



A Markov based model to estimate the number of syphilis cases among floating population



Chentong Li ^{a, *}, Jiawei Liu ^b, Yingying Zhang ^c, Huan Lei ^a, Jinhu Xu ^d, Yao Rong ^e

^a Guangdong Key Laboratory of Modern Control Technology, Institute of Intelligent Manufacturing, Guangdong Academy of Science, Guangzhou, Guangdong, 510 070, PR China

^b Department of Ecology and Evolution, University of Chicago, Chicago, IL, 60 637, USA

^c College of Science, Northwest Agriculture and Forestry University, Yangling, Shaanxi, 712 100, PR China

^d School of Sciences, Xi'an University of Technology, Xi'an, Shaanxi, 710 049, PR China

^e School of Mechanical Engineering and Automation, Harbin Institute of Technology, Shenzhen, Guangdong, 518 055, PR China

ARTICLE INFO

Article history:

Received 27 September 2021

Received in revised form 11 January 2022

Accepted 16 January 2022

Available online 24 January 2022

Handling Editor: Dr Lou Yijun

Keywords:

Syphilis

Markov process

Immigrant workers

Sexual workers

ABSTRACT

Syphilis is a sexually transmitted disease that spreads widely around the world, infecting tens of millions of people every year. In China, syphilis not only causes more than 1 million infections every year, but also has its own characteristics in spreading pattern: this disease always spreads with the migration of floating population. There have been many related investigations and studies on the transmission of syphilis with the floating population in China, but the study of quantitative modeling in this field is very limited. In this paper, based on the Markov process model and datasets collected in Zhejiang Province, China, we construct a new model to analyze the transmission and immigration process of syphilis. The results show that immigrant patients are one of the sources of infection of syphilis in Zhejiang province, and the infection rate is remarkable which should not be ignored. By using the PRCC method to analyze the relationship between parameters and infected cases, we also find two main effective measures that can control the spread of syphilis and reduce the infection rate: the self-attention of infected persons, and the use of sexual protection measures. With the increasing frequent exchanges of people among different countries and regions, studying the transmission of diseases with the floating populations has become more and more important. The method we use in this paper gives a new insight studying this issue, providing a quantitative research method using the data of diagnosed cases. All the methods and models in this paper can be extendly used in the studies of other diseases where immigrant patients should be considered.

© 2022 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Corresponding author.

E-mail address: ct.li@giim.ac.cn (C. Li).

Peer review under responsibility of KeAi Communications Co., Ltd.

1. Introduction

Syphilis is a worldwide sexually transmitted disease caused by the bacterium *Treponema pallidum*. In 2015, there were about 45.4 million people infected with syphilis in the world (Vos et al., 2016), of which 6 million were newly infected cases (Newmanjane Rowley et al., 2015), resulting in about 107,000 deaths (Wang et al., 2016). In mainland China, the first case of syphilis was found in the year 1505. By 1950, 84% of sexual workers, and 2–3% of rural residents had suffered from this disease (Chen et al., 2007). In the 1960s, the Chinese government led an elimination campaign of sexually transmitted disease, so syphilis and other sexually transmitted disease were virtually eradicated (Chen et al., 2007). However, with the relaxation of government prevention and control measures, new cases of syphilis appeared in 1979 (Chen et al., 2007). And in 2015, the annual incidence rate of syphilis in China increased to 9.81 per 100,000 people (Wong et al., 2018). Nowadays, syphilis is regarded as a Class B infectious disease in mainland China and the Ministry of Health of China has established a national-wide surveillance system to control this disease.

In China, sexual workers are one of the main sources to transmit sexually transmitted disease, such as AIDS and syphilis (Li et al., 2009; Zhou et al., 2020). Many of them come from rural areas of poor provinces, and become sex workers in the immigrant workers' gathering area of big cities in rich provinces, such as Zhejiang Province (Ma et al., 2017; Xiao et al., 2013; Zhou et al., 2020). The survey data on Ningbo city, Zhejiang province, shows that 81.09% of sexual workers comes from other provinces (Jiang et al., 2012). Another source of syphilis are the immigrant workers, who also come from rural areas of poor provinces and work in the cities of rich provinces. Many of them engage in sexual relationships with sex workers or other non-marital partners (Yang et al., 2014), which accelerates the spread of syphilis and increases the difficulty of prevention and control of this disease.

With the rapid development of China's economy and the increase in population mobility, the problem of infectious diseases carried by the floating population cannot be ignored. Some survey works (Shen & Pang, 2008; Xie, 2000) show that the floating population could accelerate the spread of many infectious diseases, such as malaria, AIDS and hepatitis B. The latest data show that those big cities or rich provinces with lots of immigrant workers (such as Beijing and Zhejiang Province) are also the places where new cases of syphilis are found (Tucker & Cohen, 2011). Although there are many data investigations and studies on the transmission of diseases with the floating population, few works have constructed models to quantitatively analyze how the floating population could influence the transmission of disease, such as the syphilis.

To compensate the limitation of theoretical models, in this paper, a method based on Markov process is introduced to estimate the floating population. Based on Markov process and Bayes' theorem, this method could generate the estimated number of infections of a disease over time (Sweeting et al., 2005). This method has long been used to estimate the HIV incidence rate with the AIDS diagnosed cases (BrookmeyerMitchell et al., 1994). In the work of Birrell et al. (Birrell et al., 2012), the authors applied this method to the new HIV, AIDS diagnoses with the help of data on CD4 counts in England and Wales. Using this method, the authors characterized the time-varying distribution of the time between infection and diagnosis, and also estimated the HIV incidence. Furthermore, in the subsequent work, Birell and his coauthors (Birrell et al., 2013) used this method to analyze the spread of HIV among the nationwide population of UK. Based on Birell's work, Sun et al. (Sun et al., 2020) further estimated the cases of HIV and AIDS within different ranges of CD4 counts in Guangxi province, China.

Overall, in this paper, we improve the traditional method and introduce a novel Markov model which can estimate the number of infected patients among the floating population. Based on this method, we divided the people newly diagnosed with syphilis into two parts: people infected in the working province and those infected in other provinces. Based on the Markov processes of those people who were infected outside the working province and the infection pattern of the syphilis, we construct a new model to describe the infection population in the floating population. In this paper, we refer Zhejiang province as the working province. Using the likelihood function and the data on diagnosed cases in Zhejiang province (Zhejiang; Zhejiang provincial burea, 1525), we estimate how the floating population can affect the spread of syphilis in Zhejiang Province. In actual calculations, we use the data from January 2009 to October 2020. The samples after October 2020 are dropped to reduce the influence of control strategies of COVID-19 on our model.

2. Method

In this section, we mainly describe how the model is constructed and show the likelihood function used in model fitting. Firstly, a Markov model which can describe the probabilities of the immigrant patients' behaviors is constructed. Based on this model, we find an equation to calculate the expected number of diagnosed cases in each month. Then based on the Poisson distribution, a likelihood function that could link the expected number of new cases and the real data is established. Finally, by the standard estimation parameter method and the likelihood function, we fit the model. All of the parameters and variables that used in this section are listed in Table 1 and Table 2.

The state probabilities of one person infected out of his working province but finally diagnosed in the working province at time t are made up of three parts: (1) the probability that a person is infected outside his working province but has not immigrated to this province; (2) the probability that a person was infected outside his working province and has already immigrated to the working province but not been diagnosed; (3) the probability that a person was infected outside his working province and has been diagnosed in the working province. Let $p_0(t)$, $p_1(t)$ and $p_2(t)$ denote these three probabilities. Then the equations can be written as,

Table 1
The table of the parameters.

| Parameters | Description |
|------------|-------------------------------------------------------------------------|
| α | The immigration rate. |
| β | The infection rate of the syphilis. |
| γ | The rate of infection ability loss due to either recovery or migration. |
| p | The parameter to evaluate the Spring Festival effect on infection rate. |
| δ | The diagnostic rate. |

$$\begin{cases} \frac{dp_0(t)}{dt} = -\alpha p_0(t), \\ \frac{dp_1(t)}{dt} = \alpha p_0(t) - \delta p_1(t), \\ \frac{dp_2(t)}{dt} = \delta p_1(t), \end{cases} \tag{1}$$

where α is the immigration rate, and δ is the diagnostic rate. The diagram of that process is shown in Fig. 1. Many of the immigrant workers or sexual workers tend to work in the working province for a long time (Yang et al., 2014). Thus in this Markov process, we don't consider the case where the person migrates back. When the person is infected at time 0, the initial conditions can be given as $p_0(0) = 1$, $p_1(0) = 0$, and $p_2(0) = 0$. Thus, the solutions of equation (1) are,

$$p_0(t) = e^{-\alpha t}, \quad p_1(t) = \frac{\alpha}{\delta - \alpha} (e^{-\alpha t} - e^{-\delta t}),$$

and

$$p_2(t) = \frac{\delta e^{-\alpha t} - \alpha e^{-\delta t}}{\alpha - \delta} + 1.$$

Therefore, the probability M of N person who come from other place and finally diagnosed at Zhejiang province after the time t could be given as,

$$\binom{N}{M} p_2(t)^M (1 - p_2(t))^{N-M},$$

which is a binomial distribution with the expectation $Np_2(t)$. This expectation and similar binomial distribution process with the same expectation formula will be used in the following equations.

One person who is infected with syphilis will hardly self-recover from this disease until the person is diagnosed and treated by the specific therapies (Clement et al., 2014). In China, syphilis is a national legally classified B infectious disease and all of the newly diagnosed syphilitics must be reported to the CDC (China, 2018). Based on this situation, the expected number of newly infected people in the patients' working province at time t can be written as the sum of the expected numbers of people who are infected by (1) the diagnosed and not healed people, (2) the newly incoming undiagnosed immigrants, and (3) the other undiagnosed people at time t :

$$I(t) = \beta S(t) \left(\int_{-\infty}^t \mu(\tau) g_{\tau,t} + \lambda(\tau) p_1(t - \tau) + I(\tau) f_{\tau,t} d\tau \right), \tag{2}$$

where $g_{\tau,t} = e^{-\gamma(t-\tau)}$ is the probability that one person is diagnosed at time τ , but will not be healed or migrate to another province at time t , and γ is the rate of infection ability loss due to either recovery or migration. $\lambda(\tau)$ is the number of immigrants who have already been infected outside their working province at time τ . And $f_{\tau,t} = e^{-\delta(t-\tau)}$ is the probability that one person is already infected at time τ but not diagnosed at time t , and β is the infection rate, respectively. The term $S(t)$ is the number of susceptible people in the immigrant patients' working province at time t and this comes from the population data in the Zhejiang Statistical Yearbook (Zhejiang provincial burea, 1525). Since the statistical yearbook includes only annual data, in order to get the monthly data, we assume that the population increase in each month is the same in the actual calculation, and then use the population interpolation of the adjacent two years to calculate the population of each month. Meanwhile, the number of susceptible people (ten million) is much larger than that of syphilis cases (thousands), thus in the actual calculation, the change in the number of susceptible people caused by infected cases is ignored. The function $\mu(t)$ represents the expected number of diagnosed people at time t and can be written as the sum of the expected numbers of residents and new immigrants who are diagnosed at time t :

Table 2
The table of the variables.

| Variables | Description |
|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------|
| $p_0(t)$ | The probability that a person is infected outside his working province but has not immigrated to this province. |
| $p_1(t)$ | the probability that a person was infected outside his working province and has already immigrated to the working province but not been diagnosed. |
| $p_2(t)$ | The probability that a person was infected outside his working province and has been diagnosed in the working province. |
| $\lambda(t)$ | The number of immigrants who have already been infected outside their working province at time t . |
| $S(t)$ | The number of the susceptible person at the patients' working province at time t . |
| $I(t)$ | The expectation number of newly infected cases at the patients' working province at time t . |
| $\mu(t)$ | The expectation number of newly diagnosed cases at the patients' working province at time t . |
| $D(t)$ | The data of newly diagnosed cases at the patients' working province at time t . |

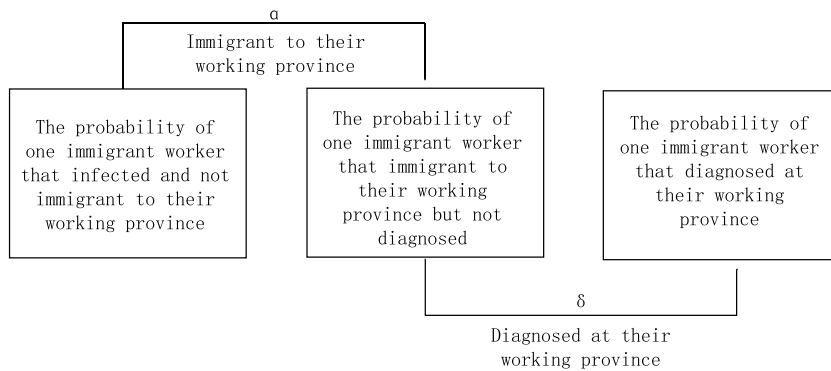


Fig. 1. The diagram of the Markov process that one infected person that finally immigrant to Zhejiang province and is diagnosed here.

$$\mu(t) = \int_{-\infty}^t I(\tau)(1 - f_{\tau,t}) + \lambda(\tau)p_2(t - \tau)d\tau. \tag{3}$$

Within the month around the Spring Festival holiday (Jiang et al., 2015; Yao et al., 2019), many workers including the sex workers will go back to their hometown and stop their work until the end of the Spring Festival. So, the infection rate β will become smaller during that period. Thus, we could rewrite the infection rate as $p(t)\beta$, where

$$p(t) = \begin{cases} p, & \text{if } t \text{ is the month around the Spring Festival holiday,} \\ 1, & \text{others,} \end{cases}$$

and $0 < p < 1$ is the parameter to evaluate the effect of the Spring Festival on infection rate.

Based on the Poisson distribution, the probability of $D(t)$ cases at time t given the model parameters could be written as,

$$P(D|\theta) = \prod_{i=1}^N \frac{\mu(t_i)^{D(t_i)}}{D(t_i)!} \exp(-\mu(t_i)),$$

where $\theta = \{\lambda(t), \alpha, \beta, \delta, \gamma, p, \mu(t), I(t)\}$ is the set of parameters and terms in model (3), N is the number of data points, and $t_i, i = 1, 2, \dots, N$ are the time points in data, respectively. Hence the likelihood function can be written as,

$$L(D|\theta) = \sum_{i=1}^N (D(t_i) \log(\mu(t_i)) - \mu(t_i) - \log \Gamma(D(t_i) + 1)), \tag{4}$$

where $\log \Gamma(x)$ is the log Gamma function. Based on equation (4), the Markov Chain Monte Carlo (MCMC) method (Bahamóndez & Fröhlich, 2019) is used to estimate the posterior distributions of parameter set θ .

In the numerical calculation, we use the date 12 months before the present time to replace the negative infinity in the integral formulas (2) and (3), so that we can get an approximate estimation while increasing the calculation efficiency. By that approximate estimation, the samples collected from January 2009 to December 2009 are used as the known diagnosed cases and the prior of the $I(t)$ (which is the initial values of the estimation), and formula (4) is used to fit the other 9 years' samples.

The function $\lambda(t)$ is considered as a piece-wise function and is assumed as a constant in every three months (Jan to March, April to June, July to September, and October to December have different constants in different years). All of the source codes are available at the authors' GitHub (<https://github.com/ChentongLi/Migration-population-estimation>).

One of the data sources is the official website of the Zhejiang Provincial Center for Disease Control and Prevention (Zhejiang). According to the requirements of China's infectious disease prevention and control, the CDC in each province needs to count and announce the number of cases infected with Class A and B infectious diseases each month, and syphilis is a Class B infectious disease. Another source on the number of susceptible people $S(t)$ comes from the official website of Zhejiang Provincial Statistics Bureau (Zhejiang provincial burea, 1525). In China, the statistical bureaus of each province announce the number of residents in the previous year at the beginning of each year.

3. Main results

Fig. 2 shows the fitted result of the model in Section 2. The black line represents the mean values of the estimated number of diagnosed patients and the black points which come from the data sets are the reported diagnosed cases. The sky-blue shadow in that figure shows the 95% confidence interval of the estimated results. This figure shows that our Markov model (1) could fit the data set very well verifying that our model could reflect the immigration and infection processes of syphilis among the immigrant workers in Zhejiang province.

Fig. 3 illustrates the estimated values of the parameters in our model. The mean value of the infection rate β is 1.66×10^{-9} per health person per month in Zhejiang province and its 95% confidence interval is $(1.43 \times 10^{-9}, 1.92 \times 10^{-9})$. Considering the whole population of Zhejiang province (around 30 million residents (Zhejiang provincial burea, 1525)), although the infection rate β is small, it can still cause more than 30,000 diagnosed patients per year. Disease prevention and control departments need to consider measures such as the distribution of condoms to reduce the infection rate and control the spread of syphilis. The mean value of the parameter p , which is used to evaluate the Spring Festival effect, is 0.12 (with the 95% confidence interval (0.03, 0.20)). It shows that the Spring Festival has played a role in reducing the spread of syphilis in China, because people at this time seldom have sexual relationships with sex workers. The mean value of the parameters α , δ and γ of the Markov process (1), are 9.95 per month (95% confidence interval (6.05, 13.88)), 1.73×10^{-2} per month (95% confidence interval $(1.64 \times 10^{-2}, 1.79 \times 10^{-2})$), and 1.92 per month (95% confidence interval (1.15, 2.71)), respectively. These parameters reflect the immigration and infection process of syphilis, and can be used in other related works on simulating the immigration of workers and syphilis transmission among healthy people.

The partial rank correlation coefficients (PRCC) between the model parameters and the mean infected cases from the year 2010–2019 (MI) are also calculated in this paper shown in Fig. 4. The samples for calculating the PRCC are collected from the 95% confidence intervals of the model parameters α , β , δ , γ and p , which are generated from the MCMC method. The result shows that the parameters δ and γ have negative effects on the mean infected number, while the other parameters have positive effects. And the parameter p , which is a value to reflect human behavior, has the greatest impact on the mean infection number over these years. This means that the most effective way to control syphilis is that everyone tries to control the behavior of themselves. The second greatest factor in disease control is the infection rate, which means measures to decrease the infection rate, such as the use of condoms, can have great impacts on the control of syphilis.

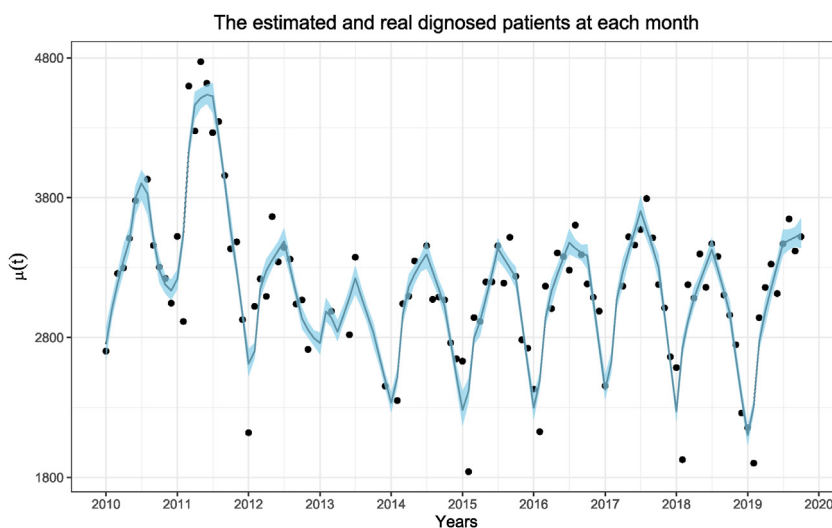


Fig. 2. The fitted number of diagnosed patients (black line) and reported number (points) in each month. The sky-blue shadow shows the 95% confidence interval of the fit results.

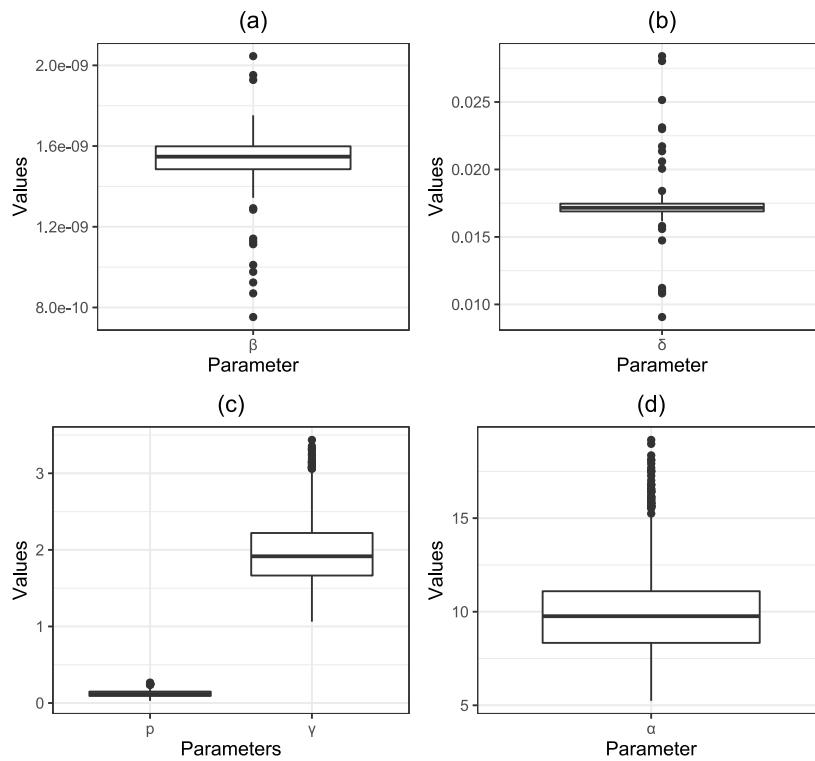


Fig. 3. The estimated parameter values of α , β , δ , γ and p .

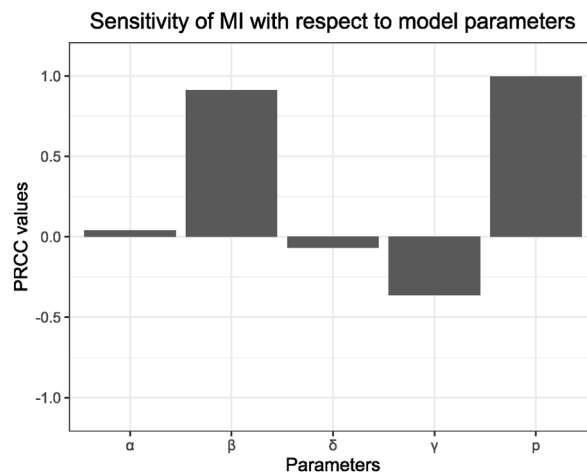


Fig. 4. The sensitivity of the mean infected cases in Zhejiang province from the year 2010–2019 with respect to the parameters α , β , δ , γ and p .

The estimated results of the number of patients infected in Zhejiang province are shown in Fig. 5. The black line shows the mean values in each month and the sky-blue shadow shows the 95% confidence intervals. This figure shows the transmission of syphilis in Zhejiang province. Compared with other years, more people were infected in Zhejiang province in 2010 and 2011. From the year 2012–2019, the number of infected people in each month doesn't show an obvious difference and the differences between years are narrow. The figure also shows the Spring Festival leads to fewer infection cases since at this time people pay more attention on their behavior.

The estimated results of the number of the immigrants who were infected out of Zhejiang province and diagnosed in Zhejiang province in every three months (the piece-wise function $\lambda(t)$) are illustrated in Fig. 6. The distribution of the immigrant syphilis patients shows no clearly regular pattern. The first three months of each year always touch the lowest

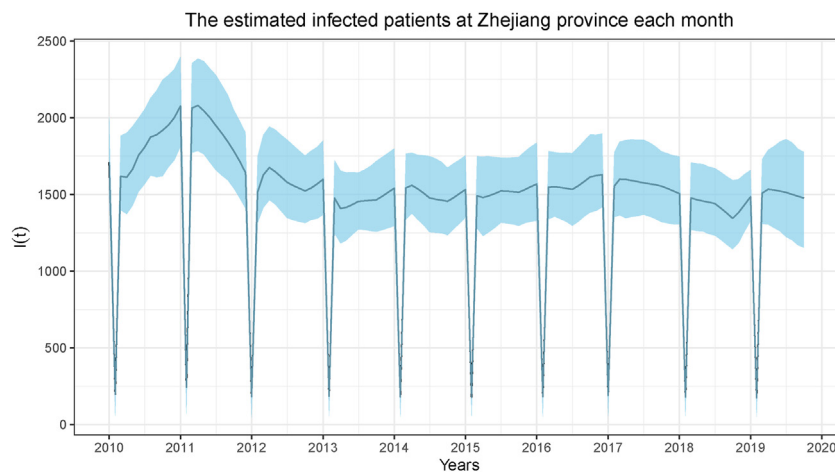


Fig. 5. The fitted number of the patients infected at Zhejiang province in each month. The mean value illustrated in black line and the sky-blue shadow shows the 95% confidence interval of the fit results.

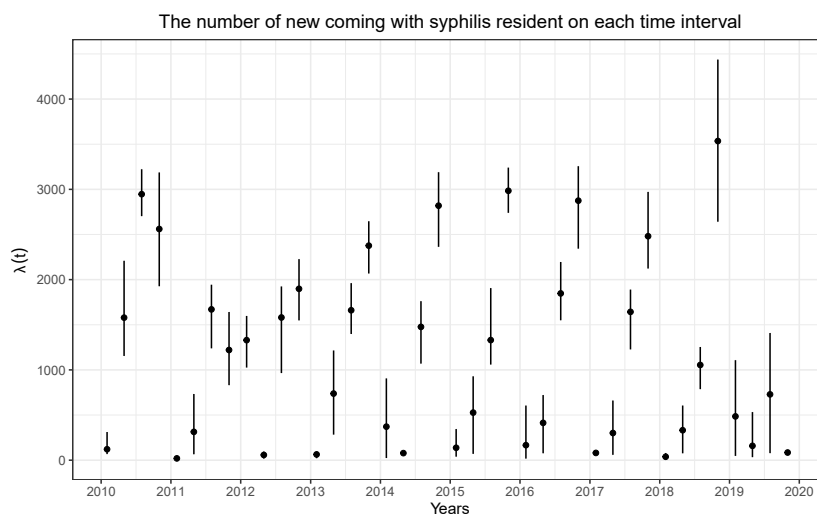


Fig. 6. The fitted number of immigrants infected out of Zhejiang province in each three months. The points show the estimated mean values and the error bars illustrate the 95% confidence intervals.

number, and the number of patients in the second half of each year is always larger than that in the first half. The number of infected people in the warm seasons is sometimes larger than that in the cold seasons. The peak of the number of immigrant syphilis patients is at the period from October to December of the year 2018, with the mean value 3535 and the 95% confidence interval (2641, 4437). Compared the result shown in this figure with the number illustrated in Fig. 5, immigrant patients are one of the main sources of syphilis in Zhejiang province, hence the national disease control department needs to carry out effective actions to supervise these key populations (especially the immigrant sex workers).

4. Discussion

In this paper, based on a Markov process method and the data on the diagnosed cases of syphilis in Zhejiang province, we mainly estimate the possible number of syphilis patients among the floating population. The result shows that immigrant patients are the main source of syphilis in Zhejiang province, implying that the government should pay more attention to the immigrant sex workers. From the estimation result of the parameters, we find that the Spring Festival has an effect of decreasing the transmission of syphilis. We also find that the infected cases in Zhejiang province stay as an almost stable distribution among different months. By the PRCC analysis, two main measures can be considered to control syphilis: (1) the

infected people pay attention to their behaviors and (2) the usage of practical ways to reduce the infection rate, such as the wide use of condoms among sex workers.

By using the Markov process and infection equations to describe the immigration and infection processes of the syphilis patients, the model is constructed within a floating population. This method can estimate not only the number of infected cases but also the number of floating patients via the data of diagnosed cases, which improves the estimation ability of the traditional method. With the rapid development of modern economy and the increasing frequent exchanges of people among different countries and regions, studying the disease status within the floating populations has become more and more important. And the method we use in this paper can give a new insight studying this problem, providing a quantitative research basis via the diagnosed cases. All of the methods mentioned in this paper can be extensively used in other disease studies where immigrant patients should be considered.

Other than the on-site investigation of the spread of diseases and the patients in the floating population (Jiang et al., 2012), a quantitative analysis on syphilis has been carried out. The number of new infection cases and patients among the immigrant population, as well as the transmission and immigration parameters have been estimated in this paper by using the data on diagnosed cases in each month. These quantitative values can be used to guide how to control diseases more accurately in reality and deepen people's understanding of diseases. Compared with the model in (Sun et al., 2020), we have innovations in model and likelihood function construction in terms of calculation details. Based on the Markov process model, we have not only considered the factor of floating population but also modified the likelihood function to make it more in line with the issues considered in this paper.

For the model described in this article, the environmental factors in the spread of the disease are ignored because it is a sexually transmitted disease. However, for air-transmitted diseases such as influenza, and vector-transmitted diseases such as dengue fever, these diseases are sensitive to the effects of temperature and humidity, and can also migrate with the migration of the population. Therefore, to meet the more practical needs for disease studies, a more comprehensive model incorporating environmental factors should be considered in future work.

Declaration of competing interest

We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Acknowledgement

This work is supported by the financial support from GDAS' Project of Science and Technology Development (2021GDASYL-202103089), China Postdoctoral Science Foundation (2021M690747), National Natural Science Foundation of China (11701445), Science and Technology Program of Guangzhou (202007040007), GDAS' Project of Science and Technology Development (2019GDASYL-0502007), Guangdong Provincial Rural Revitalization Strategy Special Fund Project (2019KJ138), and Guangdong Basic and Applied Basic Research Foundation (2019A1515110503).

References

- Bahamóndez, G. I., & Fröhlich, H. (2019). Mcmc techniques for parameter estimation of ode based models in systems biology. *Front. Appl. Math. Stat.* <https://doi.org/10.3389/fams.2019.00055>
- Birrell, P. J., Chadborn, T. R., Gill, O. N., Delpech, V. C., & De Angelis, D. (2012). Estimating trends in incidence, time-to-diagnosis and undiagnosed prevalence using a cd4-based bayesian back-calculation. *Statistical Communications in Infectious Diseases*, 4(1).
- Birrell, P. J., Gill, O. N., Delpech, V. C., Brown, A. E., Desai, S., Chadborn, T. R., Rice, B. D., & De Angelis, D. (2013). Hiv incidence in men who have sex with men in england and wales 2001–10: A nationwide population study. *The Lancet Infectious Diseases*, 13(4), 313–318.
- Brookmeyer, R., Mitchell, H. G., Mitchell, H. G., et al. (1994). *AIDS epidemiology: A quantitative approach*. Oxford University Press on Demand.
- Chen, Z.-Q., Zhang, G.-C., Gong, X.-D., Lin, C., Gao, X., Liang, G.-J., Yue, X.-L., Chen, X.-S., & Cohen, M. S. (2007). Syphilis in China: Results of a national surveillance program. *The Lancet*, 369(9556), 132–138.
- China CDC. Management of infectious disease information report. http://www.chinacdc.cn/jkzt/crb/xcrxjb/201810/t20181017_195160.html.
- Clement, M. E., Lance Okeke, N., & Hicks, C. B. (2014). Treatment of syphilis: A systematic review. *JAMA*, 312(18), 1905–1917.
- Jiang, Q., Sun, Y. L., Wang, Z., & Yin, Y. (2015). Aerosol composition and sources during the Chinese spring festival: Fireworks, secondary aerosol, and holiday effects. *Atmospheric Chemistry and Physics*, 15(11), 6023.
- Jiang, J., Wang, H., Fang, W., Sun, J., Chen, B., Bo, D., & Dong, W. (2012). Investigation on the prevalence of aids among female sex workers in a supervision place in ningbo city, zhejiang province. *Disease Surveillance*. in chinese.
- Li, S., Huang, H., Cai, Y., Xu, G., Huang, F., & Shen, X. (2009). Characteristics and determinants of sexual behavior among adolescents of migrant workers in Shanghai (China). *BMC Public Health*, 9(1), 195.
- Ma, Q., Jiang, J., Pan, X., Cai, G., Wang, H., Zhou, X., Jiang, T., & Chen, L. (2017). Consistent condom use and its correlates among female sex workers at hair salons: A cross-sectional study in zhejiang province, China. *BMC Public Health*, 17(1), 910.
- Newman, L., Jane Rowley, Vander Hoorn, S., Wijesooriya, N. S., Unemo, M., Low, N., Stevens, G., Gottlieb, S., James, K., & Temmerman, M. (2015). Global estimates of the prevalence and incidence of four curable sexually transmitted infections in 2012 based on systematic review and global reporting. *PLoS One*, 10(12), Article e0143304.
- Shen, J., & Pang, Y. (2008). Analysis on the characteristics of infectious diseases among floating population in huzhou city, zhejiang province, 2004–2006. in chinese *Disease Surveillance*, 23, 315 (20080519).
- Sun, X., Yang, W., Tang, S., Shen, M., Wang, T., Zhu, Q., Shen, Z., Tang, S., Chen, H., Ruan, Y., et al. (2020). Declining trend in hiv new infections in guangxi, China: Insights from linking reported hiv/aids cases with cd4-at-diagnosis data. *BMC Public Health*, 20(1), 1–12.
- Sweeting, M. J., De Angelis, D., & Aalen, O. O. (2005). Bayesian back-calculation using a multi-state model with application to HIV. *Statistics in Medicine*, 24(24), 3991–4007.

- Tucker, J. D., & Cohen, M. S. (2011). China's syphilis epidemic: Epidemiology, proximate determinants of spread, and control responses. *Current Opinion in Infectious Diseases*, 24(1), 50.
- Vos, T., Allen, C., Arora, M., Barber, R. M., Bhutta, Z. A., Brown, A., Carter, A., Casey, D. C., Charlson, F. J., Chen, A. Z., et al. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: A systematic analysis for the global burden of disease study 2015. *The Lancet*, 388(10053), 1545–1602.
- Wang, H., Naghavi, M., Allen, C., Barber, R. M., Bhutta, Z. A., Carter, A., Casey, D. C., Charlson, F. J., Chen, A. Z., Coates, M. M., et al. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the global burden of disease study 2015. *The Lancet*, 388(10053), 1459–1544.
- Wong, N. S., Chen, L., Tucker, J. D., Zhao, P., Goh, B. T., Poon, C. M., Yang, L., Yang, B., Zheng, H., & Huang, S. (2018). Distribution of reported syphilis cases in south China: Spatiotemporal analysis. *Scientific Reports*, 8(1), 1–9.
- Xiao, S., Wang, R., Zhong, X., Tiao, K., Fu, G., Zhang, Y., Peng, B., & Huang, A. (2013). Association rules based approach to aids knowledge acquisition among female sex workers and its relationship with demographic characteristics. in chinese *Journal of Shanghai Jiaotong University Medical Science*, 10, 80–84.
- Xie, S. (2000). Epidemiological characteristics of infectious diseases among floating population in zhejiang province from 1997 to 1998. in chinese *Journal of Zhejiang Preventive Medicine*, 6, 11–13.
- Yang, Z., Wang, X., Diao, R., Wang, M., Zhang, Y., Zheng, Y., & Li, S. (2014). Reflections on the migrant workers' sexual morality in the process of urbanization: Overview on the migrant workers non-marital sexual behaviors. in chinese *Chinese Journal of Human Sexuality*, (8), 93–96.
- Yao, L., Wang, D., Fu, Q., Qiao, L., Wang, H., Li, L., Sun, W., Li, Q., Wang, L., Yang, X., et al. (2019). The effects of firework regulation on air quality and public health during the Chinese spring festival from 2013 to 2017 in a Chinese megacity. *Environment International*, 126, 96–106.
- Zhejiang CDC. Health commission of zhejiang province. <https://wsjkw.zj.gov.cn/>.
- Zhejiang provincial bureau of statistics. Statistical yearbook of zhejiang province. <http://tjj.zj.gov.cn/col/col1525563/index.html>.
- Zhou, X., Ma, Q., Pan, X., Chen, L., Wang, H., & Jiang, T. (2020). The prevalence and correlates of oral sex among low-tier female sex workers in zhejiang province, China. *PLoS One*, 15(9), Article e0238822.