Geo-Spatial Mapping of Tuberculosis Burden in Anambra State, South-East Nigeria

Chukwuebuka Immanuel Ugwu1, Ugochukwu Chukwulobelu2, Chukwumuanya Igboekwu1, Nwamaka Emodi5, Joseph Uche Anumba3, Samuel Chika Ugwu4, Chidinma Lilian Ezeobi6, Vivian Ibeziako6, Glory Ugochukwu Nwakaogor7

1World Health Organization, Awka Field Office, Anambra, Nigeria
2Anambra State Ministry of Health, State Tuberculosis Leprosy and Buruli Ulcer Control Program, Awka, Nigeria
3Federal Ministry of Health, Neglected Tropical Diseases Control and Elimination Programme, Abuja, Nigeria
4Department of Internal Medicine, Jos University Teaching Hospital, Jos, Nigeria
5CARITAS, Anambra, Nigeria
6World Health Organization, TUB Unit, South-East Zonal Office, Enugu, Nigeria
7Healthy Environment for International Development (HEID), Anambra, Nigeria
Email: ugwuc@who.int

Abstract

Background: Anambra state in south-east Nigeria is one of the high TB burden states in the country. Despite recent improvements in TB case notification, estimates from the National Prevalence survey suggest that there is still a significant pool of missed TB cases in the state. Although active TB case finding interventions are needed at community level, information on local TB transmission hotspots is lacking. The objective of this study was to map the geo-spatial location of all TB cases detected in the state in 2019. Findings from this secondary data analysis will help to target interventions appropriately with a view to achieving better program efficiency. Method: A de-identified dataset containing descriptive physical addresses of registered TB cases in 2019 was developed. The dataset was then deconstructed and restructured using Structured Query Language in a relational data base environment. The validated dataset was geocoded using ArcGIS server geocode service and validated using python geocoding toolbox, and Google geocoding API. The resultant geocoded dataset was subjected to geo-spatial analysis and the magnitude-per-unit area of the TB cases was calculated using the Kernel Density function. TB case notification rates were also calculated and Choropleth maps were plotted to portray the TB burden as contained in the dataset. Results: Five local government areas (LGAs) (Onitsha North, Onitsha South, Idemili North, Nnewi North, Ogbaru) had spots with “Extremely high” burden with two LGAs (Onitsha North and South) accounting for the largest spots. Eight LGAs had spots with “Very high” TB burden. Also, 24 hotspots across the...
state had “High” TB burden and two LGAs (Orumba North, Orumba South) had only “Low” TB burden areas. **Conclusion:** Visualizing the heat map of TB patients has helped to identify transmission hotspots that will be targeted for case finding interventions and effort should be made to increase sensitization of the people on certain behavioural attributes that may contribute to contracting Tuberculosis.

**Keywords**

Tuberculosis Transmission Hotspots, Disease Surveillance, Community Tuberculosis Burden

### 1. Introduction

Tuberculosis (TB) is an infectious disease usually caused by *Mycobacterium tuberculosis* (MTB) bacteria. Tuberculosis generally affects the lungs, but can also affect other parts of the body [1] [2] [3]. Most infections show no symptoms in which case it is known as latent tuberculosis. About 10% of latent infections progress to active disease which, if left untreated, kills about half of those affected [4] [5]. The classic symptoms of active TB are a chronic cough with blood-containing mucus, fever, and night sweats. It was historically called consumption cough due to the weight loss. Infection of other organs causes a wide range of symptoms.

Despite increasing treatment success rates, tuberculosis (TB) continues to spread worldwide. The World Health Organization (WHO) estimates that, in 2015, 10.4 million people developed TB and more than 1 million died by the disease [6], with the burden of TB distributed unevenly in the world. The highest incidence rate of TB is being reported in Africa [7]. The United Nations’ (UN) Sustainable Development Goals (SDGs) and WHO’s End TB Strategy have a common target of ending the global TB epidemic by 2030 [8]. To achieve this ambitious target, there is a need to narrow TB interventions with consideration of the geographical inequalities prevalent in the areas with high TB burden. This approach also helps in increasing access to preventive, diagnostic and treatment services for those at the highest risk of TB. In resource-limited and high-TB burden countries such as Nigeria, regarded as one of the TB high burden Countries by WHO [9], identifying hotspot areas for targeted interventions is particularly important to optimize use of available resources and to increase efficiencies in the delivery of TB services [10]. This particular study is set in Anambra state located in the South-east region of Nigeria and is also regarded as one of the high TB burden states within the region. Several community-based active surveillance interventions have been implemented with varying degrees of success in the state. It is hoped that this Geospatial analysis of TB hotspots will help to increase the yield of TB surveillance activities in the state.
Spatial analytics of prevalence in epidemiology at national and subnational levels is useful in identifying vulnerable populations and geographic areas of TB [11]. Understanding the local dynamics in TB prevalence across geographic regions is effective in designing and implementing planned interventions focused on TB control [12]. TB data with spatial foot prints can help to plan the allocation of resources available for TB control in many settings to achieve better efficiency.

**Aim and Objectives:** The objective of this study was to map the geo-spatial location of all drug-susceptible and drug-resistant TB cases detected in 2019 in Anambra state, Nigeria and by comparing the case notification rates (cases per 100,000 population) across the Local government areas, map the TB hotspots in the state. Findings from this secondary data analysis will help to target interventions appropriately with a view to achieving better program efficiency.

**2. Methodology**

**2.1. Study Area**

Anambra is a state in southeastern part of Nigeria. Administratively the state is made up of 21 Local government areas, 181 towns and 330 political wards. Although the administrative capital city is Awka, Onitsha and Nnewi are major hubs of economic activities in the state. Boundaries are formed by Delta State to the west, Imo and Rivers States to the south, Enugu State to the east, and Kogi State to the north. The name was derived from the Anambra River (Omambala) which flows through the area and is a tributary of the River Niger.

**2.2. Ethical Approval**

Permission to conduct this secondary analysis on an already existing dataset was sought and obtained from the Research ethics unit of the directorate for planning research and statistics of the Anambra state ministry of health. The dataset was anonymized as personal identifiers were removed before the permission for use was granted.

**2.3. Data Sources**

Primary data collected over a period of One year January-December, 2019 comprising of TB clients’ addresses and places of diagnosis and treatment. The actual dates of enrolment and addresses were verified from the local government TB treatment registers wherein every enrolled client is assigned a unique treatment number. Two datasets collated in Microsoft excel file were used for geospatial analysis; one comprising case of DR-TB and the other, cases of DS-TB. Both datasets contained anonymized residential address of clients notified between January and December of 2019. The DS-TB dataset contained a total of 2411 cases/records while the DR-TB contained a total of 133 cases. The major tasks performed included geocoding of the entire dataset in order to obtain corresponding geographical coordinates used to perform subsequent geospatial
analytics and creation of thematic maps. We computed the TB case notification rate (per 100,000 population) for each LGA using the number of TB cases notified applied to the projected 2019 population of each LGA as shown in the equation below. Comparison of the CNR was made across the 21 LGAs.

\[
\text{CNR} = \frac{\text{TB Cases}}{\text{Population}} \times 100000
\]

2.4. Geocoding

The datasets were deconstructed, and restructured using Structured Query Language (SQL) in a Relational Database Environment (RDBMS). Thereafter the datasets were validated and exported into a Geographic Information System Software Environment, ArcGIS Pro-Version2.5/ArcGIS Desktop-Version 10.8. The imported data was geocoded using the ArcGIS Server geocode service and crosschecked/validated using python geocoding toolbox and Google geocoding API.

The resultant geocoded dataset (vector data) was then subjected to geospatial analysis to create visual representation of the geo-distribution of cases as well as density/heat maps to reveal hotspots for DR-TB and DS-TB Cases. Spatial analyst toolsets situated in the ArcGIS software were used to calculate the density of TB cases represented as point features in a neighborhood around each output raster cell. In other words, the Point Density tool calculated the density of point features (cases) around each output raster cell. A neighborhood was defined around each raster cell center, and the number of points that fall within the neighborhood was summed and divided by the area of the neighborhood. The mathematical formula used in calculating the point density is given below:

\[
\text{Density} = \frac{1}{(\text{radius})} \sum_{i=1}^{n} \frac{3}{\pi} \cdot \text{pop}_i \left(1 - \frac{\text{dist}_i}{\text{radius}}\right)^2
\]

where \(i = 1, \ldots, n\) are the input points. Only include points in the sum if they are within the radius distance of the \((x, y)\) location; \(\text{pop}_i\) is the population field value of point \(i\), which is an optional parameter; \(\text{dist}_i\) is the distance between point \(i\) and the \((x, y)\) location.

Furthermore, Kernel Density was used to calculate the density of point features (cases) around each output raster cell. This calculation was done to determine the magnitude-per-unit area from the point features (TB Cases) using a kernel function to fit a smoothly tapered surface to each point or polyline. For the point data in this study, the summation of densities of all points within a bandwidth around the focus point was calculated using the formula below:

\[
\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K \left(\frac{x - X_i}{h}\right)
\]

where \(h\) is the bandwidth; \(n\) number of points within the bandwidth; \(K(x)\) is the
kernel function defined for an \( x \) of \( d \) dimensions; \( \left( x - x_i \right) \) represents the Euclidean distance between each point \( I \) and the location where the density estimator is worked out.

Conceptually, a smoothly curved surface was fitted over each point. The surface value is highest at the location of the point and diminishes with increasing distance from the point, reaching zero at the Search radius distance from the point. Only a circular neighborhood was possible. At the end of the geo-processing, thematic maps were produced. The Point density analysis was done using Jenks Natural Break classification (or optimization). The Jenks classification method can be advantageous because it identifies real classes within the data. Choropleth maps method was used to portray trends found in the dataset.

### 3. Result

A total of 2411 physical addresses of DS-TB cases were geocoded to obtain corresponding geographic coordinates. Of these cases, 2295 corresponding coordinates fell into the area of study (Anambra LGAs) based on the set geocoding parameters with a hit rate of 95.2% (shown in Figure 1). With respect to the DR-TB cases, 133 addresses were geocoded; of these, 124 fell into the area of study (Anambra LGAs) with a hit rate of 93.2% based on the geocoding parameters.

The case notification rate calculations are summarized in Table 1 below.

<table>
<thead>
<tr>
<th>LGA</th>
<th>Cases Notified</th>
<th>CNR</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aguata</td>
<td>87</td>
<td>16.9</td>
<td>17</td>
</tr>
<tr>
<td>Anambra East</td>
<td>76</td>
<td>35.6</td>
<td>8</td>
</tr>
<tr>
<td>Anambra West</td>
<td>62</td>
<td>26.6</td>
<td>11</td>
</tr>
<tr>
<td>Anaocha</td>
<td>73</td>
<td>18.4</td>
<td>16</td>
</tr>
<tr>
<td>Awka North</td>
<td>38</td>
<td>24.2</td>
<td>12</td>
</tr>
<tr>
<td>Awka South</td>
<td>214</td>
<td>81.3</td>
<td>4</td>
</tr>
<tr>
<td>Ayamelum</td>
<td>87</td>
<td>39.4</td>
<td>7</td>
</tr>
<tr>
<td>Dunukofia</td>
<td>30</td>
<td>22.3</td>
<td>14</td>
</tr>
<tr>
<td>Ekwusigo</td>
<td>27</td>
<td>12.3</td>
<td>18</td>
</tr>
<tr>
<td>Idemili North</td>
<td>424</td>
<td>70.7</td>
<td>5</td>
</tr>
<tr>
<td>Idemili South</td>
<td>68</td>
<td>23.5</td>
<td>13</td>
</tr>
<tr>
<td>Ihiala</td>
<td>125</td>
<td>29.7</td>
<td>10</td>
</tr>
<tr>
<td>Njikoka</td>
<td>46</td>
<td>22.2</td>
<td>15</td>
</tr>
<tr>
<td>Nnewi North</td>
<td>183</td>
<td>83.4</td>
<td>3</td>
</tr>
<tr>
<td>Nnewi South</td>
<td>37</td>
<td>11.4</td>
<td>19</td>
</tr>
<tr>
<td>Ogharu</td>
<td>159</td>
<td>51.4</td>
<td>6</td>
</tr>
<tr>
<td>Onitsha North</td>
<td>414</td>
<td>237.9</td>
<td>1</td>
</tr>
<tr>
<td>Onitsha South</td>
<td>171</td>
<td>89.8</td>
<td>2</td>
</tr>
<tr>
<td>Orumba North</td>
<td>19</td>
<td>7.9</td>
<td>21</td>
</tr>
<tr>
<td>Orumba South</td>
<td>22</td>
<td>8.4</td>
<td>20</td>
</tr>
<tr>
<td>Oyi</td>
<td>79</td>
<td>33.8</td>
<td>9</td>
</tr>
<tr>
<td>State Total</td>
<td>2441</td>
<td>41.9</td>
<td></td>
</tr>
</tbody>
</table>
and show that Onitsha North and South LGAs, Nnewi North, Awka south and Idemili North are the high burden LGAs in the state and notified the highest cases per 100,000 population.

As shown in Figure 2 and Figure 3 below, five local government areas LGAs (Onitsha North, Onitsha South, Idemili North, Nnewi North, Ogbaru) had spots with “Extremely high” burden with two LGAs (Onitsha North and South) accounting for the largest spots. Extremely high TB burden were found mostly within the urbanized towns. Eight LGAs had spots with “Very high” TB burden. Also, 24 hotspots across the state had “High” TB burden and two LGAs (Orumba North, Orumba South) had only “Low” TB burden areas.

The childhood TB burden displayed in Figure 4 portrays a slightly different picture. Idemili North LGA is the area with the highest burden of TB amongst the 0 - 14 age group. This is followed in density by Ogbaru, Onitsha north and south LGAs. Although the four LGAs were part of the LGAs with the highest burden in the adult category, Idemili North has a higher burden than the others. Further, Nnewi north also had a high burden of childhood TB. These LGAs along with Awka South account for the top 6 ranked LGAs with the highest case notification rates in the state. Anambra west, Anambra East, and Ayamelum which ranked 11th, 8th and 7th respectively in the case notification rates have
Figure 2. Choropleth map with LGA labels of TB burden in Anambra state.

Figure 3. Choropleth map without LGA labels of TB burden in Anambra state.
Figure 4. Choropleth map of childhood TB burden in Anambra state 2019.

significantly high paediatric TB burden.

4. Discussion

This study found that about 5% of all DS-TB patients and nearly 7% of DR-TB patients reside outside Anambra state. This information is critical in planning patients’ care and arranging the schedules of hospital visits given that DS-TB treatment requires directly observed therapy which entails several hospital visits during the treatment period [13]. DR-TB clients being managed on ambulatory basis also would require regular but less frequent hospital visits for monthly check-up visits as well as visits to laboratories for monitoring of therapy. Additionally, in the wake of COVID-19 pandemic, Nigeria implemented widespread
movement restrictions and complete lockdown of inter-state travel making it potentially impossible for this group of patients to access care [14]. Without adequate planning of care for such patients, poor treatment outcomes such as loss to follow up, treatment failure, and even death could result. It could also increase the burden of drug resistance as partially treated patients could develop TB in the future and continue to transmit within their localities. This information will thus help state TB programs to make care more person-centred putting into consideration the needs of patients. Clients coming from afar for various reasons should be given alternatives of facilities closer to their dwelling places. At present, 71% of Nigerian TB patients incur catastrophic costs as a result of accessing TB care [15] [16] [17] and it is expected to be more severe for patients travelling from outside the state. Admittedly, some TB treatment centres within and outside the state have a long history of successful TB treatment outcomes and have gained popularity and patronage beyond the borders of their state of location. It is expected that such facilities in collaboration with their respective state TB programs should develop mitigation plans for cross-border clients to forestall an abrupt interruption of treatment in event of future emergencies. Although the guidance of the National Tuberculosis and Leprosy Control program provides that patients should access care from the treatment sites contiguous to their homes, it has been previously reported that some clients insist otherwise [18].

The five local governments that were identified by this study to have “Extremely high” TB burden are all mostly urban and sub-urban areas boasting high population densities [19]. Onitsha North and South LGAs with a combined land space of about 52 km² play host to the largest market in West Africa that enjoys patronage from merchants from as far as Togo, Cameroun, Ghana, and Niger amongst other countries [20]. Additionally, the rise of sub-urban sprawls around the city centres of Onitsha increases the risk of Tuberculosis transmission in those localities. Ogbaru and Idemili North LGAs were also shown to have “Extremely high” burden. The two LGAs directly border Onitsha South and North respectively. They also boast very high population densities and Idemili north with a human population of over 616,833 is the most populated LGA in the state [19]. Ogbaru on the other hand is classified as a sub-urban LGA with many peri-urban slum areas such as Okpoko, Iyiowa, and Odekpe [21] [22]. The over-crowded dwellings and other indices of deprivation prevalent in these areas significantly increase the risk of TB transmission. The above findings have weighty implications for public health broadly and specifically for disease control in the state. Collaboration between the state ministries of health, environment, town planning and trade and industry will be critical to developing strategies to make the cities safer for human activities.

Other areas such as Ihiala, Nnewi North, other parts of Idemili North, Aguata, other parts of Ogbaru, were also identified as having “very high” TB burden. This is also important to help in targeting interventions to the areas of high
transmission rates. Data driven case finding portends greater efficiency in finding more TB cases with even less resources.

Age disaggregated routine TB notification data was used to conduct spatial analysis and demonstrate the importance of monitoring hotspots of childhood TB [23]. Spatial surveillance of childhood TB cases has been likened to sentinel surveillance as it represents a more temporally relevant approximation of TB transmission [23]. Our study found that certain LGAs had more hotspots of childhood cases compared to their adult TB cases. Similar analytical methods have been used to demonstrate statistically significant cluster of TB cases in an Indian district [24]. Inclusion of a time component to the analysis would have enabled us to detect newly emerging clusters. A key challenge of geospatial studies is to demonstrate statistically significantly higher prevalence in identified clusters and to establish the mechanisms that bring about the clustering of cases. In our context, it is not completely obvious whether the higher prevalence and CNR in our hotspots are as a result of more sensitive surveillance systems and efforts or an actual increase in transmission. At a programmatic level however, community active surveillance targeting these areas for have yielded comparatively more cases. Furthermore, the active community surveillance implemented in the state prior to the spatial analysis were conducted in all the 21 LGAs of the state. It is therefore likely that the hotspots represent areas of higher transmission. On the whole, visualizing the density heat map of TB burden in the state has helped to identify transmission hotspots that are being targeted for active TB case finding interventions.

5. Conclusion

To interrupt community transmission of TB, Community-based active surveillance interventions are needed. However, these interventions are known to be capital intensive as the high number needed to test to diagnose a case of TB in the general population and other logistics challenges tend to increase the cost. A prior understanding of geo-spatial density of the missing TB cases in high-burden resource-limited settings portends better program efficiency in resource use. Although our findings have shown that urban and peri-urban slum areas with high population density harbour most of the transmission hotspots in our state, more research is needed to understand the recency of these transmissions using real time data as well as the exact mechanisms through which these high burden areas have come to be. Disease control programs in our settings and similar are likely to benefit from such efforts that improve understanding of geographical distribution of cases.

Acknowledgements

The authors hereby acknowledge the State Tuberculosis Leprosy and Buruli Ulcer control program of the Anambra State ministry of health who made the data available for the analysis. Dr. Andrew Sunday Jr. created some, and mod-
ified all the maps to meet the requirements of the journal.

**Conflicts of Interest**

The authors declare no conflicts of interest regarding the publication of this paper.

**References**


Critical Care Medicine, 170, 561-566. https://doi.org/10.1164/rccm.200401-095OC


