Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/mbe

MBE, 20(6): 10678–10693.
DOI: 10.3934/mbe.2023473
Received: 05 December 2022
Revised: 18 March 2023
Accepted: 24 March 2023
Published: 14 April 2023

Research article

Spatio-temporal distribution characteristics of the risk of viral hepatitis B incidence based on INLA in 14 prefectures of Xinjiang from 2004 to 2019

Yijia Wang[1,†], Na Xie[2,†], Zhe Wang[3], Shuzhen Ding[1], Xijian Hu[1,*] and Kai Wang[4,*]

[1] College of Mathematics and System Science, Xinjiang University, Urumqi 830017, China
[2] Xinjiang Center for Disease Control and Prevention, Urumqi 830054, China
[3] Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, China
[4] College of Medical Engineering and Technology, Xinjiang Medical University, Urumqi 830017, China

[†] These authors contributed equally.

[*] **Correspondence:** Email: wangkaimath@sina.com, xijianhu@xju.edu.cn.

Abstract: This study aimed to explore the spatio-temporal distribution characteristics and risk factors of hepatitis B (HB) in 14 prefectures of Xinjiang, China, and to provide a relevant reference basis for the prevention and treatment of HB. Based on HB incidence data and risk factor indicators in 14 prefectures in Xinjiang from 2004 to 2019, we explored the distribution characteristics of the risk of HB incidence using global trend analysis and spatial autocorrelation analysis and established a Bayesian spatiotemporal model to identify the risk factors of HB and their spatio-temporal distribution to fit and extrapolate the Bayesian spatiotemporal model using the Integrated Nested Laplace Approximation (INLA) method. There was spatial autocorrelation in the risk of HB and an overall increasing trend from west to east and north to south. The natural growth rate, per capita GDP, number of students, and number of hospital beds per 10,000 people were all significantly associated with the risk of HB incidence. From 2004 to 2019, the risk of HB increased annually in 14 prefectures in Xinjiang, with Changji Hui Autonomous Prefecture, Urumqi City, Karamay City, and Bayangol Mongol Autonomous Prefecture having the highest rates.

Keywords: hepatitis B; Bayesian spatiotemporal model; INLA method; risk factor

1. Introduction

In this study, we aimed to (i) identify the spatial and temporal patterns of HB disease from 2004 to 2019 in Xinjiang, China, (ii) determine the risk factors for HB disease, and (iii) determine how the risk

of each disease varies spatially and temporally.

HB is a widely spread infectious illness that poses a major threat to human health [1, 2]. According to a 2006 survey, hepatitis B virus (HBV) infection in China accounts for almost 1/3 of all HBV-infected individuals globally, and the HB pandemic is quite serious. The northwestern region of China has a high prevalence of HB, and its incidence is increasing [3]. Xinjiang, located in northwest China, is one of the provinces with a high prevalence of HB. In 2019, Xinjiang ranked first in the incidence of officially reported cases (176.325/100,000) and third in the death rate of viral hepatitis (0.0402/100,000) [4].

To provide a reference for effective prevention and control strategies for HB epidemics, Gao et al. [5] used descriptive epidemiological methods to analyze the prevalence of HB epidemics and discussed the distribution of patient populations and morbidity. Wang et al. [6] conducted a systematic review and meta-analysis to better understand the epidemiology of HB in the general Chinese population. Wang et al. [7] studied the nationwide spatiotemporal distribution of HB and found that Xinjiang is a high-value cluster area. Zhu et al. [8] explored the temporal trends in viral hepatitis and examined the spatial clustering of viral hepatitis in China. Many studies have demonstrated that HB endemicity is associated with socioeconomic and demographic factors, access to healthcare, etc. Sagnelli et al. [9] indicated that improved socioeconomic conditions, improved hygiene standards and reduced family size were the main reasons for the reduced incidence of HB in Italy. Ochola et al. [10] studied the relationship between HBV infection and poverty in sub-Saharan Africa and Uganda and showed a negative correlation between the wealth index and HBV infection. Kinfe et al. [11] found that educational status is a risk factor for HB infection. Zhang et al. [12] demonstrated that adequate hospital beds are essential for disease control and eradication. Some scholars have focused on the epidemiology of HB within certain regions in terms of age-specific, sex-specific, large-scale spatio-temporal distribution characteristics nationwide, or risk factors for HB infection; however, few have analyzed it from a spatio-temporal perspective and integrated the impact of risk factors on the HB epidemic in Xinjiang. Therefore, in this study, global trend analysis and spatial autocorrelation methods were used to analyze the prevalence of HB epidemics and to discuss the distribution of morbidity risk. A spatio-temporal model was developed to investigate the risk factors and spatio-temporal heterogeneity of HB incidence in Xinjiang by selecting four risk factors: number of beds (per 10,000 people), per capita gross domestic product (per capita GDP), number of school students and natural growth rate. For spatiotemporal models, non-Bayesian modeling methods include Geographically and Temporally Weighted Regression (GTWR) [13] and spatiotemporal panel data models [14]. The Bayesian spatiotemporal model can consider the prior information and uncertainty of the model parameters compared to the non-Bayesian model, and the model can study the spatiotemporal patterns of disease variation and relate them to relevant pathogenic factors. For such models, Markov Chain Monte Carlo (MCMC) methods are typically used to estimate parameters. Compared to the MCMC method, the integrated nested Laplace approximation (INLA) method has more powerful computing capabilities without losing accuracy [15]. Therefore, the INLA method was used to estimate the parameters of Bayesian spatiotemporal models. Currently, the method has been embedded in a variety of models, such as generalized linear mixed models [16] and binary logit mixed models [17], which are widely used in various fields, such as soil organic matter changes. Sun et al. [18] compared the nested Laplace approximate stochastic partial differential equation (INLA-SPDE) model with a Bayesian maximum entropy approach to predict detailed spatio-temporal variations in sampling periods, and the predic-

tions derived from INLA-SPDE were found to be more accurate. Using the INLA-SPDE method, Marie et al. [19] found that juveniles and adults of Lutjanus synagris have patchy distributions along the Rio Grande do Norte coast and are mainly influenced by environmental predictors such as water depth and sea surface temperature. Lezama-Ochoa et al. [20] used the INLA-SPDE to predict the occurrence of Mobula mobular in the Eastern Pacific Ocean (EPO). Iulian et al. [21] predicted cancer growth in space and time using a Bayesian spatiotemporal model. Froelicher et al. [22] used the spatial age-period-cohort method to study breast cancer incidence in the Portuguese region. Ben et al. [23] used R-INLA to find that weather had different effects on six honeybee diseases and that the risk of disease varied spatially. Bie et al. [24] used a spatio-temporal model and the INLA algorithm to investigate seven influences on the relative risk of tuberculosis in China and the impact of spatio-temporal distribution. The INLA method is a fast computational method for Bayesian inference proposed by Rue in 2009 [15]. The MCMC method tends to exhibit poor performance because the components of the latent field are strongly dependent on each other, and the vectors of the hyperparameters and latent field are also strongly dependent, especially when the dimension n is large [25]. The INLA method solves both problems in three steps. The first step approximates the posterior margin of the vector of hyperparameters using the Laplace approximation. The second step computes the Laplace approximation or simplified Laplace approximation for selected values of the vector of hyperparameters to improve the Gaussian approximation. The third step combines the two previous methods using numerical integration.

This article is organized as follows: Section 1 contains the Introduction. Section 2 describes the materials and methods used in the study. Section 3 presents the results from the risk factors of HB, spatial and temporal. Section 4: Discussion of the results. Finally, Section 5 concludes the paper.

2. Materials and methods

2.1. Data sources

The number of HB monthly incidence cases in 14 prefectures in Xinjiang from 2004/1 to 2019/12 (Figure 1) was obtained from the infectious disease information management system of the Xinjiang CDC. Incidence data were counted by year for use in subsequent studies.

To study the factors influencing the risk of HB, socio-demographic, socio-economic and healthcare data from 14 prefectures in Xinjiang were gathered and organized from 2004 to 2019. socio-demographic information included the natural growth rate (birth rate-death rate), the total number of students in higher education and general secondary vocational schools (persons, abbreviated as the number of students) and socio-economic information, including per capita gross domestic product (per capita GDP) (yuan). The healthcare information included the number of hospital beds per 10,000 people (abbreviated as the number of beds). The four indicators are the factors influencing the risk of HB incidence in this study and are denoted by $X_k(k = 1, ..., 4)$.

2.2. Methods

2.2.1. Statistical spatial analysis

Global trends and spatial autocorrelation analyses were used to explore the distribution characteristics of HB risk. Trend analysis is the transformation of points in the study area into a three-dimensional

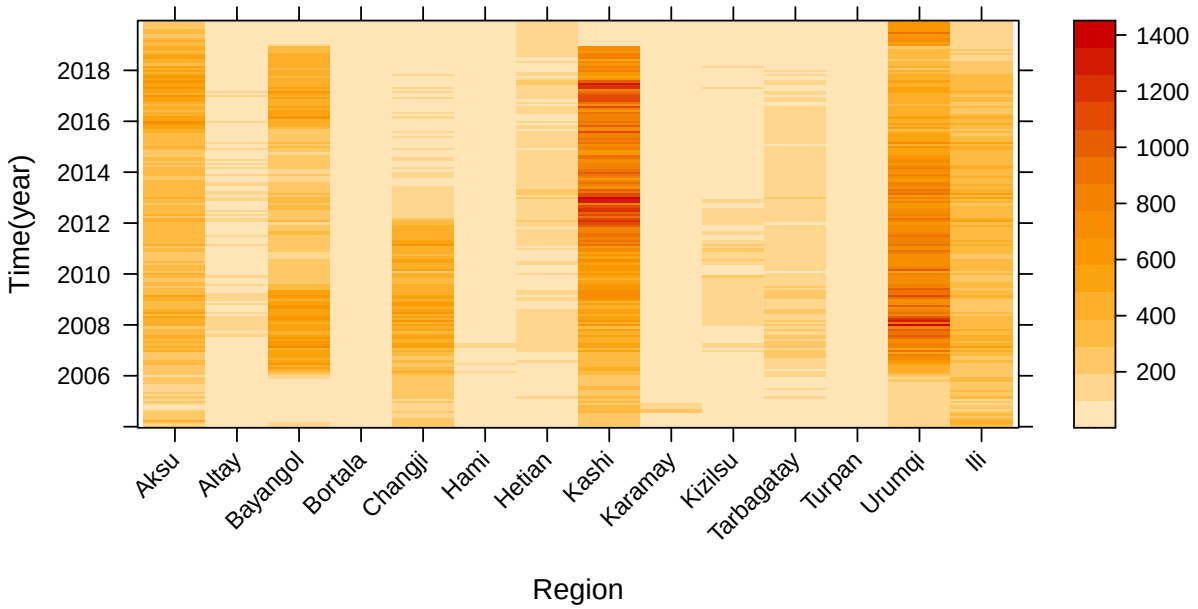

Figure 1. Distribution of the number of new HB cases per month in 14 prefectures from 2004 to 2019.

map with specific attribute values as heights and the analysis of global trends in the dataset from different perspectives.

Spatial autocorrelation reflects the spatial distribution of an attribute value within an entire region [26]. Global Moran's I was applied to the Geoda environment to examine the spatial distribution of HB prevalence in all prefectures within the study area. A global Moran's I index between -1 and $+1$ indicates whether the spatial distribution of a disease is clustered (> 0), dispersed (< 0), or random ($= 0$).

2.2.2. Bayesian spatiotemporal model

Bayesian spatiotemporal models allow for a more accurate estimation of disease risk, analysis of spatiotemporal evolution patterns and relationships with known explanatory variables, and are now widely used in disease spatial data modeling and disease mapping [27, 28].

The model assumes that Y_{it} denotes the number of HB patients in region i ($i = 1, ..., 14$) in year t ($t = 1, ..., 16$). In general, the disease incidence data obey a Poisson distribution, i.e., $Y_{it} \sim \text{Poisson}(\lambda_{it})$, then we have $E(Y_{it}) = \lambda_{it} = E_{it}\theta_{it}$, E_{it} denotes the expected number of cases in region i in year t, θ_{it} denotes the relative risk of HB incidence, denotes the ratio of the actual number of incidences to the expected number of incidences in region i in year t. Assuming that k factors influence the development of HB, a Bayesian spatiotemporal model is constructed using logit link function, which is expressed as follows:

$$\log(\theta_{it}) = b_0 + \sum_k \beta_k X_{kit} + u_i + v_i + \gamma_t + \varphi_t. \tag{1}$$

Equation (1) is denoted as Model 1, where $\log(\theta_{it})$ is a linear predictor, b_0 denotes the average log

relative risk, and X_{kit} is the kth ($k = 1, ..., 4$) covariates, β_k is the coefficient of the kth covariate. u_i obeys the proper conditional autoregressive (CAR) [29], which indicates the spatial structure effect and reflects the existence of spatial autocorrelation among neighboring spatial units of the 14 prefectures in Xinjiang. v_i follows a normal distribution, which indicates spatial unstructured effects and can reflect other random effects caused by non-spatial factors. u_i and v_i can be considered as hidden variables for region i, which are related and unrelated to the location of the region, respectively. γ_t follows the second-order wandering model [24], indicating that the temporal structure effect, representing undetermined features in the tth year, has a temporal structure. φ_t obeys the normal distribution, indicating a temporal unstructured effect, indicating that the undetermined features in the tth year do not have a temporal structure. γ_t and φ_t can be regarded as hidden variables in year t that are related and unrelated to the location of year t respectively.

According to the description above, the distributions of u_i, v_i, γ_t and φ_t take the following forms:

$$u_i \mid \boldsymbol{u}_{-i} \sim N\left(\frac{1}{M_i}\sum_{j=1}^{n} a_{ij}u_j, \frac{1}{M_i\tau_u}\right),$$
$$\gamma_t \mid \gamma_{t-1}, \gamma_{t-2} \sim N\left(2\gamma_t - 1 + \gamma_t - 2, \sigma^2\right),$$
$$v_i \sim N\left(0, \frac{1}{\tau_v}\right), \varphi_t \sim N\left(0, \frac{1}{\tau_\varphi}\right),$$

where $\boldsymbol{u}_{-i}$ denotes the remaining regions except for i, $M_i = \#N(i)$ is the number of neighbors of region i, $N(i)$ is the region adjacent to region i, τ_u is the precision of the spatially structured effect, τ_v is the precision of the spatially unstructured effect and τ_φ is the precision of the temporally unstructured effect [30].

The hidden Gaussian field contains all the hidden variables including the linear combination of the quantity to be estimated and the independent variables, so model 1's hidden Gaussian field $\boldsymbol{Z} = \{\boldsymbol{\theta}, \boldsymbol{b}, \boldsymbol{\beta}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{\gamma}, \boldsymbol{\varphi}\}$ and the model hyperparametric vector $\boldsymbol{\psi} = \{\boldsymbol{\psi}_u\}$, which is distributed a priori as $\log \boldsymbol{\psi} \sim \text{logGamma}(1, 0.001)$. It was proposed by Knorr-Held L [31] and avoids the problem of sensitive parameter estimates caused by poorly chosen hyperparameter prior distributions. The R package, INLA, was used to set the prior distribution and estimate the model.

In contrast to Model 1, Model 2 is set up without temporal and spatial effects, that is, a generalized additive regression model [32], which is expressed as follows:

$$\log(\theta_{it}) = b_0 + \sum_k \beta_k X_{kit}. \tag{2}$$

The optimal model was selected for the HB risk impact factor analysis based on the deviance information criterion (DIC) [31], which has a smaller value and better model fit.

The INLA applies a hidden Gaussian model that requires the hidden field $\boldsymbol{Z}$ to be a Gaussian Markov random field. The hidden Gaussian model is essentially a Bayesian hierarchical model that can be divided into the following structures:

The first level is the likelihood structure; the variable $\boldsymbol{y}$ follows an exponential family distribution, ψ_1 is the hyperparameter in the likelihood function and is usually the precision parameter of the distribution that $\boldsymbol{y}$ obeys

$$\mathbf{y} \mid \boldsymbol{Z}, \psi_1 \sim \prod \pi(\mathbf{y} \mid \boldsymbol{Z}, \psi_1),$$

The second level of the structure is the hidden variable distribution, which must obey a normal distribution and satisfy the conditions of a Gaussian Markov random field; ψ_2 is the hyperparameter in the $\boldsymbol{Z}$ prior distribution, which is generally the precision parameter in the prior distribution:

$$\boldsymbol{Z} \mid \psi_2 \sim N\left(\mu(\psi_2), \boldsymbol{Q}^{-1}(\psi_2)\right),$$

The third level is the prior distribution of hyperparameters $\boldsymbol{\psi}$. The hyperparameter vector is assumed to have a distribution of $\boldsymbol{\psi} \sim \pi(\boldsymbol{\psi}), \boldsymbol{\psi} = (\psi_1, \psi_2)$.

Based on the Bayesian statistical inference [33], the joint posterior distribution of the parameters and hyperparameters in the model can be obtained as follows:

$$\pi(\boldsymbol{Z}, \boldsymbol{\psi} \mid \boldsymbol{y}) \propto \pi(\boldsymbol{\psi})\pi(\boldsymbol{Z} \mid \boldsymbol{\psi}) \prod \pi\left(y_i \mid Z_i, \boldsymbol{\psi}\right),$$

The objective of the INLA is to compute the posterior marginal distributions $\pi\left(Z_i \mid \boldsymbol{y}\right)$ and $\pi\left(\psi_j \mid \boldsymbol{y}\right)$ of the parameters and hyperparameters, which can be obtained using Bayesian inference from, as follows:

$$\pi\left(Z_i \mid \boldsymbol{y}\right) = \int \pi(\boldsymbol{Z}, \boldsymbol{\psi} \mid \boldsymbol{y}) d\boldsymbol{\psi} = \int \pi\left(Z_i \mid \boldsymbol{\psi}, \boldsymbol{y}\right) \pi(\boldsymbol{\psi} \mid \boldsymbol{y}) d\boldsymbol{\psi}, \tag{3}$$

$$\pi\left(\psi_j \mid \boldsymbol{y}\right) = \int \pi(\boldsymbol{\psi} \mid \boldsymbol{y}) d\boldsymbol{\psi}_{-j}. \tag{4}$$

where $\boldsymbol{\psi}_{-j}$ denotes the remaining hyperparameters except for ψ_j. Using the INLA method, Equations (3) and (4) can be calculated quickly. The INLA method was previously described [15].

3. Results

3.1. Characteristics of the distribution of the risk of HB

Table 1 shows the risk of HB in 14 localities in Xinjiang, with specific prefectures listed in the table where the risk of HB was greater than 1 (i.e., the actual number of cases was greater than the expected number of cases). It can be seen that the high incidence of the epidemic is concentrated in Urumqi City, Kashgar Prefecture and Bayangol Mongol Autonomous Prefecture. From 2004 to 2019, the risk of HB increased from 1.07 to 1.45 in Aksu Administrative Prefecture, and from 0.86 to 1.65 in Kashgar Prefecture among all prefectures. The risk of HB in Hami City, Karamay City, Tacheng Prefecture, Turpan City and Ili Kazakh Autonomous Prefecture showed a decreasing trend. The risk of HB in Hotan Prefecture was low (approximately 0.4) in all years. No clear trend was observed in the other prefectures.

The average risk of developing HB from 2004 to 2019 was greater than 1 for Aksu Administrative Prefecture, Bayangol Mongol Autonomous Prefecture, Changji Hui Autonomous Prefecture, Kashgar Prefecture, Kizilsu Kirgiz Autonomous Prefecture and Urumqi City, and the average risk of developing HB in Bayangol Mongol Autonomous Prefecture and Urumqi City exceeded two.

The average risk of morbidity and spatial trend analyses across Xinjiang are shown in Figure 2. The x-axis points east, with the green curve indicating the east-west direction, and the y-axis points north, with the blue curve indicating the north-south direction. The z-axis is the onset risk value, which is given by the height of the rod. The risk of HB in Xinjiang tended to increase from west to east and from north to south, with a clear stepwise distribution.

Table 1. Prefectures with a risk of HB > 1 in Xinjiang from 2004 to 2019.

Years	Prefectures
2004	Aksu, Changji, Karamay, Kizilsu Kirgiz, Ili Kazakh
2005	Aksu, Changji, Karamay, Kizilsu Kirgiz, Urumqi, Ili Kazakh
2006–2007	Bayangol, Changji, Tacheng, Urumqi
2008–2010	Bayangol, Changji, Urumqi
2011	Bayangol, Changji, Kashgar, Urumqi
2012–2013	Bayangol, Kashgar, Urumqi
2014	Altay, Bayangol, Kashgar, Urumqi
2015–2019	Aksu, Bayangol, Kashgar, Urumqi

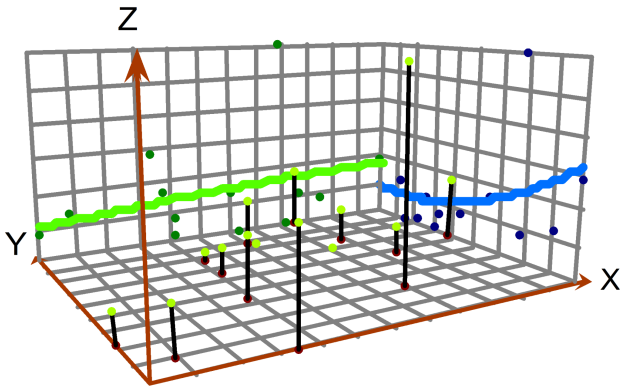

Figure 2. Analysis of spatial trends in 14 prefectures in Xinjiang from 2004 to 2019.

3.2. Spatial pattern of HB Risk

A global autocorrelation analysis was conducted on the rate of HB in Xinjiang from 2004 to 2019, in which there was a positive spatial correlation between 2006 and 2012, 2014–2015, and 2018, with a global Moran's I index ranging from 0.146 to 0.349 for each year (Table 2); the global Moran's index was greater than 0 and $p < 0.05$. Therefore, it is necessary to conduct studies using spatio-temporal models.

3.3. Model selection

Because DIC can integrate the model's goodness of fit and complexity, it is currently the most frequently used metric for comparison of model fit in Bayesian spatiotemporal models. Table 3 shows that the DIC value of Model 1 is reduced by 364,351.95, relative to Model 2, that is, Model 1 fits

Table 2. Global spatial autocorrelation analysis of HB incidence in Xinjiang from 2004 to 2019.

Years	Moran's I	p-Value	Years	Moran's I	p-Value
2004	-0.025	0.211	2012	0.206	0.048
2005	0.132	0.065	2013	0.117	0.069
2006	0.169	0.040	2014	0.164	0.049
2007	0.226	0.042	2015	0.218	0.043
2008	0.311	0.019	2016	0.094	0.084
2009	0.349	0.011	2017	0.111	0.086
2010	0.223	0.035	2018	0.146	0.043
2011	0.276	0.019	2019	0.131	0.079

Note: $Moran's\ I > 0$ and $p < 0.05$ represent the clustered pattern of HB.

better. Therefore, a Bayesian spatiotemporal model that includes spatial and temporal effects better fits the characteristics of the HB data from 2004 to 2019 in Xinjiang. Model 1 is investigated in the following sections.

Table 3. Evaluation index of risk model of HB incidence in 14 prefectures of Xinjiang.

Model	Expression	DIC
model 1	$\log(\theta_{it}) = b_0 + \sum_k \beta_k X_{kit} + u_i + v_i + \gamma_t + \varphi_t$	30,791.05
model 2	$\log(\theta_{it}) = b_0 + \sum_k \beta_k X_{kit}$	395,143

3.4. Influencing factors

Table 4 shows the posterior estimates of the regression coefficients of the influencing factors and RR. The relative risk of each factor is $RR = \exp(\beta)$ (the logarithm of θ_{it} is taken in Model 1, so the logarithm needs to be converted to an exponent when calculating relative risk, as below), where β is the regression coefficient of the relevant component. The results indicated that the risk of HB incidence was positively related to the natural growth rate and the number of beds and negatively correlated with the number of students and per capita GDP. The posterior means of the regression coefficients for the natural growth rate and the number of beds were 0.041 and 0.192, respectively, with corresponding relative risks of 1.04 (95% CI:1.04–1.04), and 1.21 (95% CI:1.20–1.22). Each unit increase in the variable was associated with a relative risk increase of 0.04 (95% CI:0.04–0.04), and 0.21 (95% CI:0.20–0.22). The posterior means of the regression coefficients for the number of students and per capita GDP were -0.042 and -0.356, respectively, with corresponding relative risks of 0.96 (95% CI:0.95–0.96), and 0.70 (95% CI:0.69–0.71). Each unit increase in the variable was associated with a relative risk decrease of 0.04 (95% CI:0.04–0.05), and 0.30 (95% CI:0.29–0.31).

3.5. Spatial distribution

The relative spatial risk is $RR_{spatial} = \exp(u + v)$, and the relative risk of the sum of spatial effects $(u + v)$ is shown in Table 5: In this study, the risk of HB incidence was divided into four levels: low, medium, medium-high and high risk. The low-risk areas were Hotan Prefecture, Turpan City,

Table 4. Posterior estimates of the regression coefficients of each variable and their RR.

variable	Mean (confidence interval)	Variance	RR (confidence interval)
natural growth rate	0.041 (0.038, 0.044)	0.001	1.04 (1.04, 1.04)
the number of students	-0.042 (-0.047, -0.036)	0.003	0.96 (0.95, 0.96)
per capita GDP	-0.356 (-0.365, -0.346)	0.005	0.70 (0.69, 0.71)
The number of beds	0.192 (0.183, 0.201)	0.005	1.21 (1.20, 1.22)

Note: $RR(relative\ risk) = exp(\beta)$, β is the regression coefficient.

Hami City and Bortala Mongol Autonomous Prefecture, whereas the high-risk areas were Changji Autonomous Prefecture, Urumqi City, Karamay City and Bayangol Mongol Autonomous Prefecture. To prevent the epidemic from becoming more serious, relevant disease control departments should develop different epidemic prevention and control measures based on the incidence risk levels, pay close attention to the development dynamics of the HB epidemic in high-risk areas, and perform better monitoring of the epidemic in the three medium-high-risk areas of Tacheng Prefecture, Kashgar Prefecture and Aksu Administrative Prefecture.

Table 5. Distribution of spatial relative risk $RR_{spatial}$.

the risk of HB incidence level	RR is located in the interval	Location prefectures			
low-risk	[0.32, 0.79]	Hotan	Turpan	Hami	Bortala
		0.32	0.64	0.72	0.79
medium-risk	(0.79, 0.98]	Ili Kazakh	Altay	Kizilsu Kirgiz	
		0.81	0.93	0.91	
medium-high-risk	(0.98, 1.35]	Tacheng	Kashgar	Aksu	
		1.03	1.15	1.09	
high-risk	(1.35, 2.45]	Changji	Urumqi	Karamay	Bayangol
		1.42	1.74	1.86	2.45

Note: $RR_{spatial} = exp(u + v)$, $(u + v)$ is the spatial effects.

3.6. Temporal distribution

The relative risk at time $RR_{temporal} = \exp(\gamma + \phi)$, which is the sum of the time impacts of $(\gamma + \phi)$ relative risk (Figure 3). The overall risk of HB in Xinjiang from 2004 to 2019 showed an increasing trend, reaching 1.12 in 2018. The unstructured effect curve fluctuates at 1, while the structured effect curve displays an upward trend. From 2004 to 2014, the risk of HB incidence ranged from 1, except in 2005 and 2013, with the lowest relative risk of 0.92 in 2009.

4. Discussion

The selection of indicators is very important and can affect the results of the study. For disease prevalence, which was initially determined by the number of cases, the method simply marks the number of new cases in each area on the map, and the indicator selected in this study is the risk of incidence (the ratio of the actual number of cases to the expected number of cases). The data were divided into intervals by the minimum, first quartile, median, third quartile and maximum values of

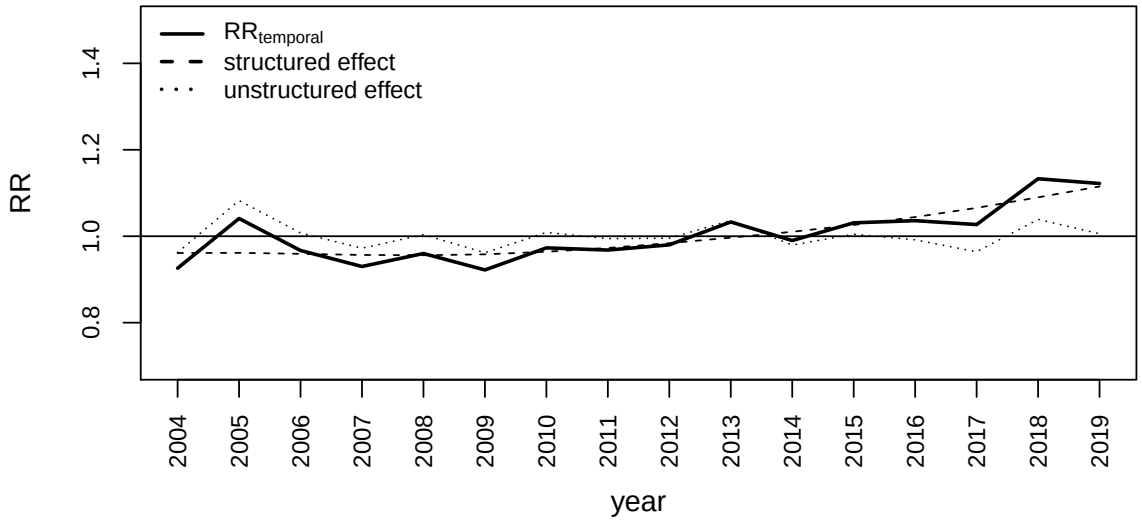

Figure 3. Distribution of $RR_{temporal}$ from 2004 to 2019.

the two indicators of incidence and risk of incidence, and the data were divided into four intervals. Using the number of new cases as the study indicator, Aksu Administrative Prefecture, Ili Kazakh Autonomous Prefecture, Kashgar Prefecture, and Urumqi City were the areas with the most severe HB epidemics. The high-risk areas were Changji Autonomous Prefecture, Urumqi City, Karamay City and Bayangol Mongol Autonomous Prefecture, with the risk of incidence used as the study index.

There were 67,199 new cases in Aksu Administrative Prefecture and 63,377 new cases in Bayangol Mongol Autonomous Prefecture in 16 years, and the incidence of HB was more serious in Aksu Administrative Prefecture. However, the average annual population of Aksu Administrative Prefecture is 2,401,336 and that of Bayangol Mongol Autonomous Prefecture is 1,278,448. Therefore, it is not appropriate to use the number of cases to determine the prevalence of HB in different populations in different regions. Therefore, this study relied on the risk of HB incidence, which was unaffected by socio-demographic characteristics of the region.

Therefore, this study used this indicator to conduct global trend and spatial autocorrelation analyses and established a Bayesian spatiotemporal model to explore the factors influencing the risk of HB disease in Xinjiang and its spatio-temporal characteristics. The results indicated that the risk of HB incidence was positively related to the natural growth rate and the number of beds, and negatively correlated with the number of students and per capita GDP.

HB is mainly transmitted from mother to child, through sexual contact and daily life contact. As the natural population growth rate increases, so does the population density, increasing the likelihood of mother-to-child and daily-life contact transmission.

The number of beds was positively associated with HB incidence, which was the opposite of the finding of the study by Sun et al. [34]. This could be explained by the fact that our study was based only on the Xinjiang region (14 prefectures), whereas Sun et al. covered all the geographical areas of

China. Medical resources and economics have been found to affect diseases [35, 36]. A high-quality and accessible healthcare system always leads to improved health status in the country, hospitals and other health facilities that are beneficial for HB patients. This problem was also found in the GDP, the level of vaccines and medical care.

The higher the per capita GDP, the lower the risk of HB incidence. Vaccination against HB is an important tool and effective measure for HB prevention and control. China's gross domestic product (GDP) has been growing rapidly in recent years, and as the economy has progressed, the risk of HB incidence has decreased as the HB vaccination rate has increased. The World Health Organization has divided the world into three endemic zones—high, medium and low—based on the differences in the prevalence of hepatitis B worldwide. Areas with hepatitis B surface antigen (HBsAg) levels above 8% are considered highly endemic, those with HBsAg levels in the range of 5–7.99% are considered moderately endemic, and those with HBsAg levels in the range of 2–4.99% are classified as low endemic areas for HB [37]. Because of vaccination and other efficient prevention and control efforts, the country moved from a highly endemic (HBsAg positivity rate: 9.8 %) to a lowly endemic zone (HBsAg positivity rate: 2.60%) for HB in 2014 [38]. This illustrates the economic benefits of HBV vaccines, epidemic prevention and control [39]. This also reflects the disparity in access to health services between poor and wealthy communities, as areas with a higher GDP are likely to have more adequate medical facilities and resources, higher vaccination rates and a lower HB incidence.

In addition, according to our findings, the number of students was negatively associated with the risk of developing HB. This may imply that education impacts the incidence of HB, and a prevention campaign on HB in schools can make students fully aware of the dangers of infectious diseases, ways of transmission and knowledge of effective self-protection, thus reducing the risk of disease. It has been found that increased awareness and education are effective in reducing the outbreak [11, 40].

The spatial autocorrelation results from 2004 to 2019 showed spatially positive correlations for HB distribution in 2006–2012, 2014–2015 and 2018, with the strongest spatial correlations found in 2009. The risk of HB varies among the regions, with an increasing trend from west to east and north to south. High-risk areas included the Changji Autonomous Prefecture, Urumqi City, Karamay City and Bayangol Mongol Autonomous Prefecture. This may be related to population density, health conditions and socioeconomic factors. The seriousness of the HB epidemic in the Bayangol Mongol Autonomous Prefecture may have been influenced by economic conditions, lack of human resources for prevention and control, inadequate medical resources and low awareness of the population about disease prevention. In addition, living habits and conditions such as transportation also contribute to the high incidence of epidemics in the southern region [41]. Urumqi City is the capital of a prefecture-level city in the Xinjiang Uyghur Autonomous Region and is an important integrated transport hub with a highly mobile population that is more likely to be exposed to HB and, therefore, at higher risk in these areas. The populations in Bortala Mongol Autonomous Region, Hami City and Turpan City account for only 2.18%, 2.56% and 2.91% of the total population of Xinjiang, respectively. Due to the sparse population and low mobility, there is a lower risk of HB infection in these areas.

From 2004 to 2019, the overall risk of HB incidence in Xinjiang increased; therefore, future attention should be paid to high-risk areas in Xinjiang, as well as targeted epidemic prevention and control to reduce the risk of HB incidence.

Analysis of the Bayesian spatiotemporal model developed in this study allows authorities to tailor HB epidemic prevention and control strategies for each region's economic development, population

size, and healthcare, thereby accelerating the achievement of the World Health Organization's goal of eliminating viral hepatitis infection by 2030, which was announced in 2016 [42].

Although this study was a spatio-temporal analysis of the risk factors of HB in the Xinjiang region, some limitations should be acknowledged. First, the study was conducted at the prefecture level because data were not available on a finer geographical scale for Xinjiang. As a result, differences within counties were not examined. Second, it is limited to using the number of people in schools to consider whether they are educated and only counts the number of people in each school in Xinjiang. The educational attainment of the entire population should be investigated. Family size, sexual transmission, and mother-to-child transmission are also influential factors in the development of HB. However, these indicators were not included in this study. If these risks of developing HB were considered, the results might have been more accurate. Further studies should explore the main types and characteristics of HB incidence in Xinjiang districts and counties; collect information on the age, sex and occupation of HB patients; analyze the factors influencing the prevalence of HB; and compare the heterogeneity among the districts and counties. It is possible to identify key populations and provide a quantitative basis for assigning appropriate prevention and control measures.

5. Conclusions

The INLA method can be used to embed spatio-temporal models. The R-INLA package can be applied to study spatial epidemiology, including the analysis of factors influencing diseases and their spatial and temporal distributions. The construction of a spatial conditional autoregressive prior in the model can effectively explore the spatial correlation and heterogeneity of the data, and the construction of a second-order wander can effectively explore the changes in the risk of HB incidence over time.

This study found that (i) there was spatial heterogeneity in the risk of HB incidence in Xinjiang; (ii) the risk of HB incidence tended to increase from west to east and from north to south; (iii) factors such as the number of hospital beds, per capita GDP, the number of students and natural growth rate may have an impact on the incidence of HB; (iv) from 2004 to 2019, the risk of HB incidence increased year by year in Xinjiang; and (v) among the 14 prefectures, the risk of incidence was highest in Changji Hui Autonomous Prefecture, Urumqi City, Karamay City and Bayangol Mongol Autonomous Prefecture.

Our findings suggest that HB control measures should focus on these factors in high-risk areas to allocate public health resources more precisely and reduce the burden of HB. Therefore, epidemic prevention and control policies should further be improved according to the risk of morbidity in each region, combined with the influencing factors.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (Grant Nos. 11961065 and 11961071), Youth Science and Technology Innovation Talent of Tianshan Talent Training Program in Xinjiang (Grant No. 2022TSYCCX0099) and the Chinese Foundation for Hepatitis Prevention and Control (YGFK20200059).

Conflict of interest

The authors declare no conflicts of interest.

References

1. Y. P. Yan, H. X. Su, Z. H. Ji, Z. J. Shao, Z. S. Pu, Epidemiology of hepatitis B virus infection in China: Current status and challenges, *J. Clin. Transl. Hepatol.*, **2** (2014), 15–22. https://doi.org/10.14218/JCTH.2013.00030
2. R. Zampino, A. Boemio, C. Sagnelli, L. Alessio, L. E. Adinolfi. E. Sagnelli et al., Hepatitis B virus burden in developing countries, *World J. Gastroenterol.*, **21** (2015), 11941–11953. https://doi.org/10.3748/wjg.v21.i42.11941
3. Y. Jiang, X. Dou, C. Yan, L. Wan, H. C. Liu, M. C. Li, et al., Epidemiological characteristics and trends of notifiable infectious diseases in China from 1986 to 2016, *J. Global Health*, **10** (2020), 020803. https://doi.org/10.7189/jogh.10.020803
4. Y. Wang, B. Wang, J. Wang, H. Asiya, X. Tang, Q. Zheng, Analysis of morbidity and mortality characteristics of notifiable infectious diseases in Xinjiang, 2019, *Bull Dis. Control Prev.*, **35** (2020), 14–19. https://doi.org/10.13215/j.cnki.jbyfkztb.2006018
5. H. Gao, W. Luan, M. Wang, Y. Dong, Epidemic characteristics of main infectious diseases in Yantai city between 2010 and 2012 and prevention strategy research, *Pak. J. Pharm. Sci.*, **29** (2016), 2191–2198.
6. H. Wang, P. Men, Y. F. Xiao, P. Gao, M. Lv, Q. L. Yuan, et al., Hepatitis B infection in the general population of China: a systematic review and meta-analysis, *BMC Infect. Dis.*, **19** (2019), 811. https://doi.org/10.1186/s12879-019-4428-y
7. Z. Wang, Y. Liu, R. Zhang, X. Gong, Y. Shi, H. Zhang, Epidemiological characteristics of hepatitis B in China from 2004 to 2013, *Chin. Gen. Pract.*, **20** (2017), 2879–2083.
8. B. Zhu, J. Liu, Y. Fu, B. Zhang, Y. Mao, Spatio-temporal epidemiology of viral hepatitis in China (2003–2015): Implications for prevention and control policies, *Int. J. Environ. Res. Public Health,* **15** (2018), 661.
9. E. Sagnelli, C. Sagnelli, M. Pisaturo, M. Macera, N. Coppola, Epidemiology of acute and chronic hepatitis B and delta over the last 5 decades in Italy, *World J. Gastroenterol.*, **20** (2014), 7635–7643. https://doi.org/10.3748/wjg.v20.i24.7635
10. E. Ochola, P. Ocama, C. G. Orach, Z. K. Nankinga, J. N. Kalyango, W. McFarland, et al., High burden of hepatitis B infection in Northern Uganda: results of a population-based survey, *BMC Public Health*, **13** (2013), 727. https://doi.org/10.1186/1471-2458-13-727
11. H. Kinfe, E. G. Sendo, K. B. Gebremedhin, Prevalence of hepatitis B virus infection and factors associated with hepatitis B virus infection among pregnant women presented to antenatal care clinics at Adigrat General Hospital in Northern Ethiopia, *Int. J. Women's Health*, **13** (2021), 119–127.

12. M. Zhang, J. Ge, Z. Lin, The impact of the number of Hospital beds and spatial heterogeneity on an SIS epidemic model, *Acta Appl. Math.*, **167** (2020), 59–73. https://doi.org/10.1007/s10440-019-00268-y

13. K. Liu, S. Yang, Q. Zhou, Y. Qiao, Spatiotemporal evolution and spatial network analysis of the urban ecological carrying capacity in the Yellow River basin, *Int. J. Environ. Res. Public Health*, **19** (2021), 229. https://doi.org/10.3390/ijerph19010229

14. Y. H. Xu, Z. L. Yang, Specification tests for temporal heterogeneity in spatial panel data models with fixed effects, *Reg. Sci. Urban Econ.*, **81** (2020), 103488. https://doi.org/10.1016/j.regsciurbeco.2019.103488

15. H. Rue, S. Martino, N. Chopin, Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, *J. R. Stat. Soc. B*, **71** (2009), 319–392. https://doi.org/10.1111/j.1467-9868.2008.00700.x

16. Y. Fong, H. Rue, J. Wakefield, Bayesian inference for generalized linear mixed models, *Biostatistics*, **11** (2010), 397–412. https://doi.org/10.1093/biostatistics/kxp053

17. L. Grilli, S. Meteli, C. Rampichini, Bayesian estimation with integrated nested Laplace approximation for binary logit mixed models, *J. Stat. Comput. Simul.*, **85** (2015), 2718–2726. https://doi.org/10.1080/00949655.2014.935377

18. X. L. Sun, B. Minasny, H. L. Wang, Y. G. Zhao, G. L. Zhang, Y. J. Wu, Spatiotemporal modelling of soil organic matter changes in Jiangsu, China between 1980 and 2006 using INLA-SPDE, *Geoderma*, **384** (2021), 114808. https://doi.org/10.1016/j.geoderma.2020.114808

19. M. C. Rufener, P. G. Kinas, M. F. Nobrega, L. D. F. Oliveira, Bayesian spatial predictive models for data-poor fisheries, *Ecol. Modell.*, **348** (2017), 125–134. https://doi.org/10.1016/j.ecolmodel.2017.01.022

20. N. Lezama-Ochoa, M. G. Pennino, M. A. Hall, J. Lopez, H. Murua, Using a Bayesian modelling approach (INLA-SPDE) to predict the occurrence of the Spinetail Devil Ray (Mobular mobular), *Sci. Rep.*, **10** (2020), 18822. https://doi.org/10.1038/s41598-020-73879-3

21. I. T. Vlad, P. Juan, J. Mateu, Bayesian spatio-temporal prediction of cancer dynamics, *Comput. Math. Appl.*, **70** (2015), 857–868. https://doi.org/10.1016/j.camwa.2015.06.006

22. J. H. Froelicher, G. Forjaz , P. S. Rosenberg, P. Chernyavskiy, Geographic disparities of breast cancer incidence in Portugal at the district level: A spatial age-period-cohort analysis, 1998–2011, *Cancer Epidemiol.*, **74** (2021), 102009. https://doi.org/10.1016/j.canep.2021.102009

23. B. Rowland, S. P. Rushton, M. D. F. Shirley, M. A. Brown, G. E. Budge, Identifying the climatic drivers of honeybee disease in England and Wales, *Sci. Rep.*, **11** (2021), 21953. https://doi.org/10.1038/s41598-021-01495-w

24. S. Y. Bie, X. J. Hu, H. G. Zhang, K. Wang, Z. Dou, Influential factors and spatial-temporal distribution of tuberculosis in mainland China, *Sci. Rep.*, **11** (2021), 6274. https://doi.org/10.1038/s41598-021-85781-7

25. H. Rue, S. Martino, Approximate Bayesian inference for hierarchical Gaussian Markov random field models, *J. Stat. Plann. Inference*, **137** (2007), 3177–3192. https://doi.org/10.1016/j.jspi.2006.07.016

26. N. Aral, H, Bakir, Spatio-temporal pattern of COVID-19 outbreak in Turkey, *GeoJournal*, **2022** (2022), 1–12. https://doi.org/10.1007/s10708-022-10666-9

27. J. Abellan, S. Richardson, N. Best, Use of space-time models to investigate the stability of patterns of disease, *Environ. Health Perspect.*, **116** (2008), 1111–1119. https://doi.org/10.1289/ehp.10814

28. A. B. Lawson, *Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology*, 3rd edition, Chapman and Hall/CRC, 2018.

29. L. K. Held, J. Besag, Modelling risk from a disease in time and space, *Stat. Med.*, **17** (1998), 2045–2060. https://doi.org/10.1002/(SICI)1097-0258(19980930)17:18<2045::AID-SIM943>3.0.CO;2-P

30. J. Teng J, S. Ding, H. Zhang, K. Wang, X. Hu, Bayesian spatiotemporal modelling analysis of hemorrhagic fever with renal syndrome outbreaks in China using R-INLA, *Zoonoses Public Health*, **70** (2023), 46–57. https://doi.org/10.1111/zph.12999

31. L. H. Knorr, Bayesian modelling of inseparable space time variation in disease risk, *Stat. Med.*, **19** (2000), 2555–2567. https://doi.org/10.1002/1097-0258(20000915/30)19:17/18<2555::AID-SIM587>3.0.CO;2-%23

32. X. Wang, Y. R. Ryan, J. J. Faraway, *Bayesian Regression Modeling with INLA*, Taylor & Francis Group an Informa Business, Chapman and Hall/CRC, (2018), 77–79. https://doi.org/10.1201/9781351165761

33. L. Bernardinelli, D. Clayton, C. Pascutto, C. Montomoli, M. Ghislandi, M. Songini, Bayesian analysis of space-time variation in disease risk, *Stat. Med.*, **14** (1995), 2433–2443. https://doi.org/10.1002/sim.4780142112

34. W. Sun, J. Gong, J. Zhou, Y. Zhao, J. Tan, A. N. Ibrahim, et al., Spatial, social and environmental study of tuberculosis in China using statistical and GIS technology, *Int. J. Environ. Res. Public Health*, **12** (2015), 1425–1448. https://doi.org/10.3390/ijerph120201425

35. X. Ma, X. F. Luo, L. Li, Y. Li, G. Q. Sun, The influence of mask use on the spread of COVID-19 during pandemic in New York City, *Results Phys.*, **34** (2022), 105224. https://doi.org/10.1016/j.rinp.2022.105224

36. G. Q. Sun, H. T. Zhang, L. L. Chang, Z. Jin, H. Wang, S. G. Ruan, On the dynamics of a diffusive foot-and-mouth disease model with nonlocal infections, *SIAM J. Appl. Math.*, **82** (2022), 1587–1610. https://doi.org/10.1137/21M141299

37. X. Liang, S. Bi, W. Yang, L. Wang, G. Cui, F. Cui, et al., Epidemiological serosurvey of hepatitis B in China declining HBV prevalence due to hepatitis B vaccination, *Vaccine*, **27** (2009), 6550–6557. https://doi.org/10.1016/j.vaccine.2009.08.048

38. F. Cui, L. Shen, L. Li, H. Wang, F. Wang, S. Bi, et al., Prevention of chronic hepatitis B after 3 decades of escalating vaccination policy, China, *Emerging Infect. Dis.*, **23** (2017), 765–772. https://doi.org/10.3201/eid2305.161477

39. Z. Y. Gong, Global trends in the progress of routine vaccination against hepatitis B in early childhood in 2003, *Dis. Surveillance*, **19** (2004), 113–115.

40. X. Ma, G. Q. Sun, Z. H. Wang, Y. M. Chu, Z. Jin, B. L. Li, Transmission dynamics of brucellosis in Jilin province, China: Effects of different control measures, *Commun. Nonlinear Sci. Numer. Simul.*, **114** (2022), 106702.

41. H. L. Li, X. L. Zhang, K. Wang, A quantitative study on the epidemic situation of tuberculosis based on thetransmission disease dynamics in 14 prefectures of Xinjiang from 2005 to 2017, *Chin. J. Infect. Control,* **17** (2018), 945–950.

42. J. Liu, W. Liang, W. Jing, M. Liu, Countdown to 2030: eliminating hepatitis B disease, China, *Bull World Health Organ*, **97** (2019), 230–238. https://doi.org/10.2471/BLT.18.219469

AIMS Press