CODEN [USA]: IAJPBB ISSN: 2349-7750

INDO AMERICAN JOURNAL OF
PHARMACEUTICAL SCIENCES

http://doi.org/10.5281/zenodo.2525792

Available online at: http://www.iajps.com ***Research Article***

METHODS USED IN THE SPATIAL ANALYSIS OF TB EPIDEMIOLOGY

[1]Dr. Ishtiaq Awais,[2]Dr. Husna Rashid,[3]Dr. Aamina Ahmed

[1]MO, Mobile Disensary 142/9L, [2]Ex House Officer,Sir Ganga Ram Hospital,Lahore,
[3]WMO,RHC Lawa.

Abstract:

Tuberculosis transmission often happens inside a household or community, leading to heterogeneous spatial patterns. Therefore, apparent spatial clustering of Tuberculosis could reflect ongoing transmission or co-location of risk factors and can vary considerably depending on the type of data available, the analysis methods employed and the dynamics of the underlying population. Thus, we aimed to review methodological approaches used in the spatial analysis of tuberculosis burden.

We conducted a systematic literature search of spatial studies of tuberculosis published in English using Medline, Embase, PsycInfo, Scopus and Web of Science databases with no date restriction from inception to 15 February 2017.The protocol for this systematic review was prospectively registered with PROSPERO.

We identified 168 eligible studies with spatial methods used to describe the spatial distribution (n = 154), spatial clusters (n = 73), predictors of spatial patterns (n = 64), the role of congregate settings (n = 3) and the household (n = 2) on tuberculosis transmission. Molecular techniques combined with geospatial methods were used by 25 studies to compare the role of transmission to reactivation as a driver of tuberculosis spatial distribution, finding that geospatial hotspots are not necessarily areas of recent transmission. Almost all studies used notification data for spatial analysis (161 of 168), although none accounted for undetected cases. The most common data visualization technique was notification rate mapping, and the use of smoothing techniques was uncommon.

The clusters of spatial were analyzed utilizing the range of methods, with the most commonly employed being Kulldorff's spatial scan statistic. In the 11 papers that compared two such methods using a single data set the clustering patterns identified were often inconsistent. Classical regression models that did not account for spatial dependence were commonly used to predict spatial tuberculosis risk.

Keywords: *Spatial analysis, Tuberculosis, Genotypic cluster.*

Corresponding author:

Dr. Ishtiaq Awais,

MO, Mobile Disensary 142/9L.

QR code

Please cite this article in press Ishtiaq Awais et al., ***Methods Used In The Spatial Analysis Of Tb Epidemiology.,*** *Indo Am. J. P. Sci, 2018; 05(12).*

INTRODUCTION:

Tuberculosis differs from other infectious diseases in several ways that are likely to influence apparent spatial clustering. For example, its long latency and prolonged infectious period allow for significant population mobility between serial cases. Thus, Mycobacterium tuberculosis infection acquired in a given location may progress to tuberculosis disease in an entirely different region, such that clustering of cases may not necessarily indicate intense transmission but could rather reflect aggregation of population groups at higher risk of disease, such as migrants. Similarly, Mycobacterium tuberculosis infection acquired from workplaces and other congregate settings can be wrongly attributed to residential exposure, as only an individual's residence information is typically recorded on tuberculosis surveillance documents in many settings (Cameletti and Finazzi, 2017).

Identifying heterogeneity in the spatial distribution of tuberculosis cases and characterizing its drivers can help to inform targeted public health responses, making it an attractive approach. However, there are practical challenges in appropriate interpretation of spatial clusters of tuberculosis. Of particular importance is that the observed spatial pattern of tuberculosis may be affected by factors other than genuine tuberculosis transmission or reactivation, including the type and resolution of data and the spatial analysis methods used. For instance, use of incidence data versus notification data could give considerably different spatial pattern, as the latter misses a large number of tuberculosis cases and could be skewed towards areas with better access to health care in high-burden settings (Erdmann, 2017).

Thus, spatial analysis using notification data alone in such settings could result in misleading conclusions. Similarly, the type of model used and the spatial unit of data analysis are important determinants of the patterns identified and their associations. That is, different Spatial resolutions could lead to markedly different results for the same dataset regardless of the true extent of spatial correlation and the effect observed at a regional level may not hold at the individual level (an effect known as the ecological fallacy). Therefore, we aimed to review methodological approaches used in the spatial analysis of tuberculosis burden (Erdmann, 2017).

METHODS AND MATERIALS:

We searched Medline, Embase, Web of Science, Scopus and PsycInfo databases from their inception to 15 February 2017 using a combination of keywords and medical subject headings (MeSH) pertaining to our two central concepts: tuberculosis and space. We refined search terms related to the latter concept after reviewing key studies, including a previous systematic review not limited to tuberculosis.

Inclusion and Exclusion Eligibility

We included peer-reviewed papers that incorporated the spatial analysis approaches described above in the study of tuberculosis. After exclusion of duplicates, titles and abstracts were screened by two researchers (DS and MK) to identify potentially eligible studies. Of these papers, articles were excluded hierarchically on the basis of article type, whether the method used could be considered spatial or not and the outcomes assessed. No exclusions were made on the basis of the outcome reported, with studies that considered incidence, prevalence or any tuberculosis -related health outcome included (Goovaerts, 2017).

Studies were excluded if the language of the publication was not English, the report was a letter, conference abstract or a review or only reported the temporal (trend) of tuberculosis. Spatial studies of non-tuberculosis mycobacteria, non-human diseases and population immunological profiles were also excluded. Full-text articles were excluded if they did not provide sufficient information on the spatial analysis techniques employed. There were no exclusions based on study setting or anatomical site of disease (Jacquez, 2017).

RESULTS:

Study characteristics

A total of 2350 records were identified from the electronic searches, of which 252 full-text articles were assessed. Of these, 168 articles met all inclusion criteria and were included in the final narrative synthesis according to Figure 1. Using a cutoff of 100 tuberculosis cases per 100,000 population in reported incidence in 2016, 111 (66%) of the studies were from low-incidence settings. All references returned by the search strategy were from the period 1982 to 2017, with 71% published from 2010 onwards. Earlier studies (predominantly in the 1980s and 1990s) tended to be descriptive visualizations, while studies in the last two decades frequently incorporated cluster detection and risk prediction. More recently, a range of statistical techniques including Bayesian statistical approaches and geographically weighted regression have become increasingly popular (Jacquez and Goovaerts, 2017).

Key objectives of included studies

Spatial analysis was applied to address a range of

objectives, according to Table 1 given below, with the commonest ones including description of the distribution (n = 135), statistical analysis of spatial clustering (n = 73) and analysis of risk factors and risk prediction (n = 64). Spatial methods were also used to determine the relative importance of transmission by comparison to reactivation as a driver of tuberculosis incidence (n = 25), the effect of tuberculosis interventions (n = 2), barriers to tuberculosis service uptake (n = 2), spatial distribution of tuberculosis -related health outcomes (mortality, default, hospitalization) (n = 5), spatial pattern of tuberculosis incidence among people living with HIV (PLHIV) (n = 4), HIV-related tuberculosis mortality (n = 4), multidrug-resistant tuberculosis (MDR-TB) drivers (n = 1), tuberculosis outbreak detection (n = 3) and drivers of spatial clustering (including the role of congregate settings, such as social drinking venues and schools) (n = 30) (Jacquez and Goovaerts, 2017).

Types of tuberculosis disease analyzed

Spatial analysis was most commonly conducted on data for all types of tuberculosis (i.e. without distinction between pulmonary or extra-pulmonary; n = 121), followed by pulmonary tuberculosis only (n = 28) and smear-positive pulmonary tuberculosis only (n = 13). Spatial analysis of multidrug-resistant tuberculosis (MDR-TB) and extensively drug-resistant tuberculosis (XDR-TB) was reported in 15 studies and one study respectively.

Data used and scale of analysis

Nearly all studies used retrospective tuberculosis program data (notifications), with the exception of five studies that used prevalence surveys and two prospectively collected data. None of the studies using notification data accounted for undetected/unreported cases. In all included studies, spatial analysis of tuberculosis was based on the individual's residence, except for three studies that explored the effect of exposure from social gathering sites (James, Yavchitz and Boutron, 2018).

Spatial analysis was generally done using data aggregated over administrative spatial units (n = 131), but the scale of aggregation differed markedly. Common spatial scales included census tract (n = 20), district (n = 15), postal code (n = 15), county (n = 15), neighbor-hood (n = 10), health area (n = 7), municipality (n = 11), state (n = 7), province (n = 6), local government area (LGA) (n = 4) and ward (n = 4). Data were analyzed at the individual level in 37 studies, while three studies were reported at a continent and country scale (Jacquez and Goovaerts, 2017).

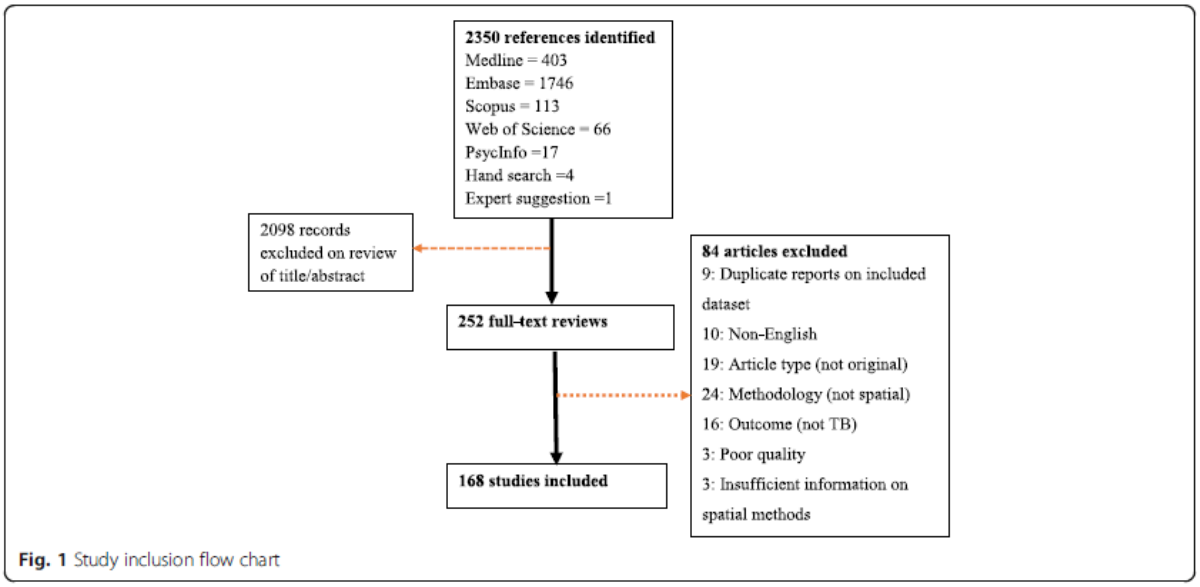

Fig. 1 Study inclusion flow chart

Table 1

Table 1 Application areas of spatial methods in TB studies

Spatial method application areas	Methods used
Spatial TB distribution or spatial clustering	Dot maps, rate maps, thematic maps, Moran's *I*, GetisOrd statistic, NNI Besag and Newel statistic, *k*-functions, spatial scan statistic
Risk factors	Bayesian CAR models, regression models (with or without including spatial terms), GWR, PCA, mixture models, spatial lag models
Monitoring spatiotemporal TB trends	Temporal trend maps
Intervention evaluation	Distance map, kernel density map
Barriers to TB care	Rate map, dot map, travel time map, distance map
TB program performance	Map (time to detection)
HIV-related TB incidence	Rate map, dot map, spatial scan statistic
TB treatment outcomes	Spatial empirical Bayes smoothing, kernel density maps, spatial scan statistic, spatial regression
Mortality related to TB/HIV coinfection	Rate map, thematic maps, Moran's *I* and spatial regression
Transmission	Dot maps (congregate settings)
	Dot maps (cases)
	Geospatial and genotypic clustering methods
Methodological	Spatial scan statistic
TB outbreak detection	Spatial scan statistic
Prevalence estimation	Model-based geostatistics
Drivers of MDR-TB	*k*-function

Methods in the spatial analysis of tuberculosis

Below mentioned Table 2 shows the range of spatial methods used. Spatial analysis was used to visualize patterns (n = 154), explore spatial clusters (n = 73) and identify risk factors for clustering (n = 64), with risk prediction undertaken by 11 studies. Of the included studies, six did not explicitly report any of these methods but reported statistical results that implied the use of these methods (James, Yavchitz and Boutron, 2018).

Table 2

Method category	Method	Number
Visualisation	Rate map	63
	Dot map	37
	SMR map	12
	Kernel density map	7
	Case counts maps	3
	Others*	17
Spatial cluster analysis	Global Moran's *I*	28
	Local Moran's *I*	14
	Kulldorff's spatial scan statistic	43
	GetisOrd statistic	12
	k-NN	8
	k-function	6
	Besag and Newell statistic	2
Statistical modelling	Bayesian CAR models	7
	Geographically weighted regression	6
	Mixture modelling	2
	Conventional logistic	15
	Conventional Poisson	5
	Conventional linear	5
	Negative binomial	1
	Factor analysis	6
	Regression models with spatial terms	9
	Spatial prediction	11

Source: (Luan, Law and Lysy, 2018)

Data visualization

Data visualization was the most consistently applied technique, with 154 of the studies using at least one data visualization method to present tuberculosis distribution and/or risk factor patterns across space (Table 1). The tuberculosis incidence rate was the commonest indicator mapped (n = 63), followed by event maps (n = 37), which were smoothed using kernel density in seven studies. Data visualization was based on standardized morbidity ratios (SMR) in 12 studies. Five studies reported maps of trends in tuberculosis incidence over time, and thematic maps were used in nine to consider the impact of risk factors on tuberculosis incidence by displaying the spatial distribution of other variables. Variables plotted included climate (n = 1), socioeconomic factors (n = 5), diabetes (n = 1) and obesity (n = 1) (James, Yavchitz and Boutron, 2018).

Approaches used to account for data sparseness

Tuberculosis is a relatively rare disease at the population level, and burden is typically expressed in terms of cases per 100,000 populations. Various approaches were used to account for this sparseness

in the number of cases, such as aggregating cases over administrative geographic levels and over time periods (ranging from 1 to 25 years). An alternative approach was rate smoothing, although this practice was rare, despite the fact that tuberculosis rates were the commonest indicators mapped. In the included studies, smoothed rates were used in six (4%) studies. Similarly, of 12 studies that analyzed SMRs, smoothed SMRs were presented in seven. In the included studies, several different data smoothing techniques were used, including fully Bayesian (n = 8), empirical Bayes (n = 4) and spatial empirical (James, Yavchitz and Boutron, 2018).

Bayes (n = 5). A significant number of visualization reports (n = 30) were not complemented by hypothesis testing, either by exploration methods or modeling approaches. In 12 studies (7%), maps were not presented, but a narrative description of tuberculosis burden or a tabular presentation of TB distribution by administrative unit was described.

Spatial cluster (hotspot) identification

Use of at least one spatial cluster identification method was reported in 73 (43%) studies, with Kulldorff's spatial scan statistic used most frequently (n = 43). Nearest neighbor index (NNI), k-function and Besag and Newell methods were reported in eight, six and two studies respectively (Table 1). The presence of overall area-wide heterogeneity was assessed most often using global Moran I (n = 28) (Luan, Law and Lysy, 2018).

False-positive clustering

Not all spatial clusters are true clusters. False-positive clusters can arise from various sources, including data and methods used and unmeasured confounding. Given that notification data were by far the most commonly used data source in the spatial analyses reviewed here, it could not be determined if these clusters represented true clusters of tuberculosis incidence or if they were caused by factors such as pockets of improved case detection. The role of differential tuberculosis detection has been documented in some studies from low-income settings, where increased spatial tuberculosis burden was linked to improved health care access. In addition, rate was the commonest disease indicator used for disease mapping, as well as cluster detection in this study. As described earlier, rates are liable to stochasticity and can lead to false-positive clustering. However, rate smoothing and stability (sensitivity) analysis of clusters identified using rates was done in only a few studies. This remains an important area of consideration in the future spatial analysis of tuberculosis as per mentioned in below Table 3.

Comparisons of spatial clusters from multiple cluster identification methods

Author, year	Methods	Outcome	Conclusion
Alene, K, 2017	Local Moran's I Getis and Ord	Clustered Clustered	50% similarity (two non-significant clusters identified by LISA)
Álvarez-Hernández, G., et al. 2010	Local Moran's I Besag and Newell	No significant Clustered	Widely conflicting
Dangisso M, et al. 2015	Getis and Ord Spatial scan statistic	Clustered Clustered	Similar overall pattern, but marked differences by years
Feske, M., et al. 2011	Getis and Ord GWR residuals	Clustered Heterogeneous	Similar overall pattern, but some local differences
Ge E, et al. 2016	Getis and Ord Spatial scan statistic	Clustered Clustered	Similar overall pattern, but differences in some locations and across time
Haase I, et al. 2007	Hotspot analysis SaTScan	Clustered Clustered	Similar overall pattern, but some local differences
Hassarangsee S, et al. 2015	LISA Spatial scan statistic	Clustered Clustered	Very similar, but not identical
Li L, et al. 2016	LISA Spatial scan statistic	No significant cluster, Clustered	Widely conflicting
Maceiel ELN, et al. 2010	LISA, Getis and Ord Model prediction	Clustered Heterogeneous	Widely conflicting
Wubuli A, et al. 2015	LISA Getis and Ord	Clustered Clustered	Similar overall pattern, but some local differences
Wang T, et al 2016	Spatial scan statistic Getis and Ord	Clustered Clustered	Similar overall pattern, but some local differences

Source: (Luan, Law and Lysy, 2018)

Results from spatial analysis

Geographic distribution of TB

The geographic distribution of tuberculosis was heterogeneous in all included studies both from low- and high-incidence settings, although no formal hypothesis testing was presented in 55 (33%). An exception was one study from South Africa that reported no significant clustering of cases among HIV patients on ART. Spatial analysis was also used to describe the drivers of drug-resistant tuberculosis, with tighter spatial aggregation of MDR-TB cases compared with non-MDR cases taken as evidence of transmission of MDR-TB. Spatial analyses into both HIV and tuberculosis investigated outcomes including HIV-associated TB incidence (n = 4) and spatial patterns of TB/HIV-related mortality (n = 4).

All such studies revealed significant spatial heterogeneity. TB/HIV-related mortality in children was linked to areas with low socio-economic status and maternal deaths. Spatial methods used to study the impact of community-based TB treatment showed marked improvement in access compared to health facility-based treatment approaches (n = 1), and similar studies demonstrated travel time and distance to be important barriers to TB control (n = 2) Source: (Luan, Law and Lysy, 2018)

Correlations with social and environmental factors

The observed spatial patterns of tuberculosis were consistently linked to areas with poverty (n = 14), overcrowding and non-standard housing (n = 9), ethnic minority populations (n = 3), population density (n = 2), low education status (n = 2), health care access (n = 3) and immigrant populations (n = 5). However, a minority of studies have also found conflicting or non-significant associations between tuberculosis and poverty, population density and unemployment.

Use of spatial methods to inform public health interventions

In addition to their use in characterizing the spatial distribution and determinants of tuberculosis, spatial methods have been used to inform tuberculosis-related public health interventions. In these studies, spatial analysis methods have proved to be attractive in guiding public health interventions, although their application to tuberculosis care beyond research is not well documented. For instance, spatial analysis techniques have been used to identify locations with a high density of tuberculosis cases (termed hotspots, although this definition was not based on spatial statistical tests) (James, Yavchitz and Boutron, 2018).

Community screening was then conducted in these areas, and its yield was compared to that from routine service provision. This GIS-guided screening was found to considerably improve the detection of individuals with latent tuberculosis infection and other infectious diseases.

DISCUSSION:

In almost all reviewed studies, retrospective program data (notifications) were used. Notification data, especially from resource-scarce settings, suffer from the often large proportion of undetected cases and are heavily dependent on the availability of diagnostic facilities. None of the spatial studies of tuberculosis that used notification data accounted for undetected cases, such that the patterns in the spatial distribution and clustering could be heavily influenced by case detection performance (James, Yavchitz and Boutron, 2018).

Hence, distinguishing the true incidence pattern from the detection pattern has rarely been undertaken, despite its importance in interpretation. The problems of undetected cases could be compounded in the spatial analysis of drug-resistant forms of tuberculosis, especially in resource-scarce settings where testing for drug-resistant tuberculosis is often additionally conditional on the individual's risk factors for drug resistance (James, Yavchitz and Boutron, 2018).

However, recently, there have been some attempts to account for under-detection in the spatial analysis of tuberculosis. A Bayesian geospatial modeling approach presented a framework to estimate tuberculosis incidence and case detection rate for any spatial unit and identified previously unreported spatial areas of high burden. Another approach is to estimate incidence using methods such as capture-recapture and mathematical modeling. If case detection rate is truly known for a defined region, incidence can be calculated as notifications divided by case detection rate, although this is rarely if ever the case.

Spatial analysis using prevalence data could also be considered in areas where such data are available. In relation to the data problems outlined above, spatial analysis of tuberculosis could benefit from the use of model based geo-statistics, which is commonly used in other infectious diseases, although there are few studies that consider Mtb. In particular, measurement of tuberculosis prevalence is impractical to perform at multiple locations due to logistic reasons. Therefore, model-based geo-statistics can be used to

predict disease prevalence in areas that have not been sampled from prevalence values at nearby locations at low or no cost, producing smooth continuous surface estimates (Jacquez and Goovaerts, 2017).

Bayesian smoothing techniques can mitigate the problems of stochastically unstable rates from areas with small population, although such techniques were not widely used in the included studies and so false spatial clustering remains an important consideration. The less frequent use of rate smoothing techniques in the spatial analysis of tuberculosis could have various explanations, including lack of software packages that are easily accessible to the wider user. It may also be that most spatial analyses of tuberculosis are based on data aggregated over larger geographic areas from several years, such that the problem of statistical may not be a major problem, although this was not explicitly discussed in the included studies (Jacquez and Goovaerts, 2017).

CONCLUSIONS:

A range of spatial analysis methodologies have been employed in divergent contexts, with virtually all studies demonstrating significant heterogeneity in spatial tuberculosis distribution regardless of geographic resolution. Various spatial cluster detection methods are available, although there is no consensus on how to interpret the considerable inconsistencies in the outputs of these methods applied to the same dataset. Further studies are needed to determine the optimal method for each context and research question and should also account for unreported cases when using notifications as input data where possible. Combining genotypic and geospatial techniques with epidemiologically linkage of cases has the potential to improve understanding of tuberculosis transmission.

REFERENCES:

1. Cameletti, M. and Finazzi, F. (2017). An analysis of temporal and spatial patterns in Italian hospitalization rates for multiple diagnosis. *Spatial and Spatio-temporal Epidemiology*, 19(11), pp.37-45.
2. Erdmann, C. (2017). Spatial Analysis of Disease Patterns Using Cartograms: Methods and Issues. *Epidemiology*, 18(Suppl), p.S181.
3. Goovaerts, P. (2017). Geostatistical analysis of health data with different levels of spatial aggregation. *Spatial and Spatio-temporal Epidemiology*, 3(1), pp.83-92.
4. Jacquez, G. (2017). Cluster morphology analysis. *Spatial and Spatio-temporal Epidemiology*, 1(1), pp.19-29.
5. Jacquez, G. and Goovaerts, P. (2017). The emerging role and benefits of boundary analysis in spatio-temporal epidemiology and public health. *Spatial and Spatio-temporal Epidemiology*, 1(4), pp.197-200.
6. James, A., Yavchitz, A. and Boutron, I. (2018). Importance of the methods used to support the node-making process in network meta-analysis. *Journal of Clinical Epidemiology*, 101, pp.125-126.
7. Lee, D. (2011). A comparison of conditional autoregressive models used in Bayesian disease mapping. *Spatial and Spatio-temporal Epidemiology*, 2(2), pp.79-89.
8. Luan, H., Law, J. and Lysy, M. (2018). Diving into the consumer nutrition environment: A Bayesian spatial factor analysis of neighborhood restaurant environment. *Spatial and Spatio-temporal Epidemiology*, 24(2), pp.39-51.
9. Moraga, P. (2017). SpatialEpiApp : A Shiny web application for the analysis of spatial and spatio-temporal disease data. *Spatial and Spatio-temporal Epidemiology*, 23(12), pp.47-57.
10. Orndahl, C. and Wheeler, D. (2018). Spatial analysis of the relative risk of suicide for Virginia counties incorporating uncertainty of variable estimates. *Spatial and Spatio-temporal Epidemiology*, 27(4), pp.71-83.
11. Sadatsafavi, M., Shahidi, N., Marra, F., FitzGerald, M., Elwood, K., Guo, N. and Marra, C. (2017). A statistical method was used for the meta-analysis of tests for latent TB in the absence of a gold standard, combining random-effect and latent-class methods to estimate test accuracy. *Journal of Clinical Epidemiology*, 63(3), pp.257-269.
12. Selvin, S. and Merrill, D. (2017). A Spatial Analysis. *Epidemiology*, 13(2), pp.151-156.
13. Souris, M. and Bichaud, L. (2011). Statistical methods for bivariate spatial analysis in marked points. Examples in spatial epidemiology. *Spatial and Spatio-temporal Epidemiology*, 2(4), pp.227-234.
14. Sparks, C. (2017). Violent crime in San Antonio, Texas: An application of spatial epidemiological methods. *Spatial and Spatio-temporal Epidemiology*, 2(4), pp.301-309.