

# A Preliminary Study to Review the Impact of Lifestyle on the Occurrence of Cancer

Chetna\*, Dr. K. R. Ramkumar\*\* & Dr. Shaily Jain\*\*\*

\*C G C, Landran, Punjab, chetna279@gmail.com

\*\*Chitkara University, Rajpura, Punjab, k.ramkumar@chitkara.edu.in

\*\*\*Chitkara University, Baddi, Himachal Pradesh, shaily.jain@chitkarauniversity.edu.in

Received Sept 16, 2017

Accepted Oct. 15, 2017

## ABSTRACT

*Purpose* – The objective of this paper is to review the impact of lifestyle which include habits and exposure of the ethnic population to any chemical compounds which leads to cancer. *Design/methodology/approach* – The review has been done to study the impact of lifestyle on the occurrence of cancer. Urbanization, increased life expectation and embracing western lifestyles activate the occurrence of cancer disease. Sometimes geographically conditions also propel the occurrence. So many researchers tried to recognize these risk factors so that it can be avoided to decrease the cancer mortality rate to minimum extent. *Results* – The results obtained show the impact of smoking, diet and exposure to chemical substance on the occurrence of cancer. By appropriately detecting the risk factors using data mining tools, these risks can be avoided or at least minimized to avoid further expansion of this malign disease. *Conclusion* – Different lifestyle, habits and exposure to certain chemical compounds were seen to be important risk factors for causing cancer.

**Keywords:** Data mining, lifestyle, cancer, tandoor, cooking bread, taluk.

## INTRODUCTION

As Countries tend to become more industrialized and people become more aged, the frequencies of occurring diseases generally increase. The disease which occurs because of this reason is called lifestyle diseases or non communicable diseases. These diseases include heart disease, hypertension, cancer, obesity, type II diabetes, mental disorders and many other diseases (Meng *et al.*, 2013; Amin *et al.*, 2013). In accordance with the World Health Organization(WHO) also, the increased wave of spreading non communicable diseases in developing as well as developed countries is because of dietary and lifestyle changes(WHO,2003).

Cancer is a potentially critical ailment which is the resultant of abnormal cell functioning which causes extensive masses of anomalous cells whose escalation raze normal tissues and can extend to crucial organs followed by even patient's death. If it is detected at an earlier stage, this disease is curable.

Although Cancer was earlier treated clinically and biologically but during recent years analytical studies have become more common. There is good progress in diagnosing and treatment of cancer disease, still it poses a considerable menace to our society. This is the most frequent reasons for deaths after heart diseases in the whole world (Jemal *et al.*,2007). The expected world's population will be 7.5 billion by 2020 in which prediction is that, 15.0 million will be diagnosed to be cancer cases (Brayand *et al.*,2006) . In India these number will be approximately about 2.5 million in which 8, 00,000 will be fresh cases and 5,50,000 will be the number of death cases per annum . With the survey of historical data it has been found that the figure of cancer cases has drastically expanded over the time and is continue to do so with 96.4% for males and 88.2% for females out of 100,000 cases will be analyzed by the end of the last century (Rao *et al.*,1998; Nandakumar *et al.*,2001;Dinshaw *et al.*, 1999).

Investigation was done on 70% of cancer cases during the last decade and was successful with existence of only some patients. In the Indian sub-continent also, the degree of cancer patient is increasing due to impartial living standards and lack of proper medical care (Ali *et al.*,2001). The risk of developing cancer among people is due to variation in the ecological factors (Kotnis *et al.*,2005) such as smoking cigarette (Wynder *et al.*,1974)or following some particular dietary patterns. For example people in Australia highly suffer from skin cancer because of their high vulnerability to the sun and due to the ease of availability of tobacco and majority of the persons are having the habit of chewing it, people living in India develops oral cancer 25% times the average rate. In fact the India has the maximum number of cases nearly about 75,000 to 80,000 new cases added every year.

When the need to extract the amount of valuable knowledge from the ocean of data arises then data mining (Moro *et al.*,2016; Chen *et al.*,2016) come into being . Data Mining can be defined as analyzing data from different perspectives and realistic predictions be made by summarizing it (Goebel *et al.*,1999; Wu *et*

al.2002 ). The objective of this study was to review the work done by different researchers in the field of cancer according to the lifestyle of the patients.

**CANCER TYPES**

Cancer is classified on the basis of the type of cells in which they build up. Cancer which appears from epithelial cells is called carcinomas that form the exterior layer of some internal organs of the skin. Carcinomas cancer is most prevalent type (Figure 1). Other type of cancer is Leukemias that affects the blood and blood forming organs such as the lymphatic system, bone marrow and the spleen. This cancer affects the resistant power of our body. Sarcoma is also one of the kind of cancer which grow in connective tissues.

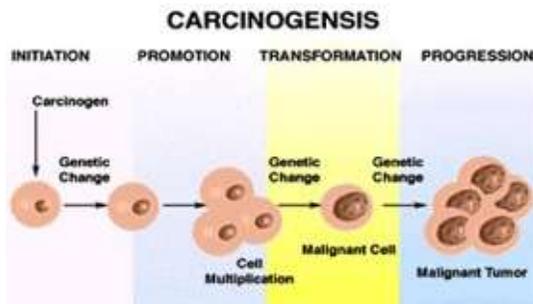


Figure 1: site: www.medscape.com source: cancer control @2007

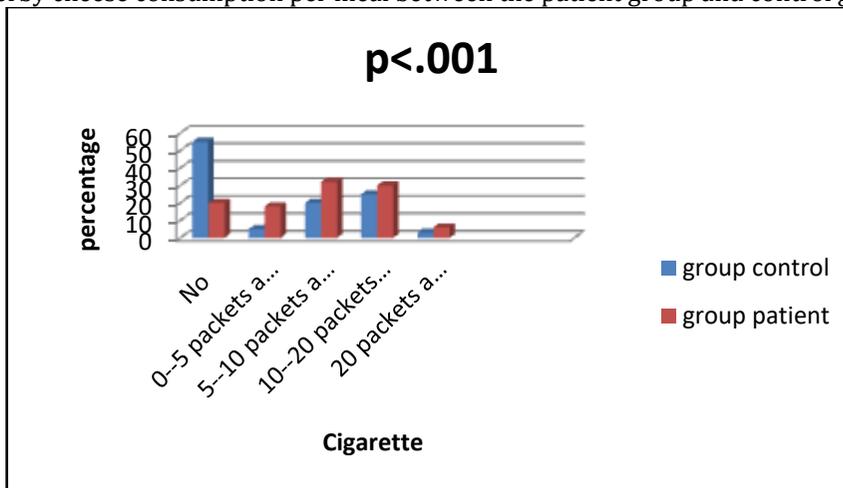
**LITERATURE REVIEW**

Sebahattin Celik *et al.* (2014) in their paper “The Relationship between Eating and Lifestyle Habits and Cancer in Van Lake Region: Another Endemic Region for Esophageal and Gastric Cancers” intended to estimate the association between esophageal and gastric cancer which is generally seen in Van Lake region and their traditional consumption habits. They also created control group having same gender, age and socio-economic features equivalent to patient group. Their study came to conclusion that Cheese consumption and smoking were higher in patient group. So heavy smoking, cooking bread on tandoor and high intake of herby cheese were found to be significant risk factors for gastric and esophageal cancers.

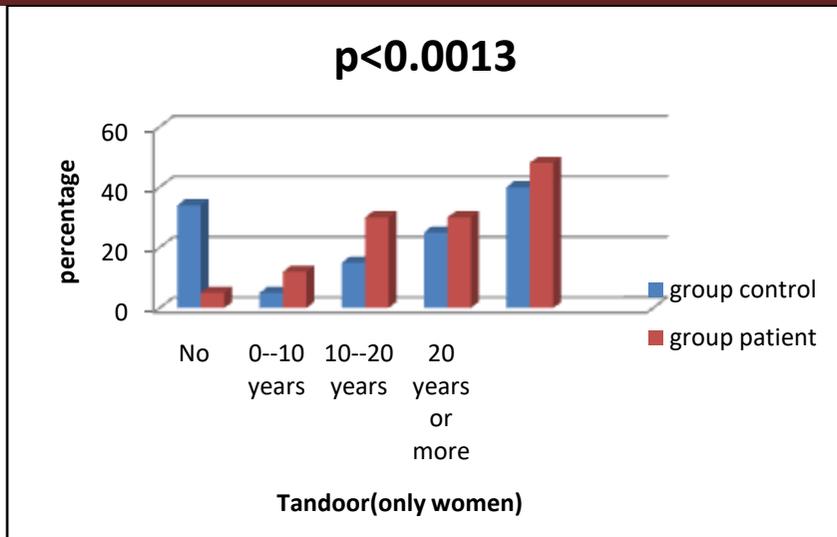
**Table I- Comparison between patient and control groups in terms of herby cheese and tea consumption**

	Patient group Mean(%95-CI)*	Control group Mean(95%-CI)*	P
Amount of Herby Cheese(gram/day)	238.05(217.85-258.26)	85.00(68.71-101.29)	<0.001
Tea(cc/day)	929.38(841.49-1017.28)	1045.70(966.04-1125.36)	0.055

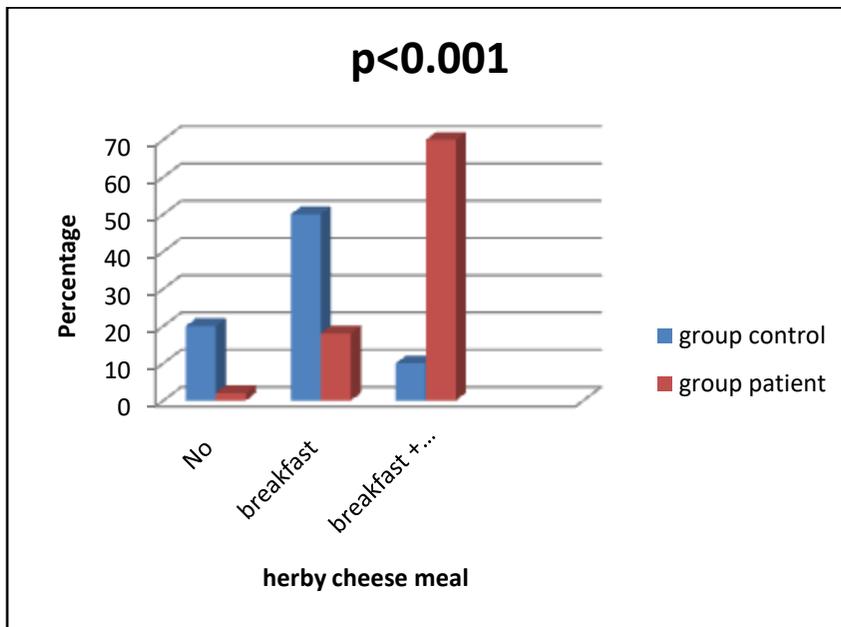
Table I above states that the patient group usually has higher intake of herby cheese whereas there is no significant difference exists as far as Tea is concerned. Figure 2 concludes the effect of smoking, tandoor exposure and herby cheese consumption per meal between the patient group and control group.



(a)



(b)



(c)

Figure 2: Comparison between the patient and control groups in terms of smoking (a), tandoor exposure (b), and herby cheese consumption per meal (c).

They investigated the features of lifestyle and eating habits of Van Lake region and reach to a conclusion that there exist a negative relationship between their habits and esophageal and gastric cancers and taking collective measures towards these diseases can only lower the frequency and mortality rates of this deadly disease.

Arnīs Kirshners *et al.* (2015) in their paper “Gastric Cancer Risk Analysis in Unhealthy Habits Data with Classification Algorithms” initially took 136 attributes which after preprocessing reduced to finally 20 attributes and 742 records. Their medical history and high intake of promoted by other habits like: smoking, consuming alcohol, salt. Figure 3 shows analysis of the obtained results. experimental evaluation determined that C4.5 algorithm as the most accurate algorithm which proves that the infection risk which leads to *H. pylori* development is working environment and physical activity. The ethnic group which shows positive results worked in an environment where they come into direct contact with chemical compounds as well as heavy workload. Another factor is lack of physical activity.

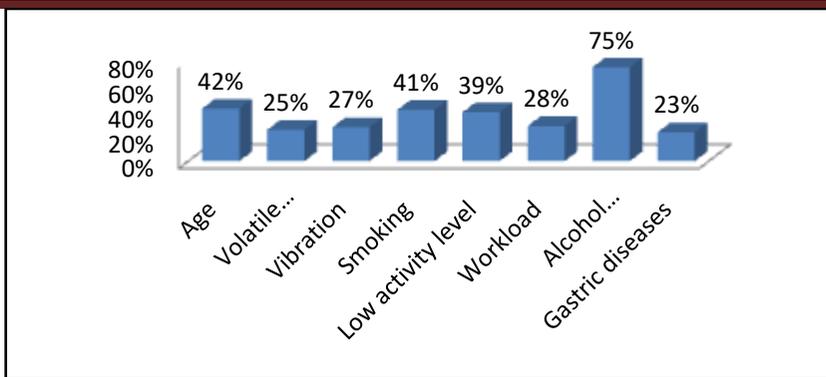


Figure 3: Analysis of the obtained results.

These statistics states that the maximum number of respondents are those who consume alcohol besides that the persons' who face hazardous work and volatile compounds at work also be recommended for H. pylori test which could facilitate early detection and prevention of gastric diseases and cancer(Ramachandran *et al.*,2014).

Harathi Parasur Babu *et al.* (2015) in their paper "Spatial data mining using association rules and fuzzy logic for autonomous exploration of geo-referenced cancer data in Western Tamilnadu, India" stated that data mining with association rules can be used in medical databases for recognizing the hidden disease and their causative agents. The database which they used is divided into taluks where each taluk is related with demography related cancer data and its respective lifestyle.

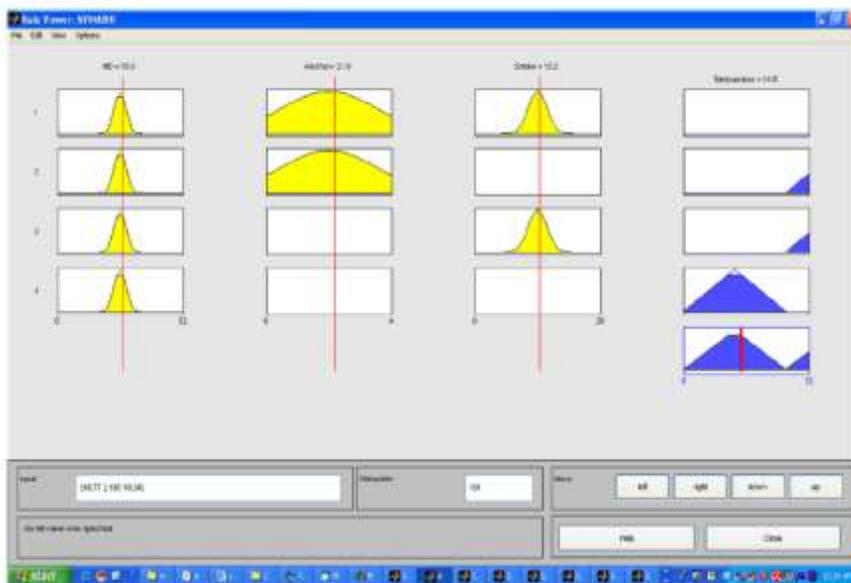


Figure 4: Different Lifestyle's Normal distribution curve and seriousness in Avinashi taluk

In Figure 4 the first three columns represents the antecedents and the fourth column represents the consequent. The variables values are written on the top of each graph which clearly shows the involvement of these particular attributes which causes cancer. Their research helped in the western region of Tamilnadu to recognize the demographical pattern of different lifestyle in occurring of cancer.

Most of the cancer patients at younger age only, use to consume tobacco and alcohol. The researchers' in their paper proved that there exist a relationship between heavy smoking and alcohol consumption in adolescent cancer patients.

**Discussion on the findings:**

So focus on two areas should be there to reduce morbidity and mortality of patients suffering from oral cancer, these areas are: primary avoidance by minimizing risk factors and early recognition. Generally people prefer visiting their physician rather than a dentist but physician are less expected to detect the patient for oral cancer. This is the moral duty of all the primary-care providers for counseling patients

about habits that make them more prone to develop this type of cancer. Habits include tobacco use or excessive alcohol consumption, and if there is some suspicious oral lesion refer patients to an appropriate consultant for further treatment.

Oral cancer gets least amount of attention despite being the most accessible and visible part of human body. It accounts to 30 percent of mortality rate in India, adding to approximately 1 million per year (Uppal *et al.*, 2017). Because of the lack of understanding and alertness, oral cancers, as a group, are also not cured to as much in the post treatment or therapy phase. India is definitely not in protected region because of the easy access to tobacco or other narcotics (as per the Global Adult tobacco Survey) in the country; however, death rate because of oral cancer is preventable. Early detection, accurate treatment plan and its execution, followed by proper post treatment care, can help the person to restore his regular lifestyle in most cases.

## CONCLUSION

Due to the data disparity, it is a complex task to extract useful knowledge from such collections. Data mining is therefore highly effective method to sort out the huge amount of data related problems by the researchers (Goebel *et al.*, 1999; David *et al.*, 2013). Data mining using association rule combined with fuzzy logic is mostly functional in medicine especially in cancer treatments (Delen *et al.*, 2009; Priyanga *et al.*, 2013). During medical activities, large amount of data is generated and to get a comprehensive and systematic amount of data, data mining is highly used. To fulfill the requirements of a wider and deeper level of management, Data mining is the best solution.

In conclusion, this paper is to review the risk factors associated with lifestyle habits which leads to cancer. The World Health Organization (WHO) affirmed that as far as cancer is concerned at least one-third cases are preventable and if they are diagnosis early about 40% of them can be cured. Some of the most common risks are heavy smoking, alcohol consumption, low socioeconomic status, intake of heavy salty foods, consuming very hot drinks, high frequency of *HP* and indoor cooking bread on tandoor. Results also revealed that there is six times increase in the total cases if there will be no control on tobacco smoking and consumption of alcohol.

In the world, many persons are suffering from cancer. Most of the persons even do not know that they are suffering from this disease. After completely affected, this disease has no remedy. So the prediction of having the cancer at an early stage plays a significant role in the diagnosis process (Kharya *et al.*, 2012). If the prediction quality is good, it is believed that various data mining tools (Vital *et al.*, 2015; Naik *et al.*, 2016) are beneficial not only for the patients but also help doctors in making decision with patient consent. This is done by making them understand the risks involved in particular procedure of treatment depending upon each patient's condition.

There is a clear indication that the amount of cancer patients in India is also increasing every year. Depending on the region and severity of disease, several methods are recommended like surgery, chemotherapy or radiation therapy. These are given in solitude or in combination manner. Proper care which involves oncologist, clinical nutritionist and physiotherapists all are required for finest outcome. Cessation of smoking is an important part of treatment that may involve medication for tobacco de-addiction. The awareness must be created among the public about this deadly disease and its prevention. Diet and living style must be taken care of by each and every Indian, to curb the spreading of this deadly disease. One must discourage adolescents from adopting with noxious lifestyles and inappropriate habits which may lead to harmful results; instead they must be encouraged to eat plenty of fresh fruits and green leafy vegetables (five to six portions per day).

## REFERENCES

1. Ahmed, K., Abdullah-Al-Emran, A. A. E., Jesmin, T., Mukti, R. F., Rahman, M., & Ahmed, F. (2013). "Early Detection of Lung Cancer Risk using Data Mining", *Asian pacific journal of cancer prevention*, vol.14,no.1, pp. 595-598.
2. Ali, I., Wani, W. A., & Saleem, K. (2011). "Cancer Scenario in India with Future Perspectives", *Cancer therapy*, vol. 8, no. 1,pp. 56-70.
3. Aly, H. F. (2012). "Dietary habits and relation to cancer disease in different population", *Archives in Cancer Research*, vol. 1, no. 1.
4. Amin, S. U., Agarwal, K., & Beg, R. (2013). "Genetic Neural Network based Data Mining in Prediction of Heart Disease using Risk Factors", *Information & Communication Technologies (ICT)*, pp. 1227-1231.
5. Babu, H. P., Selvaraj, J., Ramachandran, S., Marimuthu, P. D. and Somanathan, B. (2015). " Spatial Data Mining using Association Rules and Fuzzy Logic for Autonomous Exploration of Geo-referenced Cancer Data in Western Tamilnadu, India", *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 4,no. 1,pp. 1-11.
6. Celik, S., Yilmaz, E. M., Özden, F., Kotan, C., & Okut, H. (2015). "The relationship between eating and lifestyle habits and cancer in Van Lake region: another endemic region for esophageal and gastric cancers", *Journal of cancer epidemiology*, vol. 2015, no. 254823

7. Cerquitelli, T., Chiusano, S. and Xiao, X. (2016). "Exploiting Clustering Algorithms in a Multiple-Level Fashion: A Comparative Study in the Medical Care Scenario", *Expert Systems with Applications*, vol. 55, pp. 297-312.
8. David, S. K., Saeb, A. T. and Al Rubeaan, K. (2013). "Comparative Analysis of Data Mining Tools and Classification Techniques using WEKA in Medical Bioinformatics", *Computer Engineering and Intelligent Systems*, vol. 4, no. 13, pp. 28-38.
9. Delen, D. (2009). "Analysis of cancer data: a data mining approach", *Expert Systems*, vol. 26, no. 1, pp. 100-112.
10. Deoskar, P., Singh, D., & Singh, A. (2013). "Mining Lung Cancer Data and Other Diseases Data using Data Mining Techniques: A Survey", *International journal of computer engineering & technology (IJCET)*. Journal Impact Factor, vol. 4, no. 2, pp. 508-516.
11. Einsele, F., Sadeghi, L., Ingold, R., & Jenzer, H. (2015). "A study about discovery of critical food consumption patterns linked with lifestyle diseases using data mining methods", *In Proceedings of the International Conference on Health Informatics*, Lisbon, Portugal, pp. 239-245.
12. Fouad, M. M., Oweis, N. E., Gaber, T., Ahmed, M., and Snasel, V. (2015). "Data Mining and Fusion Techniques for WSNs as a Source of the Big Data", *Procedia Computer Science*, vol. 65, pp. 778-786.
13. Jemal, A., Siegel R., Ward E., Murray T., Xu J. and Thun MJ. (2007). "Cancer Statistics", *CA Cancer J Clin* 57, pp. 43-66.
14. Jeske, D. R., Longmate, J., Katheria, V. and Hurria, A. (2016). "Mining Branching Rules from Past Survey Data with an Illustration Using a Geriatric Assessment Survey for Older Adults with Cancer", *Algorithms Research*, vol. 9, no. 2, pp. 33.
15. Kaladhar, D. S. V. G. K., Chandana B. and P. B. Kumar (2011). "Predicting Cancer Survivability using Classification Algorithms", *LMT*, vol. 34, no. 65.7, pp. 96-106.
16. Kirshners, A., Polaka, I. and Aleksejeva, L. (2015). "Gastric Cancer Risk Analysis in Unhealthy Habits Data with Classification Algorithms". *Information Technology and Management Science*, vol. 18, no. 1, pp. 97-102.
17. Meng, X. H., Huang, Y. X., Rao, D. P., Zhang, Q. and Liu, Q. (2013). "Comparison of Three Data Mining Models for Predicting Diabetes or Prediabetes by Risk Factors", *The Kaohsiung journal of medical sciences*, vol. 29, no. 2, pp. 93-99.
18. Moro, S., Rita, P., & Vala, B. (2016). "Predicting Social Media Performance Metrics and Evaluation of the Impact on Brand Building: A Data Mining Approach", *Journal of Business Research*, vol. 69, no. 9, pp. 3341-3351.
19. Naik, A. and Samant, L. (2016). "Correlation Review of Classification Algorithm Using Data Mining Tool: WEKA, Rapidminer, Tanagra, Orange and Knime", *Procedia Computer Science*, vol. 85, pp. 662-668.
20. Priyanga, A., & Prakasam, S. (2013). "Effectiveness of Data Mining Based Cancer Prediction System (DMBCPS)", *International Journal of Computer Applications*, vol. 83, no. 10.
21. Ramachandran, P., Girija, N., & Bhuvaneshwari, T. (2014). "Early Detection and Prevention of Cancer using Data Mining Techniques", *International Journal of Computer Applications*, vol. 97, no. 13.
22. Rathore, S., Hussain, M., Ali, A., & Khan, A. (2013). "A Recent Survey on Colon Cancer Detection Techniques", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 10, no. 3, pp. 545-563.
23. Senturk, Z. K., & Kara, R. (2014). "Breast Cancer Diagnosis via Data Mining: Performance Analysis of Seven different algorithms", *Computer Science & Engineering*, vol. 4, no. 1, pp. 35.
24. Thangaraju, P., Karthikeyan, T. and Barkavi, G. (2014). "Mining Lung Cancer Data for Smokers and Non-Smokers by Using Data Mining Techniques", *International Journal of Advanced Research in Computer and Communication Engineering* vol. 3.
25. Uppal, S and Puri, Abhishek (2017), "Watch that mouth, literally", *The Tribune Trends*, 21 January, pp. 3.
26. Vidhu, M. R. and Kiruthika, M. S. (2016). "A New Feature Selection Method for Oral Cancer Using Data Mining Techniques", *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 5, no. 1.