International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Measuring India's Tobacco-to-Cardic Disease Pathway

Lekha Somaiya[1], Lekha Somaiya[2], Vijay Kumar H[3], Vijay Kumar H[4]

[1,2,3,4]School of CSA, Reva University, Bengluru, India

Abstract

This paper gives an overview of the link between tobacco consumption and cardiovascular heart disease (referred to as CVDs in the paper) in India. It therefore aims to look at how time of tobacco exposure affects the health of the heart using a combination of medical and analytical understanding. Apart from literature survey, the paper also suggests for future work on AI based prediction models, Behavioral and Policy level intervention analysis. It shows how moving away from simples associations, techniques that can be used for evidence based prevention.

Keywords: Tobacco-Use, Cardiovascular Disease, Risk Assessment, Public Health, India, Tobacco Consumption, Cardiac Risk Factors, Epidemiology, Statistical Modeling, Predictive Model, Data Mining, Machine Learning, Logistic Regression, Random Forest, Health Informatics, Risk estimation, Preventive Care.

Introduction:

India is facing increasing burden of cardiovascular diseases. They now contribute significantly to both mortality as well as morbidity. Use of tobacco in any form both smoking as well as smokeless tobacco has been identified as one of the most important preventable risk factors. Despite, campaigns for awareness and policies for abatement of tobacco its consumption continues various socio-economic groups. This study will try to understand the pathway leading to cardiac diseases due to use of tobacco especially exploring the various biological, behavioral socio- economic factors involved. Basis for a comprehensive framework for early identification and prevention will be laid out.

Background and Context:

Various forms of consumption of tobacco exist in India; so it is all too easy to categorize and various forms of the tobacco consumption. These include cigarettes, bidis and smokeless tobacco such as gutka and paan, which make risk calculation difficult as the effects on the system will be different. Chronic cardiovascular disease (CVD) includes heart attack, high blood pressure, stroke and heart failure. Due to the slow development of most of these conditions early diagnosis of risk is very important, and the impacts of smoking on the development of the CVD are also gradual. Smoking hastens this process through blood vessel damage, oxidative stress and inflammation.

Literature-Review:

Is known that tobacco abuse is associated with Cardiovascular disease (CVD) totally established in the previous studies. Smokers are much more likely to develop the CVD than nonsmokers. Studies have

indicated that smokeless tobacco also causes the risk of high blood pressure and arterial damage. Socioeconomics also have an impact on the rate of tobacco consumption. Materials fromstatistical learningto machine learning, were used in the past few years and applied to prediction of disease risk. But most of them did not include behavioral and social factors in the models, our research is even more meaningful..

METHODOLOGY:

This study has decided to use the simple and logical approach;how tobacco production might lead to the development of CVDs in India, to explore for this pathway. The pathway is not simple nor hypothetical, it is simply displayed to ensure the pathway easy for you to understand and until be implement not only in Malaysia, but also other Asian countries with can provide their own necessary local data. The data is also collected from WHO report, GATS India and many research papers. The important variables included are age, duration of tobacco use, BMI and blood pressure. The pathway simply appears as:
Tobacco Use->Internal Damage->Physical Changes->Increased Risk-> Heart Disease.
In addition, an amount of toxins will be absorbed into the body. These toxins will affect the blood vessels by blocking the blood passage and squeezing the heart. If the tobacco consumption is going on, hypertension and blocked blood vessels will accumulate which can cause cardiovascular diseases.Statistical methods Comparison and pattern recognition are used to process the data.

A. Data Sources:

The evidence is based upon desk review, or secondary data research from valid and widely accepted sources. For example – national health surveys, papers on international health, and published research papers. Significant sources used are:
Global Adult Tobacco Survey (GATS) India
As reported by WHO,
Indian Heart Association data
An academic article in a peer-reviewed journal regarding the use of tobacco or other heart disease.
This wide array of sources gives us a picture of the trends in tobacco use, demographic transitions, and CVH measures in the country.

B. Selection of Key Variables:

To study the link between tobacco and cardiac disease, the cofounding interested variables with medical significance and previous knowledge were selected. The variables can be grouped into three categories:

1. Behavioral Factors

- Quantity (how much cigarette you use; how many cigarettes/day and/week)
- The type of tobacco (smoked or smokeless)
- Time of use (years)

2. Health and Physiological Factors

- Body Mass Index (BMI)
- Bloodpressure levels.
- General condition of the cardiovascular system

3. Demographic and Social Factors

- Age and gender
- Income and education level
- Urban or rural background

The decision to include these variables results in a3 more realistic, and hopefully more useful, picture of what these risk factors look like rather than just giving abstract data.

C. Development of Risk Pathway:

In order to provide a clear explanation of how tobacco uses affects heart health, a conceptual pathway (developed using) is constructed as follows:

Tobacco Use -> Internal Injury -> Physical Changes -> Increase Risks-> Congenital Heart Defects

- It introduces contaminations into the body.
- that these chemicals injure blood vessels and load the heart
- *This leads, in its turn, to the formation of blood pressure and blocked arteries.
- At alarming rates, we see that rates of cardiovascular diseases increase sharply.

For this simple path it is easier to see the trend in risk in a fairly uncomplicated way.

D. Analytical Approach:

The study employs a field study combination of four basic analytical methods to interpret the data:

-- The descriptive analysis to see the trend of the tobaccos consumed

Tobacco-using and non-using comparisons' assessments of atypicality

Correlation based understanding to observe relationship, relationships between the variables We want from results is not only we want to generate the result,

What we do we want to be explaining the simplified to use in the reality.

E. Risk Categorization:

To make the findings more practical, individuals are grouped into three levels: Low Risk Moderate Risk High Risk

This classification can also be used to define different categories of people who need to be put forward.

F. Reason for Choosing:

This Approach This approach has been designed to be simple, transparent and useful. While predictions derived using more sophisticated machine learning algorithms are likely to be more accurate; they can more difficult to interpret.

Tobacco-to-Cardiac Risk Pathway:

The pathway from tobacco use to heart disease can be understood in stages:

Tobacco exposure ; Biological impact ; Physiological changes ; Increased risk ; Cardiovascular-disease.

Risk Analysis:

Risk analysis in is aiming to identify those factors or variables which have greater probability of causing CVD. The principal risk factors are:

Tobacco used duration How often they eat Food:

A) Age, gender.

Lifestyle factors and eating and physical activity activities

* Socio-economic background. It should be noted that these variables have the potential to be associated with one another, for instance a long life history of tobaccosmoking coupled with a sedentary lifestyle, will significantly increase a person's opportunity of cardiac risk.

Future Scope and Research Directions:

Essentially, it provides an insight into how tobacco use lead to cardiovascular diseases but there definitely exist much scope of improvement and expansion in this work.

Major future directions include using machine learning and artificial intelligence based models to improve the accuracy of cardiac risk prediction. Such models when used for larger, real-time data can help identify those at high risk more quickly.

Another improvement would be to use in the future wearable devices such as watches and fitness trackers. These devices can measure a number of parameters in the body including heart rate and blood pressure and so could potentially be used to measure risk in real time rather than data already collected.

It can also be predicted in future using personalized risk assessment system, which will give the prediction of different risk for different individual according to the different lifestyle, general health, cigarette intake pattern.

Additionally, through a study on behavioral tendencies and social causes behind the persistent consumption of tobacco despite the knowledge of its harmful effects, it might result in more effective awareness initiatives.

Finally, there is scope to developsimulation models at policy levelthat would allow exploring the impact of different government interventions, such as road shows and taxation, on consumption and thus public health. Overall, healthcare, data analysis and behavioral science need to go hand in hand for future work.

DISSCUSION:

The study shows that the use of tobacco is an important factor in increasing the risk of CVD in India. But the effects are not same in all groups. Indian Urban population groups have an increased risk due to same popularity of other risk factors in other urban and all rural population groups. The shows importance of behavioral factors also into risk analysis. Most of the existing methods account heavily for behavior and largely ignore the social factors affecting consumption of tobacco.

CONCLUSION:

A. The paper packs a complete model of the mechanism through which use of tobacco results in CVDs in India. Results lends support to the use of tobacco as a high risk factor as well as a means of early prevention. With future oriented advancements with data oriented techniques and state of the art technology can deliver an even higher level prevention and treatment. This problem has to be addressed by the health system, policy maker and the people.

Algorithm :

Algorithm 1: Logistic Regression-Based Cardiac Risk Prediction

Algorithm 2: Logistic Regression for Cardiac Risk Prediction

A. Objective

In the present case the Logistic Regression simply takes the predictors and attempts to judge whether an individual goes into the high risk or the low risk group.

B. Mathematical Formulation

Logistic Regression models the probability of an outcome using the sigmoid function: Where:

- = Target variable (0 = Low Risk, 1 = High Risk)
- = input features (Age, Tobacco Duration, BMI, Blood Pressure,...)

- = Intercept
- Model coefficients.

C. Algorithm Steps

1. Data Preparation

- Prepare dataset with health and tobacco variables
- Deal with missing values and normalize data.

2. Feature Selection

- Select other important features for example, age, period of smoking, BMI, blood pressure.

3. Model Training

- Apply the training data to the fit the Logistic Regression model28.
- Parameter estimation with the method of Maximum likelihood estimation (MLE)

1. Prediction

- Calculate probability 12345,6 7.3 Transition from m to p in using Weights and Tables 8 Assume p5 if p5 40 or 50. If p 4 in (10), then it is 10 61 while in (11) it is 010 4.61. Since p5 50 40 00, it is more profitable to use 4.43 in (11). Therefore, p5 50.
- If , then Classify as High Risk, otherwise Low Risk

2. Model Evaluation

- Evaluate using Accuracy, Precision, Recall, and ROC-AUC

D.Interpretation :

Logistic Regression in the sense that it gives information about whether a particular variable is damaging or helpful.

For example, the coefficient for cigarettes duration, the higher the contribute to cardiac.

E. Advantages

- Conclusions are very easy to draw
- That isok(+k, x) (or, equivalently, isok (positive, x)) is a candidate for binary classification. where k (+k, x) is one of the feature we want to use for classification andisok (+k, x) returns 0 if the featurek (+k, x) is positive and 1 otherwise.
- Accommodates to medical datasets
- Algorithm 2: Random Forest for Risk Analysis and Feature Importance

A. Objective

Random Forest can be used to improve prediction accuracy and assess the importance of the most significant variables to the outcome of cardiovascular disease.

B. Mathematical Formulation

Random Forest combines predictions from multiple decision trees: Where:

- The; B and E equal final prediction
- = One of the predicted classes by the decision tree
- = Number of trees in a species

C. Algorithm Steps

1. Bootstrap Sampling

- Produce several new sets of data by randomly drawing values with duplicate samples from the initial set

2. Tree Construction

- For each dataset: Randomly choose a set of features18. Build a decision tree by splitting criterion(Gini Index or Entropy).

3. Prediction Aggregation

- Train trees? of the prediction: to average of--training data is: 00 0 CB. Predict: ; test: 40 0 CB; So, need to find a value of that minimizes path of test data.
- Use majority voting for classication

4. Feature Importance Calculation

- Quantify the average reduction in impurity over all splits in each tree.

D. Feature Importance Formula Where:

- = the importance of a feature

E. Interpretation

Random Forest aggregates to determine which factors are "most important" contributing to risk of cardiac disease.

The following would be an illustration that rendered tobacco duration and blood pressure from the most strongest predictor variables.

F. Advantages

- Very high accuracy
- Does not impose certain 'functional' forms on the relationships of interest

The algorithm provides feature importance (very useful for analysis)

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation .

1. World Health Organization, "WHO global report on tobacco epidemic," Geneva, Switzerland, 2023.
2. Ministry of Health and Family Welfare, Government of India, "Global Adult Tobacco Survey (GATS) India Report," New Delhi, India, 2017.
3. G. A. Roth et al., "Global burden of cardiovascular diseases and risk factors, 1990–2019," J. Am. Coll. Cardiol., vol. 76, no. 25, pp. 2982–3021, 2020.
4. Centers for Disease Control and Prevention, "Smoking and cardiovascular disease," CDC Report, 2022.
5. D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, Applied Logistic Regression, 3rd ed. New York, NY, USA: Wiley, 2013.
6. L. Breiman, "Random forests," Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.
7. P. Jha and R. Peto, "Global effects of smoking, of quitting, and of taxing tobacco," N. Engl. J. Med., vol. 370, no. 1, pp. 60–68, 2014.
8. S. Gupta, C. Subramoney, R. B. Sekhar, and K. S. Reddy, "Tobacco use and cardiovascular diseases in India," Lancet, vol. 370, pp. 182–183, 2007.
9. K. S. Reddy and S. Yusuf, "Emerging epidemic of cardiovascular disease in developing countries," Circulation, vol. 97, no. 6, pp. 596–601, 1998.
10. A. Teo et al., "Tobacco use and risk of myocardial infarction in 52 countries," Lancet, vol. 368, no. 9536, pp. 647–658, 2006.

11. M. Ezzati and A. D. Lopez, “Regional, disease-specific patterns of smoking-attributable mortality,” Lancet, vol. 362, no. 9382, pp. 847–852, 2003.
12. S. Basu, D. Stuckler, G. McKee, and M. Galea, “The health effects of tobacco consumption in India,” BMJ, vol. 345, 2012.
13. R. Prabhakaran, P. Jeemon, and S. Roy, “Cardiovascular diseases in India: current epidemiology and future directions,” Circulation, vol. 133, no. 16, pp. 1605–1620, 2016.
14. T. W. Valente, “Social network influences on tobacco use,” Ann. Rev. Public Health, vol. 31, pp. 193–211, 2010.
15. J. Esteva et al., “A guide to deep learning in healthcare,” Nat. Med., vol. 25, pp. 24–29, 2019.