

Spatial Distribution of Cardio-Vascular Diseases in India

Ayantika Biswas (✉ ayantika111@gmail.com)

International Institute for Population Sciences

Shri Kant Singh

International Institute for Population Sciences

Jitendra Gupta

International Institute for Population Sciences

Research Article

Keywords: Cardio-vascular diseases, socio-demographic factors, Lifestyle, Moran's I, spatial autocorrelation, autoregression

Posted Date: June 4th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-561563/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Objective: Cardio-vascular Diseases (CVDs) are a leading cause of death and disease burden across the world, and the burden is only expected to increase as the population ages. The objective of this paper is to explore the patterns of CVD risk factors among women in the late reproductive ages (35-49 years) across 640 districts in India, and investigate the association between area-level socioeconomic factors and CVD risk patterns., using a nationally representative sample of 239,729 women aged 35–49 years from all 36 States/UTs under NFHS-4 (2015–16).

Methods: Age-standardized prevalence of CVDs have been calculated, along with 95% CI among women in their late reproductive ages (35–49 years) in India. The spatial dependence and clustering of CVD burden has been examined by Moran's I indices, bivariate Local Indicator of Spatial Autocorrelation (LISA) cluster and significance maps. Ordinary Least Square (OLS) regression has been employed with CVD prevalence as the outcome variable. To consider for spatial dependence, Spatial Autoregressive (SAR) models have been fitted to the data. Diagnostic tests for spatial dependence have also been carried out to identify the best fit model.

Results: Higher values of Moran's I imply high spatial autocorrelation in CVD among districts of India. Smoking, alcohol consumption, hailing from a Scheduled Caste background, more than 10 years of schooling, as well as urban places of residence appeared as significant correlates of CVD prevalence in the country. The spatial error model and the spatial lag model are a marked improvement over the OLS model; among the two, the spatial error model emerging to be the most improved of the lot.

Conclusions: A broader course of policy action relating to social determinants can be a particularly effective way of CVD risk addressal. Social policy interventions related to health like reduction in inequalities in factors like education, poverty, unemployment, access to health-promoting physical or built-environments are crucial in tackling the long-term effects of CVD inequalities between geographical areas.

Background

In the last few years, spatial analysis has gained relevance in epidemiological studies and the identification and management of risk factors associated with diseases [1,2]. A majority of these studies employ spatial statistics to reveal factors, playing an essential role in decision-making, planning interventions, and distributing available resources. Although particularly useful for infections requiring vectors [3,4], they are also valuable for studying cardio-vascular diseases (CVDs) and other non-communicable disorders [5,6,7]. The techniques employed in spatial analysis can be difficult to explain to a non-specialist, but its outcome, e.g., a map, can be immediately grasped by a non-specialist. Different spatial techniques are helpful at different levels, and they are complementary to each other.

Cluster detection is an important epidemiological tool, as it can help identify factors spatially closely related to a disease. Positive spatial autocorrelation (SA) implies the rates for a given phenomenon are

likely to be similar for neighboring areas compared to the rates of geographically distant regions [8,9]. In the case of health variables, too, the variables tend to be spatially correlated [10,11] with their underlying related factors, which is because of the high probability of closely situated areas having similar underlying factors with regards to different phenomena [9].

There is a difference between the global and local situation when it comes to clustering. When locating global clusters, the focus is on their existence, not location [12,13]; local cluster analysis aims to quantify SA and clustering within small geographical units in the study area [13]. Moran's I is a commonly used spatial statistic for the detection of global clustering [14], with another tool called local indicators of spatial association (LISA) used for finding local clustering [15,16]. Next, regression techniques are employed to determine possible associations between variables of interest and decipher the strength and direction of the association of the relation. Other techniques utilized for estimating spatial regression are the ordinary least squares (OLS) [17], bivariate LISA [18,19], the generalized additive model [20], the spatial lag model [21], and the spatial error model.

CVDs are a leading cause of death and disease burden across the world, and the burden is only expected to increase as the population ages [22,23,24]. Most commonly used CVD risk prediction algorithms have been derived from the Framingham Risk Equation (FRE), used in general practice to assess risk for individual patients [25]. The trend in primary prevention of CVDs has traditionally been to depart from the relative risk factor assessment but treat these factors as absolute CVD risk [26]. The most effective prevention strategies require knowledge and a contextualized understanding of people, communities, environments, as well as variations in CVD risk. Despite the availability of clinically proven CVD risk factor assessment tools, the most at-risk populations rarely take part in such assessments until disease progression is well underway. Though imprecise proxies for risk can be wielded for community-level risk estimation, a considerable knowledge gap persists due to the unavailability of fine-grained population tools to predict "hotspots" for the future CVD risks from general practice clinical data [27].

There are a few studies that have attempted to inspect the spatial variation of NCD risk at a smaller geographic scale worldwide. Noble et al. examined the feasibility of mapping chronic disease risk among the general population. They created a small-area map of diabetes risk from general practice clinical records in the UK [28].

Factors of importance for cardiovascular disease spatial distribution patterns

Socioeconomic status

The association between socioeconomic status and CVD is chronicled in multiple studies [29]. Spatial analysis has only added to the extant knowledge. In Australia, CVD clustering tended to occur in relatively disadvantaged areas [27]. Even in Harris, Texas, geographically weighted regression depicted a correlation between CVD mortality and social deprivation at the community level [17]. Similar results have been found in Strasbourg, France, where high-risk clusters of myocardial infarction (MI) were seen to accumulate in economically disadvantaged areas, despite good access to health amenities [30].

Socioeconomic status can be a significant determinant of the opportunity of access to health services. Clustering of heart disease mortality before any attempt at transport has been observed in areas of lower socioeconomic status and household amenities [31]. Mapping diseases at the district level and risk analysis with rapid inquiry facility (RIF) techniques showed people residing in highly-deprived areas showed a low relative frequency of prescription statin treatments [32].

Education level

It has been found in a study by Pedigo and colleagues [33] that neighborhoods from eastern Tennessee, USA had high-risk clusters of neighborhoods prone to stroke and MI mortality, along with a high prevalence of low educational level. These results have been replicated in another study conducted in Brazil, reporting spatial clustering for ischemic heart disease mortality and relating it with illiteracy in the study population [18].

Rural vs urban residency

Some characteristic peculiarities are unique to rural and urban areas, which may play a role in CVD development. A recent study conducted in Peru highlighted the clustering of obesity among urban children, while the prevalence was low among rural areas [34]. A study examining MI and stroke determinants in Tennessee found rural residency to be an important factor [35]. There is significant clustering among the rural Taiwanese population, characterized by the underutilization of cardiovascular drugs, which is further connected to cardiology specialists' low presence in some areas [36]. Hence, rural places of residence can select for inadequate access to health services, impeding timely disease management. Spatial regression analysis carried out in Taiwan has further revealed that mortality due to heat or cold waves is more rampant in rural areas, owing to limited access to medical facilities and resources, as opposed to urban metropolitan areas [37].

Alcohol intake

High alcohol intake and heart disease are connected, which is, in fact, also corroborated by a study from Chile, where a particular study region, which was a high-risk zone for CVD deaths, was also associated with alcohol consumption. These results are also supported by the presence of two clusters in the Valparaíso and Biobío areas; the alcohol consumption in Biobío has been among the highest in Chile for the last 45 years [38].

Smoking

Smoking in terms of risk factor clustering has been observed in both United States [39] and China [40]. It has also been observed when studying the CVD incidence among Asian and Caucasian populations [41].

The main objective of this study has been to explore the patterns of CVD risk factors among women in the late reproductive ages (35-49 years) across districts in India, and investigate the association between area-level socioeconomic factors and CVD risk patterns. Hence, the production of fine-grained maps of

CVD risk is possible through this approach for clinicians and policymakers to use, enabling geographic targeting of community interventions for CVDs.

Method

The CVD prevalence within a region can be modeled as a spatial process [42], as can all the demographic and socioeconomic variables associated with the disease prevalence. These observed processes are likely to exhibit spatial dependence, as well as non-stationarity.

In case of disease prevalence, individuals living in close proximity to each other tend to have similar socio-demographic characteristics like age, income, access to healthcare facilities, and, hence, similar disease prevalence, thereby begetting positive spatial dependence. The process is non-stationary, though, given the fact that the disease prevalence is not constant over space: prevalence rates vary from young and wealthy areas to retirement establishments (inconstant mean), variability within a young neighborhood greater than in a relatively older populace (inconstant variance), also, the spatial extent of the spatial dependence varying across regions, from densely populated areas to suburbs (inconstant covariance).

When modeling for the aforementioned spatial process, procedures should be designed with a view to reduce model variance, while also considering for spatial dependence and non-stationarity. Also, the fact that these spatial processes entail both the properties' simultaneous occurrence is worth noticing. Despite the known effects of this relationship [43], most of the existing advanced spatial techniques address only one of the properties: more specifically, spatial autoregressive methods [44] focus on spatial dependence, while disregarding non-stationarity; and local, or geographically weighted methods [45] focusing on non-stationarity while disregarding spatial dependence.

The present study is limited to applying spatial autoregressive procedures; the analytical implementation begins with an examination of spatial dependence. Spatial autocorrelation measures based on the Moran's I [46] are commonly used to test for clustering tendency of medical data, even while analysing for multivariate specifications [47]. Throughout this study, the traditionally used spatial autocorrelation tool Moran's I [46] has been implemented. Even though the authors are well aware of the limitations of the index [48] the interpretation of the Moran's I results can be used to assess the presence and magnitude of spatial dependence. A single index for exploratory analysis, individual variables, and model residuals is vital to decipher spatial dependence, as seen in the current data.

Computation of this index requires the specification of a spatial dependence model, defined by a spatial weight matrix, which can be something as simple as a binary construct, or a more complex specification, which might include various types of weights accounting for distance decay effects.

The identification and quantification of a geographical variable's spatial clustering have been a central issue in all confirmatory and exploratory spatial research. Moran's I provides summary statistics for overall spatial clustering [14,49,50,51,52]. Local Indices, e.g., local Moran's I_i , allow for exploring local

disparities in spatial dependence by quantifying each area's relative contribution to the global measure [53,15,44]. These measures are part of a broader attempt to *spatialise* general statistics, in view of the knowledge and recognition of regular statistical assumptions not applicable to spatial data. For example, data in geographically referenced datasets are not independent, as is generally assumed in statistical analysis, but are influenced by each other, a phenomenon termed as spatial autocorrelation [54]. Spatial distributions are prone to significant local variations, giving rise to discrete spatial patterns in the study area (spatial heterogeneity or non-stationarity) [44,55]. Univariate spatial association measures exclusively on the spatial clustering of observations pertaining to a single variable, while employing a bivariate spatial association ensures deciphering the relationship between two variables, in purview of the topology of the observations. Hence, we are able to *parameterize the bivariate spatial dependence* [56].

Spatial contiguity can be accounted for in various ways [57]: a common way being the definition of k orders of spatial neighbors or the specification of a threshold distance, or even a method based on shared borders (for areal units only). Some methods are reliant on the spatial units' topology, but the computation of spatial neighbours is a far more general method. In any case, the extent of spatial dependence can be defined, either via a maximum distance parameter, or by a maximum number (k) of nearest neighbors.

Spatial autoregressive methods use generalized least squares (GLS) and maximum likelihood (ML) models; the covariance structure generally explained by a conditional autoregressive (CAR), a simultaneous autoregressive (SAR), or a moving average (MA) specification. Generally, a constant covariance structure is assumed, and a spatial weight matrix dictates the spatial units which are spatially dependent [42]. The model is specified by the following equation:

$$Y = X\beta + \rho WY + \varepsilon$$

Where ρ (rho) is the autoregressive parameter and W is the spatial weight matrix.

The autoregressive parameter is a correlation coefficient, ranging between -1 and +1. The definition of the spatial weight matrix entails the same specifications as the spatial autocorrelation index. A backward method of model selection is conducted for all the regressions. Once each regression is specified, the spatial autocorrelation index is calculated on the regression individuals in the aforementioned process. The spatial weight matrix plays an important role in the methodology defined, eventually influencing the spatial autocorrelation index value, as well as the efficacy of the spatial autoregressive models. Spatial matrix definition remains subjective, owing to its resting on an estimate of the spatial dependence in the spatial processes involved [58].

Incorporation of spatial dependence in a regression model can be done in one of two ways; spatial error model or a spatial lagged model. Tools are available assessing which of the models is a better fit for the data, namely, the Lagrange Multiplier tests.

The difference between these models is technical and conceptual. A spatial error model basically implies that the:

"spatial dependence observed in our data does not reflect a truly spatial process, but merely the geographical clustering of the sources of the behavior of interest. For example, citizens in adjoining neighborhoods may favor the same (political) candidate not because they talk to their neighbors, but because citizens with similar incomes tend to cluster geographically, and income also predicts vote choice. Such spatial dependence can be termed attributional dependence" [59].

On the other hand, a spatially lagged model incorporates spatial dependence by adding a "spatially lagged" variable y on the right hand side of the regression equation, which, essentially considers a spatially lagged "dependent" variable among its explanatory factors. Hence, the values of CVD in the neighboring areas of observation $n \sim i$ is an important predictor of CVD in each individual area $n \sim i$. In other words, this implies that spatial dependence may be resulting from a process such as the diffusion of behavior between neighboring units:

"If so the behaviour is likely to be highly social in nature, and understanding the interactions between interdependent units is critical to understanding the behavior in question. For example, citizens may discuss politics across adjoining neighbors such that an increase in support for a candidate in one neighborhood directly leads to an increase in support for the candidate in adjoining neighborhoods" [59].

Data

India's standard Demographic Health Survey (DHS), also known as the National Family Health Survey (NFHS) is a country-wide large-scale, multi-round survey conducted in a representative sample of households. It is conducted in collaboration with International Institute for Population Sciences, Mumbai, India; ORC Macro, Calverton, Maryland, USA, and the East-West Centre, Honolulu, Hawaii, USA. The survey is aided by United States Agency for International Development (USAID), and supported by United Nations Children's Fund (UNICEF). The data generated by the survey is available at the DHS official website (https://dhsprogram.com/data/dataset_admin/index.cfm).

NFHS-4 (2015-16) [60] collected information from 601,509 households and 699,686 women aged 15-49 years and men aged 15-54 years. Information on extant as well as emerging health issues, including the prevalence of CVDs (Heart disease and hypertension), is available for each of the 640 Indian districts. However, the estimates available for men and women were provided at the state level. The study aims to explore the spatial differentials of CVDs among women in their late reproductive years (35-49 years).

For the analyses presented in the current study, the variables have been normalized, using the total resident state women population aged 35-49 years as denominators. This normalization renders all the variables becoming rates, instead of numbers. The data has been stratified by age, and the outliers have been removed, resulting in 239,729 women as the sample size.

For carrying out spatial analysis, districts have been set as the unit of analysis, for which the shapefile for 640 districts has been generated. Next, to examine the spatial dependence and clustering of CVD burden, Moran's I indices, and bivariate Local Indicator of Spatial Autocorrelation (LISA) cluster and significance maps have been generated.

Ordinary Least Square (OLS) regression has been employed with CVD prevalence as the outcome variable. To consider for spatial dependence, Spatial Autoregressive (SAR) models have been fitted to the data. Diagnostic tests for spatial dependence have also been carried out to identify the best fit model. A larger value of Lagrange's Multiplier (LM), a larger R² value, and a lower value of Akaike Info Criterion (AIC), indicates better suitability of the model applied.

Data visualization has been completed using STATA version 15.0 (StataCorp, College Station, TX, USA). GeoDa [61] version 1.12.1.129 has been employed for performing spatial analyses. All the estimates provided in the current study have been appropriately weighted by sampling weights provided in NFHS-4 (2015-16) data.

Results

The present study depicting the spatial variation in the prevalence of CVDs is based on 239,729 women in their late reproductive ages (35-49 years) from 36 states/Union Territories of India.

The descriptive statistics encompassing prevalence results for various socio-demographic and behavioral indicators are portrayed in Table 1. Results from NFHS-4 data show 17.4 percent of women in the late reproductive ages of 35-49 years are currently suffering from CVDs. A higher prevalence of CVDs is found among urban women (21.9 percent), as compared to rural (14.9 percent) women. Women with more than 10 years of schooling have a higher prevalence of CVD-affected individuals (22 percent), as compared to their counterparts with less than 10 years of schooling (18.7 percent), or even those who have not had any formal education (14.7 percent). Women practising Islam (20.3 percent) or religions other than Hinduism or Islam (19.5 percent) have a higher prevalence of CVDs, whereas Hindu women have a lower CVD prevalence (16.8 percent). As the wealth index progresses from least well-off to most well-off, the prevalence of CVDs also gradually increases from 10.3 percent for the least well-off category to 22.8 percent for the most well-off category. Non-smoking women have a higher CVD prevalence (17.4 percent), as compared to their counterparts who smoke (15.9 percent), as do women who do not consume alcohol (17.5 percent), as compared to their counterparts who do (13.9 percent).

Table 1 Demographic/Socio-economic/Behavioural Characteristics and age-standardized prevalence of CVDs along with 95% CI among women in their late reproductive ages (35–49 years) in India, 2015–16

Demographic/Socio-economic/Behavioural Characteristics	Weighted Prevalence (Frequency)	Age-standardized prevalence [C.I.]
Age		
35-39	14.5 (89535)	
40-44	17.09 (76804)	
45-49	20.65 (73390)	
Place of Residence		
Urban	21.58 (86909)	21.87 [21.16-22.58]
Rural	14.73 (152820)	14.85 [14.53-15.18]
Highest educational level		
No formal education	14.78 (109817)	14.67 [14.29-15.05]
<10 yr schooling	18.21 (79402)	18.70 [18.19-19.21]
>10 yr schooling	20.93 (50511)	21.97 [21.20-22.75]
Religion		
Hindus	16.67 (196381)	16.82 [16.45-17.19]
Muslims	19.85 (28919)	20.27 [19.47-21.07]
Others	19.31 (14429)	19.45 [18.46-20.45]
Wealth Index		
Poorest	10.2 (40583)	10.28 [9.87-10.70]
Poorer	13.81 (44944)	13.93 [13.43-14.42]
Middle	17.06 (47827)	17.20 [16.59-17.81]
Richer	20.06 (51164)	20.26 [19.56-20.96]
Richest	22.64 (55210)	22.77 [22.04-23.50]
Smoke		
No	17.24 (235209)	17.43 [17.10-17.77]
Yes	15.94 (4520)	15.84 [14.26-17.42]
Drinks alcohol		
No	17.28 (235224)	17.46 [17.13-17.80]
Yes	14.07 (4505)	13.89 [12.35-15.44]
Total	17.22	17.39 [17.06-17.72]

Table 2 depicts global Moran's I results, which measures the Moran's I test statistic for spatial autocorrelation. Higher values of Moran's I imply high spatial autocorrelation in CVD among districts of India. In case of CVD data from NFHS-4, the Moran's I value of 0.49 is significantly high. Diagnosing spatial autocorrelation as an issue and the errors (residuals) being related systematically amongst themselves implies the usage of a different method to study these phenomena: the spatial regression approach.

For bivariate association between CVDs and independent variables, Moran's I values are 0.45 for women with no formal education and 0.45 for women with more than 10 years of schooling, 0.42 for rural places of residence, 0.36 for hailing from a Scheduled Caste or Other Backward Castes, signifying high spatial autocorrelation among different Indian districts w.r.t. CVD cross-classified by different socio-demographic characteristics.

Table 2 Moran's I statistics showing spatial dependence of CVDs by different household and background characteristics across districts of India, 2015–16

Indicators	Univariate Moran's I	Bi-variate Moran's I
Urban	0.26	0.34
Rural	0.42	0.42
No formal education	0.43	0.45
<10 yr schooling	0.43	0.45
>10 yr schooling	0.30	0.33
SC	0.25	0.36
ST	0.11	0.25
OBC	0.29	0.36
Non-SC/ST/OBC	0.15	0.19
Consume Alcohol	0.08	0.08
Smoke	0.21	0.25
Overall	0.49	

Table 3 presents the results of the Ordinary Least Square (OLS) model, Spatial Lag Model (SLM), and Spatial Error Model (SEM). Concentrating on the OLS model, it is found that smoking, alcohol consumption, hailing from a Scheduled Caste background, more than 10 years of schooling, as well as

urban places of residence appeared as significant correlates of CVD prevalence in the country. However, it would be early to conclude before model diagnostics are taken into consideration.

Table 3 Summary of spatial error model and maximum likelihood estimation for CVD by some selected background characteristics across different districts of India, 2015-16

Indicators	OLS Model	Sig. value	Spatial Lag Model	Sig. value	Spatial Error Model	Sig. value
	Coefficients		Coefficients		Coefficients	
Urban	0.071	0.000	0.064	0.000	0.056	0.004
>10 yr schooling	0.164	0.000	0.063	0.027	0.102	0.013
SC	0.264	0.000	0.209	0.000	0.240	0.000
ST	0.074	0.147	0.125	0.003	0.120	0.003
OBC	0.055	0.278	0.114	0.006	0.120	0.003
Non-SC/ST/OBC	0.075	0.154	0.133	0.002	0.139	0.001
Consume Alcohol	0.092	0.040	0.050	0.165	0.029	0.550
Smoke	0.390	0.000	0.178	0.015	0.128	0.190
AIC	4582.970		4376.950		4364.270	
Lambda					0.663	
Rho			0.601			
R-squared	0.173		0.449		0.470	

A Moran's I score of 0.49 ($p < 0.001$) is highly significant, indicating a strong spatial autocorrelation of the residuals. Both Lagrange Multiplier (lag) and Lagrange Multiplier (error) are significant, pointing towards spatial dependence. Also, Akaike Information Criterion (AIC), measuring the relative goodness of fit for the models, prefers the model with the least value of AIC. The co-efficient of determination (R-squared = 0.173 for OLS) does not indicate OLS to be the best fit for the given CVD data.

Once the presence of spatial dependence has been established, the spatial lag model is applied with maximum likelihood approach, the results of which are presented in table 3. An additional indicator in the form of spatial lag coefficient (ρ) appears in the model. It shows the spatial dependence inherent in the sample data, calculating the average influence of neighboring observations on a particular

observation. Inclusion of this criterion and the fit of the spatial lag model also converts into a higher R-squared value. The effects of other independent variables remain virtually the same.

Next, the spatial error model has been employed. In the SEM, an additional term emerges, i.e., the coefficient on the spatially correlated errors (λ). It has been found to have a positive effect and is highly significant, too. Hence, the general fit improved, as evidenced by the R-squared value.

Both the spatial models are a marked improvement over the OLS model; among the two, the spatial error model emerging to be the most improved of the lot. Variables significant in the OLS model like alcohol consumption were no more significant in the spatial models. Other variables like urban places of residence, hailing from a Scheduled Caste or Tribe or Other Backward caste background, as well as smoking (significant in the SLM but eventually insignificant in the SEM) have all been positively correlated with CVD prevalence in the regions, controlling for selected socio-demographic and behavioural variables.

Bivariate LISA (Cluster and significance) results:

Figure 1 demonstrates bivariate LISA (cluster and significance) map of CVD prevalence across 640 districts in India. There are four hotspot clusters consisting of 73 districts concurrent with high prevalence of CVDs as well as high prevalence of women without any formal education (Fig. 1a). One of the clusters is majorly in Jammu & Kashmir, consisting of the districts like Kupwara, Bandipore, Udhampur, Doda, Kishtwar, Punc, Rajouri, Srinagar, Pulwama, Anantnag. Another cluster consists of parts of Punjab and Himachal Pradesh, with districts like Jalandhar, Una, Ludhiana, Bilaspur, Mandi. Other clusters in Haryana, Andhra Pradesh, Tamil Nadu consisted of districts like Sripotti, Sriramulu, Nellore, Thiruvallur, Kanchipuram, Thirivarur, Nagapattinam, Thottukud, Tirunelveli. So, the clusters are mainly in the northern central region, and additionally comprise of the south-eastern coastal districts. Similarly, there are 117 districts with low CVD prevalence and low percentage of women with no formal education majorly spread across western and central India like Barmer, Jalor, Udaipur, Pratapgarh, Sawai Madhopur, Mahoba, Fatehpur, Hamirpur, Chhindwara, Betul. These areas are broadly constituted by districts from northern and central Gujarat, southern Uttar Pradesh, southern and eastern Rajasthan, Chhattisgarh, southern Bihar, south-eastern Maharashtra.

Also interesting are those district clusters which have high CVD prevalence among women, but low percentage of women without formal education like Aurangabad and Nagpur. Almost 14 districts have low CVD prevalence but high percentage of women without formal education, like Nalgonda, Chittoor, Kolam, Idukki.

In case of bivariate results depicting CVD among women aged between 35-49 years and those with less than 10 years of schooling (Fig. 1b), the high-high hotspot cluster in Kashmir remained almost similar to the previous scenario of high CVD prevalence concurrent with a high percentage of women without any formal education. Other districts with the same situation are Bilaspur, Solan, Ludhiana, Bulandshahr,

Kaithal, Karnal, Gonda, Maharajganj, Guntur, Sri Potti, Sriramulu, Nellore, Thiruvallur, Kancheepuram, Ramanathapuram, Virudhanagar, Thottukud, Madurai, Sivaganga, Pudukottai, Thanjavur.

The 106 districts with low CVD prevalence as well as low prevalence of women with less than 10 years of schooling consisted of districts like Kachchh, Banas Kantha, Sirohi, Udaipur, Uttarkashi, Rudraprayag, Chamoli, Tehri Garhwal, Raipur, Bilaspur, Mahasamund, Hingoli, majorly constituted by southern Rajasthan, northern coastal Gujarat, almost the whole of Uttarakhand, southern Uttar Pradesh, southern Bihar, Chhattisgarh, central and eastern Maharashtra.

The high CVD prevalence along with low prevalence of women with less than 10 years of schooling consists of districts like Pune, Nagpur, Neemuch.

Low CVD prevalence along with high prevalence of women with less than 10 years of schooling is found in 19 districts, a few of which are Erode, Salem, Idukki, Hanumangarh.

On the basis of classification by caste (Fig. 1c), bivariate results between CVD among women aged between 35-49 years and women from OBCs show high prevalence of both the phenomena in 63 districts, a few of which are Kupwara, Punch, Srinagar, Ludhiana, Ambala, Baghpat, Sonipat, Panipat, Nellore, Vellore, Ramanathapuram, Virudhanagar, Madurai, Sivaganga, from Jammu, and a few districts of Punjab, Haryana, and coastal Tamil Nadu down south.

Low prevalence of CVD in 35-49 yr women as well as hailing from OBCs has been observed in Nashik, Ahmednagar, Osmanabad, Bijapur, Jaipur, Sirohi, Udaipur, Rajsamand, and a majority of the districts of Chhattisgarh.

There are 15 districts apiece with low prevalence of CVD-affected women but high percentage of women from OBCs and vice versa.

Bivariate results of CVD among women aged between 35-49 years and women from Scheduled Caste households (Fig. 1d) show high-high clusters in 56 districts like West District in Sikkim, West Garo Hills, Tawang, Bghpat, Sonipat, Jind, Kancheepuram, Tiruvallur.

Low-low clusters are mainly found in the central Indian districts like Bid, Parbhani, Hingoli, Rajnandgaon, Durg, Kabeerdham.

There are a few districts like Neemuch, Dumka, East Khasi Hills with High CVD prevalence, but low percentage of women from Scheduled Caste households.

Women aged between 35-49 years with CVD and hailing from rural areas (Fig. 1e) have high-high clusters in 70 districts like Ferozpur, Jalandhar, Ludhiana, Fatehgarh, Kaithal, Guntur, Nellore, and most parts of Tamil Nadu.

Low-low clusters have been found in 112 districts like Koppal, Bellary, Bagalkot, Gadag, Bijapur etc. High CVD prevalence but low rural residential percentage has been found in 12 districts like Dumka, Jamui,

Bokaro, Nagpur, Kalahandi.

In case of bivariate results of CVDs among women and women who smoke (Fig. 1f), high high clusters have been found in Prakasam, Mahbubnagar, Krishna, Guntur, Thiruvarur, Perambalu. 39 districts with low-low clusters are found, a few of which are Malappuram, Ramanagara, Gadag, Solapur, Bid, Latur, Parbhani.

Khardha, Jagatsinghpur, Nagpur, Bhadrak are a few of the 26 districts found with high CVD prevalence but low prevalence of women who smoke.

There has been a shift in the location of cold spots in this case, as compared to the previous LISA results with a majority of them shifting predominantly into the northern Deccan area, constituted by districts from southern Madhya Pradesh like Betul, west Nimar, Balaghat, Mandla, western Maharashtra like Amaravati, Aurangabad, Buldana, Solapur, Bid, Latur, Gadchiroli, Chandrapur, Chhattisgarh like Durg, Rajnandgaon, Kabeerdham, Dhamtani, Raipur, Bastar.

Discussion

The association between cardio-vascular diseases and quite a few socio-demographic variables have been analysed for India. Socioeconomic inequalities are well-recognised in chronic disease awareness and management [62].

The present study shows 17.4 percent of women in the age-group of 35-49 years to be suffering from cardio-vascular diseases (hypertension or heart diseases). Differences in prevalence have been observed among rural informal areas and urban formal areas, with a higher prevalence of CVDs observed among urban women as compared to their rural counterparts. These findings are in line with extant literature for India where ischemic heart diseases (IHD) prevalence increased from 2% to $\approx 14\%$ from 1960 to 2013 in urban India [63,64] while it more than quadrupled from 1.7% to 7.4% between 1970 and 2013 in rural India [65,66].

Education-related disparities have been observed in awareness, treatment, and control of CVDs, as found in the current study, where women educated for more than 10 years have a higher CVD prevalence as compared to their lesser educated counterparts; this may be explained by lifestyle and behavioural factors like stress, physical inactivity, and a sedentary lifestyle [66].

A higher CVD prevalence has also been found among the women hailing from the richest wealth quintile as compared to the other wealth quintiles, corroborating a positive social gradient still prevalent in the CVD prevalence in India among women in later reproductive ages (35-49 years).

The statistical analysis made important contributions in evaluating the important components of CVD prevalence using advanced geo-spatial techniques. The disparities at the district-level are conspicuous and this kind of analysis can help in prioritising those areas which require attention. Hotspots and cold spots of CVD prevalence have been found across the country. The coastal districts from south-eastern

India like parts of Tamil Nadu, Kerala, and Andhra Pradesh, parts of Haryana, Punjab, western Uttar Pradesh house a majority of the hotspots. Most of the cold spots exist in Rajasthan, Gujarat, Maharashtra, Madhya Pradesh, basically western and central parts of India.

One of the key findings in the study is regarding the almost constant placement of hotspots and cold spots when CVDs are spatially patterned upon cross-classification, be it by years of schooling, place of residence, smoking habits, or even by background.

The incidence of CVDs has largely been an urban phenomenon, but a recent rise in rural incidences have been reported, too [64]. The current study corroborates these findings with almost 70 districts clustering to form such hotspots. According to the INTERHEART study, standard risk factors like smoking, abnormal lipids, hypertension, diabetes, central obesity, sedentary lifestyle, as well as increasing psychosocial stress, as well as dwindling consumption of fruits and vegetables contribute to more than 90% explanation of acute CHD events in south Asians [67]. Epidemiological studies have also reiterated the fact that all the major risk factors are increasing in India [65]. Tobacco production as well as consumption in various forms has increased substantially. There is significant clustering in northern parts and southern parts of the country when smoking is considered as a factor for assessing spatial patterns, also in agreement with the findings from the second and third rounds of National Family Health Survey conducted in 1998-99 and 2005-06, according to which smoking habits have increased among young subjects (20-35 years) [64].

Target areas coming up in studies of this sort should direct the next course of action to be undertaken in terms of morbidity combat. Region-specific steps are another area of action which can help in addressing the issue. The risk factors, aiming which can control the disease prevalence in one area, might not be effective in combatting an increased CVD prevalence in another area. In some parts, an increased proportion of women living in rural residences has been found to be having a high prevalence of CVDs, too, and clustered significantly in northern districts of states like Haryana, Punjab, Jammu, but there do exist 16 districts with high proportion of women hailing from rural areas with low prevalence of CVDs like Chittoor, Nalgonda, Erode, Namakkal, Idukki. Program-based interventions can be instrumental in creating momentum in CVD control schemes and awareness regarding behavioural changes to delay the onset of symptoms.

Conclusion

The main objective of this study has been to explore the patterns of CVD risk factors among people across districts in India, as well as investigate the association between area-level socioeconomic factors and CVD risk patterns. Hence, the production of fine-grained maps of CVD risk is possible through this approach for use by clinicians and policy makers, enabling geographic targeting of community interventions for CVDs. In case of disease prevalence, individuals living in close proximity to each other tend to have similar socio-demographic characteristics like age, income, access to healthcare facilities, and, hence, similar disease prevalence; thereby giving rise to positive spatial dependence. Throughout

this study, the traditionally used spatial autocorrelation tool Moran's I has been implemented. Ordinary Least Square (OLS) regression has been employed with CVD prevalence as the outcome variable. To consider for spatial dependence, Spatial Autoregressive (SAR) models have been fitted to the data. Diagnostic tests for spatial dependence have also been carried out to identify the best fit model. A larger value of Lagrange's Multiplier (LM), a larger R^2 value, as well as a lower value of Akaike Info Criterion (AIC), indicates better suitability of the model applied.

A higher prevalence of CVDs is found among urban women (21.9 percent), as compared to rural (14.9 percent) women. Non-smoking women have a higher CVD prevalence, as compared to their counterparts who smoke, as do women who do not consume alcohol, as compared to their counterparts who do. A Moran's I score of 0.49 ($p < 0.001$) is highly significant, indicating a strong spatial autocorrelation of the residuals. The spatial error model has been found to be the best-fit model for describing the characteristics affecting CVD. Variables which were significant in the SEM were urban places of residence, hailing from a Scheduled Caste or Tribe or Other Backward Castes, as well as smoking (significant in the SLM but eventually insignificant in the SEM) have all been positively correlated with CVD prevalence in the regions, controlling for selected socio-demographic and behavioural variables. In case of bivariate results of CVDs among women and women who smoke, high high clusters have been found in Prakasam, Mahbubnagar, Krishna, Guntur, Thiruvarur, Perambalu. 39 districts with low-low clusters are found, a few of which are Malappuram, Ramanagara, Gadag, Solapur, Bid, Latur, Parbhani. Khardha, Jagatsinghpur, Nagpur, Bhadrak are a few of the 26 districts found with high CVD prevalence but low prevalence of women who smoke. Most of the hotspots exist in Jammu, Uttarakhand, Punjab, Haryana, coastal Andhra Pradesh, Tamil Nadu, NCT, while the majority of the cold spots for all the bivariate LISA results are clustered in the stretch of area encompassing districts of Gujarat, Rajasthan, southern Uttar Pradesh, Chhattisgarh, eastern Maharashtra, northern Karnataka, and western Odisha.

Studies of this kind highlighting geographical disparities can rightly shift the focus on rural-urban differentials, provincial or district-level inequalities, hence demonstrating the need for targeted action and population-wide interventions to reduce CVD burden as well as associated behavioural risks.

Globalization and urbanisation have been working at the macro-societal level, leading to the developing world increasingly being subjected to risky behaviour like smoking, drinking, low physical activity, as well as unhealthy food habits. Limited access to healthcare facilities, public health education, and prevention plans as compared to their counterparts in the developed world further compound the problem [68,69].

Disparities in CVD prevalence and CVD health point towards a deeper problem. The need of the hour is comprehensive tobacco control policies, smoking cessation programs, increased access to medical facilities, physical activity campaigns via Information Education Communication efforts are all methods to decreasing the CVD risk, all the while targeting the disadvantaged areas of the country. A broader course of policy action relating to social determinants can be a particularly effective way of CVD risk addressal [70,71,72]. Social policy interventions related to health like reduction in inequalities in factors like education, poverty, unemployment, access to health-promoting physical or built-environments are crucial in tackling the long-term effects of CVD inequalities between geographical areas.

Declarations

Authors' Contributions

SKS and AB conceived the idea. JG, SKS, and AB designed the experiment and analysed it, AB interpreted the results and drafted the manuscript. All the authors read and approved the final manuscript.

Funding

This research has not received any specific grant from any funding agency, commercial entity or any not-for-profit organization.

Availability of data and materials

The data is available online for public use at the Demographic Health Survey website.

Ethics approval and consent to participate

This study utilised a secondary data set and it has no identifiable information of the survey participants. This dataset is easily available on the public domain for research purpose, hence, no approval was required from any institutional review board as there was no question of human subject protection which arose from this study.

Consent for publication

This study is based on a secondary dataset. The conducting organisation of the survey has already taken care of consent of the persons being surveyed during the interviews with women, which is easily accessible from the source [59].

Competing interests

The authors declared that they have no competing interests.

Acknowledgements

The authors are thankful to all the co-ordinators of National Family Health Survey (NFHS-4), 2015–16 for their relentless support in providing community-based estimates at National and Sub-national levels.

Author details

1. International Institute for Population Sciences, Govandi Station Road, Deonar, Mumbai 400088, India.
2. Department of Mathematical Demography & Statistics, International Institute for Population Sciences, Govandi Station Road, Deonar, Mumbai 400088, India.

References

1. Graham AJ, Atkinson PM, Danson FM. Spatial analysis for epidemiology.
2. Rezaeian M. Application of geographical sciences and technologies to investigate health problems in the Eastern Mediterranean Region. *EMHJ-Eastern Mediterranean Health Journal*, 15 (6), 1564-1569, 2009. 2009.
3. Anno S, Imaoka K, Tadono T, Igarashi T, Sivaganesh S, Kannathasan S, Kumaran V, Surendran SN. Space-time clustering characteristics of dengue based on ecological, socio-economic and demographic factors in northern Sri Lanka. *Geospatial health*. 2015 Nov 26.
4. Bergquist R. Climate and the distribution of vector-borne diseases: what's in store?. *Geospatial health*. 2017 May 8;12(1).
5. Oliveira A, Cabral AJ, Mendes JM, Martins MR, Cabral P. Spatiotemporal analysis of the relationship between socioeconomic factors and stroke in the Portuguese mainland population under 65 years old. *Geospatial health*. 2015 Nov 4.
6. Park SY, Kwak JM, Seo EW, Lee KS. Spatial analysis of the regional variation of hypertensive disease mortality and its socio-economic correlates in South Korea. *Geospatial health*. 2016 May 31.
7. Bascuñán MM, Quezada CR. Geographically weighted regression for modelling the accessibility to the public hospital network in Concepción Metropolitan Area, Chile. *Geospatial health*. 2016 Nov 22.
8. Griffith DA. Spatial autocorrelation. A Primer (Washington, DC, Association of American Geographers). 1987.
9. Rezaeian M, Dunn G, St Leger S, Appleby L. Geographical epidemiology, spatial analysis and geographical information systems: a multidisciplinary glossary. *Journal of Epidemiology & Community Health*. 2007 Feb 1;61(2):98-102.
10. Lorant V, Thomas I, Deliege D, Tonglet R. Deprivation and mortality: the implications of spatial autocorrelation for health resources allocation. *Social science & medicine*. 2001 Dec 1;53(12):1711-9.
11. Sofianopoulou E, Pless-Mulloli T, Rushton S. Use of spatial autocorrelation to investigate clustering of health deprivation. *Epidemiology*. 2006 Nov 1;17(6):S95.
12. Aamodt G, Samuelsen SO, Skrondal A. A simulation study of three methods for detecting disease clusters. *International journal of health geographics*. 2006 Dec;5(1):1-1.
13. Jacquez GM. Spatial cluster analysis. *The handbook of geographic information science*. 2008 Apr 15;395(416).
14. Moran PA. The interpretation of statistical maps. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1948 Jul;10(2):243-51.
15. Anselin L. Local indicators of spatial association—LISA. *Geographical analysis*. 1995 Apr;27(2):93-115.
16. Bailey TC, Gatrell AC. *Interactive spatial data analysis*. Essex: Longman Scientific & Technical; 1995 Oct.

17. Ford MM, Highfield LD. Exploring the spatial association between social deprivation and cardiovascular disease mortality at the neighborhood level. *PloS one*. 2016 Jan 5;11(1):e0146085.
18. de Andrade L, Zanini V, Batilana AP, de Carvalho EC, Pietrobon R, Nihei OK, de Barros Carvalho MD. Regional disparities in mortality after ischemic heart disease in a Brazilian state from 2006 to 2010. *PloS one*. 2013 Mar 19;8(3):e59363.
19. Martinez AN, Mobley LR, Lorvick J, Novak SP, Lopez AM, Kral AH. Spatial analysis of HIV positive injection drug users in San Francisco, 1987 to 2005. *International journal of environmental research and public health*. 2014 Apr;11(4):3937-55.
20. Hastie TJ. *Generalized additive models*. Routledge; 2017 Nov 1.
21. Levine N, Kim KE, Nitz LH. Spatial analysis of Honolulu motor vehicle crashes: II. Zonal generators. *Accident Analysis & Prevention*. 1995 Oct 1;27(5):675-85.
22. Australian Institute of Health. *Australia's health 2012: the thirteenth biennial health report of the Australian Institute of Health and Welfare*. AIHW; 2012.
23. Mathers CD, Vos ET, Stevenson CE, Begg SJ. The burden of disease and injury in Australia. *Bulletin of the World Health Organization*. 2001;79:1076-84.
24. Australian Institute of Health and Welfare 2011. *Cardiovascular disease: Australian facts 2011*. Cardiovascular disease series. Cat. no. CVD 53. Canberra: AIHW.
25. Kannel WB, McGee D, Gordon T. A general cardiovascular risk profile: the Framingham Study. *The American journal of cardiology*. 1976 Jul 1;38(1):46-51.
26. Nelson MR, Doust JA. Primary prevention of cardiovascular disease: new guidelines, technologies and therapies. In reply. *The Medical journal of Australia*. 2014 Feb 1;200(3):148-.
27. Bagheri N, Gilmour B, McRae I, Konings P, Dawda P, Del Fante P, Van Weel C. Peer Reviewed: Community Cardiovascular Disease Risk From Cross-Sectional General Practice Clinical Data: A Spatial Analysis. *Preventing chronic disease*. 2015;12.
28. Noble D, Smith D, Mathur R, Robson J, Greenhalgh T. Feasibility study of geospatial mapping of chronic disease risk to inform public health commissioning. *BMJ open*. 2012 Jan 1;2(1).
29. Backholer K, Peters SA, Bots SH, Peeters A, Huxley RR, Woodward M. Sex differences in the relationship between socioeconomic status and cardiovascular disease: a systematic review and meta-analysis. *J Epidemiol Community Health*. 2017 Jun 1;71(6):550-7.
30. Kihal-Talantikite W, Weber C, Pedrono G, Segala C, Arveiler D, Sabel CE, Deguen S, Bard D. Developing a data-driven spatial approach to assessment of neighbourhood influences on the spatial distribution of myocardial infarction. *International journal of health geographics*. 2017 Dec;16(1):1-8.
31. Pathak EB, Reader S, Tanner JP, Casper ML. Spatial clustering of non-transported cardiac decedents: the results of a point pattern analysis and an inquiry into social environmental correlates. *International journal of health geographics*. 2011 Dec;10(1):1-1.
32. Boruzs K, Juhász A, Nagy C, Ádány R, Bíró Relationship between statin utilization and socioeconomic deprivation in Hungary. *Frontiers in pharmacology*. 2016 Mar 24;7:66.

33. Pedigo A, Aldrich T. Neighborhood disparities in stroke and myocardial infarction mortality: a GIS and spatial scan statistics approach. *BMC public health*. 2011 Dec;11(1):1-3.
34. Hernández-Vásquez A, Bendezú-Quispe G, Díaz-Seijas D, Santero M, Minckas N, Azañedo D, Antiporta DA. Spatial analysis of childhood obesity and overweight in Peru, 2014. *Revista peruana de medicina experimental y salud pública* 2016 Jul 1;33(3):489-97.
35. Odoi A, Busingye D. Neighborhood geographic disparities in heart attack and stroke mortality: comparison of global and local modeling approaches. *Spatial and spatio-temporal epidemiology*. 2014 Oct 1;11:109-23.
36. Cheng CL, Chen YC, Liu TM, Yang YH. Using spatial analysis to demonstrate the heterogeneity of the cardiovascular drug-prescribing pattern in Taiwan. *BMC public health*. 2011 Dec;11(1):1-9.
37. Wu PC, Lin CY, Lung SC, Guo HR, Chou CH, Su HJ. Cardiovascular mortality during heat and cold events: determinants of regional vulnerability in Taiwan. *Occupational and environmental medicine*. 2011 Jul 1;68(7):525-30.
38. Castillo-Carniglia Á, Kaufman JS, Pino P. Geographical distribution of alcohol-attributable mortality in Chile: a Bayesian spatial analysis. *Addictive behaviors*. 2015 Mar 1;42:207-15.
39. Li C, Ford ES, Mokdad AH, Balluz LS, Brown DW, Giles WH. Clustering of cardiovascular disease risk factors and health-related quality of life among US adults. *Value in Health*. 2008 Jul;11(4):689-99.
40. Gu D, Gupta A, Muntner P, Hu S, Duan X, Chen J, Reynolds RF, Whelton PK, He J. Prevalence of cardiovascular disease risk factor clustering among the adult population of China: results from the International Collaborative Study of Cardiovascular Disease in Asia (InterAsia). *Circulation*. 2005 Aug 2;112(5):658-65.
41. Peters SA, Wang X, Lam TH, Kim HC, Ho S, Ninomiya T, Knuiman M, Vaartjes I, Bots ML, Woodward M. Clustering of risk factors and the risk of incident cardiovascular disease in Asian and Caucasian populations: results from the Asia Pacific Cohort Studies Collaboration. *BMJ open*. 2018 Mar 1;8(3):e019335.
42. Cressie N. *Statistics for spatial data* New York.
43. Tiefelsdorf M. Misspecifications in interaction model distance decay relations: A spatial structure effect. *Journal of Geographical Systems*. 2003 May 1;5(1):25-50.
44. Anselin L. *Spatial econometrics: methods and models*. Springer Science & Business Media; 2013 Mar 9.
45. Fotheringham AS, Brunsdon C, Charlton M. *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons; 2003 Feb 21.
46. Getis, A. (2008). A history of the concept of spatial autocorrelation: A geographer's perspective. *Geographical analysis*, 40(3), 297-309.
47. Lin G, Zhang T. Loglinear residual tests of Moran's I autocorrelation and their applications to Kentucky breast cancer data. *Geographical Analysis*. 2007 Jul;39(3):293-310.

48. Li H, Calder CA, Cressie N. Beyond Moran's I: testing for spatial dependence based on the spatial autoregressive model. *Geographical Analysis*. 2007 Oct;39(4):357-75.
49. Geary RC. The contiguity ratio and statistical mapping. *The incorporated statistician*. 1954 Nov 1;5(3):115-46.
50. Goodchild MF. *Spatial autocorrelation*. Geo Books; 1986.
51. Griffith DA. *Spatial autocorrelation. A Primer* (Washington, DC, Association of American Geographers). 1987.
52. Odland J. *Spatial autocorrelation*. SAGE Publications, Incorporated; 1988.
53. Getis A, Ord JK. The analysis of spatial association by use of distance statistics. In *Perspectives on spatial data analysis 2010* (pp. 127-145). Springer, Berlin, Heidelberg.
54. Anselin L. The Moran scatterplot as an ESDA tool to assess local instability in spatial. *Spatial Analytical*. 2019;111.
55. Haining R. *Spatial data analysis in the social and environmental sciences*. Cambridge University Press; 1993 Aug 26.
56. Lee SI. Developing a bivariate spatial association measure: an integration of Pearson's r and Moran's I. *Journal of geographical systems*. 2001 Dec 1;3(4):369-85.
57. Getis A, Aldstadt J. Constructing the spatial weights matrix using a local statistic. *Geographical analysis*. 2004 Jul;36(2):90-104.
58. Bertazzon S, Olson S, Knudtson M. A spatial analysis of the demographic and socio-economic variables associated with cardiovascular disease in Calgary (Canada). *Applied Spatial Analysis and Policy*. 2010 Mar;3(1):1-23.
59. IIPS. *National Family Health Survey (NFHS-4), 2015–16*. International Institute for Population Sciences (IIPS), Mumbai, India. 2017.
60. Darmofal D. *Spatial analysis for the social sciences*. Cambridge University Press; 2015 Oct 29.
61. Anselin L, Syabri I, Kho Y. GeoDa: an introduction to spatial data analysis. In *Handbook of applied spatial analysis 2010* (pp. 73-89). Springer, Berlin, Heidelberg.
62. Di Cesare M, Khang YH, Asaria P, Blakely T, Cowan MJ, Farzadfar F, Guerrero R, Ikeda N, Kyobutungi C, Msyamboza KP, Oum S. Inequalities in non-communicable diseases and effective responses. *The Lancet*. 2013 Feb 16;381(9866):585-97.
63. Praveen PA, Roy A, Prabhakaran D. Cardiovascular disease risk factors: a childhood perspective. *The Indian Journal of Pediatrics*. 2013 Mar;80(1):3-12.
64. Gupta R. Recent trends in coronary heart disease epidemiology in India. *Indian heart journal*. 2008 Mar 1;60(2 Suppl B):B4-18.
65. Gupta R, Joshi P, Mohan V, Reddy KS, Yusuf S. Epidemiology and causation of coronary heart disease and stroke in India. *Heart*. 2008 Jan 1;94(1):16-26.
66. Gupta R, Sharma KK, Gupta BK, Gupta A, Gupta RR, Deedwania PC. Educational status-related disparities in awareness, treatment and control of cardiovascular risk factors in India. *Heart Asia*.

2015 Jan 1;7(1):1-6.

67. Joshi P, Islam S, Pais P, Reddy S, Dorairaj P, Kazmi K, Pandey MR, Haque S, Mendis S, Rangarajan S, Yusuf S. Risk factors for early myocardial infarction in South Asians compared with individuals in other countries. *Jama*. 2007 Jan 17;297(3):286-94.
68. Mendis S, Puska P, Norrving B, World Health Organization. *Global atlas on cardiovascular disease prevention and control*. World Health Organization; 2011.
69. Laslett LJ, Alagona P, Clark BA, Drozda JP, Saldivar F, Wilson SR, Poe C, Hart M. The worldwide environment of cardiovascular disease: prevalence, diagnosis, therapy, and policy issues: a report from the American College of Cardiology. *Journal of the American College of Cardiology*. 2012 Dec 25;60(25S):S1-49.
70. Singh GK, Siahpush M. Increasing inequalities in all-cause and cardiovascular mortality among US adults aged 25–64 years by area socioeconomic status, 1969–1998. *International Journal of Epidemiology*. 2002 Jun 1;31(3):600-13.
71. Department of Health. *Tackling Health Inequalities: 10 Years On*. London, UK: Health Inequalities Unit, Department of Health; 2009.
72. Shaw M, Dorling D, Gordon D, Smith GD. *The widening gap: health inequalities and policy in Britain*. Policy Press; 1999 Nov 20.

Figures

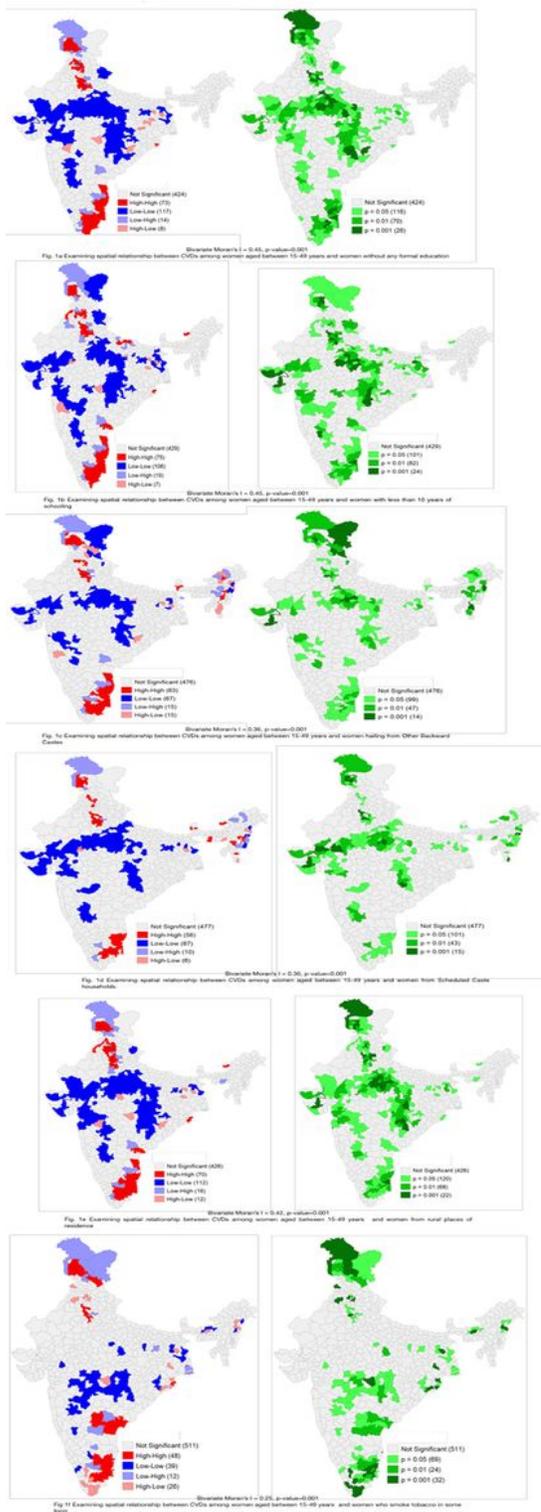


Figure 1

Bivariate LISA Maps and significance maps depicting spatial clustering and spatial outliers of CVDs across 640 districts of India, 2015–16 Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.