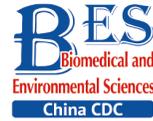


Letter to the Editor

**Spatial-temporal Analysis of Tuberculosis at the Community Level in the Baoshan District, Shanghai 2014–2019***MA Zhi Tao¹, WANG Na¹, HE Fan¹, LIU Xiao Feng^{1,#}, and SHEN Xin^{2,3,#}

Mycobacterium tuberculosis is the primary bacteria that leads to human tuberculosis (TB), which usually affects the pulmonary organs. TB is one of the top 10 causes of death worldwide; however, China is one of eight countries that together account for two-thirds of the total number of cases globally. The traditional analysis of TB is limited by discovering outbreaks or clusters in the spatial-temporal dimension^[1]. The spatial-temporal analysis could be a more effective method based on Monte Carlo statistics for detecting underlying outbreaks of TB at the community level, however, few studies have been conducted with the space-time scan analysis at the community level in China.

This study gathered TB data from the National Notifiable Infectious Disease Reporting Information System (NNIDRIS), which was established by the Chinese Center for Disease Control and prevention (China CDC) for surveillance of diseases. The purely temporal scan, spatial autocorrelation, and space-time scan statistical analysis was used to search for TB clusters at the community level in spatial-temporal dimension from 2014 to 2019 in the Baoshan district, Shanghai. Eventually, various control measures were adopted based on the analysis combining the classic analysis with space-time scan statistical analysis, which was utilized to rank the different risks of communities distributed in the clusters classified.

Shanghai is a municipality with 16 districts distributed in east of China, one of which is the Baoshan district that governs 12 communities, is located in north of Shanghai, and contains an area of 424.56 km² (Supplementary Figure S1, available in www.besjournal.com). In 2019, the population in Baoshan district numbered 2,080,963.

Totally 3,343 TB cases in the study were pulmonary tuberculosis (PTB) cases except MDR-TB (Multi-drug resistance TB), Mono-resistance TB and Poly-resistant TB, resided in Baoshan district from 2014 to 2019, which were confirmed with criteria announced in 2008 by the Ministry of Health of the People's Republic of China^[2]. The PTB cases in this study were clinically-diagnosed and laboratory-confirmed by the TB-designated hospital in Shanghai, and reported by the medical staff who were specialized in notification and submitting of TB to the NNIDRIS. The data of population and social-health-economy characteristics distributed in the 12 communities was gathered from the Statistical Yearbooks of Baoshan district^[3], such as the sex, age, residence address, current address, onset date. The notifiable date of confirmed PTB cases were screened from the system. The administrative code of current address at the community level was consistent with the area code from the polygon map based on a 1:1,000,000 scale layer, and the community's central point were selected as the representation of a potential cluster, the longitude and latitude coordinates of which were determined through the Google geocoding service and toolbox of Geoprocessing in ArcGIS v.10.8 (ESRI Inc, Redlands, CA, USA).

The dimensions of purely temporal scan and spatial autocorrelation analyses were confined to temporal and spatial measures, respectively. To merge the spatial and temporal dimensions, we utilized Kulldorff's space-time scan statistic for space-time scan analysis, which could be realized by SaTScanTM v9.6^[4]. Kulldorff's space-time scan statistic is an analysis for detecting and evaluating local clusters in a study area with spatial-temporal dimension. The statistic produces a cylindrical

doi: 10.3967/bes2022.062

*This work was supported by the National Natural Science Foundation of China [81872679 to SHEN Xin] and the Shanghai Municipal Project for Academic Leaders in Public Health [GWV-10.2-XD23 to SHEN Xin]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

1. Department of Tuberculosis and HIV/AIDS Prevention, Baoshan Centers for Disease Control and Prevention, Baoshan 201901, Shanghai, China; 2. Division of Tuberculosis and HIV/AIDS Prevention, Shanghai Municipal Center for Disease Control and Prevention, Shanghai 200336, China; 3. Shanghai Institutes of Preventive Medicine, Shanghai 200336, China

window containing a circular geographic base and height linking to time scanning. The cylindrical window moves in space along with time for searching likely clusters through a 999-times replicated simulation. The hypothesis was shaped in this study to determine whether the risk of being a confirmed TB case in communities inside the cluster(p) appear larger compared to the one outside(q). The spatial-temporal risk can be assessed through the null hypothesis ($H_0:p = q$) and the alternative hypothesis ($H_1:p > q$). The relative risk (RR) was utilized to estimate the risks inside and outside the cluster, which is calculated with the following equation: $RR = \frac{p}{q}$. The log likelihood ratio (LLR) was used to calculate each simulative cylindrical window, and the sum of the $LLRs$ ($\sum_i LLR_i$) in which the observed cases were larger than the one with expected cases among 999 simulative replications applied to assess the probability of the cluster that appeared. The maximum sum of $LLRs$ ($Max \sum_i LLR_i$) with more expected TB cases indicated the most likely cluster; the second largest LLR indicated the second likely cluster. The relative likelihood (RL) evaluates the significance of the TB cluster, which is the value of the LLR inside the cluster to the overall $LLR(L_0)$. The RL produced as the cylindrical window move in this study and the rank of the RL was generated to assess the statistical significance of the TB cluster at the 0.05 level ($\alpha = 0.05$). In our space-time statistical analysis, limits were set such that the expected TB cases were no more than the total TB cases and the scan time range was no more than one-half of the study period in case that $Max \sum_i LLR_i$ did not reflect the difference between the risk of infection inside of the cluster and the risk outside for overlapping most communities and the study period.

The PTB is a high burden disease, the spread and treatment of which are strongly influenced by social-economy status^[5]. Previous studies explored the risks of PTB in association with economy status and demography characteristics inside the cluster in the study area on the prefecture level^[1]. We involved the indicators of social-economy status and population structure, such as annual average values of GDP per capita, native-total population ratio (NTR), and the lowest society guarantees rate (LSGR) over 6 years span. The risk rank of every community was defined as a 3-1 values based on corresponding RR values through the space-time scan statistical analysis over 6 years span, which depends on every community's

locations inside the most likely cluster ($RR > 1$), inside the second-like cluster ($0.54 < RR < 1$), and outside the cluster ($RR < 0.54$) respectively. The Spearman rank correlation test was used to explore the correlations of the risk and social-economy-demography influencing factors with R [v4.1.0; R Foundation for Statistical Computing, Vienna, Austria (<https://www.r-project.org/>)].

A total of 3,343 TB cases were included in the study. The average annual notification rate in the Baoshan District was 27.09/10⁵ from 2014 to 2019. The number and notification rate of male cases were both larger than the female cases from 2014 to 2019 and the ratio of males to females was roughly greater than 2 (Supplementary Table S1, available in www.besjournal.com). A larger notification rate appeared among the cases > 60 years of age compared to the other age groups. In addition, the number and notification rate of native cases were larger than migrant cases. Supplementary Figure S2 (available in www.besjournal.com) presents the distribution of PTB cases in the Baoshan district from 2014 to 2019. Few regular patterns appeared and the peak of distribution mostly presented in the summer after April. This study showed that the characteristics of the 20–30 year age group, > 60 year age group, and natives possessed a larger proportion compared to other characteristics in this study, which supports the findings of previous studies; specifically, the high prevalence existing in the elderly and native populations, and the elderly population presented a limited immunity and a worse treatment outcome compared to the other age groups, which may possess a higher risk in transmission of PTB^[6].

Figure 1 presents the results of purely temporal scan and seasonal temporal analysis. The former showed that a temporal cluster started from April 17th to May 1st, 2014 (Figure 1A), the RR and LLR of which were 2.36 and 15.08, respectively. Seasonal temporal analysis indicted a seasonal cluster from March to August (Figure 1B), accordingly, the RR and LLR were 1.20 and 13.43, respectively. The P values of purely temporal scan and seasonal temporal analysis were both < 0.001. The seasonal cluster of PTB was from March to August, considering the uncertain incubation period of TB, the onset time of which started nearly 3 months later, the case usually infect from December last year to next May, which corresponded to the seasonal pattern analysis with the ARIMA model in Li et al.^[7] There is evidence that exposure to air pollution with particulate matter with an aerodynamic diameter $\leq 10 \mu\text{m}$ (PM_{10}),

sulfur dioxide (SO₂), and nitrogen dioxide (NO₂) increase the PTB risk^[8]. A previous study revealed that ambient air pollutants are associated with the seasonality of TB in Shanghai, China^[9]. The pure temporal scan analysis showed that a temporal cluster started from April 17th to May 1st, 2014, and the specific period was consistent with the seasonal pattern from March to August, which indicated the rising notification rates of PTB cases and peak cluster in the study area for uncertain incubation of TB, exposure to air pollution, crowding in airtight room and delay on seeing a physician during the Lunar New Year^[8].

The spatial distribution was presented with a notification rate from 2014 to 2019 at the community level, and the cut-off values of the legends during the 6 years were aggregated and adjusted to a unique value in one legend (Supplementary Figure S3, available in www.besjournal.com). The highest-level notification rates were distributed in Wusong and Gaojing communities in 2014, Zhangmiao community in 2015 and 2018, and Luojin community in 2019, respectively. The notification rates of Luodian, Youyi, Miaohang, Yanghang, and Dachang communities were relatively lower.

The result of global spatial autocorrelation analysis indicated the possible clusters with Global

Moran's over 6 years (Supplementary Table S2). No apparent cluster with statistical significance was found through global spatial autocorrelation analysis.

Supplementary Figure S3 showed the clusters in spatial-temporal dimension, which presented at the community level annually from 2014 to 2019. Five statistically significant clusters with P values < 0.05 were found among a total of 12 clusters; the RR values were all > 1 (Table 1). An outbreak of PTB occurred among students at a school on September 28, 2015^[10] which located in the community inside one of five statistically significant clusters through the annual space-time scan analysis in the Baoshan district. A statistically significant cluster suggested a potential outbreak at the community level inside the cluster, and the space-time scan analysis proved to be helpful and effective in detecting a real outbreak. When a statistically significant cluster was found through annual space-time scan analysis, the measures of investigation and controlling intervention adopted among the communities inside the cluster could be more effective and accurate during the early stage of the outbreak compared to the general intervention in the entire district with less resource consumption.

PTB is a disease that is strongly influenced by social, economic, demographic characteristics^[5], then the GDP per capita, NTR, and LSGR are involved. The Spearman's rank correlation test was used to test the associations of social, economic, and demographic characteristics with ranked risk base on the space-time scan analysis over a 6-year span in the study area.

Consequently, the results of the Spearman's rank correlation test showed that the annual average GDP per capita ($r_{k-GDP} = -0.670, P < 0.05$), NTR ($r_{k-MTR} = 0.658, P < 0.05$), and LSGR ($r_{k-LSGR} = 0.703, P < 0.05$) presented associations with risk rank of the communities at a statistically significant level in the study area (Supplementary Table S3). The most likely cluster included four communities, and the value of RR was > 1 with a P value < 0.05. The secondary likely cluster was located in one community, and the RR value was < 1 with a P value = 0.047. The communities distributed in the most likely cluster and secondary likely cluster possessed a lower GDP per capita, higher NTR, and higher LSGR compared to the communities outside, which is consistent with the results of the Spearman's rank correlation test (Figure 2).

Bases on the findings in this study, 4 measures were appropriate for the study area below: 1)

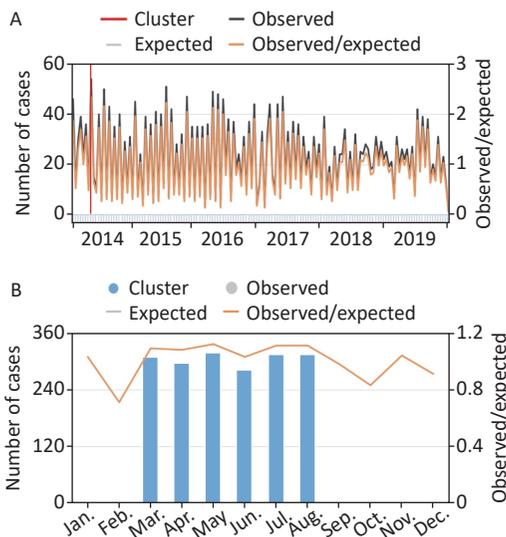


Figure 1. (A) Purely temporal scan analysis graph showing the mostly likely cluster in the time dimension. (B) Seasonal temporal analysis graph presenting the most likely seasonal cluster with the seasonal pattern among 6 years. Graphs were from the HTML produced by SatScan.

merging a neo-approach like spatial-temporal analysis into active surveillance for detecting a potential outbreak; 2) enhancing personal protection, like wearing a mask outdoors from

winter-to-next spring, based on the seasonal cluster discovered; 3) spreading PTB-related knowledge among a native population based on the social characters associated with the risk of PTB; and 4)

Table 1. Annual space-time scan statistical analysis of PTB in Baoshan district, Shanghai from 2014 to 2019

Year	Involved communities [†]	Period	Observed case	Expectedcase	P	LLR	RR
2014	ZM, GJ, SN, YH, YY, WS	5/1–6/4	157	109.40	0.0029*	11.52	1.593
	GU	9/18–10/2	13	3.68	0.1350	7.16	3.59
2015	LD, YH, YP, MH, LJ, GC	6/20–7/4	30	10.74	0.0025*	11.90	2.90
	ZM	1/1–6/4	38	19.20	0.1310	7.48	2.05
2016	GJ, ZM, SN, YH, YY, WS	4/21–5/5	28	10.03	0.0082*	11.06	2.88
2017	ZM	7/5–9/2	23	7.32	0.0058*	10.88	3.24
	LD, LJ, YP	11/17–12/1	14	4.36	0.2120	6.77	3.27
2018	LD, YP, GC	1/1–1/20	21	8.91	0.3550	6.06	2.41
	ZM	1/1–5/5	29	14.28	0.3620	6.03	2.09
2019	LD, LJ	6/20–7/4	12	2.49	0.0100*	9.46	4.91
	DC, MH, ZM	6/20–9/2	56	35.00	0.4130	5.77	1.67
	YH, SN, WS, YY	2/20–3/6	15	6.03	0.7490	4.77	2.53

Note. [†] Abbreviation for community; WS, Wusong; YY, Youyi; DC, Dachang; LD, Luodian; YH, Yanghang; YP, Yuepu; SN, Songnan; GJ, Gaojing; MH, Miaohang; LJ, Luojin; ZM, Zhangmiao; GC, Gucun; LLR, log likelihood ratio; RR, relative risk; *P < 0.05.

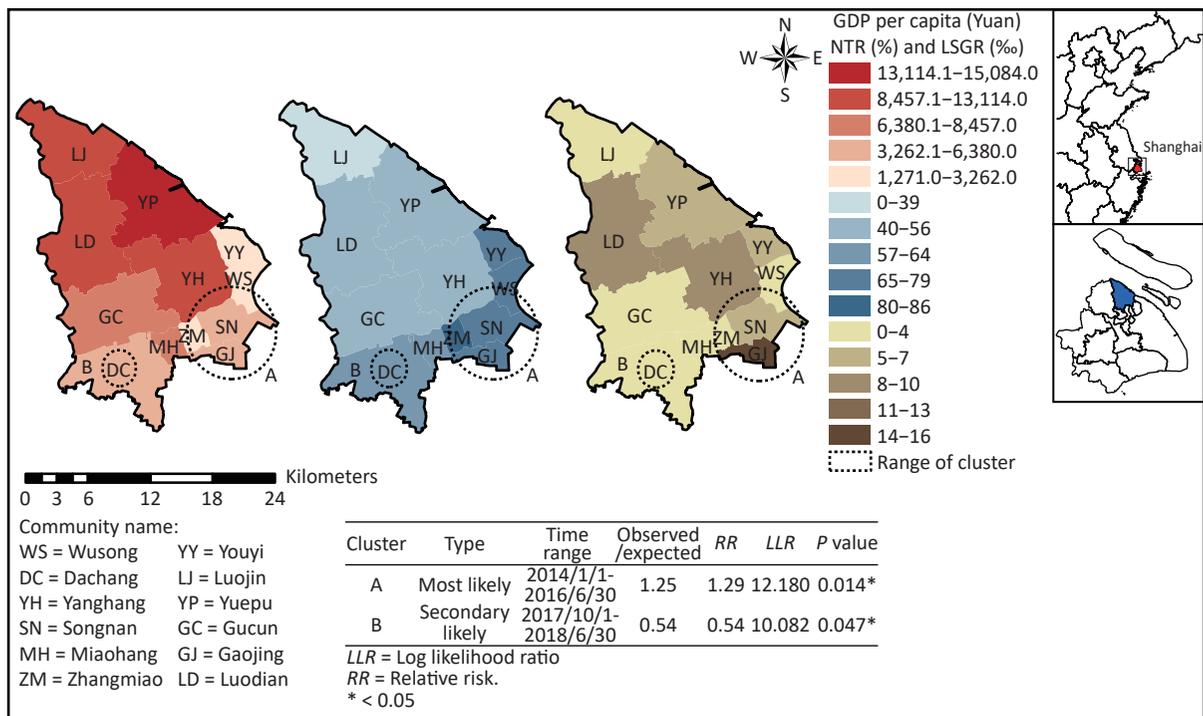


Figure 2. The PTB clusters through space-time scan statistical analysis and distributions of annually average GDP per capita (left), native-total population ratio (middle) and lowest society guarantees rate (right) over a 6-year span in the Baoshan district, Shanghai.

improving economic status and LSGR distributed in low-GDP communities based on the economic factor identified significantly associated with risk. We recommend the 4 measures prior to the stakeholders or politicians of high risk communities inside the statistically significant clusters, aiming to control the spreading trend of a potential outbreak and impacting the range of PTB risk. Eventually, by combining the spatial-temporal analysis with active surveillance for factor analysis, a latent outbreak or cluster will be handled at the community level with less time compared to the classic method of surveillance-investigation-controlling.

There were some limitations for this study. First, the sample size at the community level was confined, by which the power of the test may have been affected. Second, the variables involved for identifying association with risk of PTB was limited and more potential factors with a statistically significant trend may exist.

In conclusion, this study detected a seasonal cluster of PTB ranges from March to August, which conformed the pattern of air pollutants. The PTB risk-rank cluster was discovered, and the GDP per capita, NTR, and LSGR were identified as being significantly associated with PTB risk rank of communities in the Baoshan district of Shanghai. We recommend 4 measures prior to the high risk communities to control the next potential TB outbreak, and the effectiveness of the 4 measures being implemented could be verified. This study will bring a new hypothesis: Whether the analysis and corresponding measures can be applied to other airborne diseases, like Covid-19, which will be determined in the future.

Acknowledgements This study was supported by the Center of Disease Control and Prevention in Baoshan district, Shanghai. The authors would like to thank Jeanette J. Rainey of the United States Centers for Disease Control and Prevention (US CDC) for technical support. We appreciate Dr. Lei Zhou from the Chinese Center for Disease Control and Prevention for her support and help with the methodology and study design. All authors are grateful for medical staffs' dedications to TB control, protection of the health in the Tuberculosis

Prevention and Clinic Departments of the Baoshan district.

Conflicts of interest All authors declare that they have no competing interests.

Author contributions ZM and XS designed the research. ZM, NW, and FH collected data. ZM wrote the first draft of the manuscript and performed the analysis. ZM prepared all figures and tables. ZM and XS interpreted the results. ZM, XS, and XL reviewed and revised the manuscript. All authors read and accepted the final manuscript.

[#]Correspondence should be addressed to LIU Xiao Feng, Tel: 86-21-33796707; E-mail: liuxiaofeng@bscdc.org.cn; SHEN Xin, PhD, Tel: 86-21-62758710; E-mail: shenxin@scdc.sh.cn

Biographical note of the first author: MA Zhi Tao, Male, born in 1988, MPH, majoring in disease control and epidemiology.

Received: November 29, 2021;

Accepted: April 11, 2022

REFERENCES

1. Liu MY, Li QH, Zhang YJ, et al. Spatial and temporal clustering analysis of tuberculosis in the mainland of China at the prefecture level, 2005-2015. *Infect Dis Poverty*, 2018; 7, 106.
2. Ministry of Health of the People's Republic of China. Diagnostic criteria for pulmonary tuberculosis (WS 288-2008). 2008. (In Chinese)
3. Baoshan Statistics Bureau, Baoshan investigation team of National Bureau of Statistics. Baoshan statistical yearbook 2020. Baoshan Statistics Bureau. 2021. (In Chinese)
4. Kulldorff M. SaTScan User Guild for v. 9.6. 2018.
5. World Health Organization. Global tuberculosis report 2020. <https://www.who.int/publications/i/item/9789240013131>. [2020-10-15].
6. Oshi DC, Oshi SN, Alobu I, et al. Profile and treatment outcomes of tuberculosis in the elderly in southeastern Nigeria, 2011-2012. *PLoS One*, 2014; 9, e111910.
7. Li YS, Zhu LM, Lu W, et al. Seasonal variation in notified tuberculosis cases from 2014 to 2018 in eastern China. *J Int Med Res*, 2020; 48, 300060520949031.
8. Xiang K, Xu ZW, Hu YQ, et al. Association between ambient air pollution and tuberculosis risk: a systematic review and meta-analysis. *Chemosphere*, 2021; 277, 130342.
9. Wang H, Tian CW, Wang WM, et al. Temporal cross-correlations between ambient air pollutants and seasonality of tuberculosis: a time-series analysis. *Int J Environ Res Public Health*, 2019; 16, 1585.
10. Yang XT, Xiang LH, Liu XF, et al. The investigation of an outbreak on TB at school in Baoshan district, Shanghai. *Chin J School Health*, 2016; 37, 1902-4. (In Chinese)