A Study On Spatial Variations In Temporal Trends Of Dengue Incidences In Tamil Nadu, India

Jaisankar R, Kesavan J

Abstract: Finding disease clusters is an important facet of epidemiology, which is very important to develop surveillance systems that are being used to detect disease outbreaks. In this context, more meaningful information from the collected data can effectively be obtained using statistical analysis with the geo-visualization approach by adding spatial parameters through which one can identify the spatial patterns or space or space-time clusters of the disease. It is well known that the disease incidence changes over time and it is possible that these temporal trends are different in different geographical regions and the knowledge about such differences may provide additional information which may aid for successful disease prevention and control measures. In this paper, Spatial Variations in Temporal Trends (SVTT) analysis using the Poisson model assumption is used for the identification of trends in incidence over time for disease clusters of dengue incidences in Tamil Nadu. Maps of the geographical location of disease clusters are provided to enhance the understanding of results of statistical analysis.

Keywords: Disease temporal trends, spatial variations, scan statistics, dengue incidence.

1. INTRODUCTION

Spatial analyses in epidemiology utilizes geo-referenced data for finding three interrelated objectives, namely, the mapping of diseases and their determinants, the study of geographical correlations and disease clustering analyses. Spatial and space–time clustering analysis detects event clusters which are hotspots with a number of events significantly higher than that would be expected. These methods are applied various arena of sciences, such as ecology, criminology, environmental sciences and medicine identifying areas that differ in some aspect. In the recent past, spatio-temporal clustering techniques applied to the realm of health and epidemiology leading to potential achievements towards the monitoring and prevention of diseases. Spatial variation in temporal trends is relatively new approach which identifies the clustering of geographical locations according to a common trend that is significantly different from trends in outside areas and hence proven to be highly useful for epidemiological surveillance and disease control. Both Space–time clustering and spatial variations in temporal trends play significant roles in identifying areas with distinct epidemiological behavior that is not explained by randomness, providing more meaningful information to public health decision-makers for taking more appropriate strategic interventions.

- *Jaisankar R, Professor, Department of Statistics, Bharathiar University, Coimbatore, Tamil Nadu, India. (E-mail: r_jaisankar@rediffmail.com)*
- *Kesavan J, Research Scholar, Department of Statistics, Bharathiar University, Coimbatore, Tamil Nadu, India. (E-mail: kesavanstat@gmail.com)*

2. REVIEW OF LITERATURE

Scan statistics methodology was introduced in the field of health sciences by Naus (1965a, 1965b). Spatial and spatiotemporal extensions of the methods have been introduced by Kulldorff (1997) and Kulldorff et al. (1998, 2006). The scan statistics for spatial variations in temporal trends (SVTT) are designed for the detection of clusters of areas with unusual different temporal trends was introduced by Kulldorff(2010) which is based on scan statistics and uses a Poisson regression with time as independent variable to estimate the disease trend. As type of estimation makes impossible to detect points in time where the tendency changes and gives low power in some situations. Kulldorff (2013) proposed to new methodology called quadratic SVTT which is modification of the linear method where the trend estimation procedure is changed by introducing a new explanatory variable, time squared, to the regression model used in the linear method. Paula Moraga et al., (2013) have studied spatial variations in temporal trend method with a quadratic function and applied the same to cervical cancer mortality in the area of United States. Areias C et. al., (2015) used the technique of space time clustering and spatial variation in temporal trends for pulmonary tuberculosis incidences in Portugal. In present work deals with the spatial variations in temporal trends for dengue incidences occur in Tamil Nadu with assumption of linearity.

3. METHODOLOGY

Spatial Variations in Temporal Trends (SVTT) Method

There are several spatio-temporal models for disease mapping are available for investigating emerging spatial patterns and temporal trends of disease risks, have been proposed. Their bases are either a parametric description of time trends, independent risk estimates for every period, autoregressive approaches or the definition of the joint covariance matrix for all the periods as a Kronecker product of matrices. A natural model for disease mapping in space-time scenario is given by,

$$Y_{ij} \sim P_0\left(E_{ij} \times \theta_{ij}\right)$$

where Y_{ij} and E_{ij} denotes the observed and the expected number of cases of disease in area *i* and time period *j* respectively and θ_{ij} is the relative risk in area *i* and time period *j*. There are three groups of components for $\log\left(\theta_{ij}\right)$, namely, the spatial group where (A_i), the temporal group (B_j), and the space-time interaction group (C_{ij}) which can be described as

$$\log\left(\theta_{ij}\right) = \alpha_0 + A_i + B_j + C_{ij},$$

These can be defined as follows: $A_i = \phi_i$, where ϕ_i is an area random effect, $B_j = \beta t_j$ where βt_j is linear trend term in time is t_j and $C_{ij} = \delta_i t_j$, where $\delta_i t_j$ is an interaction random effect between area and time. By using these hierarchical models we borrow information from neighboring areas for improving the local estimates by taking in to account of spatial dependence between the relative risks. Further these models increase flexibility by introducing an unstructured exchangeable component that models uncorrelated noise. Estimation of covariate effects, predict missing data and handle multiple diseases can effectively be done through these models.The spatial, temporal and spatio-temporal scan statistics are used to detect spatial, temporal and spatio-temporal disease clusters effectively and hence highlighting the areas where the numbers of cases are significantly greater than expected. Apart from these, the scan statistics for spatial variations in temporal trends (SVTT) are used for the detection and inference of any zone with significantly different temporal trend. This method takes a fixed temporal period of interest where the disease trend is assessed and gradually scans a spatial window centered in each location and with different sizes. The disease trend in each of these windows is then estimated. The window whose trend that mostly differs from the rest of the trends is considered as having unusual different trend and its statistical significance is assessed. The detection and inference of any geographical location with exceptionally increasing or decreasing linear trend can be identified through the linear SVTT method using a special type of scan statistic. Here, the estimation of trend is done using a Poisson regression with time as independent variable, the number of events as dependent variable and the time changing population size as offset. These estimated trends are then used to adjust the expected number of cases for each location and time. The adjustment will be different inside and outside the window due to the different estimated trends. Then the likelihood for this window is calculated using the new expected counts and the maximum likelihood over all windows is found. The maximum of this is then compared with the maximum likelihoods from a large set of random data. The analysis is conditioned on the overall trend, since the prime objective is only in the difference in the trends between areas. This is done in the randomization step, for each time, a spatial location according to the background population size at that time is randomized and not by randomizing the observed times. The quadratic SVTT method is also available (Moraga P and Kulldorff M, 2013) which consists of similar steps as in the case of linear SVTT method but the only difference in the way the trends are estimated. A huge number of windows are constructed at first. Then, the trends inside and outside are estimated and the likelihood is computed for each window. The window with maximum likelihood is taken-up as the most likely cluster and its *p*-value is obtained using Monte Carlo hypothesis testing. In this work spatial variation in temporal trends has been found under the assumption of linear trend functions.

4. DATA COLLECTION AND PROCESSING

Data collection

Data for the analysis was collected from Public Health and Preventive Medicine, Tamil Nadu, Health Department. The data covering details on dengue cases in Tamil Nadu, District-wise which are available for the years 2007 to 2018.

Data Pre-processing

For the analysis coordinate information about the geographical location are collected where the cases have occurred, which includes the details about location name, latitude and longitude of the location of the cases.

Map Digitization

Geo-referencing process is used to assign real-world coordinates to each pixel of the raster using QGIS. In the presented work, scanned map of Tamil Nadu with district wise is digitized by obtaining coordinates from the markings on the map image itself. Using these GCPs (Ground Control Points), the image is warped and it is made to fit within the chosen coordinate system.

5. APPLICATION TO DENGUE INCIDENCE IN THE TAMIL NADU

Total number of locations taken up is 32 districts of Tamil Nadu and total number of cases reported is 61295 from January 2007 to October 2018. The Population averaged over time is 73379945 and on an average 7 individuals were affected by dengue among every 100000 people per annum and time trend annual increased in 17.408 %. The analysis of detection of clusters is done by SaTScan software with the assumption that the number of events in a geographical location is Poisson distributed. The shape file of the clusters is generated and saved by the package is easily imported in Google Earth so that interpretations can be made easily. The disease clusters are detected with the cluster radius as 1 kilometers and the proposed visualization method maps are generated using QGIS softwareOn the basis data collected from 2007 to 2018 the infected cases are highest for the year 2017, almost in all the districts. While making comparison among the districts Tirunelveli is with highest incidence and Chennai is placed as second. The proportion of males affected by dengue is higher than that of females.The spatial variations in temporal trends occurred in dengue incidence in the Tamil Nadu during the period 2007 to 2018 are studied by SVTT analysis. The linear SVTT method to detect groups of counties with temporal trends is used. The method identifies 5 significant groups. For each of these detected groups, the table 1 appendix below shows its population,

number of observed and expected cases, risk, LLR and p-value together with the parameter estimates of the trends inside and outside the groups. The groups of counties detected are labeled with the numbers 1 to 5 according to their significance, with the group 1 being the most significant and the group 4 the less significant. Figure 1show appendix, it depicts the significant detected clusters of districts in the Tamil Nadu map.The following are the details of detected clusters: The locations included in the most likely cluster are Thoothukkudi and Tirunelveli with number of cases observed to be 10375 and expected number of cases as 4081.86, inside and outside time trend annually increased by 25.753% and 15.958%, highest among detected clusters, which is found to be statistically significant and hence the occurrences are not by chance.The next significant cluster is found in Tiruppur with observed and expected number of cases as 2834 and 1974.55, inside and outside time trend annually increased by 28.711% and 16.942% respectively. The third significant cluster is found in Viluppuram, Cuddalore, Tiruvannamalai, Ariyalur, Kancheepuram with observed and expected number of cases as 4955 and 11149.43 respectively and inside and outside time trend annually increased by 25.122% and 16.830% respectively. The fourth significant cluster is found in Salem with observed and expected number of cases as 2677 and 3007.33, inside and outside time trend annually increased by 26.894% and 16.830% respectively. The fifth significant cluster is found in Theni with observed and expected number of cases as 2825 and 1052.11, inside and outside time trend annually increased by 24.079% and 17.127% respectively.

6. SUMMARY AND CONCLUSION

It is observed that incidence is increasing overall but not in the same way in all areas. On an average 7 individuals were affected by dengue among every 100000 people per annum and the annual increase in the trend is 17.408 %. The most likely cluster are Thoothukkudi and Tirunelveli with number of cases observed to be 10375 and expected number of cases as 4081.86 and inside time trend annually increased by 25.753% , which is found to be statistically significant.The spatial and Geographical visualization approaches are developed in this paper using the spatial variation in temporal trend in cluster detection method which provides an efficient representation of the results of statistical analysis in geographical space. These maps show density of cases within each district, Clusters with high rates of disease and statistically significant clusters. Proactive actions can be taken to prevent disease outbreaks, using there results. The present study is helpful in identifying the hot spot districts of an epidemic.

REFERENCES

[1] Areias C., Briz T and Nunes C. 2015. Pulmonary tuberculosis space-time clustering and spatial variation in temporal trends in Portugal, 2000-2010: an updated analysis. Epidemiolog & Infection. 143:3211-3219.

[2] Gail, M., Krickeberg, K., Samet, J., Tsiatis, A., Wong, W. and Tango, T. 2009. Statistics for biology and health, Springer.

[3] Kulldorff, M. 1997. A spatial scan statistic. Communication Statistics Theory and Methods 26(6):1481-1495.

[4] Kulldorff, M. 2018. SaTScan User Guide v9.6. Martin Kulldorff and Information Management Services Inc., http://www.satscan.org.

[5] Kulldorff, M., Athas F. William, Feuer J. Eric, Miller A. Barry and Key R. Charles. 1998. Evaluating cluster alarms: a space-time scan statistic and brain cancer in los alamos, new mexico, American Journal of Public Health 88(9):1377-1380.

[6] Kulldorff, M., Feuer J. Eric, Miller A. Barry and Freedman S. Laurence. 1997. Breast cancer clusters in the northeast united states: a geographic analysis. American Journal of Epidemiology 146(2): 161-170.

[7] Kulldorff, M., Heffernan Richard, Hartman Jessica, Assunc Renato and Mostashari Farzad. 2005. A space time permutation scan statistic for disease outbreak detection. plos journals 2(3):216-224.

[8] Kulldorff, M., Huang, L., Pickle, L. and Duczmal, L. 2006. An elliptic spatial scan statistics. Wiley InterScience 25:3929-3943.

[9] Kulldorff. M, Athas. W, Feuer. E, et al. 1998. Evaluating cluster alarms: a space-time scan statistics and brain cancer in LosAlamos. *Am J Public Health 88(12): 1377–1380*.

[10] Moraga P, Kulldorff M. 2013. Detection of spatial variations in temporal trends with a quadratic function. Statistical Methods in Medical Research, pp. 1-16.

[11] Tatlonghari Jayson A. Rick. 2009. GIS Manual using Quantum GIS, pp. 23-36.

APPENDIX

Table 1: *Detected groups of counties with Dengue Incidence trend in Tamil Nadu State over the period 2007 to 2018. For each of the detected groups, population, number of observed and expected cases, risk, LLR, p-value and parameter estimates of the trends inside and outside the group.*

Cluster	Location	Population	Obs.	Exp.	Risk	Inside			Outside			LLR	p-value
						Trend	Intercept	Linear	Trend	Intercept	Linear		
1	Thoothukkudi	4895409	10375	4081.86	2.86	25.75	-0.65	0.23	15.96	-1.08	0.15	252.78	0.001
2	Tiruppur	2360417	2834	1974.55	1.46	28.71	-1.42	0.25	16.94	-1.04	0.16	103.05	0.001
3	Viluppuram	13342771	4955	11149.43	0.40	25.12	-2.36	0.22	16.83	-0.90	0.16	94.24	0.001
4	Salem	3588336	2677	3007.33	0.89	26.89	-1.78	0.24	17.03	-1.02	0.16	71.55	0.001
5	Theni	1258648	2825	1052.11	2.77	24.08	-0.48	0.22	17.13	-1.07	0.16	40.34	0.001

Figure 1. Detected areas with Dengue Incidences Trends in the Tamil Nadu state over the period 2007 to 2018