

Received: 2019.02.17

Accepted: 2019.03.21

Published: 2019.07.30

# Trends in the Epidemiology of Sexually Transmitted Disease, Acquired Immune Deficiency Syndrome (AIDS), Gonorrhea, and Syphilis, in the 31 Provinces of Mainland China

Authors' Contribution:

Study Design A

Data Collection B

Statistical Analysis C

Data Interpretation D

Manuscript Preparation E

Literature Search F

Funds Collection G

CDEF 1 **Xuechen Ye**

BCDF 2 **Jie Liu**

AE 3 **Zhe Yi**

1 Department of Social Medicine, School of Public Health, China Medical University, Shenyang, Liaoning, P.R. China

2 Department of Health Statistics, School of Public Health, China Medical University, Shenyang, Liaoning, P.R. China

3 Department of Prosthodontics, School of Stomatology, China Medical University, Shenyang, Liaoning, P.R. China

**Corresponding Author:** Zhe Yi, e-mail: zheyi\_2000@sina.com

**Source of support:** This study was funded by the National Natural Science Foundation of China (Grant numbers 71673301, 71473269, and 81273186)

**Background:** This study aimed to investigate trends in the epidemiology of the leading sexually transmitted diseases (STDs), acquired immune deficiency syndrome (AIDS), gonorrhea, and syphilis, in the 31 provinces of mainland China.

**Material/Methods:** This retrospective study analyzed the incidence data of STDs from official reports in China between 2004 and 2016. The grey model first order one variable, or GM (1,1), time series forecasting model for epidemiological studies predicted the incidence of STDs based on the annual incidence reports from 31 Chinese mainland provinces. Hierarchical cluster analysis was used to group the prevalence of STDs within each province.

**Results:** The prediction accuracy of the GM (1,1) model was high, based on data during the 13 years between 2004 and 2016. The model predicted that the incidence rates of AIDS and syphilis would continue to increase over the next two years. Cluster analysis showed that 31 provinces could be classified into four clusters according to similarities in the incidence of STDs. Group A (Sinkiang Province) had the highest reported prevalence of syphilis. Group B included provinces with a higher incidence of gonorrhea, mainly in the southeast coast of China. Group C consisted of southwest provinces with a higher incidence of AIDS.

**Conclusions:** The GM (1,1) model was predictive for the incidence of STDs in 31 provinces in China. The predicted incidence rates of AIDS and syphilis showed an upward trend. Regional distribution of the major STDs highlights the need for targeted prevention and control programs.

**MeSH Keywords:** **Cluster Analysis • Forecasting • Incidence • Sexually Transmitted Diseases**

**Full-text PDF:** <https://www.medscimonit.com/abstract/index/idArt/915732>

 2695

 5

 4

 30



## Background

Acquired immune deficiency syndrome (AIDS) due to human immune deficiency virus (HIV) infection, syphilis, and gonorrhea are the three major sexually transmitted diseases (STDs) in China, and are legally notifiable infectious diseases according to the Law of the Peoples' Republic of China on Prevention and Control of Infectious Diseases [1]. Clinically diagnosed STDs, or those diagnosed by laboratory investigations, have a mandatory requirement to be reported to national surveillance systems in China [1]. Surveillance data from 2004 to 2013 in China showed that HIV infection and syphilis had the most rapid increase in incidence, with an annual increase in prevalence of 16.3% (95% CI, 11.5–21.2), and 16.3% (95% CI, 13.8–18.8) respectively, and gonorrhea had the most rapid rates of decline with an annual change of –8.5% (95% CI, –11.7 to –5.1) [2]. The increased incidence of HIV and syphilis was attributed to a variety of reasons, including internal labor migration, unsafe sexual practices, and lack of consulting healthcare services [3–5].

Monitoring and evaluating the increasing incidence of STDs, particularly the ability to predict the prevalence of STDs, is essential to control infectious disease [1]. Several methods have been used to detect and predict the prevalence of infectious disease over time, including the autoregressive integrated moving average (ARIMA) model, the back propagation neural network (BPNN) model, and the grey model (GM) [6,7]. The ARIMA model assumes a linear relationship between dependent and independent variables and a constant standard deviation in a time series [8]. However, the actual surveillance data often presents complex non-linear relationships and non-stationary distributions over time [8,9]. The BPNN model requires sufficient training data and longer training times to obtain an accurate and generalized model and may be associated with a predisposition to falling into local minima in the training process [6]. Difficulties in the collection of detailed information and complex mathematical operations also limit the application of the BPNN model.

Compared with the ARIMA and BPNN methods, the grey model requires a smaller sample size, less stringent data distribution, and it is easier to analyze [10]. Deng first proposed the grey system theory in 1982 through the analysis of systems with partially known, incomplete, or uncertain information [11]. The grey model can be built to aid predictions and decision-making in the face of limited information and knowledge [11]. Recently, the grey model has been applied to the analysis of data involving the environment, energy, and finance, with satisfactory reported results [12]. The grey model can also predict the epidemiology of infectious diseases, including typhoid fever and echinococcosis [7,10]. The grey model can predict the incidence of infectious disease and facilitate intervention

strategies and was chosen for this study to predict the prevalence of STDs in China.

To identify and monitor the incidence of STDs an understanding of the regional distribution of disease incidence is required [13]. In China, the occurrence of STDs has significant disparity across regions, with some infections reaching epidemic levels in certain parts of the country [14,15]. The regional epidemiology of disease must be considered when healthcare services, including administration and allocation of resources to implement prevention and control of STDs.

Comprehensive STD control programs are urgently required to reduce the spread of STDs in China. Therefore, this study aimed to investigate trends in the epidemiology of the leading sexually transmitted diseases (STDs), acquired immune deficiency syndrome (AIDS), gonorrhea, and syphilis, in the 31 provinces of mainland China between 2004 and 2016.

## Material and Methods

### Study design and data access

This retrospective study required no ethical approval as anonymized retrospective data were collected and analyzed from existing resources. The incidence data of legally notifiable infectious diseases in mainland China were publicly accessible from the official website of the National Health and Family Planning Commission of the Peoples' Republic of China [16].

### Data collection

The incidence of acquired immune deficiency syndrome (AIDS) due to human immune deficiency virus (HIV) infection, syphilis, and gonorrhea in mainland China from 2004 to 2016 were obtained from official reports released by the National Health and Family Planning Commission of the People's Republic of China [16]. The reported incidence data of 31 provinces, or autonomous regions and municipalities, were also collected.

### Data analysis using the grey prediction model, GM (1,1)

The grey model was designed to make predictions from partially known or limited data [12]. The grey model first order one variable, or GM (1,1), time series forecasting model for epidemiological studies is a more specific and widely used grey prediction model [7]. In this study, the reported incidence rates of AIDS, syphilis, and gonorrhea from 2004 to 2016 were collected to establish the GM (1,1) model for incidence prediction. The following steps explain how the model was developed.

The original data sequence was presented as Eq (1):

$$x^{(0)} = \{x^{(0)}(0), x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\} \quad (1)$$

where,

$$x^{(0)}(i) \geq 0, I = 0, 1, 2, \dots, n.$$

The accumulated generating operation (AGO) was then used to generate the first order accumulative data sequence of  $\chi^{(0)}$ :

$$x^{(1)} = \{x^{(1)}(0), x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\} \quad (2)$$

where,

$$x^{(1)}(k) = \sum_{i=0}^k x^{(0)}(i), k = 0, 1, 2, \dots, n.$$

The adjacent neighbor mean sequence of  $\chi^{(1)}$  was expressed as:

$$z^{(1)} = \{z^{(1)}(1), z^{(1)}(2), \dots, z^{(1)}(n)\} \quad (3)$$

where,

$$z^{(1)}(k) = 0.5 \times [x^{(1)}(k) + x^{(1)}(k-1)], k = 1, 2, \dots, n.$$

Next, the basic grey differential equation of GM (1,1) was established as:

$$x^{(0)}(k) + a \times z^{(1)}(k) = b \quad (4)$$

The whitening differential equation of GM (1,1) represented the first-order linear differential equation based on  $\chi^{(1)}(k)$ , given as:

$$dx^{(1)}/dt + ax^{(1)} = b \quad (5)$$

where,  $a$  and  $b$  were defined as the model parameters and estimated using the ordinary least square (OLS) method. The estimation of  $a$  and  $b$  using the OLS method was as follows:

$$\hat{u} = \begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y \quad (6)$$

where,

$$Y = \begin{bmatrix} x^{(0)}(1) \\ x^{(0)}(2) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, B = \begin{bmatrix} -z^{(1)}(1) & 1 \\ -z^{(1)}(2) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \quad (7)$$

With the obtained parameters  $a$  and  $b$  substituted into Eq (5), the time response function of GM (1,1) was denoted as follows:

$$\hat{x}^{(1)}(k) = [x^{(0)}(0) - \frac{b}{a}] e^{-ak} + \frac{b}{a}, k = 0, 1, 2, \dots, n \quad (8)$$

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \quad (9)$$

where,

$\hat{\chi}^{(0)}(k+1)$  represents the predicted value.

After establishing the GM (1,1), its accuracy was examined to extrapolate the predicted values.

During the error analysis of the grey system, the late error detection method was used to quantify the prediction error and to assess its accuracy. This method used two important indices, the posterior error ratio, termed C, and the small error probability termed P. Between the actual data and the predicted values, the residual error was denoted as:

$$\varepsilon^{(0)}(k) = x^{(0)}(k) - \hat{x}^{(0)}(k), k = 1, 2, 3, \dots, n \quad (10)$$

where, C was the ratio between the standard deviation of the residual error ( $S_e$ ) and the standard deviation of original data ( $S_x$ ). The specific equation was given as:

$$C = S_e / S_x \quad (11)$$

where,

$$S_e = \sqrt{\frac{\sum (\varepsilon^{(0)}(k) - \bar{\varepsilon})^2}{n}}, S_x = \sqrt{\frac{\sum (x^{(0)}(k) - \bar{x})^2}{n}} \quad (12)$$

P was calculated as follows:

$$P = P \{ | \varepsilon^{(0)}(k) - \bar{\varepsilon} | < 0.6745 S_x \} = m/n \quad (13)$$

where,  $m$  was the total number of cases that  $| \varepsilon^{(0)}(k) - \bar{\varepsilon} | < 0.6745 S_x$ .

A lower C-value and a higher P-value indicated a better forecasting result. The combination of P and C defined four levels of forecasting accuracy (Table 1). Statistical analysis was performed using Predictive Analytics SoftWare (PASW) Statistics version 18.0 for Windows.

### Hierarchical cluster analysis

Hierarchical cluster analysis was used to classify the variables according to differences or similarities [17]. In this study,

**Table 1.** Four levels of forecasting accuracy of the grey model.

Level	C-value	P-value
Good (G)	C ≤ 0.35	P ≥ 0.95
Acceptable (A)	0.35 < C ≤ 0.50	0.80 ≤ P < 0.95
Pass (P)	0.50 < C ≤ 0.65	0.70 ≤ P < 0.80
Unacceptable(U)	C > 0.65	P < 0.70

**Table 2.** Actual and predicted values of the incidence of reported cases of acquired immune deficiency syndrome (AIDS) in China using the GM (1,1) model from 2004 to 2018 (1/100,000).

Year	k	$\chi^{(0)}(k)$	$\hat{\chi}^{(0)}(k)$	$\varepsilon^{(0)}(k)$	$\varepsilon^{(0)}(k)-\bar{\varepsilon}$
2004	0	0.23	–	–	–
2005	1	0.43	0.66	–0.23	0.01
2006	2	0.51	0.79	–0.28	–0.05
2007	3	0.74	0.95	–0.21	0.02
2008	4	0.76	1.14	–0.38	–0.15
2009	5	1	1.37	–0.37	–0.14
2010	6	1.2	1.64	–0.44	–0.21
2011	7	1.53	1.97	–0.44	–0.21
2012	8	3.11	2.37	0.74	0.97
2013	9	3.12	2.85	0.27	0.51
2014	10	3.33	3.42	–0.09	0.15
2015	11	3.69	4.11	–0.42	–0.18
2016	12	3.97	4.93	–0.96	–0.73
2017	13	–	5.92	–	–
2018	14	–	7.11	–	–

$\chi^{(0)}(k)$  – actual data of reported incidence of acquired immune deficiency syndrome (AIDS) at year k;  $\hat{\chi}^{(0)}(k)$  – predicted value of reported AIDS incidence at year k;  $\varepsilon^{(0)}(k)$  – residual error between  $\chi^{(0)}(k)$  and  $\hat{\chi}^{(0)}(k)$ ;  $\varepsilon^{(0)}(k)-\bar{\varepsilon}$  – deviation from the average of residual error.

the reported incidence rates of AIDS, syphilis, and gonorrhea across the 31 Chinese provinces were introduced into a hierarchical cluster analysis. Given the significant differences between the incidence rates of the three STDs, all data were first Z-score standardized. The between-group linkage method was used for the clustering algorithm, and the squared Euclidean distance was selected to measure the similarity coefficient. In the analysis dendrogram, closely connected provinces were identified to be more similar than those further apart. Provinces in the same cluster had similar STD epidemiological characteristics and provinces between clusters were more variable.

## Results

### The predicted incidence of acquired immune deficiency syndrome (AIDS) in China

The reported incidence of acquired immune deficiency syndrome (AIDS) in China from 2004 to 2016 is shown in Table 2 and Figure 1. The incidence of AIDS continually increased between 2004 to 2016. The grey model first order one variable, or GM (1,1), time series forecasting model for epidemiological

studies was used to predict the reported incidence of AIDS by the following equation:

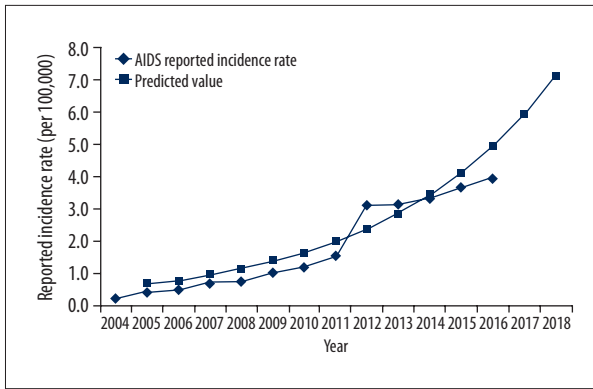
$$\hat{\chi}^{(0)}(k+1)=0.66e^{0.183k}$$

The C-value of the GM (1,1) calculated over the 13-year period was 0.30, and the grey forecasting accuracy was good. The  $S_{\chi}$  value was 1.34, and the number of cases that  $|\varepsilon^{(0)}(k)-\bar{\varepsilon}| < 0.6745S_{\chi}$  was 11. The P-value of GM (1,1) was 92%, and the grey forecasting accuracy was acceptable. The predicted incidence obtained using the GM (1,1) model closely agreed with the actual data, suggesting the model fitted well. The results predicted that the incidence of AIDS maintained an upward trend in China, and the annual reported incidence rates from 2017 and 2018 were 5.92/100,000 and 7.11/100,000, respectively.

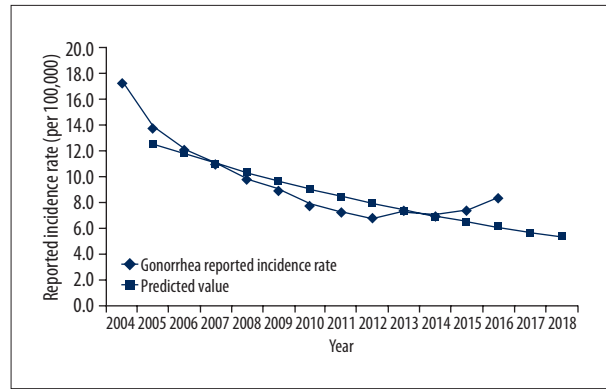
### Predicted incidence of gonorrhea in China

The reported incidence of gonorrhea in China from 2004 to 2016 is shown in Table 3 and Figure 2. The forecast equation of GM (1,1) was applied for data analysis.

$$\hat{\chi}^{(0)}(k+1)=12.52e^{-0.065k}$$



**Figure 1.** Actual and predicted reported incidence of acquired immune deficiency syndrome (AIDS) in China using the GM (1,1) model, from 2004 to 2018.



**Figure 2.** Actual and predicted reported incidence of gonorrhea in China using the GM (1,1) model, from 2004 to 2018.

**Table 3.** Actual and predicted values of the incidence of reported cases of gonorrhea in China using the GM (1,1) model from 2004 to 2018 (1/100,000).

Year	$k$	$\chi^{(0)}(k)$	$\hat{\chi}^{(0)}(k)$	$\varepsilon^{(0)}(k)$	$\varepsilon^{(0)}(k) - \bar{\varepsilon}$
2004	0	17.34	—	—	—
2005	1	13.79	12.52	1.27	1.24
2006	2	12.14	11.73	0.41	0.38
2007	3	11.08	10.99	0.09	0.05
2008	4	9.90	10.30	-0.40	-0.43
2009	5	9.02	9.65	-0.63	-0.67
2010	6	7.91	9.04	-1.13	-1.17
2011	7	7.31	8.47	-1.16	-1.20
2012	8	6.82	7.94	-1.12	-1.16
2013	9	7.36	7.44	-0.08	-0.12
2014	10	7.05	6.97	0.08	0.04
2015	11	7.36	6.53	0.83	0.79
2016	12	8.39	6.12	2.27	2.23
2017	13	—	5.74	—	—
2018	14	—	5.38	—	—

$\chi^{(0)}(k)$  – actual data of reported incidence of gonorrhea at year  $k$ ;  $\hat{\chi}^{(0)}(k)$  – predicted value of reported incidence of gonorrhea at year  $k$ ;  $\varepsilon^{(0)}(k)$  – residual error between  $\chi^{(0)}(k)$  and  $\hat{\chi}^{(0)}(k)$ ;  $\varepsilon^{(0)}(k) - \bar{\varepsilon}$  – deviation from the average of residual error.

The C-value of the GM (1,1) model was 0.33 and the grey forecasting accuracy was good. The  $S_{\chi}$  value was 3.04, and the number of cases that  $|\varepsilon^{(0)}(k) - \bar{\varepsilon}| < 0.6745S_{\chi}$  was 11. Therefore, the P-value of GM (1,1) was 92%, and the grey forecasting accuracy was acceptable. The results predicted that the incidence of gonorrhea followed a decreasing trend in China, and the annual reported incidence rates from 2017 to 2018 were 5.74/100,000 and 5.38/100,000, respectively.

### Predicted incidence of syphilis in China

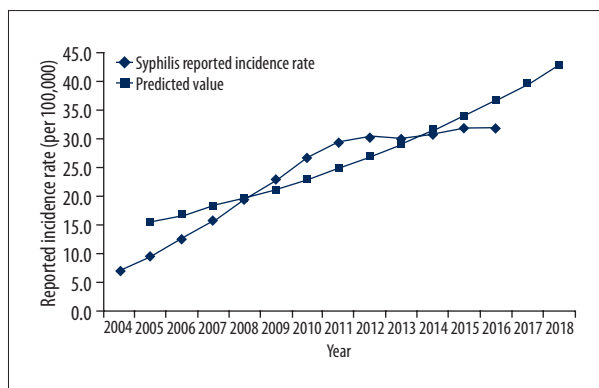
The reported incidence of syphilis in China from 2004 to 2016 is shown in Table 4 and Figure 3. The forecast equation of GM (1,1) was applied to analyze the data.

$$\hat{\chi}^{(0)}(k+1) = 15.56e^{0.078k}$$

**Table 4.** Actual and predicted values of the incidence of reported cases of syphilis in China using the GM (1,1) model from 2004 to 2018 (1/100,000).

Year	$k$	$\chi^{(0)}(k)$	$\hat{\chi}^{(0)}(k)$	$\varepsilon^{(0)}(k)$	$\varepsilon^{(0)}(k)-\bar{\varepsilon}$
2004	0	7.12	–	–	–
2005	1	9.67	15.56	–5.89	–5.49
2006	2	12.80	16.83	–4.03	–3.62
2007	3	15.88	18.19	–2.31	–1.91
2008	4	19.49	19.67	–0.18	0.23
2009	5	23.07	21.26	1.81	2.21
2010	6	26.86	22.99	3.87	4.28
2011	7	29.47	24.85	4.62	5.02
2012	8	30.44	26.87	3.57	3.98
2013	9	30.04	29.05	0.99	1.40
2014	10	30.93	31.41	–0.48	–0.07
2015	11	31.85	33.95	–2.10	–1.70
2016	12	31.97	36.71	–4.74	–4.33
2017	13	–	39.69	–	–
2018	14	–	42.91	–	–

$\chi^{(0)}(k)$  – actual data of reported incidence of syphilis at year  $k$ ;  $\hat{\chi}^{(0)}(k)$  – predicted value of reported incidence of syphilis at year  $k$ ;  $\varepsilon^{(0)}(k)$  – residual error between  $\chi^{(0)}(k)$  and  $\hat{\chi}^{(0)}(k)$ ;  $\varepsilon^{(0)}(k)-\bar{\varepsilon}$  – deviation from the average of residual error.



**Figure 3.** Actual and predicted reported incidence of syphilis in China using the GM (1,1) model, from 2004 to 2018.

The C-value of the GM (1,1) model was 0.39, and the grey forecasting accuracy was acceptable. The  $S_{\chi}$  value was 8.68, and the number of cases that  $|\varepsilon^{(0)}(k)-\bar{\varepsilon}| < 0.6745S_{\chi}$  was 12. Therefore, the P-value of the GM (1,1) model was 100%, indicating good forecasting accuracy. The results predicted that the incidence of syphilis would increase in China, and the annual reported incidence rates from 2017 to 2018 were 39.69/100,000 and 42.91/100,000, respectively.

**The reported incidence of the three STDs in China in 2016**

The reported incidence of AIDS, syphilis, and gonorrhoea across the 31 provinces in 2016 are shown in Table 5. Guangxi, Yunnan, Sichuan, and Chongqing reported a higher incidence of AIDS. Areas with a high incidence of gonorrhoea included Zhejiang, Shanghai, Hainan, Guangdong, and Fujian. Sinkiang, Zhejiang, Fujian, and Shanghai had a higher reported incidence of syphilis than other provinces.

**Cluster analysis of the three STDs in China in 2016**

The dendrogram (Figure 4) shows the cluster analysis results of the incidence of STDs across the provinces of China in 2016. The 31 provinces were classified into four groups according to the degree of similarity in reported incidence. In combination with the results of Table 5, the classification results showed that group A (Sinkiang) had the highest reported incidence of syphilis. Group B (Fujian, Hainan, Guangdong, Zhejiang, and Shanghai) had the highest incidence of gonorrhoea. Group C (Chongqing, Guizhou, Guangxi, Yunnan, and Sichuan) had the highest incidence of AIDS.

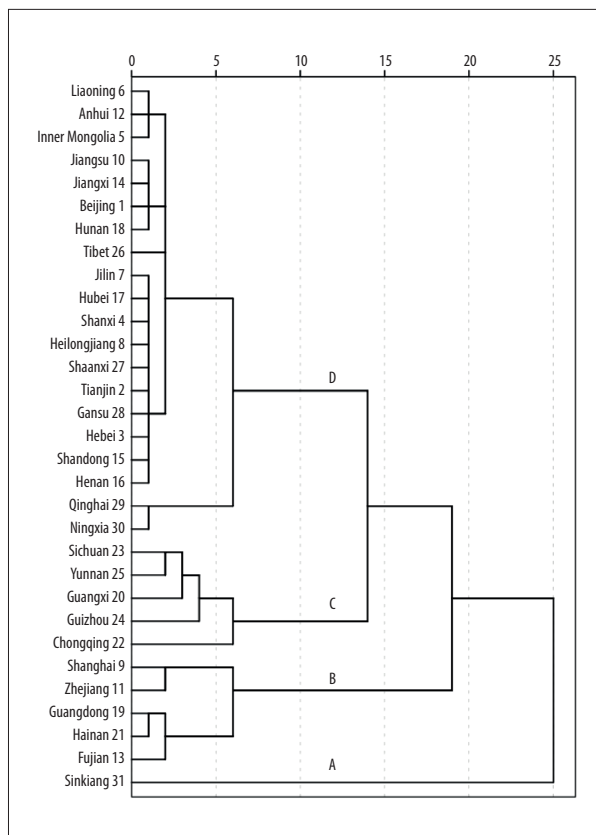
**Table 5.** The reported incidence of acquired immune deficiency syndrome (AIDS), gonorrhea, and syphilis in China in 2016 (1/100,000).

Region	AIDS	Gonorrhea	Syphilis
Beijing	3.62	6.58	22.92
Tianjin	1.79	2.06	16.6
Hebei	0.79	1.91	12.88
Shanxi	1.45	3.57	25.3
Inner Mongolia	1.14	8.91	38.12
Liaoning	2.31	5.85	37.41
Jilin	1.99	4.97	18.36
Heilongjiang	1.48	4.01	24.3
Shanghai	2.27	28.03	57.07
Jiangsu	2.02	9.03	29.7
Zhejiang	3.37	32.75	62.12
Anhui	2.17	6.05	37.14
Fujian	2.43	15.42	58.51
Jiangxi	3.01	8	28.7
Shandong	0.74	4.15	15.54
Henan	3.17	3.13	15.44
Hubei	2.18	4.74	21.67
Hunan	4.15	4.82	33.43
Guangdong	3.8	19.8	48.73
Guangxi	12.48	10.23	15.41
Hainan	1.92	20.36	50.69
Chongqing	10.2	6.79	54.86
Sichuan	11.16	3.56	28.35
Guizhou	7.42	6.41	30.37
Yunnan	12.04	8.64	35.96
Tibet	1.57	1.39	35.25
Shaanxi	2.13	4.53	26.09
Gansu	1.55	3.24	18.61
Qinghai	2.97	2.52	48.47
Ningxia	1.26	4.72	56.7
Sinkiang	8.14	7.29	89.05

## Discussion

This study aimed to predict the incidence of acquired immune deficiency syndrome (AIDS) due to human immune deficiency virus (HIV) infection, syphilis, and gonorrhea using data reported from 2004 to 2016, and to classify the regional distribution of these sexually transmitted diseases (STDs) in China. The incidence rates of AIDS and syphilis may increase for a further two years, indicating that AIDS had reached epidemic levels in China. The factors associated with the increasing trend include engagement in unsafe sexual practices, especially among men

who have sex with men (MSM), and female sex workers (FSWs) and their clients [18,19]. Of the newly reported annual cases of HIV/AIDS, cases transmitted by homosexual sex increased from 2.5% in 2006 to 27.6% in 2016 [20,21]. Previous studies have also shown that that rural-to-urban migration, social stigma, and lack of healthcare-seeking behavior expand the spread of HIV/AIDS and syphilis [3–5,18]. Also, HIV and syphilis can be co-transmitted, which potentially further drives the growth of the epidemic [22]. There is an urgent need for the promotion of interventions such as condom use, voluntary counseling and testing, and health education regarding safe sex and STDs to



**Figure 4.** Hierarchical cluster analysis dendrogram of the reported incidence of the three sexually transmitted diseases (STDs), acquired immune deficiency syndrome (AIDS), gonorrhea, and syphilis in the 31 provinces of mainland China in 2016.

control HIV/AIDS and syphilis [18,23]. Since syphilis and HIV share similar modes of transmission and similar risk behaviors, their control interventions should be combined [22].

Cluster analysis classified the provinces into 4 groups based on the degree of similarities of the characteristics of the incidence of STDs. The provinces in each group were geographically adjacent. Group C included Chongqing, Guizhou, Yunnan, and Sichuan in southwest China that reported the highest incidence of AIDS. Provinces in the Southwest border of China are adjacent to the Golden Triangle Drug Trade in Southeast Asia. Yunnan is renowned for drug trafficking and drug consumption, which may explain the high rates of HIV/AIDS [24]. HIV infection in China among injecting drug users was identified in Yunnan in 1989. The transmission rates of HIV among injecting drug users residing near major routes of drug-trafficking in Guangxi, Sichuan and other provinces increased rapidly [25]. It has previously been suggested that injecting synthetic drugs, such as ketamine and ecstasy, increases high-risk sexual practices, and facilitated the spread of HIV to spread from intravenous drug users to their sexual partners [24,26].

Based on this information, the government should strengthen anti-drug law enforcement to control drug use as an effective response to the HIV/AIDS epidemic in Southwest areas [27].

Group B included provinces with a high reported incidence of gonorrhea, which were mainly located in the Southeast coastline of China that has a developed economy and mobile population. Previous studies identified high-incidence clusters of gonorrhea that were primarily distributed in the eastern Yangtze River Delta (Zhejiang and Shanghai) and southern Pearl River Delta (Guangdong), which are the most developed coastal areas of China [28]. This regional pattern was associated with mass rural-to-urban migration [14]. A previously published systematic review showed that the prevalence of gonorrhea amongst migrants was 13.6 greater (range, 5.8–32.1) than the odds of infection amongst the general Chinese population [4]. Several migrants who were unmarried or away from their spouses have flocked to wealthy cities for better employment, with increased rates of unsafe sexual behaviors including unprotected sex and multiple sexual partners [18,28]. Furthermore, the high rates of migration have facilitated the spread of gonorrhea across prosperous regions in China. Therefore, gonorrhea control measures including health education, the promotion of condom use, and early screening and treatment should be increased in the Southeast provinces and targeted to key groups including migrant workers and female sex workers [14,28].

Group A included the Sinkiang province, which had the highest reported incidence of syphilis. This province is located in northwest China and borders many Asian countries. Regional or spatial analysis studies have also indicated that some regions at high risk of syphilis were present in less developed areas and inland areas, including Sinkiang [15]. Approximately half of the syphilis cases in Sinkiang international ports originated from ethnic minorities, particularly the Uyghurs, accounting for 37.88% of the population [29]. In remote areas near the Chinese border, the spread of syphilis may be associated with different cultures and customs, inadequate sexual knowledge, and high-risk behaviors among a multi-ethnic population [3,29]. Therefore, targeted interventions should consider ethnic minorities, and should include bilingual health education promotion pamphlets both in the Chinese Han language and in ethnic minority languages [29,30]. Also, the quality and allocation of medical resources should be optimized in distant inland areas to reduce the incidence of syphilis effectively [15].

This study had several limitations. In contrast to nationwide time trend analysis for 13 years, and for provincial analysis, this study analyzed only incidence information over one year at the provincial level. In future studies, time trends and the geographic distribution of incidence at the sub-provincial level should be assessed to identify the diversity of the epidemiology of STDs in China in more detail.



## Conclusions

This study showed that the GM (1,1) model was an effective method to predict the incidence of sexually transmitted diseases (STDs). The incidence rates of acquired immune deficiency syndrome (AIDS) and syphilis in China are predicted to continue to rise, and the rates of STD were unequally distributed across mainland China. Therefore, control and prevention programs should consider the geographic distribution of STDs. Specifically, intervention programs must acknowledge the epidemiology and increasing incidence of AIDS in the Southwest

provinces, gonorrhoea in the Southeast coastline provinces, and syphilis in Sinkiang province of northwest China.

## Acknowledgments

We gratefully acknowledge Haiqiang Guo for the practical advice on analysis of the data.

## Conflict of interest

None.

## References:

1. Vlieg WL, Fanoy EB, van Asten L et al: Comparing national infectious disease surveillance systems: China and the Netherlands. *BMC Public Health*, 2017; 17: 415
2. Yang SG, Wu J, Ding C et al: Epidemiological features of and changes in incidence of infectious diseases in China in the first decade after the SARS outbreak: an observational trend study. *Lancet Infect Dis*, 2017; 17: 716–25
3. Lin CC, Gao X, Chen XS, Chen Q, Cohen MS: China's syphilis epidemic: A systematic review of seroprevalence studies. *Sex Transm Dis*, 2006; 33: 726–36
4. Zou X, Chow EP, Zhao PC et al: Rural-to-urban migrants are at high risk of sexually transmitted and viral hepatitis infections in China: A systematic review and meta-analysis. *BMC Infect Dis*, 2014; 14: 490
5. Xu JJ, Yu YQ, Hu QH et al: Treatment-seeking behaviour and barriers to service access for sexually transmitted diseases among men who have sex with men in China: A multicentre cross-sectional survey. *Infect Dis Poverty*, 2017; 6: 15
6. Zhang XY, Liu YY, Yang M et al: Comparative study of four time series methods in forecasting typhoid fever incidence in China. *PLoS One*, 2013; 8: e63116
7. Yang XB, Zou JJ, Kong DG, Jiang GF: The analysis of GM (1,1) grey model to predict the incidence trend of typhoid and paratyphoid fevers in Wuhan City, China. *Medicine*, 2018; 97: e11787
8. Wang YW, Shen ZZ, Jiang Y: Comparison of ARIMA and GM (1,1) models for prediction of hepatitis B in China. *PLoS One*, 2018; 13: e0201987
9. Kane MJ, Price N, Scotch M, Rabinowitz P: Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. *BMC Bioinformatics*, 2014; 15: 276
10. Zhang LP, Wang L, Zheng YL et al: Time prediction models for echinococcosis based on gray system theory and epidemic dynamics. *Int J Environ Res Public Health*, 2017; 14: 262
11. Deng JL: Introduction to grey system theory. *J Grey Syst*, 1989; 1: 1–24
12. Yin MS: Fifteen years of grey system theory research: A historical review and bibliometric analysis. *Expert Syst Appl*, 2013; 40: 2767–75
13. Law DC, Serre ML, Christakos G et al: Spatial analysis and mapping of sexually transmitted diseases to optimise intervention and prevention strategies. *Sex Transm Infect*, 2004; 80: 294–99
14. Yin F, Feng ZJ, Li XS: Spatial analysis of county-based gonorrhoea incidence in mainland China, from 2004 to 2009. *Sex Health*, 2012; 9: 227–32
15. Yin F, Feng Z, Li X: Spatial analysis of primary and secondary syphilis incidence in China, 2004–2010. *Int J STD AIDS*, 2012; 23: 870–75
16. National Health and Family Planning Commission of the Peoples Republic of China. China health statistical yearbook Available from: URL: <http://www.nhfpc.gov.cn/zwgkzt/tjnj/list.shtml>
17. Qian SS, Guo W, Xing JN et al: Diversity of HIV/AIDS epidemic in China: A result from hierarchical clustering analysis and spatial autocorrelation analysis. *AIDS*, 2014; 28: 1805–13
18. Zhao YP, Luo TY, Tucker JD, Wong WC: Risk factors of HIV and other sexually transmitted infections in China: A systematic review of reviews. *PLoS One*, 2015; 10: e0140426
19. Tucker JD, Chen XS, Peeling RW: Syphilis and social upheaval in China. *N Engl J Med*, 2010; 362: 1658–61
20. National Health and Family Planning Commission of the Peoples Republic of China. 2015 China AIDS Response Progress Report. Available from: URL: [http://www.aidsdatahub.org/sites/default/files/publication/China\\_narrative\\_report\\_2015.pdf](http://www.aidsdatahub.org/sites/default/files/publication/China_narrative_report_2015.pdf)
21. NCAIDS, NCSTD, China CDC. Update on the AIDS/STD epidemic in China in December, 2016. *Chin J AIDS STD*, 2017; 23: 93
22. Li YB: Epidemiological characteristics of HIV/AIDS combined with syphilis infection. *Occup and Health*, 2017; 33: 1665–67
23. Du MR, Zhao J, Zhang JX et al: Depression and social support mediate the effect of HIV self-stigma on condom use intentions among Chinese HIV-infected men who have sex with men. *AIDS Care*, 2018; 30: 1197–206
24. Huang MB, Ye L, Liang BY et al: Characterizing the HIV/AIDS Epidemic in the United States and China. *Int J Environ Res Public Health*, 2016; 13: 30
25. Wang L, Guo W, Li DM et al: HIV epidemic among drug users in China: 1995–2011. *Addiction*, 2015; 110(Suppl. 1): 20–28
26. Ritchwood TD, Ford H, DeCoster J et al: Risky sexual behavior and substance use among adolescents: A meta-analysis. *Child Youth Serv Rev*, 2015; 52: 74–88
27. Ma Y, Du CH, Cai T et al: Barriers to community-based drug dependence treatment: implications for police roles, collaborations and performance indicators. *J Int AIDS Soc* 2016;19(4 Suppl3): 20879.
28. Cao WT, Li R, Ying JY et al: Spatiotemporal distribution and determinants of gonorrhoea infections in mainland China: A panel data analysis. *Public Health*, 2018; 162: 82–90
29. Guliner M, Zhu L, Wang LB, Zhou XB: Epidemiological study of syphilis infection in Xinjiang international port from 2001 to 2010. *Chin J AIDS STD*, 2013; 19: 138–39
30. Liao KJ, Zhang SK, Liu M et al: Seroepidemiology of syphilis infection among 2 million reproductive-age women in rural China: A population-based, cross-sectional study. *Chin Med J (Engl)*, 2017; 130: 2198–204