



ELSEVIER

Contents lists available at [SciVerse ScienceDirect](http://www.sciencedirect.com)

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Predicting the HIV/AIDS epidemic and measuring the effect of mobility in mainland China

Yanni Xiao^{a,*}, Sanyi Tang^b, Yicang Zhou^a, Robert J. Smith^c, Jianhong Wu^d, Ning Wang^{e,**}

^a Department of Applied Mathematics, Xi'an Jiaotong University, Xi'an 710049, PR China

^b College of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710062, PR China

^c Department of Mathematics and Faculty of Medicine, The University of Ottawa, 585 King Edward Ave, Ottawa, ON, Canada K1N 6N5

^d Centre for Disease Modelling, York University, York, Toronto, ON, Canada M3J 1P3

^e National Center for AIDS/STD Prevention and Control, Chinese Center for Disease Control and Prevention, 155 Changbai Rd, Beijing 102206, PR China

HIGHLIGHTS

- ▶ Assess the HIV/AIDS epidemic in mainland China accurately.
- ▶ The mobility patterns for HIV-positive individuals or AIDS patients are obtained.
- ▶ Address effect of population mobility on the HIV epidemic in mainland China.
- ▶ The community-based support and care program is suggested.

ARTICLE INFO

Article history:

Received 4 March 2012

Received in revised form

27 September 2012

Accepted 29 September 2012

Available online 9 October 2012

Keywords:

Mathematical model

HIV infection

The basic reproduction number

Networks

Interventions

ABSTRACT

HIV has spread widely in mainland China, but there is significant geographic variation in the severity of the epidemic. We aimed to assess the HIV/AIDS epidemic in mainland China accurately, and address the effect of population mobility on it. Markov-Chain Monte-Carlo simulations and Latin Hypercube Sampling were used to estimate the basic reproductive ratio and its sensitivity to parameter variations. We estimated a mean reproduction number of 1.708 (95% CI 1.440–1.977). Our analysis using national surveillance data indicates that HIV-positive individuals most likely move from economically developed regions to regions with more numerous HIV cases, while mobility of AIDS patients likely flows in the opposite direction, due to the current policy that AIDS patients must return to their registered residence to receive free antiretroviral therapy. Our results based on a spatially stratified population dynamical model show increasing mobility rates of HIV/AIDS cases can have a significant effect on the number of HIV/AIDS cases per province and has the potential to decrease the overall number of HIV/AIDS cases in the country. We recommend that the community-based HIV/AIDS support and care program should be implemented by some local governments (especially in epidemically severe areas) to mitigate HIV infections in China.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Since the first case of acquired immunodeficiency syndrome (AIDS) in China was reported in 1985, human immunodeficiency virus (HIV) has spread to all of mainland China (CMH, 2005, 2009; Wu et al., 1995). According to estimates by experts in 2011 from the Chinese Center for Disease Control and Prevention, the accumulated number of people living with HIV/AIDS (PLWHIV) has reached 780,000, of which 154,000 were living

with AIDS (CMH, 2011). Moreover, comprehensive data from the south of China showed that HIV has advanced from high-risk groups to others in the general population (Lu et al., 2008). To prevent the epidemic from spreading and to mitigate the impact of AIDS, enhanced HIV/AIDS prevention and control has been implemented at a national level in recent years (CMH, 2005, 2009). HIV/AIDS prevention and control efforts in China have made significant progress. However, many problems and numerous challenges remain (Cui et al., 2009). How to assess the situation accurately and evaluate effectiveness of various prevention and control measures from the standpoint of public health policy are key to realize the objectives stated in the China Long and Medium Term Plan on HIV/AIDS Prevention and Control (SCD, 1998; SCP, 2006).

* Corresponding author. Tel.: +86 29 826 639 38; fax: +86 29 82 668 559.

** Corresponding author.

E-mail addresses: yxiao@mail.xjtu.edu.cn (Y. Xiao), wangnbj@163.com (N. Wang).

The status of the HIV/AIDS epidemic is quite different in different areas (CMH, 2009). Until the end of 2009, provinces with more than 50,000 reported HIV infections were Yunnan, Guangxi, Sichuan and Guangdong, while Qinghai and Tibet had less than 200 infections. There are key factors such as needle sharing among intravenous drug users (IDUs), unsafe sexual behaviors and high population mobility that affect the HIV/AIDS epidemic in mainland China (CMH, 2009). In particular, according to the current policy of HIV treatment, AIDS patients have to return to their registered residence (defined as province where an individual gets his/her certification of birth in China's household system) to get free antiretroviral treatment (ART), irrespective of where they live. This significantly increases the movement of AIDS patients. Moreover, statistical data show that population mobility has recently become much more frequent between and within provinces (RCMPD, 2010). Whether increased mobility contributes to an increase in disease cases or affects the global outcomes of the geographic distribution is unclear (Smith et al., 2009). If the current variation of the HIV epidemic among provinces is changed by the increasing movement, then the current targeting of interventions in some (or all) provinces may need to be modified.

The primary objective of our study is to understand the epidemic trend in China and to accurately predict the future. To do this we follow a modeling approach for stratification of the population (Anderson and May, 1991) according to the clinical progression of disease and epidemiological status of the individuals. These models (homogeneous and heterogeneous) are fitted to data on the number of people living with HIV or AIDS in mainland China, and hence the basic reproduction number, a crucial quantity for identifying the intensity of interventions required to control an epidemic, and important epidemiological parameters are estimated (Heffernan et al., 2005; Velasco-Hernandez et al., 2002). The second purpose of this study is to analyze the spatial characteristic of the HIV/AIDS epidemic and the mobility patterns of HIV/AIDS cases, based on the national surveillance system, and to address the effect of mobility on the HIV/AIDS epidemic in mainland China. To this end, a metapopulation framework (Levin, 1974) with realistic network structures is formulated and parameterized. The impact of spatial-related interventions on China's HIV infection is then assessed. It is hoped that the outcomes will improve our understanding of the HIV epidemic in China and the implications of increasing movements in HIV/AIDS cases for the spread of HIV/AIDS, and potentially assist in designing effective public-health intervention strategies.

2. Materials and methods

2.1. Data

We obtained data on cases of HIV infection and AIDS in mainland China from current surveillance systems. Note that the data are not publicly available secondary data. Network reporting has covered all counties and this surveillance system gives real-time statistics. For each diagnosed HIV/AIDS case, the current surveillance system has 74 items including name, ID card number, gender, occupation, diagnosis date, date of confirmation, current residence, registered permanent residence, etc. We focus on the information on differences in the current residences, registered residences and the reported provinces for diagnosed HIV-positive individuals or AIDS patients to determine their mobility patterns and use information on the number of annual reported cases of HIV/AIDS to try to accurately assess the HIV/AIDS epidemic in China.

2.2. The model

We used network and compartmental models to understand the spread and control of HIV in China. We initially made an assumption of homogeneous mixing among the entire high-risk population for simplicity. However, the high-risk Chinese populations in different regions exhibit various mixing patterns (Cui et al., 2009). Furthermore, population mobility has recently become much more frequent so spatial-related intervention strategies may be necessary in mainland China (CMH, 2009). We then formulated a metapopulation model to describe the dynamics of HIV infection among patches. We hope to use both simple and complex deterministic models to comprehensively understand the HIV epidemic in China and to explore policy-related questions, including investigation of the impact and effectiveness of a variety of HIV prevention interventions in different provinces in mainland China.

Based on characteristics of the nationwide database, we formulated a baseline model that reflects some key epidemiological properties of the HIV/AIDS epidemic and then implemented public-health interventions (methadone treatment, condoms and sterile syringes). Here, we only focus on high-risk population groups rather than the general population since not all individuals are exposed to the HIV virus. The underlying structure of the model comprises classes of individuals who are high-risk susceptibles (S), HIV infected but not yet diagnosed (I), diagnosed HIV-positive individuals who have not yet progressed to AIDS (D_I), and those with clinical AIDS (D_A) (Fig. 1). For simplicity, we suppose that all AIDS patients have been diagnosed due to their having developed clinical symptoms of AIDS (de Arazoza et al., 2000; Williams and Anderson, 1994). The model equations are

$$\begin{cases} \frac{dS}{dt} = U - (1-\pi)c\beta\frac{I+qD_I+\epsilon D_A}{N}S - (\mu+d)S \\ \frac{dI}{dt} = (1-\pi)c\beta\frac{I+qD_I+\epsilon D_A}{N}S - (\delta+d)I \\ \frac{dD_I}{dt} = \rho\delta I - (\xi+d+\alpha_I)D_I \\ \frac{dD_A}{dt} = (1-\rho)\delta I + \xi D_I - (d+\alpha_A)D_A \end{cases} \quad (1)$$

where $N = S + I + D_I + D_A$ represents high-risk population size. People enter into the susceptible class at a rate U , exit high-risk behavior at a constant rate μ , and become infected at a rate $(1-\pi)c\beta(I+qD_I+\epsilon D_A)S/N$. Here, d is the natural death rate; β denotes the probability of transmission per high-risk behavior (sexual action or needle sharing); c represents the contact rate, i.e., the rate of acquisition of new sexual or needle-sharing partners; π represents the degree of intervention (e.g., due to condom use or methadone treatment); q and ϵ are the modification factors accounting for varying levels of the activity and infectiousness of the diagnosed HIV-positive individuals and the

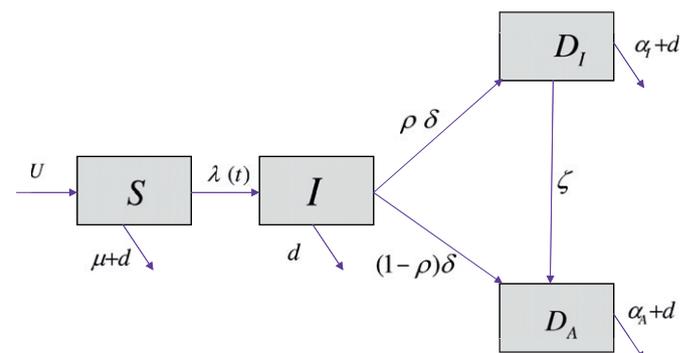


Fig. 1. Flow diagram of the model.

AIDS patients; δ denotes the diagnosis rate; ρ is the proportion of diagnosed individuals who have not yet progressed to AIDS; ζ is the rate of progression from HIV diagnosis to the AIDS class; α_i and α_A denote additional death rates for the diagnosed HIV-positive individuals and for those living with AIDS. Although individuals with HIV do not die directly from the disease, suicide rates are high among HIV-positive individuals in China (Jin et al., 2006). The model is referred to as the national model, illustrated in Fig. 1 and the definition of parameters is given in Table 1.

To evaluate the effectiveness of different spatially relevant interventions and the impact of population mobility, we extended our national model to a metapopulation model, where coupling among patches is through movement across mobility networks (Smith? et al., 2009; Levin, 1974; Tang et al., 2010). In each patch (province), the progression of the disease is tracked using a model structure similar to model (1). Transmission among patches is represented by mobility of individuals from one province to another on the mobility networks for HIV-positive individuals or for AIDS patients (and their corresponding mobility matrices) obtained from surveillance data. Here, we assume that undiagnosed HIV-positive individuals follow a similar network structure to diagnosed HIV-positive individuals but with a higher dispersal rate (denoted by a modification factor $\eta \geq 1$).

The dynamics of HIV-positive individuals for the i th province (patch) is as follows:

$$\begin{cases} \frac{dS_i}{dt} = U_i - \beta_{0i} \frac{I_i + \varrho_i D_i + \epsilon_i D_{Ai}}{N_i} S_i - (\mu_i + d) S_i \\ \frac{dI_i}{dt} = \beta_{0i} \frac{I_i + \varrho_i D_i + \epsilon_i D_{Ai}}{N_i} S_i - (\delta_i + d) I_i + \eta \sum_{j=1}^n M_{ij}^I I_j \\ \frac{dD_{ii}}{dt} = \rho_i \delta_i I_i - (\zeta_i + d + \alpha_{ii}) D_{ii} + \sum_{j=1}^n M_{ij}^I D_{ij} \\ \frac{dD_{Ai}}{dt} = (1 - \rho_i) \delta_i I_i + \zeta_i D_{ii} - (d + \alpha_{Ai}) D_{Ai} + \sum_{j=1}^n M_{ij}^A D_{Aj} \end{cases} \quad (2)$$

where $i = 1, 2, \dots, n$ (here $n = 31$), $\beta_{0i} = (1 - \pi_i) c_i \beta_i$. The definitions of parameters are the same as those in model (1) but with subscript i to indicate the i th patch. Note that on the basis of surveillance data, we can easily determine the number of cases who move from region i to region j , denoted by \bar{C}_{ij} , the element of row i and column j of the matrix \bar{G} (from the networks N_{rc}^I, N_{rr}^A obtained in the following section). We let the entry G_{ij} be a

fraction of the number of PLWHIV in province i who flow into region j , that is,

$$G_{ij} = \frac{\bar{C}_{ij}}{\text{number of PLWHIV in province } i}$$

Then G_{ij} denotes the transition rate from region i to region j and $G_{ii} = 0$. Let matrix E represent the emigration, and is thus a diagonal matrix. Here we have neglected the death rates and birth rates of individuals during the dispersal process, i.e., $E_{ii} = \sum_{j=1}^n G_{ij}$. Hence dispersal among patches is governed by the matrix $-E + G^T$, denoted by matrix M^I (i.e., $M^I = G^T - E$, G^T is the transpose of matrix G) and we have $\sum_{i=1}^n M_{ij} = 0$ for all $j = 1, 2, \dots, n$. Similarly, we can get dispersal matrix M^A for AIDS patients. Note that we do not consider mobility of susceptible individuals since we have no reliable data for movement of susceptibles, and further we will show that this ignorance of the mobility of susceptibles is reasonable. If we consider the dynamics of HIV infection for province i in isolation we simply assume none mobility rates in model (2), which is actually the same as the structure of model (1) but with different parameter values.

2.3. Statistical analysis

On the basis of our current surveillance system using the Kaplan–Meier method, we obtained the mean estimate for the time from HIV diagnosis to AIDS (from the reported HIV-positive individuals between 1990 and 1998 who died from AIDS-related illness) to be 8.6583 years, so we have $\zeta = 0.116$. Similarly, we obtained the additional death rate for the diagnosed HIV-positive individuals to be 0.172, and the disease-related death for the AIDS patients to be 0.318. Note that AIDS patients are inactive in most rural areas where the health care system is poor and discrimination is serious, whereas they may be active in big cities where there are good health medical systems and little discrimination (Zhu et al., 2012; Yang et al., 2005). Moreover, most majority of AIDS patients live in the rural areas (Cui et al., 2009), and consequently we assume that the modification factor in the transmission coefficient for the AIDS patients ϵ is initially set to be zero (Williams and Anderson, 1994; de Arazoza et al., 2000; McCluskey, 2003) and the effect of varying it on outcomes will be analyzed by sensitivity analysis.

Table 1
Parameters and initial data.

Parameters	Definition	Mean value	(Estimated) std	Source
β_0	Transmission coefficient, $\beta_0 = \beta c(1 - \pi)$	0.386	0.054	Calculated
β	Transmission probability of HIV per high-risk behavior	–	–	–
c	Contact rate per year	–	–	–
π	Protection rate due to interventions	–	–	–
ϱ	Modification factor in transmission coefficient of diagnosed HIV-positive individuals	0.491	0.285	MCMC
ϵ	Modification factor in transmission coefficient of AIDS patients	0	–	de Arazoza et al. (2000)
U	Recruitment rate	4.545×10^6	1.281×10^6	MCMC
d	Natural death rate	0.0149 year^{-1}	–	Li et al. (2008)
μ	Exit rate	0.051 year^{-1}	0.029	MCMC
δ	Diagnosis rates	0.304 year^{-1}	0.050	MCMC
ρ	Proportion of HIV-diagnosed individuals who have not yet progressed to AIDS	0.864	0.03	MCMC
ζ	Rate of progression to AIDS	0.116 year^{-1}	–	Database
α_i	Disease-related death for diagnosed HIV-positive individuals	0.172 year^{-1}	–	Database
α_A	Disease-related death for AIDS patients	0.318 year^{-1}	–	Database
$S(0)$	Initial number of susceptibles	7.121×10^6	1.633×10^6	MCMC
$I(0)$	Initial number of undiagnosed HIV-positive individuals	193,940	28,843	MCMC
$D_i(0)$	Initial number of diagnosed HIV-positive individuals	89,766	–	Database
$D_A(0)$	Initial number of AIDS patients	30,577	–	Database

Note: Mean value of the basic reproduction number R_{0N} is estimated as 1.708 with standard deviation 0.054. Mean value for transmission coefficient β_0 is calculated from the formula of R_0 and the estimates of parameters.

We used an adaptive Metropolis–Hastings (M–H) algorithm to carry out the Markov-Chain Monte-Carlo (MCMC) procedure (see parameter estimation in Appendix B), and, after a burn-in period of 500,000 iterations, the next 1,000,000 samples provided our estimates (Haario et al., 2006). Here the Geweke convergence diagnostic method was employed to assess convergence of chains (Geweke, 1992). We chose the year 2005 as a starting point since the consistent surveillance and testing policy has been implemented in mainland China since then. We consider the following two steps to parameterize the national model (1). Firstly, to estimate the mean R_{0N} and its standard deviation without estimating the initial susceptible population size and to reduce the number of parameters that need to be estimated to a minimum, we note that in the case where the size of the high-risk population is large compared to the size of reported HIV/AIDS cases, S/N approximately equals 1 and hence we have the reduced model (B.1) given in Appendix B. Using large sample realization, we initially fit the reduced model (B.1) to data on the numbers of individuals living with HIV or AIDS for the whole country from 2005 to 2009 and get the estimates for the basic reproduction number R_{0N} , the modification factor in the transmission coefficient of the diagnosed HIV-positive individuals ϱ , the diagnosis rate δ , the proportion of HIV-diagnosed individuals who have not yet progressed to AIDS ρ and the initial value of infected individuals $I(0)$. Secondly, once we had estimates for all parameters involved in the reduced model (B.1), the same procedure was used to estimate the rest of the unknown parameters (i.e., $S(0)$, U and μ) by fitting the national model (1) to the data. The mean values of the estimated parameters and their standard deviations are listed in Table 1.

To estimate the province specific parameter values we did not directly parameterize the metapopulation model because there were too many unknown parameters. We estimated such parameters by employing a model for each province (patch) that does not account for the mobility coupling using the same method as in parameterizing the national model (1), based on the surveillance data for each province. Similarly we estimate the reproduction number of disease transmission for each province (R_{0j}), four parameter values (including the δ_j , ϱ_j , ρ_j , and U_j) and the initial data for susceptible and undiagnosed HIV-positive individuals ($S(0)$ and $I(0)$). We kept the rest of the parameters (such as the rate of progression to AIDS, natural death rate, death rate for HIV, and death rate for AIDS) fixed and they were set to be the same as those for the national model listed in Table 1 (see details for parameter estimation in Appendix B). The province specific parameter values are given in Table S1 in the Supplementary Material.

3. Results

We initially analyzed the spatial characteristics of the HIV epidemic in mainland China and the mobility patterns of HIV-positive individuals and AIDS patients. We then focused on understanding the epidemic, predicting evolutionary trends and identifying the key factors that increase the HIV epidemic using our national model and spatially stratified population dynamical model.

3.1. The spatial characteristics of the epidemic

National case-reporting data show that annual reported cases of HIV/AIDS in China are still increasing (Fig. 2A). According to Chinese HIV/AIDS case reports, the number of reported HIV/AIDS cases in 2009 was 56,027, of which 12,633 were AIDS patients. There has been a sharp rise in the annual reported HIV/AIDS cases since 2004 (as shown in Fig. 2A), mainly due to intensive

screening of former plasma donors (FPDs) as well as expanded and strengthened surveillance and testing (CMH, 2005). Note that in the early mid-1990s commercial plasma donation was promoted in rural areas among peasants to supplement incomes and some collectors returned pooled red blood cells from blood type-matched donors after removing the plasma. This unsanitary practice resulted in a large number of HIV infections among the FPDs. Until 2004 mass HIV screenings in known FPD regions were conducted and tens of thousands of infected FPDs were identified (Wu et al., 2006; Dou et al., 2010).

In contrast to the disease pattern in many countries, HIV in China initially spread in rural areas before moving to the urban areas (Cui et al., 2009). HIV infection in China exhibits great variation in patterns among provinces. For example, in 2009, Yunnan, Guangxi, Sichuan and Guangdong had over 5000 reported HIV/AIDS cases, accounting for 57.28% of all reported HIV/AIDS cases (Fig. 2C). The annual reported HIV/AIDS cases in Henan and Yunnan provinces exhibited a sudden increase in 2004 due to the large-scale HIV testing campaign (CMH, 2005), while the annual reported HIV/AIDS cases gradually increased in other provinces such as Jiangsu, Zhejiang and Beijing (Fig. 2D). We define GNH to be the 10 provinces with the greatest number of individuals living with HIV/AIDS (see Table 2), six have over 20,000: Yunnan (61,438), Guangxi (41,111), Henan (32,335), Guangdong (27,305), Xinjiang (24,530) and Sichuan (23,106), accounting for 77.22% of all reported HIV/AIDS cases.

In other provinces, the epidemic is less severe: Qinghai and Tibet had less than 200 reported HIV/AIDS cases by the end of 2009. Estimating the exponential growth rate since 2005 for each province on the basis of the cumulatively reported HIV/AIDS cases quantifies the power of the disease to spread now. The top 10 FGH (currently fastest-growing HIV) provinces are Tianjin, Inner Mongolia, Liaoning, Heilongjiang, Jiangsu, Zhejiang, Chongqing, Sichuan, Tibet and Ningxia (see Table 2). Note that Sichuan and Chongqing belong to both GNH and FGH, and should thus be aggressively targeted.

The difference between the reported province (where a patient has been diagnosed and reported) and current residence for the annual reported HIV cases forms the mobility pattern among provinces (denoted by the network N_{rc}^l and shown in Fig. 3A). Fig. 3B shows the number of diagnosed HIV-positive individuals in 2009 who were diagnosed in region i and are currently living in region j . This represents the number of cases moving from region i to region j and can act as the element of row i and column j of the mobility matrix \bar{G} . Hence, the mobility matrix \bar{G} for HIV-positive individuals can govern the dispersal among 31 provinces, autonomous regions and municipalities.

Note that the mobility network N_{rc}^l is asymmetric, directional and varies in each reported year. The detailed outflow/inflow links or number of immigration/emigration cases for the top 10 provinces are given in Table 3. It shows that Sichuan, Xinjiang, Henan, Yunnan, Guizhou, Chongqing and Guangxi had over 20 sources for immigration (inflow links), while Guangdong, Beijing, Yunnan, Zhejiang, Jiangsu and Shanghai had over 20 emigration destinations (outflow links) in the year 2009. Moreover, Table 3 indicates that Sichuan (resp. Xinjiang, Henan) had 1034 (resp. 313, 264) reported HIV immigration cases, while Guangdong (resp. Yunnan, Zhejiang) had 586 (resp. 268, 257) reported HIV emigration cases (see details in Table 3 and Fig. 3B). Note that Beijing, Shanghai and inshore provinces such as Guangdong, Zhejiang and Jiangsu, which have more outflow links (or cases), belong to the economically developed regions. It follows from Table 2 that the top 10 GNH provinces include Sichuan, Xinjiang, Henan, Yunnan, Guangxi and so on, who have more inflow links (or cases) according to Table 3. By analyzing the number of links/edges of a node (representing a province) and number of

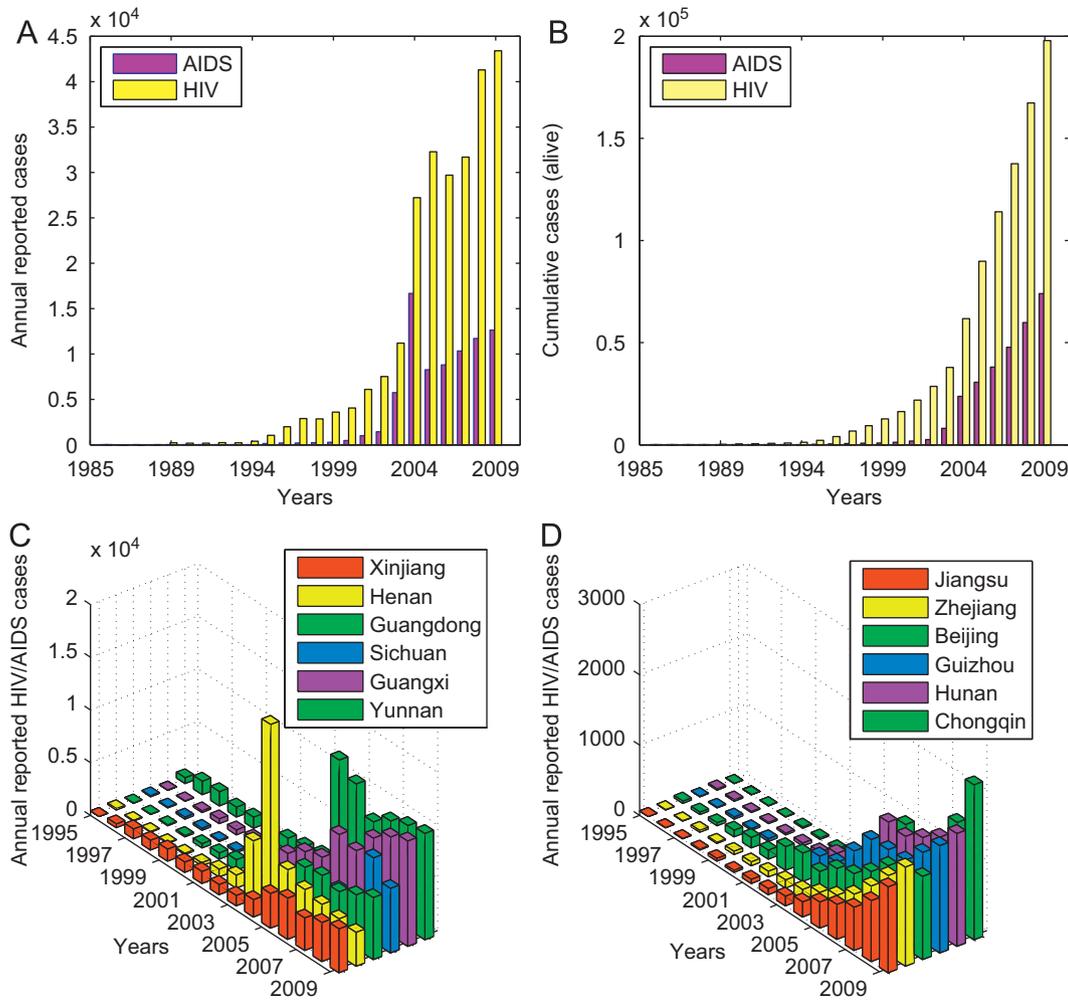


Fig. 2. (A) Annual reported HIV or AIDS cases for mainland China; (B) cumulative number of individuals living with HIV or AIDS in mainland China; (C) annual reported HIV/AIDS cases for high-prevalence provinces and (D) annual reported HIV/AIDS cases for low-prevalence provinces.

Table 2
Four types of regions.

Types	Provinces	Fraction (%)
GNH provinces (top 10)	Beijing, Henan, Hunan, Guangdong, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Xinjiang	86.26
FGH provinces (top 10)	Tianjin, Inner Mongolia, Liaoning, Heilongjiang, Jiangsu, Zhejiang, Chongqing, Sichuan, Tibet	15.42
PMH provinces	Jinlin, Hunan, Guangxi, Sichuan, Guizhou, Yunnan, Xinjiang	60.0
MNS provinces	Henan, Guangxi, Yunnan, Tibet	49.67

Note: GNH, greatest number of HIV/AIDS cases; PMH, positive mobility-induced HIV provinces; FGH, fastest-growing HIV provinces; MNS, mobility nonsensitive provinces. We define a province as mobility nonsensitive if doubling the dispersal rates results in the number of individuals living with HIV/AIDS varying by less than 2%.

immigration/emigration cases of a province for the network N_{rc}^I , we obtained that HIV-positive individuals tend to emigrate from economically developed regions to regions with more numerous HIV cases.

According to the national policy that HIV/AIDS patients are only able to obtain free treatment and drugs at the treatment center local to their registered hometown (Zhang et al., 2005). Also, the rigidity of China’s household system prevents patients from changing their residence status for medical reasons (Jia et al., 2011). As a result, many HIV/AIDS patients return to their registered residence (where they obtained their certification of birth) to get free treatment. Thus, if an AIDS patient’s residence (province i) is different from his/her registered residence (j), he/she has to return to the registered hometown to get free ART. Then, the difference between the registered residence and

the reported province for the annual reported AIDS cases forms the mobility pattern of AIDS patients among provinces (denoted by the network N_{rr}^A , and see details in Fig. 4A and B). By analyzing the network N_{rr}^A we similarly give the detailed outflow/inflow links or number of immigration/emigration cases for the top 10 provinces, see details in Table 4. It follows from Table 4 and Fig. 4B that AIDS patients more likely to move from the epidemically severe areas to economically developed areas. Further, based on the network N_{rr}^A we can get the mobility matrix \bar{C} for the AIDS patients used in Section 2 for the metapopulation model (2).

3.2. The overall estimation for mainland China

Using the next generation matrix (Diekmann and Heesterbeek, 2000; Van den Driessche and Watmough, 2002) we can easily get

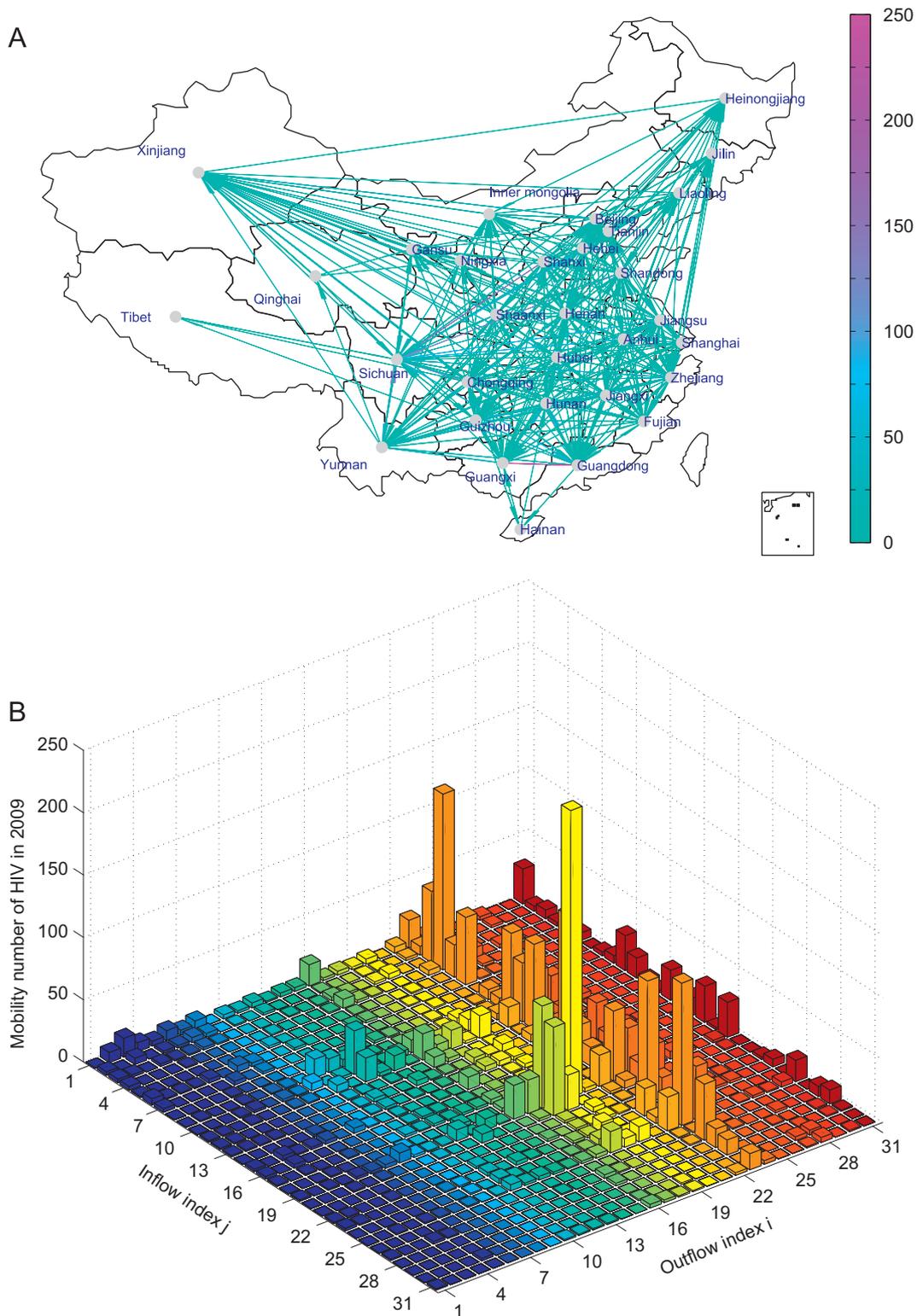


Fig. 3. (A) Mobility network of the reported HIV-positive individuals in the year 2009. Each edge with direction denotes the mobility from the tail of an arrow to the head of an arrow. The color (from blue to purple) of the edges represents the intensity of the mobility flow. (B) Mobility numbers of the reported cases of HIV-positive individuals among 31 provinces, autonomous regions and municipalities. The height of each vertical pole in the plane with outflow index i and outflow index j ($i, j = 1, \dots, 31$) represents the number of cases moving from region i to region j , denoted by \bar{c}_{ij} , and is the element of row i and column j of the mobility matrix \bar{C} for HIV-positive individuals. We index the 31 provinces of mainland China as follows: 1. Beijing, 2. Tianjin, 3. Hebei, 4. Shanxi, 5. Inner Mongolia, 6. Liaoning, 7. Jilin, 8. Heilongjiang, 9. Shanghai, 10. Jiangsu, 11. Zhejiang, 12. Anhui, 13. Fujian, 14. Jiangxi, 15. Shandong, 16. Henan, 17. Hubei, 18. Hunan, 19. Guangdong, 20. Guangxi, 21. Hainan, 22. Chongqing, 23. Sichuan, 24. Guizhou, 25. Yunnan, 26. Tibet, 27. Shaanxi, 28. Gansu, 29. Qinghai, 30. Ningxia, and 31. Xinjiang. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

Table 3
Mobility network of HIV-positive individuals (N_{rc}) in 2009.

Top 10 provinces having most contacts (numbers)		Top 10 provinces having most emigration/immigration (numbers)	
Output	Input	Output	Input
Guangdong (27)	Sichuan (27)	Guangdong (586)	Sichuan (1034)
Beijing (25)	Xinjiang (24)	Yunnan (268)	Guangxi (313)
Yunnan (23)	Henan (23)	Zhejiang (257)	Xinjing (264)
Zhejiang (21)	Yunnan (22)	Jiangsu (219)	Yunnan (228)
Jiangsu (21)	Guizhou (21)	Shandong (192)	Hunan (218)
Shanghai (19)	Chongqin (21)	Shanxi (176)	Guizhou (210)
Hubei (18)	Guangdong (21)	Beijing (166)	Chongqin (131)
Shandong (18)	Guangxi (20)	Guangxi (162)	Henan (131)
Fujian (18)	Shaanxi (17)	Chongqin (139)	Anhui (83)
Sichuan (17)	Hubei (17)	Anhui (101)	Guangdong (76)

the basic reproduction number for the national model (1),

$$R_{0N} = \beta_0 \left(\frac{1}{d + \delta} + \frac{\rho \delta}{(d + \delta)(d + \xi + \alpha_i)} + \frac{\epsilon \delta (\xi + (d + \alpha_i)(1 - \rho))}{(d + \delta)(d + \xi + \alpha_i)(d + \alpha_A)} \right) \quad (3)$$

with transmission coefficient $\beta_0 = (1 - \pi)c\beta$. The epidemiological interpretation of the above formula term-by-term is given in Appendix A. Since the structure of province-specific model is similar to that for the nation model (1), we can similarly derive the patch reproduction number R_0^j for each province j ($j = 1, \dots, 31$) in isolation. Note that provinces with a greater reproduction number have greater exponential growth, and hence belong to the defined FGH provinces. Further, the next generation method applied to the metapopulation model (2) yields an expression for the reproduction number R_{0P} , which is the spectral radius of the matrix $F_P V_P^{-1}$ given in Appendix A. It is interesting to note that the basic reproduction number R_{0P} for the metapopulation system is bounded above and below by the minimum and maximum patch reproduction numbers R_0^j ($j = 1, \dots, 31$), that is,

$$\min\{R_0^j, j = 1, \dots, 31\} \leq R_{0P} \leq \max\{R_0^j, j = 1, \dots, 31\}$$

and approaches a finite value as the rates of movement increase and approach infinity (McCormack and Allen, 2007). So far, whether the limit value of the basic reproduction number R_{0P} for the metapopulation system as mobility rates approach infinity is its counterpart for the homogeneous mixing or not remains unclear, but our numerical studies show that these two values are close to each other.

Based on the data of number of individuals living with HIV (but not AIDS) or AIDS by year and based on our national model (1), we initially estimated the mean reproductive number (R_{0N}) as 1.708 (95% CI 1.440–1.977), the mean transmission coefficient β_0 as 0.386 and mean diagnosis rate δ as 0.304 for the period 2005–2009. Note that, although our model does not include a treated class, our estimate of a near 50% lower infectivity of diagnosed individuals reflects both reduced risk behavior and reduced infectivity due to treatment. The estimates for other parameters are listed in Table 1. Numerical simulations using the aforementioned parameter values give the best fit with data together with the uncertainties (Fig. 5). We illustrate the 50%, 90%, 95% and 99% limits of the posterior uncertainty (from the darkest to the lightest areas) due to the model parameters. The accuracy and fit diagnostics of the model can be studied by considering the width of these regions. Fig. 5 predicts that the estimated number of people living with HIV/AIDS will reach 1,636,200 in 2015 in mainland China if the current surveillance, testing and interventions are unchanged. In addition, the high-risk population size is estimated and shown in Fig. S1 in the Supplementary Material on the basis of model (1).

On the basis of the number of individuals living with HIV (not AIDS) or AIDS by year for each province, we estimated mean values of parameters for each province and derived the goodness of fit for 31 provinces (Fig. 6A–C). We then derived the patch reproduction number R_0^j ($j = 1, \dots, 31$) for each province. Provinces with a greater reproduction number have greater exponential growth, and hence belong to the defined FGH provinces. Further, the next-generation method applied to the mobility network N_{rc}^I for HIV-positive individuals and the network N_{rr}^A for AIDS patients produces $R_{0P} = 2.218$ for the metapopulation model (for the relative mobility for the undiagnosed HIV-positive individuals $\eta = 2$), which is greater than the estimate from the national model (1).

We simulate the aforementioned 31-patch model based on the parameter values for each province and the mobility networks N_{rc}^I and N_{rr}^A obtained above. Fig. 6D shows the estimated number of individuals living with HIV/AIDS (i.e., summation of $I_j + D_{ij} + D_{Aj}$) for the metapopulation model (green curve), 31 isolated models (red curve) and the national model (1) (blue curve). The estimated number of HIV/AIDS individuals will reach 1,524,800 for the metapopulation model and 1,566,800 for the 31 isolated models by the year 2015, which are less than the 1,636,200 predicted by the national model (1). Comparing predictions based on the heterogeneous model with mobility couple (1,524,800) and without mobility coupling (1,566,800) suggests that although the estimated number of total HIV/AIDS individuals becomes less due to mobility coupling, the difference is less than 3% and could be within the range of uncertainty in such simulations. This indicates that mobility coupling barely affects the estimated number of total HIV/AIDS individuals.

3.3. Effect of mobility on the HIV epidemic

Note that population mobility is still rising and will continue to increase in the near future. Generally speaking, economic disparities between geographic areas do cause great increases in population mobility (RCMPD, 2010; Wu et al., 2007). Other important factors that contribute to population mobility are disparities in medical resources and lack of liaison between the medical care systems of different provinces. Moreover, a number of HIV/AIDS cases are lost to follow-up and a number of HIV-positive individuals are undiagnosed according to the nationwide database (CMH, 2009). In addition, according to the new policy on free treatment (CMH, 2009) HIV-positive individuals with $CD4^+$ T cell counts lying between 200 and 350 cells/ μ l could get free drugs in their hometown, which results in an increase in mobility of HIV-positive individuals. Hence, the real mobility rates are much higher than those we obtained from the database. It is then reasonable to consider the effect of increasing movement on the HIV epidemic. To investigate the impact of population mobility on the epidemic in mainland China, we simulated the metapopulation model with 31 provinces with increasing mobility rates, while retaining the mobility network structure.

Numerical simulations show that increasing dispersal rates can either increase the estimated number of individuals living with HIV/AIDS for some provinces (denoted by PMH, positive mobility-induced HIV) such as Sichuan, Hunan, Guizhou and Xinjiang, or decrease the estimated HIV/AIDS cases for other provinces such as Beijing, Shanghai, Zhejiang and Guangdong (see Fig. 7A). Table 5 gives an example that illustrates variations of the estimated number of individuals living with HIV/AIDS by 2015 for some provinces and for the whole country with increasing mobility rates. It follows that the estimated number of individuals living with HIV/AIDS by 2015 for Sichuan increases from 268,170 to 297,880, while the estimate for Shanghai decreases from 24,540 to 22,910 when mobility rates double.

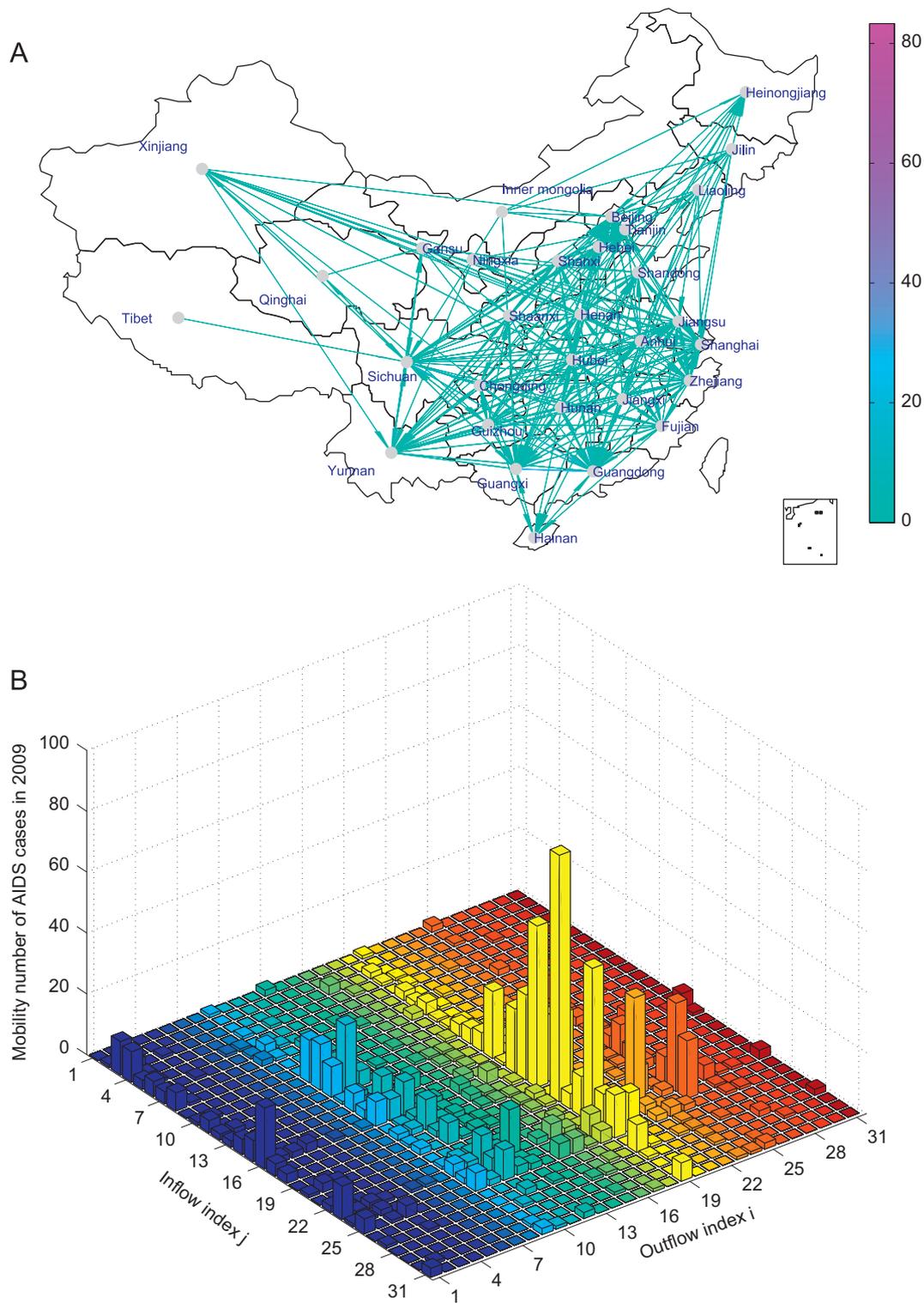


Fig. 4. (A) Mobility network of the reported number of individuals with AIDS in the year 2009. (B) Mobility numbers of the reported cases of AIDS patients among 31 provinces, autonomous regions and municipalities. See Fig. 3 for details and refer to the legend of Fig. 3 for the index number of the provinces in mainland China.

We note that there are some provinces like Guangxi and Yunnan whose estimates slightly change with varying mobility rates, and we call these provinces mobility nonsensitive (MNS) provinces. It is interesting to note that increasing dispersal rates does result in HIV infection declining overall for the whole country. In particular, increasing dispersal rates by 2, 4, 6 and 8 times from the baseline values decreases the estimated HIV/AIDS cases by 1.45%, 3.74%, and 5.38% by the year 2015. This interesting result is

associated with the mobility pattern: outflow/inflow contacts (degrees) and mobility rates. Thus, for a given mobility structure, increasing mobility rates may cause HIV infections to decrease overall, whereas the prospects for increased mobility limiting HIV is quite limited.

Increasing mobility rates from the baseline values while keeping the network structure unchanged can either increase or decrease the estimated HIV/AIDS cases for a particular province

(as shown in Fig. 7), depending on the inflow and outflow of contacts and mobility rates. It follows from Fig. 7B and C that PMH provinces are associated with regions with more inflow contacts or higher movement rates, while an opposite situation holds for the NMH (negative mobility-induced HIV) provinces. Then it is not surprising that, given the network structure, the increasing of mobility rates increases the number of HIV-positive individuals in PMH provinces and decreases the number of

HIV-positive individuals in NMH provinces. Note that Fig. 7A shows that the number of NMH provinces is greater than the number of PMH provinces; moreover, if the majority of HIV-positive individuals move from economically developed areas to the regions with more numerous HIV cases, then the potential to spread is reduced (Xiao et al., 2011). Hence, the overall number of HIV/AIDS cases could be decreased by increasing the mobility rates.

Note that the effect of varying mobility rates is solely induced by the movement of HIV-positive individuals since AIDS patients are assumed not to contribute to infection (i.e., $\epsilon = 0$). When infections caused by AIDS patients are not negligible, it is interesting to examine how increasing mobility rates of AIDS patients influences the overall number of HIV/AIDS cases. This issue is related to the current policy that AIDS patients must return to their registered residence in order to receive free ART (Zhang et al., 2005). Numerical studies (results not shown) indicate that increasing mobility rates of AIDS patients, while keeping the mobility structure unchanged and fixing the movement of HIV-positive individuals, only slightly affects the overall number of HIV/AIDS cases even when the modification factor ϵ in the transmission coefficient of AIDS patients increases and reaches the value of that (ϱ) for diagnosed HIV-positive individuals. This result indicates that our assumption of $\epsilon = 0$ in model equations is reasonable. Moreover, this implies that the current policy does not cause more infection, and hence the question of

Table 4
Mobility network of AIDS patients (N_{rr}) in 2009.

Top 10 provinces having most contacts (numbers)		Top 10 provinces having most emigration/immigration (numbers)	
Output	Input	Output	Input
Sichuan (24)	Guangdong (24)	Sichuan (198)	Guangdong (346)
Yunnan (19)	Beijing (23)	Hunan (111)	Beijing (112)
Henan (19)	Shanghai (20)	Guangxi (106)	Yunnan (110)
Guizhou (16)	Yunnan (19)	Henan (104)	Guangxi (94)
Anhui (15)	Guangxi (19)	Yunnan (84)	Shanghai (84)
Hunan (14)	Zhejiang (16)	Guizhou (64)	Zhejiang (82)
Chongqing (13)	Sichuan (15)	Anhui (56)	Jiangsu (48)
Guangxi (13)	Jiangsu (15)	Jiangxi (54)	Chongqing (43)
Hubei (13)	Hubei (14)	Hubei (48)	Fujian (37)
Shaanxi (10)	Fujian (13)	Chongqing (47)	Sichuan (25)

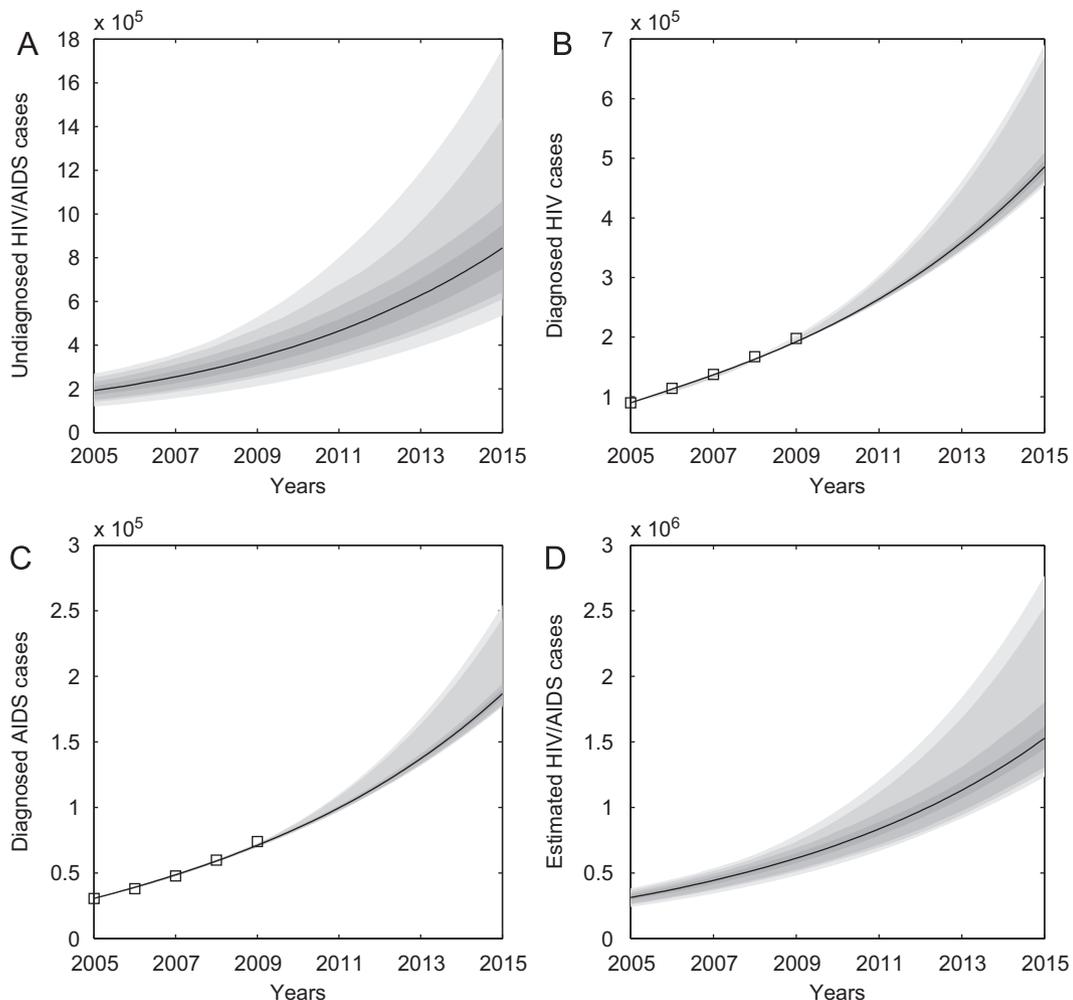


Fig. 5. Goodness of fit and prediction of HIV/AIDS trends, together with uncertainties, until 2015. (A) Undiagnosed individuals living with HIV/AIDS; (B) diagnosed individuals living with HIV; (C) AIDS patients; and (D) estimated number of HIV/AIDS cases. Squares represent the cumulative number of people living with HIV or AIDS by year. The solid lines show the median fits. The areas from the darkest to the lightest correspond to the 50%, 90%, 95% and 99% posterior limits of the model uncertainty.

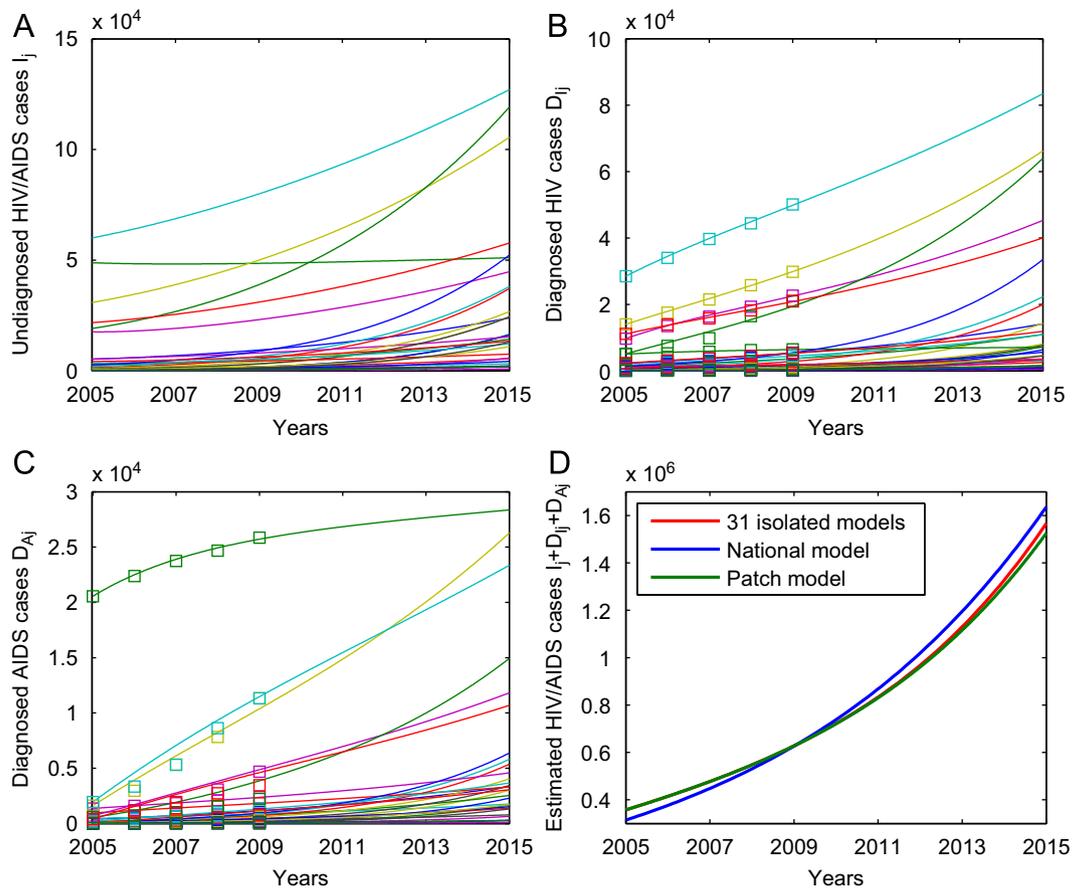


Fig. 6. Goodness of fit and prediction of HIV/AIDS trends for all 31 regions. (A) Undiagnosed individuals living with HIV/AIDS; (B) diagnosed individuals living with HIV; (C) AIDS patients; and (D) estimated number of HIV/AIDS cases. Squares represent the number people living with HIV or AIDS by year. The solid lines show median fits. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

whether AIDS patients should return to their registered residence to get free ART is insignificant, in terms of reducing the overall HIV/AIDS infection.

3.4. Sensitivity analysis

To examine the sensitivity of our results to parameter variation, we used Latin Hypercube Sampling (LHS) and partial rank correlation coefficients (PRCCs) to examine the dependence of R_{ON} and R_0^j , the reproduction numbers for the national and meta-population models, respectively, on each parameter. LHS is a statistical sampling method that allows for an efficient analysis of parameter variations across simultaneous uncertainty ranges in each parameter (Blower and Dowlatabadi, 1994). PRCC, showing which parameters have the largest influence on model outcomes, is calculated using the rank-transformed LHS matrix and output matrix (Marino et al., 2008). We used 1000 simulations per run. A uniform distribution function was used and tested for significant PRCCs for all parameters with wide ranges, such as $d \sim U(0.01, 0.02)$, $\epsilon \sim U(0, 1)$ and other parameters were varied between 0.1 and 1.

Fig. 8 shows the PRCC results which illustrate the dependence of R_{ON} on each parameter. We considered absolute values of PRCC > 0.4 as indicating an important correlation between input parameters and output variables, values between 0.2 and 0.4 as moderate correlations, and values between 0 and 0.2 as not significantly different from zero. The parameters with the most impact on the reproduction number R_{ON} are the transmission coefficient β_0 and the diagnose rate δ . In particular, it is noted that the PRCC value for the modification factor ϵ is less than 0.2, which

indicates that the modification factor in transmission coefficient of AIDS patients, ϵ , has only a small impact on the outcome. Thus, our decision to assume $\epsilon = 0$ is further justified. Moreover, PRCC scatter plots of the reproduction number R_{ON} on each parameter are given in Fig. S2 in the Supplementary Material (Marino et al., 2008). It follows from Fig. S2 that R_{ON} is also highly dependent on changes in the transmission coefficient and diagnosis rate.

Similarly, LHS and PRCCs can also be used to examine the sensitivity of the reproduction number for each province (R_0^j , $j = 1, \dots, 31$) to parameter variations, and we omit here. However, in order to show relatively realistic variation in the reproduction number for each province R_0^j ($j = 1, \dots, 31$) we produce boxplot for R_0^j on the basis of Markov Chains obtained from parameter estimation processes and generate Fig. 9. We also include variation in the reproduction number R_{ON} for the national model based on Markov Chains in Fig. 9 for comparison purposes. It shows that the patch-reproduction number R_0^j ($j = 1, \dots, 31$) for individual regions may have very little or a great deal of variation, while the reproduction number R_{ON} for the national model has only a small variance overall. However, in all cases, the individual reproduction numbers are greater than 1.

3.5. Effect of intervention programs

The exogenous policy interventions aim at increasing the frequency of condom use in sexual acts and improving clean needles programs or methadone treatment, which decreases the transmission coefficient β_0 by increasing the protection rate π . Moreover, strengthening education programs, resulting in a decrease in recruitment rate

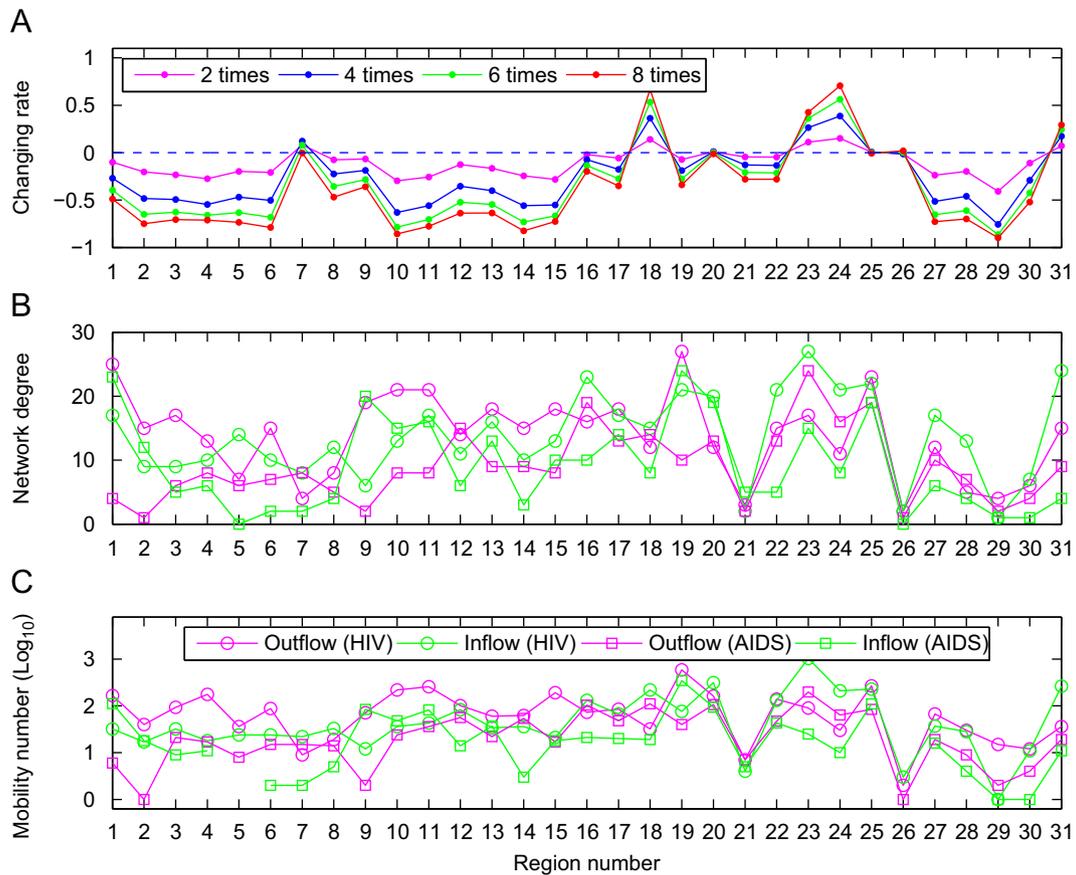


Fig. 7. The effects of increasing the mobility rate on the proportional change in the estimated HIV/AIDS cases for each province. (A) Proportional change in the estimated HIV/AIDS cases when increasing mobility rates by factors of 2, 4, 6 and 8; (B) outflow/inflow edges for each province; and (C) outflow/inflow numbers for each province. Refer to the legend of Fig. 3 for the index number of the provinces in mainland China.

Table 5
The predicted HIV/AIDS cases by 2015 for some provinces and the whole country.

Provinces	Baseline mobility rate M_{ij}^l	$2M_{ij}^l$	$4M_{ij}^l$	$6M_{ij}^l$	$8M_{ij}^l$
Beijing	34,080	30,630	24,910	20,600	17,430
Shanghai	24,540	22,910	19,970	17,570	15,720
Henan	83,190	81,530	77,010	71,900	66,790
Hunan	37,730	43,030	51,430	57,830	82,860
Guangdong	87,340	81,040	70,770	63,150	57,630
Sichuan	268,170	297,880	338,700	364,720	382,390
Yunnan	235,830	237,170	237,510	236,210	234,170
Xinjiang	125,300	134,330	147,130	155,720	161,960
Whole country	1,524,800	1,502,700	1,467,800	1,442,700	1,424,600

U or contact rate c , is another strategy aimed at reducing the high-risk population size and high-risk behaviors. Due to a lack of information on the transmission probability β , contact rate c and the protection rate π we have estimated the transmission coefficient β_0 for simplicity and accuracy, and consequently we mainly vary the transmission coefficient β_0 to assess the impact of strengthening prevention and control strategies on the HIV epidemic.

Initially ignoring spatial disparity in interventions, we investigated how changing certain control parameters affects the evolution of the number of infected individuals over a certain period by simulating the national model (1). Numerical simulations show that strengthening a control strategy (i.e., decreasing β_0) and increasing surveillance and testing (i.e., increasing δ) greatly slow down the speed of HIV evolution (see Fig. S3 in the Supplementary Material).

In particular, the estimated number of individuals living with HIV/AIDS by the year 2015 can be controlled at around 1,112,200 and decreased by 32.01% if we simultaneously decrease the transmission coefficient β_0 by 15% and increase the diagnosis rate δ by 15% (see Fig. S3 in the Supplementary Material). Decreasing the recruitment rate U or increasing the exit rate μ greatly affects the susceptible population size but hardly influences HIV infection (shown in Fig. S4 in the Supplementary Material). This implies that varying the size of the high-risk population has little effect on the HIV epidemic during a certain period. This is partly because the number of infected individuals is sufficiently small compared to the number of susceptibles, so a small variation in susceptibles barely affects the epidemic in the limited time. Another reason is that the HIV epidemic in mainland China is still very much in the exponential growth phase on the basis of the epidemic curve, hence it has not yet reached the critical point and only then will the size of the susceptible pool become a limiting factor.

Considering the significant differences in HIV infection among provinces, interventions should be tailored for each province. To assess the impact and efficacy of spatial-related interventions on China's HIV epidemic, we calculated the variation in the estimated number of PLWHIV for the different control measures. We considered three scenarios: decreasing transmission coefficients β_{0j} (associated with increasing the rate of condom use or methadone treatment) by 10% from the baseline values for (1) all 31 provinces (denoted by the baseline strategy, S1) and (2) only GNH provinces, FGH provinces, or the PMH provinces (denoted by the regional strategy, S2). Strategy S3 was characterized by an additional 10% decrease in the transmission coefficients β_{0j} from the baseline

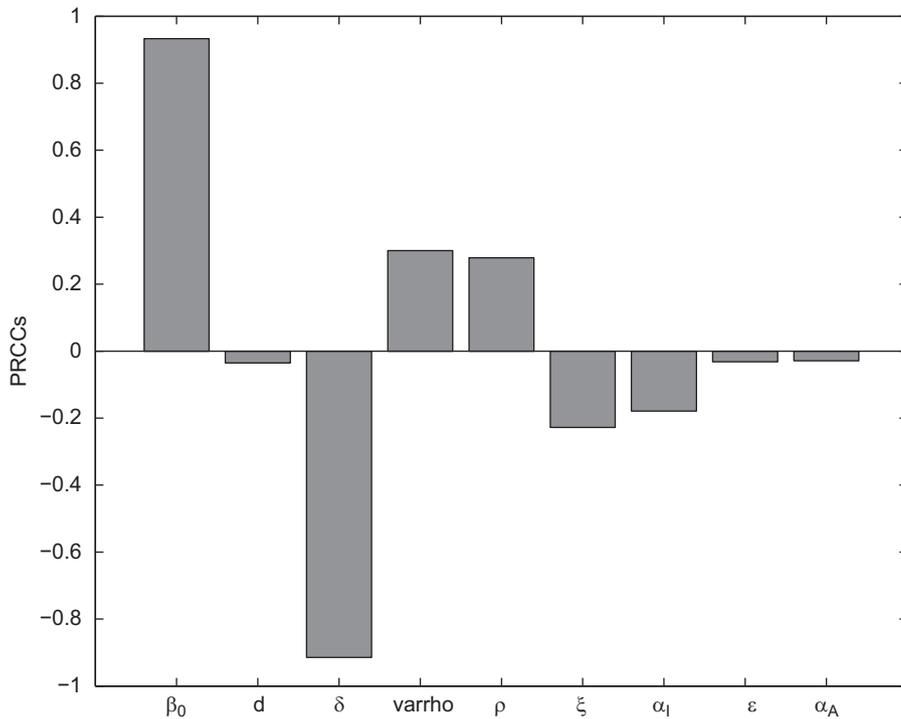


Fig. 8. Partial rank correlation coefficients illustrating the dependence of R_{0N} for the national model on each parameter.

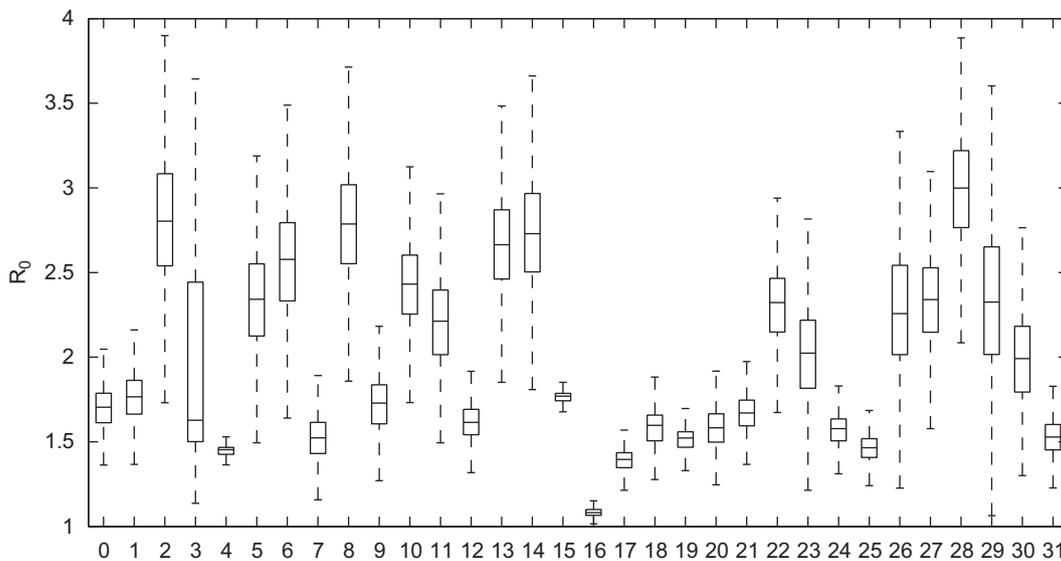


Fig. 9. Boxplot graph showing variation in the reproduction numbers R_{0N} for the national model and the patch reproduction numbers R_0^j ($j = 1, \dots, 31$) for each province. The index zero is reserved for the national model; the remaining indices are listed in Fig. 3.

values for GNH provinces, FGH provinces, or the PMH provinces on the basis on strategy S1.

Using the baseline parameter values we obtained the estimated number of PLWHIV in 2015 as 1,524,800. Under the strategy S1, the model predicts the estimated number of PLWHIV in 2015 will decrease by 18.1%. Strategy S2 indicates that the intervention procedure implemented in the top 10 GNH provinces, the top 10 FGH provinces and all PMH provinces results in the overall estimated HIV/AIDS cases declining by 12.5%, 7.41%, and 9.14%, respectively, by the year 2015, (see details in Table 6). It is interesting to note that the PMH provinces, top 10 GNH provinces and top 10 FGH provinces account for 60%, 86.26%, 15.42% of the total HIV infections (from 2009 surveillance data), respectively. Unsurprisingly, strengthening interventions for all

31 provinces (S1 strategy) is most effective in terms of minimizing the estimated number of PLWHIV. However, strengthening interventions in the top 10 FGH is somewhat effective since it involves relatively fewer high-risk individuals.

Moreover, for the enhanced strategy S3, additional 10% in decreasing transmission coefficients β_{0j} in the top 10 GNH provinces, top 10 FGH provinces, or PMH provinces will lead to a decrease in the number of PLWHIV in the whole country by 27.5%, 23.3% or 25.0%, respectively (shown in Table 6). Comparing strategy S3 with strategy S1 and strategy S2 indicates that in order to mitigate HIV infections in mainland China, a consistent prevention and control plan for all 31 provinces should be proposed, but also we recommend that more attention should be paid to the GNH provinces, FGH provinces and particular PMH provinces if significant mobility occurs in the future.

Table 6
Comparison of different strategies.

Strategies	All provinces	GNH	FGH	PMH	Fraction (%)	Estimated number of PLWHIV in 2015 (f_c)	Percentage decrease ($(f_b - f_c)/f_b$ (%))
S1	✓				100	1,250,200	18.1
S2 ₁		✓			86.26	1,333,800	12.5
S2 ₂			✓		15.42	1,411,700	7.41
S2 ₃				✓	60.00	1,385,400	9.14
S3 ₁	✓	✓				1,105,300	27.5
S3 ₂	✓		✓			1,169,000	23.3
S3 ₃	✓			✓		1,143,500	25.0

Note: ✓ represents decreasing transmission coefficients β_{0j} by 10% in the corresponding regions. We denote the estimated number of PLWHIV in 2015 for the baseline situation by f_b , which is 1,524,800. S1: 10% decrease in transmission coefficient for all provinces; S2_j (j = 1,2,3): 10% decrease in transmission coefficient for GNH, FGH and PMH provinces; and S3_j (j = 1,2,3): additional 10% decrease in transmission coefficient for GNH, FGH and PMH provinces.

4. Discussion

We formulated a national model in which homogeneous mixing is assumed among the entire high-risk population, as well as a metapopulation model in order to study the HIV/AIDS epidemic in mainland China on the basis of surveillance data. The metapopulation model was proposed in order to explore the impact of population mobility and evaluate the effectiveness of spatially relevant interventions. We proposed two modeling structures to estimate HIV/AIDS infection, including estimation of the reproduction number, prediction of the HIV epidemic and evaluation of effectiveness of various intervention strategies, in order to comprehensively assess the HIV/AIDS epidemic in mainland China.

Using the data on the number of individuals living with HIV (not AIDS) or AIDS by year in mainland China, we obtained estimates of the reproduction number, intervention parameter values and the high-risk population size. Our estimated reproduction number from the national model (1) with homogeneous mixing is 1.708 (95% CI 1.440–1.977). Given significant spatial variation of HIV infection, we estimated the reproduction number R_{OP} under heterogeneous mixing as 2.218, which is greater than its value under homogeneous mixing. Note that the estimated reproduction number under heterogeneous mixing decreases and approaches a finite value (close to R_{ON}) as the rates of movement increase and approach infinity (McCormack and Allen, 2007). The discrepancy in the two estimates may be partly attributable to incorporating heterogeneity, which almost always increases the value of the reproduction number in a metapopulation model. The estimates are less than those obtained in a published study of the data among homosexual/bisexual men from European countries: France (3.38–3.81), Western Germany (3.43–4.08) and UK (3.38–3.96) (Nishiura, 2010).

Analyzing the national surveillance data can illuminate the characteristics of mobility patterns of the diagnosed HIV or AIDS cases. The directed mobility network for HIV-positive individuals indicates that these individuals likely move from economically developed regions to regions with more numerous HIV cases (see Table 3). When these individuals develop AIDS, they likely return to their registered residence (i.e., their home town) to receive free ART treatment since, according to central government policy, AIDS patients can only receive free therapy in their registered residences (Zhang et al., 2005). Thus, AIDS patients move in the opposite direction (see Table 4).

This policy on free ART treatment may cause more mobility of AIDS patients among provinces; our results, however, indicate

that increasing mobility rates of AIDS patients barely affects HIV infection, while increasing mobility rates of HIV-positive individuals likely lowers the overall HIV infection rate. Note that an enhancement of mobility of HIV-positive individuals may not be applied as an intervention measure. But this conclusion can suggest that the GNH provinces should be highly targeted due to the mobility characteristics of HIV-positive individuals. Moreover, this result is also associated with the new pattern of community-based HIV/AIDS support and care program introduced by some local governments (especially in epidemically severe areas). This new program results in HIV/AIDS individuals organizing together to self-rescue and self-help, aiming towards improving quality of life for HIV/AIDS individuals (Wang et al., 2007; Zhang et al., 2008). The organization pattern implemented in epidemically severe areas could reduce the outflow/emigrated contacts (or mobility rates) which, according to our results, will likely mitigate the spread of the HIV epidemic.

Ignoring infections caused by AIDS patients would yield lower estimates for the basic reproduction number R_{ON} (or R_{OP}), which may underestimate the disease spread. However, sensitivity analysis (shown in Fig. 8) demonstrates that variations in ϵ have little effect on R_{ON} . Further numerical simulations indicate that when the modification factor ϵ is nonzero, increasing mobility rates of AIDS patients, while keeping the mobility structure unchanged and fixing the movement of HIV-positive individuals, only slightly affects the overall number of HIV/AIDS cases in the near future. Therefore, the assumption of the modification factor ϵ being set to be zero is justified. This result further implies that although many of the infections may not be caused by the current free ART program that AIDS patients should return to their registered residence to get free ART, how to allocate the treatment resources is an important issue for a future study if the central government keeps the current free ART program.

It should be noted that it would be ideal to directly parameterize the metapopulation model. However, there are 248 parameters (8×31 , with eight parameters for each province) that need to be estimated, so we restricted our estimates to the metapopulation model without mobility coupling for each province. Note also that our metapopulation model does not incorporate the mobility of susceptible individuals due to variation in the high-risk population size barely affecting the HIV epidemic (see Fig. S4 in the Supplementary Material), and surveillance data show the mobility rates for HIV-positive individuals or AIDS patients among provinces are extremely small. Therefore, although mobility coupling affects estimates of parameters, its effect is very weak.

Our results have a number of limitations, which should be acknowledged. Estimates for this study were obtained based on the reported (national or provincial) data. Applying them to high-risk groups should be done with caution, since we ignored the specific mode of transmission in our models. Estimates using simplified models ignoring low-risk groups may underestimate the disease spread. Although we have provided rationale for the assumption that AIDS patients are noninfectious, the assumption may not be true for all AIDS patients and will be examined in the future. Our theoretical result that increasing mobility rates results in an overall decline of HIV infections indicates that further investigation of national mobility by policymakers in China is warranted. Although this conclusion may lack generality, it describes the reality for mainland China and suggests which regions should be highly targeted, and supports the community-based HIV/AIDS care programs implemented by some local governments (Wang et al., 2007; Zhang et al., 2008). Despite these caveats, our results strongly suggest that the enhanced interventions and the community-based HIV/AIDS support and care program should be widely implemented in epidemically severe areas to effectively mitigate HIV infections in China.

Acknowledgments

We would like to thank Prof Robert A. Cheke and Prof Hulin Wu for their kind help and comments. We also thank the referees for their careful reading of the original manuscript and many valuable comments that greatly improved the presentation of this paper. This work was supported by the National Mega-project of Science Research No. 2012ZX10001-001 (N.W., Y.X.), by the National Natural Science Foundation of China (NSFC, 81161120403 (S.T.), 11171268 (Y.X.), and by the Fundamental Research Funds for the Central Universities (08143042 (Y.X.), GK 201003001 (S.T.)). R.J.S. is supported by an NSERC Discovery Grant, an Early Researcher Award and funding from MITACS. J.W. is supported by the Canadian Institute of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC), and Mathematics of Information Technology and Complex Systems (MITACS).

Appendix A. The reproduction number of the national model

We calculated the basic reproduction number for the national model using the next-generation matrix (Diekmann and Heesterbeek, 2000; Van den Driessche and Watmough, 2002). This reproduction number, the spectral radius of the matrix $F_N V_N^{-1}$, is given by

$$\begin{aligned}
 R_{ON} &= \rho(F_N V_N^{-1}) \\
 &= \beta_0 \left(\frac{1}{d+\delta} + \frac{\rho\delta}{(d+\delta)(d+\xi+\alpha_I)} \right. \\
 &\quad \left. + \frac{(1-\rho)\delta\epsilon}{(d+\delta)(d+\alpha_A)} + \frac{\rho\delta}{d+\delta} \cdot \frac{\xi}{d+\xi+\alpha_I} \cdot \frac{\epsilon}{d+\alpha_A} \right) \\
 &= \beta_0 \left(\frac{1}{d+\delta} + \frac{\rho\delta}{(d+\delta)(d+\xi+\alpha_I)} + \frac{\epsilon\delta(\xi+(d+\alpha_I)(1-\rho))}{(d+\delta)(d+\xi+\alpha_I)(d+\alpha_A)} \right) \tag{A.1}
 \end{aligned}$$

where the matrices F_N and V_N are

$$F_N = \begin{pmatrix} \beta_0 & \beta_0\varrho & \beta_0\epsilon \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad V_N = \begin{pmatrix} \delta+d & 0 & 0 \\ -\rho\delta & d+\xi+\alpha_I & 0 \\ -(1-\rho)\delta & -\xi & d+\alpha_A \end{pmatrix}$$

Note that each term of the aforementioned expression for R_{ON} has clear epidemiological interpretation. A single HIV infected individual, with the controlled contact rate β_0 and mean duration $1/(d+\delta)$, gives a contribution of $\beta_0/(d+\delta)$. A fraction $\rho\delta/(d+\delta)$ goes from I to the diagnosed class D_I , with the controlled contact rate $\varrho\beta_0$ and mean duration $1/(d+\xi+\alpha_I)$, giving a contribution of $\varrho\beta_0\rho\delta/((d+\delta)(d+\xi+\alpha_I))$. A fraction $(1-\rho)\delta/(d+\delta)$ directly goes from I to the (diagnosed) AIDS class D_A , with the controlled contact rate $\epsilon\beta_0$ and mean duration $1/(d+\alpha_A)$, giving a contribution of $\epsilon\beta_0(1-\rho)\delta/((d+\delta)(d+\alpha_A))$. A fraction $\rho\delta\xi/((d+\delta)(d+\xi+\alpha_I))$ indirectly goes from I to the (diagnosed) AIDS class D_A via class D_I , with the controlled contact rate $\varrho\beta_0$ and mean duration $1/(d+\alpha_A)$, giving a contribution of $\epsilon\beta_0\rho\delta\xi/((d+\delta)(d+\xi+\alpha_I)(d+\alpha_A))$. The sum of these individual contributions is R_{ON} .

Using the next-generation matrix we give an expression for the reproduction number for this metapopulation model,

$$\begin{aligned}
 R_{OP} &= \rho(F_P V_P^{-1}) = \rho(A_1 B_1^{-1} - A_2 C_2^{-1} C_1 B_1^{-1} - A_3 D_3^{-1} \\
 &\quad \times (D_1 B_1^{-1} - D_2 C_2^{-1} C_1 B_1^{-1}))
 \end{aligned}$$

where

$$\begin{aligned}
 F_P &= \begin{pmatrix} A_1 & A_2 & A_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad V_P = \begin{pmatrix} B_1 & 0 & 0 \\ C_1 & C_2 & 0 \\ D_1 & D_2 & D_3 \end{pmatrix} \\
 B_1 &= \begin{pmatrix} \delta_1+d-\eta M_{11}^I & -\eta M_{12}^I & \cdots & -\eta M_{1n}^I \\ -\eta M_{21}^I & \delta_2+d-\eta M_{22}^I & \cdots & -\eta M_{2n}^I \\ \vdots & \vdots & \ddots & \vdots \\ -\eta M_{n1}^I & -\eta M_{n2}^I & \cdots & \delta_n+d-\eta M_{nn}^I \end{pmatrix} \\
 C_2 &= \begin{pmatrix} \zeta_1+d+\alpha_{I1}-M_{11}^I & -M_{12}^I & \cdots & -M_{1n}^I \\ -M_{21}^I & \zeta_2+d+\alpha_{I2}-M_{22}^I & \cdots & -M_{2n}^I \\ \vdots & \vdots & \ddots & \vdots \\ -M_{n1}^I & -M_{n2}^I & \cdots & \zeta_n+d+\alpha_{In}-M_{nn}^I \end{pmatrix} \\
 D_3 &= \begin{pmatrix} d+\alpha_{A1}-M_{11}^A & -M_{12}^A & \cdots & -M_{1n}^A \\ -M_{21}^A & d+\alpha_{A2}-M_{22}^A & \cdots & -M_{2n}^A \\ \vdots & \vdots & \ddots & \vdots \\ -M_{n1}^A & -M_{n2}^A & \cdots & d+\alpha_{An}-M_{nn}^A \end{pmatrix}
 \end{aligned}$$

$$\begin{aligned}
 A_1 &= \text{diag}(\beta_{01}, \beta_{02}, \dots, \beta_{0n}) \\
 A_2 &= \text{diag}(\beta_{01}\varrho_1, \beta_{02}\varrho_2, \dots, \beta_{0n}\varrho_n) \\
 A_3 &= \text{diag}(\beta_{01}\epsilon_1, \beta_{02}\epsilon_2, \dots, \beta_{0n}\epsilon_n) \\
 C_1 &= -\text{diag}(\rho_1\delta_1, \rho_2\delta_2, \dots, \rho_n\delta_n) \\
 D_1 &= -\text{diag}((1-\rho_1)\delta_1, (1-\rho_2)\delta_2, \dots, (1-\rho_n)\delta_n) \\
 D_2 &= -\text{diag}(\zeta_1, \zeta_2, \dots, \zeta_n)
 \end{aligned}$$

Appendix B. Parameter estimation

The variance of measured components, $D(t)$ and $D_A(t)$, was given by an inverse gamma distribution with hyper-parameters (0.01, 4), where 0.01 is the initial error variance which is updated by the inverse gamma distribution (see <http://www.helsinki.fi/~mjlaine/mcmc/>), and the small MCMC package provided in this website was used to estimate the parameters. When estimating unknown parameters and initial values for model (1), the following prior information were given: $\beta_0 \in (0, 1)$, $\varrho \in (0, 1)$, $\delta \in (0, 1)$, $\rho \in (0, 1)$, $I(0) \in (D_I(0)+D_A(0), 5(D_I(0)+D_A(0)))$ with mean $2(D_I(0)+D_A(0))$ and standard variance $0.4(D_I(0)+D_A(0))$; the proposed density was chosen to be a multivariate normal distribution. These ranges were used to ensure good convergence of the MCMC chain.

Note that in the case that the number of susceptibles is large compared to the number of HIV/AIDS cases S/N approximately equals to 1 and hence we have the following reduced model

$$\begin{cases} \frac{dI}{dt} = R_{ON} \frac{I + \varrho D_I + \epsilon D_A}{P_0} - (\delta + d)I \\ \frac{dD_I}{dt} = \rho\delta I - (\zeta + d + \alpha_I)D_I \\ \frac{dD_A}{dt} = (1-\rho)\delta I + \zeta D_I - (d + \alpha_A)D_A \end{cases} \tag{B.1}$$

where

$$P_0 = \left(\frac{1}{d+\delta} + \frac{\rho\delta}{(d+\delta)(d+\xi+\alpha_I)} + \frac{\epsilon\delta(\xi+(d+\alpha_I)(1-\rho))}{(d+\delta)(d+\xi+\alpha_I)(d+\alpha_A)} \right)$$

By fitting this reduced model to the number of individuals living with HIV or AIDS from 2005 to 2009 for the whole country we estimate mean values of R_{ON} , ϱ , δ , ρ and $I(0)$ and their standard deviations. Once we had estimates for these parameters involved in model (B.1), the same procedure was used to estimate the rest of the unknown parameters (i.e., $S(0)$, U and μ) by fitting the national model (1) to the data. Note that here the transmission

coefficient β_0 is actually calculated from the formula of R_{ON} and other parameter estimates.

We also estimate the unknown parameters for 31 isolated provinces in mainland China using the same methods as above based on the number of individuals living with HIV or AIDS for each province from 2005 to 2009. For simplicity, we keep the parameters such as rate of progression to AIDS (ξ_j), natural death rate (d_j), death rate for HIV (α_I), death rate for AIDS (α_A) fixed and set them to be the same as those for the national model listed in Table 1. We estimate the reproduction number of disease transmission (R_{0j}), four parameters such as ρ_j , δ_j , q_j , U_j and initial data for susceptible and undiagnosed HIV-positive individuals ($S_j(0)$ and $I_j(0)$). Note that here we assume the exit rate for each province (μ_j) to be the same and let them be the value for the national model given in Table 1. Since we have no precise information on the size of the number of susceptible individuals for each province, it is impossible to simultaneously estimate the values for recruitment rate U_j or the exit rate μ_j . In fact, the difficulty also arose when parameterizing the national model, but then we could refer to the reference (CMH, 2009) for the estimation of high-risk population to get suitable values for the recruitment rate and the exit rate for the national model. The estimates of parameters for each province are listed in Table S1 in the Supplementary Material.

Appendix C. Supplementary data

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.jtbi.2012.09.037>.

References

- Anderson, R.M., May, R.M., 1991. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press.
- Blower, S.M., Dowlatabadi, H., 1994. Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. *Int. Stat. Rev.* 2, 229–243.
- China Ministry of Health and UN Theme Group, 2005. *A Joint Assessment of HIV/AIDS Prevention, Treatment and Care in China*.
- China Ministry of Health and UN Theme Group, 2011. *A Joint Assessment of HIV/AIDS Prevention, Treatment and Care in China*.
- China Ministry of Health and Joint UN Program on HIV/AIDS WHO, 2009. *Estimates for the HIV/AIDS Epidemic in China*.
- Cui, Y., Adrian, L., Wu, Z., 2009. An overview of the history of epidemic of and response to HIV/AIDS in China: achievements and challenges. *Chin. Med. J.* 122 (19), 2251–2257.
- de Arazoza, H., Lounes, R., Hoang, T., Interlan, Y., 2000. Modeling HIV epidemic under contact tracing. The Cuban case. *J. Theor. Med.* 2, 267–274.
- Diekmann, O., Heesterbeek, J.A.P., 2000. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*. John Wiley, Chichester.
- Dou, Z., Chen, R.Y., Wang, Z., Ji, G., et al., 2010. HIV-Infected Former Plasma Donors in Rural Central China: From Infection to Survival Outcomes, 1985–2008. *PLoS ONE* 5 (10), e13737.
- Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Bernardo, J.M., Berger, J., Dawid, A.P., Smith, A.F.M. (Eds.), *Bayesian Statistics*, vol. 4. Oxford University Press, Oxford, pp. 169–193.
- Haario, H., Laine, M., Mira, A., Saksman, E., 2006. DRAM: efficient adaptive MCMC. *Stat. Comput.* 16, 339–354.
- Heffernan, J.M., Smith?, R.J., Wahl, L.M., 2005. Perspectives on the basic reproductive ratio. *J. R. Soc. Interface* 2 (4), 281–293.
- Jia, Z.W., Wang, L., Chen, R.Y., et al., 2011. Tracking the evolution of HIV/AIDS in China from 1989–2009 to inform future prevention and control efforts. *PLoS ONE* 6 (10), e25671. <http://dx.doi.org/10.1371/journal.pone.0025671>.
- Jin, H., Atkinson, J.H., Xin, Y., et al., 2006. Depression and suicidality in HIV/AIDS in China. *J. Affect. Disord.* 94 (1), 269–275.
- Levin, S.A., 1974. Dispersion and population interactions. *Am. Nat.* 108, 207–228.
- Li, J., Luo, C., de Klerk, N., 2008. Trends in infant/child mortality and life expectancy in indigenous populations in Yunnan Province, China. *Aust. N. Z. J. Public Health* 32, 216–223.
- Lu, L., Jia, M.H., Ma, Y.L., Yang, L., Chen, Z.W., et al., 2008. The changing face of HIV in China. *Nature* 455, 609–611.
- Marino, S., Iain, B., Hogue, I.B., Ray, C.J., Kirschner, D.E., 2008. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J. Theor. Biol.* 254, 178–196.
- McCormack, R.K., Allen, L.J.S., 2007. Multi-patch deterministic and stochastic models for wildlife diseases. *J. Biol. Dyn.* 1, 63–85.
- McCluskey, C.C., 2003. A model of HIV/AIDS with staged progression and amelioration. *Math. Biosci.* 181, 1–16.
- Nishiura, H., 2010. Correcting the actual reproduction number: a simple method to estimate R_0 from early epidemic growth data. *Int. J. Environ. Res. Public Health* 7, 291–302.
- Report on China's Migrant Population Development, 2010. The Department of Services and Management of Migrant Population of National Population and Family Planning Commission of China.
- State Council Document, 1998. *The National Long and Medium Term Plan for Preventing and Controlling the Spread of HIV/AIDS (1998–2010)*, No. 38.
- State Council of the People's Republic of China, 2006. *Regulations on AIDS Prevention and Treatment*. Decree of the State Council of the People's Republic of China, No. 457.
- Smith?, R.J., Li, J., Gordon, R., Heffernan, J.M., 2009. Can we spend our way out of the AIDS epidemic? A world halting AIDS model. *BMC Public Health* 9 (Suppl. 1), S15.
- Tang, S.Y., Xiao, Y.N., Yang, Y.P., et al., 2010. Community-based measures for mitigating the 2009 H1N1 pandemic in China. *PLoS ONE* 5, 1–11. (e10911).
- Van den Driessche, P., Watmough, J., 2002. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.* 180, 29–48.
- Velasco-Hernandez, J.X., Gershengorn, H.B., Blower, S.M., 2002. Could widespread use of combination antiretroviral therapy eradicate HIV epidemics? *Lancet Infect. Dis.* 2, 487–493.
- Wang, G.F., Hu, S.Y., Yuan, F., Li, M.Y., Yang, G.T., 2007. Study on new pattern of community-based HIV/AIDS support and care in Guiyang: self-rescue and self-help. *Chin. J. Nat. Med.* 9 (5), 391–394.
- Williams, J.R., Anderson, R.M., 1994. Mathematical models of the transmission dynamics of human immunodeficiency virus in England and Wales: mixing between different risk groups. *J. R. Stat. Soc. Ser. A* 157, 69–87.
- Wu, Z., Sullivan, S.G., Wang, Y., Rotheram-Borus, M.J., Detels, R., 2007. Evolution of China's response to HIV/AIDS. *Lancet* 369, 679–690.
- Wu, Z., Liu, Z.Y., Detels, R., 1995. HIV-1 infection in commercial plasma donors in China. *Lancet* 346, 61–62.
- Wu, Z., Sun, X., Sullivan, S.G., Detels, R., 2006. Public health, HIV testing in China. *Science* 312, 1475–1476.
- Xiao, Y., Zhou, Y., Tang, S., 2011. Modelling disease spread in dispersal networks at two levels. *IMA J. Math. Appl. Med. Biol.* 28, 227–244.
- Yang, Y., Zhang, K.L., Wang, K.R., Li, Z.R., Yan, H.W., Zhao, Y.T., Gu, Q., Zhang, R.D., 2005. Study on quality of life and related factors among people living with HIV/AIDS. *Chin. J. AIDS STD* 11, 244–246.
- Zhang, F., Pan, J., Yu, L., 2005. Current progress of China's free ART program. *Cell Res.* 15, 877–882.
- Zhang, C.M., Zheng, J.Z., Wang, H.Q., 2008. Thoughts & recommendation on AIDS epidemic and care. *J. Chin. Med. Res.* 8 (4), 309–311.
- Zhu, Q.W., Fang, P.A., Liao, Q., Yang, Y., Cheng, H.Y., 2012. Investigation on related stigma and discrimination for HIV infected and AIDS patients. *Chin. Nurs. Res.* 26, 214–216.