SHORT COMMUNICATION

SPATIAL CLUSTER ANALYSIS OF NEW AND RELAPSED CASES OF PULMONARY TUBERCULOSIS BY DISTRICT: PAKISTAN 2015

Masood Ali Shaikh, Naeem Abbas Malik*

Independent Consultant, Karachi, *Institute of Geo-Information and Earth Observation, PMAS Arid Agriculture University Rawalpindi-Pakistan

Tuberculosis (TB) is a major public health problem in Pakistan. There are no studies in Pakistan on TB disease mapping and spatial analysis at the country level. This study was conducted to map and identify clusters of TB incidence rates at the district level for all provinces of Pakistan, including Islamabad for all new and relapsed cases of pulmonary TB, for the year 2015. The district level TB rate for new and relapsed cases ranged from 4.7 to 422.6 per 100,000. Global Moran's I was 0.25238, with a pseudo *p*-value of 0.001, indicating clustering in the data. Local Moran's I, i.e., Local indicators of Spatial Association identified 11 districts as high-high clusters, and 20 districts as low-low clusters. Better understanding of these hot and cold spot districts would be helped by the availability and analysis of TB data at the more granular level of Union Councils in the country.

Keywords: Tuberculosis; Epidemiology; GIS; Pakistan

Citation: Shaikh MA, Malik NA. Spatial cluster analysis of new and relapsed cases of pulmonary Tuberculosis by district: Pakistan 2015. J Ayub Med Coll Abbottabad 2019;31(2):293–5.

INTRODUCTION

Tuberculosis (TB) is a chronic bacterial, infectious airborne disease that can affect any part of the human body, mainly the lungs.[1] "Ensure healthy lives and promote well-being for all at all ages" is goal number 3 of the Sustainable Development Goal (SDG), which entails 13 targets.[2] The target 3.3 states that "By 2030, end the epidemics of AIDS, tuberculosis, malaria and neglected tropical diseases and combat hepatitis, water-borne diseases and other communicable diseases".[2]

According to the World Health Organization's latest report on TB, the 'Global Tuberculosis Report 2017': globally, TB was the ninth leading cause of mortality, with an estimated 10.4 million people succumbing to TB in 2016.[3] However alarmingly, 56% of these cases were from five countries that included Pakistan.[3] The 2016 Pakistan TB mortality rate per 100,000 populations was estimated to be 23 with uncertainty interval of 18–29 in 2016; this rate excludes individuals with combined HIV and TB mortality.[3] While the 2016 Pakistan incidence rate per 100,000 population was estimated to be 268 with uncertainty interval of 174 to 383; this rate includes individuals with combined HIV and TB.[3] Few studies have been published from Pakistan on the use of Geographic Information Systems (GIS) for disease mapping and aspatial analysis in the context of TB. However, these studies have been limited to the city of Karachi only, and do not study clustering of TB incidence rates by administrative subdivisions.[4–6]

Spatial epidemiology harnesses the power of Geographic Information Systems for visualizing and analysing geographic variations in the distribution of diseases. There are no studies in Pakistan on TB disease mapping and spatial analysis at the country level. This study was conducted to map TB incidence rates at the district level for all provinces of Pakistan, including Islamabad for all new and relapsed cases of pulmonary TB, either laboratory or clinically diagnosed cases, for the year 2015, i.e., the year for which latest data are made publicly available. To better understand TB disease epidemiology in terms of distribution and clustering of TB incidence rates at the district level in Pakistan. No ethical approval was required for this study; as only publicly available data were used for analysis.

METHODS AND RESULTS

Pulmonary Tuberculosis (TB) data entailing the number of all new and relapsed cases (bacteriological confirmed or clinically diagnosed) by district, reported during the year 2015, were downloaded from the National Tuberculosis Control Program's (NTCP) website.[7] Year 2015 is the latest year for which data are available and freely downloadable in Excel format. TB data were joined with spatial file of districts, using statistical analysis program R version 3.4.2, mapping was done using ArcGIS 10.5. While GeoDa version 1.12 was used for spatial autocorrelation analysis. Since the TB data for geographical regions for the disputed Kashmir area were not available, we excluded this data from inclusion. However, Azad Kashmir data were included. While the TB data from FATA although available, could not be joined with the spatial data – as the geographical regions by which TB data were available did not match with the GIS file – and were also excluded. In total, 132 districts including Islamabad were used for the analysis.

The TB data file downloaded from NTCP also included population for each district, which was used as a denominator for the calculation of incidence rate for the year 2015. While all new and relapsed cases of pulmonary TB, either laboratory or clinically diagnosed cases, were used as numerator for the calculation of cases per 100,000 cases by district. Proportion of bacteriologically confirmed cases as a percent of all TB cases reported was also calculated. With first order queen contiguity weights, univariate global Moran's I and local Moran's I were used to determine evidence for spatial clustering at the country level, and the locations of clusters, respectively. The district level TB rate for new and relapsed cases ranged from 4.7 to 422.6 per 100,000 populations. The number of districts where proportion of all new and relapsed cases of pulmonary TB were laboratory (bacteriologically) confirmed was 33.0% or less, was ten. While in 94 districts this proportion ranged from 33.01% to 66.0%. In 28 districts, this proportion ranged from 66.01–100%. However, out of 132 districts in 2015, only one district this proportion was 100%. Figure-1 shows the map of new and relapsed cases of pulmonary TB incidence rates per 100,000, by district as well as the proportion of laboratory confirmed cases, in 2015.

Global Moran's I was 0.25238, with a pseudo p-value of 0.001 with permutation test set at 999. Local Moran's I, i.e., Local indicators of Spatial Association (LISA) identified 11 districts as high-high clusters, and 20 districts as low-low clusters. As spatial outliers, 8 districts were identified as low-high clusters and 3 districts as high-low clusters. LISA identified 26 districts with a statistical significance for spatial clustering/outliers at the *p*-level of <0.05, 8 districts at the *p*-level of <0.01 and another 8 districts at the level of <0.001. Figure-2 shows the output of global and local Moran's I analysis.

The right map shows clusters, left map shows statistical significance of each cluster, and the Global Moran's I scatter plot is shown in the bottom middle.

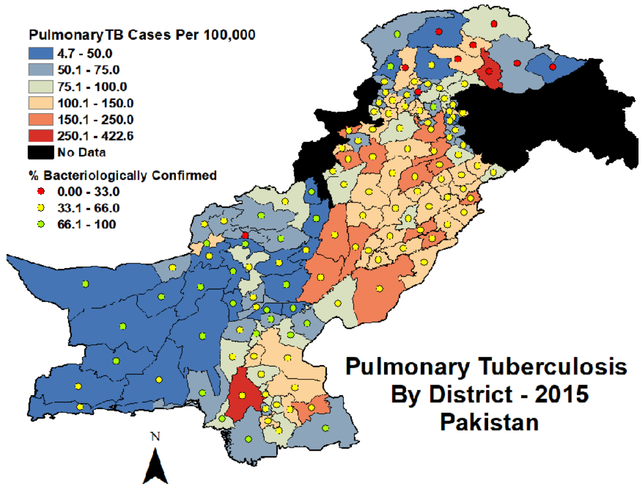

Figure-1: Pulmonary tuberculosis incidence rates and proportion of laboratory confirmed cases, by district in 2015.

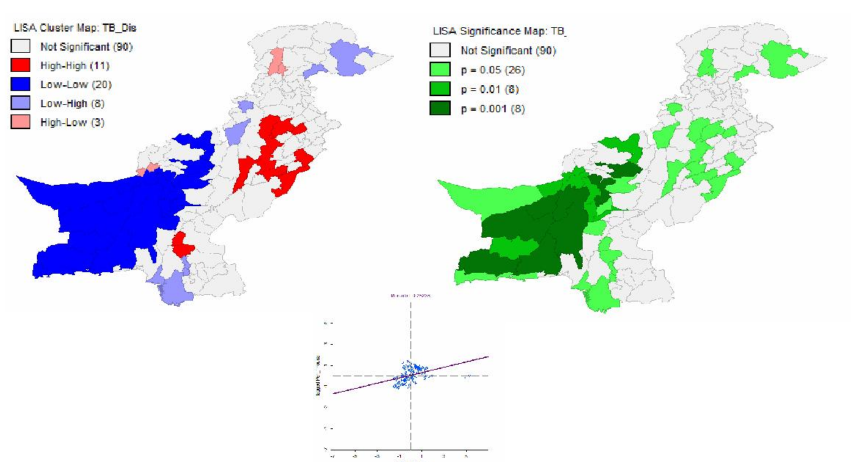

Figure-2: Local indicators of Spatial Association (LISA) maps and Global Moran's I scatter plot

DISCUSSION

Geographic Information Systems (GIS) and spatial analysis help well understand disease epidemiology in terms of distribution and clustering by place/space. Clustering denotes a pattern in data that would not exist under the assumption of spatial randomness. Mapping of 2015 district level data for new and relapsed cases of pulmonary TB show lower incidence rates per 100,000 in most Balochistan districts and higher rates in Punjab districts. In most Balochistan districts, proportionally more TB cases were bacteriologically confirmed than in any other province. Detection of district level incidence rate hot and cold spots, i.e., contiguous districts with higher or lower incidence rates, was done by cluster analysis. The high-high cluster, i.e., contiguous districts with high incidence rate were identified in the nine, north-central, central, and south-central districts of Punjab that included Bahawalnagar, Hafizabad, Jhang, Khushab, Muzafargarh, Okara, Sargodha, Vehari, and Khannewal districts. While a cluster of two high-high districts in Sindh were identified that included Naushahro Feroz, and Shaheed Benazirabad districts. Out of the 30 Balochistan districts, 19 formed a cluster of low-low, i.e., contiguous districts with lower incidence rates. This cluster included one Sindh district of Shahdad Kot that is contiguous i.e. shares border with Balochistan.

Spatial epidemiological analysis helps identify patterns/clustering. However, once clusters are identified, there is a need to better understand as to why these places/districts are different in some respect. Robust, effective, and better managed TB surveillance systems identifying cases in some districts versus others could be one possible explanation. As identification of fewer incident/replaced cases does not necessarily suggest lower number of such cases in the cold-spots districts.

In this study, we used LISA for identification of clusters. However, other methods like Kulldorff's spatial scan statistic are also available. Since these cluster detecting methods mathematically differ, they may produce different results. Future studies need to deploy varied spatial cluster methods and factor element of time in the analysis as well, i.e., spatio-temporal analysis. This would help better understand trends over space as well as time. Finally, for more granular analysis, availability of data for TB cases and baseline at-risk populations at the Union Council level would provide even better insights and help reduce the TB burden in the country.

REFERENCES

1. Basic TB Facts. Tuberculosis CDC. [Internet]. 2018 [cited 2018 April 29]. Available from: https://www.cdc.gov/tb/topic/basics/default.htm
2. Health - United Nations Sustainable Development. [Internet]. [cited 2018 April 29]. Available from: https://www.un.org/sustainabledevelopment/health/.
3. Global tuberculosis report 2017. [Internet]. [cited 2018 April 29]. Available from: http://www.who.int/tb/publications/global_report/en/
4. Burke F, Hamza S, Miandad M, Khan I, Nawaz-ul-Huda S, Naqvi SG. Ecological Analysis of Tuberculosis Patients' – A Case Study of Karachi. Am J Health Res 2016;4(3):39–45.
5. Miandad M, Burke F, Nawaz-ul-Huda S, Azam M. Geodemographic analysis of tuberculosis patients in Karachi, Pakistan. Hum Geogr--J Stud Res Hum Geogr 2015;9(2):166–82.
6. Miandad M, Burke F, Nawaz-ul-Huda S, Azam M. Tuberculosis incidence in Karachi: A spatio-temporal analysis. Geogr-Malays J Soc Space 2014;10(5):1–8.
7. National TB Control Program-Pakistan. National TB Data. [Internet]. [cited 2018 Mar 11]. Available from: http://ntp.gov.pk/national_data.php

Received: 25 June, 2018	Revised: --	Accepted: 30 December, 2018

Address for Correspondence:
Masood Ali Shaikh
Email: masoodali1@yahoo.com