution from wood stoves is a particular concern in the winter when cold, stagnant air and temperature inversions limit air movement. Communities located in valleys are more strongly affected. As wood burning increases on cold, clear, calm nights, smoke is unable to rise and disperse. Pollutants are trapped and concentrated near the ground, and the small size of the particles allows them to seep into houses through closed doors. In addition to its potential health impacts, wood smoke contributes to the unpleasant brown haze we often experience on winter mornings. Regional haze reduces visibility and obscures our enjoyment of scenic vistas. The air quality on cold, calm nights in low-lying valleys of New Hampshire is periodically very poor. This poor air quality is a serious concern for scientists and state officials; computer models have connected 80 percent of the particulate matter in the air in some areas of the state to wood stoves. Although the source has been identified as wood stoves, in order to meet federal and state regulations for ambient air quality, the local industries will likely have very stringent emission standards imposed on them. These emissions standards will be costly to meet and

could drive local industries out of business. Therefore, to keep local businesses operating, it is important to realize the consequences of wood stove emissions and do something about it [1].

In [10]**,** analyzed 37 published epidemiological studies of the risk of lung cancer (4626 cases) in non-smokers. The review found that the excess risk of lung cancer in life-long non-smokers who lived with a smoker was 24 per cent (95% confidence interval: 13% to 36%). Adjustment for factors such as diet had little overall effect. Tobacco specific carcinogens in the blood of the non-smokers provided clear evidence of the effect of passive smoking. In addition, the study found a dose-response relationship between a non-smoker's risk of lung cancer and the number of cigarettes and years of exposure to the smoker. The authors concluded that: "*The epidemiological and biochemical evidence on exposure to environmental tobacco smoke, with the supporting evidence of tobacco specific carcinogens in the blood and urine of non-smokers exposed to environmental tobacco smoke, provides compelling confirmation that breathing other people's tobacco smoke is a cause of lung cancer*" [10].

TABLE I. SYMPTOMS AND SYNDROMES OF LUNG CANCER [9]

Vein and artery (vascular) syndromes: Certain types of lung tumors can affect the circulatory system as follows:	Thrombophlebitis, an inflammation in an artery or vein caused by a blood clot Arterial thrombosis, a blood clot in an artery, Nonbacterial thrombotic endocarditis (NBTE), a deposition of material onto the valves of the heart
Pulmonary symptoms: The symptoms of lung cancer that are clearly and directly related to lung function are:	Coughing, Bloody sputum, Shortness of breath, Chest pain, Hoarseness, Paralysis of the diaphragm, either symptomless or perceived as shortness of breath, Wheezing or vibrating breathing noises (stridor), Recurrent pneumonia or bronchitis, Difficulty swallowing (dysphagia)
Symptoms in other organs:	Swelling of the face, arms, and neck. Pancoast syndrome
Symptoms of Pancoast	Horner's syndrome, including weak or drooping eyelid, lessened or no perspiration on one side of the face, and a smaller pupil in one eye, Pain in the shoulder, Weakening of hand muscles, Destruction of bone, which might be perceived as bone pain, Headache, Weakness, numbness, or paralysis, Dizziness, Partial loss of vision, Bone or joint pain, Abdominal pain upon probing Unexplained weight loss, Loss of appetite, Unexplained fever, Yellowing of the skin (jaundice), Fluid in the chest or abdomen (effusion, ascites)
Cardiac symptoms and Rare symptoms of metastasis	Lumps in or beneath the skin, Protrusion of the eyes, Eyelid tumors, Perforation of the bowel experienced as severe abdominal pain with fever, Acute pancreatitis experienced as severe abdominal pain and swings in blood sugar levels• A lump in a salivary gland, A lump in one tonsil, A breast lump, Bleeding in the gastrointestinal tract detectable with fecal occult testing
Symptoms linked to tumor metabolism: The most common paraneoplastic syndromes in non-small cell lung cancer patients are:	Hypercalcemia (high blood calcium levels), Excess growth of certain bones, especially in fingertips (hypertrophic osteoarthropathy), Blood clots, Excess breast growth in men (gynecomastia)
Blood (hematologic) syndromes:	Excessive numbers of platelets (thrombocytosis), causing blood clots, Too few or poorly functioning platelets, causing bruises and tiny "blood blisters" (thrombocytopenic purpura) Irregular growth, appearance, or untimely destruction of red blood cells (polycythemia, haemolytic anemia, red-cell aplasia) causing fatigue, Abnormal levels or functions of blood proteins (dysproteinemia) causing fatigue, headache, or dizziness, Abnormal white blood cell production or function (leukemoid reaction, eosinophilia)
Skin (cutaneous) syndromes: Various skin conditions are associated with some cases of lung cancer. These conditions are thought to be caused by aberrant behavior of white blood cells and their antibodies reacting either to the tumor or to substances produced by the tumor:	Chronic inflammation of the skin (dermatomyositis), Itching (pruritis), Chronic red patches (erythema multiforme), Dark patches (hyperpigmentation), Hives (urticaria), Scaly patches (scleroderma), New tissue growth in fingertips (digital clubbing)
Musculoskeletal syndromes: Certain lung tumors can cause the following symptoms by producing hormones or causing unusual tissue growth:	Inflamed muscles, skin, and subcutaneous tissue (polymyositis-dermatomyositis), Abnormal bone growth in certain bones, Softening of the bones (osteomalacia), Various painful or dysfunctional muscular symptoms (myopathy)
Nervous system (neurologic) syndromes: Some or most of the paraneoplastic syndromes that affect the nervous system appear to be caused by an attack of antibodies against healthy tissue. Antibodies are white blood cell proteins that normally sequester and immobilize foreign tissue such as viruses or bacteria.	Cerebellar degeneration, causing problems with balance or unsteady arm and leg movements. Lambert-Eaton myasthenic syndrome (LEMS), Peripheral neuropathy, evidenced by pain, tingling, or numbness in hands or feet, Encephalopathy, an infection or inflammation in the brain, Myelopathy, presenting as bone pain or tenderness or bone marrow dysfunction, Psychosis, mimicking mental illnesses such as schizophrenia or bipolar disorder, Dementia, Pseudo-obstruction of the bowel, including nausea, vomiting, abdominal pain, and changed bowel habits. Rarely, visual changes (retinopathy)
Hormonal (endocrine) syndromes: Certain types of lung cancers are capable of producing hormones that act on organs within the brain (the pituitary and hypothalamus) or upon other organs such as the kidneys, adrenal glands, thyroid, ovaries, or testes. Symptoms include:	Fatigue, Loss of appetite, Muscle cramps or weakness, Nausea, Vomiting, Restlessness and confusion, Seizures, Coma, Respiratory arrest, High blood calcium levels (hypercalcemia).
Cushing's syndrome, an overproduction of adrenal hormones by cancerous tissue. The most common symptoms in lung cancer patients are:	Weight loss, Fluid retention (edema), Muscle weakness (myopathy), High blood pressure (hypertension), Less common Cushing's symptoms in lung cancer patients are stretch marks (striae),

A case-control study was conducted to determine whether household exposure to environmental tobacco smoke is associated with an increased risk for lung cancer in pet dogs. Lung cancer cases and controls with other forms of cancer were obtained from two veterinary teaching hospitals during 1985–1987. Exposures assessed included the number of smokers in the household, the amount smoked, and the proportion of time spent indoors by the pet. A weak relation was found for exposure to a smoker in the home (odds ratio = 1.6, 95% confidence interval 0.7–3.7), after controlling for confounding in stratified analyses. Strong evidence for a further increase in risk associated with more than one smoker in the home was not found, nor was a significant trend observed for increasing number of packs of cigarettes smoked per day or an exposure index based on number of smokers in each household, packs smoked per day, and the proportion of time the dog spent within the home. However, skull shape appeared to exert effect modification; the risk was restricted to breeds with short and medium length noses (odds ratio = 2.4, 95% confidence interval 0.7–7.8). Despite the inconclusive findings of the current study, epidemiologic studies in pet animals may add to our understanding of environmental tobacco smoke effects in human populations [11].

III. Materials And Methods

Information and knowledge management concepts are progressively evolving in our day-to-day activities in technology, process and people. Knowledge management is a science of policy that focus on information harnessing, storage, processing, sharing and management in an efficient and economic manner. In this work, knowledge management concept has been extensively used from planning process up to the final framework designed.

A. Planning the awareness

Increasing public understanding of the risks of Carbon Monoxide (CO), Sulfur dioxide (SO_2), Particulate matter (PM), Nitrogen dioxide (NO_2) Lead, Second-hand smoke (SHS), Radon, Volatile Organic Compounds (VOCs), asbestos poisoning and taking sensible precautions could dramatically reduce this risk [12].

There are signs to note an incomplete combustion occurrences and may result in the production of CO with other harmful carcinogenic by-products:

a. yellow or orange rather than blue flames (except fuel effect fires or flueless appliances which display this colour flame)

b. soot or yellow/brown staining around or on appliances

c. pilot lights that frequently blow out

d. increased condensation inside windows

There are a number of simple steps that gas consumers can take to keep themselves safe.

Carbon monoxide can be produced by any combustion appliance, including those that burn fossil fuels e.g. oil, wood and coal. Major attention should be put on the installation processes of any of these appliances and its management and maintenance should be done by a competent person.

Information obtained from the awareness processes is used to creating databases for the population, where the populace are considered to be at high risk based on the nature of work, workplace and the habiting environment. The populace are either considered as active or passive smokers but are inhabitants of the targeted environment.

The objectives of the awareness planning programme include (Indian Cancer Society, 2004);

1. To create awareness of the disease.
2. To help people recognize the early signs and symptoms of cancer, thus enabling them to seek treatment at an early stage. The program encourages the participants to seek prompt medical attention for symptoms which may include lumps, sores, bleeding, hoarseness, weight loss and persistent indigestion/cough/pain, etc.
3. To educate people about the key risk factors of cancer since more than 30% of cancer cases could be prevented by modifying lifestyle or avoiding the key risk factors.
4. To inform people about the importance of cancer check-ups at an early stage and to motivate them to avail of cancer check-up services at a very nominal cost through the Cancer Detection Centre and Mobile Cancer Detection Centre of ICS.

B. Prevention and Diagnosing Processes

Cancer has been a leading cause of death worldwide. In 2004, it accounted for 7.4 million deaths (about 13% of all deaths). Cancer often creates fear which comes out of ignorance and misconception. If there is a lifestyle adjustment and key risk factors are avoided, more than 30% of cancer cases can be prevented. Many cases of cancer could be reduced if they are treated and detected early enough before getting out of hand [13].

More also, people can reduce the amount of smoke from their wood stoves by the use of low-emission stoves that are certified, learning how to operate them properly, while using good quality firewood. This will improve combustion efficiency, reduce emissions, help protect public health and the environment, and save fuel costs [1].

For patients whose symptoms are similar to that of cancer patients, several diagnosing techniques proposed can be followed religiously. The adopted modernized techniques for diagnosing cancer include; imaging techniques, surgical resection of early stage lung cancer. However, it is discovered chest X-Ray (CXR), sputum cytology, computed tomography (CT), fluorescence endoscopy and low-dose spiral CT (LDCT) has not improved survival except for the report in 2010 by the National Lung Screening Trial (NLST), that showed a 20 percent mortality reduction in high risk participants screened with LDCT compared to those screened with CXRs. Furthermore, serum bio-markers for detection of lung cancer using free circulating DNA and RNA, exosomal microRNA, circulating tumour cells and various lung cancer specific

antigens have been studied extensively and novel screening methods are being developed with encouraging results [5].

C. Integrating the awareness, screening and ICT Systems

The integrated framework is seen in a cycle of Knowledge management which comprises of the Screening exercise, ICT and People.

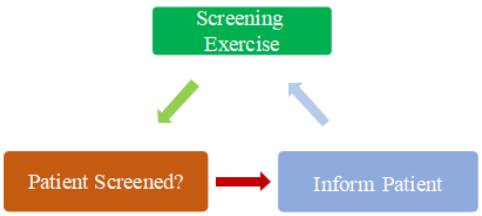

Figure 4. Knowledge Management Cycle of Lung cancer Awareness program (Screening exercise, ICT and People)

D. Creating The First Database

The names of all at risk individual is collated and stored in a centralised data base. The database is managed and maintained by the organization charged with the responsibility of monitoring and control of the awareness program. To ensure proper management of the system, the Health care personnel should be trained in ICT and Database management system The Database constitutes the names of all the screened and unscreened patients that make up of that entire population. Each name is coded in relation to time of registration, alphabetical order of names, kinds of occupation, age brackets, patient's nature of moral habits and locational or residential indices.

The moral habit of the individual populace is taken and is encoded as follow: "**X**", "**Y**" and "**XY**". "**X**" denotes a non-smoking moral habit, "**Y**" denotes an active smoking moral habit and "**XY**" denotes a close relationship between an active and a non-smoker. This is the category of the passive smokers. This work considers that a non-smoker individual may be exposed to smokes from an active smoker if they are both living, working and have to do with each other.

The feature matrix when working with our database set of "p" features variable and "n" observations can be expressed as in equation (1):

$$A = \begin{pmatrix} 1 & a_{11} & \cdots & a_{1p} \\ 1 & a_{21} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_{n1} & \cdots & a_{np} \end{pmatrix} \quad (1)$$

Noting that the *ith* observation for the j^{th} feature are denoted by a_{ij}.

We are letting $a_{i0} = 1$, as a convention.

For a_i, the observation for the i^{th} value may be expressed as shown in equation (2)

$$a_i = \begin{bmatrix} 1 \\ a_{i1} \\ a_{i2} \\ . \\ . \\ a_{ip} \end{bmatrix} \quad (2)$$

For the i^{th} observation, the predicted response is represented by $h(a_i)$

The moral habit parameters, "X", "Y" and "XY" form our target output, y_i, using one-vs-rest multi-class-classification

We have, from linear regression, the prediction of $h(a_i)$ may be expressed as:

$$h(a_i) = \eta_0 + \eta_1 a_{i1} + \eta_2 a_{i2} + \ldots + \eta_p a_{ip} \quad (3)$$

Where $\eta_0, \eta_1, \ldots, \eta_p$ are the regression coefficients. The matrix vector for the regression coefficient can be written as:

$$\eta = \begin{bmatrix} \eta_0 \\ \eta_1 \\ \eta_2 \\ \vdots \\ \eta_p \end{bmatrix} \quad (4)$$

From the foregoing, h(ai) can be expressed as:

$$h(a_i) = \eta^T a_i \quad (5)$$

In an expanded form, equation (5) can be expressed in the form of equation (6)

$$h(a_i) = g(\eta^T a_i) = \frac{1}{1 + e^{-\eta^T ai}} \quad (6)$$

From equation (6), we obtain equation (7) which is the logistic (or the Sigmoid function) applied in our classification. The value of g(z) is always bounded between "0" and "1"

$$g(z) = \frac{1}{1 + e^{-z}} \quad (7)$$

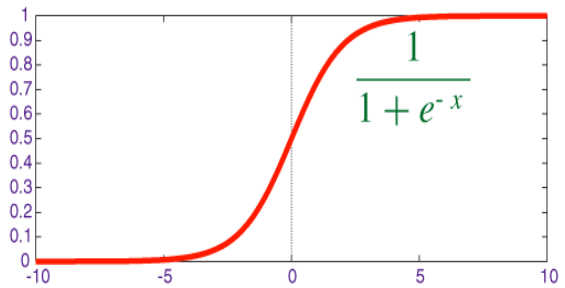

Figure 5. Logistic function (Source: artint.info)

Let our conditional probabilities for two (2) labels (0 and 1) (that is, One-Vs-Rest) for the i^{th} observation be defined as:

$$p(y_i = 1 | a_i; \eta) = h(a_i) \quad (8)$$

or

$$p(y_i = 0 | a_i; \eta) = 1 - h(ai) \quad (9)$$

Combining equations (8) and (9) we have:

$$p(y_i | a_i; \eta) = (h(a_i))^{yi} (1 - h(ai))^{1-yi} \quad (10)$$

From equation (10), the output could be one of the moral habit parameters: "X", "Y" and "XY" using a One-Vs-Rest classification.

Using our training database, our likelihood of parameters are expressed as follows:

$$L(\eta) = \prod_{i=1}^{n} P(y_i | a_i; \eta) \quad (11)$$

Alternatively, equation (11) can be expanded to equation (12)

$$L(\eta) = \prod_{i=1}^{n} (h(a_i))^{y_i} (1 - h(a_i))^{1-y_i} \quad (12)$$

If we take the log of equation (12) it results to equation (13)

$$l(\eta) = \sum_{i=1}^{n} y_i \log(h(a_i)) + (1 - y_i) \log(1 - h(a_i)) \quad (13)$$

Using the log likelihood equation of (13), the cost function can be expressed as in equation (14)

$$J(\eta) = \sum_{i=1}^{n} -y_i \log(h(a_i)) - (1 - y_i) \log(1 - h(a_i)) \quad (14)$$

By the Gradient Descent algorithm, the partial derivatives of *J(ŋ) w.r.t.*each η_j *E* ŋ can be summarized as in equation (15)

$$\frac{\partial J(\eta)}{\partial \eta_j} = (h(a) - y) a_j \quad (15)$$

Where the response vector and the predicted response vector are y and h(a) respectively. Values for the j^{th} feature is represented by the vector set a_j.

An iteration of equation (16) will help to obtain min J(ŋ)

$$\eta_j := \eta_j - \pi \sum_{i=1}^{n} (h(a_i) - y_i) a_{ij} \quad (16)$$

All η_j values are simultaneously updated.

E. Creating the Second Database

During the first awareness campaign, the entire population is to be tested and screened. This is expected because the awareness program focus is to be able to carry out early detection and eradication activities through screening and diagnosing of lung cancer. In the case some of the populace have not been tested and screened, the information is flagged in the first data base created and a reminder message is sent to the category of people still to be tested. This first data base is usually located at the monitoring and control offices while the second data base is located at the screening centers.

F. Information and Knowledge Management Concept

The sharing of information is done using information technology most readily available. In this paper, the use of the internet and mobile network carrier is employed in information storage, processing and dissemination. For privacy concerns, only the intended receptor of information as regards his/her awareness, screening results and further action to be taken is alerted via information technology tools including e-mail, text-messages and phone calls at a scheduled time frame.

The time frame for the subsequent awareness and screening activities for those whose initial screening results were negative to any signs of lung cancer is stipulated for a period of three years. Due to the sustainability of the proposed framework, it does including regular check-up programme via the information processing and sharing with the initial population whose initial screening results were negative to lung cancer. This work has considered the time lapse (stipulated three years – a window period), and it is aware of the likelihood of a patient developing a lung cancer at an early stage due to any of the risk factors considered in the paper within this time frame. At an early stage, the treatment and management of a patient with lung cancer that is characterized by metastatic tendency has been proves to showing great success rate, the purpose for which the proposed framework is designed

G. Designing Lung Cancer Integrated Framework

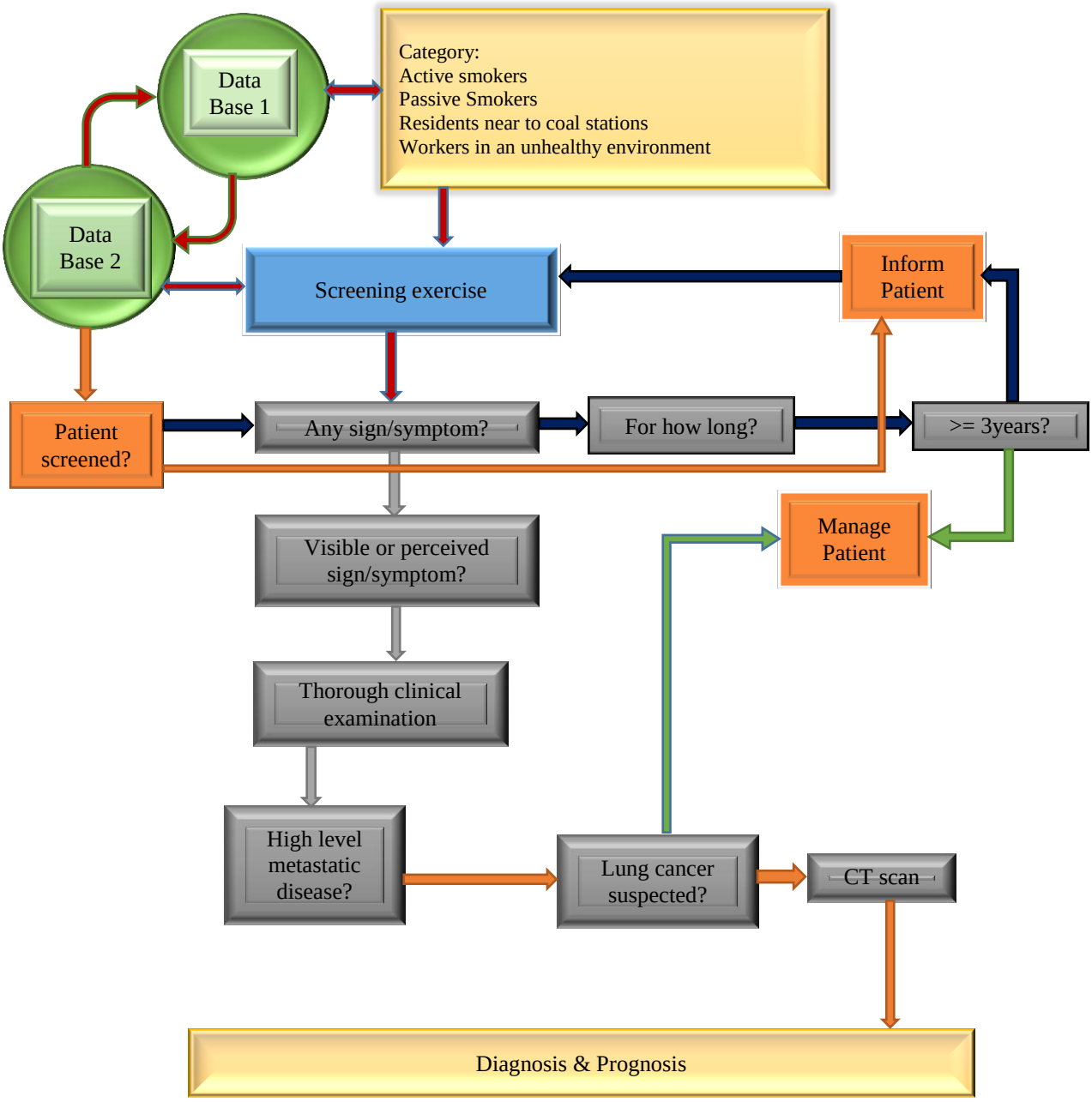

Figure 6. An integrated framework for the Awareness, Screening and Detection and Diagnosis of Lung cancer

IV. Results And Discussions

The implementation of the proposed framework will comprise of smart based decision support system (artificial intelligence –AI) and with programmable initiative such as python. As soon as a patient is tested and screened, his/her information is stored at the screening center database before it is being uploaded into the centralized database. The awareness and screening programme is designed to last throughout the year. This is to ensure that the programme is sustainable. At time interval basis, untested and unscreened people receive awareness text messages, email and/or calls to further intensify the publicity of lung cancer and the need for screening which is early detection, treatment and management of lung cancer.

The algorithm and the framework have shown robustness, accuracy and efficiency. The integrated framework can be used by organisations, NGOs and other stakeholders to create a sustainable action plans for early detection, diagnosis, treatment and management of lung cancer appropriately.

V. CONCLUSION

In conclusion, by creating and designing an integrated, robust and sustainable framework that intensify the publicity and awareness of lung cancer in Nigeria, early detection, treatment and cancer management (prognosis) through screening and diagnosis is ensured. Consequently, increased mortality rate due to lung cancer could be reduced drastically especially among the rural settlers. The framework adapted evolutional information technology (IT) applications and decision support initiative (artificial intelligence – AI) to enhance clinical knowledge sharing among the biomedical professional. This provide easy access to secured clinical information system towards promote patients' data sharing. The framework could be available to non-governmental NGOs and other stakeholders to create a sustainable action plan in lung awareness cancer diagnosis and prognosis activities; and promote clinical knowledge gap sharing.

REFERENCES

[1] State of new harmsphire, 2012. *Wood Stoves and Air Pollution - State of New Hampshire,* New Hampshire: ARD - 36.

[2] Gary, B., 2007. Numerous scientific studies report serious adverse health effects from breathing woodsmoke that comes from residential wood combustion (RWC).

[3] U.S Environmental Protection Agency, 2010. *Quantitative Risk and Exposure Assessment for Carbon Monoxide,* North Carolina: EPA-452/R-10-009.

[4] American cancer society, 2015. *Lung Cancer (Small Cell),* Atlanta, Ga: American cancer society.

[5] Xiang, Dong; Bicheng, Zhang; Donald, Doll; Kui, Shen; Goetz, Kloecker; Carl, Freter, 2013. *Lung cancer screening: from imaging to biomarker,* Louisville: Biomarker Research.

[6] Trevor, R., Rupert, S., Phil, M. and Angela, D., 2010. *Early Detection of Lung Cancer,* Doncaster: Doncaster Primary Care Trust, White Rose House.

[7] Office of the Director, National Cancer Institute, 2010. *Scientific Framework for Small Cell Lung Cancer (SCLC).*

[8] Lawrence Berkley National Laboratory, 2010. *Indoor Volatile Organic Compounds and Health.* [Online] Available at: http://www.iaqscience.lbl.gov/voc-introduction.html [Accessed 16 February 2020].

[9] O'Reilly Media, 2000. *Lung Cancer Resources.* [Online] Available at: www.oreilly.com/onconurse [Accessed 12 March 2020].

[10] Hackshaw, A., Law, M. and Wald, N., 1997. The accumulated evidence on lung cancer and environmental tobacco smoke.

[11] John, R. S., Dunn, K., Gregory, O. K. and Cheryl, H. K., 1992. Passive Smoking and Canine Lung Cancer Risk. *American Journal of Epidemiology, Volume 135, Issue 3,* pp. 234-239.

[12] U.S. Environmental Protection Agency, 2009. *The Inside Story: A Guide to Indoor Air Quality.* [Online] Available at: http://www.epa.gov/iaq/pubs/insidest.htm [Accessed 31 December 2009].

[13] Indian Cancer Society, 2004. *Cancer Awareness and Education Program.* [Online] Available at: http://www.indiancancersociety.org [Accessed 04 May 2020].

[14] https://www.geeksforgeeks.org/understanding-logistic-regression/?ref=rp

How to Cite this Article:

Apena, W. O. (2021). Development of Lung Cancer Awareness Framework in Nigeria: Decision Support System and Knowledge Management (KM). International Journal of Science and Engineering Investigations (IJSEI), 10(108), 1-9. http://www.ijsei.com/papers/ijsei-1010821-01.pdf