

RESEARCH

Open Access



# Spatiotemporal dynamics and potential ecological drivers of acute respiratory infectious diseases: an example of scarlet fever in Sichuan Province

Cheng Li<sup>1†</sup>, Rujun Liao<sup>2†</sup>, Wenhui Zhu<sup>1</sup>, Guiyu Jiang<sup>1</sup>, Yumeng Wang<sup>1</sup>, Ling Li<sup>2</sup>, Tao Zhang<sup>1\*</sup> and Qiang Lv<sup>2\*</sup>

## Abstract

**Object:** Scarlet fever is an acute respiratory infectious disease that endangers public health and imposes a huge economic burden. In this paper, we systematically studied its spatial and temporal evolution and explore its potential ecological drivers. The goal of this research is to provide a reference for analysis based on surveillance data of scarlet fever and other acute respiratory infectious illnesses, and offer suggestions for prevention and control.

**Method:** This research is based on a spatiotemporal multivariate model (Endemic-Epidemic model). Firstly, we described the epidemiology status of the scarlet fever epidemic in Sichuan Province from 2016 to 2019. Secondly, we used spatial autocorrelation analysis to understand the spatial pattern. Thirdly, we applied the endemic-epidemic model to analyze the spatiotemporal dynamics by quantitatively decomposing cases into endemic, autoregressive, and spatiotemporal components. Finally, we explored potential ecological drivers that could influence the spread of scarlet fever.

**Results:** From 2016 to 2019, the incidence of scarlet fever in Sichuan Province varied much among cities. In terms of temporal distribution, there were 1–2 epidemic peaks per year, and they were mainly concentrated from April to June and October to December. In terms of transmission, the endemic and temporal spread were predominant. Our findings imply that the school holiday could help to reduce the spread of scarlet fever, and a standard increase in Gross Domestic Product (GDP) was associated with 2.6 folds contributions to the epidemic among cities.

**Conclusion:** Scarlet fever outbreaks are more susceptible to previous cases, as temporal spread accounted for major transmission in many areas in Sichuan Province. The school holidays and GDP can influence the spread of infectious diseases. Given that covariates could not fully explain heterogeneity, adding random effects was essential to improve accuracy. Paying attention to critical populations and hotspots, as well as understanding potential drivers, is recommended for acute respiratory infections such as scarlet fever. For example, our study reveals GDP is positively

<sup>†</sup>Cheng Li and Rujun Liao contributed equally to this work.

\*Correspondence: scdxzhangtao@163.com; 913154470@qq.com

<sup>1</sup> Department of Epidemiology and Health Statistics, West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, Sichuan, China

<sup>2</sup> Sichuan Center for Disease Control and Prevention, Chengdu, Sichuan, China



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

associated with spatial spread, indicating we should consider GDP as an important factor when analyzing the potential drivers of acute infectious disease.

**Keywords:** Spatiotemporal analysis, Potential ecological drivers, Scarlet fever, Acute respiratory infections

## Background

Acute respiratory infection (ARI) is one of the leading causes of morbidity and mortality worldwide. Every year, approximately 156 million instances of acute respiratory infections are reported, with children under the age of 5 and those over the age of 65 being the most vulnerable [1]. According to the estimate of the World Health Organization (WHO) [2], ARIs rank as the fourth-highest global cause of mortality and resulting in nearly 3 million deaths worldwide in 2016 (40 deaths per 100,000). Acute respiratory infections such as scarlet fever [3], Ebola [4], and coronavirus disease 2019 (COVID-19) [5] have historically posed a substantial threat to public health and imposed a significant economic burden.

While it is critical to have heightened vigilance regarding emerging high-risk respiratory infections, we should not lose sight of those re-emerging infectious diseases with high incidence and outbreaks. A typical example of emerging infectious diseases is COVID-19, which caused a heavy burden on public health and the economy. As of 1 July 2022, there have been 545,226,550 confirmed cases of COVID-19, including 6,334,728 deaths [6]. Another typical example of the re-surfaced traditional infectious disease is scarlet fever. It was associated with high levels of morbidity and mortality when epidemics were common in the 18th and 19th centuries in Europe and the USA [7]. Although this disease nearly disappeared during the twentieth century, many countries have recently experienced a re-emergence of scarlet fever. This global re-emergence of scarlet fever has caused more than 600,000 cases around the world [8]. Scarlet fever is re-emerging as a public health threat.

However, developing vaccines, viral mutations, and the growing problem of antibiotic-resistant bacteria pose remarkable challenges in the prevention of ARIs such as scarlet fever [9]. Therefore, investigations on epidemiology and spatiotemporal dynamics are still critically needed to offer the scientific foundation for disease control and prevention. The spatiotemporal studies of infectious diseases often based on surveillance data and easily accessible variables. At present, there are two types of methods for modeling infectious disease surveillance data [10]: one is the mechanistic model, and the other is the statistical model. Mechanistic models, such as SIR models, build on extensive knowledge of basic epidemiology and aim to capture

the important mechanisms of disease transmission. Statistical models can be used to explore and explain statistical variability in data. But the former approach ignores the dynamics of space [11], and the latter (statistical models) could not capture specific characteristics of infectious disease [12].

In order to solve the above problems and to better explore the infectious disease on surveillance data, Held [13] et al. proposed a spatio-temporal model (also known as Epidemic-Epidemic Model, EE model) on the basis of branching process, combining the mechanistic model and statistical model. This model decomposes the disease incidence into an “endemic” and an “epidemic” part (e.g., epidemic with region and epidemic among regions), which can simulate the spread of infectious disease and quantify the effect of covariates in different components. This modeling framework has been used widely to analyze other diseases such as tuberculosis [14], influenza [15], dengue [16], and so on.

As for the spatiotemporal study of scarlet fever, most studies concentrated on dynamics of the epidemic part (e.g., epidemic from temporal aspect and epidemic from spatial aspect) [17–20], ignoring the endemic part (e.g., indirect transmission from the environment). And also they did not qualify the effect of covariates in different dimensions (i.e., time, space, and endemic). In recent years, more and more study shows ecological factors could be a strong force in transmission [21]. However, previous research on scarlet fever focused on meteorological factors and failed to explore the association between ecological factors and spread [18, 19, 22]. To better understand the spatiotemporal dynamics of scarlet fever and explore potential ecological drivers in different dimensions, we applied the EE model to visualize space-time dynamics and describe the factors related to transmission.

Most spatiotemporal research of ARIs is based on the surveillance data of infectious diseases and a limited number of covariates, the urgency of the analysis, the accessibility of covariates, and the result of instability are the main challenges. In this paper, we took scarlet fever as an example, using the EE model to decompose the transmission into three dimensions: endemic, epidemic within city, and epidemic among cities. We also explore the effects of ecological factors in these dimensions. We hope our study can provide a reference for scarlet fever and other ARIs.

## Method

### Data

#### Case report

According to Chinese Infectious Diseases Law, when doctors identify any probable, clinical, or laboratory-confirmed case of scarlet fever, they need to report the case to the Notifiable Infectious Diseases Reporting Information System (NIDRIS) within 24 h [23].

#### Data collection

The data of reported cases (including the patient's age, sex, occupation, and address) of scarlet fever were collected by the Nationwide Notifiable Infectious Diseases Reporting Information System (NIDRIS), which was used under license and not publicly available. We extracted data on scarlet fever in 2016–2019 of the Sichuan province from the NIDRIS, and the data included the number of cases, and patient data on age, sex, and date of disease onset, diagnosis, and address limited to cities. Because all the data in our study were anonymous and without access to the identity information of the patients, and hence, informed consent were waived by the ethics committee of the Sichuan Center for Disease Control and Prevention. The research has been approved by the ethics committee of the Sichuan Center for Disease Control and Prevention (SCCDCIRB-2022-185).

Additionally, we collected data on ecological factors from the *Sichuan Statistical Yearbook* (<http://tjj.sc.gov.cn/scstjj/c105855/nj.shtml>), which included population, population density, number of health institutions, number of primary schools, number of kindergartens, passenger-kilometers of Highways and Gross Domestic Product (GDP).

#### Study area

Sichuan province, located in the hinterland of southwest China, between 26°03' -34 °19' N and 97°21' -108 °33' E, is located in the upper reaches of the Yangtze River and covers an area of 486,000 km<sup>2</sup>, ranking fifth in China. Sichuan consists of 21 cities and prefectures, with the fourth largest population and the sixth GDP in China. The development of social and economic conditions varies between cities and prefectures in the province.

## Statistical analysis

### Descriptive statistics

We cleaned and analyzed the data using R version 4.0.2. Count data were expressed by frequency and incidence. Firstly, we described the characteristics of population distribution and spatiotemporal distribution of scarlet

fever. Secondly, we applied Global Moran's *I* statistics to calculate the spatial autocorrelation [24]. The formula is defined as:

$$I = \frac{n \sum_{i,j} \omega_{ij} (X_i - \bar{X})(X_j - \bar{X})}{\sum_{i,j} \omega_{ij} \sum_{i=1}^n (X_i - \bar{X})^2} \quad (1)$$

where n denotes the number of spatial units,  $\omega_{ij}$  denotes the weight matrix,  $X$  denotes the variable of interest (i.e., incidence), and  $\bar{X}$  denotes the mean value of all units. The value for Moran's *I* can range from -1 to 1 where:

- i) 1: The variable of interest is perfectly dispersed;
- ii) 0: The variable of interest is randomly dispersed;
- iii) 1: The variable of interest is perfectly clustered together.

Thirdly, we used Anselin's Local Moran's *I* (local indicators of spatial association, LISA) test statistics to explore the clusters or outliers in the study area. For the spatial unit *i*, the formula is defined as [25]:

$$I_i = \frac{X_i - \bar{X}}{S^2} \sum_j \omega_{ij} (X_j - \bar{X}) \quad (2)$$

where  $S^2$  denotes the variance, the value of  $I_i$  can disclose spatial clusters or outliers. There are five categories according to the value: high-high, low-low, high-low, low-high, and non-significant. The high-high and low-low areas represent spatial clusters, while the high-low and low-high areas were the outliers. The spatial analysis was conducted by GeoDa 1.20.

### Transmission decomposition

We adopted the Endemic-epidemic model developed by Held and Paul [26] to decompose the incidence level. The formulas are as below:

$$\begin{aligned} Y_{i,t} | Y_{i,t-1} &\sim NegBin(\mu_{it}, \psi) \\ \mu_{it} &= v_{it}e_{it} + \lambda_{it}Y_{i,t-1} + \phi_{it} \sum_{j \neq i} \omega_{ij} Y_{j,t-1} \end{aligned} \quad (3)$$

Infectious disease counts  $Y_{i,t}$  in the region  $i=1,\dots,21$  during weeks  $t=1,\dots,208$ .  $Y_{i,t}$  follows a negative binomial distribution with the mean  $\mu_{it}$  and overdispersion parameter  $\psi > 0$ . We decomposed the transmission into three component as follows:

- (1) endemic component ( $v_{it}e_{it}$ ): infections from outside the study area or from indirect transmission (i.e., infected by bacteria from environment)
- (2) autogressive component ( $\lambda_{it}Y_{i,t-1}$ ): reproduction of scarlet fever within city *i* (i.e., epidemic within city)

- (3) neighborhood component ( $\phi_{it} \sum_{j \neq i} \omega_{ij} Y_{i,t-1}$ ): transmission from other regions except city  $i$  (i.e., epidemic among cities).

In Eq. (3),  $v_{it}$  is log-linear predictor of the endemic component that, multiplied by an offset such as population  $e_{it}$ , could describe incidence due to sociodemographic variation. The coefficient  $\lambda_{it}$  represents the transmission of infections from the past period, and the coefficient  $\phi_{it}$  quantifies the contribution of spatial transmission capturing infections from other cities. We defined the spatial weight matrix as

$$\omega_{ij} = O_{ij}^{-\rho} \quad (4)$$

where  $O_{ij}$  represents the path distance between cities  $j$  and  $i$ , and  $\rho$  is the decay parameter can be estimated from the data. The path distance  $O_{ij}$  is on an ordinal scale based on the adjacency index [27].

Meanwhile, we added seasonal variation and covariates related to ecological situation as follows:

$$\begin{cases} \log(v_{it}) = \alpha_0 + \alpha_i + \kappa_{it}^T \alpha + \sum_S \gamma_s \sin(\theta t) + \delta_s \cos(\theta t) \\ \log(\lambda_{it}) = \beta_0 + \beta_i + u_{it}^T \beta \\ \log(\phi_{it}) = \gamma_0 + \gamma_i + Z_{it}^T \gamma \end{cases} \quad (5)$$

$$\gamma_s \sin(\theta t) + \delta_s \cos(\theta t) = A_s \sin(\theta t + \varphi) \quad (6)$$

$$A_s = \sqrt{\gamma_s^2 + \delta_s^2} \quad (7)$$

Where  $\alpha_0$ ,  $\beta_0$ , and  $\gamma_0$  are intercepts;  $\kappa_{it}^T$ ,  $u_{it}^T$  and  $Z_{it}^T$  are covariates. Due to the covariates in our study could not explain the spatiotemporal heterogeneity, we introduce random effects  $\alpha_i$ ,  $\beta_i$  and  $\gamma_i$  with  $\alpha_i \sim N(0, \sigma_\alpha^2)$ ,  $\beta_i \sim N(0, \sigma_\beta^2)$ ,  $\gamma_i \sim N(0, \sigma_\gamma^2)$  in the model;  $\alpha$ ,  $\beta$ ,  $\gamma$  denotes the coefficient of covariates. Seasonal terms  $\left\{ \sum_S \gamma_s \sin(\theta t) + \delta_s \cos(\theta t) \right\}$  reflect seasonally varying incidence [28], and  $A_s$  is the amplitude of the corresponding sine-cosine curve and  $\theta = \frac{2\pi}{52}$ .

According to previous research [27], the power-law algorithm was superior to other methods for spatial weight. Relevant parameter estimation was based on penalized maximum likelihood ratio method [29]. Given that the Akaike information criterion (AIC) is inappropriate to compare models that include random effects [30], we used the ranked probability scores (RPS) and logS scores (logS) [31] to select the best model and quantify the effect of covariates among the three components. These scores measure differences between the predicted distribution  $P$  of the fitted model and the observed value  $y$ . The lower the score is, the better the fit. We first chose

the best-fixed effect model by AIC, then we added the random effects and use logS and RPS to evaluate the final model. The model framework is based on the R package "surveillance" [32]. We also used the packages including "spdep", "sf", "ggplot2" and "dplyr" to run endemic-epidemic model.

### Exploration of ecological effects

As an ecological study, we also aim to provide clues from the perspective of correlation between covariates and the spread of infectious diseases. For the analysis of surveillance data, indicators of ecological factors are easy to gather. We took scarlet fever in Sichuan as an example to explore the effects of ecological factors on three components of transmission. However, as with most, this study could not include all factors, so we introduced random effects in our model to alleviate the uncertainty caused by covariates.

Our model incorporated multiple ecological covariates. Table A1 in the Additional file describes the meaning and values of covariates that could influence the spread of the infectious disease [33–37]. We added the covariates in three components based on epidemiology knowledge. First, we use the stepwise regression method to explore the optimal fixed effect model by AIC. We began with an intercept-only model (model 1) with a population offset in the endemic component and spatial weight matrix based on the power law. Secondly, we sequentially added ecological covariates in the three components. Thirdly, we added random effects to capture more realism in our model. Last, we chose the final model by comparing logS and RPS.

## Results

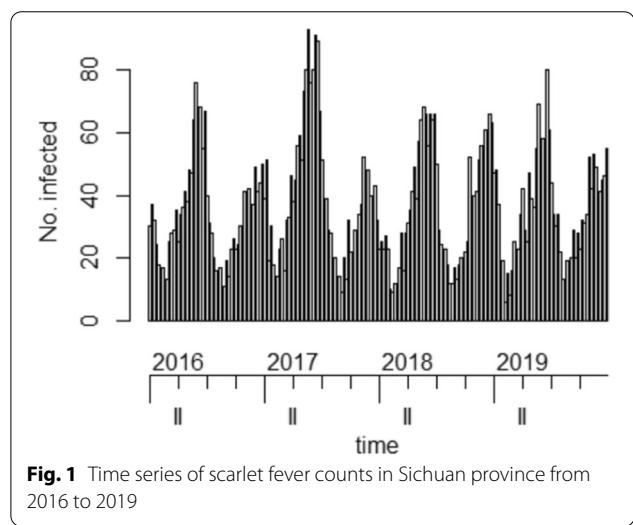
### Basic characteristic

Table 1 summarizes the basic characteristic of patients. A total of 7356 scarlet fever cases were reported from 2016 to 2019 in Sichuan province. Males outnumbered females. (4388 vs. 2968 cases). The youngest patient was confirmed 1 day after birth, and the oldest patient is 59 years old. The majority of recorded cases (7160 cases, 97.34%) came from the native population rather than the floating population (i.e., a large and increasing number of migrants without local household registration status). The incidence of the 4–6 years group was the highest of all age groups (Table A2 in Additional file), which is more than 1 per 100,000 population each year.

Figure 1 shows the temporal distribution of Sichuan province from 2016 to 2019. It also can be seen that there is a certain seasonality and periodicity between 2016 and 2019, with 1–2 epidemic peaks every year. The peak period of incidence was mainly concentrated in the second quarter of the year (April to June). A few areas

**Table 1** Description of scarlet fever cases in Sichuan Province from 2016 to 2019

	Total N(%)	Year			
		2016 N(%)	2017 N(%)	2018 N(%)	2019 N(%)
Gender					
Female	2968 (40.35)	657 (38.47)	805 (40.21)	754 (41.27)	752 (41.34)
Male	4388 (59.65)	1051 (61.53)	1197 (59.79)	1073 (58.73)	1067 (58.66)
Age (year)					
0–3	1287 (17.50)	306 (17.92)	356 (17.78)	298 (16.31)	327 (17.98)
4–6	3890 (52.88)	916 (53.63)	1096 (54.75)	1003 (54.90)	875 (48.10)
7–9	1670 (22.70)	390 (22.82)	453 (22.62)	399 (21.84)	428 (23.53)
10–12	296 (4.03)	64 (3.75)	72 (3.60)	86 (4.71)	74 (4.07)
13–15	106 (1.44)	15 (0.88)	16 (0.80)	19 (1.04)	56 (3.08)
16–69	107 (1.45)	17 (1.00)	9 (0.45)	22 (1.20)	59 (3.24)
Source of cases					
Other province	57 (0.77)	11 (0.64)	16 (0.80)	11 (0.61)	19 (1.04)
Foreign country	1 (0.01)	1 (0.06)	0 (0.00)	0 (0.00)	0 (0.00)
The same city	7160 (97.34)	1675 (98.07)	1933 (96.55)	1728 (97.53)	1770 (97.30)
Other cities	138 (1.88)	21 (1.23)	53 (2.65)	34 (1.86)	30 (1.66)
Total	7356	1708	2002	1827	1819



occasionally have a small peak period in the fourth quarter (October to December), such as Mianyang in 2018 (Fig. A1 in Additional file).

Scarlet fever cases were found in 21 cities and prefectures across Sichuan province, with the number of cases varying in different cities (Fig. A1 in Additional file). To be specific, Chengdu ranked first, with a total of 2428 cases, accounting for about one-third of the total cases. Leshan, Mianyang, and Liangshan were the following three cities (/autonomous prefectures). Ya'an and Ganzi had the fewest number of cases, with no more than 10

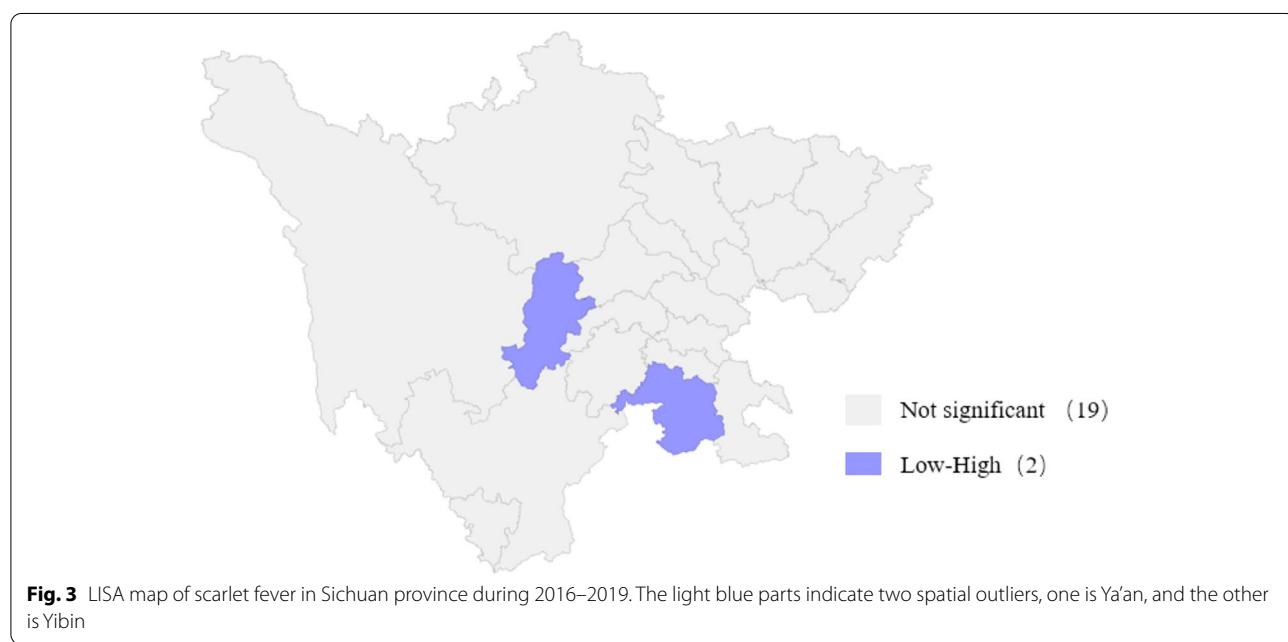
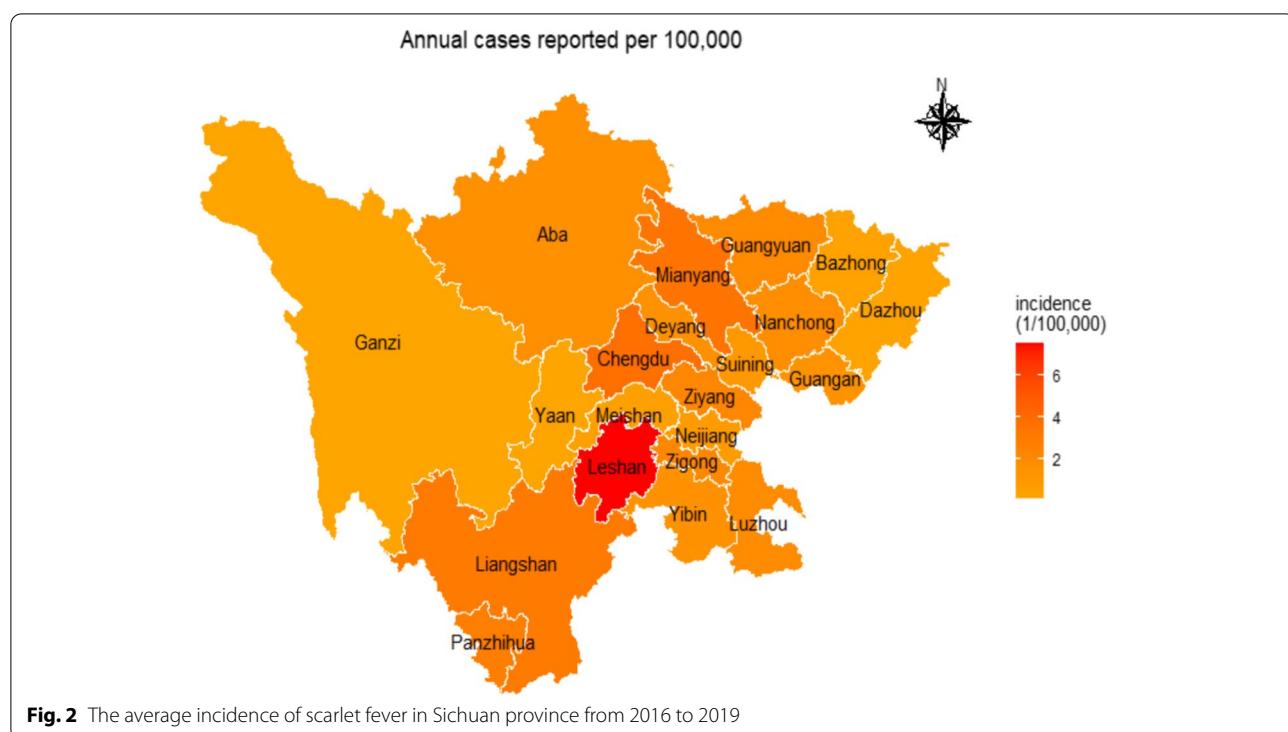
cumulative cases in the 4 years. The annual average incidence was 1.868 cases (per 100,000 population).

Figure 2 shows the spatial distribution. The incidence of scarlet fever varies greatly in different areas. Cases were mainly concentrated in Leshan, Mianyang, Chengdu, and Aba Tibetan Autonomous Prefecture. Among them, Leshan had the highest average annual incidence of 7.556 cases (/100,000 population), whereas Ganzi Tibetan Autonomous Prefecture had the lowest average annual incidence rate of 0.0825 cases (/100,000 population).

In our study, the global Moran's  $I = -0.0512$  ( $Z = -0.013$ ,  $P = 0.484$ ), indicated no significant spatial autocorrelation of scarlet fever incidence during 2016–2019 in Sichuan province. For Local Moran's  $I$ , as shown in Fig. 3, there were only two low-high outliers with light blue: Ya'an and Yibin, where the given cities evidenced low incidences but were surrounded by high incidence areas.

#### Transmission decomposition

Figure 4 shows the averaged proportions of fitted components all over weeks. It can be seen that the autoregressive component and endemic component have a certain proportion, whereas the proportion of the spatiotemporal component is very low. This is consistent with the results spatial autocorrelation, which means scarlet fever is less affected by spatiotemporal transmission. According to the average proportions shown in Fig. 4, we classified three types of transmission as follows:



- (1) The main transmission is resulted from endemic, including Ganzi, Aba, Ya'an, Meishan, Dazhou, and Bazhong.
- (2) The main transmission is epidemic within city, including Mianyang, Chengdu, Leshan, and Liangshan.
- (3) The main transmission are resulted from epidemic within city and endemic, including Ziyang, Neijiang, Zigong, Yibin, Luzhou, Suining, Nanchong, Guang'an, Guangyuan, Deyang, and Panzhihua.

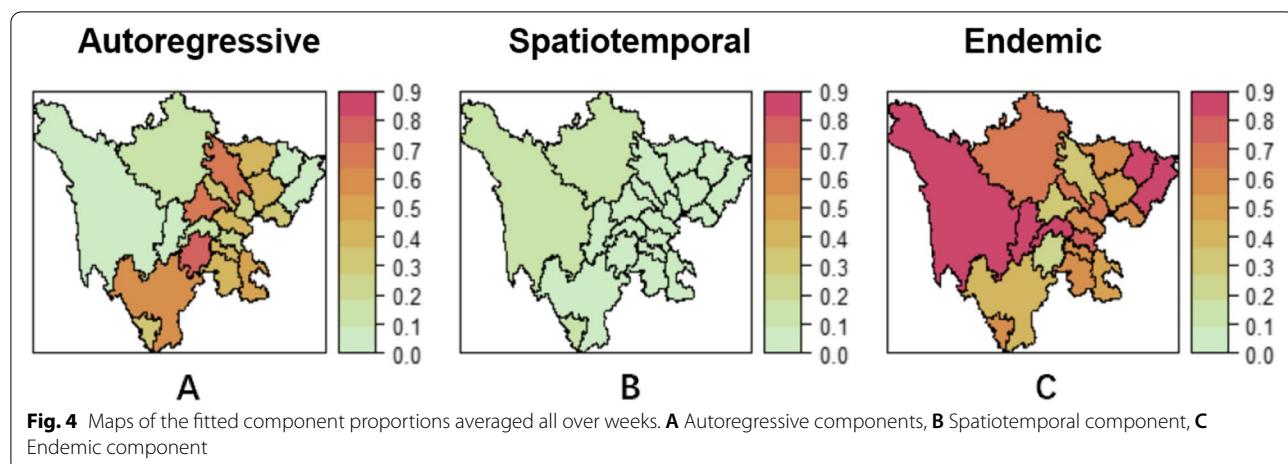


Fig. A2 (Additional file) presented fitted values for 21 cities (/prefectures) and Fig. A3 in the Additional file reveals the seasonality of the endemic mean. We took the six cities of Liangshan, Mianyang, Dazhou, Bazhong, Yibin, and Suijing as examples (Fig. 5). The incidence was quite low in Dazhou and Bazhong, and the transmission was predominantly composed of endemic and a little spatiotemporal component. In high-incidence cities such as Liangshan and Mianyang, scarlet fever cases were mainly influenced by transmitted cases from previous periods. In Suining and Yibin, the transmission was composed of endemic and autoregressive components

#### Association with ecological factors

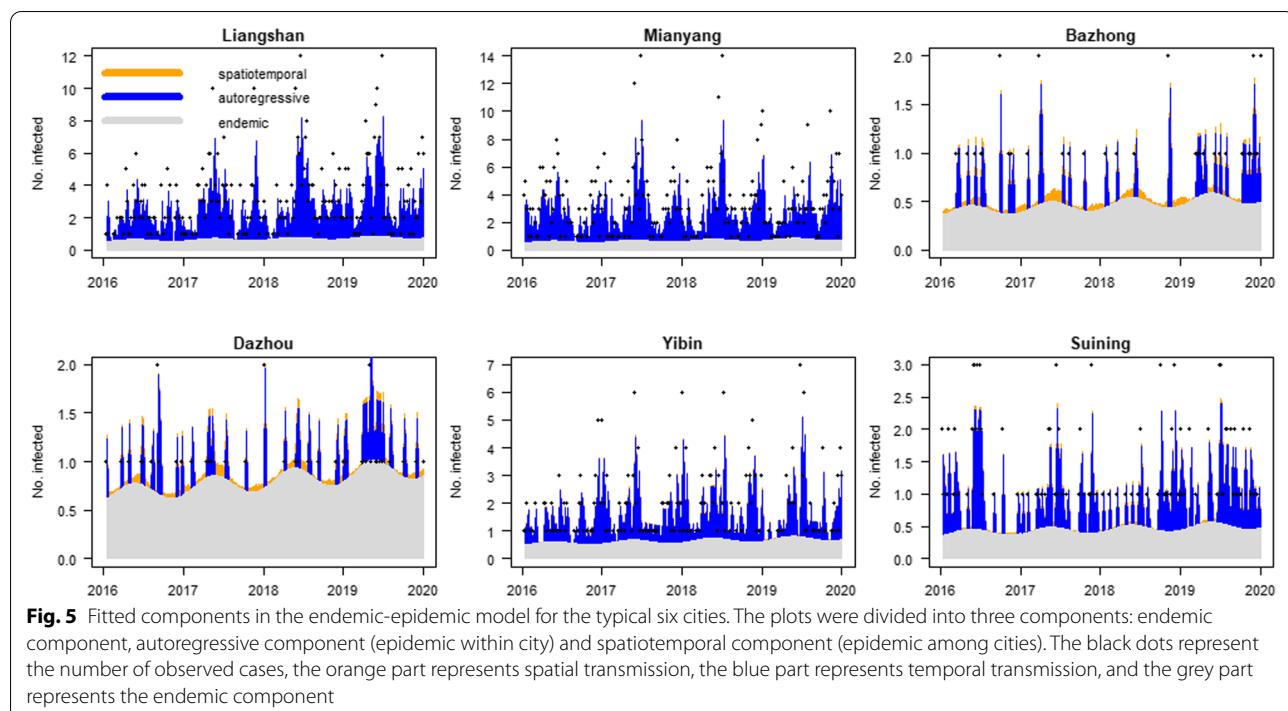
Table A3 in the Additional file presents a summary of the model selection and comparison process. The model 5 yielded the lowest AIC (from 11903 in model 1 to 11,604.41 in model 5 without random effects, in Additional file, Table A3). In the final model, the autoregressive component included population density, school holidays, GDP, number of health institutions, number of preschools, and number of kindergartens. The spatiotemporal component consisted of population density, GDP, and passenger kilometers of Highway. Endemic component comprised school holidays, number of preschools and number of kindergartens.

Because ecological factors could not fully explain the spatiotemporal variation of transmission mechanisms and incidence levels, we allowed the intercept (mean levels of  $\nu_{it}$ ,  $\lambda_{it}$  and  $\phi_{it}$ ) in the three components of infections to vary for each city as random effects. The results of the final model were shown in Table 2. Table 2 shows the relative risk (RR) and 95% confidence interval (CI) for each explanatory variable in different components. The overdispersion parameter is 0.1053, which means our data of surveillance is discrete. The weight decay parameter is

0.0108, indicating there is a low correlation among cities. Fig. A4 and Fig. A5 in the Additional file shows the spatial connectivity weights and weight matrix, respectively. School holidays is a protective factor in both endemic component and autoregressive component (RR = 0.4712, 95%CI: 0.2776–0.7999; RR = 0.5299, 95%CI: 0.3883–0.7233). With each standard level increase in GDP, the contribution to the spatial transmission was higher by 2.6 folds (RR = 2.6898, 95% CI: 1.1516–6.2822).

#### Random effects

Random effects are useful if the regions exhibit heterogeneous incidence levels not explained by observed covariates [10]. As covariates in our study could not account for spatiotemporal heterogeneity, we introduced random effects in our fixed effect model to capture the heterogeneity in different components. The variances of random effects were 0.163 (0.086, 0.195), 1.3057 (0.8895, 1.8535), and 1.5493 (1.1173, 2.0107) in autoregressive component, spatiotemporal component, and endemic component, respectively. The estimates of random effects was summarized in Table A 5 (Additional file). As shown in Fig. 6, in terms of endemic component, the random effects of Panzhihua, Chengdu, Mianyang, and Dazhou are all greater than 1. In the autoregressive component, the random effect of Panzhihua, Aba, Luzhou, and Chengdu was greater than 1, while in the spatiotemporal component, only Zigong has a random effect greater than 1. The value above (or below)1 indicates that the average incidence rate in one area is higher (or lower) than in other areas [38]. Given the number of reported infections, this could be interpreted as a tendency to produce more or fewer cases in a particular area.



## Discussion

We analyzed the reported cases of scarlet fever in Sichuan from 2016 to 2019. Then explored the transmission dynamic rules based on three dimensions (i.e,

**Table 2** Coefficient estimates from the endemic-epidemic model

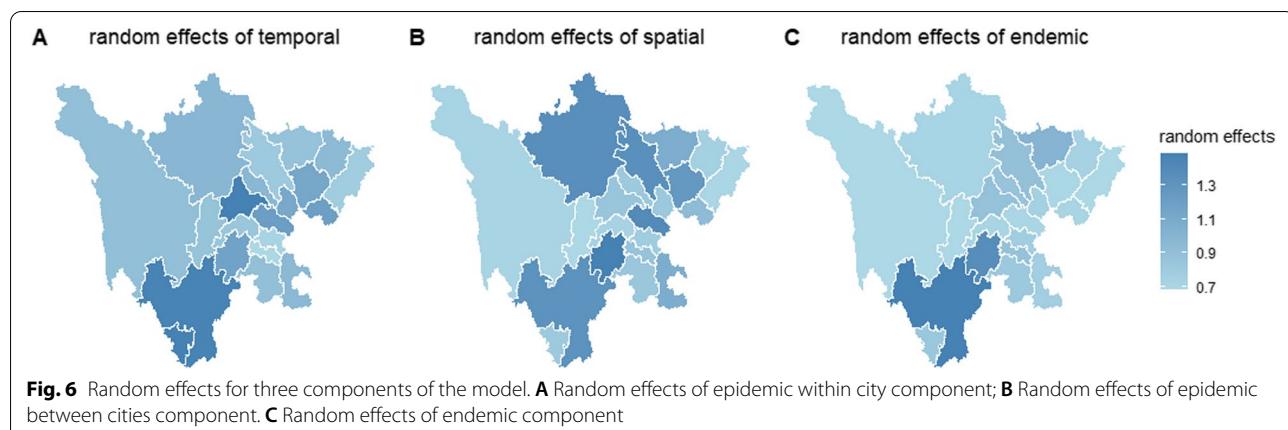
Parameter	RR	95%CI	P value
Epidemic within the city (autoregressive)			
Intercept	0.1588	(0.0034, 7.4853)	–
Density	0.8162	(0.5192, 1.2831)	0.379
School holidays	<b>0.4712</b>	(0.2776, 0.7999)	< 0.001
GDP	1.6331	(0.9146, 2.9160)	0.097
number of health institutions	0.8328	(0.2562, 2.7073)	0.761
number of kindergartens	1.0954	(0.4324, 2.7752)	0.848
number of primary schools	0.7951	(0.3146, 2.0093)	0.628
Epidemic among cities (spatiotemporal)			
Intercept	0.0024	(0.0001, 2.5915)	–
Density	1.1721	(0.6524, 2.1058)	0.595
GDP	<b>2.6898</b>	(1.1516, 6.2822)	0.022
Passenger-kilometers of highways	0.7528	(0.4971, 1.1400)	0.180
Endemic			
Intercept	0.9965	(0.9929, 1.0001)	–
School holidays	<b>0.5299</b>	(0.3883, 0.7233)	< 0.001
number of kindergartens	2.1440	(0.6288, 7.3111)	0.223
number of primary schools	0.4996	(0.1780, 1.4021)	0.188

Statistically significant covariates are in bold

endemic, epidemic within city and epidemic among cities). Moreover, we studied the potential ecological factors that could influence the spread of scarlet fever. Our study could improve epidemiological understanding of scarlet fever, depict the epidemiology status of the scarlet fever epidemic, capture the transmission pattern of scarlet fever, and provide clues for the future prevention and control of acute infectious diseases such as scarlet fever.

In our study, males outnumbered females in all patients, according to statistics, which is consistent with previous studies [20, 39]. This can be explained by the risk of exposure, for example, men have more outdoor activities and unhygienic behaviors than women [21]. Our study indicated over 50% of cases focused on the age group 4–6, which may be due to the lack of herd immunity to scarlet fever in the children aged 3–6 [39]. To prevent outbreaks in these vulnerable groups, there is an urgent need to strengthen the education on health and the environment of sterilization. Furthermore, we discovered that the major peak of incidence occurred in the summer, with a slight surge in the winter, which is typical of most respiratory diseases [15]. This suggests personal protection is primary during the peak season.

Our results are consistent with previous studies, the overall incidence of scarlet fever rises in recent years, and the southeast area has more cases than the northwest [40]. We also found that the incidence was randomly dispersed during 2016–2019 in Sichuan Province, this may



be because the susceptible population is mainly children and teenagers, whose mobility in space are limited.

Although our results showed there were no spatial autocorrelation of scarlet fever in Sichuan in 2016–2019, we still conducted spatiotemporal analysis for three reasons: Firstly, previous simulation study demonstrated that significant covariate effect in multiple regression but not in univariate regression [41]. The multiple regression results only depends on univariate screening may be dramatically biased and misleading. Moreover, from a methodological perspective, the spatio-temporal model itself has the property of “borrowing strength” [42]. Our data contained both temporal and spatial information, and neglecting spatial information will lead to deviations. Secondly, combined with the actual situation of scarlet fever transmission in Sichuan Province, we found that transmission among cities had occurred. Finally, we compared the difference between including and not including the spatial component (Table A3 and Table A4 in additional file) in the spatiotemporal analysis. The results showed that the model with the spatial influence had a better fit with the same direction of estimates, so we considered it necessary to do spatiotemporal analysis.

As for the components of transmission of scarlet fever, endemic and autoregressive components accounted for major proportions of transmission in many areas in Sichuan province. This could be due to acute respiratory infections causing symptoms quickly, making relocation to another place unlikely [43]. Because scarlet fever is easily influenced by prior instances, it is critical to enhancing the early prediction and early warning of local cases, and it is recommended to detect and control outbreaks as early as possible to limit secondary cases. At the same time, controlling the epidemic of scarlet fever not only requires good local case management but also

entails joint efforts with adjacent areas to reduce the spread of the epidemic.

The disparity in ecological conditions will finally cause health inequality [36]. In our study, we found that ecological factors could influence the transmission of scarlet fever, such as the school holidays could restrict the spread of scarlet fever. As school vacations play a significant role in the epidemiology of various infectious diseases [35], the incidence was lower on holidays because of fewer opportunities for contagiousness. Although researchers proved GDP was associated with the transmission of infectious diseases [37], the effect has not been quantified. Our finding suggests GDP could contribute to the transmission in the spatial aspect, which is probably because GDP can boost transmission by altering other factors such as population migration [33]. We also should consider GDP as an important factor when analyzing the potential drivers of acute infectious disease.

To improve the accuracy of our model, we incorporated random effects. If regions showed heterogeneity that cannot be explained by covariates, random effects may be able to capture this unexplained element of the model [44], especially in the multiple regions with the obvious distinction. One reason is that the incidence of scarlet fever in Sichuan province varies greatly among regions. For example, Chengdu has about one-third of the cases in the whole province, while Ganzi has no more than 10 cumulative cases in the 4 years. The other reason is that the ecological factors in our study cannot fully account for the heterogeneity among regions and components. Therefore, we used the random effect to explore heterogeneity in different components and regions. For the surveillance data of infectious diseases, the heterogeneity may be caused by missed diagnosis [45] and incomplete covariates.

Compare to traditional mathematical models, the EE model can study spatiotemporal dynamics from three dimensions: endemic, epidemic within city, and epidemic among cities. And also qualifies the effect of covariates from the above aspects. It could provide more information, such as the main transmission of scarlet fever in this paper—cases are predominantly affected by local past cases rather than cases in other areas and so on. In addition, it allowed random effects to capture more realism in the model.

There are several limitations in our study. As for the surveillance data, not all infected individuals seek medical care [45]. As for the model, although adding random effects can improve the representativeness of results [46], it could lead to more complexity in the model. Further simulation studies are needed to explore relatively optimal models with appropriate numbers of random effects.

With the prevalence of new respiratory infections and the global “resurgence” of historical infectious diseases such as scarlet fever, it is critical to optimize public health prevention and control strategies in the absence of targeted vaccines. Globally, acute respiratory infections are the major reason for morbidity and mortality in children under 5 [43]. Scarlet fever epidemics have historically occurred every 5–6 years, presumably due to the establishment of herd immunity in vulnerable groups and many influence factors [47]. There is currently no solid solution to eradicate acute respiratory infections. We must pay close attention to key populations and hotspots, as well as understand the elements that contribute to the spread of ARI. At the same time, a thorough understanding of the components of the transmission can help to strategically allocate resources to high-priority geographic areas and formulate specific public health interventions for varied regions. For example, if the epidemic is affected by cases in the previous period, early prediction and early warning of local epidemic need to be strengthened to reduce secondary cases. If the endemic component plays a significant role in transmission, it is necessary to enhance patient management and environmental disinfection to reduce the follow-up impact. To control the spread of the disease, not only do we need local case management but also entail joint prevention and control with neighboring regions can reduce the spread of the epidemic.

## Conclusion

By analyzing the data of scarlet fever from 2016 to 2019 in Sichuan Province, we explored the transmission pattern of this infectious disease from the endemic, temporal and spatial dimensions. Furthermore, we established the ecological factors that could influence transmission. Our work could benefit to optimize strategies for the

prevention, detection, and management of scarlet fever and ARI.

In this paper, we took scarlet fever in Sichuan province as an example to explore the spatiotemporal transmission rules and potential ecological drivers. Firstly, We discovered that the majority of reported cases were transferred via local previous cases as well as external or indirect transmission through environmental sources. Secondly, we also found ecological factors could also influence the spread of acute respiratory infections such as scarlet fever. For example, school holidays could reduce transmission whereas GDP could boost the spread of scarlet fever. According to our findings, we should focus more on improving environmental disinfection and dealing with earlier occurrences. Because most ARIs lack viable vaccinations, public health initiatives remain critical in combating them. To maximize methods for the prevention, identification, and management of acute respiratory infectious illnesses, a focus on critical populations and hotspots is required, as well as an understanding of how ecological variables promote the spread of acute respiratory infectious diseases.

## Abbreviations

ARI: Acute respiratory infection; COVID-19: Coronavirus disease 2019; GDP: Gross Domestic Product; AIC: Akaike information criterion; RPS: Ranked probability scores; logS: logS scores; RR: Relative risk; CI: Confidence interval.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12889-022-14469-y>.

**Additional file 1: Table A1.** The meaning and value of covariates.

**Table A2.** Demographic characteristics of scarlet fever in Sichuan Province, 2016–2019. **Fig. A1.** Time series of scarlet fever counts in 21 cities (/prefectures) of Sichuan province from 2016 to 2019. **Figure A2.** The fitted values for 21 cities (/prefectures) in Sichuan Province during 2016–2019.

**Fig. A3.** The estimated multiplicative effect of seasonality on the endemic mean. **Table A3.** Model selection and comparison. **Table A4.** Model without spatial term. **Table A5.** The estimates of random effects. **Fig. A4.** Spatial connectivity weights. **Fig. A5.** The matrix of cities showing the connectivity weights.

## Acknowledgments

We are grateful to all the students who participated in data collection and processing.

## Authors' contributions

CL put forward analysis ideas, RJL collected and provided data, and CL and RJL analyzed the data and were the main contributors to write the manuscript. All authors read and approved the final manuscript, and WHZ, GYJ, YMW and LL helped to revise the manuscript. TZ and QL has made important contributions to analysis and manuscript preparation.

## Funding

This research work was funded by Sichuan Science and Technology Program (grant numbers 2022YFS0229, 2020YFS0015, 2020YFS0091, 2021YFS0001-LH), Health Commission of Sichuan province (grant number 20PJ092), National Natural Science Foundation of China (grant numbers 81602935), Chongqing Science and Technology Program (grant number cstc2020jscx-cylhX0003),

Chengdu Science and Technology Program (grant number 2021-YF05-01585-SN), Sichuan University (grant numbers 2018hhf-26), Central government funding items (grant numbers 2021zc02) and Liangshan Prefecture Center for Disease Control and Prevention (grant numbers H210322). The funders played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

#### Availability of data and materials

The data that support the findings of this study are available from Sichuan Center for Disease Control and Prevention but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Sichuan Center for Disease Control and Prevention.

#### Declarations

##### Ethics approval and consent to participate

All the data involved in this research were taken from observational records in the Nationwide Notifiable Infectious Diseases Reporting Information System (NIDRIS), which was used under license and not publicly available. And no experiments were implemented by the researchers, so informed consent were waived by the ethics committee of the Sichuan Center for Disease Control and Prevention. All methods were carried out in accordance with relevant guidelines and regulations. The research has been approved by ethics committee of the Sichuan Center for Disease Control and Prevention (SCCDCIRB-2022-185).

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare that they have no competing interests.

Received: 5 May 2022 Accepted: 17 October 2022

Published online: 21 November 2022

#### References

- Gonzalez-Bandala DA, Cuevas-Tello JC, Noyola DE, Comas-Garcia A, Garcia-Sepulveda CA. Computational forecasting methodology for acute respiratory infectious disease dynamics. *Int J Environ Res Public Health.* 2020;17(12):4540.
- Global Health Estimates 2016: deaths by cause, age, sex, by country and by region, 2000–2016. [http://www.who.int/healthinfo/global\\_burden\\_disease/estimates/en/index1.htm](http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.htm). Accessed 7 Aug 2022.
- Brouwer S, Lacey JA, You Y, Davies MR, Walker MJ. Scarlet fever changes its spots. *Lancet Infect Dis.* 2019;19(11):1154–5.
- Lekone PE, Finkenstadt BF. Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. *Biometrics.* 2006;62(4):1170–7.
- Gold MS, Sehayek D, Gabrielli S, Zhang X, McCusker C, Ben-Shoshan M. COVID-19 and comorbidities: a systematic review and meta-analysis. *Postgrad Med.* 2020.
- WHO Coronavirus (COVID-19) Dashboard. <https://covid19.who.int/>. Accessed 1 Jul 2022.
- Drug, Therapeutics B. Managing scarlet fever. *BMJ.* 2018;362:k3005.
- Scarlet fever is reemerging as a public health threat. <https://www.consumeraffairs.com/news/scarlet-fever-is-reemerging-as-a-public-health-threat-100720.html>. (Accessed 15 Nov 2022).
- Jacobs J. Homeopathic prevention and Management of Epidemic Diseases. *Homeopathy.* 2018;107(3):157–60.
- Held L, Hofmann M, Hohle M, Schmid V. A two-component model for counts of infectious diseases. *Biostatistics.* 2006;7(3):422–37.
- Riley S. Large-scale spatial-transmission models of infectious disease. *Science.* 2007;316:1298–301.
- O'Neill PD. Introduction and snapshot review: relating infectious disease transmission models to data. *Stat Med.* 2010;29(20):2069–77.
- Held L, Höhle M, Hofmann MW. A statistical framework for the analysis of multivariate infectious disease surveillance counts. *Stat Model.* 2005;5:187–99.
- Zuo Z, Wang M, Cui H, Wang Y, Wu J, Qi J, et al. Spatiotemporal characteristics and the epidemiology of tuberculosis in China from 2004 to 2017 by the nationwide surveillance system. *BMC Public Health.* 2020;20(1):1284.
- Chiavenna C, Presanis AM, Charlett A, de Lusignan S, Ladhani S, Pebody RG, et al. Estimating age-stratified influenza-associated invasive pneumococcal disease in England: a time-series model based on population surveillance data. *PLoS Med.* 2019;16(6):e1002829.
- Cheng Q, Lu X, Wu JT, Liu Z, Huang J. Analysis of heterogeneous dengue transmission in Guangdong in 2014 with multivariate time series model. *Sci Rep.* 2016;6:33755.
- Liu Y, Ding H, Chang ST, Lu R, Zhong H, Zhao N, et al. Exposure to air pollution and scarlet fever resurgence in China: a six-year surveillance study. *Nat Commun.* 2020;11(1):4229.
- Rao HX, Li DM, Zhao XY, Yu J. Spatiotemporal clustering and meteorological factors affected scarlet fever incidence in mainland China from 2004 to 2017. *Sci Total Environ.* 2021;777:146145.
- Zhang Q, Liu W, Ma W, Shi Y, Wu Y, Li Y, et al. Spatiotemporal epidemiology of scarlet fever in Jiangsu Province, China, 2005–2015. *BMC Infect Dis.* 2017;17(1):596.
- Chen H, Chen Y, Sun B, Wen L, An X. Epidemiological study of scarlet fever in Shenyang, China. *BMC Infect Dis.* 2019;19(1):1074.
- Buckee C, Noor A, Sattenspiel L. Thinking clearly about social aspects of infectious disease transmission. *Nature.* 2021;595(7866):205–13.
- Jiang F, Wei T, Hu X, Han Y, Jia J, Pan B, et al. The association between ambient air pollution and scarlet fever in Qingdao, China, 2014–2018: a quantitative analysis. *BMC Infect Dis.* 2021;21(1):987.
- Law of the People's Republic of China on the Prevention and Treatment of Infectious Diseases. [http://www.npc.gov.cn/zgrdw/npc/ztxw/crbfz/2003-04/01/content\\_1804055.htm](http://www.npc.gov.cn/zgrdw/npc/ztxw/crbfz/2003-04/01/content_1804055.htm). Accessed 1 Jul 2022.
- Banerjee S. Spatial Data Analysis. *Annu Rev Public Health.* 2016;37(1):47–60.
- Anselin L. Local indicators of spatial association—LISA. *Geogr Anal.* 1995;27(2):93–115.
- Paul M, Held L, Toschke AM. Multivariate modelling of infectious disease surveillance data. *Stat Med.* 2008;27(29):6250–67.
- Geilhufe M, Held L, Skrovseth SO, Simonsen GS, Godtliebsen F. Power law approximations of movement network data for modeling infectious disease spread. *Biom J.* 2014;56(3):363–82.
- Held L, Paul M. Modeling seasonality in space-time infectious disease surveillance data. *Biom J.* 2012;54(6):824–43.
- Meyer S, Held L. Power-law models for infectious disease spread. *Ann Appl Stat.* 2014;8(3):1612–39.
- Ondrikova N, Clough HE, Cunliffe NA, Iturriza-Gomara M, Vivancos R, Harris JP. Understanding norovirus reporting patterns in England: a mixed model approach. *BMC Public Health.* 2021;21(1):1245.
- Held L, Meyer S, Bracher J. Probabilistic forecasting in infectious disease epidemiology: the 13th Armitage lecture. *Stat Med.* 2017;36(22):3443–60.
- Meyer S, Held L, Höhle M. Spatio-temporal analysis of epidemic phenomena using the R package surveillance. *J Stat Softw.* 2017;77(11):1–55.
- Beltran RM, Holloway IW, Hong C, Miyashita A, Cordero L, Wu E, et al. Social determinants of disease: HIV and COVID-19 experiences. *Curr HIV/AIDS Rep.* 2022;19(1):101–12.
- Santos-Vega M, Martinez PP, Pascual M. Climate forcing and infectious disease transmission in urban landscapes: integrating demographic and socioeconomic heterogeneity. *Ann NY Acad Sci.* 2016;1382(1):44–55.
- Eames KT, Tilston NL, Edmunds WJ. The impact of school holidays on the social mixing patterns of school children. *Epidemics.* 2011;3(2):103–8.
- Zhou M, Guo W. Social factors and worry associated with COVID-19: evidence from a large survey in China. *Soc Sci Med.* 2021;277:113934.
- Zhang T, Yin F, Zhou T, Zhang XY, Li XS. Multivariate time series analysis on the dynamic relationship between class B notifiable diseases and gross domestic product (GDP) in China. *Sci Rep.* 2016;6:29.
- Ssentongo P, Fronterre C, Geronimo A, Greybush SJ, Mbabazi PK, Muwawala J, et al. Pan-African evolution of within- and between-country COVID-19 dynamics. *Proc Natl Acad Sci U S A.* 2021;118(28):e2026664118.

39. Liu Y, Chan T-C, Yap L-W, Luo Y, Xu W, Qin S, et al. Resurgence of scarlet fever in China: a 13-year population-based surveillance study. Lancet Infect Dis. 2018;18(8):903–12.
40. Liao RJ, Liu RY, Yang CH. Epidemiological analysis of scarlet fever in Sichuan Province from 2009 to 2018. Infect Dis Inform. 2019;32(04):322–324+328.
41. Feng G, Peng J, Tu D, Zheng JZ, Feng C. Two paradoxes in linear regression analysis. Shanghai Arch Psychiatry. 2016;28(6):355–60.
42. Zhang T, Zhang X, Ma Y, Zhou XA, Liu Y, Feng Z, et al. Bayesian spatio-temporal random coefficient time series (BaST-RCTS) model of infectious disease. Math Biosci. 2014;258:93–100.
43. Shi T, McAllister DA, O'Brien KL, Simoes EAF, Madhi SA, Gessner BD, et al. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. Lancet. 2017;390(10098):946–58.
44. Herzog SA, Paul M, Held L. Heterogeneity in vaccination coverage explains the size and occurrence of measles epidemics in German surveillance data. Epidemiol Infect. 2011;139(4):505–15.
45. Bracher J, Held L. A marginal moment matching approach for fitting endemic-epidemic models to underreported disease surveillance counts. Biometrics. 2021;77(4):1202–14.
46. Data Analysis in R. [https://bookdown.org/steve\\_midway/DAR/infographics.html](https://bookdown.org/steve_midway/DAR/infographics.html). Accessed 1 Aug 2022.
47. You Y, Davies MR, Protani M, McIntyre L, Walker MJ, Zhang J. Scarlet fever epidemic in China caused by streptococcus pyogenes serotype M12: epidemiologic and molecular analysis. EBioMedicine. 2018;28:128–35.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](http://biomedcentral.com/submissions)

