

Article

Cancer Incidence in Algeria: Fuzzy Inference System Modeling

Bouharati Khaoula^{1,2}, Bouharati Imene^{3,4}, Boucenna Nassim⁴, lakehal abdelhak², Amarouche Soumaya² Laouamri Slimane^{1,4}, Nadir Boussouf².

¹Laboratory of Health and Environment. Faculty of Medicine, UFAS, Setif1, Algeria

²Faculty of Medicine, Constantine University, Algeria

³Laboratory of intelligent systems, UFAS Ferhat Abbas Setif University, Algeria

⁴Faculty of Medicine, UFAS Ferhat Abbas Setif University, Algeria

*Correspondence: Bouharatik@gmail.com

Abstract: Background: Cancer surveillance data provide information on the incidence and trends of cancer in the population level. Analyzing cancer trends according to these characteristics plays an important role in cancer surveillance. Knowledge of the causes of cancer allow better prevent the appearance of it. A large number of epidemiological evidence supporting the effect of smoking on the causes of cancer there is strong evidence supporting a role for smoking in the etiology of cancers. Alcohol appears to interact with the tobacco significantly and can be considered a risk factor in the development of cancers. Obesity which is now well recognized as a public health problem increases the risk of developing cancers. All these factors are characterized by uncertainty, complexity and imprecision. **Methods:** In this study, we propose an analysis of these factors based on the principles of fuzzy logic inference system. The data were collected from WHO data. As this technique addresses the uncertain, its application in this area is perfectly adequate. **Results:** A database is established, after the analysis system is done, it will be possible to read the prevalence of cancer by introducing randomly the values in inputs variables. **Conclusion:** like cancer has become a national scourge, this application allows predicting the impact of it just from the introduction inputs variables such as BMI, degree of physical activity, tobacco and sex.

Keywords: Epidemiology, Prevention, Cancer, Risk factors, Fuzzy logic

How to cite this paper: Khaoula, B., Imene, B., Nassim, B., Abdelhak, L., Soumaya, A., Slimane, L., & Boussouf, N. (2021). Cancer Incidence in Algeria: Fuzzy Inference System Modeling. *Current Research in Public Health*, 1(1), 1-7. Retrieved from <https://www.scipublications.com/journal/index.php/crph/article/view/45>

Received: June 15, 2021

Accepted: July 23, 2021

Published: July 24, 2021



Copyright: © 2021 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Knowledge of the causes of cancer allow better prevent the appearance of it. Cancer surveillance data provide information on the incidence and trends of cancer in the population level. Analyzing cancer trends according to these characteristics plays an important role in cancer surveillance.

Often, individual cases captured in registry data are missing information on these important variables [1]. Several risk factors contribute to apparition cancer. Certain factors such as genes, diet and environmental exposures may alter susceptibility to cancer in tobacco users. Among these factors, tobacco, alcohol, obesity, physical activity. Several other factors which are also sufficiently studied constitute a risk of cancer. The amount of missing information may vary between subgroups and can change over time [2,3]. Where necessary the estimates provided have been derived from multiple sources, depending on each indicator and on the availability and quality of data. Statistical and health information systems are weak and the underlying empirical data may not be available or may be of poor quality.

Because of the weakness of the underlying empirical data in many countries, a number of the indicators presented here are associated with significant uncertainty [4]. In this case it is very difficult to use classical mathematical modeling. Classical logic does not allow working with uncertainty in the information when knowledge about the behavior of the systems is imprecise. The proposed system utilizes fuzzy logic. Classical logic does

not allow working with uncertainty in the information when knowledge about the behavior of the systems is imprecise. A fuzzy system was constructed with three inputs parameters and one output expressing the number of cases. This technique allows us to establish a fuzzy program so far, is a numeric and symbolic terms of number of cancer type recorded; using the fuzzy inputs data in the universe of discourse (tobacco, obesity and physical inactivity). Once the established system, it allows to predict the impact of each input and its effect on the output parameter. Assess the degree of impact allows us to define the entire factor that has the greatest impact in predicting the onset of what type of cancer and its impact. The result is the contribution of the set of input variable, taking into account inaccuracies and the complexity involved in the process.

2. Cancer risk factors

Tobacco cancer is just one of the ways in which tobacco kills, and interventions to reduce tobacco use will have much broader benefits than just in terms of cancer. The health effects of tobacco—overwhelmingly cardiovascular diseases, cancers, and respiratory diseases—and the benefits of stopping tobacco use are considered together, not separately [5]. Tobacco products contain more than 50 established or identified carcinogens and these may increase risk of cancer by causing mutations that disrupt cell cycle regulation, or through their effect on the immune or endocrine systems. A large number of epidemiological evidence supporting the effect of smoking on the causes of cancer there is strong evidence supporting a role for smoking in the etiology of cancers respiratory tract, including the oral cavity, pharynx, larynx and lung, as well as in the bladder, pancreas, kidney and renal pelvis [6]. Tobacco can also be a cause of esophageal cancer and stomach. Some 55 chemical compounds from 2000 have been identified in tobacco leaves, some of which are released by smoking and were evaluated by the International Agency for Research on Cancer (IARC) as showing «sufficient evidence of carcinogenicity» in the laboratory animals or humans [7]. It may also play a role in the onset of cancer of the oral cavity, larynx and pharynx [8] and a liver cancer [9]. Tobacco-related cancers are a burden. However, these cancers are preventable. Tobacco smoking provides sufficient evidence to establish a causal association between cigarette smoking and cancer of the nasal cavities and paranasal sinuses, nasopharynx, stomach, liver, kidney (renal cell carcinoma) and uterine cervix, and for adenocarcinoma of the esophagus and myeloid leukemia [10].

Obesity which is now well recognized as a public health problem increases the risk of developing cancers. Some systematic review and meta-analyses assessed the strength of associations between body mass index and common cancers such as breast, endometrial, colon and adenocarcinoma of esophagus. However, a regular physical activity and a limited caloric intake are probably safe in healthy subject to prevent cancer and also in cancer survivors [11]. Weight gain and obesity represent approximately 20% of all cancer cases. In particular esophagus cancer, thyroid, colon, kidney, liver, melanoma, multiple myeloma, rectum, gall bladder, leukemia, lymphoma, and prostate in men; and post-menopausal breast cancer and endometrial cancer in women [12]. Many prospective epidemiological studies have shown a direct link between being overweight and cancer, it was estimated that about 20% of all cancers are caused by excess weight, more study of women have shown that nearly half of all cancer cases can be attributed to obesity. The biological mechanisms responsible for stronger association of BMI with some cancer subtypes are not known, but this information might have implications in determining the patient's risk and the subsequent recommendations [13]. The relationship between obesity and the development of cancer is multifactorial and involves a network of metabolic and immunological factors. The differential role of biological mechanisms in various cancer sites should be defined and mechanisms related to the imbalance of adipokines by obesity

should be investigated [14]. In cancer prevention, it is suggested daily exercise, there have been several studies showing that exercise can reduce the risk of breast cancer and that there are benefits to both physical and mental to patients with cancer. It was also correlated with a decreased risk of both the colon and breast cancer development, and probably with less risk of endometrial cancer [15]. Observational studies have assessed the association between physical activity and breast cancer risk. Although most studies reported that high physical activity is associated with decreased risk, some reported the contrary, whereas others reported no relation. The lack of consistency may be attributable to differences in methods for assessing physical activity such as the types of physical activity covered (eg, occupational, recreational, and household). Studies also vary greatly as to the ages at which the physical activity level was assessed [16]. The past two decades of the epidemiological studies have produced a large amount of evidence on the benefits of moderate to vigorous physical activity in relation to risk of cancer at several organs and other chronic diseases. A commonly used the simplest categorization of physical activity, considered in epidemiological studies, identifies the activities at leisure-time (recreational physical activity); household activity, transport activity (traveling to and from work), and occupational activity. Recreational physical activity includes sports, conditioning exercise and other activities [17].

3. Materials and Methods

3.1. Fuzzy logic inference

The fuzzy logic approaches, a sub-field of intelligent systems, are being widely used to solve a wide variety of problems in medical, biological and environmental applications. One of the most important areas of application of fuzzy set theory as developed by Zadeh, 1965 [19] is Fuzzy Rule-Based System. These fuzzy logic systems constitute an extension of the classical rule-based systems, because they deal with “if-then” rules whose antecedents and consequences are composed of fuzzy logic statements, instead of classical logic ones. In a broad sense, a fuzzy rule-based system is a rule-based system where fuzzy logic is used as a tool for representing different forms of knowledge about a problem, as well as for modeling the interactions and relationships that exist between its variables. Due to this property, fuzzy logic principles have been successfully applied to a wide range of problems in different domains for which uncertainty and vagueness emerge in varying ways. Fuzzy modeling, fuzzy control [20] and fuzzy classification are the most common applications. Fuzzy logic deals with reasoning on a higher level, using linguistic information acquired from domain experts. The above-mentioned capabilities make fuzzy logic a very powerful tool to solve many medical problems, where data may be complex or in an insufficient amount. The fuzzy logic concept provides a natural way of dealing with problems where the source of imprecision is an absence of sharply defined criteria rather than the presence of random variables [21]. The fuzzy approach considers cases where linguistic uncertainties play some role in the control mechanism of the phenomena concerned [22]. Fuzzy inference systems (FIS) are powerful tools for the simulation of nonlinear behaviors with the help of fuzzy logic and linguistic fuzzy rules [23, 24]. For example, there is not a straight-line relationship between the socioeconomic, age and breast cancer incidence. In this study, we take to decision algorithms using the engine that makes inferences on a fuzzy rule system. For all the algorithms presented below there is a common rule form for rules that associate a vector.

$a=(a(1),a(2),\dots,a(n))$ with a diagnosis. Further, we assume the following general form of the k th rule in the system.

If $a(1)$ is A_{1k} AND ...AND $a(n)$ is $A_{n,k}$ THAN b is B_k . where $A_{i,k}$, are fuzzy sets (whose membership functions are designated by $(\mu_{A_i,k})$ that correspond to the nature of particular observations (for simplicity we assume the sets to be triangular fuzzy numbers).

3.2. Fuzzy Logic Modeling

A most studies interest exists for evaluating effect of tobacco, obesity and physical inactivity impact of the type of cancer. In our case, we can introduce the relationship between these parameters as input and risks of damage by cancer as output variable. A fuzzy logic system is established. The input variables are sex, smoking, alcohol, physical activity and obesity. The output variables are cancer type and incidence cancer for males and females (Figure 1).

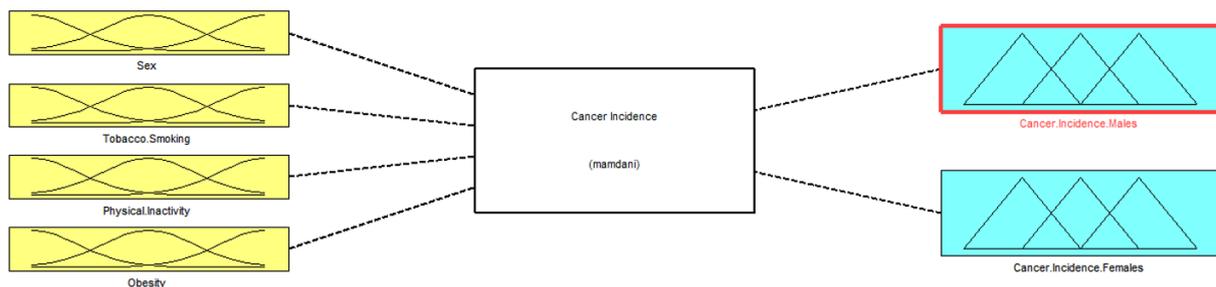


Figure 1. System block diagram with four inputs and two outputs

3.2.1. Fuzzy Rules

The rules determined by the choice of the fuzzy membership function are defined for each input variable. In general form, each fuzzy rule is written as were A_1 and A_2 are the fuzzy sets that describe the nature of the inputs, such as small smoker, Average smoker and Heavy smoking. The linguistic control rules of this system are given by:

If X_1 IS $X_1(1)$ and X_2 IS $X_2(2)$ and... X_n IS $X_n(n)$ Than Y_1 is $Y_1(1)$.

3.2.2. Fuzzyfication of Inputs

The variable 'Sex' is not fuzzy. We attribute value '1' for males and value '2' for females. The data for the other inputs were classified into three linguistic categories:

The variable 'Tobacco smoking' is fuzzyfied on [Small smoker, Average smoker and Heavy smoking] according to the percentage of that category with respect to the total number of population (Figure 2). As it is impossible to define clear boundary between these three categories, we considered them as fuzzy variables and therefore we created fuzzy intervals between the different membership functions to overcome these uncertainties. In the same way, we fuzzyfied other inputs (physical inactivity and obesity).

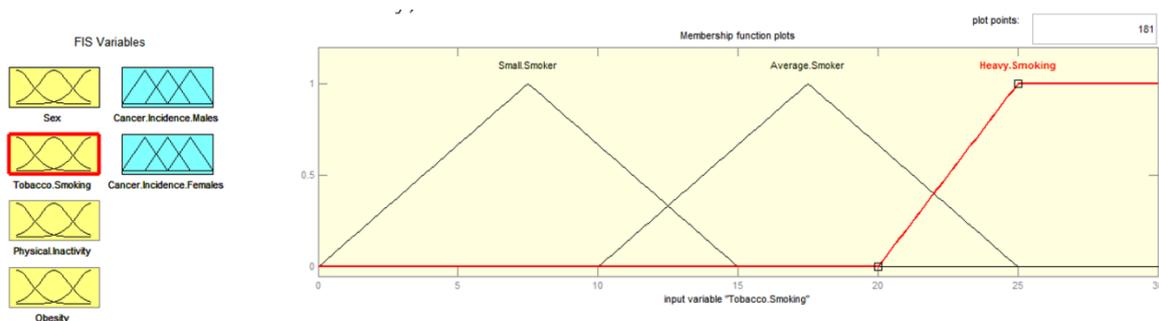
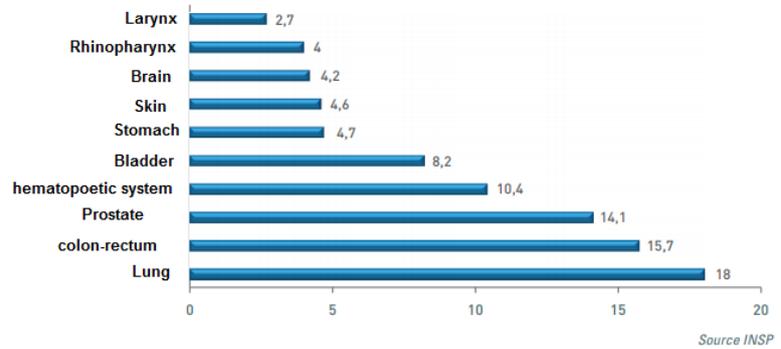


Figure 2. Fuzzyfication of input variable "Tobacco smoking"

3.2.3. Representation of Outputs

The data for the output (the type of cancer incidence) is represented by the type of cancer provided by sound recorded in connection with the rate input variables (Figure 3).



The most common forms of cancer in men and women are collected from [25], and [26].

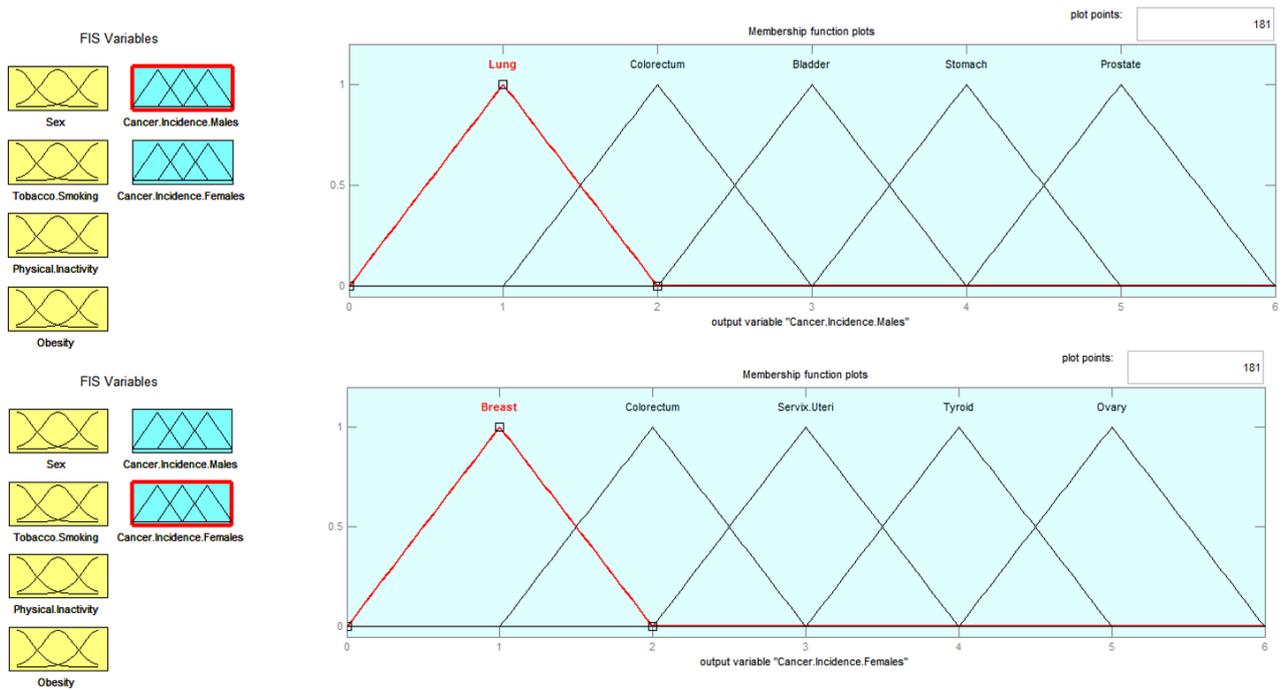


Figure 3. Representation of output variables

3.2.4. Defuzzifier

This system has two outputs that describe for males and females the type and the incidence of cancer. In fact, it explains the contribution of each factor on the rate that induces a direct effect. The crisp value outputs are given by the average method, which has the mathematical expression: $(\sum Si.Ri)/(Ri)$. The defuzzification process follows after estimating its input values.

4. Results

The effect of risk factors on the degree of impact on different aspects is based on a fuzzy logic model. It is designed for the measurement of different parameters. This system consists of four input variables. The rule base of this system is used to determine the output parameter values for each cancer type, according to the input values. MATLAB simulation is used by applying rules. The result of the fuzzy program so far is a numeric and

symbolic terms; using the fuzzy inputs data in the universe of discourse (Sex, tobacco smoking, physical activity and obesity). As the input parameters are characterized by uncertainty, we believe that this tool is very adequate. We emphasize that our fuzzy system is not meant to replace or substitute for an experienced physician; on the contrary, we envisage that the fuzzy logic system should be viewed as a decision support in the most accurate.

Once the established system, it allows to predict the impact of each input and its effect on the output parameter. Assessing the degree of impact allows us to define the set the factor that has the greatest impact in the fight against cancer. The result is the contribution of the set of input variable, taking into account inaccuracies and the complexity involved in the process. Figure 4 shows an example using the MATLAB-rule viewer and simulation result. At each level corresponding to the use of factors (Sex, tobacco smoking, physical activity and obesity); the result can be displayed to the linguistic term output and the number of cases of the corresponding nature of cancer and its incidence rate. It is sufficient to set at random values of the input variables to automatically and instantaneously read the corresponding cases in the system output.

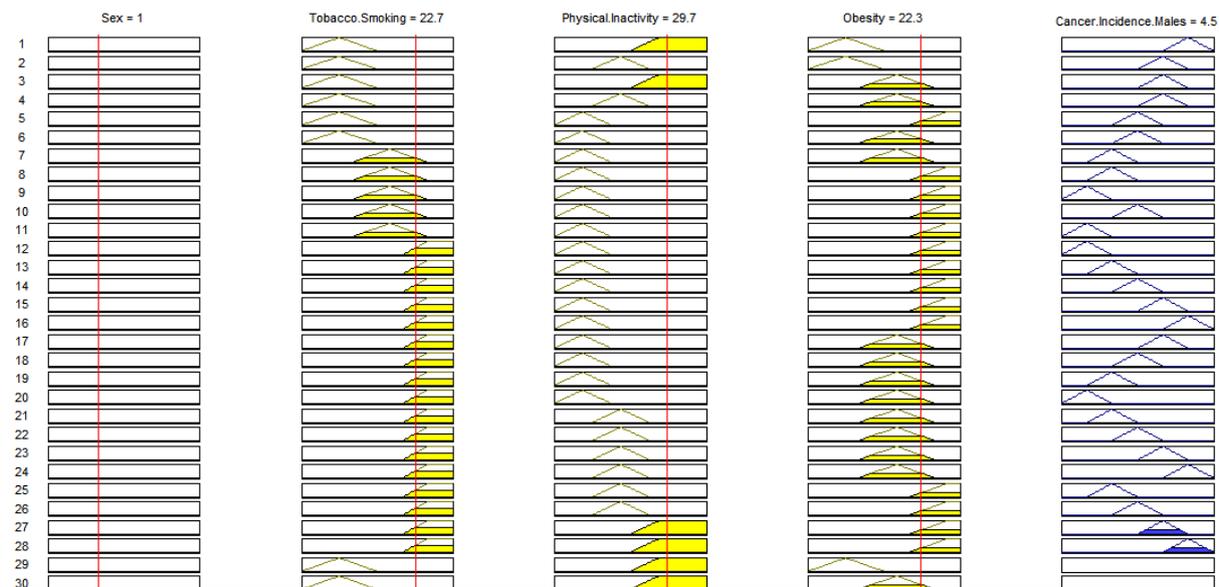


Figure 4. Application Example: Attribution random variable inputs and direct reading of the output variables

5. Conclusion

The artificial intelligent system using fuzzy logic method could extend our understanding of factors affects on the nature of the cancer and its incidence rate. Like other factors are involved in the process and that are not considered in this study, in addition to imprecision already considered in the data analysis, the system is expandable to other factors. The goal of this study is to design and perform a pilot investigation which will provide preliminary data. Modern methods of computational intelligence such as fuzzy logic are used to achieve the medical complex analysis.

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional

Informed consent was obtained from all individual participants included in the study.

Key points

The intelligent software created in this study could be used for giving an idea about the cancer nature expected with its incidence rate.

References

- [1] Howlader N., Noone A.M., Yu M., and Kathleen A. Cronin. Use of Imputed Population-based Cancer Registry Data as a Method of Accounting for Missing Information: Application to Estrogen Receptor Status for Breast Cancer. *American Journal of Epidemiology*. 2012, Vol. 176, No. 4
- [2] Kogevinas M, Pearce N, Susser M, Boffetta P. Social Inequalities in Cancer. Lyon, France: International Agency for Research on Cancer. *IARC scientific publication* 1997 eds, 138.
- [3] Rose DP, Royak-Schaler R. Tumor biology and prognosis in black breast cancer patients: a review. *Cancer Det. Prev.*2001. 25:16–31.
- [4] Krieger N. Exposure, susceptibility, and breast cancer risk: a hypothesis regarding exogenous carcinogens, breast tissue development, and social gradients, including black/white differences, in breast cancer incidence. *Breast Cancer Res T.* 1989, 13:205–223.
- [5] World Health Organization. WHO Library Cataloguing-in-Publication Data. *World health statistics 2015. NLM classification: WA 900.1.* 2015
- [6] Cancer control opportunities in low- and middle-income countries/Committee on Cancer Control in Low- and Middle-Income Countries, Board on Global Health; Frank A. Sloan and Hellen Gelband, editors. by the National Academy of Sciences. 2007
- [7] Kuper H., Adami H-O. & Boffetta P. Tobacco use, cancer causation and public health impact. *Journal of Internal Medicine*; 2002, 251: 455–466
- [8] International Agency for Research on Cancer. Tobacco: a Major International Health Hazard. Lyon: IARC,
- [9] Baron JA, Rohan TE. Tobacco. In: Schottenfeld D, Fraumeni JF. (1996). Jr., eds. *Cancer Epidemiology and Prevention*, 2nd edn. New York: *Oxford University Press*, 1986, 269–89.
- [10] Kuper HE, Tzonou A, Kaklamani E et al. Tobacco smoking, alcohol consumption and their interaction in the causation of hepatocellular carcinoma. *Int J Cancer*. 2000 Feb 15;85(4):498-502.
- [11] Sasco A.J., M.B. Secretan, K. Straif. Tobacco smoking and cancer: a brief review of recent epidemiological evidence. *Lung Cancer*. Volume 45, Supplement 2, August, 2004. Pages S3S9
- [12] Schlienger JL, Luca F, Vinzio S, Pradignac A. Obesity and cancer. *Rev Med Interne*. Sep2009; 30(9):776-82.
- [13] Kathleen Y.Wolin, kenneth carson, Graham A. Colditz. Obesity and Cancer. *The Oncologist*; 15:556–565 causation of hepatocellular carcinoma. *Int J Cancer*; 201085: 498–502.
- [14] Giovanni De Pergola and Franco Silvestris. Obesity as a Major Risk Factor for Cancer. *Journal of Obesity*. Volume 2013 Article ID 291546, 2013, 11 pages
- [15] Michela Ceschi, Felix Gutzwiller, Holger Moch, Monika Eichbolze, Nicole M. Probst-Hensch. *Epidemiology and pathophysiology of obesity as a cause of cancer*. *Swiss MED Wikly*, 2007, 137:50-56
- [16] McTiernan A. "Mechanisms linking physical activity with cancer," *Nature Reviews Cancer*, 2008, vol. 8, no. 3, pp. 205–211,
- [17] Evelyn M. Monninkhof, Sjoerd G. Elias, Femke A. Vlems, Ingeborg van der Tweel, A. Jantine Schuit, Dorien W. Voskuil, and Flora E. van Leeuwen. (). Physical Activity and Breast Cancer. *A Systematic Review*. *Epidemiology* 2007. Volume 18, Number 1, January
- [18] Joanna Kruk, Urszula Czerniak. Physical Activity and its Relation to Cancer Risk: Updating the Evidence. *Asian Pacific Journal of Cancer Prevention*. 2013. 14 (7), 3993-4003
- [19] Zadeh L. Fuzzy sets. *Information and Control*. 1965. 8:338–353.
- [20] Pedrycz, W. Fuzzy modelling: paradigms and practice. *Kluwer Academic Press*.1996
- [21] Driankov D., Hellendoorn H., Reinfrank M. The Mathematics of Fuzzy Control. In: An Introduction to Fuzzy Control. *Springer*, Berlin, Heidelberg. 1993. https://doi.org/10.1007/978-3-662-11131-4_2
- [22] Chi, Z.; Yan, H.; Pham, T. Fuzzy algorithms: with applications to image processing and pattern recognition. *World Scientific*.1996<https://doi.org/10.1142/3132> | October 1996. Pages: 240
- [23] Bouharati S., Benmahammed K., Harzallah D. & El-Assaf Y.M. Application of artificial neuro-fuzzy logic inference system for predicting the microbiological pollution in fresh water. *J. of Applied Sciences*. 2008. 8(2):309-315.
- [24] Demir F, Korkmaz KA. Prediction of lower and upper bounds of elastic modulus of high strength concrete. *Constr Build Mater*. 2008. 22(7):1385–93.
- [25] Allag F., Zegadi R., Bouharati S., Tedjar L., Bouharati I. Dynamic of air pollution and its effect on newborns: Analysis using fuzzy logic inference system. *Wulfenia Journal*, 2013. part no. 2. pp. 18-25.
- [26] World Health Organization – *Cancer Country Profiles*, 2014
- [27] 25. https://extranet.who.int/ncdccs/Data/DZA_B5_plan_national_cancer.pdf.