

## **Spatial Modelling of Cardiovascular Mortality Patterns: A GIS based Study of Rural Bist Doab Region (Punjab, India)**

**Anandvir Kaur Saini**

*(Department of Geography, Panjab University, Chandigarh, India)*

---

**Abstract:** Population health is a major area of concern for researchers in developing nations with large demographic base. Most of the geographical studies done in India have focused on modelling communicable diseases only. The importance of constructing spatial models for non-communicable diseases has risen tremendously in recent times due to their ever-increasing burden. The present paper attempts to create a regression model of cardiovascular mortality in Bist Doab region of Punjab (India). A composite risk factor score was calculated block-wise and was used as an explanatory variable for the observed spatial patterns of cardiovascular mortality. The regression modelling was done using Ordinary Least Squares method of Spatial Analyst tool in ArcGIS 9.3. The resultant regression coefficient was 0.44 and the model was capable of explaining nearly 35% of variation in the observed spatial patterns of cardiovascular deaths. The model was statistically significant at 0.05 level. Given the complex aetiology of cardiovascular diseases and the lack of block-wise data on several key risk factors, it was concluded that this model displayed considerably good performance. Finally, Moran's I test was used to check spatial autocorrelation to ensure that the residuals are not spatially autocorrelated.

**Keywords:** Cardiovascular diseases, GIS, Punjab, Regression model, Risk factors.

---

### **I. Introduction**

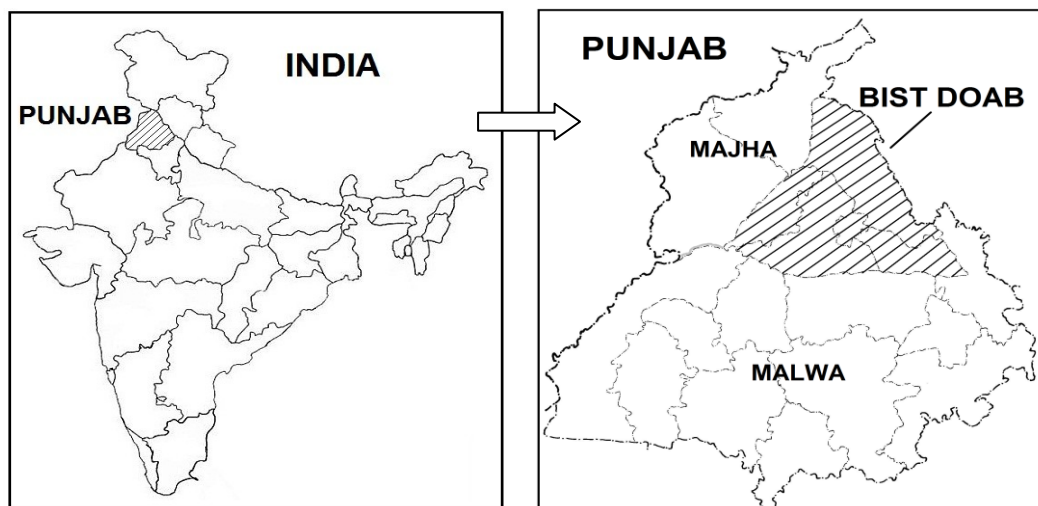
The importance of constructing spatial models for non-communicable diseases has risen tremendously in recent times due to their ever-increasing burden. The spatial etiology of a disease can be investigated by discovering where the disease occurs and subsequently relating its spatial characteristics to the relevant features of physical and human environment. Out of all diseases, cardiovascular diseases are the world's largest killers that claim 17.1 million lives every year, representing 29% of all global deaths. About 82% of the deaths caused by cardiovascular diseases occur in low- and middle-income countries [1]. These diseases are typically multi-causal in which several factors are implicated.

The problem of cardiovascular diseases in India has assumed epidemic proportions, earning it the dubious sobriquet of being the 'Heart Disease Capital of the World' [2]. The World Health Report (2002) has projected cardiovascular diseases to be the largest cause of death and disability in India by 2020, which will account for 42% of the total mortality from all causes [3]. Within the country, the state of Punjab has the highest proportion (approximately 50%) of cardiovascular deaths out of the total mortality. Punjab can be divided into three distinct cultural regions namely, Majha, Malwa and Bist Doab (Map 1 and 2). The present paper aims to construct a regression model of cardiovascular mortality recorded in rural areas of the state's Bist Doab region specifically.

Bist Doab is a natural region which extends from 30°57' N to 32°7' N latitude and 75°4' E to 76°38' E longitude. It is bounded by river Beas in the north-west, Satluj in the south and Shiwalik hills in the east. The region accounts for nearly 20% of Punjab's total population and around 68% of its people live in rural areas. In terms of administration, the region has four districts namely, Jalandhar, Hoshiarpur, Kapurthala and S.B.S. Nagar (Nawanshahar), which encompass 30 Community Development Blocks. There are 3,528 villages, 35 towns and 2 cities (Jalandhar and Hoshiarpur) in the region. The present study pertains to the rural areas of the region.

### **II. Data and Methods**

The study has been conducted for the year 2009 and the analysis was done at block-level. The cause-specific mortality data for this study has been obtained from village-wise death registers available in the Office of Registrar (Births and Deaths) of all the four districts of the Bist Doab region and was aggregated at block-level. Apart from it, data on various risk factors of cardiovascular diseases has been compiled from 'Block at a Glance' publication of Economic and Statistical Organization (Punjab), Website of Excise and Taxation Department (Punjab) and Defense Mapping Agency Topographic Center, Washington, D.C. (U.S.A.).



Map 1: India - Location of Punjab

Map 2: Punjab - Location of BistDoab

The following methods and techniques were applied for analysis and representation of data:

**1. Proportional Mortality Rate:** The Proportional Mortality Rate was calculated block-wise using the formula given below and the values were plotted on map using Choropleth technique in ArcGIS 9.3 software.

$$\text{Proportional Mortality Rate} = \frac{\text{Number of deaths from Cardiovascular Diseases} \times 100}{\text{Number of deaths from All Causes}}$$

**2. Cardiovascular Risk Index:** This index was developed to study the extent of risk to cardiovascular health in different parts of the study region by exploring the composite spatial distribution of various risk factors. The indicators used in the formulation of this index were chosen after a thorough review of literature and keeping in view the availability of relevant data. The cardiovascular risk index (CRI) consisted of the following five components:

- (i.) **Alcohol Vends Index (AVI):** It has been observed that the mere presence or absence of an alcohol vend in a village directly affects the availability and level of alcohol consumption in the surrounding area. Thus the Alcohol Vends Index was used to portray the availability of alcohol in the region. The indicator used for its calculation was 'number of males per alcohol vend' in each block. Lower the number of males per alcohol vend in an area, higher will be the risk to cardiovascular health owing to higher per capita availability of alcohol. Alcohol Vends Index (AVI) was arrived at using the following formula:

$$\text{AVI} = \frac{\max(x) - x}{\max(x) - \min(x)}$$

- (ii.) **Milch Cattle Index (MCI):** This index reflected the availability of milk and milk products in the study area. Since block-wise data on availability or consumption of milk and milk products was not available, so the 'number of milch cattle per person' was used as a surrogate indicator such that higher the number of milch cattle per person, higher will be the risk of developing cardiovascular ailment due to greater fat consumption. Milch Cattle Index (MCI) was calculated as follows:

$$\text{MCI} = \frac{x - \min(x)}{\max(x) - \min(x)}$$

- (iii.) **Health Services Index (HSI):** Health Services Index was calculated from the indicator named 'population served per health institution'. Lower the population served per health institution in an area, lower is the risk to cardiovascular health. Since the data on the number of private health institutions was not available at block-level, so only government ones were considered. The formula used for computing Health Services Index (HSI) was:

$$\text{HSI} = \frac{x - \min(x)}{\max(x) - \min(x)}$$

- (iv.) **Average Slope Index (ASI):** The role of living on moderate altitudes with uneven terrain in promoting physical activity has been well accepted in existing literature. The eastern parts of Bist Doab region are hilly while the central and western parts are flat plains. Thus physiography of the region has significant influence on the spatial distribution of cardiovascular mortality. Robinson's method of slope analysis was applied on contour map of the region for calculating the block-wise average slope in degrees. Higher the average slope of an area, lower will be the risk of acquiring cardiovascular diseases, owing

to the greater level of physical activity involved. The Average Slope Index (ASI) was calculated as given below:

$$ASI = \frac{\max(x) - x}{\max(x) - \min(x)}$$

- (v.) **Industrial Units Index (IUI):** The industries present in an area are a marker of physically inactive workforce and in some cases, high levels of air pollution as well. Therefore, higher the number of industrial units per hectare, higher is the risk to cardiovascular health. The Industrial Units Index was calculated from this indicator using the following formula:

$$IUI = \frac{x - \min(x)}{\max(x) - \min(x)}$$

Thus, the overall Cardiovascular Risk Index (CRI) was calculated as follows:

$$CRI = \frac{(AVI + MCI + HSI + ASI + IUI)}{5}$$

A choropleth map was prepared to depict the block-wise distribution of CRI and this score was eventually used in the regression analysis of cardiovascular mortality.

**3. Regression Analysis:** The regression analysis was done at block level using Ordinary Least Squares (OLS) method in ArcGIS 9.3 software. The OLS tool created a single regression equation as follows:

$$Y = a + bX + E$$

where, Y = dependent variable

X = independent / explanatory variable

a = intercept

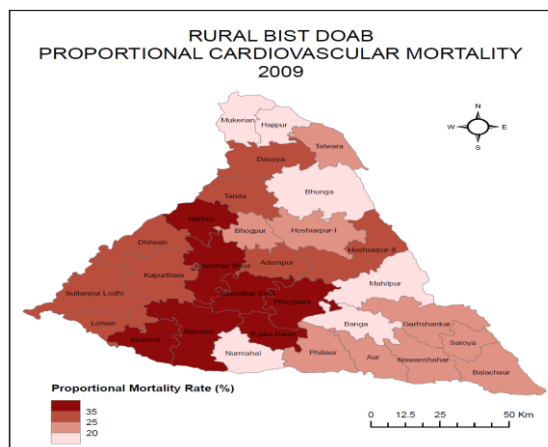
b = regression coefficient

E = random error / residual

The Proportional Mortality Rate (PMR) was the dependent variable and Cardiovascular Risk Index (CRI) was used as the explanatory variable. The value and sign of regression coefficient denotes the strength and type of relationship the explanatory variable has to the dependent variable. In case of linear regression, the regression coefficient is a constant that represents the rate of change of dependent variable as a function of change in explanatory variable. The intercept represents the expected value for the dependent variable if the explanatory variable is zero. The residuals denote the difference between observed and predicted values of the dependent variable and are shown in the OLS output map as under predictions (where the actual proportion of cardiovascular mortality is lower than predicted) and over predictions (where the actual proportion of cardiovascular deaths is higher than predicted). In a properly specified regression model, the over and under predictions display a random spatial distribution. Moran's-I spatial autocorrelation test was used to ensure that the over/under predictions do not portray spatial clustering. If the over/under predictions tend to cluster, it implies that some key explanatory variable is missing. The results of regression analysis were interpreted from the numeric output of OLS method.

### III. Spatial Variation In Proportional Mortality Rate

The total number of deaths from all causes occurring in rural areas of Bist Doab region in 2009 was 24,573. Out of these, 6,796 deaths (nearly 28%) were caused due to cardiovascular disorders. Approximately 65% of the persons who died from cardiovascular diseases were males, while the rest 35% were females. The distribution of cardiovascular mortality observed at block level shows the following spatial patterns (Map 3):

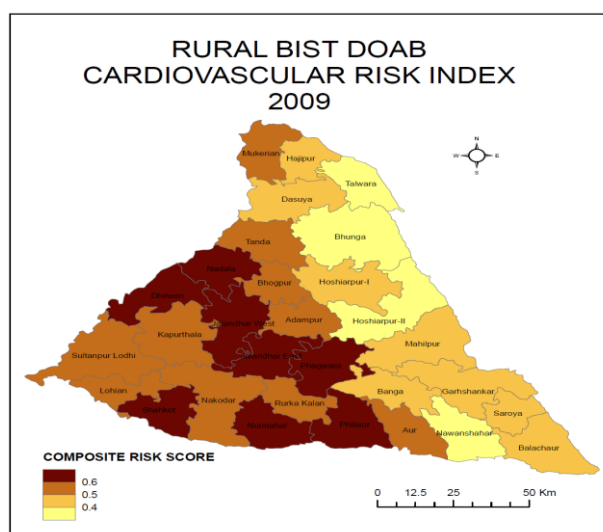


Map 3: Rural Bist Doab - Proportional Cardiovascular Mortality, 2009

- (i) **Areas of Very High PMR (above 35%):** The central parts of the region portray a very high rate of mortality (PMR above 35%), which include the blocks of Nadala (53%), Phagwara (45%), Nakodar (45%), Jalandhar East (43%), Shahkot (41%), Rurka Kalan (37%) and Jalandhar West (36%).
- (ii) **Areas of Moderately High PMR (25 to 35%):** The eastern parts and some blocks of the western parts of the region reflect moderately high cardiovascular mortality (PMR from 25 to 35%), which consist of Tanda (32%), Dasuya (31%), Dhilwan (30%), Adampur (29%), Lohian (28%), Hoshiarpur-II (27%), Sultanpur Lodhi (27%) and Kapurthala block (26%).
- (iii) **Areas of Moderately Low PMR (20 to 25%):** The blocks falling in the category of moderately low cardiovascular mortality (PMR from 20 to 25%) are largely situated in the south-eastern parts of the region and some of them lie in the east-central parts. These blocks include Saroya (25%), Garhshankar (24%), Nawanshahar (24%), Balachaur (23%), Hoshiarpur-I (22%), Phillaur (22%), Bhogpur (22%), Talwara (21%) and Aur (20%).
- (iv) **Areas of Low PMR (below 20%):** The category of low cardiovascular mortality (PMR below 20%) encompasses most of the blocks lying in the eastern hilly parts of the study area and one block of the south. These include Nurmahal (19%), Bhunga (19%), Mahilpur (18%), Mukerian (18%), Hajipur (17%) and Banga block (11%).

#### IV. Geographical Distribution Of Cardiovascular Risk Index

The pattern of cardiovascular risk index displays the distribution of geographically antecedent sources of major risk factors of cardiovascular diseases. The index has great variation from one part of study area to the other, as discussed below (Map 4):



Map 4: Rural Bist Doab - Cardiovascular Risk Index, 2009

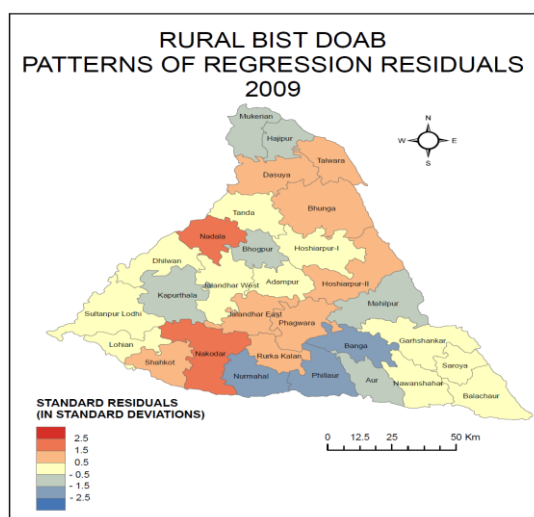
- (i) **Areas having High CRI (above 0.6):** The level of risk is high (CRI above 0.6) in and around the central parts of the region encompassing the blocks of Nadala (0.73), Phagwara (0.70), Phillaur (0.69), Jalandhar East (0.69), Shahkot (0.67), Dhilwan (0.65), Jalandhar West (0.63) and Nurmahal (0.62).
- (ii) **Areas having Moderately High CRI (0.5 to 0.6):** The risk score is moderately high (0.5 to 0.6) in the adjoining areas lying both to the east and west of this high risk axis. This group includes the blocks of Rurka Kalan (0.59), Nakodar (0.59), Kapurthala (0.58), Adampur (0.58), Bhogpur (0.57), Lohian (0.57), Sultanpur Lodhi (0.56), Tanda (0.52), Mukerian (0.52) and Aur (0.51).
- (iii) **Areas having Moderately Low CRI (0.4 to 0.5):** The third category of moderately low risk score (0.4 to 0.5) is spread over the south-eastern parts and some blocks of the north-eastern parts. The blocks falling in this group include Banga (0.47), Garhshankar (0.47), Mahilpur (0.46), Hajipur (0.45), Dasuya (0.42), Balachaur (0.42), Hoshiarpur-I (0.41) and Saroya (0.41).
- (iv) **Areas having Low CRI (below 0.4):** The spatial expanse of the areas recording low cardiovascular risk score (below 0.4) is spread over the eastern blocks of Nawanshahar (0.38), Hoshiarpur-II (0.37), Talwara (0.23) and Bhunga (0.20).

#### V. Regression Model of Cardiovascular Mortality

In order to explain the variation in cardiovascular mortality, a regression model has been created at block level by considering Proportional Mortality Rate (PMR) as the dependent variable and Cardiovascular

Risk Index (CRI) as the explanatory variable. The model has been used to determine the spatial dependence of variation in cardiovascular mortality on the devised cardiovascular risk score. The value of regression coefficient comes out to be +0.44. The sign of the coefficient reflects positive linear direction of the relationship. This means that as the cardiovascular risk score goes up, the rate of cardiovascular mortality also goes up.

The regression coefficient is statistically significant at 0.05 level. The coefficient of determination (R-squared) is 0.35 which implies that the resultant model explains about 35% of the variation in cardiovascular mortality. The performance of the model could not be enhanced due to serious limitation of data on important behavioural risk factors of cardiovascular diseases like smoking, physical inactivity etc.



Map 5: Rural Bist Doab - Patterns of Regression Residuals, 2009

The spatial autocorrelation of the regression residuals was checked using Moran's I test. When a regression model is performing well, no structure is displayed by the residuals and they reflect a random spatial distribution. The results of test for spatial autocorrelation reveal that the value of Moran's I index is 0.16 and the over- and under-predictions of the regression model do not cluster spatially (Map 5). Thus the composite cardiovascular risk index devised in this paper has been successful in explaining 35% of the variation in cardiovascular mortality observed at block level.

## VI. Conclusion

The present research work has its own peculiar strengths and limitations. The paper attempted to fill the glaring research gap on geographical patterns of cardiovascular mortality in Punjab state of India, which holds the top slot in prevalence of cardiovascular diseases among all states of the country. In order to sum up the effect of major cardiovascular risk factors, the present research work devised a composite risk index. The main advantage of using this index was that it summarized the geographical antecedents of some of the important risk factors proven to influence cardiovascular health of people. Additionally, it was easier to interpret one number than finding patterns across multiple indicators. Cardiovascular diseases are extremely multi-causal in nature.

Therefore, the construction of a reasonably good regression model using individual independent variables was not an easy task. Hence the composite index was calculated to overcome this problem and facilitate the regression analysis of cardiovascular mortality patterns.

The main limitation of this study was the unavailability of spatial data on key cardiovascular risk factors like physical inactivity, smoking, alcohol consumption, intake of fats, psychosocial stress etc. Therefore, indirect measures and surrogate indicators were used to determine the extent of risk to cardiovascular health in various parts of the region. However, the paper provides a strong basis for exploring deeper research questions in this direction. In order to determine why the composite index was rising or falling in spatial terms and what to do in response, there is a need to go back to the factors upon which the index was based. For addressing this issue, the individual determinants of cardiovascular diseases should be dealt with separately. This raises important agenda for future research on the geography of cardiovascular diseases.

Keeping in view the complex aetiology of cardiovascular diseases and the lack of block-wise data on several key risk factors, it can be concluded that the regression model constructed in this paper displayed considerably good performance. The value of the resultant regression coefficient turned out to be 0.44 and the model was capable of explaining nearly 35% of variation in the observed spatial patterns of cardiovascular

mortality in the Bist Doab region of Punjab. The regression model was statistically significant at 0.05 level. In Health Geography, the potential of GIS should be tapped in building such models for studying the causal factors of both infectious and non-infectious diseases. Additionally, the spatial visualization of disease patterns can provide valuable inputs to rural-targeted health policies.

#### **Acknowledgements**

The author duly acknowledges the financial support provided by University Grants Commission (India) in the form of Senior Research Fellowship for carrying out the present research work.

#### **References**

- [1] World Health Organization, Cardiovascular diseases ([http://www.who.int/cardiovascular\\_diseases/en](http://www.who.int/cardiovascular_diseases/en), 2010)
- [2] The Tribune, Why India must work harder at heart health (Chandigarh, Dec 10, 2010)
- [3] World Health Organization, The world health report ([http://www.who.int/whr/2002/en/whr02\\_en.pdf](http://www.who.int/whr/2002/en/whr02_en.pdf), 2002)