Journal of Sciences, Islamic Republic of Iran 33(2): 149 - 156 (2022) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

Statistical Analysis of Spatial Distribution of Tuberculosis in Niger State from 2016-2020

S.O. Balogun, E.S. Oguntade, D.M. Oladimeji*

Department of Statistics, Faculty of Sciences, University of Abuja, Abuja, Nigeria

Received: 20 December 2021 / Revised: 25 June 2022 / Accepted: 20 July 2022

Abstract

Tuberculosis (TB) remains one of the 21st century's crucial public health problems. Today, it is the second leading cause of death from a single infectious disease agent. Spatial information on health-related diseases like TB is not well documented in Nigeria. The study of Spatial Distribution of TB in Niger State is necessitated for broad and more comprehensive coverage of TB cases within the State. As such, there is a need to carry out a spatial auto-correlation analysis of the entire 25 Local Government Areas (LGAs) of the State to explore the epidemiological data, distribution, and pattern of infection from 2016 to 2020. This study investigated the spatial auto-correlation structure with the Moran Index, Anselin's local Moran's I, Getis-ord local statistics and kriging interpolation methods. The result of the Analysis of Variance (ANOVA) showed that there was a significant difference in infection patterns between LGAs. This result implied variability in the number of cases across different local governments within the study periods. Although the considerable number of TB-reported instances, the analysis based on Moran's I and global G revealed no clustering pattern in the data. The local indicator of spatial association result showed that the number of cases was random across the LGAs; there were no significant clusters. Kriging interpolation method identified Kontogora, Bosso and Chanchaga LGAs as hotspots of TB cases in Niger State. This study provides information that can assist policymakers in rationally planning targeted interventions to effectively control TB while addressing the underlying socio-economic risk factors in Niger State. This study has shown that despite data limitations, Geostatistical approaches are viable for understanding interpolation patterns.

Keywords: Moran's I; Spatial auto-correlation; Global statistic; Kriging; Tuberculosis.

Introduction

Tuberculosis (TB) is a communicable disease caused by the bacterium *Mycobecterium Tuberculosis.* Today, it is the second leading cause of death from a single infectious disease agent after malaria [1]. Persons who are ill with pulmonary TB are often infectious and can spread the disease through coughing, sneezing or simply talking, as these acts propel TB into the air.

Globally, there were about 1.7 billion estimated

* Corresponding Author: Tel: 2347032325611; Email:oladimejidamilare@yahoo.com

cases of TB in 2018 with over 10 million incident TB cases and about 1.5 million TB deaths [2]. These cases and deaths were rampart among the poor and marginalized communities with deprived access to good medical facilities and health cares. Over 25% of these deaths occurred in Africa in which Nigeria and other sub-Saharan African countries are central. TB is an infectious disease that spread rapidly from person to person through coughing and sneezing especially in crowded areas and communities. Nigeria, South Africa and five other countries of the world accounted for about 64% of the global TB cases in 2016 [3].

There has been an increasing interest in the application of spatial data analysis methods to the study of the spatial distribution of disease risk and occurrence called spatial epidemiology [4]. The spatial epidemiology covers three main essentials which are; visualization of spatial data using map (disease mapping), exploration of patterns and relationships and modelling of spatial data in such a way to test hypothesis and search for explanations and relationships using statistical models [5].

Studies have explored the usage of spatial autocorrelation indices and interpolators (such as kriging and inverse distance weight) to identify cluster patterns and prediction in different locations respectively for diseases [6]. The application of statistical tools that measure differences between two closed pixels and their means to give measure of local similarities or dissimilarity in health and environmental sciences are common phenomena. Ibrahim *et al.* (2015) identified clusters and spatial patterns in TB prevalence using spatial autocorrelation indices [7]. Likewise, the most commonly used spatial interpolator (kriging) was used due to the fact that it minimizes variance estimates through weighted moving average [8]. For instance, Singh *et al*. (2014) used different kriging methods to predict the number of people infected by dengue fever in four districts in India [6].

Spatial epidemiology has therefore provided an opportunity to better understand geographical variation of disease, explore local risk factors that may be responsible for the observed pattern, identify and detect disease clusters or high risk areas that may require targeted intervention for control of the disease and by using advanced spatial statistical analysis can predict future occurrence of the disease and how different strategies will impact the control of diseases [5]. The study of spatial distribution of TB in Niger State is necessitated as a result of the study carried out in Suleja and Minna which are mostly of relatively small size of TB patients [9]. As such there is a need to carry out spatial auto-correlation analysis of the entire 25 LGAs of the State with the aims of statistically analyzing the TB's spreading, prevalent and dissimilarity pattern from 2016 to 2020. Therefore, this study aimed at exploring suitable spatial statistical techniques to analyze TB cases and pattern in Niger State, Nigeria.

The remainder of this article is structured as follows. Materials and methods including study area, source of data and spatial analysis methods such as spatial autocorrelation indices and Kriging Interpolation technique are described in Sections 2. Also, Section 3 presents the results and discussion of findings. Finally, the study concludes in Section 4.

Materials and Methods

1. Study Area

Niger State is situated in the north-central geopolitical zone of Nigeria with Minna as its capital city. Other major cities in the State include Suleja, Bida and Kontagora. Established in 1976, Niger State was created out of the defunct North-Western States. It is the largest State in Nigeria with a vast land mass of about 76,363km^2; approximately 8.6 million hectares constituting about 9.3% of the total land area of Nigeria. Lying on latitude 3.20° East and longitude 11.30° North, the State shares a country border with the Republic of Benin (West) and bounded to the south by the river Niger. The state borders with Niger State include the Federal Capital Territory (FCT), Zamfara, Kebbi, Kwara and Kaduna. The official population of Niger State was 3,954,772 people at the 2006 population and housing census [10]. The vegetation of Niger state is Savannah. The climate of the study area is similar to that of FCT, Abuja which is tropical: non-arid climate with only two seasons throughout the year: wet and dry [11].

2. Data Source

The data used for this study were reported TB cases in the Niger State, Nigeria. The data were collected from the National TB Control Program, Nigeria. It was based on yearly record of the reported cases for the state aggregated as a single entity representing each of the 25 LGAs. Specifically, the data for the year 2016, 2017, 2018, 2019 and 2020 were used for spatial pattern analysis of TB prevalence. The data can be accessed through Nigeria Federal Ministry of Health via these link www. health.gov.ng or www.leprosy-information.org.

3.Global Spatial Autocorrelation

Using the available data, the spatial autocorrelation analysis was performed. The contiguity-based

neighbour is defined as first-order queen's contiguity, such that only local government sharing borders with each other were considered neighbours. The row standardized weight matrix is created and the Moran's I test for global spatial autocorrelation was determined at p-value ≤0.05. The Global Moran's I statistic is used to test the null hypothesis that there is no spatial autocorrelation of the TB cases, The Global Moran's I is given as: (Equation 1).

$$I = \frac{1}{S_x^2}.\frac{\sum_{i=1}^{n}\sum_{(j=1;i\neq j)}^{n} w_{ij}\,(x_i-\bar{x})(x_j-\bar{x})}{\sum_{i=1}^{n}\sum_{(j=1;i\neq j)}^{n} w_{ij}} \quad (1)$$

where, $S_x^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$, $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ and $w_{ij} = \begin{cases} 1, if\ i,j\ are\ adjacent\ neighbours, \\ 0, \qquad otherwise \end{cases}$

w_{ij} is the spatial weight between feature i and j, x_i is the number of TB cases and n is the total number of TB cases at the geographical locations. Global Moran's I test was carried out under the assumptions of randomization, normality and Monte Carlo simulation [12]. The value of the Global Moran's I ranged from -1 to +1. A value close or equal to +1 indicate a positive spatial autocorrelation (spatial clustering) while a value close to or equal to -1 showed a negative spatial autocorrelation while a value of 0 signified no spatial autocorrelation (complete spatial randomness). The significant global autocorrelation occurred when the p-value is less than 0.05. The Z score is calculated as $Z(I)=\frac{I-E(I)}{\sqrt{Var(I)}}$; where $E(I)=\frac{-1}{(n-1)}$. After the Z score is obtained from Z(I) and E(I), then a critical value is tested. This test will reject the initial hypothesis if $|Z(I)|>Z_{(\alpha)}$.

4.Local Indicator of Spatial Autocorrelation

Anselin (1995) proposed the use of Local Indicators of spatial association (LISA) for a localized form of exploratory spatial analysis [13]. Autocorrelation for spatial data can also be measured at the local level to evaluate the extent of autocorrelation within local neighborhoods. The importance of the LISA test was to detect or identify regions of spatial clustering around an observation. Local measures capture some local spatial variation and spatial dependency while global measurements provide only one set of values the represent the extent of spatial autocorrelation across the entire study area [14]. Local Moran's I is computed using (Equation 2).

$$I_i = \frac{x_i-\bar{x}}{{S_i}^2}\sum_{i=i\ j\neq i}^{n} w_{ij}\,(x_i-\bar{x}) \quad (2)$$

where x_i and x_j are the attribute for feature i and j, $\bar{x}$ is the mean of corresponding attribute, w_{ij} is the spatial weight between feature i and j, n is the number of features, and

$S_i^2 = \frac{\sum_{j=1 j\neq i}^{n} w_{ij}}{n-1} - \bar{x}^2$.

The advantage of local spatial autocorrelation is that it can identify spatial clusters at a more local spatial unit than relying on Global Moran's I statistics which is a single global autocorrelation measure.

5.Getis-Ord Local Statistic

Local G statistics or Getis-Ord (Gi*) is used to find out the spatial association of the high and low values of the feature. Ord and Getis (1995) introduced a family of G statistics, that can be used as measures of spatial association in a number of circumstances [15].

Getis-Ord local Statistic is used to detect the presence of local clustering, it also provides additional information about the intensity and stability of core hotspot/cold spot clusters [16]. The Z score assigned to each district identified the presence and intensity of local clusters of hot sport and cold spot of TB incidence relative to the hypothesis of spatial randomness. The Getis-Ord Gi index is calculated using (Equation 3).

$$Gi^* = \frac{\sum_{j=1}^{n} wijxj - X\sum_{j=1}^{n} wij}{S\sqrt{\frac{\left[n\sum_{j=1}^{n} Wij^2 - \left(\sum_{j=1}^{n} Wij\right)^2\right]}{n-1}}} \quad (3)$$

Where x_j is TB incidence for LGA j, w_{ij} is the spatial weight between district i and j, n is the total number of district and $\bar{x} = \frac{\sum_{j=1}^{n} x_j}{n}$ and $S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n}} - (\bar{x})^2$

6. Kriging Interpolation Technique

The method of kriging is used extensively in spatial interpolation problem and has found many applications in geology, mining industries, engineering, environmental sciences and medicines [17]. According to Lundberg (1999), kriging is a Geostatistical technique for spatial interpolation or prediction that gave the best unbiased linear predictor based on the observation of a spatial process [18]. It is a statistical method for interpolating the value of a random field at an unobserved sites or locations based on available surrounding measurements or observed spatial data sets [6]. Ordinary Kriging was preferred to other types of kriging because it predict an estimate for un-sampled location by assuming a constant mean in the local neighborhood of each estimation location, which is a characteristic of focal diseases like TB. It uses a semi-variogram model to measure spatial autocorrelation between pairs of prevalence rate as follows [19].

$$Y_h = \frac{1}{2n_h}\sum_{i=1}^{n_h}\{Z_x - Z_{x+h}\}^2 \quad (4)$$

where n_h is the number of pairs of sample locations, Z_x and Z_{x+h} are the prevalence rates at any two locations x and $x+h$ separated by distance h,

calculations of Y_h are repeated for 2h, 3h,4h, …, kh.

Results

1. Distribution and Prevalence Pattern of Tuberculosis Cases in Niger State

Table 1 presents the summary statistics of TB cases in Niger State. In 2016, the total number of TB cases recorded was 1678 across the twenty-five LGAs of Niger state. The highest number of cases was recorded in Chanchaga LGA while the lowest number of cases was in Edati LGA. In the year 2020, the total number of cases recorded was 3630 with 146 cases as the average number of cases for the year. The number of cases recorded were on the increase within the study period (Table 1).

Figure 1 displays the box-plot of cases of TB over time. The box plot indicated the number of reported cases across the years for the study area. This raw data together with corresponding population data were used to obtained the TB prevalence. The patterns were alternating over the years. There were few reported cases in 2017 which might due to medical treatment or awareness campaign against the disease by various government or private corporations. But in 2020, the prevalence escalated to over 500 cases (Figure 1).

Table 2 shows the years and the prevalence cases of TB over the years in Niger State. The prevalence of TB in the year 2016 was 0.0304 while the prevalence of TB in year 2020 was 0.0575. Based on this result, the highest prevalence of the TB in Niger State was observed in the year 2020 while the lowest prevalence of TB was in the year 2016. This result depicts an upward increase in the prevalence of TB in Niger state which may be due to a number of factors. These factors may range from traditional belief, in-adherence to medications, access to adequate medical facilities among others.

The results as revealed in Table 3 indicated a significant difference between LGAs and the number of TB cases across the different year (p-value <0.001). This implied that the number of TB cases varies from LGAs to LGAs for different years under study.

2. Spatial Autocorrelation Analysis and Kriging Interpolation of Tuberculosis Cases in Niger State

The result of Moran's I across the different period showed a non-clustering pattern in the prevalence of TB

Table 1. Distribution of Tuberculosis Cases in Niger State (2016-2020)

Year	2016	2017	2018	2019	2020
Min	10	6	17	29	28
Max	283	272	241	298	505
Mean	68.375	79.625	82.792	114.458	145.75
Median	44	45	52	74	77
Standard Deviation	63.041	70.762	67.519	84.505	140.576
Total Case	1678	1964	2042	2864	3630

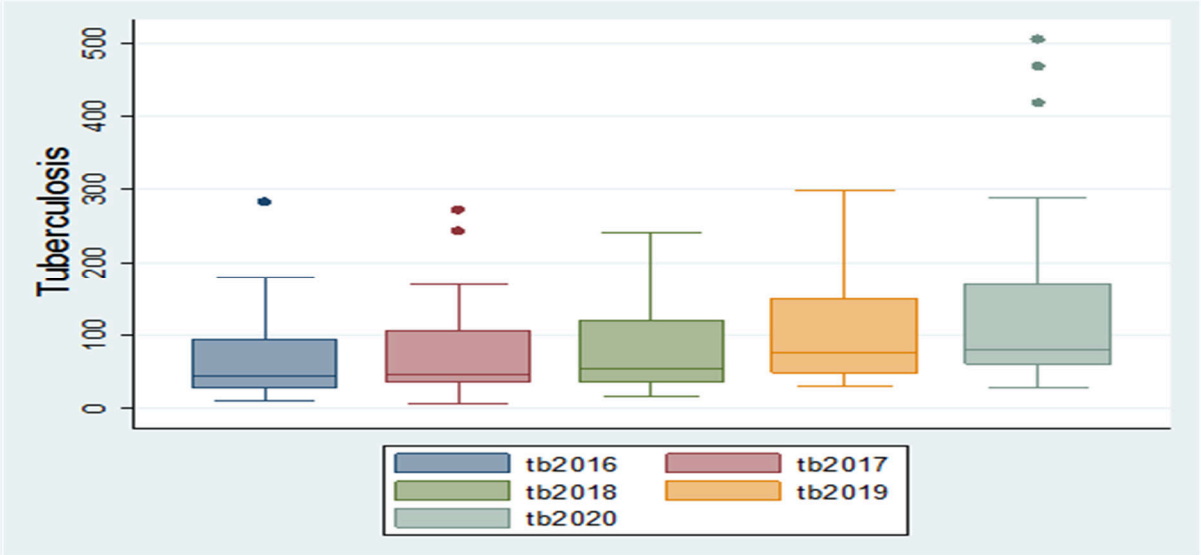

Figure 1. Box-Plot of the Distribution of Tuberculosis in Niger State (2016-2020)

Table 2. Prevalence cases of Tuberculosis (2016-2020)

Year	Prevalence
2016	0.030
2017	0.034
2018	0.035
2019	0.047
2020	0.058

across various LGAs. The observed values of Moran' I was all negative for the years under consideration (Table 4). The value of spatial autocorrelation cannot display spatial pattern information in a particular area so that a visual display of the tendency for spatial relationship is needed by using the LISA.

Cluster Detection and Hotspot Mapping Using LISA measures characterize individual units as to whether cluster or non-clustering pattern occurs in TB dataset. Figure 2 reveals that all the number of cases were random across the LGAs. There were no significant clusters as shown by the LISA.

Local Indicator of Spatial Association

Table 5 gives the result of Gestid-Ord, a global measure of spatial autocorrelation. Z statistic and P-values were used as a measure of statistical significance. The results indicated no clustering patterns in the number of TB cases across the various LGAs in Niger State since all the p-value exceeded the 5% level of significance.

Table 3. Comparing cases of Tuberculosis across the Local Government and Years

ANOVA Source of Variation	SS	Df	MS	F	P-value
LGA	732590.100	24	30524.589	14.640	<0.001
YEAR	102456.100	4	25614.032	12.2848	<0.001
Error	200162.300	96	2085.024		
Total	1035209	124			

ANOVA: Analysis of Variance; SS: Sum of Squares; MS: Mean Square; Df: Degree of Freedom.

Table 4. Testing Global Measures for Spatial Autocorrelation of Tuberculosis Prevalence (2016-2020)

Moran'i	2016	2017	2018	2019	2020
Moran'si	-0.029	-0.053	-0.069	-0.081	-0.101
Expected Index	-0.042	-0.042	-0.042	-0.042	-0.042
Standard Deviation	0.037	0.040	0.041	0.042	0.037
Z Score	0.351	-0.275	-0.659	-0.929	-1.590
P-Value	0.736	0.780	0.506	0.352	0.093

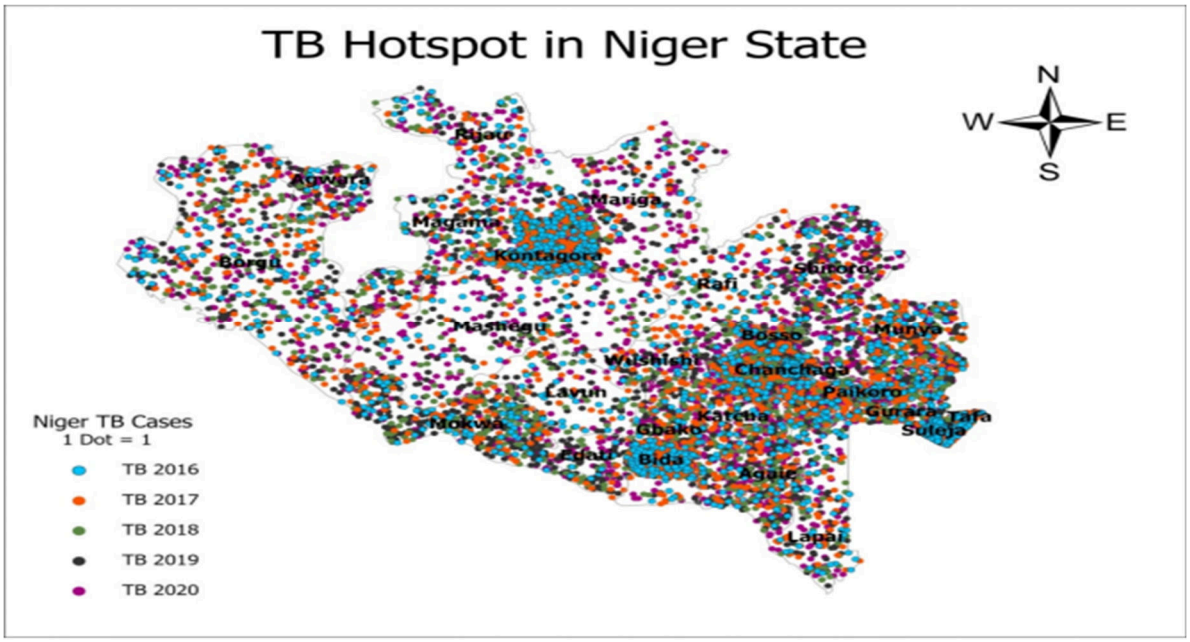

Figure 2. A Graphical Display of Tuberculosis Hot spot in Niger State

Table 5. Gestid-Ord or General G

G. ord	2016	2017	2018	2019	2020
Observation.G	0.133	0.217	0.301	0.253	0.333
Expected. G	0.187	0.187	0.187	0.187	0.187
Variance	0.023	0.027	0.027	0.030	0.024
Z score	-0.355	0.183	0.695	0.381	0.942
P-value	0.724	0.827	0.491	0.673	0.115

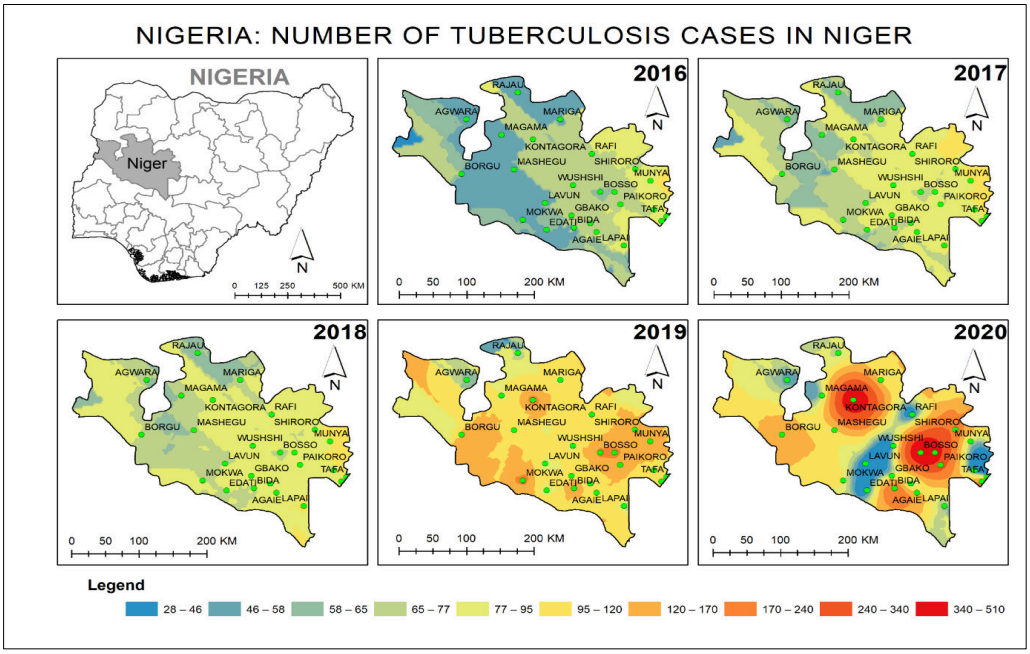

Figure 3. Spatial Map of Tuberculosis Cases in Niger State Nigeria (2016-2020)

Kriging Interpolation

The total number of 12178 cases of TB were observed from 2016-2020 in all the LGAs. Figure 3 reveals the continues surface generated by ordinary Kriging of TB prevalence in the study areas. The generated surface revealed an increasing TB cases over the years. The increased in TB cases across the State became noticeable in the year 2019. In year 2020, the map indicated TB hotspots in Niger State. These hotspots include Kontogora, Bosso, Chanchaga and Masehegu LGAs while Mokwa, Lavun and Tafa LGAs were non-hotspot location in the State with low TB cases. The overall view of the spatial map suggests a no spatial clustering pattern in the study areas within the study period.

Discussion

The issue that emerged from this study is the reliability of global measures to test clustering pattern using disaggregated data. The use of global measure has been widely used in health research to detect spatial pattern and clusters [7]. In this study, ANOVA, Moran's I, Gestis-Ord and Kriging were employed. The result of the ANOVA showed that there was a significant difference between LGAs which implies that the number of TB cases varies from LGAs to LGAs. The plausible explanation for this may be due to the presence of natural bodies in some LGAs while others may not have such features. These discrepancies in TB cases across different LGAs could be due to the geographical features of the villages where topography like hill and water body play a vital role in the disease transmission. The common natural features in Niger State includes presence of water body (River Niger) in Borgu, water falls in Gurara and mountains in Suleja LGA while places like Kontagora has a plain ground. Study has identified such features as some of the major determinants of disease incidence and disease spread [20].

Despite significant number of TB prevalence cases in some of the LGAs, the analysis of Moran's I and global G did not reveal any clustering pattern for the data. This result disagrees with a related study in Kebbi State, Nigeria [7]. The authors reported a low clustering patterns of TB cases in Kebbi State, Nigeria. Likewise,

a similar study in south western Nigeria also observed an alternating pattern in infection pattern which agrees with this current study in Niger State [21].

There were increased in TB cases across the State with highest cases predicted in LGAs with high population across the state like Kontogora, Bosso and Chanchaga while LGAs like Mokwa, Lavun and Tafa with low population densities had low to medium predicted TB cases. The overall view of the spatial map suggests a no spatial clustering pattern in the study areas within the study period. This result aligns with the one reported in Kuala Lumpur, Malaysia [22]. The authors identified a high populated area as the hotspot of TB cases. The possible explanation for this may be due to lack of proper medical awareness, poverty (as majority of the inhabitants are peasant farmers and local fishermen) and cultural belief among others. The authors have identified this in literature as a big determinant to curb the disease incidence and spread [11].

Conclusion

This study involves the use of Geostatistical approach and spatial analyses which have been applied to many epidemiological researches to analyze and more clear display of spatial pattern of TB incidence. Like other State, Niger State is also experiencing cases of TB. The Spatial interpolation model for predicting the TB prevalence employed in the study was for identifying spatial pattern of TB. This study has shown that despite data limitation, Geostatistical approaches are quite viable for the understanding interpolation pattern via the use of Arc-GIS. Thus, the outcome of this study is quite imperative for enhancing policy and decision-making health service provision. The result of the study suggests that there was heterogeneity of spatial pattern of TB within the LGAs. The findings, in terms of the presence of hot spot of TB in the LGAs can help government and health officer to improve on TB finding in Niger State by expansion of quality assured TB microscopy services to all LGAs taking into consideration the terrain and distribution of these services to ensure equitable access to the Arc-GIS diagnostic services.

Despite the success of the study of TB prevalence cases, we recognized the limitation of the dataset is still something to worry about. The lack of monthly cases of TB in each local government and the coordinate locations for each case is one of the starting points.

References

1. Zumla A, George A, Sharma V, Herbert RHN, Oxley A, Oliver M. The WHO 2014 global tuberculosis report—further to go. Lancet Glob Health. 2015;3(1):e10-e12.
2. Centre for Disease Control and Prevention, Global Health Tuberculosis. Available online from: http//www.cdc.gov.globalhealthnewstoom/topic/tb/index.html(Accessed on 2nd March, 2022).
3. World Health Organization Tuberculosis (TB). WHO Africa. Available from: https//www.afro.who.int.health-topic/tuberculosis-tb//(Accessed on 2nd March, 2022).
4. Ostfeld RS, Glass GE, Keesing F. Spatial epidemiology: an emerging (or re-emerging) discipline. Trends Ecol Evol. 2005;20(6):328-336.
5. So FFM, Lai P. Spatial epidemiology of asthma in Hong Kong. GIS for Health and the Environment. Springer; Chapter 2038. 2007;154-169:.
6. Singh DS, Saranya S, Mary MI. Kriging Approach for an Epidemic Data Analysis. Int J Sci Innov Math Res. 2014;2(6):530-536.
7. Ibrahim S, Hamisu L, Gulma UL. Spatial Pattern of Tuberculosis Prevalence in Nigeria: A Comparative Analysis of Spatial Autocorrelation Indices. Am J Geogr Inf Syst. 2015;4(3):87-94.
8. Fu WJ, Jiang PK, Zhou GM, Zhao KL. Using Moran ' s I and GIS to study Sthe spatial pattern of forest litter carbon density in a subtropical region of southeastern China. Biogeosciences. 2014;11(8):2401–2409.
9. Sani RA, Garba SA, Oyeleke SB, Abalaka ME. Prevalence of Pulmonary Tuberculosis (PTB) in Minna and Suleja Niger State, Nigeria. Am J Med. 2015;5(6):287-291.
10. National Population Commission's Population Figures for the Nigeria States for 2006 Population and Housing Census, 2006; NPC: Abuja, Nigeria.
11. Oguntade ES, Shohaimi S, Nallapan M, Lamidi-sarumoh AA, Salari N. Statistical Modelling of the Effects of Weather Factors on Malaria Occurrence in Abuja, Nigeria. Int J Environ Res Public Health. 2020;17(3474):1–12.
12. Comber A, Brunsdon C. A spatial analysis of plant phenophase changes and the impact of increases in urban land use. Int J Climatol. 2015;35(6):972-980.
13. Anselin L. Local indicators of spatial association—LISA. Geographic Analys. 1995;27(2):93-115.
14. Mathur M. Spatial autocorrelation analysis in plant population: An overview. J Appl Nat Sci. 2015;7(1):501-513.
15. Ord JK, Getis A. Local spatial autocorrelation statistics: distributional issues and an application. Geographic Analys. 1995;27(4):286-306.
16. Anselin L, Getis A. Spatial statistical analysis and geographic information systems. Ann Region Sci. 1992;26(1):19-33.
17. Salahi-Moghaddam A, Khoshdel A, Dalaei H, Pakdad K, Nutifafa GG, Sedaghat MM. Spatial changes in the distribution of malaria vectors during the past 5 decades in Iran. Acta Tropica. 2017;166:45–53.
18. Lundberg L. Ordinary Kriging Prediction Intervals Based on SemiparametricBootstrap. J Math Stat. 2019;22.
19. Luo W, Taylor M, Parker S. A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales. Int J Climatol. 2008;28(7):947-959.

20. Haddawy P, Hasan AHMI, Kasantikul R, Lawpoolsri S, Sa-angchai P, Kaewkungwal J, et al. Spatiotemporal Bayesian networks for malaria prediction. Artif Intell Med. 2018;84:127–138.
21. Cadmus CA, Agada II, Onoja IS. Risk Factors Associated with Bovine Tuberculosis in some Selected Herds in Nigeria. Trop Anim Health Prod. 2010;42:547-549.
22. Rajab NA, Hashim N, Rasam ARA. Spatial Mapping and Analysis of Tuberculosis Cases in Kuala Lumpur, Malaysia, 2020; IEEE 10th International Conference on System Engineering and Technology (ICSET), 38-43.