Spatial clustering of drug-resistant tuberculosis in Hunan province, China: an ecological study

Kefyalew Addis Alene,1,2 Zuhui Xu,3 Liqiong Bai,4 Hengzhong Yi,5 Yunhong Tan,5 Darren Gray,6 Kerri Viney,6,7,8 Archie CA Clements1,2

ABSTRACT

Objective This study aimed to investigate the spatial distribution of drug-resistant tuberculosis (DR-TB) in Hunan province, China.

Methods An ecological study was conducted using DR-TB data collected from the Tuberculosis Control Institute of Hunan Province between 2012 and 2018. Spatial clustering of DR-TB was explored using the Getis-Ord statistic. A Poisson regression model was fitted with a conditional autoregressive prior structure, and with posterior parameters estimated using a Bayesian Markov chain Monte Carlo simulation, to quantify associations with possible risk factors and identify clusters of high DR-TB risk.

Results A total of 2649 DR-TB patients were reported to Hunan Tuberculosis Control Institute between 2012 and 2018. The majority of the patients were male (74.8%, n=1983) and had a history of TB treatment (88.53%, n=2345). The proportion of extensively DR-TB among all DR-TB was 3.3% (95% CI 2.7% to 4.1%), which increased from 2.8% in 2012 to 4.4% in 2018. Of 1287 DR-TB patients with registered treatment outcomes, 434 (33.8%) were cured, 198 (15.3%) completed treatment, 92 (7.1%) died, 108 (8.3%) had treatment failure and 455 (35.3%) were lost to follow-up. Half (50.9, n=655) had poor treatment outcomes. The annual cumulative incidence rate of notified DR-TB increased over time from 0.25 per 100 000 people in 2012 to 0.83 per 100 000 people in 2018. Substantial spatial heterogeneity was observed, and hotspots were detected in counties located in the North and East parts of Hunan province. The cumulative incidence of notified DR-TB was significantly associated with urban communities.

Conclusion The annual incidence of notified DR-TB increased over time in Hunan province. Spatial clustering of DR-TB was detected and significantly associated with urbanisation. This finding suggests that targeting interventions to the highest risk areas and population groups would be effective in reducing the burden and ongoing transmission of DR-TB.

INTRODUCTION

Tuberculosis (TB) remains a significant global public health problem that kills more than one million people every year and over two billion people during the last two centuries.1 2 In 2018, one-third of the world’s population were estimated to have latent TB infection, and approximately 10 million people developed TB disease.3 The United Nations (UN) Sustainable Development Goals and the WHO’s End TB Strategy have a common target of ending the global TB epidemic by 2030.3 4 However, the emergence of multidrug-resistant tuberculosis (MDR-TB), defined as TB that is resistant to at least isoniazid and rifampicin, has become a major obstacle to achieve these ambitious targets. Globally, approximately 3.4% of new TB cases and 18% of previously treated cases have MDR-TB or rifampicin-resistant TB (MDR/RR-TB).5 About half a million new cases of MDR/RR-TB were estimated to occur in 2018.6

China has the second-largest global burden of DR-TB (after India), accounting for 14% of the global TB burden.7 The country is striving to reduce the magnitude of this burden in line with the targets of the global End-TB Strategy and the national TB strategic plan.8 DR-TB in China is unevenly distributed by geographic area and population subgroup, with a higher incidence reported among males, minority ethnic groups, those living in poverty and people with chronic medical
conditions such as HIV and diabetes mellitus. The identification of potential risk factors and hotspot areas where TB is concentrated should allow policy makers to implement targeted interventions aimed at DR-TB prevention and management. This might be particularly important in high DR-TB burden countries such as China as the epidemic is concentrated in certain areas. However, the spatial distribution of DR-TB has not been studied in some provinces of China, including in Hunan province.

Hunan is one of the provinces in China with a high burden of DR-TB. The proportions of DR-TB among all patients with TB and previously treated patients with TB in Hunan were 10.5% and 28.8%, respectively. Our previous studies showed that successful treatment outcomes for DR-TB patients in Hunan Province are low at 57%, suggesting that additional efforts are needed to control and prevent the disease. To inform these efforts, the distribution of DR-TB in terms of place and time as well as the drivers of transmission should be investigated and documented. This study explored the spatial distribution of DR-TB and identified the areas where DR-TB is concentrated in Hunan province. The study aims to provide additional evidence for policy-makers that can help develop targeted interventions to reduce the burden of DR-TB in the province.

METHODS
Study area
This study was conducted in Hunan province, 1 of the 22 provinces in China, located in South Central China along the middle reaches of the Yangtze River. Hunan is the seventh largest province in China with a total population of approximately 72 million people and with a total area of 211,800 square kilometres. Hunan province is administratively divided into fourteen prefecture-level divisions (13 prefecture-level cities and one autonomous prefecture); and further subdivided into 122 county-level divisions (35 districts, 17 county-level cities, 63 counties, 7 autonomous counties). Those are in turn divided into 2,587 township-level divisions (1,098 towns, 1,158 towns, 98 ethnic townships, 225 subdistricts and 8 district public offices).

DR-TB diagnosis and treatment in Hunan province
DR-TB diagnosis and treatment regimen in Hunan Province has been described detailly elsewhere. Briefly, of the 131 counties in Hunan province, 32 counties are able to provide comprehensive diagnostic services, which include culture. However, drug susceptibility testing (DST) is mainly carried out in the Hunan Chest Hospital. Thus, sputum specimens from all culture-positive TB patients from throughout the province are referred to the Hunan Chest Hospital for DST. In the hospital, phenotypic DST based on solid and liquid culture techniques, and molecular methods using line probe assays as well as Xpert MTB/RIF are performed. Patients with confirmed DR-TB are admitted to the MDR-TB treatment centre at the Hunan Chest Hospital for treatment and management. Patients were treated with an individualised treatment regimen based on their DST results and history of previous TB treatment. The regimen at the time of this study includes an injectable agent, a fluoroquinolone, para-aminosalicylic acid, prothionamide, pyrazinamide, clarithromycin, ethambutol or cycloserine. The duration of treatment was 24 months for patients with MDR-TB and 30 months for patients with extensively DR-TB. The injectable drugs were used for a minimum of 6 months for MDR-TB patients and 12 months for XDR-TB patients. Patients were admitted to the hospital for 1–2 months during the intensive phase. During this time, patients received psychological support and counselling from hospital nurses. When the patients are medically fit, they were treated as outpatients. They received support from trained family members or from trained supervisors in the community and returned to the hospital once a month for a drug refill until the end of treatment.

Study design and data sources
The study was conducted at the county level, using an ecological study design. DR-TB data were obtained from the internet-based TB management information system managed by the Hunan Tuberculosis Control Institute. The study included information from patients with RR-TB, MDR-TB and XDR-TB that were reported to the Hunan Tuberculosis Control Institute, during the period 2012–2018, as part of routine TB reporting. Population estimates were obtained from the Hunan Bureau of Statistics. A wide range of socioeconomic, environmental, and healthcare access data were also obtained from Hunan Bureau of Statistics website to be used as possible covariates. The TB data obtained from the Hunan Tuberculosis Control Institute were aggregated at the county level and were linked with the possible covariates using ArcGIS Geographical Information System software (ESRI, Redlands, CA).

Data analysis
The overall crude incidence rate of notified DR-TB was calculated by taking the total number of DR-TB cases reported during the study period as the numerator and the mid-point total population during the same time period as the denominator. The standardised morbidity ratio (SMR) was then calculated for each county using the formula: $Y_i = \frac{O_i}{E_i}$; where $Y_i$ is the SMR in the county, $O$ is the observed number of DR-TB cases in the county and $E$ is the expected number of DR-TB cases in the county across the study period. The expected number of DR-TB cases for each county was calculated by multiplying the mid-point population of each county by the overall crude DR-TB incidence rate for the study area and period.

Spatial analysis
Spatial autocorrelation was explored using Moran’s I statistic and the Getis-Ord statistic. The Moran’s I statistic was used to assess the presence, strength and
direction of spatial autocorrelation over the whole study area and to test the assumption of spatial independence in the implementation of the spatial pattern analysis. The spatial clustering analysis using Getis-Ord Gi* statistic identified hotspot counties where DR-TB were clustered—targeted interventions might be important to reduce the burden of the disease in these hotspot counties. These analyses were conducted using tools provided in ArcGIS.12

In conducting a multivariable Bayesian spatial analysis, we assumed that the number of people with DR-TB at the county level followed a Poisson distribution. An initial bivariate analysis was conducted and variables with \( p < 0.2 \) were selected for the multivariable Bayesian spatial model. We constructed three different models using WinBUGS V.1.4.3 software (Medical Research Council Biostatistics Unit, Cambridge, UK). All variables selected from the bivariate analysis were incorporated as fixed effects in all three models. In the first model (model I), unstructured random effects were added. The assumptions were that: (1) the number of DR-TB cases in a county was independent of the number of DR-TB cases in other counties in the province, after accounting for the covariates and (2) the variance of random effects was homogeneous across the study area. To handle possible spatial dependency of cases and the violation of homogeneity of variance within each county due to the spatial nature of the data, a second model (model II) containing spatially structured random effects was constructed. These random effects used a Bayesian smoothing conditional autoregressive (CAR) structure.13 Finally, a third model (model III) containing the covariates and both the unstructured and spatially structured random effects, was constructed as follows:

\[
Y_i \sim \text{Poisson}(\mu_i)
\]
where \( Y_i \), the observed number of DR-TB cases in county \( i \), was assumed to follow a Poisson distribution with mean \( \mu_i \); and the log of the mean was modelled as:

\[
\log(\mu_i) = \log(E_i + \alpha + \sum_{k=1}^{K} \beta_k \times X_k + U_i + V_i)
\]

where \( E_i \) is the expected number of DR-TB cases in county \( i \), \( \alpha \) is the intercept, \( \beta_k \) is the coefficient for covariate \( X_k \) (where \( K \) is the number of covariates), \( U_i \) are the unstructured random effects and \( V_i \) are the spatially structured random effects.

The spatially structured random effects (\( V_i \)) were computed using a CAR structure by including an adjacency matrix to determine the spatial relationships between each pair of counties. The adjacency matrix for each county was generated using ArcGIS. A weight of 1 was given if two counties were neighbouring and a weight of 0 was given if two counties were not neighbouring. Two counties were considered to be neighbouring if they shared the same edges or corners (i.e., queen contiguity). Prior probability distributions for the coefficients (\( \beta \)), \( U_i \) and \( V_i \) were assumed to have normal distributions with a mean=0 and a precision (i.e., inverse of variance)=1 x 10^{-6}. For the intercept (\( \alpha \)) a flat prior distribution was used (i.e., a non-informative, improper prior with bounds \(-\infty \) and \( +\infty \)). The prior for the precision of the unstructured and spatially structured random effects was assigned a non-informative gamma distribution with shape and scale parameters=0.001. The posterior distribution of the parameters was estimated from the prior and data likelihood information using a Bayesian Markov Chain Monte Carlo simulation approach with Gibbs sampling employed by WinBUGS. The models were run for 100,000 iterations and convergence was successfully achieved after 80,000 iterations for each of the models. Convergence of the model was checked by visual inspection of posterior kernel densities and history plots, and the posterior distribution of each parameter was stored for summary measures such as the posterior mean, SD and the 95% credible interval (CrI). The deviance information criterion (DIC) was also stored for model selection, where a lower DIC value was used to indicate the best-fitting, most parsimonious model. Choropleth maps of the raw data and model random effects were created using the ArcGIS software to show the spatial clustering of DR-TB in Hunan province before and after accounting for the covariates and to indicate hotspot areas for targeted interventions. Since the covariates have different units and scales of measurement, we standardised all the variables to a z-scale on the basis of their mean and SD but converted the relative risks back to the original scale for presentation of the results.

**Patient and public involvement**

No patients or public were involved in this research.

**RESULTS**

**Demographic characteristics of DR-TB**

A total of 323,340 TB cases were reported to Hunan TB Control Institute from 104 counties between 2013 and 2018. Of these, 2,676 cases were DR-TB and included in the analysis. Twenty-seven DR-TB cases were from other provinces and were therefore excluded from the study.
Of the 2649 included DR-TB patients, 1983 (74.8%) were male and 1883 (71.0%) were farmers by occupation. The mean age of patients was 46.5 years (SD ±15.16, range 10–90 years). The largest number of cases was reported between September and December in every year, which accounted for 46.4% (n=1229) of the cases over the study period. Table 1 summarises the demographic and clinical characteristics of patients with DR-TB in Hunan province.

Clinical characteristics of DR-TB
The majority of DR-TB patients (88.5%, n=2345) had a history of first-line or second-line TB treatment and 11.4% (n=304) were new cases of DR-TB. The number of new DR-TB cases increased from 15 cases in 2012 to 114 cases in 2018, and over two-thirds of all the cases (37.5%) were reported in 2018. The proportion of XDR-TB among all DR-TB was 3.3% (n=89) (95% CI 2.7% to 4.1%) overall, and increased from 2.8% in 2012 to 4.4% in 2018. The mean age of patients with XDR-TB was higher at 49.5 years (SD ±15.4 years) than the mean age of patients with MDR-TB at 46.2 years (SD ±15.1 years) (p=0.05). Half (50.0%; n=1337) of the patients were taking treatment at the time of this study; treatment outcomes were not registered for 25 patients (0.95%). Of 1287 DR-TB patients with registered treatment outcomes, 434 (33.8%) were cured, 198 (15.3%) completed treatment, 92 (7.1%) died, 108 (8.3%) had treatment failure and 455 (35.3%) were lost to follow-up. Half (50.9% n=655) of the patients had poor treatment outcomes.

Incidence rate of notified DR-TB
The 6-year incidence rate of notified DR-TB in Hunan province was 3.9 per 100 000 population. The overall drug-resistant testing coverage in Hunan Province has increased from 0.4% in 2013 to 1.3% in 2018 (online supplemental figure S1). Similarly, the annual incidence rate of notified DR-TB increased significantly over time from 0.25 per 100 000 population in 2012 to 0.83 per 100 000 population in 2018 (p=0.002) (figure 1). The highest SMR) and incidence of notified DR-TB over the 6-year period were reported in Huarong county (12.6 per 100 000 population) and Hengshan county (11.0 per 100 000 population) (figure 2). The incidence rate of notified DR-TB at prefecture level is presented in online supplemental figure S2.

Spatial clustering of DR-TB
Substantial spatial heterogeneity in DR-TB was observed, based on our clustering analyses. Using the Getis–Ord Gi* statistic, 12 counties were identified as hotspots for DR-TB in Hunan province and 8 counties were cold spots (figure 3). The hotspot counties were in the north and east parts of Hunan province, while the cold spot counties were located in the south and west of the province. Anselin Local Moran’s I statistic showed high-high and low-low clustering of DR-TB at county-level in Hunan province (online supplemental figure S3). Important changes in the spatial clustering of notified DR-TB was observed overtime in Hunan province (online supplemental figure S4).

Ecological-level factors associated with spatial clustering of DR-TB
Based on the lower DIC value identified from the multivariate Bayesian regression models, the model which contained the spatially structured random effect only (model II) was the best fit. The best-fitting model indicated that the incidence rate of notified DR-TB was significantly associated with urban communities (ie, percentage of urban residence in the counties (%)) (RR: 1.02; 95%CrI: 1.01 to 1.20). Several ecological level factors such as sex, residence, gross domestic product and the mortality rate as well as healthcare access variables such as the number of health institutions (ie, hospitals and clinics) and number of medical personnel, and climatic variables such as monthly average temperature, annual precipitation and annual mean sunshine hours in a county were not significantly associated with the incidence rate of notified DR-TB in Hunan province (table 2). After accounting for the ecological-level factors in the model, the posterior mean of spatially structured random effects was found to be clustered in the province (figure 4). This indicates

Figure 3 Spatial clustering of drug-resistant tuberculosis (DR-TB) incidence in Hunan province, 2012–2018, based on the Getis-Ord Gi* statistic.
that a substantial amount of county level heterogeneity in DR-TB remained unexplained by the ecological level factors included in our models. Geographical distribution of healthcare access variables such as number of health institutions (online supplemental figure S5), number of beds (online supplemental figure S6) and number of medical personnel (online supplemental figure S7) are presented in online supplemental file 1.

**DISCUSSION**

Our study detected spatial clustering in the incidence of notified DR-TB in Hunan Province, China. Counties located in the north and west part of Hunan province had high incidence of notified DR-TB. Previous studies conducted in South Africa, Moldova and Ethiopia as well as studies conducted in other parts of China have also showed spatial clustering of DR-TB at the subnational level. The substantial heterogeneity in DR-TB indicates an imperative for more regular surveillance to detect emerging and existing hotspots for targeted interventions to prevent DR-TB, especially in settings where there are limited resources. Understanding the spatial distribution of DR-TB and identifying areas with the highest incidence of notified DR-TB is a first step towards reducing the burden of DR-TB.

Our study also showed that the annual incidence of notified DR-TB in Hunan province steadily increased over time, from 0.25 per 100 000 population in 2012 to 0.83 per 100 000 population in 2018. There could be two reasons for this finding. The first reason could be due to a real increase in numbers of patients with DR-TB (ie, primary DR-TB) or due to the development of drug-resistant strains during a course of first-line TB treatment (ie, acquired or secondary DR-TB). Since the number of patients who were lost to follow-up in Hunan province was high at 35% in this study and 27% in a previous study, the risk of primary transmission of DR-TB in the community as well as the development of acquired DR-TB is high. Whole genome sequencing would be important to investigate this and to identify the main sources of DR-TB transmission in the province. The second reason could be due to improved diagnostic services and the introduction of rapid molecular diagnostic methods such as line probe assays and Xpert MTB/RIF. The impact of new molecular diagnostic test

| Table 2 | A multivariate Bayesian poisson regression model with spatially structured and spatially unstructured random effects for DR-TB in Hunan province, China, 2012–2018 |
|---|---|---|---|
| Variables | Spatially unstructured model | Spatially structured model | Both spatially structured and unstructured model |
| | (RR (95%CrI)) | (RR (95%CrI)) | (RR (95%CrI)) |
| **Socioeconomic and demographic factors** | | | |
| Proportion of male in a county | 0.96 (0.87 to 1.07) | 0.93 (0.82 to 1.07) | 0.96 (0.86 to 1.08) |
| Percentage of urban residence in the counties | 1.09 (1.08 to 1.30) | 1.02 (1.01 to 1.20) | 1.04 (1.02 to 1.25) |
| GDP of the county (10 000 yuan per person per year)* | 1.01 (0.87 to 1.17) | 1.00 (0.87 to 1.15) | 1.01 (0.88 to 1.17) |
| Annual birth rate in the county (%) | 0.97 (0.85 to 1.12) | 1.00 (0.85 to 1.17) | 0.98 (0.85 to 1.15) |
| Annual death rate in the county (%) | 1.01 (0.84 to 1.20) | 1.04 (0.88 to 1.23) | 1.02 (0.85 to 1.22) |
| **Healthcare access** | | | |
| Contraceptive use rate of the county (%) | 0.93 (0.82 to 1.05) | 0.97 (0.84 to 1.11) | 0.94 (0.83 to 1.08) |
| No of institutions per 10 000 population in a county | 1.01 (0.86 to 1.19) | 0.94 (0.80 to 1.09) | 0.98 (0.83 to 1.16) |
| No of medical personnel per 10 000 population in county | 0.89 (0.75 to 1.05) | 0.94 (0.80 to 1.09) | 0.91 (0.77 to 1.08) |
| **Climatic factors** | | | |
| Monthly average temperature (°C) | 1.13 (1.00 to 1.28) | 1.00 (0.85 to 1.18) | 1.08 (0.92 to 1.24) |
| Annual total precipitation (mm) | 0.99 (0.85 to 1.14) | 1.08 (0.88 to 1.33) | 1.01 (0.85 to 1.23) |
| Monthly sunshine hours (kJ/m²/day) | 1.15 (1.00 to 1.32) | 1.02 (0.80 to 1.26) | 1.11 (0.90 to 1.32) |
| **Heterogeneity** | | | |
| Variance of spatially unstructured random effect (σ²) | 4.56 (3.05 to 6.73) | | 6.21 (3.47 to 72.77) |
| Variance of spatially structured random effect (σ²) | | 1.10 (0.73 to 1.64) | 6.39 (1.11 to 729.60) |
| Intercept (alpha) | −0.14 (−0.25 to −0.03) | −0.12 (−0.19 to −0.06) | −0.13 (−0.23 to −0.04) |
| DIC | 689.9 | 498.1 | 603.2 |

*GDP (¥10 000 per person per year; ¥1=US$0.15).

CrI, credible interval; DIC, deviance information criterion; GDP, gross domestic product.
Using routinely collected surveillance data from the provincial TB Programme, we produced maps showing the counties where DR-TB were clustered. These may help Provincial authorities to design strategies and target interventions for the prevention and control of TB and DR-TB. However, the study used secondary data available from the Hunan TB Control Institute, which may not reveal the true burden of the disease in the community because in some counties with high DR-TB, only those cases that are at highest risk of having DR-TB may receive a DST. Future studies should incorporate national TB and DR-TB survey, as well as whole genome sequencing data to better identify geographical areas at increased risk of DR-TB.

**Conclusion**

In this study, we found that the annual incidence rate of notified DR-TB increased over time and spatial clustering of DR-TB was detected in the North and East parts of Hunan province. This finding suggested that targeting interventions to highest risk areas and population groups would help to reduce the burden and to prevent the ongoing transmission of DR-TB.

**Author affiliations**

1Faculty of Health Sciences, Curtin University, Bentley, Western Australia, Australia
2Telethon Kids Institute, Nedlands, Western Australia, Australia
3Department of Tuberculosis Control, Tuberculosis Control Institute of Hunan Province, Changsha, Hunan, China
4Department of Director's Office, Hunan Tuberculosis Control Institute, Changsha, Hunan, China
5Department of MDR-TB, Internal Medicine, Hunan Chest Hospital, Changsha, China
6Research School of Population Health, Australian National University, Canberra, ACT, Australia
7Public Health Sciences, Karolinska Institute, Stockholm, Sweden
8School of Public Health, The University of Sydney, Sydney, New South Wales, Australia

**Acknowledgements**

The authors would like to thank those who were participating in the data translation from Mandarin to English.

**Contributors**

KAA, KV, DG, ACAC and ZX conceived and designed the study. KAA, HY, LB, YT and ZX supervised the data translation process. KAA and ZZ prepared the draft manuscript and KV improved the drafted manuscript. KV and ACAC advised the overall data analysis and the development of the manuscript. All the authors involved with critical revision of the manuscript and approved the final version.

**Funding**

This study was funded by the China in the World Grant (CW) 2019 support. Grant number is not applicable.

**Disclaimer**

The contents are the responsibility of the authors and do not necessarily reflect the views of the funder. The funder had no role in study design, data collection and analysis, interpretation of data, decision to publish, or in writing the manuscript.

**Competing interests**

None declared.

**Patient consent for publication**

Not required.

**Ethics approval**

Ethical clearance was obtained from the Australian National University (HRE11087) and Curtin University Human Research Ethics Committees (HRE2019-0581). Additional approval for the study and permission to access the DR-TB data was obtained from the Hunan Tuberculosis Control Institute. Since we used data aggregated at county level. All methods were performed in accordance with the approved protocol by the ethics committee.

**Provenance and peer review**

Not commissioned; externally peer reviewed.


ArcGIS Desktop: Release 10. 6.1 [program], USA 2015.


Supplementary information

Figure S1: Prevalence of drug-resistant testing coverage in Hunan Province
**Figure S2:** The incidence rate of drug-resistant tuberculosis in Hunan Province at prefecture-level, between 2012 and June 2018.

**Figure S3:** Spatial clustering of drug-resistant tuberculosis at county-level in Hunan Province, between 2012 and June 2018, based on Local indicators of spatial association using Anselin Local Moran’s I statistic.
Figure S4. Spatial clustering of drug-resistant tuberculosis in Hunan Province from 2012 to 2018, based on the Getis-Ord Gi* statistic.
Figure S5: Number of health facilities per 10 000 population in Hunan Province at country level
Figure S6: Number of beds per 10 000 population in Hunan Province at country level

Figure S7: Number medical personnel per 10 000 population in Hunan Province at country level