ijer.skums.ac.ir

Pattern mining analysis of pulmonary TB cases in Hamadan province: Using space-time cube

Sayed Mehdi Hosseini[1], Masoud Parvin[2], Masoud Bahrami[1], Manoochehr Karami[3,4], Meysam Olfatifar[1*]

[1]*Student Research Committee, Epidemiology Dept., Hamadan University of Medical Sciences, Hamadan, I.R. Iran;* [2]*Student Research Committee, Dentistry Dept., Hamadan University of Medical Sciences, Hamadan, I.R. Iran;* [3]Social Determinants of Health Research Center, *Hamadan University of Medical Sciences, Hamadan, I.R. Iran;* [4]*Epidemiology Dept., Hamadan University of Medical Sciences, Hamadan, I.R. Iran.*

Received: 4/Nov/2016 Accepted: 11/Dec/2016

ABSTRACT

Background and aims: One of the most common approach to understanding spatial and temporal trends of event data is to break it up into a series of time snapshots. Therefore space-time cube method applied in order to portray the likely trend in occurrence of the pulmonary tuberculosis (TB) cases.

Methods: In this study, information of all patients with pulmonary TB recorded in surveillance system of Hamadan University of Medical Sciences from 2005 to 2013 years were studied. After geocoding the residence location address of pulmonary TB cases and converted to point layer, the space-time cube was used to detect likely trends in occurrence of tuberculosis. Then, based on the space-time cube results the Emerging Hot Spot Analysis was run to clustering hot and cold spots.

Results: There was significant increasing trend in occurrence of pulmonary TB cases. The statistic trend was 2.1871 and P-value was 0.0287, as well as 36 hot spots locations was detected that have been form approximately in central areas of province.

Conclusion: Significantly increasing trend in occurrence of TB cases and existence of hot spot, especially intensifying hot spots in central areas of province can represent pay more attention to this disease in mentioned areas in order to detect the change in epidemiological face and to implement suitable prevention programs.

Keywords: Trend, Space time cube, Pulmonary Tuberculosis, Hamadan.

Original article

INTRODUCTION

Tuberculosis (TB) is an infectious disease that affects the lives of millions people, especially people who live in developing countries. The importance of the problem is vary in different environments, which may be due to socio-economic

[*]**Corresponding author:** Meysam Olfatifar, Student Research Committee, Epidemiology Dept., Hamadan University of Medical Sciences, Hamadan, I.R. Iran, Tel: 00989189732971, E-mail: ol.meysam92@gmail.com

conditions, poverty, overcrowding, poor access to services, socio-cultural barriers and HIV infection.[1]

Each year, in world more than 9 million people fell ill with TB and more than 1.7 million people died from the disease.[2] TB is a public health threat over the world, that mostly affecting the lower income countries.[3,4] So that more than 95% of TB deaths occur in low and middle income countries and the third cause of death in 15-55 years old women.[4]

Among other forms of TB, pulmonary tuberculosis is avoidable if proper actions of health promotion, prevention and treatment do not perform.[5] However, detection of geographic areas with on-going transmission of disease, using GIS and spatiotemporal statistical analyses, have become inevitable.[6] studies that use spatial analysis methods in order to understanding TB and other problems have increased.[7,8] Because GIS is a method used to assess the effect of environmental and geographical factors on diseases and their aggregation and distribution locations. It also reduces the health costs related to diagnosis, treatment, as well as prevention.[9] Despite all of the GIS benefits, so far a study have not been done in Hamadan province to study spatial clustering of tuberculosis. Accordingly, this study was aimed to investigate likely trend in occurrence of pulmonary tuberculosis cases in Hamadan province from 2005 to 2013.

METHODS

Avaliable data on 1057 TB cases (include: age, gender, location, condition, etc.) from 2005 to 2013 years was obtained from the Deputy Health affiliated to Hamadan University of Medical Sciences. The information of 728 pulmonary tuberculosis cases was studied, that of these 65 cases due to incomplete information about the residence address and seven cases due to non-Iranian citizenship (lived in Malayer County) excluded from the study. The geographical coordinates (longitude and latitude) of patient's residence identified using Google Earth software. Coordinate information of patients converted to point layer data (using Projected Coordinate System: WGS_1984_UTM_Zone_39N).

Moreover the space-time cube tool was used to investigate likely trend of pulmonary TB point data.[10] This method aggregated the data point in small space-time bins and organized them in a data cube (netCDF file). In this space-time cube the y-axis represents the time and the x axis represent the place, latitude and longitude of bins is determined by the investigator, we assigned these parameter respectively a seven month and 6000 meters. The number of data points for each small bin count and likely existence trend in data for each bin during the time in each location calculated using the Mann-Kendall statistic. This statistic is a rank correlation analysis for the bin count values and their time sequence, meaning that compares bin count for each time period in each location (if the first is smaller than second, the result is +1 and vice versa and if two period be equal, the result is zero) and finally, sum the results of compared each pair time period.

Then, the created cube was used to run Emerging Hot Spot Analysis. This method considers the count value for each space-time bin within the context of the count values for neighbouring space-time bins and calculated the Getis-Ord GI* (Hot Spot Analysis) statistic for each bin and classified the study areas in 16 categories (new, consecutive, intensifying, persistent, diminishing, Sporadic, Historical and no trend) of hot or cold spots.

RESULTS

Totally 728 cases of pulmonary TB (68%) from all of the 1057 cases recruited

for analysis. The distribution of the pulmonary and extra-pulmonary cases and number of diagnosed cases according to county are shown in Table 1 and Figure 1. Hamadan province had the most of pulmonary and extra pulmonary tuberculosis cases Table1 and in both cases most of TB cases were in 2010 year Figure 1.

Table 1: Distribution of pulmonary and extra pulmonary TB cases in Hamadan province, in duration of study

County	Pulmonary	Extra pulmonary
Hamadan	270	169
Bahar	55	19
Toyserkan	37	25
Asadabad	53	19
Razan	37	8
Famenain	16	2
Malayer	114	35
Nahavand	60	32
Kabudrahang	86	20
Total	728	329

Geographical coordinates of 656 patient's residence showed in Figure 2a. The space time cube contains point counts for 754 locations over 16 time step intervals. Each location is 6000 meters by 6000 meters spanning an area 156000 meters west to east and 174000 meters north to south. Each of the time step intervals is 7 months in duration so the entire time period covered by the space time cube is 112 months. Of the 754 total locations, 142 (18.83%) contain at least one point for at least one time step interval. These 142 locations comprise 2272 space time bins of which 418 (18.40%) have point counts greater than zero. There is a statistically significant increase in point counts over time; trend statistic and P-values were respectively 2.1871 and 0.0287. Characteristics of the space-time cube have been shown in Table 2.

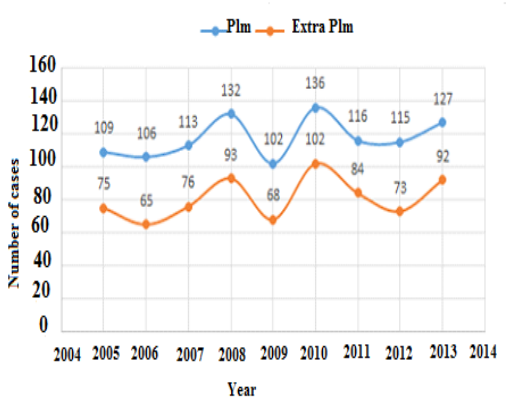

Figure 1: The number of pulmonary and extra pulmonary TB cases in Hamadan province by year

Table 2: Characteristics of created space-time cube of pulmonary TB cases in Hamadan province

Place	Number or percent	Time	Number/time (week or month)	Cube extent across space	Coordinates (meters)
Total number of location	754	Time step interval	1 week	Min X	206576.1330
Location with at least one point	142	First time step interval	2005/1/18 to 2005/1/25	Min Y	3774028.6592
Associated bins	66314	Last time step interval	2013/12/24 to 2013/12/31	Max X	359823.7299
Non zero (sparseness)	0.95	Number of time steps	467	Max Y	3947606.6193
Rows	29	Columns	26	Total bins	352118

Emerging Hot Spot Analysis parameter (neighborhood distance and neighborhood time step intervals) were set respectively 18000 meters and one interval (7 months). The space time cube contained point counts for 582 locations over 16 time step intervals, so that in total 9312 space-time bins were analyzed (multiplying number of time steps in number of locations). Table 3 summarized results of the Emerging Hot Spot Analysis, 36 of 582 locations were hot spot, of these six locations were consecutive hot spot (a single uninterrupted run of hot time step intervals, comprised of less than 90% of all intervals), 18 locations were intensifying hot spot (at least 90% of the time step intervals are hot, and becoming hotter over time) and 12 locations were Sporadic hot spot (some of the time step intervals are hot) respectively Table 3 and Figure 2b. The west of Hamadan city, south-east of Bahar and north of Toyserkan city are location that mentioned hot spot have been formed Figure 2b.

Table 3: Results of Emerging Hot Spot Analysis of pulmonary TB cases in Hamadan province

Type	Hot	Cold
Consecutive	6	0
Intensifying	12	0
Persistent	0	0
Diminishing	0	0
Sporadic	18	0
Oscillating	0	0
Historical	0	0

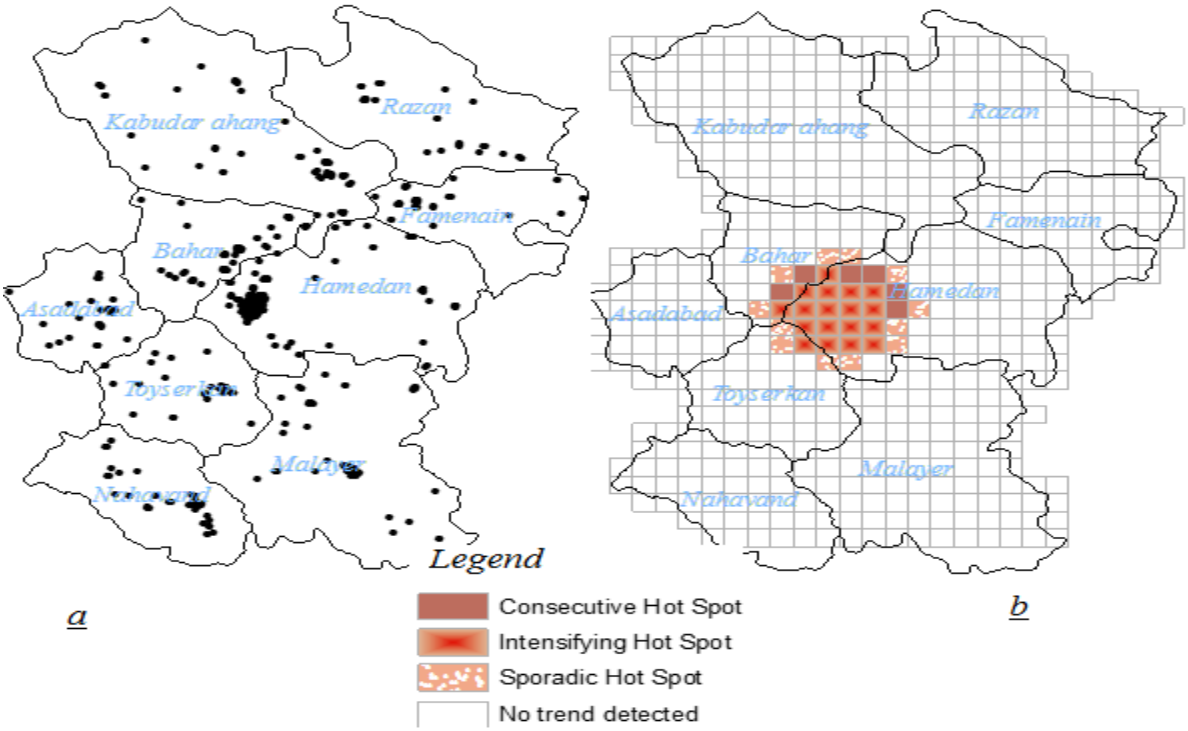

Figure 2: Geographical location of Pulmonary TB related hot spots in Hamadan province

a: Each dot presents a case, the dots are the actual cases locations; b: Emerging Hot Spots Analysis of pulmonary TB cases in Hamadan province.

DISCUSSION

In this study, a significant increasing trend over time was observed and various hot spots locations were detected in some areas of province especially central areas, which can indicate pay more attention to disease and set up an active surveillance system for early detection of possible changes, because regarding to current trend and intensifying hot spot definition (becoming hotter over time), the number of cases would be increased more than existence estimates.

To compare with other studies, we could not find similar articles that used method applied in this work, with search in PubMed and Google scholar in context of tuberculosis and other diseases but find articles that use hot spot analysis and other similar epidemiological methods (such as scan statistic and global and local Moran I Indices) in Iran and other regions of the world. On other hand, it was tryed to focus on similar published studies about tuberculosis in Hamadan province. For example, Liu and et al. studied pulmonary TB cases in Gansu and showed that hot spot areas were mainly in Hexi area, Linxia, part of Dingxi.[11] Yazdani-Charati and et al. studied Tb cases in north of Iran using Moran index and showed that high risk cores (hot spots) of Tb incidence were identified in Mazandaran province.[12] Ghanbarnezhad and et al. used Moran index to study spatial distribution of TB and HIV co-infection in south of Iran and detected significant clusters of TB/HIV incidence in Hormozgan. Olfatifar and et al. in two relatively difference study used scan statistic to detect space-time clusters of pulmonary tuberculosis and concluded that there was significant space-time clusters in Hamadan province.[13-15]

The limitations of this study were, lack of access to information of pulmonary tuberculosis cases for 2014 and 2015 years, changes in the resident location of cases during the study and thus, may distort the results, lack of access to appropriate map to geocoding the residents locations addresses of pulmonary tuberculosis cases and use Google Earth to find geographical coordinates and ultimately due to lack of access to the ArcGIS pro software.[16] Results didn't show in three-dimensional view.

This study knows geographic information system as efficient and low-cost means to study distribution of effecting health factors that can be used easily and quickly in different levels of the health system, because has some special characteristics such as: Visualization, speed in analysis, precise and so on, and can reduce the health costs related to diagnosis, treatment, and prevention. The WHO announcement and conducted studies in Iran and other regions of the world and what is in context of tuberculosis and other diseases is consistent with this conclusion.[17-21] Therefore, the use of geographical information systems along with other molecular and spatial methods can be an efficient tool to monitor transfer of certain *Mycobacterium tuberculosis* species and improves access to treatment through population-based programs in rural areas and the slums.[22-24] By portraying information can help epidemiologists and other researchers to obtain relevant information about the epidemic outbreaks and other public health concerns. Finally, it is needed to be mentioned that to better interpretation of results and to better understanding of TB epidemiology more similar studies needed to be conducted with regarding to environmental and spatial influences factors berceuse in some studies has been showed that TB has a seasonality pattern and latitude has effect on this pattern.[25,26]

CONCLUSION

Significantly increasing trend in occurrence of TB cases and existence of hot

spot, especially intensifying hot spots in central areas of province highligh these areas to detect change in epidemiological profile and implement suitable prevention programs.

CONFLICT OF INTEREST

Authors have no conflicts of interest.

ACKNOWLEDGEMENTS

This work has been supported by Student Research Center, Hamadan University of Medical Sciences. The study was funded by Vice-chancellor for Research and Technology, Hamadan University of Medical Sciences with grant number of 931212.

REFERENCES

1. Lopez De Fede A, Stewart J, Harris M, Mayfield-Smith K. Tuberculosis in socio-economically deprived neighborhoods: Missed opportunities for prevention. Int J Tuberc Lung Dis. 2008; 12(12): 1425-30.
2. Koh GC, Hawthorne G, Turner AM, Kunst H, Dedicoat M. Tuberculosis incidence correlates with sunshine: An ecological 28-year time series study. PloS one. 2013; 8(3): e57752.
3. Sia IG, Wieland ML. Current concepts in the management of tuberculosis. Mayo Clinic Proceedings; 2011: Elsevier.
4. WHO. Tuberculosis. Available from: http: //www.who.int/mediacentre/factsheets/fs104/en/.
5. Santos Neto M, Yamamura M, Garcia MCdC, Popolin MP, Rodrigues LBB, Chiaravalloti Neto F, et al. Pulmonary tuberculosis in Sao Luis, State of Maranhao, Brazil: Space and space-time risk clusters for death (2008-2012). Rev Soc Bras Med Trop. 2015; 48(1): 69-76.
6. Nunes C. Tuberculosis incidence in Portugal: Spatiotemporal clustering. Int J Health Geogr. 2007; 6(1): 30.
7. Caprarelli G, Fletcher S. A brief review of spatial analysis concepts and tools used for mapping, containment and risk modelling of infectious diseases and other illnesses. J Parasitol. 2014; 141(05): 581-601.
8. Chan-Yeung M, Yeh A, Tam C, Kam K, Leung C, Yew W, et al. Socio-demographic and geographic indicators and distribution of tuberculosis in Hong Kong: A spatial analysis. Int J Tuberc Lung Dis. 2005; 9(12): 1320-6.
9. Tabatabaee H, Zadeh JH, Nia FY, Akbarpoor MA, Javanmard S, Honarvar B. Spatio-temporal pattern of tuberculosis in the regions supervised by Shiraz University of Medical Sciences 2006-2012. Iran J Public Health. 2015; 44(9): 1253.
10. Create space time cube. Available from: http: //desktop.arcgis.com/en/desktop/latest/tools/space-time-pattern-mining-toolbox/create-space-time-cube.htm.
11. Liu X, Gou F, Ren X, Liu D, Zheng Y, Wei K, et al. Spatial-temporal specific incidence of pulmonary tuberculosis in Gansu, 2009-2013. Zhonghua Liu Xing Bing Xue Za Zhi. 2015; 36(5): 465-9.
12. Yazdani-Charati J, Siamian H, Kazemnejad A, Vahedi M. Spatial clustering of tuberculosis incidence in the North of Iran. Glob J Health Sci. 2014; 6(6): 288.
13. Ghanbarnezhad A, Roustazadeh A, Alizadeh A, Abbastabar H, Nazarnezhad M, Mohseni S. Spatial distribution of TB and HIV co-infection in south of Iran. Ayer. 2015; 4: 177-83.
14. Olfatifar M, Karami M, Hosseini SM, Parvin M. Clustering of pulmonary tuberculosis in Hamadan province, western Iran: A population based cross sectional study (2005-2013). J Res Health Sci. 2016; 16(3): 166-9.
15. Olfatifar M, Hosseini S, Bahrami M, Karami M, Parvin M. Spatiotemporal analysis of pulmonary tuberculosis in Hamadan province: Using population

attributable risk. Iran J Epidemiol. 2017; 13(1): In press.
16. Visualizing the space time cube. Available from: http:// desktop.arcgis.com/en/desktop/ latest/tools/space-time-pattern-mining-toolbox/ visualizing-cube-data.htm.
17. Partilla M. The uses of mapping in improving management and outcomes of tuberculosis control programs: An overview of available tools. Manag Sci Health. 2008; 24(6): 1-18.
18. Goswami ND, Hecker EJ, Vickery C, Ahearn MA, Cox GM, Holland DP, et al. Geographic information system-based screening for TB, HIV, and syphilis (GIS-THIS): A cross-sectional study. PloS one. 2012; 7(10): e46029.
19. Osooli M, Haghdoost A, Yarahmadi S, Foruzanfar M, Dini M, Holakouie Naieni K. Spatial distribution of congenital hypothyroidism in Iran using geographic information system. Iran J Epidemiol. 2009; 5(1): 1-8.
20. Ricketts TC. Geographic information systems and public health. Annu Rev Public Health. 2003; 24(1): 1-6.
21. Rushton G, Elmes G, McMaster R. Considerations for improving geographic information system research in public health. Urisa-Washington DC. 2000; 12(2): 31-50.
22. Moonan PK, Bayona M, Quitugua TN, Oppong J, Dunbar D, Jost KC, et al. Using GIS technology to identify areas of tuberculosis transmission and incidence. Int J Health Geogr. 2004; 3(1): 23.
23. Richardson M, van Lill SW P, van der Spuy GD, Munch Z, Booysen C, Beyers N, et al. Historic and recent events contribute to the disease dynamics of Beijing-like *Mycobacterium tuberculosis* isolates in a high incidence region. Int J Tuberc Lung Dis. 2002; 6(11): 1001-11.
24. Tanser F, Wilkinson D. Spatial implications of the tuberculosis DOTS strategy in rural South Africa: A novel application of geographical information system and global positioning system technologies. Trop Med Int Health. 1999; 4(10): 634-8.
25. MacLachlan JH, Lavender CJ, Cowie BC. Effect of latitude on seasonality of tuberculosis, Australia, 2002–2011. Emerg Infect Dis. 2012; 18(11): 1879-81.
26. Parrinello C, Crossa A, Harris T. Seasonality of tuberculosis in New York City, 1990-2007. Int J Tuberc Lung Dis. 2012; 16(1): 32-7.

How to cite the article: Hosseini SM, Parvin M, Bahrami M, Karami M, Olfatifar M. Pattern mining analysis of pulmonary TB cases in Hamadan province: Using space-time cube. Int J Epidemiol Res. 2017; 4(2): 111-117.