

# An Epidemiological Study of Cervical and Breast Screening in India: District-Level Analysis

Monica Monica

International Institute for Population Sciences

Raman Mishra (✉ [raman.mishra10@gmail.com](mailto:raman.mishra10@gmail.com))

International Institute for Population Sciences <https://orcid.org/0000-0001-8257-6019>

---

## Research article

**Keywords:** Cervical Screening, Breast Screening, Spatial Analysis, India

**Posted Date:** April 24th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-24243/v1>

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

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on October 7th, 2020. See the published version at <https://doi.org/10.1186/s12905-020-01083-6>.

# Abstract

**Background:** Breast cancer and cervical cancer, the most common forms of cancer in women worldwide, are on a fast and steady rise, accounting for more deaths in women than any other cancer in the developing world. Cancer screening tests are an important tool to combat cancer-related morbidity and mortality. The World Health Assembly, in its agenda for cancer prevention and control in the context of an integrated approach, urges governments and the WHO to accelerate action to achieve Goal 3.4 of the Sustainable Development Goals (SDG 3.4) in order to reduce premature mortality from non-communicable diseases including cancer by one-third by 2030 (WHO, 2016).

**Methods:** Until recently, there was no evidence pertaining to screening for cervical and breast cancers at the national level. The National Family Health Survey, 2015-16 (NFHS-4), for the first time in the NFHS series, collected information on examination of the breast and the cervix from over 699,000 women in the age group 15-49 years. For the present study, the data was aggregated for all 640 districts in India. Moran's Index was calculated to check for spatial autocorrelation. Univariate Local Indicators of Spatial Association (LISA) maps were plotted to look for spatial dependence associated with the uptake of screening practices. The spatial error model was employed to check for spatial magnitude and direction.

**Results:** The factors associated with uptake of cervical and breast screening at the district level were: belonging to a general caste, residing in rural areas, being currently married, and being well-off economically. The study provides spatial inference by showing geographical disparities in screening, this study highlights the importance of ensuring a region-specific and organ-specific approach toward control and prevention of cancer.

**Conclusions:** As part of the National Health Mission, the Indian government for the first time has launched population-based prevention, screening, and control programs for cancers of the cervix and the breast. The spatial analysis presented here may add essential information in policy-making. By showing geographical disparities in screening, this study highlights the importance of ensuring a region-specific and organ specific approach toward control and prevention of cancer.

## Background

Cancer is increasingly being recognized as a major cause of mortality and morbidity, with approximately 18.1 million new cases reported in 2018 (WHO, 2018). The World Health Organization (WHO) projects that the number of global cancer deaths will rise by 45% between 2008 and 2030 (WHO, 2019). The rising burden of the mortality from cancer is likely to be fivefold greater in the low-income countries compared to established market economies (Rastogi et al. 2004). The economic burden of cancer is significant and rising. In 2010, the total annual economic cost of cancer was estimated at approximately US\$ 1.6 trillion (Stewart & Wild, 2014), thus threatening health budgets at all income levels and causing financial distress for individuals and families.

Breast cancer and cervical cancer, the most common forms of cancer in women worldwide, too are on a fast and steady rise, accounting for more deaths in women than any other cancer in the developing world (WHO, 2018). Statistics suggest that about 527,624 and 1,671,149 new cases of cervical and breast cancers are added every year. To this, India contributes about 122,844 cervical cancer and 144,937 breast cancer cases every year (Ferlay et al. 2015). India accounts for nearly one-third of the global cervical cancer deaths, with women facing a 2.5% cumulative lifetime risk and 1.4% cumulative death risk from cervical cancer (WHO/ICO, 2007). Earlier, cervical cancer was the most common cancer throughout the nation, but now the incidence of breast cancer has surpassed it and is the leading cause of death (Karthigeyan, 2012). A vital observation here is that breast and cervical cancers are curable if diagnosed at an early stage. These cancers are preventable with access to high-quality care, periodic screening tests, and regular follow-up (WHO, 2018).

The World Health Assembly (WHA 70.12), in its agenda for cancer prevention and control in the context of an integrated approach, urges governments and the WHO to accelerate action to achieve Goal 3.4 of the Sustainable Development Goals (SDG 3.4) in order to reduce premature mortality from non-communicable diseases including cancer by one-third by 2030 (WHO, 2016). The strategies to reduce the high burden of cervical and breast cancers include risk factor intervention, vaccination, screening, and early diagnosis (Viens, et al. 2017). Effective screening is the first step toward reducing the burden of cervical and breast cancers. Screening is defined as “the systematic application of a test or an inquiry to identify individuals at sufficient risk of a specific disorder to warrant further investigation or direct preventive action among persons who have not sought medical attention on account of symptoms of that disorder (Britain, 1998).” On the other hand, screening uptake refers to the proportion of persons eligible to be screened within a population who have been both invited for screening and have received an adequate screening during a specified period (Beining, 2012). Experience from the developed world shows that effective population-based screening programmes can easily reduce the incidence of cervical and breast cancers. Mortality rates from cervical and breast cancers can also be reduced by such interventions (Hermann et al. 2018; Kitchener et al. 2006). Despite the clear and proven benefits of population-based screening programs, screening for cervical and breast cancers in low income countries, including India, remains a challenge.

Until recently, there was no evidence pertaining to screening for cervical and breast cancers in India. The National Family Health Survey, 2015-16 (NFHS-4), for the first time in the NFHS series, collected information on examination of the breast and cervix from over 699,000 women ages 15–49 (IIPS and ICF 2017). The availability of such information in NFHS-4 provided us with a great opportunity to analyze the levels and patterns in the screening for cervical and breast cancers in India at the state and district levels.

Some past studies, mostly conducted in developed country settings, have identified a number of socio-economic, demographic, bio-medical, and residence-related characteristics that are associated with the screening of the cervix and breast. The likelihood of a woman receiving a Pap test, or a clinical breast examination, depends on many aspects such as age, marital status, income level, education, and health status. Women with higher education, higher incomes, and greater insurance coverage are more likely to

undergo cervical and breast cancer screening services (Lin, 2008). Employed females are more inclined to go for screening because of their higher opportunity cost, higher incomes, and ability to afford out-of-pocket expenditure (Wu, 2003). On the other hand, rural women are less likely than urban women to go for cervical and breast screening (Coughlin et al. 2008; McLafferty et al. 2011; Beining 2012). Studies of breast and cervical screening show that women with greater access to health care, such as those with health insurance, are opting to have screening tests (Selvin and Brett 2003; Wu, 2003). The risk of infection with HPV and also the risk of cervical cancer depends on the number of sexual partners, age at first intercourse, and sexual behavior of the woman's male partner/s (Bosch et al. 1997). Additional risk indicators for cervical cancer are number of live births, long-term use of oral contraceptives, and cigarette smoking (Franco et al. 2001). Risk factors, other than socio-economic and demographic characteristics, accountable for breast cancer are alcohol, obesity, longer use of oral contraceptives, early onset of menstrual periods, etc. (CDC, 2018). Studies also suggest that health policies and quality of the health care system influence cervical and breast cancer screening behaviors (Akinyemiju et al. 2015; Yang et al. 2011).

## Rationale For Spatial Dependence

Even though there has been an increase in the prevalence of cancer among the female population, most of the research on cancers in females is concentrated only on the incidence and mortality rates of cervical and breast cancers. A review of cancer screening-related literature in India reveals that the spatial perspective of cancer screening has not been explored yet. The present study attempts to address some of these research gaps and aims to develop a comprehensive assessment of the geographical distribution of cervical and breast screening at the district level so as to capture the precise picture of the screening practices.

It is crucial to note that any aggregation of socioeconomic, demographic or health variables over a geographic space tends to manifest a spatial pattern or spatial clustering. In such a case, spatial autocorrelation creates a problem for statistical testing as the autocorrelated data violates the assumptions of classical statistics, one of them being the independence of the observations (Legendre, 1993). Such regression analyses, which ignore the spatial correlations, lead to incorrect inference of the estimated regression coefficients by narrowing the confidence intervals (Huque et al., 2014). This limitation can, however, be overcome through the use of geospatial models (Fotheringham *et al.* 2002; Legendre et al. 2002).

## Data Source

This study used data from the fourth round of the National Family Health Survey (NFHS, 2015-16) conducted in all the 29 states and 6 union territories of India, divided administratively into 640 districts. The district-level data for all the states and union territories was aggregated for the present study.

## Methods

The outcome variables used in the analysis were cervical and breast screening. The data of women undergoing cervical and breast screening in the age group 15-49 was aggregated at the district level. The study assessed the variations in cervical and breast screening through a set of independent factors.

We assessed the following predictors influencing the uptake of cervical screening: 1) covered by insurance, 2) having multiple sexual partners, 3) consuming tobacco in any form, 4) using oral contraceptives, and 5) having parity greater than three. The determinants influencing the uptake of breast screening included: 1) being obese, 2) using oral contraceptives, 3) consuming tobacco in any form, 4) covered by insurance, and 5) consuming alcohol.

Other common socio-demographic factors included in the analysis were: status of literacy (literate), marital status (currently married), religion (Hindu), caste (general), area of residence (rural), and economic status (rich). All the variables were aggregated from the individual level to the district level.

## Analytic Strategy

1. First, the ordinary least squares (OLS) model was estimated to explore the relationship between cervical and breast screening and their independent variables respectively.
2. Second, tests for the spatial dependence were performed.
3. Third, spatial regression model was employed to study the spatial effect on screening.

## Estimation Of Spatial Association

We used R software to generate descriptive maps of cervical and breast screening across the districts of India. We then generated the spatial weights for the calculation of the spatial autocorrelation statistics. Contiguity-based spatial weights were used, as our objective was to understand the spatial interdependence between the dependent variable and a set of independent variables in the neighboring regions (districts). Within the spatially contiguous weights, we chose queen's weight, which works on the principle that at least one point on the boundary of a polygon is within the snap distance of at least a point of its neighbor. Finally, we used geo-spatial techniques, such as Moran's I statistics, Univariate LISA, and geospatial regression, to address the research questions.

## Global And Local Spatial Autocorrelation

Moran's I is the measure of global spatial autocorrelation. The magnitude of Moran's I was estimated by using the "moran.test" function. A significance level of P-value<0.05 was used to assess the spatial autocorrelation. The main idea behind spatially autocorrelated data is that values are not independent of space. This concept is based on the first law of geography proposed by Waldo Tobler, according to which,

“Everything is related to everything else, but near things are more related than distant things.”

(see Formula 1 in the Supplementary Files)

Where  $z_i$  is the deviation of an attribute for feature  $i$  from its mean ( $x_i - \bar{X}$ ),  $w_{ij}$  is the spatial weight between feature  $i$  and  $j$ ,  $n$  is equal to the total number of features, and  $S_0$  is the aggregate of all the spatial weights. The Moran's  $I$  score ranges from -1 (dispersed) to 1 (clustered). A value of 0, or very close to 0, refers to random distributions. Positive autocorrelation suggests that points with similar attribute values are closely distributed in space, whereas negative spatial autocorrelation suggests that closely associated points are more dissimilar in spatial terms. By applying the Monte Carlo simulation computational technique, Moran's  $I$  was permuted 999 times to determine the significance using multiplication.

The global spatial autocorrelation does not reveal existence of regional spatial patterns. Therefore, in order to visualize spatial clustering, Local Indicators of Spatial Autocorrelation (LISA) maps were created using local Moran's index calculations (Anselin, 1995). LISA statistic was calculated for each observation and cluster, with the significance level at  $P < 0.05$ . The LISA statistic gives an indication of the extent of significant spatial clustering of similar or dissimilar values around a spatial feature. It is provided by the following formula: (see Formula 2 in the Supplementary Files)

The parameters for the LISA statistics are the same as those for Moran's  $I$ . In fact, the sum of the LISA statistics for all spatial features is proportional to the global Moran's  $I$ . A positive  $I_i$  value indicates spatial clustering of similar values around a spatial characteristic, whereas negative values indicate a clustering of dissimilar values around a spatial feature.

Univariate Local Indicators of Spatial Association (LISA) measures the correlation of neighborhood values around a specific spatial location. It determines the extent of spatial non-stationarity and clustering present in the data. It is given by: (see Formula 3 in the Supplementary Files)

Thus, the data set and the shapefiles were imported to R studio to calculate Moran's  $I$  and generate detailed Local Indicators of Spatial Association (LISA) maps to study the spatial variations and conduct the spatial analysis. Then, the spatial error model was used to scrutinize for the presence of a spatial relationship, which shows that a value observed in one location depends on the values found at the nearby sites, indicating a spatial dependence. Spatial data may show spatial dependence in the variables and error terms. When spatial dependence is present in the error term, a spatial autoregressive specification for this dependence is typically assumed.

Spatial error model—Incorporates spatial effects through error term

$$Y = X\beta + \epsilon$$

$$\epsilon = \lambda W \epsilon + \xi$$

$\epsilon$  is the vector of error terms, spatially weighted using the weight matrix ( $W$ )

$\lambda$  is the spatial error coefficient

$\xi$  is a vector of uncorrelated error terms

If there is no spatial correlation between the errors, then  $\lambda = 0$ .

The spatial error model tells us only that there is an unexplained spatial structure to the residuals, not what caused them. It may offer better estimates of the model parameters and their statistical significance, but it does not presuppose any particular spatial process generating the patterns in the screening values. A different model that explicitly tests for whether the screening at a point is functionally dependent on the values of neighboring points is the spatially lagged model (Harris, 2013). It is given by: (see Formula 4 in the Supplementary Files)

where  $\mathbf{y}$  is the endogenous variable,  $\mathbf{X}$  is a matrix of exogenous variables, and  $\mathbf{W}$  is the spatial weights matrix.

## Results

### Geographical Distribution of Cervical and Breast Screening

Figures 1 and 2 display the prevalence of cervical and breast screening across the districts of India. According to the NFHS report, 22% of women have undergone a cervical examination, whereas the corresponding figure for breast examination is 10%. The pattern of cervical screening indicates that the southern region, Kerala particularly, has a major contribution followed by districts from Maharashtra. The distribution curve of the prevalence of cervical screening and the districts shows that the majority of the districts fall in the range of 10–20%. A majority of the districts in Kerala have a high uptake of breast cancer screening, whereas North Goa has the maximum share. The distribution curve exhibits that the majority of the districts have fall in the range of 0–10%.

### Descriptive Statistics

Table 1 presents the summary statistics of the variables used in the analysis. The average values of cervical screening and breast screening across the districts were 21.96% and 9.66% respectively. Little variation was seen in the average values of women having multiple partners, those having parity more than 3, and those consuming alcohol (SD = 5%, 8%, 6%). Hindu women and those who were literate, currently married, and residing in rural areas had mean proportion values above 70% across the districts of India.

Table 1  
Summary Statistics of Covariates N = 640

Characteristics	Mean	Median	Standard Deviation	Characteristics	Mean	Median	Standard Deviation
Cervical Screening	21.96	18.47	14.62	Rich	38.71	33.98	24.26
Breast Screening	9.66	6.71	8.47	Oral Contraception	14.13	9.87	12.26
Literate	72.39	73.85	14.18	Tobacco	7.13	3.66	10.78
Currently Married	71.96	72.78	5.54	Insurance	18.96	10.99	20.13
Hindu	74.58	85.33	27.69	Multiple Partners	4.20	2.81	4.75
General Caste	21.77	17.48	18.29	Parity > 3	13.14	12.37	7.58
Rural	71.57	77.98	21.66	Alcohol	2.59	0.28	6.43

Table 2 illustrates Moran's I values for the dependent and independent variables incorporated in the study. Moran's I values for cervical and breast screening were 0.61 and 0.55 respectively, indicating a high spatial autocorrelation across the districts of India. Moran's I for the independent variables ranged between 0.81 (for districts with percentage of tobacco consumption) and 0.41 (for districts with percentage of women having multiple partners).

Table 2  
Moran's I for covariates

Characteristics	Moran's I	Characteristics	Moran's I
Cervical Screening	0.610	Rich	0.702
Breast Screening	0.555	Oral Contraception	0.765
Literacy	0.690	Tobacco	0.813
Currently Married	0.570	Insurance	0.735
Hindu	0.749	Alcohol	0.580
General caste	0.553	Parity > 3	0.753
Rural	0.418	Multiple Partners	0.413
Note: above values are significant at p value < 0.01			

## Univariate LISA Maps

To begin with, we plotted the values of the local I statistics (Fig. 3a & Fig. 4a). This was followed by plotting the significant local clusters of cervical and breast screening respectively (Fig. 3b & Fig. 4b). A highly dense clustering of cervical screening could be seen in the districts of Kerala, Maharashtra, Assam, Punjab, Jammu and Kashmir, and West Bengal (Fig. 3b). A few clusters were also observed in the districts of Madhya Pradesh and Uttar Pradesh. As seen in Fig. 4b, clustering of breast screening could be found in the districts of Kerala, Tamil Nadu, Karnataka, Maharashtra, Punjab, and Jammu and Kashmir.

## Ordinary Least Squares and Spatial Regression Models

The OLS regression model was employed to examine the determinants of cervical screening in India. The results are presented in Table 3. Higher socio-economic status was found to be significantly positively associated with cervical screening across the districts, as was having insurance. This was also the case among districts having higher proportion of currently married women, belonged to a general caste, and resided in rural areas in the districts. In contrast, use of oral contraceptives, having multiple partners, having parity above three, and being a Hindu woman were seen to have a significant negative association with cervical screening in the districts.

Table 3  
**Factors influencing Cervical and Breast Screening  
(Ordinary Least Square Model)**

<b>Characteristics</b>	<b>OLS Cervical</b>	<b>OLS Breast</b>
Literate	-0.015	0.130 ***
Currently Married	0.262 ***	0.162 ***
Hindu	-0.096 ***	-0.054 ***
General Caste	0.103 ***	0.054 ***
Rural	0.127 ***	0.023
Rich	0.251 ***	0.134 ***
Oral Contraception	-0.571 ***	-0.245 ***
Tobacco Consumption	-0.006	-0.145 ***
Insurance	0.079 ***	0.034 **
Multiple Partners	-0.394 ***	-
Parity > 3	-0.403 ***	-
Alcohol	-	-0.015
Obese	-	-0.325***
Model Estimates		
Multiple R Squared	0.389	0.342
Adjusted R Squared	0.379	0.331
AIC	4968	4314.4
Note: *** p < 0.001, ** p < 0.01		

Table 3 shows the effect of female literacy on breast screening at the district level, which at 0.130 was found to be positively significant. This implies that education of women and preventive health care seeking behavior goes hand in hand. Districts with higher percentage of currently married women were 0.16 times more likely to go for breast examination. A similar positive association was observed among districts with higher percentage of general caste women, women from economically prosperous districts, districts with higher insured women, and women from rural areas. Use of oral contraceptives (-0.24) and consumption of tobacco (-0.15) had a significant negative association with breast screening. Even adiposity and breast screening shared a negative association (-0.32) with each other at the district level.

Next, while examining the diagnostics of the OLS model, the variance inflation factor (VIF) detected multicollinearity in the regression analysis. The VIF estimates showed that there was very little

multicollinearity among the independent variables (Appendix 1). Moran's I values of 0.422 ( $p < 0.01$ ) and 0.344 ( $p < 0.01$ ) for cervical and breast screening respectively indicated spatial autocorrelation of the residuals. The Akaike Information Criterion (AIC) is a measure of the relative goodness of fit of a statistical model. For a set of models, the preferred model is the one with the minimum AIC value. The coefficient of determination ( $R^2 = 0.379$  and  $0.331$  for cervical and breast screening respectively) indicated that the OLS model was not a best fit.

Table 4 shows that the Lagrange Multiplier (error) and the Lagrange Multiplier (lag) were significant and indicated the presence of a spatial dependence in the cervical and breast screening data. The robust error and the robust lag tests were also significant.

Table 4  
**Spatial Dependence for Cervical and Breast Screening**

Test Statistics	Cervical	Breast
Lagrange Multiplier Error	293.65 ***	194.99 ***
Lagrange Multiplier Lag	266.96 ***	209.79 ***
Robust Lagrange Multiplier Error	37.28 ***	7.00 **
Robust Lagrange Multiplier Lag	10.59 **	21.79 ***
SARMA	304.24 ***	216.79 ***
Note: *** $p < 0.001$ , ** $p < 0.01$		

After establishing the presence of spatial dependence, the spatial lag model was employed. The spatial lag coefficient parameter  $\rho$  measures the average effect on observations by their neighboring observations and thus reflects the spatial dependence inherent in the data. It was found to be statistically significant and had a positive effect for both the models (cervical and breast screening). As a result, the general model fit improved, as indicated by the higher values of Log likelihood. However, the significance of Breusch-Pagan test and the Likelihood Ratio test of spatial lag dependence for both the models revealed that even though the introduction of the spatial lag term led the model fit to improve, it could not eliminate the presence of spatial effects.

Thereafter, we employed the spatial error model. The model provides  $\lambda$  (Lambda) as a coefficient on the spatially correlated errors. The model was highly significant, with a positive effect for both the models (cervical and breast screening). As a result, the model fit improved, as indicated by the higher values of Log Likelihood. The Breusch-Pagan test and the Likelihood Ratio test of spatial error dependence were significant, indicating that the spatial effects in both the models (cervical and breast screening) were still present. However, both the spatial error and the spatial lag models were an improvement on the OLS model. The spatial error model appeared to fit the data better among all as the AIC score was lower and the log likelihood value was greater for the spatial error model employed for both cervical and breast screening. The residuals maps of OLS and spatial error model for cervical (Fig. 5a & 5b) and breast

screening (Fig. 6a & 6b) indicated model improvement. The amount of clustering of the residuals reduced (the residuals appeared to be more randomly distributed), and the Moran's I of the spatial error residuals was reduced from 0.422 to -0.039 for cervical screening and from 0.344 to -0.047 for breast screening. The maps indicated that the problem of spatial autocorrelation amongst the residuals was mainly solved by the spatial error model. Following this, we proceeded to the analysis, considering the coefficients of the spatial error model.

The results shown in Table 5 for the spatial error model demonstrate a statistically significant spatial autocorrelation ( $\lambda = 0.690$ ) for cervical screening. Proportion of women having multiple partners (-0.18) and using oral contraceptives (-0.17) were negatively associated with women taking up cervical screening at the district level. The same was found to be the case among districts with percentage of Hindu women. A significant positive association with cervical screening was found in districts where women are insured (0.09), were currently married, and districts with higher general caste female population. Women who resided in rural districts and those who belonged to higher economic classes also shared a positive association at the district level.

Table 5  
Factors Influencing Cervical Screening (Spatial Error & Spatial Lag Model)

Characteristics cervical	Spatial Error	Spatial Lag
Literate	-0.068	0.009
Currently Married	0.302 ***	0.214 ***
Hindu	-0.083 ***	-0.064 ***
General Caste	0.083 ***	0.073 ***
Rural	0.156 ***	0.087 ***
Rich	0.332 ***	0.159 ***
Oral Contraception	-0.176 *	-0.266 ***
Tobacco Consumption	-0.014	0.039
Multiple Partners	-0.182 *	-0.206 **
Parity > 3	-0.057	-0.082
insurance	0.091 ***	0.058 ***
Model Estimates		
$\lambda$	0.690***	-
$\rho$	-	0.590***
LR Test Value	235.8 ***	215.12 ***
Log-Likelihood	-2353.1	-2363.4
AIC	4734.2	4754.8
Studentized Breusch-Pagan	22.19**	29.42***
Note: *** p < 0.001, ** p < 0.01, * p < 0.1		

Table 6  
**Factors influencing Breast Screening (Spatial Error & Spatial Lag Model)**

Characteristics breast	Spatial Error	Spatial Lag
Literate	-0.011	0.056 **
Currently Married	0.116 **	0.117 **
Hindu	-0.038 **	-0.036 ***
General Caste	0.064 ***	0.046 ***
Rural	0.062 ***	0.019
Rich	0.179 ***	0.085 ***
Oral Contraception	-0.134 **	-0.137 ***
Obese	-0.285 ***	-0.222 ***
Tobacco Consumption	-0.0732 *	-0.054 **
Alcohol	-0.005	-0.013
Insurance	-0.005	0.017
Model Estimates		
$\lambda$	0.620***	-
$\rho$	-	0.534***
LR Test Value	164.3 ***	161.5 ***
Log-Likelihood	-2062.04	-2063.4
AIC	4152.1	4154.9
Studentized Breusch-Pagan	35.3***	30.52***
Note: *** p < 0.001, ** p < 0.01, * p < 0.1		

Table 7  
**Spatial Autocorrelation of Residuals for Cervical and Breast Screening**

Residuals	Moran's I	
	Cervical	Breast
OLS	0.422 ***	0.344 ***
Error	-0.039	-0.047
Note: *** = p < 0.01 ** = p < 0.05 * = p < 0.1		

The spatial error model employed for breast screening indicated a statistically significant spatial autocorrelation, with  $\lambda = 0.620$ . Districts with obese women are negatively associated (-0.29) with the uptake of screening. A similar association was observed among for districts with Hindu women, those who used oral contraceptives, and those who consumed tobacco. A significant positive association (0.12) with uptake of breast screening was observed among districts with currently married women, those residing in rural areas, those belonging to a general caste, and those who were economically well-off.

The spatial autocorrelation ( $\lambda$ ) came out to be statistically significant in the spatial error model, indicating that the relationship between screening and the independent variables at the macro level (districts) may be misleading if spatial clustering is ignored. The spatial regression analysis enabled us to examine the spatial relationships of cervical and breast screening with their respective independent variables at the district level. It also helped us identify the determinants promoting the spatial pattern; in other words, factors that would help explain why and where screening was high.

## Discussion

This study showed variable patterns of cervical and breast screening with regard to geographic differences in India. Studies in the past may have examined the determinants of screening at the national and state levels in India, but to our knowledge, no study has captured district-level spatial variations in cervical and breast screening. Districts are the smallest administrative unit in India. Analysis at this level yield meaningful insights. The primary objective of the study was to examine the spatial patterns of screening and its determinants. One of the principal findings was the significant positive association between having insurance and undergoing cervical screening at the district level in both the models. Health care coverage may affect the decision to undergo screening since those who are protected for such procedures pay less out of pocket than those whose costs are not adequately covered (Jepson, R. et al., 2000). It's worth noting that the marital status of women has a considerable role in influencing their decision of undergoing screening. For both the models, cervical and breast screening showed a significant positive association among currently married females. Similar associations have been documented in other studies as well (Lin, S. J., 2008). Another crucial finding that emerged from our analysis was the statistically significant and positive association between high socioeconomic status and uptake of cervical and breast screening. This strongly resonates with the fact that the economic status of a woman profoundly influences her decision to undergo screening (Lin, S. J. 2008; Wu, S., 2003).

Clear and distinct spatial clusters in the different districts of Kerala and Maharashtra, covering nearly the whole state for both cervical and breast screening (NFHS-4, 2018), were hard to miss. This significant result may be attributed to the various steps taken by the Kerala state health department. Doctors in Kerala are sent to every home to screen women above the age of thirty years for breast and cervical screening. The Kerala Police and the Swasthi Foundation, in association with the Aster Med-city a leading quaternary care hospital in Kerala – launched “Rakshaka Raksha,” a series of free camps for the state police force to screen them for cancer and lifestyle diseases (Aster Medcity, 2017). Kerala was the first

state in the Indian union to formulate a cancer control program along the guidelines of the WHO as early as in 1988 (called 10-year action plan) (Nair, Mk, 2002). Even the panchayats in the state envisage cancer control activities as part of their People's Plan Programs. Kerala has turned out to be a role model for other states, with its focus on preventive health measures, in this case screening. Similarly, in few districts of Maharashtra screening programs are run by hospital, "Tata Memorial Hospital" is screening females for cervical cancer since 1998 and non-profit organization is helping through program like "Prashanti Cancer Care Mission" for breast cancer among females.

## Conclusion

As part of the National Health Mission, the Indian government for the first time has launched population-based prevention, screening, and control programs for cancers of the cervix and the breast. The spatial analysis presented here may add essential information in policy-making. By showing geographical disparities in screening, this study highlights the importance of ensuring a region-specific and organ-specific approach toward control and prevention of cancer. This is in line with the current government priorities. What is novel in the context of India is the evidence for distinct geographic patterns in the prevalence of screening practices among women, suggesting a potential role of factors such as accessibility, affordability, and availability as well as environmental factors at the district level, beyond any individual-level risk factor.

This study draws attention, in particular, to the lack of evidence for demonstrating and understanding existing disparities and the need to promote research to fill the gaps. This study is essential in both respects. Effective implementation of population-based screening programs is the need of the hour and could be a way of improving the health outcomes of women. Screening may, therefore, be an effective way of avoiding stigmatization of women appearing for gynecological examinations. This will also minimize the impact of social inequalities on health outcomes among women from different social backgrounds. The triple-A approach, that is, accessibility, affordability, and availability of health care facilities best resonates in the case of cervical and breast screening. Modeling the cost-effectiveness of this whole affair is another critical component to deal with when there is a rising global concern over the high levels of out-of-pocket expenditure on health.

This study has some limitations, and future research should be encouraged in that direction. Firstly, the National Family Health Survey provides data for women in the reproductive age group of 15–49 only. This prevented the study from analyzing women undergoing screening beyond this age group. Secondly, the scope of the paper limited us from providing any information as to whether women undergo screening of their own volition or due to external factors like government interventions. Thirdly, women opting for screening include both those who like to practice preventive behavior as well as those who are suffering from the disease itself. It is difficult to draw any inference due to data limitation.

## Abbreviations

LISA: Local Indicators of Spatial Autocorrelation

OLS: Ordinary Least Square Regression

VIF: Variance Inflation Factor

AIC: Akaike Information Criterion

NFHS: National Family Health Survey

## Declarations

**Ethical approval:** No ethical approval was required as this study is based on survey data available in public domain.

**Competing interest:** We declare that we have no competing interest.

**Funding:** No funding was received for this study from any funding agencies or organizations

**Authors' Contribution:** M and RM have conceptualized the idea of the research paper, M and RM were equally involved in analyzing, drafting and editing of the paper. Both the authors have read and approved the manuscript

## References

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, 27(2), 93-115.

Akinyemiju, T. F., McDonald, J. A., & Lantz, P. M. (2015). Health care access dimensions and cervical cancer screening in South Africa: analysis of the world health survey. *BMC public health*, 15(1), 382.

Aster Medcity (2017), <http://astermedcity.com/events-news-and-gallery//Readmore/aster-medcity-champions-cancer-screening-with-kerala-police-and-swasthi-fou/38965> [last Accessed 21st Dec 2018]

Beining, R. (2012). Screening for cervical cancer: an exploratory study of urban women in Tamil Nadu. *India [Ph.D. thesis] Iowa City, IA: University of Iowa.*

Bosch, F. X., Munoz, N., & Shah, K. (1997). Social differences in sexual behavior and cervical cancer. *IARC scientific publications*, (138), 309-317.

Britain, G. (1998). *National Screening Committee: First Report of the National Screening Committee.* Department of Health.

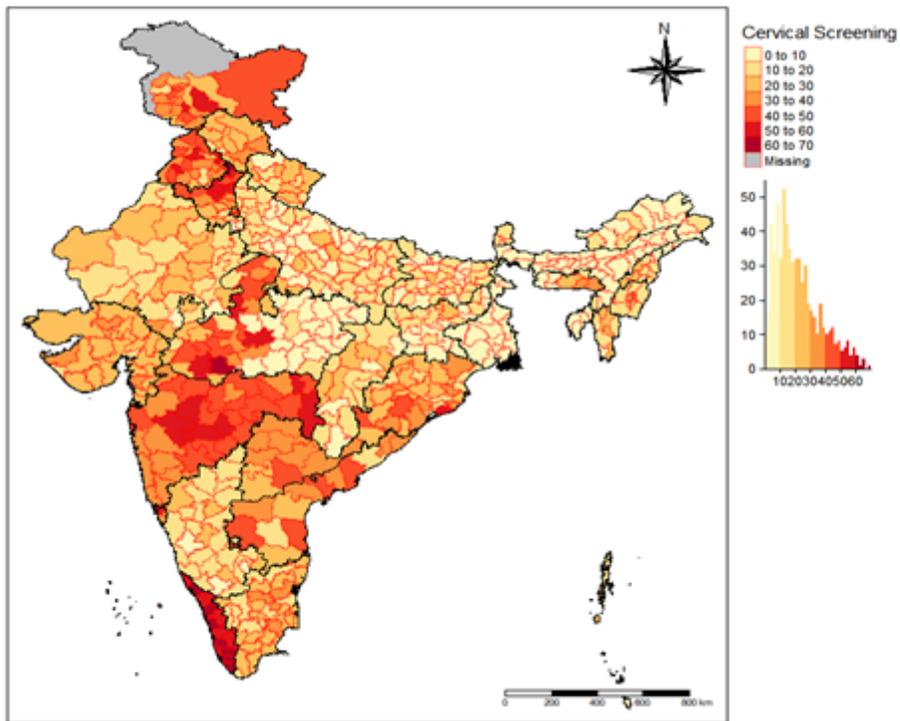
Centers for Disease Control and Prevention, (2018).

[https://www.cdc.gov/cancer/breast/basic\\_info/risk\\_factors.htm](https://www.cdc.gov/cancer/breast/basic_info/risk_factors.htm) [Accessed on 25 May 2018]

- Coughlin, S. S., Leadbetter, S., Richards, T., & Sabatino, S. A. (2008). Contextual analysis of breast and cervical cancer screening and factors associated with health care access among United States women, 2002. *Social science & medicine*, 66(2), 260-275.
- Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., ... & Bray, F. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer*, 136(5).
- Fotheringham, A. S., Brunson, C., & Charlton, M. (2003). Geographically weighted regression: the analysis of spatially varying relationships. *John Wiley & Sons*.
- Franco, E. L., Duarte-Franco, E., & Ferenczy, A. (2001). Cervical cancer: epidemiology, prevention and the role of human papillomavirus infection. *Canadian Medical Association Journal*, 164(7), 1017-1025.
- Harris, R. (2013). An Introduction to Mapping and Spatial Modelling R. *University of Bristol*
- Herrmann, C., Vounatsou, P., Thürlimann, B., Probst-Hensch, N., Rothermundt, C., & Ess, S. (2018). Impact of mammography screening programmes on breast cancer mortality in Switzerland, a country with different regional screening policies. *BMJ open*, 8(3), e017806.
- Huque, M. H., Bondell, H. D., & Ryan, L. (2014). On the impact of covariate measurement error on spatial regression modelling. *Environmetrics*, 25(8), 560-570.
- IIPS & ICF. (2017). National Family Health Survey 4 (NFHS-4) 2015–16, India.
- Jepson, R., Clegg, A., Forbes, C., Lewis, R., Sowden, A., & Kleijnen, J. (2000). The determinants of screening uptake and interventions for increasing uptake: a systematic review.
- Kaarthigeyan, K. (2012). Cervical cancer in India and HPV vaccination. *Indian journal of medical and paediatric oncology: official journal of Indian Society of Medical & Paediatric Oncology*, 33(1), 7.
- Kitchener, H. C., Castle, P. E., & Cox, J. T. (2006). Achievements and limitations of cervical cytology screening. *Vaccine*, 24, S63-S70.
- Legendre, P. (1993). Spatial autocorrelation: trouble or new paradigm?. *Ecology*, 74(6), 1659-1673.
- Legendre, P., Dale, M. R., Fortin, M. J., Gurevitch, J., Hohn, M., & Myers, D. (2002). The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography*, 25(5), 601-615.
- Lin, S. J. (2008). Factors influencing the uptake of screening services for breast and cervical cancer in Taiwan. *The journal of the Royal Society for the Promotion of Health*, 128(6), 327-334.
- McLafferty, S., Wang, F., Luo, L., & Butler, J. (2011). Rural-urban inequalities in late-stage breast cancer: spatial and social dimensions of risk and access. *Environment and Planning B: Planning and Design*, 38(4), 726-740.

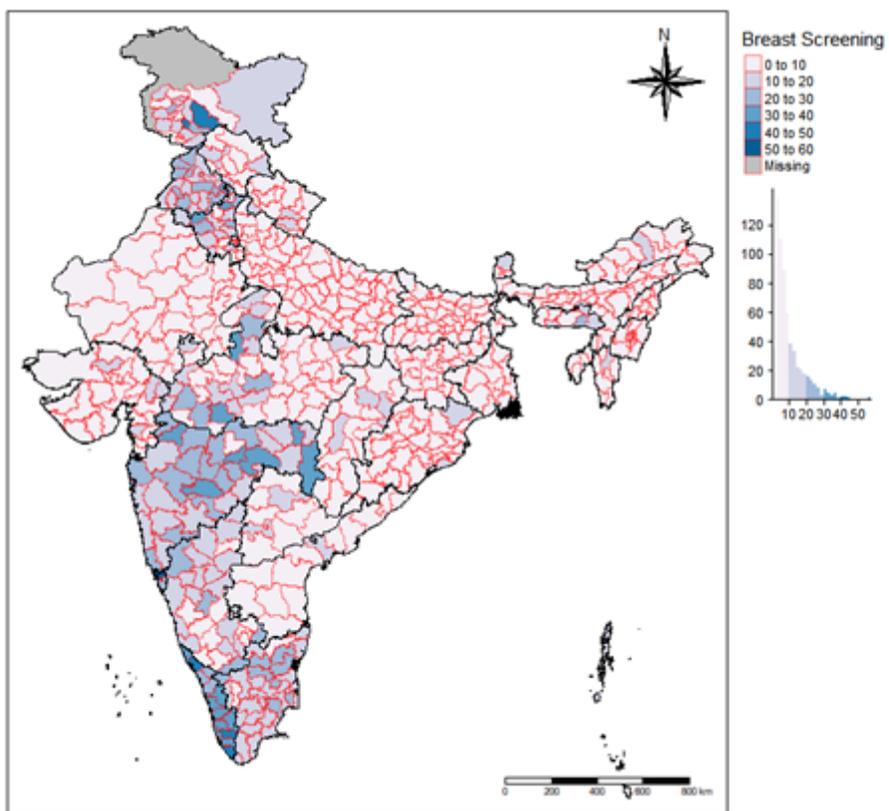
- Nair, M. K. (2002). Cancer Control in Kerala, India. In *50 years of Cancer Control in India* (pp. 114-128). National Cancer Control Programme, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India.
- Prasad, J. B., & Dhar, M. (2017). Tobacco use in India and its states: Burden of smoking and smokeless forms of tobacco (2015-25) and its predictors. *Journal of Cancer Policy, 14*, 21-26.
- Rastogi, T., Hildesheim, A., & Sinha, R. (2004). Opportunities for cancer epidemiology in developing countries. *Nature Reviews Cancer, 4*(11), 909.
- Selvin, E., & Brett, K. M. (2003). Breast and cervical cancer screening: sociodemographic predictors among White, Black, and Hispanic women. *American journal of public health, 93*(4), 618-623.
- Stewart, B. W. K. P., & Wild, C. P. (2014). World cancer report 2014. Vol. 3. Lyon: IARC .
- Viens, L., Perin, D., Senkomago, V., Neri, A., & Saraiya, M. (2017). Questions about cervical and breast cancer screening knowledge, practice, and outcomes: A review of demographic and health surveys. *Journal of Women's Health, 26*(5), 403-412.
- WHO, (2016). Cancer prevention and control in the context of an integrated approach [http://apps.who.int/gb/ebwha/pdf\\_files/EB140/B140\\_31-en.pdf](http://apps.who.int/gb/ebwha/pdf_files/EB140/B140_31-en.pdf) [Last Accessed on 21<sup>st</sup> December 2018]
- WHO/ICO Information Centre on HPV and Cervical Cancer (HPV Information Centre), (2007). *Summary report on HPV and cervical cancer statistics in India*. Available from. :<http://www.who.int/hpvcentre> .
- World Health Organization, (2019) cancer Fact-Sheet <https://www.who.int/cancer/resources/keyfacts/en/> [Last Accessed on 19 December, 2019]
- World Health Organization, (2018). International Agency for Research on cancer <https://www.who.int/cancer/PRGlobocanFinal.pdf> [Last Accessed on 19 December, 2019]
- Wu, S. (2003). Sickness and preventive medical behavior. *Journal of health economics, 22*(4), 675-689.
- Yang, T. C., Matthews, S. A., & Hillemeier, M. M. (2011). Effect of health care system distrust on breast and cervical cancer screening in Philadelphia, Pennsylvania. *American journal of public health, 101*(7), 1297-1305.

## Figures



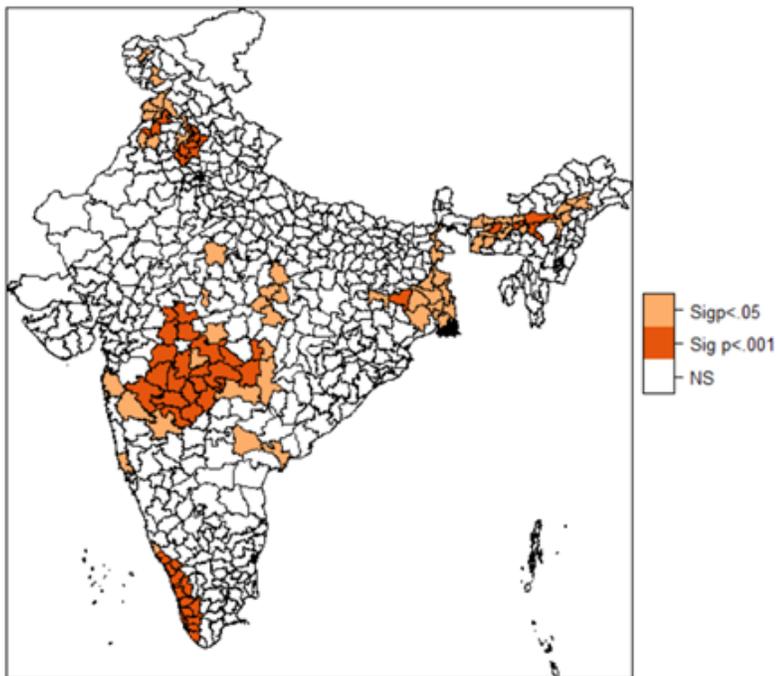
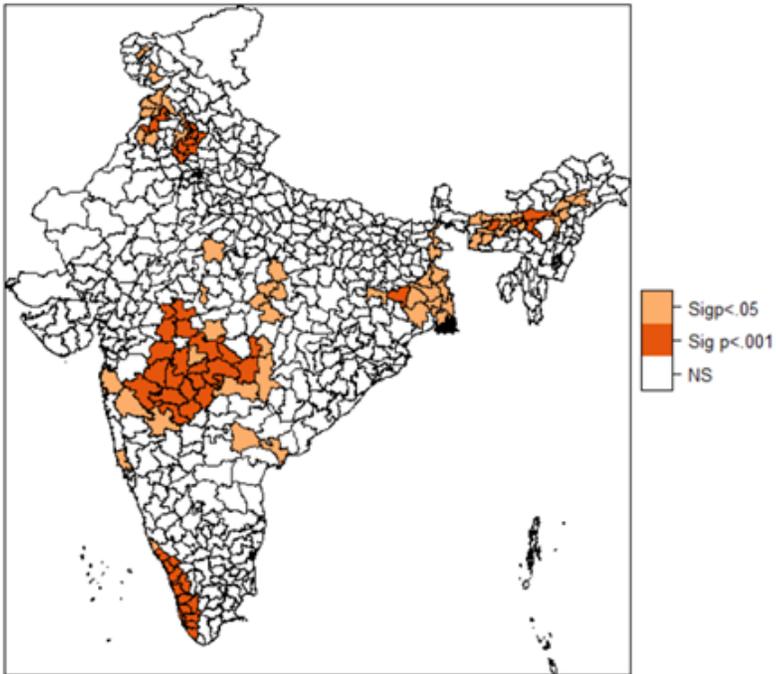
**Figure 1**

Prevalence of Cervical Screening



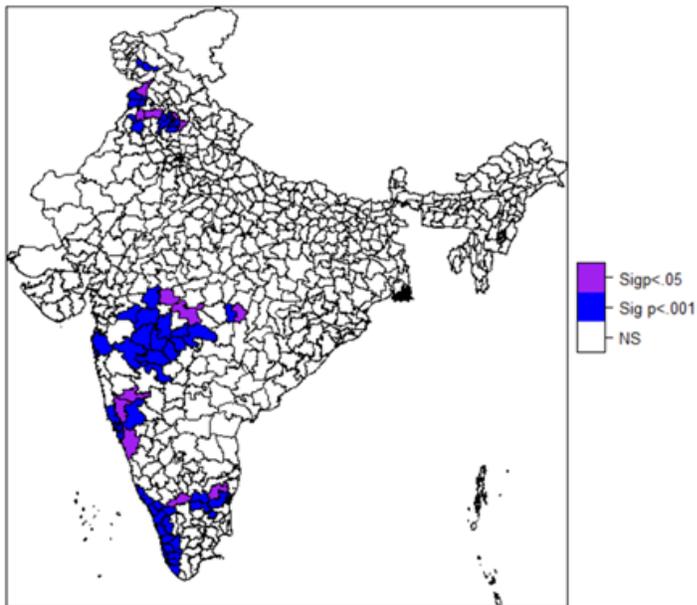
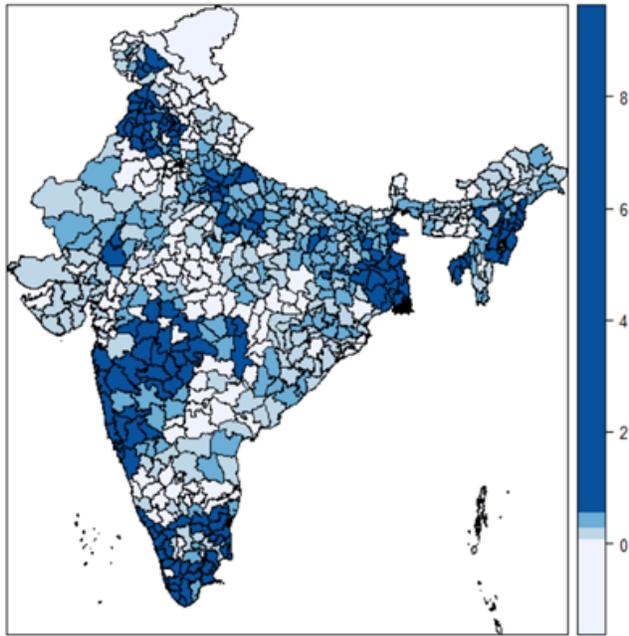
**Figure 2**

# Prevalence of Breast Screening



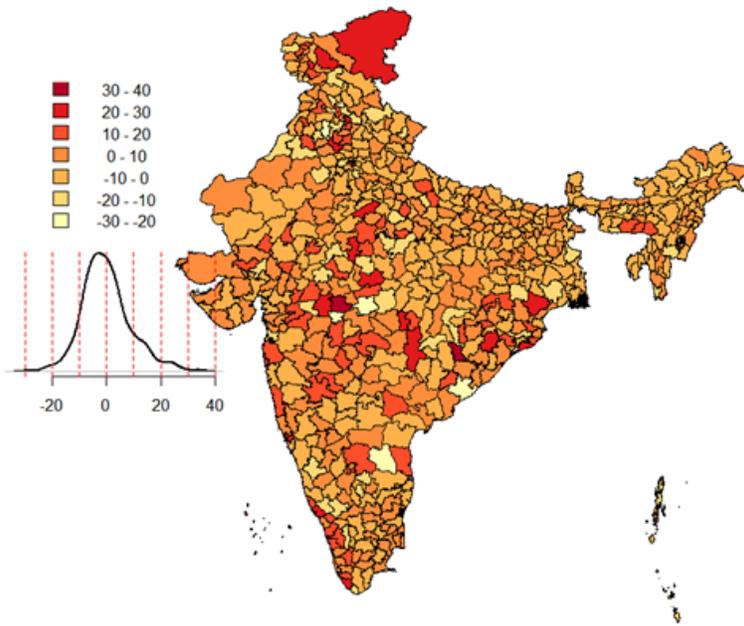
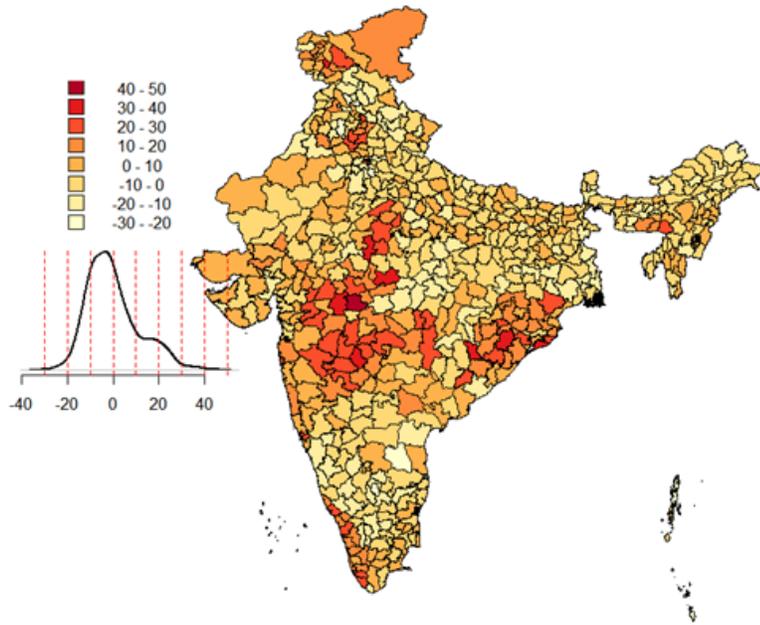
**Figure 3**

a: Local Moran's I- Cervical Screening. b: Significant Cluster Map- Cervical Screening



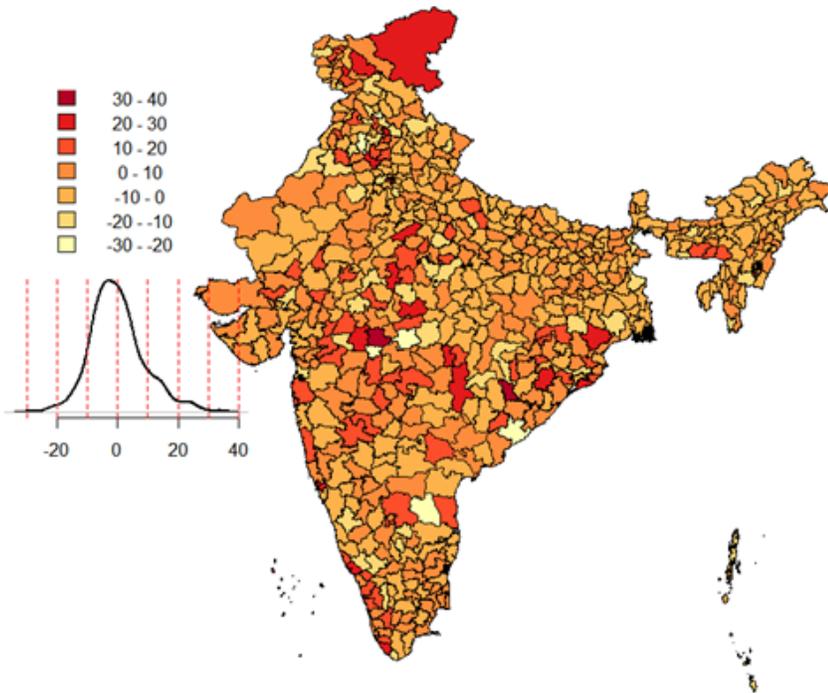
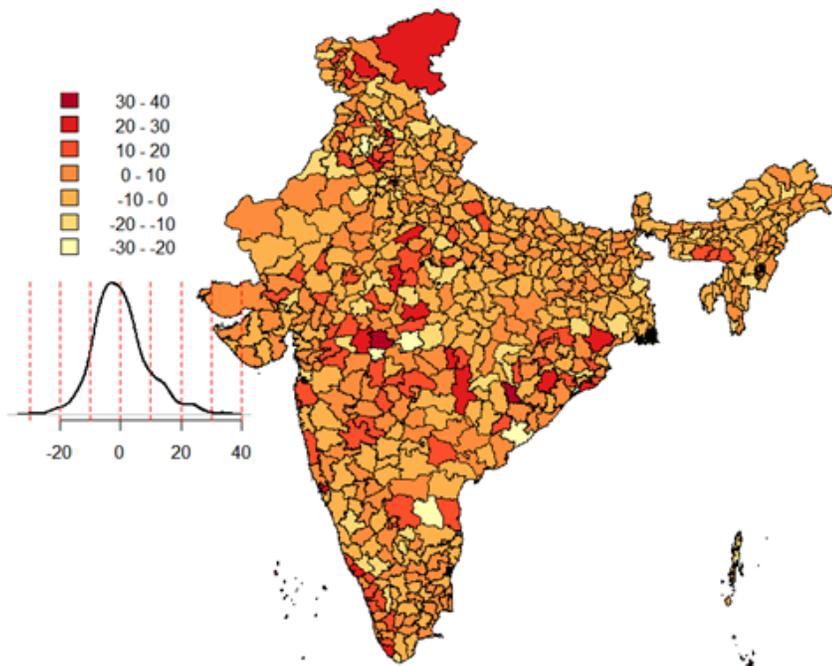
**Figure 4**

a: Local Moran's I- Breast Screening. b: Significant Cluster Map- Breast Screening



**Figure 5**

a: OLS Residual Map - Cervical Screening. b: Spatial Error Model Residual Map - Cervical Screening



**Figure 6**

a: OLS Residual Map - Breast Screening. b: Spatial Error Model Residual Map - Breast Screening

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementary.pdf](#)
- [Formulas.pdf](#)