

Specific Urban Elements Identified for Tuberculosis Epidemic in Guangzhou by Using Geographical Detector

HongYan Ren (✉ renhy@igsnr.ac.cn)

IGSNRR CAS: Institute of Geographic Sciences and Natural Resources Research Chinese Academy of Sciences

Weili Lu

IGSNRR CAS: Institute of Geographic Sciences and Natural Resources Research Chinese Academy of Sciences

Xueqiu Li

Guangzhou Center for Disease Control and Prevention

Hongcheng Shen

Guangzhou Chest Hospital

Research Article

Keywords: Tuberculosis, Geographical detector, Specific urban elements, pairwise interaction, Guangzhou

Posted Date: December 8th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1136808/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: The prevalence of tuberculosis (TB) in China has heavily affected people's health for decades, which has been widely investigated for the rural regions and west parts. However, its spatial features in urban areas remain little understood. Thus, this study aims to identify its spatial differentiations and their influencing factors in highly urbanized region on a fine scale.

Methods: Together with the TB cases in 2017 obtained from Guangzhou Institute of Tuberculosis Control and Prevention, in total 18 socioeconomic and environmental variables were included in this study. Two spatial analysis tools were respectively applied to select the relative appropriate spatial scale (global Moran's I), and to identify specific urban factors (the Geographical detector) for this epidemic in the central four districts of Guangzhou.

Results: The 2 km × 2 km grid was determined as the most appropriate spatial scale due to its relatively higher spatial autocorrelation (Moran's I=0.33, Z=4.71). At this spatial level, the TB epidemic in the four central districts was obviously closely associated with most of socioeconomic factors ($0.31 < r < 0.76$) at the significance level of 0.01. By contrast, among environmental factors, only the concentration of fine particulate matter (PM_{2.5}) correlated with this epidemic ($r=0.21$) at the significance level of 0.05. Similarly, according to the q-values derived from geographical detector analysis, socioeconomic factors posed stronger impacts ($0.08 < q < 0.57$) on the spatial differentiations of TB prevalence than those of environmental variables ($0.06 < q < 0.27$). Furthermore, 153 pairs of variables presented more powerful explanatory abilities for this epidemic's spatial disparities due to their notable enhancements of q-values ($7.3\% < sq < 311.6\%$) caused by the pairwise interactions.

Conclusion: The spatial heterogeneity of TB prevalence was remarkably influenced by a series of specific urban elements and their pairwise interactions across the central region of Guangzhou. We accordingly suggest that more attentions should be paid to the areas with pairwise interactions of these specific urban elements in this city. This study would provide meaningful clues for local authorities making more targeted interventions on this disease in China's municipal areas featured by both high urbanization and severe tuberculosis.

Introduction

Tuberculosis (TB) is a chronic infectious disease that is caused by *Mycobacterium tuberculosis* and mainly erodes the lungs [1]. In 2019, more than 1.4 million people died of TB in the world, highlighting that this disease is a worldwide public health problem [2]. There were more than 0.8 million newly reported cases in China in 2019, accounting for 8.4% across the world [3]. It is urgent to accurately identify its potential influencing factors for hygienic authorities making targeted interventions on this disease.

M. tuberculosis is spread into the air by patients through many actions such as talking, sneezing, and spitting [4]. It adheres to saliva and fine particles, which can cause infection if inhaled by susceptible individuals [4]. Many studies have been conducted on TB's risk factors, including the pathogen, susceptible populations, socioeconomic elements, and environmental conditions [5–11]. Environmental factors affect the TB

prevalence mainly by influencing the survival and spread of *M. tuberculosis* in the air. Among them, high temperature and humidity appropriately contribute to its survival and spread [12]. Airborne particulate matter provides adsorption sites for *M. tuberculosis*, thus prolonging its suspension in the air [13]. Precipitation and vegetation impede its airborne transmission by inhibiting the dispersal of its carriers (e.g., saliva and particulate matter) [14]. With comparison to environmental conditions, socioeconomic factors, such as population density, household income, the allocation of medical resources, and public traffic system (e.g., roads, bus stops, and subway stations), often affect the TB prevalence by influencing the exposure probability of susceptible populations and treatment of existing TB cases [8, 15, 16]. These earlier studies revealed that this epidemic is often influenced by multiple factors. However, the relationship between factors' interaction and TB epidemics has received very little attention.

Meanwhile, a series of epidemiological investigations have been conducted to detect the dominant influencing factors for TB epidemics in some endemic regions at various spatial scales (such as the country, province, city, county or district, township or street, village, and regular grid) [5, 6, 8, 11, 15, 16, 17, 18], from which a number of different dominant factors were identified due to the spatial scale effect. For example, some environmental conditions (i.e., the meteorological conditions, air quality) were determined as the key factors for TB epidemic's spatial disparities in China at the city level [19], and the spatial patterns of TB prevalence in the Beijing-Tianjin-Hebei region were mainly influenced by the county-level socioeconomic variables (e.g., income level, healthcare) [18]. Till the minimal spatial scale (1km × 1km grids) at present, the distribution of TB cases in Shandong Province during 2009-2011 was spatially clustered and closely associated with the transport network, especially in the densely populated region [16]. Despite of these different factors for TB prevalence, most of previous downscaling studies referred to some special factors closely correlating with the local population status (e.g., the high-risk population, population density and mobility, public transportation, and so on). In particular, small spatial scales are often considered as the final or basic nodes for making targeted intervention strategies and implementing prevention and control measures. However, the spatial heterogeneity of TB prevalence in highly urbanized region and its association with some potential urban elements and their interactions remain little understood on a fine scale.

Therefore, this study employed two spatial analysis tools to characterize the spatial patterns of TB prevalence in the central districts of Guangzhou, and to identify specific urban elements from a series of potential socioeconomic and environmental factors for this epidemic at a fine scale. The aim was to provide effective guidance for hygienic authorities making targeted interventions on this disease in Guangzhou City and China's municipal regions featured by both high urbanization and severe TB epidemic.

Methods

Study area

Guangzhou City is a typical representative of China's megacities with its high urbanization and frequent international trades [20]. The characteristics of its subtropical monsoon climate are obvious: warm and rainy, enough light and heat, an annual average temperature of 21 – 23 °C, and an average annual precipitation of 1800 mm. There are four central districts in Guangzhou, including Yuexiu, Haizhu, Tianhe, and Liwan featured

by vigorous economic activities, comprehensive public facilities, dense population, and convenient transportation [21]. The regional location is shown in Fig. 1.

Data collection

The TB epidemic data for this study were obtained from the Guangzhou Institute of Tuberculosis Control and Prevention, and included TB case data in 2016 – 2017 (taking new cases reported in 2017 as the dependent variable in this study, and existing cases reported in 2016 as a potential influencing factor). The TB case information includes age, sex, permanent residence address, occupation, as well as time of onset and diagnosis. The residence address information of reported cases, after desensitization, was used in conjunction with geocoding (restapi.amap.com/v3/geocode) and coordinate deviation correction to produce case data for a spatial point layer using ArcGIS 10.3 (ESRI, Redlands, CA, USA) software, by which the ratio of the TB cases counts to the population in 2015 was calculated at a fine scale so as to indicate the TB incidence rates across the central districts. In 2017, TB cases in the Guangzhou region were at a relatively high level, with a total of 14100 cases reported, of which 4313 were from the central area, accounting for 30.6%, while its area only accounts for 4.5% of the entire Guangzhou city.

According to the previous studies, we gathered 18 potential factors from various data sources and then categorized them into two groups as listed in Table 1. Among 11 socioeconomic factors, the population-related elements included the TB cases in the previous year and the 1 km×1 km gridded population density, economic situation was termed by the 1 km×1 km gridded gross domestic product per capita, and medical resources were indicated by officially appointed medical institutions. Road network, bus stops, and subway stations were selected for representing the condition of public traffic system. With regard to potential influences of land use, the percentages of four typical kinds in the central area, such as residential, commercial service, public service, and urban villages, were also calculated in this study. In addition to above socioeconomic variables, five environmental conditions in 2017, including monthly mean values of normalized difference vegetation index and fine particulate matters concentration, as well as average climatic conditions (e.g., temperature, precipitation, and humidity) from March to June, were simultaneously considered as potential variables influencing the TB prevalence.

Table 1
Data collection and resources in this study

Data groups	Selected variables by previous studies	Data type	Source
Socioeconomic Factors	TB incidence rates in 2016 (Incid-2016) [8, 15]	Vector	Guangzhou Center for Disease Control and Prevention
	Population density (Pop) in 2015 [8, 11, 14, 15]	Raster(1km)	Resource and Environment Science and data center (https://www.resdc.cn/)
	Gross domestic product per capita (GDP) in 2015 [6, 14]	Raster(1km)	
	Officially appointed medical institutions (Hosp) [4, 22]	Vector	Guangzhou Municipal People's Government(http://www.gz.gov.cn/)
	Density of road network (Road_net) [16]	Vector	Open Street Map (http://download.geofabrik.de/)
	Numbers of subway stations (Subway)[16]	Vector	
	Counts of bus stops (Bus) [16]	Vector	
	Percentage of residential land (Residential) [15]	Vector	Tsinghua University (http://data.ess.tsinghua.edu.cn) [23]
	Percentage of commercial service land (Commercial) [15]	Vector	

Data groups	Selected variables by previous studies	Data type	Source
	Percentage of land for public services (Pub-serv) [15]	Vector	
	Percentage of urban village area (UV) [24]	Vector	Our earlier study [25]
Environmental conditions	Monthly average of normalized difference vegetation index (NDVI) [14]	Raster(1km)	MODIS (https://modis.gsfc.nasa.gov/)
	Monthly average of fine particulate matters concentration (PM2.5) [14]	Raster(1km)	Socioeconomic Data and Applications Center (https://sedac.ciesin.columbia.edu/data/sets/browse)
	Average temperature from March to June (Temp) [12, 14]	Raster(1km)	China Meteorological Data Service Center (http://data.cma.cn/)
	Average precipitation from March to June (Prec) [12, 14]	Raster(1km)	
	Average humidity from March to June (Humi) [12, 14]	Raster(1km)	
Note: For the data of officially appointed medical institutions, medical points that do not serve the community (only designated personnel) were removed according to their service recipients. The officially appointed medical institutions were then divided into outpatient (Hosp11), inpatient and outpatient (Hosp12) types according to the range of medical institution services.			

For the convenience of statistical and spatial analysis, the dependent (newly reported TB cases in 2017) and 18 independent variables, with different data types (vector) or diverse resolutions (raster) as given in Fig. 2,

were summarized and aggregated into a uniform spatial scale by using the spatial join, zonal statistics, and field calculation tools in ArcGIS 10.3 software.

Spatial scale

In the previous epidemiological studies, the administrative divisions (e.g., counties, districts, towns, or streets) were often determined as research units, which was very likely to cause difficulties of conducting temporal analysis and comparison due to the constant changes of these administrative divisions or irregular shapes. To avoid this phenomenon, replacing them by the regular grids is an appropriate solution [26]. In particular, these spatial grids are gradually considered as the final nodes for making targeted intervention strategies and allocating prevention and control resources [27].

In this study, a series of 1×1 km - 5×5 km spatial grids were constructed. The optimal grid scale characterizing the spatial pattern of TB epidemic was selected based on Moran's I and Z-scores [28]. Moran's I was calculated as follows.

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n \omega_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n \omega_{ij} \sum_{i=1}^n (x_i - \bar{x})^2}$$

Where n is the number of grids in the study area, x_i and x_j represent the TB incidence rates in grids i and j, respectively. ω_{ij} is the spatial weight. Global Moran's I is generally tested by the Z-score/P-value, and the value varies from -1 to 1. A higher Moran's I (larger Z-score and proper P-value) indicates greater similarity among attributes between adjacent spatial grids, which reveals that the TB epidemic is clustered in the region, whereas a low negative value indicates dissimilarity between adjacent grids and shows that the TB epidemic is discretely distributed in the region. In this study, Moran's I and Z-scores of the TB incidence rates with different grid sizes were used to assess the optimal grid scales of the regional TB epidemic. Global Moran's I was performed using ArcGIS 10.3.

The geographical detector analysis of univariate and bivariate

The geographical detector (GD) [29] is a statistical tool for detecting the spatial heterogeneity and its determinants (<http://geodetector.cn/>). In this study, GD was used to detect the influence of the socioeconomic and eco-environmental factors on TB prevalence. The method assumes that if the selected factors are associated with TB prevalence, they have a similar spatial distribution. The formula is as follows:

$$q = 1 - \frac{1}{N\sigma^2} \sum_{h=1}^L N_h \sigma_h^2$$

where N and σ^2 are the total counts of grid units and the variance of TB incidence. L is the number of layers in which the study area is divided. $h = 1, 2, 3, \dots, L$. $q \in [0, 1]$, means that the explanatory power of the detection factors for the spatial heterogeneity of TB prevalence is 100*q%.

The interaction detection of GD is also used to determine the explanatory power of the interaction between the two factors on the spatial heterogeneity of TB prevalence. The interaction effects are judged by the following

rules:

$$\text{Enhance, nonlinear} - : q(X_1 \cap X_2 = X_3) > q(X_1) + q(X_2)$$

$$\text{Independent: } q(X_1 \cap X_2 = X_3) = q(X_1) + q(X_2)$$

$$\text{Enhance, bi} - : q(X_1 \cap X_2 = X_3) > \text{Max} [q(X_1), q(X_2)]$$

$$\text{Weaken, uni} - : \text{Min}(q(X_1), q(X_2))$$

$$\text{Weaken, nonlinear: } q(X_1 \cap X_2 = X_3) < \text{Min}(q(X_1), q(X_2))$$

The enhancement effect of the interaction of variables is evaluated using the indicator sq with the following equation.

$$sq = \frac{q_{X3} - \text{Max}(q_{X1}, q_{X2})}{\text{Max}(q_{X1}, q_{X2})} * 100\%$$

Where $q(X_1)$, $q(X_2)$, $q(X_1 \cap X_2 = X_3)$ are the calculated q values of factors X_1 , X_2 , and X_3 . $\text{Min}(q(X_1), q(X_2))$, $\text{Max}(q(X_1), q(X_2))$ denote the minimum and maximum values of q corresponding to X_1 and X_2 . The larger sq indicates that the greater the enhancement in the ability to explain the spatial heterogeneity of TB incidence when the two factors interact.

Results

Epidemiological characteristics

There were 4,313 TB cases newly reported in the four central districts (e.g., Yuexiu, Tianhe, Haizhu, and Liwan), accounting for about 30% of the total (14,100 cases) in Guangzhou in 2017. As shown in Table 2, more than 72% of the total TB cases were reported in the occupations of household and unemployment (33.1%), retired (23.9%), and commercial services (15.1%). Among the four age groups, the 19 – 45 years old population accounted for the highest proportion, and the young (0 – 18 years old) population had the lowest. These results showed that the TB prevalence in the four central districts was obviously featured by the distributions of age, and occupation.

Table 2
The age, and occupation distributions of TB cases in the four central districts.

Occupation	Guangzhou City		The four central region		
	No. of cases	Percentage [†]	No. of cases	Percentage [†]	Percentage [‡]
Household and unemployment	4176	29.6	1428	33.11	34.20
Farmer	2322	16.5	98	2.27	4.22
Retired	1675	11.9	1030	23.88	61.49
Worker	1333	9.5	276	6.40	20.71
business service	1155	8.2	652	15.12	56.45
Unknown	1122	8	182	4.22	16.22
Labor	627	4.4	26	0.60	4.15
Others	822	5.8	242	5.61	29.44
Student	491	3.5	212	4.92	43.18
Cadre	246	1.7	105	2.43	42.68
Catering service	131	0.9	62	1.44	47.33
	14100	100	4313	100.00	30.59
Age					
0~18	380	2.7	125	2.9	32.89
19~45	6751	47.9	1933	44.8	28.63
46~60	3519	25	1066	24.7	30.29
> 60	3450	24.4	1189	27.6	34.46
	14100	100	4313	100	30.59
Note: Among the occupations, other occupations include teachers, fishermen, herders, etc. with less than 100 cases in the whole Guangzhou city. † represents the number of cases with this attribute as a percentage of the total number of cases. ‡ represents the proportion of cases in the central districts to the number of cases in the corresponding occupation or age group in Guangzhou					

In addition to above features similar to those of the whole Guangzhou, several unique characteristics were also observed. The counts of TB cases of either over 60 years old or the occupation of household and unemployment almost accounted for one third of those in the whole city. Meanwhile, about 40% of either the student or the cadre (and office clerk) TB cases in this city were reported within the four districts. Among the TB cases in the business and catering service (1286 cases) of Guangzhou, more than a half was located in the central districts (714 cases). These analyses indicated that the study area could not only represent the features of TB prevalence in the whole city but also possess its own characteristics.

According to the Moran's I values derived from global spatial autocorrelation analysis (Additional file 1), the TB prevalence in the central area was obviously spatially differentiated. Of the five grid scales varying from 1km to 5km, the 2km × 2km grid possessed the best ability of characterizing the spatial distribution of TB prevalence across the central region. For this point, the following analyses were then conducted on the 2km × 2km grid scale.

Individual effect of selected variables

As illustrated in Fig. 2, the spatial distributions of 18 variables included in this study were obviously featured on the 2km × 2km grid scale. In particular, these high TB incidence regions were opportunely surrounded by some grids with relatively high TB incidence rates in 2016, more officially appointed medical institutions, convenient traffic system (more bus stops or subway stations), larger population density, higher PM_{2.5} concentration, and lower NDVIs. These results implied that the spatial patterns of gridded TB incidence rates were very likely to be associated with those of selected 18 factors.

According to the correlation coefficients (Table 3), the gridded TB incidence rates in 2017 were remarkably associated with the majority of independent variables. Among them, most of the socioeconomic factors presented significantly positive relationships with this epidemic ($0.37 < r < 0.76$, $P < 0.001$) except UV, GDP and Pub-serv. By contrast, this epidemic was only closely correlated with PM_{2.5} concentrations ($r < 0.21$, $P < 0.0.01$) among the five environmental factors. As a whole, the TB prevalence across the four districts of Guangzhou tended to be more heavily affected by socioeconomic factors rather than environmental conditions.

Table 3
Correlation coefficients between TB incidence and variables and the q-values derived from geographical detector analysis.

Variable Type	Variable Name	R	q values
Socioeconomic variables	Incid-2016	0.76***	0.57***
	Medical resources	0.67***	0.45***
	Hosp11	0.65***	0.49***
	Hosp12	0.54***	0.38***
	Bus	0.69***	0.44***
	Road-net	0.47***	0.26**
	Pop	0.31***	0.23***
	Subway	0.42***	0.19**
	Commercial	0.37***	0.16*
	Residential	0.37***	0.14**
	GDP	0.16	0.11*
	UV	0.17*	0.08
	Pub-serv	-0.01	0.02
Environmental variables	Temp	0.01	0.27***
	Prec	0.06	0.15***
	Humi	-0.07	0.10*
	PM _{2.5}	0.21**	0.08
	NDVI	-0.11	0.06

Note: r is the Pearson correlation coefficient. ***, **, and * indicates this value is significant at the level of 0.01, 0.05, and 0.10. Incid-2016:TB incidence in the previous year; Pop: Population density; GDP: Gross National Product per capita; Hosp: Counts of officially appointed medical institutions; Hosp11: outpatient hospitals of the officially appointed medical institutions; Hosp12: inpatient and outpatient hospitals of the officially appointed medical institutions; Road_net: Road network density; Subway: Counts of subway stations; Bus: Counts of bus stops; Residential: Percentage of residential land area; Commercial: Percentage of commercial service land; Pub-serv: Percentage of land for public services; UV: Percentage of urban village area; NDVI: Annual monthly average normalized differential vegetation index; PM_{2.5}: Annual monthly average PM_{2.5} concentration; Temp: Average temperature from March to June; Prec: Average precipitation from March to June; Humi: Average humidity from March to June.

As given in Table 3, some overall results were also derived from univariate analyses of the 18 factors by using the geographical detector. In terms of q-values, most of the socioeconomic variables with the exception of UV (q=0.08, $P>0.10$) and Pub-serv (q=0.02, $P>0.10$) possessed powerful ability ($0.11<q<0.57$, $P<0.10$) of interpreting the spatial differentiation of TB epidemic. In particular, three socioeconomic factors (e.g., the TB incidence rates in the previous year (2016), the counts of officially appointed medical institutions, and the

number of bus stops), accounted for more than 44% of the spatial heterogeneity of TB incidence rates in 2017 across the four central districts. In comparison, the environmental variables (i.e., the monthly averages of temperature, precipitation, and humidity) presented lower explanatory abilities ($0.10 < q < 0.27$, $P < 0.10$), although they were not closely associated with the TB incidence rates. These results further illustrated that socioeconomic factors posed greater impacts on the spatial heterogeneity of TB prevalence in the central region of Guangzhou.

Influences of pairwise interactions

According to the q_3 values varying from 0.16 to 0.89 (the lower left half in Fig. 3), the 153 pairwise interactions among the 18 variables posed much more powerful abilities of interpreting the spatial differentiations of TB prevalence than those of 18 individuals (the diagonal grids from the left top to the right bottom in Fig. 3). With regards to the sq values ranging from 7.34–311.63% (the higher right half in Fig. 3), the 153 pairwise interactions could be divided into three groups: 68 pairs ($sq < 50\%$), 49 pairs ($50\% < sq < 100\%$), and 36 pairs ($sq > 100\%$), among which about 55.56% of the total pairs owned notable enhancements ($sq > 50\%$). These results showed that the explanatory abilities of all the selected factors were remarkably strengthened by the pairwise interactions.

Meanwhile, the values of q_3 and Maximum (q_1 , q_2) were respectively used to classify these 153 pairwise interactions into five groups and three grades, yielding nine subgroups as given in Table 4. Within these three Maximum (q_1 , q_2) grades, there were 40 (below 0.2), 29 (0.2 – 0.4), and 31 (0.4 – 0.6) pairs of the selected factors possessing moderate enhancements of explanatory abilities, accounting for about 72.73%, 53.70%, and 93.94% of the second, third, and fourth q_3 grade, respectively. It meant that the explanatory abilities of potential factors were more likely to be enhanced to the slightly higher levels during the pairwise interacting.

Table 4
The numbers and sq values of 153 pairwise interactions in the nine subgroups

		Group 1	Group 2	Group 3	Group 4	Group 5	
Grade 1	Counts of pairs	1	40	4	\	\	45
	sq (%)	159.74	49.32-283.88	159.09-311.63	\	\	49.32-311.63
Grade 2	Counts of pairs	\	15	29	2	\	46
	sq(%)	\	21.4-66.13	17.32-121.17	67.04-145	\	21.4-145
Grade 3	Counts of pairs	\	\	21	31	10	62
	sq(%)	\	\	9.47-31.67	7.34-65.11	42.94-57.73	7.34-65.11
		1	55	54	33	10	153
Note: The q3 groups: below 0.2, 0.2 – 0.4, 0.4 – 0.6, 0.6 – 0.8, and beyond 0.8. The Maximum (q1, q2) grades: below 0.2, 0.2 – 0.4, and 0.4 – 0.6.							

Discussion

By using the geographical detector and other spatial analysis tool, a series of specific urban elements, as well as their pairwise interactions, were identified for the spatial heterogeneity of this epidemic across the four central districts in Guangzhou, from which some notable findings were achieved. This study would provide some useful clues for local authorities making more targeted interventions on this disease in Guangzhou and similar municipal regions of China.

Occupational difference of the cases was an obvious characteristic of TB prevalence. It has been reported that the farmers and workers accounted for the largest proportion of TB cases in some regions of China (e.g., north-east Yunnan Province, and Xi'an) [30–32]. On the contrary, these occupations failed to rank the first in some highly urbanized regions (e.g., Guangzhou and Foshan) while other occupations (i.e., the household and unemployed) possessed relatively higher percentages [33, 34], which was also observed in the central part of Guangzhou for the household and unemployed occupation (33.11%), the retired population (23.88%), as well as the business service (15.12%). However, the TB prevalence in this region was remarkably featured by its higher percentages of retired (61.49%), business (56.45%), catering service (47.33%), students (43.18%), and cadres' (42.68%) TB cases in the corresponding occupations in the whole city, which may be owed to its functional attributes (e.g., residential, commercial, educational, and service) [7, 21]. Meanwhile, this epidemic in the central region was also characterized by its slightly higher percentages of TB cases in different age groups, especially in the 60~ years old population (34.46%) due to the increasingly aging population [35, 36]. It can be thus seen that the TB prevalence in the four central districts possessed its own unique

epidemiological characteristics in addition to those in the whole city. Accordingly, we suggest that the features of TB epidemic closely related to the social and economic status should be heavily considered in order to draw up regional intervention plans (e.g., adequate propaganda and education on these specific occupations and population) for containing this disease.

Previous studies have already pointed out that the dominant influencing factors on the prevalence of some infectious diseases tended to change due to the varying research units [37, 38]. In our study, the 2km × 2km grid was chosen as the appropriate spatial scale, on which the spatial clustering distribution of TB prevalence across the central region were clearly distinguished, especially in the western part of Tianhe District and the junction area between Haizhu, Liwan, and Yuexiu districts owing to the grids with relatively higher incidence rates. Moreover, the spatial relation between the TB incidence rates and most of the selected factors were also easily observed so that the detection of potential influences on the TB prevalence's spatial differentiations was sufficiently conducted for identifying the specific urban elements in this study area. In a word, the choice of an appropriate spatial scale is fundamental for the identifications of spatial differentiation of the TB epidemic, as well as its influencing factors in the target region.

It has been revealed that the local TB prevalence was often determined by some socioeconomic indices including the population at risk of spreading this disease, population density and mobility, traffic system, economic status, and medical level at fine scales [6, 11, 15, 16]. Similar findings were obtained in our study that three socioeconomic variables (i.e., the incidence rate in the previous year (2016), the counts of officially appointed medical institutions, and the number of bus stops) posed larger impacts on the spatial differentiation of this epidemic across the central region of Guangzhou. There was a fourfold increase of transmission risk from some TB patients to their close contacts causing high exposure of the susceptible population to this disease [8], which may be a reasonable explanation for the strongest interpreting ability of TB incidence rate in the previous year. Another possible interpretation is that the recurrence of previous TB cases after treatment due to the increasing drug resistance of *M. tuberculosis* was very likely to increase the prevalence [39, 40], especially in the regions with high incidence rates in the previous year. As far as the count of officially appointed medical institutions is concerned, its heavy influences on the TB epidemic were probably correlated with being representative places for gathering various patients, including potential TB patients and susceptible people with low immunity [22]. Besides, the number of bus stops termed for public transportation condition was another non-neglected influencing factor for the TB epidemic because higher probability of the contacts among individuals and population mobility tended to be increased by the convenient public traffic system [16]. In general, the selection of these potential variables in this study was reasonable for identifying the dominant factors of TB epidemic in the central region. Then, we cautiously suggest that the treatment of current TB cases, as well as more effective methods dealing with the drug resistance, need to be considered first for reducing their potential impacts on the next year's TB prevalence, and that more resources should be well allocated for implementing the hospital infection management and reinforcing the propaganda and education for the individuals who often visits the hospitals or takes the buses.

In comparison to the individuals, their explanatory abilities were very likely to be extremely enhanced by their pairwise interactions[41, 42]. Rasam et al., and Ge et al., also found that the interactions, between public traffic system, population density, and urban functional areas, posed much higher abilities of accounting for the

spatial differentiations of TB prevalence than that of each individual [16, 15]. Our study obtained similar findings that the explanatory ability of each factor interpreting the spatial heterogeneity of TB epidemic in the central region of Guangzhou was extremely enhanced due to the pairwise interactions. In particular, the contributions of these weaker factors ($q < 0.2$) tended to be significantly enhanced while interacting with bus stops, officially appointed hospitals (i.e., Hosp, Hosp11, and Hosp12), and the incidence rate in previous year. Among these individuals, UV, termed for the widely distributed units with crowded population in the low buildings clustered in the study area [24], is a typical urban element posing impacts on the transmission of *M. tuberculosis* and TB infection [1, 43]. In general, the pairwise interactions made great contributions to the interpretation of the TB epidemic's spatial differentiations across the four central districts. We intensively recommend that these weaker elements must be comprehensively considered as necessary targets for making effective interventions, and that comprehensive prevention and control measures should be meticulously implemented in these regions with these paired elements for containing this disease in Guangzhou.

A few limitations should be mentioned here. First, although public transportation termed by bus stops and subway stations were included in this study, the mobility of the population was not adequately considered due to the difficulty of collecting the dynamics of population movement, which might be addressed in the future through obtaining and processing either cell phone data or public transportation smart cards. Second, although the geographical detector tool was employed to identify the specific influencing elements on this disease, it failed to explore the strength of pairwise interactions and its difference, for which an appropriate indicator need be developed in this tool so as to quantify the interactions. Finally, the TB case data in ten years or more longer period should be obtained in the future so as to further consolidate the current findings based on only one year's data.

Conclusion

In summary, a series of socioeconomic and environmental factors, as well as their pairwise interactions, were identified as specific urban elements posing remarkable impacts on the TB prevalence's spatial differentiations across the four central districts of Guangzhou. We accordingly suggest that more attentions should be paid to the regions with pairwise interactions of these specific elements in Guangzhou. This study would provide meaningful clues for local authorities making more targeted interventions on this disease in China's municipal areas featured by both high urbanization and severe TB epidemic.

Abbreviations

TB: Tuberculosis; *M. tuberculosis*: *Mycobacterium tuberculosis*; Incid-2016: TB incidence in the previous year; Pop: Population density; GDP: Gross National Product per capita; Hosp: Counts of officially appointed medical institutions; Hosp11: outpatient hospitals of the officially appointed medical institutions; Hosp12: inpatient and outpatient hospitals of the officially appointed medical institutions; Road_net: Road network density; Subway: Counts of subway stations; Bus: Counts of bus stops; Residential: Percentage of residential land area; Commercial: Percentage of commercial service land; Pub-serv: Percentage of land for public services; UV: Percentage of urban village area; NDVI: Annual monthly average normalized differential vegetation index;

PM2.5: Annual monthly average PM2.5 concentration; Temp: Average temperature from March to June; Prec: Average precipitation from March to June; Humi: Average humidity from March to June.

Declarations

Funding

HY R received the financial support from the National Natural Science Foundation of China (Grant NO.42071136 Grant No.41571158).

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

H-YR and W-LL conceived and designed the experimental concept. X-QL and H-CS helped collect data. H-YR and W-LL analyzed the data and drafted the paper. H-YR, X-QL and H-CS revised the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests

Author details

¹State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; ²College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China; ³Guangzhou Institute of Tuberculosis Control and Prevention, Guangzhou 510000, China; ⁴Guangzhou Chest Hospital, Guangzhou 510000, China.

References

1. Tao W. The Spatial Epidemiology of Tuberculosis in LinYi City, Shandong Province [Master]: Shandong University; 2012 (in Chinese).

2. Qian L, Wei-Xi J, Hui Z, Jun C, Sheng-Lan T, Wei-Bing W. Multi-source financing for tuberculosis treatment in China: key issues and challenges. *Infectious Diseases of Poverty*. 2021;10(02):44-9
3. Jeremiah C, Mishal K, Francine N, Eleni A, Fatima R, Peter M, et al. Global Tuberculosis Report 2020 - Reflections on the Global TB burden, treatment and prevention efforts. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases*. 2021
4. Coulter C. Infection control guidelines for the management of patients with suspected or confirmed pulmonary tuberculosis in healthcare settings. *Commun Dis Intell Q Rep*. 2016;40(3):E360
5. Zhezhe C, Dingwen L, Virasakdi C, Jinming Z, Mei L, Jing O, et al. Spatiotemporal patterns and ecological factors of tuberculosis notification: A spatial panel data analysis in Guangxi, China. *PloS one*. 2019;14(5)
6. Changmin I, Youngho K. Spatial pattern of tuberculosis (TB) and related socio-environmental factors in South Korea, 2008-2016. *PloS one*. 2021;16(8)
7. Erjia G, Xin Z, Xiaomeng W, Xiaolin W. Spatial and temporal analysis of tuberculosis in Zhejiang Province, China, 2009-2012. *Infectious Diseases of Poverty*. 2016;5(1)
8. Rasam ARA, Shariff NM, Dony JF, Othman F. Spatial and Statistics for Profiling Risk Factors of Diseases: A Case Study of Tuberculosis in Malaysia. *IOP Conference Series: Earth and Environmental Science*. 2019;385(1):012037 (9pp)
9. Khaulasari H. Modelling mixed geographically weighted poisson regression for tuberculosis disease in Surabaya. *Journal of Physics Conference Series*. 2020;1490:012004
10. Rasam ARA, Dony JF, Shariff NM. Local Spatial Knowledge for Eliciting Risk Factors and Disease Mapping of Tuberculosis Epidemics. *Environment-Behaviour Proceedings Journal*. 2020;5(S12):45-51.
11. Hailu DM, Gemechu DD, Bernt L. Identifying geographical heterogeneity of pulmonary tuberculosis in southern Ethiopia: a method to identify clustering for targeted interventions. *Global health action*. 2020;13(1)
12. Mingyao C, Miaoying Y, Adenay-Erkin. Analysis on influence of climatic factors on the incidence of tuberculosis. *Chines Journal of Disease Control & Prevention*. 2014;18(11):1051-4 (in Chinese).
13. Lai T-C, Chiang C-Y, Wu C-F, Yang S-L, Liu D-P, Chan C-C, et al. Ambient air pollution and risk of tuberculosis: a cohort study. *Occupational and Environmental Medicine*. 2016;73(1)
14. Lijia Y. Study on Spaial Distribution Charasteristics and Relative Factors of Tuberculosis [Master]: Chinese Center for Disease Control and Prevention; 2020 (in Chinese).
15. Rasam ARA, Shariff NM, Dony JF. Geospatial-Based Model for Diagnosing Potential High-Risk Areas of Tuberculosis Disease in Malaysia. *MATEC Web of Conferences*. 2019;266
16. Ge E, Lai P-C, Zhang X, Yang X, Li X, Wang H, et al. Regional transport and its association with tuberculosis in the Shandong province of China, 2009–2011. *Journal of Transport Geography*. 2015;46
17. Liu MY, Li QH, Zhang YJ, Ma Y, Liu Y, Feng W, et al. Spatial and temporal clustering analysis of tuberculosis in the mainland of China at the prefecture level, 2005-2015. *INFECTIOUS DISEASES OF POVERTY*. 2018;7
18. Wang L, Xu CD, Hu MG, Qiao JJ, Chen W, Li T, et al. Spatio-temporal variation in tuberculosis incidence and risk factors for the disease in a region of unbalanced socio-economic development. *BMC PUBLIC HEALTH*. 2021;21(1)

19. Hao L, Miao G, Mingxin Z. Spatio-temporal Distribution of Tuberculosis and the Effects of Geographical Environmental Factors in China. *BMC Infectious Diseases*. 2021
20. Cheng C, Hongyan R, Lu Z, Dafang Z. Street Space Quality Evaluation in Yuexiu District of Guangzhou City based on Multi-feature Fusion of Street View Imagery. *Journal of Geoinformation Science*. 2020;22(06):1330-8 (in Chinese).
21. Shuhui H. The study on the Characteristics of Commercial Space Agglomeration in the Inner City of Guangzhou [Master]: South China University of Technology; 2019 (in Chinese).
22. Ito Y, Nagao M, Iinuma Y, Matsumura Y, Mishima M. Risk factors for nosocomial tuberculosis transmission among health care workers. *American Journal of Infection Control*. 2016;44(5)
23. Gong P, Chen B, Li X, Liu H, Wang J, Bai Y, et al. Mapping essential urban land use categories in China (EULUC-China): preliminary results for 2018. *Science Bulletin*. 2020;65(3):182-7
24. Ren H, Wu W, Li T, Yang Z. Urban villages as transfer stations for dengue fever epidemic: A case study in the Guangzhou, China. *PLOS Neglected Tropical Diseases*. 2019;13(4)
25. Zhao L, Ren H, Cui C, Huang Y. A Partition-Based Detection of Urban Villages Using High-Resolution Remote Sensing Imagery in Guangzhou, China. *Remote Sensing*. 2020;12(14):2334
26. Hongyan R, Duanyang X, Xiaoming S, Jianwei X, Dafang Z, Gonghuan Y. Characterisation of gastric cancer and its relation to environmental factors: a case study in Shenqiu County, China. *International journal of environmental health research*. 2016;26(1)
27. LI Yingyue LZ. Review on the Development of Digital City Management in China. *Intelligent Building & City Information*. 2017(02):28-32 (in Chinese).
28. Anselin L, Getis A. *Spatial Statistical Analysis and Geographic Information Systems*. *Advances in Spatial Science*. 2010
29. Wang J-F, Li X-H, Christakos G, Liao Y-L, Zhang T, Gu X, et al. Geographical Detectors-Based Health Risk Assessment and its Application in the Neural Tube Defects Study of the Heshun Region, China. *International Journal of Geographical Information Science*. 2010;24(1)
30. Huang L, Abe EM, Li X-X, Bergquist R, Xu L, Xue J-B, et al. Space-time clustering and associated risk factors of pulmonary tuberculosis in southwest China. *BioMed Central*. 2018;7(1)
31. Xiaoting D, Xiaomei X, Rui W, Chaofeng M, Jifeng L, Yongmin X. Analysis of the epidemiological characteristics of Tuberculosis in Xi'an City from 2005 to 2014. *Modern Preventive Medicine*. 2016;43(16):3033-5+43 (in Chinese).
32. Qingang W, Dejiang L, Hongyun G, Weiguo T, Xiaoling C, Chunrong L, et al. Analysis of epidemiological characteristics of pulmonary tuberculosis in Shenzhen from 2007 to 2016. *Journal of Tropical Medicine*. 2018;18(01):86-9 (in Chinese).
33. Hongcheng S, Yuhua D, Guifeng W, Yu L, Tiegang L, Jianxiong L. Epidemiological characteristics of newly reported pulmonary tuberculosis from 2008 to 2018 in Guangzhou. *Chinese Journal of Antituberculosis*. 2020;42(04):322-6 (in Chinese).
34. Wei W, Yinong Y, Zhilong W, Qianhong Z, Liang C, Fei H, et al. Analysis of the incidence trend and epidemic characteristics of tuberculosis in Foshan, 2010-2020. *Disease Surveillance*. 1-8 (in Chinese).

35. Zhou C, Tong X, Wang J, Lai S. Spatial differentiation and the formation mechanism of population aging in Guangzhou in 2000-2010. *Geographical Research*. 2018;37(1):103-18
36. Zhou C, Li Y, Tong X. Spatial change and influence factors of population aging in Guangzhou during 2000-2010. *Acta Scientiarum Naturalium Universitatis Sunyatseni*. 2016;55(1):114-22
37. Danson FM, Graham AJ, Pleydell DRJ, Campos-Ponce M, Giraudoux P, Craig PS. Multi-scale spatial analysis of human alveolar echinococcosis risk in China. *Parasitology*. 2004;127(S1):S133-S41
38. Rican S, Salem G, Jouglu E. Spatial variation of lung cancer death in France and spatial scales: importance of changing scale in health geography. *Revue d'Épidémiologie et de Santé Publique*. 1999;47(5):411-20
39. Pramudita A, Rumende CM, Findyartini A. Fixed-dose combination antituberculosis therapy as a risk factor for tuberculosis recurrence: an evidence-based case report. *Acta Medica Indonesiana*. 2017;49(2):175
40. Mirsaeidi M, Sadikot R. Patients at high risk of tuberculosis recurrence. *International Journal of Mycobacteriology*. 2018;7(1):1-
41. Tian M, Dong J, Mengmeng H, Peiwei F, Shize Z, Gongsang Q, et al. Geographical Detector-based influence factors analysis for Echinococcosis prevalence in Tibet, China. *PLoS neglected tropical diseases*. 2021;15(7)
42. Zheng C, Tao L, Li X, Wang J, Ma W. Individual and Interactive Effects of Socio-Ecological Factors on Dengue Fever at Fine Spatial Scale: A Geographical Detector-Based Analysis. *International Journal of Environmental Research and Public Health*. 2017;14(7):795
43. Lai Keng, Wei X, Yu Hua D, Yu L, Gui Feng W, Jie Ying Y, et al. Spatiotemporal characteristics on smear-positive pulmonary tuberculosis in Guangzhou city from 2014 to 2019. *Journal of Tropical Medicine*. 2021;21(06):782-5+815 (in Chinese).

Figures

Figure 1

Illustration of study area with the spatial distribution of TB cases in 2017

Figure 2

Spatial distribution of TB and included 18 variables in this study

Figure 3

Illustration of the q_3 and sq values for the pairwise interactions among the selected factors Note: Incid-2016: TB incidence in the previous year; Pop: Population density; GDP: Gross National Product per capita;

Hosp: Counts of officially appointed medical institutions; Hosp11: outpatient hospitals of the officially appointed medical institutions; Hosp12: inpatient and outpatient hospitals of the officially appointed medical institutions; Road_net: Road network density; Subway: Counts of subway stations; Bus: Counts of bus stops; Residential: Percentage of residential land area; Commercial: Percentage of commercial service land; Pub-serv: Percentage of land for public services; UV: Percentage of urban village area; NDVI: Annual monthly average normalized differential vegetation index; PM2.5: Annual monthly average PM2.5 concentration; Temp: Average temperature from March to June; Prec: Average precipitation from March to June; Humi: Average humidity from March to June.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1.docx](#)