Parameter Estimation of Mathematical Models:
Estimation of the Burden of HIV Epidemics as a Case Study

by

Zhimin Su

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Applied Mathematics

Department of Mathematical and Statistical Sciences

University of Alberta

Abstract

Mathematical models are widely used to describe dynamics in various fields. In practice, it is necessary and important to determine model parameters based on existing data. A major challenge for parameter estimation under modeling framework lies in non-identifiability issue: parameter values on a curve or a multidimensional surface in the parameter space produce almost the same observable model outputs. A variety of techniques and methodologies for resolving non-identifiability have been proposed from different disciplines, such as mathematics, statistics and engineering. The existing methods can inform us whether there is non-identifiability issue or not, if there is, we are suggested to fix some least identifiable parameters such that all remaining parameters can be uniquely estimated. However, it is not always possible to fix some least identifiable parameters such as transmission coefficients in disease models. In this case it is desirable to investigate dependencies among model parameters.

Dependencies among model parameters are related to linear dependencies among the columns of the Jacobian matrix of observable model outputs with respect to model parameters. Due to the existence of numerical error, it is not possible to observe exact linear dependencies among the columns of the Jacobian matrix. Instead, some nearly linear dependencies can be observed. These nearly linear dependencies are the potential exact linear dependencies when numerical error is not present.

In this thesis, a matrix decomposition method was proposed to detect and resolve non-identifiability issue by checking nearly linear dependencies among the columns of

the Jacobian matrix. Our method can inform us how many nearly linear dependencies exist and which columns are involved in each nearly linear dependency.

Our method for diagnosing non-identifiability was applied to several HIV datasets from Chinese Center for Disease Control and Prevention (China CDC) to produce HIV assessment for China. We also demonstrated the applicability of our new method for diagnosing non-identifiability for a simple one-group model and a complex multi-group model.

Preface

The method described in Chapter 3 of this thesis was developed in collaboration with other graduate students in Dr. Li's group (R. de Boer and M. Akinwumi). I was responsible for literature searching, computer simulation and proposing the method. The models described in Chapter 4 and 5 of this thesis were developed as part of an international research collaboration, led by Professor M. Li at University of Alberta, with Professor Y. Shao at Chinese Center for Disease Control and Prevention. I was responsible for model analysis including parameter estimation, sensitivity and uncertainty analysis, and model simulation.

Project 2 in Chapter 4 of this thesis has been published as Z. Su, C. Dong, P. Li, H. Deng, Y. Gong, S. Zhong, M. Wu, Y. Ruan, G. Qin, W. Yang, Y. Shao and M. Li, "A mathematical modeling study of the HIV epidemics at two rural townships in the Liangshan Prefecture of the Sichuan Province of China" Infectious Disease Modeling, 2016. I was responsible for model analysis, model simulations, and results interpretation. W. Yang, Y. Shao and M. Li were responsible for study design and planning. C. Dong, P. Li, H. Deng, Y. Gong, S. Zhong, M. Wu, Y. Ruan, G. Qin, W. Yang, and Y. Shao were responsible for data collection.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this chapter, I will introduce the challenge in parameter estimation—non-identifiability issue, and conduct a literature review for existing methods resolving non-identifiability. At the end of the chapter the outline for this thesis will be discussed.

Ordinary differential equations (ODE) are widely used to describe dynamics in various fields [1-6]. Much attention has been paid to the parameter estimation problem: using data for observable model outputs to estimate model parameters that characterize the system. A major challenge of parameter estimation lies in non-identifiability issue, which means that parameter values on a curve or a multidimensional surface in the parameter space produce almost the same observable model outputs. Identifiability analysis can be found in various fields [7-12], such as mathematics, engineering and statistics. A variety of techniques and methodologies are developed in these disciplines.

1.1 Existing methods for detecting and resolving non-identifiability issue

In reality, before parameter estimation methods being applied to the system to estimate model parameters based on data, a serious obstacle to overcome is how to verify all model parameters are uniquely identifiable. If not all model parameters are identifiable, how can we resolve the problem?

In this section, I will introduce the most common methods used in detecting and resolving non-identifiability issue.

Monto Carlo Method

Monto Carlo method for identifiability analysis requires us to perform parameter estimation repeatedly based on simulated data, if the variance for one model parameter is larger than some threshold, this parameter is considered as non-identifiable [13,14]. In general, a Monte Carlo method for identifiability analysis can be outlined as follows:

i. Determine the value for model parameter p_0, which will be used to form the simulated data. Parameter value for p_0 can be obtained from model fitting using the data we have, or from literature search.

ii. Numerically solve the differential equation with parameter value fixed at p_0 to get the observable model output at the time points we already set.

iii. Generate N simulated data for the observable model output by adding the solution we obtained in step 2 to some noise variable.

iv. Fit the model to the simulated data we obtained in step 3 and define the best-fit parameter as $p_i, i = 1,2, \dots, N$.

v. Calculate the relative estimation error (ARE) for each parameter using the following definition.

$$\mathrm{ARE} = 100\% * \frac{1}{N}\sum_{i=1}^{N} \frac{\left|p_0^k - p_i^k\right|}{\left|p_0^k\right|},$$

where p_0^k is the k-th parameter in p_0 and p_i^k is the k-th parameter in p_i.

It is definitely true that ARE for each parameter will increase if we raise the level of the noise variable. If we vary the noise variable within some reasonable domain, some parameters would have a very large ARE while others would not. Commonly, before performing the Monte Carlo method, people would define some threshold, if the ARE for some parameter is larger than the threshold, the parameter is considered as non-identifiable, and otherwise the parameter is identifiable. There is no common rule for setting the value for the threshold, therefore the final decision is highly dependent on the investigator's experience.

Graph Method

The main idea for the graph method is to explore whether the error between observable model output and the data flattens out when we change model parameter values. We change one parameter's value within its domain at one time, for each sample value of the above parameter, if we can find corresponding values for the remaining parameters such

that the error mentioned above flattens out, this parameter is considered as non-identifiable [15]. The procedure can be summarized as following:

i. Determine the range for each model parameter $p_{i,}$ $i = 1, 2, \dots, m$.

ii. Assume each parameter following uniform distribution in the corresponding range.

iii. Pick up N samples for each parameter from the corresponding uniform distribution $p_i^k, k = 1, 2, \dots, N$.

iv. Fix the value of parameter p_i at $p_i^k, k = 1, 2, \dots, N, i = 1, 2, \dots, m$, fit the model to data to obtain best-fits for all remaining parameters and calculate the least error.

v. Draw the graphs for least errors vs parameters. The x-axis is the value for parameter p_i and the y-axis is the corresponding least error.

If for parameter p_i, with the variation of p_i, the least error flattens out, then parameter p_i is not identifiable. If for parameter p_j, with the variation of p_j, the least error fluctuates in a large range, then parameter p_j is identifiable.

Correlation Method

The central idea of the correlation method is to investigate the dependency between two model parameters by calculating the correlation coefficient of two columns of the Jacobian Matrix J of observable model outputs with respect to model parameters at discrete time points [16-20]. The correlation coefficient is given by:

$$\mathrm{corr}(J_{.i}, J_{.j}) = \frac{\mathrm{cov}(J_{.i}, J_{.j})}{\sigma(J_{.i})\sigma(J_{.j})},$$

where $J_{.i}, J_{.j}$ are the i-th and j-th column of J. If the calculated correlation coefficient between any two columns is close to one, these two corresponding parameters can not be identified together. However, such a conclusion involves two parameters. Is it possible to determine which parameter in this pair is more non-identifiable and should be fixed

before model fitting? Quaiser and Mönnigmann [21] proposed the concept of total correlation to solve this question,

$$c_i^{tot} = \sum_{j=1, j\neq i}^{m} |\,\mathrm{corr}(J_{.i}, J_{.j})| * I\left(\left|\mathrm{corr}(J_{.i}, J_{.j})\right| \geq 1-\delta\right),$$

where I denotes the indicator function, and $\delta \in (0,1)$ is the cut-off value specified by the user. The parameter with the highest total correlation is most non-identifiable one and is the first candidate to be fixed from the model. At last, we should fix all parameters' values whose total correlation exceeds some threshold.

Orthogonal Method

The basic idea of this approach is to examine the dependencies among model parameters by calculating the perpendicular distance from one column vector in J to the space spanned by other column vectors in J [22-24]. The orthogonal method can be outlined in the following algorithm:

i. $J_I = \emptyset, J_U = \{J_{.i}\}, i = 1, 2, \ldots, m, k = 1.$

ii. $\| J_{.r1} \|_2^2 = \max_{i\in\{1,2,\ldots m\}}(\| J_{.i} \|_2^2),\ J_I = \{J_{.r1}\}, J_U = J_U \backslash \{J_{.r1}\}.$

iii. For each vector $J_{.rh} \in J_U$, calculate ${J_{.rh}}^{\perp} = J_{.rh} - {J_{.rh}}^{proj(J_I)}$, where ${J_{.rh}}^{proj(J_I)}$ is the projection for vector $J_{.rh}$ onto the space spanned by all vectors in J_I, k=k+1.

iv. If ${J_{.rk}}^{\perp} = \max {J_{.rh}}^{\perp} > \delta$, $J_h \in J_U$, $J_I = \{J_{.r1}, \ldots, J_{.rk}\}$, $J_U = J_U \backslash \{J_{.rk}\}$, return to step 3; If ${J_{.rk}}^{\perp} = \max {J_{.rh}}^{\perp} < \delta$, stop.

Before performing the iteration, the cut-off value δ is specified, if the perpendicular distance from one vector to the space spanned by all vectors in J_I is less than δ, the corresponding parameter is considered as non-identifiable. Instead of using the perpendicular distance as a standard to make a decision, we can also employ the angle between one vector with a space. If the angle is too small, the corresponding parameter is non-identifiable.

Since there is no common rule for determining the cut-off value δ , the number of non-identifiable parameters is highly dependent on the operator's experience. Therefore,

instead of determining the non-identifiable parameters set, Quaiser and Mönnigmann [21] proposed to rank all the parameters based on the values of norms or angles. Based on that ranking we can choose to fix several least non-identifiable parameters, so that all remaining ones are identifiable.

Eigenvalue Method

Non-linear least squares or maximum likelihood parameter estimation problem amounts to minimize the function:

$$\phi(p) = \frac{1}{2}\sum_{j=1}^{q}\sum_{i=1}^{n}(y_j(t_i, p) - \gamma_j(t_i))^2, \tag{1.11}$$

where $y_j(t_i, p)$ is observable model output $y_j$at time t_i with parameter value p, $\gamma_j(t_i)$ is the data at time t_i corresponding to observable model output y_j.

The Hessian matrix of equation (1.11) can be approximated by

$$H_{kl} = \sum_{i,j}\frac{\partial y_j(t_i, p)}{\partial p_k}\frac{\partial y_j(t_i, p)}{\partial p_l},$$

$$=(J^T J)_{kl}.$$

Assume λ_r and u_r are r-th eigenvalue and the corresponding eigenvector of matrix H respectively. The eigenvalues are ordered $\lambda_1 > \lambda_2 > \cdots > \lambda_m$, and the eigenvectors are normalized with $u_r^T u_r = 1$.

If p^* is the parameter value that minimizes equation (1.11), we consider the change of $\phi(p)$ in a direction αu_r for some real constant α

$$\phi(p^* + \alpha u_r) = \phi(p^*) + \frac{1}{2}\alpha^2 u_r^T H u_r$$

$$= \phi(p^*) + \tfrac{1}{2}\alpha^2 \lambda_r,$$

which implies that if some λ_r is equal to 0, $\phi(p)$ will not change in direction u_r. While due to the existence of numerical error, exact 0 eigenvalue is seldom observed, instead are some very small eigenvalues.

The existence of a small eigenvalue λ_r implies one direction to which $\phi(p^*)$ almost does not change. The largest entry in u_r implies the parameter that is most non-identifiable, since a huge jump for that parameter has little effect on $\phi(p^*)$ [25-27].

Based on the above theoretical analysis, the algorithm proceeds as follows:

i. Set I={1,…,m}, and $U = \emptyset$.

ii. If I is empty, stop. All model parameters are non-identifiable.

iii. Fix the parameters $p_k, k \in U$ to the searched literature values and consider only the $p_k, k \in I$ to be the variable.

iv. Using least squares method to obtain the estimate for parameters $p_k, k \in I$, and set it as p^*. Calculate the eigenvalues λ_j and eigenvectors u_j for the corresponding J^TJ. Assume the eigenvalues are ordered $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{n_I}$ and the eigenvectors are normalized.

v. If $\lambda_1 \geq \varepsilon$, stop. All parameters p_k for all $k \in I$ are identifiable and others are non-identifiable.

vi. If $\lambda_1 < \varepsilon$, select k such that $|u_k^1| = \max(|u_1^1|, |u_2^1|, \ldots, |u_{n_I}^1|)$. Remove k from the set I, add k to the set U, and return to step 2.

Here ε is the threshold for identifying whether the eigenvalue is small or not, and it is commonly determined by the operator.

The parameters are ordered from least identifiable to most identifiable as the sequence they are removed from set I. We can choose to fix all parameters' values in set U, so that all remaining parameters in set I can be uniquely determined.

For the Monte Carlo and Graph methods, they will inform us whether there is non-identifiability issue or not, if there is, they can inform us which parameters are non-identifiable. However, no information is obtained about the dependencies among model parameters. Therefore, we have to select different model parameters' combinations, fixing the remaining parameters at some reasonable values, to see if all parameters in the

combination are identifiable using these two methods. The computational cost for detecting and resolving non-identifiability issue using these two methods is very high.

For the correlation method, it checks the linear dependency between two columns of J, it can not detect the linear dependency among more than three columns of J. In some cases, this limitation will prevent us from fully identifying the non-identifiable model parameters.

For the orthogonal and eigenvalue methods, both of them can inform us of the ranking of model parameters from least identifiable to most identifiable one. We are suggested to fix some of the least identifiable model parameters such that all the remaining ones can be uniquely estimated. However, it is not always possible to fix some least identifiable parameters such as transmission coefficients in disease models. In this case, it is desirable to determine which parameters are involved in each dependency. For each dependency, we can choose to fix one parameter's value which can be searched from literature, so that the corresponding dependency disappears, and the remaining parameters can be uniquely estimated.

In my thesis, I will describe a new method for detecting non-identifiability in model fitting, the Matrix Decomposition Method, which will avoid the shortcomings of the previous methods as described above. The new method will inform us of:

- The number of parameters that are non-identifiable for a given dataset.
- Relationships among the non-identifiable parameters.

1.2 Thesis outline

In the remainder of the thesis, a new method for resolving non-identifiability issue arising from parameter estimation procedure in modeling work will be developed and applied to various HIV datasets from China CDC. In chapter 2, we introduced non-identifiability issue in modeling work and the general procedure for parameter estimation. In chapter 3, our method for resolving non-identifiability issue was discussed. Our method includes two steps: singular value decomposition, which will inform us at most how many

parameters can be fitted together, and variance decomposition, which will inform us of the flexibility we have in fixing model parameters' values. The singular value decomposition and variance decomposition technique was first introduced to deal with the co-linearity problem in multiple linear regression. They have not been applied to deal with non-identifiability problem before. In chapter 4, our method was applied to a single group SIDT model to fit various HIV datasets from China CDC to estimate model parameters. In chapter 5, our method was applied to a multi-group SIDT model to fit various HIV datasets. Although the model becomes more complicated, our method was still effective in resolving non-identifiability issue, and the applicability of our method is demonstrated. Finally, chapter 6 contains conclusions as well as future directions for our work.

In this chapter, I talked about the challenge problem in parameter estimation—non-identifiability issue, and introduced various existing methods for resolving non-identifiability. Some of the existing methods have high computational cost since we need to perform model fitting repeatedly. Others can inform us whether there is non-identifiability issue or not, if there is, they can output a ranking for model parameters from the least identifiable one to the most identifiable one, and we are suggested to fix some least identifiable parameters such that all remaining model parameters can be estimated uniquely. While it is not always possible to fix some least identifiable parameter such as transmission coefficient in disease models. In this case, it is desirable to investigate the dependencies among model parameters. At the end of this chapter I introduced the outline of my thesis and our method for resolving non-identifiability issue briefly. Our method included two steps, the first step can inform us how many dependencies existing among model parameters and the second step can inform us which parameters are involved in each dependency. In this way, instead of fixing some least identifiable parameters, we can choose to fix parameters whose values can be searched

from other sources such that all dependencies disappear and all remaining parameters can be uniquely fitted.

Chapter 2

General Procedures for Parameter Estimation and Non-identifiability

In this chapter, I will introduce the general procedure for parameter estimation, and the challenge problem in the general procedure—non-identifiability issue.

An ordinary differential equation model can be expressed as a system of nonlinear differential equations:

$$\frac{dx}{dt} = f(x(t), u(t), t, \theta), \tag{2.1}$$

$$y(t, \theta) = h(x(t), u(t), \theta), \tag{2.2}$$

where $t \geq 0$ is the time variable, $f \in C^1(R^d \times R^j \times R_+ \times R^m)$ is the vector field, $x \in R^d$ represents the state of the system, $\theta \in R^m$ is an m dimensional vector of parameters, $y \in R^q$ is the observable model output vector and $u \in R^j$ is the model input vector. For the parameter estimation problem, θ is unknown and has to be estimated based on the data about y at discrete time points. Parameter estimation is a crucial step in using modeling method to perform estimation and prediction.

2.1 General procedures for parameter estimation

Step 1: Produce point estimation for model parameters using the Nonlinear Least Squares method. Parameter values for which observable model outputs y best fit the data can be selected using the least squares routine in MATLAB by minimizing the error between observable model outputs y and the data [28].

Step 2: Produce 95% confidence intervals for model parameters using the Bayesian-based Markov chain Monte Carlo (MCMC) method [29].

Posterior distribution of parameter θ_i , which is the i-th component of parameter vector θ was obtained from data γ using formula

$$p(\theta_i | \gamma) = \frac{p(\gamma|\theta_i)p(\theta_i)}{\int p(\gamma|\theta_i)p(\theta_i)d\theta_i}, \tag{2.3}$$

where $p(\theta_i)$ is the prior distribution for the parameter θ_i, $p(\gamma|\theta_i)$ is the likelihood function for observing data γ given the value of parameter θ_i.

Computing the likelihood $p(\gamma|\theta_i)$ requires some additional assumptions. Commonly it is assumed that the errors introduced in collecting data are normally distributed, that is,

$$\gamma = y(\theta) + \varepsilon,$$

where $y(\theta)$ is model output given the value of parameter θ , ε is a random variable with $\epsilon \sim MN(0, \Sigma)$, $\Sigma = \mathrm{diag}(\sigma^2, \sigma^2, \cdots, \sigma^2)$. Values for the remaining parameters were fixed at the point estimation obtained from Step 1. It is also customary to assume that σ^2 is a random variable itself with $\sigma^2 \sim \mathrm{Inv}\,\Gamma(\alpha, \beta)$. Based on these assumptions, the likelihood function in the posterior can be written as

$$\begin{aligned} p(\gamma|\theta_i) &= \int p(\gamma|\theta_i, \sigma^2)\, p(\sigma^2) d\sigma^2 \\ &= \int \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{SSE(\theta_i)}{2\sigma^2}} \frac{\beta^\alpha}{\Gamma(\alpha)}\ \sigma^{-2(\alpha-1)}\ e^{-\frac{\beta}{\sigma^2}}\ d\sigma^2 \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{2\pi}\right)^{\frac{N}{2}} \frac{\Gamma(\frac{N}{2}+\alpha)}{\left(\frac{SSE(\theta_i)+2\beta}{2}\right)^{(\frac{N}{2}+\alpha)}}\ , \end{aligned} \tag{2.4}$$

where $SSE(\theta_i)$ is the sum of squared errors between data γ and model output $y(\theta)$, and N is the number of data points.

If the differential equation model is highly nonlinear, and a close form of $p(\theta_i|\gamma)$ can not be obtained analytically. We resort to numerical solutions, using the Metropolis-Hastings algorithm [30] of the Markov chain Monte Carlo (MCMC) sampling, to approximate the posterior distribution.

Metropolis-Hastings algorithm has a starting point $\theta_{i(0)}$, a new sample point $\theta_{i(1)}$ is selected from a specified proposal distribution. This new sample point is either accepted or rejected based on the proposal distribution and the posterior distribution $p(\theta_i|\gamma)$. Subsequent sample points $\theta_{i(i)}$ are selected using $\theta_{i(i-1)}$ as the starting point, for i = 1,2,3 ….

The ratio used to determine whether the new sample is accepted or rejected is

$$A(\theta_{i0}, \theta_{i1}) = \frac{p(\theta_{i1}|\gamma)f(\theta_{i0}|\theta_{i1})}{p(\theta_{i0}|\gamma)f(\theta_{i1}|\theta_{i0})}, \quad (2.5)$$

where $p(\theta_i|\gamma)$ is the posterior distribution to be sampled, and $f(\theta_{i1}|\theta_{i0})$ is the proposal distribution given the starting point θ_{i0}. The point θ_{i1} is accepted with probability

$$\min\big(1, A(\theta_{i0}, \theta_{i1})\big).$$

Substituting equation (2.3) into (2.5) results in:

$$A(\theta_{i0}, \theta_{i1}) = \frac{p(\gamma|\theta_{i1})p(\theta_{i1})f(\theta_{i0}|\theta_{i1})}{p(\gamma|\theta_{i0})p(\theta_{i0})f(\theta_{i1}|\theta_{i0})} \quad (2.6)$$

Proposal distribution determines the size of jump the algorithm will have when taking new sample points. If the size of the jump is too small, the new sample will always be accepted, but it will take a long time to obtain enough samples covering the posterior distribution. If the size of jump is too large, the new sample will always be rejected. We need to choose a proposal distribution to balance these two concerns.

For our studies, we assume both the proposal distribution $f\big(\theta_{i(i)}|\theta_{i(i-1)}\big), i = 1,2,3 \ldots$ and the prior distribution follow Gamma distribution, with its mean fixed at $\theta_{i(i-1)}$ and the point estimation we obtained for θ_i, respectively [30].

Step 3: Perform uncertainty analysis for model predictions using Latin Hypercube Sampling (LHS) method [31]. Baseline model predictions are computed using the best-fit parameter values. To estimate uncertainty in model predictions, prediction intervals with high and low estimates are produced by varying parameters within their confidence intervals. The LHS method, which partitions the probability density function into N regions with equal area and randomly picks one sample in each region, is commonly used to produce samples of parameter values. LHS method can cover the distribution that is being sampled with relatively few sample points. In our studies, we assumed parameters following triangle distributions with peaks fixing at the point estimation we obtained in Step 1, lower bound and upper bound fixing at the those bounds we obtained in Step 2. Prediction intervals contain 95% of the model outputs using sampled parameter values.

Step 4: Perform sensitivity analysis. For each of the important model outcomes, local sensitivity analysis can be carried out using one-at-a-time method [32, 33] at the best-fit

parameter values, which changes one parameter at a time to estimate the effect on the model outcome. The analysis identifies model parameters that are highly sensitive and that are not sensitive. Based on the analysis, we can design interventions to change the value for highly sensitive model parameters, so that model outcomes we are interested in can vary a lot. The analysis can also inform us of model parameters that are not sensitive, which are the candidates to be fixed before model fitting.

For each of the important model outcomes, global sensitivity analysis can also be carried out by calculating the partial rank correlation coefficient [31, 34] between parameters and model outcome. The use of partial correlation discounts the effect of other parameters, and rank correlation assesses the strength of monotonicity between two variables. Global sensitivity analysis is a supplement to local sensitivity analysis, since it can indicate whether the behavior we observed in local sensitivity analysis can be extended over a range of parameter values.

Step 5: Conduct model validation using coefficient of determination and leave-m-out cross-validation method. In the Nonlinear Least squares fitting, goodness of the fit is measured using the R^2 ratio

$$R^2=1-\frac{E(y,\hat{y})}{E(y,\bar{y})},$$

where $E(y,\hat{y})$ is the mean squared error between data y and model output $\hat{y}$, $E(y,\bar{y})$ is the error between data y and the mean of data $\bar{y}$ [28]. A high rating R^2 implies an excellent fit between model output and data. When the leave-m-out cross-validation is applied, m data point is randomly selected and set aside, the remaining data points are used for model fitting and the data point set aside is used for model validation [35]. The consistency between model predictions and the validation point provides a validation for our model.

2.2 Non-identifiability in parameter estimation

A major challenge in the above procedure lies in Step 1 and 2, due to the non-identifiability issue: infinite parameter values giving almost the same observable model

outputs. Non-identifiability is common in modeling infectious diseases with an asymptomatic state, such as HIV, Hepatitis B, influenza, since no data is usually available on asymptomatic infections [36]. This phenomenon often occurs because multiple parameters have similar effects on the observable model outputs. But they do not necessarily result in the same model behaviors. In fact, the behavior of unobserved parts of the model can be very different.

Although authors differ in their definitions of non-identifiability, all capture the same ideas. In this thesis, we will focus on the following definition of non-identifiability [37].

Definition 2.1: A parameter p_i , i=1, 2,…m, of model (2.1-2.2) is called locally structurally identifiable at p^* if, there exists a neighborhood $V(p^*)$, for all model admissible input $u(t)$, and all admissible values $p' \in V(p^*)$,

$$y(t, p^*) = y(t, p')$$

implies

$$p_i{}^* = p_i{}'.$$

Parameters $p_i, i = 1,2, \dots, m$ are non-identifiable if they are not identifiable.

In this chapter, I first talked about the general procedure for parameter estimation. The Nonlinear Least squares method was used to obtain the point estimation of model parameters, the Bayesian-based Markov chain Monte Carlo (MCMC) method was used to obtain the 95% confidence intervals of model parameters, Latin Hypercube Sampling (LHS) method was used to perform uncertainty analysis for model predictions, one-at-a-time method and calculating the partial rank correlation coefficient method were used to perform sensitivity analysis for model parameters, and leave-m-out cross validation method was used to validate the model. Challenge for the general procedure lies in obtaining the point estimation for model parameters due to the existence of non-

identifiability issue: infinite model parameters' values have almost the same observable model outputs.

Chapter 3

A New Method for Diagnosing Non-identifiability

In this chapter, I will introduce our method, the Matrix Decomposition Method, for detecting and resolving non-identifiability issue.

Data is typically available at finite time points $t_1,\ t_2, \cdots t_n$. Equation $y(t, p) = y(t, p^*)$ in Definition 2.1 at finite time points $t_1,\ t_2, \cdots t_n$ is a system of nonlinear equations of the form:

$$\begin{aligned} y(t_1, p) &= y(t_1, p^*)\,, \\ y(t_2, p) &= y(t_2, p^*)\,, \\ &\cdots\cdots \\ y(t_n, p) &= y(t_n, p^*)\,. \end{aligned} \tag{3.1}$$

The Jacobian matrix J of $y(t_i, p), i = 1,2, \cdots, n$, with respect to parameter p at p^* is:

$$J = J(t_1,\ t_2,\ldots, t_n, p^*) = \frac{\partial y}{\partial p}\ (t_1,\ t_2,\ldots, t_n, p^*)\,,\ \ p \in R^m\,, y \in R^q.$$

By the Implicit Function Theorem, if matrix J has full rank m, then the columns of J are linearly independent, and the system (3.1) has a unique solution p in a neighborhood of p^*. As a result, parameter p can be uniquely determined and the parameter p is locally identifiable at p^*. If $\text{rank}(J) = r < m$, then there are $m - r$ linear dependencies among the columns of J, and solutions to (3.1) form a $m - r$ dimensional sub-manifold of the parameter space R^m in a neighbourhood of p^*. We need to fix the values of $m - r$ parameters, to uniquely determine the remaining r parameters. In this case, it is desirable to determine which columns are involved in each linear dependency. For each linear dependency, we can choose to fix one parameter's value, so that the corresponding linear dependency disappears, and the remaining r parameters can be uniquely estimated.

Numerical error and noise in data can interfere with detection of exact linear dependencies among columns of J. Our method, the Matrix Decomposition Method, are

designed to overcome this difficulty. The singular value decomposition of matrix J can detect how many nearly linear dependencies exist among the columns of J, and the variance decomposition can inform us which columns are involved in each nearly linear dependency. These nearly linear dependencies are the potential exact linear dependencies in the Jacobian matrix J when numerical errors and noise are not present.

3.1 Singular value decomposition for the Jacobian matrix J

Singular value decomposition is the factorization of a matrix. Let A be a real m*n matrix, where $m \geq n$. A can be decomposed as:

$$A = U\Sigma V^T,$$

where U is $m * n$, Σ is $n * n$, V is $n * n$, $U^T U = V^T V = VV^T = I_n$, and $\Sigma = \mathrm{diag}(\sigma_1, \dots, \sigma_n), \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. The diagonal entries $\sigma_i, i = 1,2, \dots, n$ of Σ are called singular values of A, and the decomposition is known as singular value decomposition of A [38].

Singular value decomposition has various applications, such as computing the pseudoinverse, signal processing and determining the rank, range, null space of a matrix.

The singular value decomposition for Jacobian matrix J,

$$J = M * S * N^T,$$

where J is nd*m, M is nd*m and column orthogonal, N is m*m, both row and column orthogonal, and $S = \mathrm{diag}\{\mu_1, \mu_2, \dots, \mu_m\}$ is a diagonal matrix of singular values of J, with $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_m$.

It can be shown that $J^T J = NSM^T MSN^T = NS^2 N^T$, or postmultiplying by N, that $J^T JN = NS^2$. We recognize that the columns of N are eigenvectors of $J^T J$, and the diagonal elements in S are the positive square roots of the eigenvalues of $J^T J$. Because the columns of N are orthogonal, $N_i^T J^T JN_i = N_i^T \mu_i^2 N_i = \mu_i^2$, where N_i is the i-th column of N. Therefore, $\| J * N_i \| = \mu_i$, each small μ_i identifies a nearly linear dependency, and the elements in N_i inform us of the columns that are involved in this near dependency. By convention, μ_i is considered as small if the corresponding condition index $v_i = \mu_1/\mu_i > 30.$.

The limitation for using N_i to determine which columns are involved in the nearly linear dependency lies in that, it is not possible to have a common convention setting how small is small for elements in N_i. Therefore we resort to variance decomposition to solve this problem.

3.2 Variance decomposition for the Jacobian matrix J

Variance decomposition involves decomposing a variance into different parts, and for each part, it is related to one factor we are interested in. Therefore, the effect for each factor on the variance will be quantified. For example, in linear regression, the variation for dependent variable can be decomposed into two parts, one part is related to the independent variables (the model) and the other part is related to the noise in collecting data. From this decomposition, we can explain how much of the variation in the dependent variable is caused by the change of the independent variables, and how much is caused by the noise.

We will use variance decomposition method to obtain more information on nearly linear dependencies among non-identifiable parameters [39]. Considering a linear system $J\beta = \epsilon$, where $\beta \in R^m, \epsilon \sim N_{nq}(0, \Sigma), \Sigma_{nq*nq} = \mathrm{diag}(\sigma^2, \sigma^2, \cdots, \sigma^2)$, is the noise due to computation error or from data. A large variance of certain parameters in the linear regression indicates nearly linear dependencies among columns of J.

The variance-covariance matrix for $\hat{\beta}$, can be written as $V(\hat{\beta}) = \sigma^2(J^TJ)^{-1} = \sigma^2 N S^{-2} N^T$. Thus the variance of the kth coefficient $\hat{\beta}_k$ is $\mathrm{var}(\hat{\beta}_k) = \sigma^2 \sum_j \frac{n_{kj}{}^2}{\mu_j{}^2}$. Since these $\mu_j{}^2$ appear in the denominator, those components associated with nearly linear dependencies will be large relative to the other components.

Consider the variance-decomposition matrix $\pi = (\pi_{jk})$, where $\pi_{jk} = \frac{n_{kj}{}^2/\mu_j{}^2}{\sum_j(n_{kj}{}^2/\mu_j{}^2)}$, $k, j = 1, \ldots, m$. The k^{th} column of matrix (π_{jk}) represents the fractions of the variance $\mathrm{var}(\hat{\beta}_k)$ that are attributed to each of the singular values. The j^{th} row

of (π_{jk}) represents the fractions of all variances attributed to the j^{th} singular value. The s^{th} column in J, where $s = 1,2,\cdots m$, is considered as involved in a nearly linear dependency, if $\hat{\beta}_s$, where $s = 1,2,\cdots m$, has most of its variance appear in rows of the matrix (π_{jk}) associated with small singular values. By convention, if a coefficient has more than 80% of its variance relating to one small singular value, the corresponding column is considered involving in the nearly linear dependency related to this small singular value.

In this chapter, I have discussed our method for detecting and resolving non-identifiability. Our method is designed to detect dependencies among model parameters. Based on implicit function theorem, detecting dependencies among model parameters are equal to detecting linear dependencies among the columns in Jacobian matrix of observable model outputs with respect to model parameters at discrete time points. Due to the existence of numerical error, exact linear dependencies is seldom observed, instead are nearly linear dependencies, which are the potential exact linear dependencies when there is no numerical error. Our method can overcome this problem, and it has two steps—singular value decomposition and variance decomposition. The first step—singular value decomposition can inform us how many nearly linear dependencies existing among the columns of the Jacobian matrix, and the second step—variance decomposition can inform us which parameters are involved in each nearly linear dependency. Based on the analysis from these two steps, we can choose to fix some model parameters' values such that all remaining ones can be uniquely estimated.

Chapter 4

The Single Group SIDT Model for Assessment of HIV Datasets from China CDC

In this chapter, we will build a single group SIDT model to study HIV datasets from China CDC. Our method discussed in Chapter 3 will be applied to detect and resolve non-identifiability in the model based on different HIV datasets. At last, based on the uniquely estimated model parameters' values, the HIV epidemic in each region will be discussed.

Estimation of the burden of HIV epidemics is an important annual or biannual task for public health agencies. Understanding the past dynamics and predicting future trends of the number of new HIV infections (incidence), number of people living with HIV (prevalence), and number of HIV/AIDS related deaths is essential for assessment of HIV interventions and for allocation of resources. Due to the long and variable incubation period of HIV infection, many people infected with the virus are not reported or aware of their infection. The undiagnosed HIV infected population is unknown to HIV surveillance, and an estimation of the size of undiagnosed HIV population is key to estimate the true burden of HIV epidemics.

Several methods exist for the estimation of yearly number of new HIV infections, the number of people living with HIV, and number of HIV related deaths using either surveillance data or prevalence data [40-55].

Back-calculation based methods use HIV diagnosis data and AIDS diagnosis data to construct historical HIV incidence curve. HIV and AIDS diagnosis data is routinely collected by the public health agency. However, the interpretation for these data needs much attention since there is a delay between HIV infection and HIV diagnosis, AIDS diagnosis. The newly diagnosed patients might be infected several years ago. The diagnosis of HIV positive patients depends on several factors including the exposure to

HIV testing and the individual's personal will for HIV testing. The method was originally designed to use AIDS diagnosis data but has been recently developed to use HIV diagnosis data [40-45]. The method assumes a distribution for the time from infection to diagnosis, which incorporates various motivations for HIV testing such as the development of symptoms or exposure to HIV testing. Determination of the distribution needs the introduction of other data sources. The methods are in uses of the recent infection fraction data from newly diagnosis in specifying the distribution.

Back-calculation based statistical methods are commonly used by public health agencies to estimate HIV incidence in developed countries where HIV epidemics have stabilized. However, HIV epidemic transmission in China has not stabilized liked that in Canada (Figure 4.1). On the other hand, Back-calculation method only estimates HIV incidence in the past and it does not make future predictions.

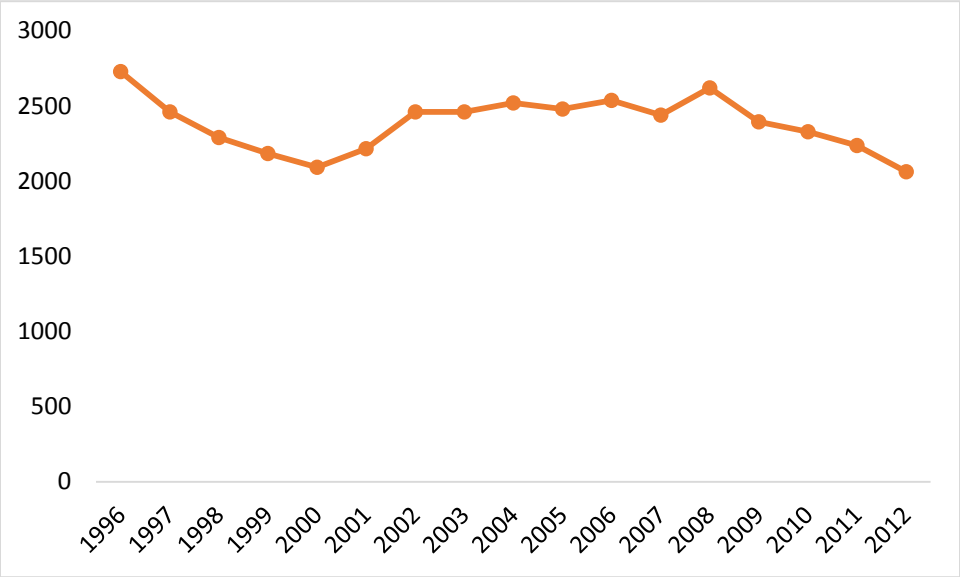


Figure 4.1(a): Yearly number of new HIV reports in Canada

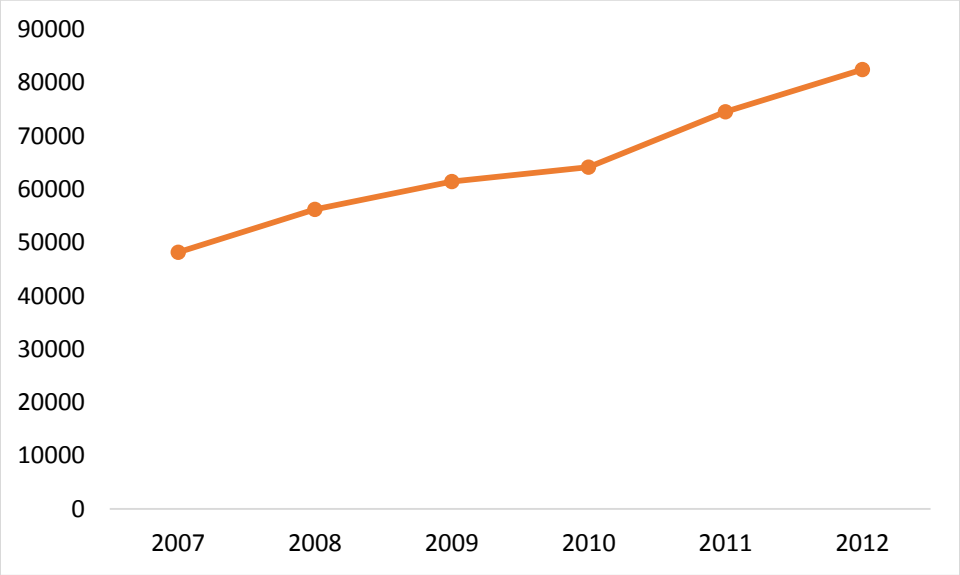


Figure 4.1(b): Yearly number of new HIV reports in China

In developing countries, the Workbook method for the estimation of HIV prevalence is recommended by UNAIDS/WHO [46-50]. HIV prevalence data and estimated size of high-risk groups from the Workbook approach are linked to the Spectrum/ Estimation and Projection Package (EPP) package to produce estimations for HIV incidence and HIV/AIDS related deaths. Prevalence surveys used in the Workbook Method are prone to

sampling biases. If prevalence surveys use a sampling approach targeting in a group with the highest levels of risk, then applying this prevalence to the entire subpopulation will overestimate HIV prevalence. Dividing a risk subpopulation into high-risk and low-risk subgroups and treating them as separate risk groups may reduce this bias, but requires accurate estimation of the respective sizes and HIV prevalence of the separate risk groups.

Mathematical models have been used for HIV estimation exercises [51-55]. Mathematical models, when fitted to surveillance data, can estimate HIV epidemics in the past, predict the epidemics in the future, and evaluate the effectiveness of intervention strategies. Findings from the modeling exercise can provide evidence to inform HIV policy.

4.1 Model description

We developed a SIDT compartmental model using differential equations for the HIV transmission dynamics. The compartments S, I, D and T are chosen based on the natural history of HIV infection and the available data. Susceptible people (S) become infected through contacts with HIV positive people. HIV positive people (I) are diagnosed through HIV testing. HIV positive people who are diagnosed (D) are then enrolled into the ART treatment programs (T). The transfer diagram for our model is shown in Figure 4.2.

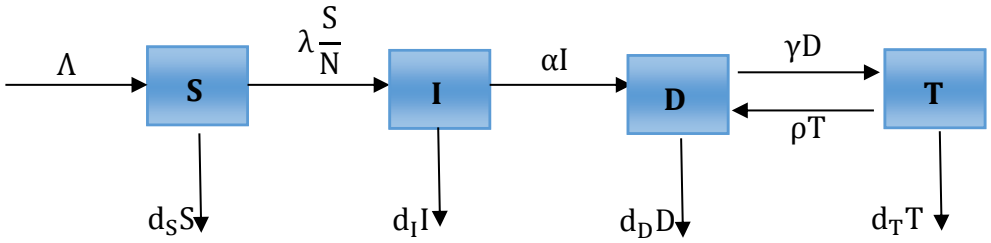


Figure 4.2: Transfer diagram for the single group HIV transmission model

Based on the transfer diagram in Figure 4.2, the model is described by the following system of nonlinear differential equations:

$$\dot{S}(t) = \Lambda - \frac{S(t)}{N(t)}\left(\beta_I I(t) + \beta_D D(t) + \beta_T T(t)\right) - d_S S(t)$$

$$\dot{I}(t) = \frac{S(t)}{N(t)}\left(\beta_I I(t) + \beta_D D(t) + \beta_T T(t)\right) - \alpha(t) I(t) - d_I I(t)$$

$$\dot{D}(t) = \alpha(t) I(t) + \rho T(t) - \gamma(t) D(t) - d_D D(t)$$

$$\dot{T}(t) = \gamma(t) D(t) - \rho T(t) - d_T T(t).$$

At a time t, the number of susceptible people is denoted by $S(t)$, the number of HIV positive people who are not diagnosed is denoted by $I(t)$, the number of diagnosed HIV positive people who are not under treatment is denoted by $D(t)$, and the number of diagnosed HIV positive people that are under treatment is denoted by $T(t)$. The sum $N(t) = S(t) + I(t) + D(t) + T(t)$ denotes the total population. The time unit used in the model is per year to align with the available data.

Parameter Λ is the influx of susceptible, d_S, d_I, d_D and d_T are death rates specific to compartments S, I, D and T. Parameter ρ is the combined rate for treatment failure and loss to follow-up, and β_I, β_D and β_T are transmission coefficients for compartment I, D and T, respectively. The HIV incidence in the model is expressed as $\lambda S(t)/N(t)$, with $\lambda = \beta_I I(t) + \beta_D D(t) + \beta_T T(t)$.

Terms $\alpha(t)I(t)$ and $\gamma(t)D(t)$ in the model are the annual number of new reports for HIV and AIDS combined and the number of new treatment enrollments, respectively. With the implementation of the "Four Frees and One Care" program in 2003, the Chinese government has rapidly scaled up HIV testing and ART treatment. This was reflected by an increase in the number of people tested for HIV and in the number of treatment centers. To correctly adjust for the increase in new HIV tests and ART treatments, we used a time-dependent diagnosis rate $\alpha(t)$ and treatment enrolment rate $\gamma(t)$.

4.2 Project 1: The national HIV dataset of China

Heroin use first emerged as a public problem in China's border region with Myanmar in Yunnan province in the late 1980s [56, 57], because of China's close to the world's major heroin producing area of the golden triangle. The first HIV outbreak among injection

drug users (IDUs) in China was reported in Yunnan in 1989 [58]. By the end of 2002, all 31 of China's provinces, municipalities and autonomous regions reported HIV infection among this population [59]. Sharing needles and syringes are the predominant mode of HIV transmission; with IDUs accounting for 71% of all HIV/AIDS reported cases in China [59]. Meanwhile, the role of plasma donation in HIV transmission was found in 1995. The plasma donors could be infected either through the use of contaminated blood collection equipment or the re-infuse of pooled buffy coats. Although the unregulated plasma collection was eradicated by the end of 1995, tens of thousands of paid commercial blood donors have been infected with HIV in the central provinces of Henan, Anhui, Hubei, and Shanxi constitutes a second important endemic in China [60]. Due to China adopted an open door policy in 1978, commercial sex activities have flourished across the country. The Chinese Public Security Office estimated that there are 4 to 6 million female sex workers nationwide, an increase of 160-fold in number compared with year 1985 [61]. HIV transmission is shifting from injection drug users (IDU) and illegal plasma/blood collecting practices to populations at risk through unprotected sex, either through heterosexual contacts or unprotected homosexual sex between men, accounting for nearly half of all new infections in 2007 [62]. The proportion of reported HIV cases among MSM has increased eight-fold from 0.4% in 2005 to 16.1% in 2011 [63]. MSM has emerged as a high-risk group in the nation. These new changes pose greater challenges than ever before to China's AIDS control efforts.

In response to the growing burden of HIV/AIDS, Chinese government has launched a four-free and one-care policy, including free HIV screening test and treatment. Understanding of epidemic trend patterns of HIV infection and annual new infections is important to allocate resource and evaluate efforts. Since China CDC established the national HIV epidemiology and antiretroviral treatment databases [64, 65], which recorded all HIV/AIDS reported cases, death and treatment data. Data from these web-based databases is recorded in real time and updated timely. It provides unique insight for

understanding the current and future HIV epidemic in China. The aim of our study is to apply the method of dynamic mathematic modeling to carry out estimation and projection of people living with HIV/AIDS (PLHIV), HIV/AIDS death and annual new HIV infections.

Data sources: Aggregated yearly HIV/AIDS surveillance data was from the National HIV/AIDS Surveillance Database of the China Center for Disease Control and Prevention. Aggregated yearly data on ART treatment was from the National Treatment Database. Surveillance and treatment data was collected from publications of the China Center for Disease Control and Prevention from 2005-2012 [66]. Yearly population data on national population was from the National Census.

Parameter estimation and fitting: Demographic parameters Λ and d_S are estimated by fitting the equation of the total population $N' = \Lambda - d_S N$ to the national census data of China from 2005 to 2012. Values of parameters d_D, d_T, ρ, and the form of function $\gamma(t)$ are estimated directly from the surveillance and treatment data. Function $\alpha(t)$ was obtained from HIV testing data. Values of other parameters, including transmission coefficients β_I, death rate d_I, the initial population size I_0 for the undiagnosed compartment at the beginning of our fitting (end of 2007), and β_D, β_T, cannot be estimated directly from the data and need to be obtained through model fitting.

We applied the correlation method, the eigenvalue method we discussed in Chapter 1 and our Matrix Decomposition method discussed in Chapter 3 to deal with the non-identifiability issue among these model parameters.

The threshold δ for the correlation method was set as 0.2. If the correlation coefficient of two columns in the Jacobian Matrix was larger than 0.8, these two corresponding parameters were considered as highly correlated. Based on this setting, the total correlation of each model parameter was calculated:

$$c_{\beta_I}^{tot} = 2.989, c_{\beta_D}^{tot} = 2.992, c_{\beta_T}^{tot} = 2.973, c_{d_I}^{tot} = 2.985, c_{I_0}^{tot} = 0.$$

The first model parameter we are suggested to fix is β_D, which has the largest total correlation. Following [69, 70], there is a reduction of transmissibility for people in compartments D, due to awareness of their HIV positive status, namely $a = \beta_D/\beta_I < 1$. A study in the United States [69] showed that high-risk sexual behaviors were reduced by 53% when people became aware of their HIV positive status, compared to people who were unaware. We conservatively assumed $a = 0.75$.

We refitted the model to data to estimate the total correlation coefficients for the remaining model parameters: $c_{\beta_I}^{tot} = 1.991, c_{\beta_T}^{tot} = 1.979, c_{d_I}^{tot} = 1.988, c_{I_0}^{tot} = 0$. In this step we are suggested to the fix the value of model parameter β_I. However, it is not possible to fix the transmission coefficient β_I. We need to resort to other methods to resolve the non-identifiability issue.

We next applied the eigenvalue method to diagnose non-identifiability issue in our model. The smallest eigenvalue for the matrix J^TJ is less than 0.001, and the corresponding eigenvector was [0.797 0.556 0 -0.231 -0.0097], which indicated that the least identifiable model parameter is β_I. We are suggested to fix the value for β_I. Since the value of β_I can not be reasonably fixed, we resort to other method to deal with the non-identifiability issue.

We applied our Matrix decomposition method to resolve the non-identifiability issue. The singular value decomposition informed us that there are two condition indexes that are larger than 30, therefore there are two dependencies among these model parameters and we need to fix two parameters' values before model fitting, and the variance decomposition informed us that parameters $\beta_I, \beta_D, \beta_T$, and d_I are involved in these two dependencies (table 4.1).

Condition Index	Proportions of Variance				
v_i	$var(\beta_I)$	$var(d_I)$	$var(I_0)$	$var(\beta_D)$	$var(\beta_T)$
1	0	0	0.0086	0	0
4	0	0	0.7122	0	0
22	0	0	0.2736	0	0.0293
268	0	0	0.0056	0.9996	0.9707
38937	1.0000	1.0000	0	0.0004	0

Table 4.1 Results of Singular value and variance decomposition analysis of the national HIV dataset for China

We estimated the value of d_I using survival data in the literature for newly infected treatment naive people [67, 68]. We preselected the ratio among β_I and β_D as we did above, $\beta_D = 0.75\beta_I$.

After fixing these two parameters' values, the remaining parameters β_I, β_T and the initial value I_0 can be uniquely estimated. The best-fit for β_T is much larger than the best-fit for β_I, which is unrealistic since the viral load for people under treatment is much lower than the viral load for undiagnosed HIV patients. Even if we set the ratio $b = \beta_T/\beta_I < 1$ and redo the fitting to estimate parameter b, the best-fit for b is 1. Therefore we fixed the value for b before model fitting. Studies [71-76] demonstrated that the transmissibility for HIV positive people under ART treatment was reduced by 92%, compared to HIV positive people not receiving ART treatment due to the reduced viral load. We conservatively assumed $b = 0.1$.

The Nonlinear Least Squares method [28] was applied to find the point estimates for model parameters β_I and I_0, which minimize the sum of squared error between model output and the available surveillance and treatment data. More specifically, the following data are used in our modeling fitting:

- Annual number of new diagnosis of HIV and AIDS combined from 2005 to 2012;
- Annual number of death due to HIV and AIDS among diagnosed from 2005 to 2012;
- Annual treatment enrollment from 2005 to 2012;
- Annual number of death among diagnosed people and people in treatment from 2005 to 2012.
- Annual number of treatment failure and loss to follow-up from 2005 to 2012
- The number of total population from 2005 to 2012

Here model output y is a vector with six components:

$$y^{diag}(t) = \int_{t-1}^{t} \alpha(\tau) I(\tau) d\tau, t = 2005, \dots, 2012$$

$$y^{diagdeath}(t) = \int_{t-1}^{t} d_D D(\tau) d\tau + \int_{t-1}^{t} d_T T(\tau) d\tau, t = 2005, \dots, 2012$$

$$y^{treat}(t) = \int_{t-1}^{t} \gamma(\tau) D(\tau) d\tau, t = 2005, \dots, 2012$$

$$y^{treatdeath}(t) = \int_{t-1}^{t} d_T T(\tau) d\tau, t = 2005, \dots, 2012$$

$$y^{treatfail}(t) = \int_{t-1}^{t} \rho T(\tau) d\tau, t = 2005, \dots, 2012$$

$$y^{pop}(t) = S(t) + I(t) + D(t) + T(t), t = 2005, \dots, 2012$$

Bayesian-based MCMC method [29, 30] was applied to calculate the 95% confidence intervals for model parameters β_I, d_I, d_D, d_T, ρ and I_0. Each time we calculate the 95% confidence interval for one parameter fixing all remaining parameters at the point estimates we obtained.

Model validation: Validation of our model was carried out using two independent approaches. First, the goodness of fit ratio [28] R^2=0.92 indicates an excellent fit between our model output and the data. Second, leave-one-out cross-validation method [35] was applied to further verify the consistency of the model fit. One data point was arbitrarily picked up for validating and the remaining ones were used for model fitting.

The model prediction was consistent with the validation point.

Sensitivity analysis and uncertainty analysis: We carried out local sensitivity analysis of our model predictions for new HIV infections of year 2010, using one-at-a-time method at the best-fit parameter values [32, 33]. Our analysis showed that the top three most sensitive parameters were transmission coefficients β_I, β_D, and death rate d_I for the undiagnosed population.

For uncertain analysis, we computed our base model prediction using the best-fit parameter values from the Nonlinear Least squares fitting. To produce prediction intervals with high and low estimates, we allowed all parameters to vary within their confidence intervals, and applied Latin Hypercube sampling (LHS) [31] to produce 20,000 samples of parameter values. Our prediction intervals contained 95% of the 20,000 model outputs using sampled parameter values.

Results

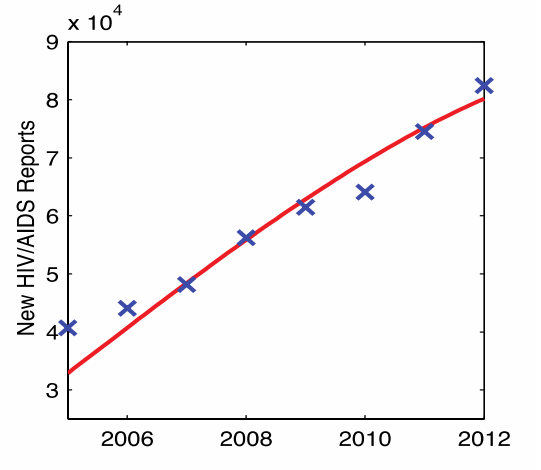
The point estimates for all model parameter together with the corresponding confidence intervals are shown in table 4.2.

Parameters	Description	Best-fit value	95% CI	Source
β_I	Transmission coefficients for I	0.14789	[0.1226, 0.16999]	Fitting
β_D	Transmission coefficients for D	0.1109	[0.09195, 0.1274925]	[69, 70]
β_T	Transmission coefficients for T	0.01479	[0.01226, 0.016999]	[71-76]
d_S	Death rate for S	0.007*	[0.00592, 0.00812]	Fitting
d_I	Death rate for I	0.06301*	[0.05959, 0.06372]	[67, 68]
d_D	Death rate for D	0.05308*	[0.05013, 0.05367]	Data
d_T	Death rate for T	0.04166*	[0.03973, 0.04248]	Data
ρ	Treatment drop-out rate	0.03744*	[0.03131,0.0436746]	Data
I_0	Undiagnosed HIV positive population in 2007	417990	[376864, 456273]	Fitting
$\alpha(t)$	Time-dependent diagnosis rate of undiagnosed	$0.017035\,t - 34.065$*		Data
$\gamma(t)$	Time-dependent treatment enrollment rate	$0.008112\,t - 16.204$*, 2005-2009, $0.034576\,t - 69.359$*, 2010-2020.		Data
R^2	Goodness of fit ratio	0.92		Fitting

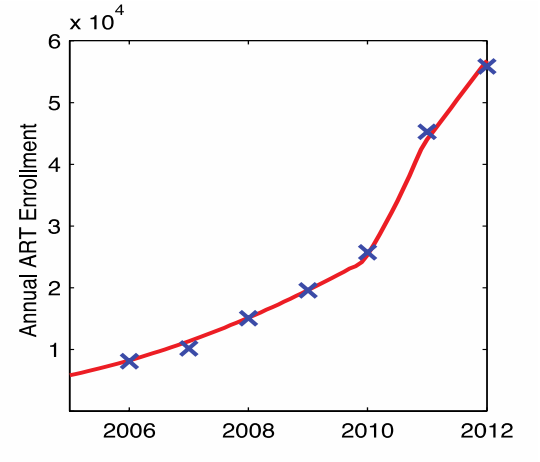
Table 4.2. Model parameter and their best-fit values together with 95% confidence intervals.

Notes: * The unites for these parameters are 1/person/year

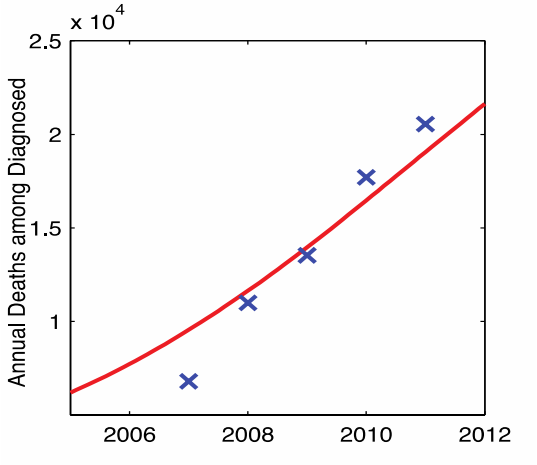
The model fitting graphs are show in figure 4.3.



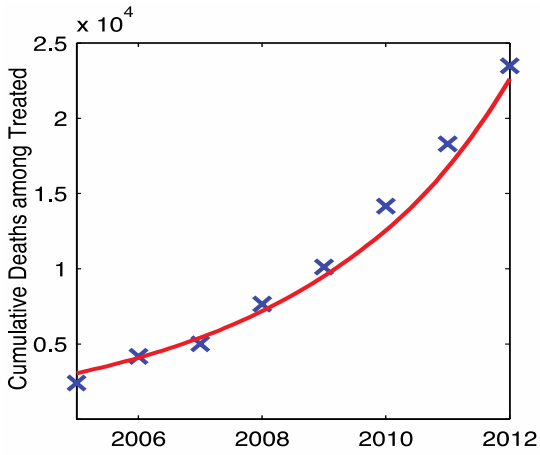
(a) Annual reported HIV/AIDS cases



(b) Annual ART enrollment



(c) Annual deaths among diagnosed



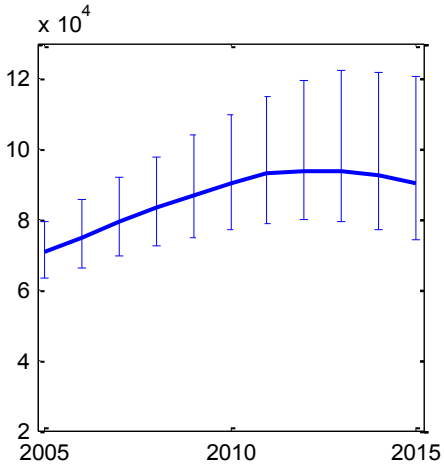
(d) Cumulative deaths among treated

Figure 4.3 Comparison of model outcomes to the surveillance and treatment data.

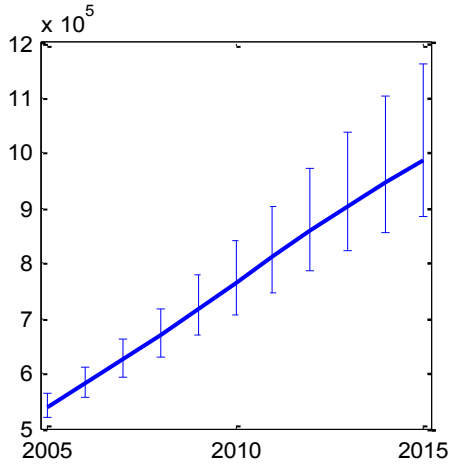
The fitting graphs imply a high goodness of fit ratio, which is equal to 0.92. This provides a validation for our model.

We have estimated three main indicators: annual number of new HIV infections, people living with HIV (PLHIV), and HIV/AIDS related death for the national population. The trends for HIV/AIDS during 2007 and 2012 were estimated using the CDC surveillance data. Estimations for the year 2005 and 2006 were back calculated

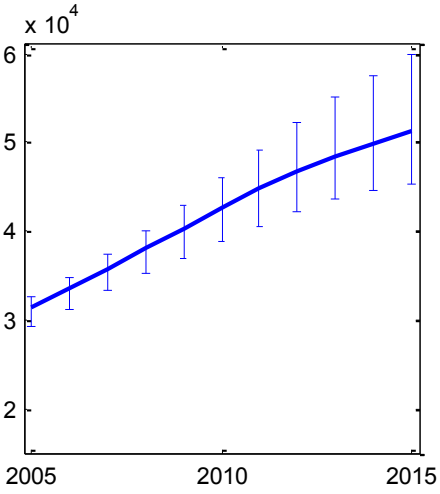
from fitted model. Projections for the period of 2013 and 2015 were done using the best-fit parameter values while assuming the diagnosis rate and treatment enrollment rate both continue to increase linearly with the same pre-2013 trends. High and low estimates are the 95% prediction intervals by sampling parameter values from their 95% confidence intervals using Latin Hypercube Sampling, which was conducted under the assumption that each parameter follows triangle distribution with peak fixing at the corresponding point estimate, two end points fixing at the lower bound and upper bound for the corresponding 95% confidence interval. The results are shown in figure 4.4.



Annual number of new HIV infections



Number of people living with HIV (PLHIV)



Number of deaths among PLHIV

Figure 4.4 Estimations and predictions for annual HIV infections, PLHIV, HIV/AIDS deaths. Vertical bars indicate 95% prediction intervals.

Our model estimations show that the annual new HIV infections has risen from 70,557 in 2005 to 93842 in 2012 (Figure 4.4 (a)), with the annual percentage increase slowing down from 6% in 2006 to 1% in 2012. This is a strong indication that the HIV epidemic in China has slowed down significantly in recent years. If the momentum of scaling-ups of new HIV diagnosis and of ART treatment is maintained through 2015, our models projection shows that the new infections will start to level off and decline from 2012. Even the high estimates of our prediction interval show a turning point for new infections as early as 2014.

Estimated total number of people living with HIV and AIDS (PLHIV), including both diagnosed and undiagnosed, has shown a steady and almost linear increase from 539,925 in 2005 to 859,452 in 2012. The annual percentage of increase has declined from 7.7% in 2005 to 5.8% in 2012. The model projection for PLHIV in 2015 is 985,971, with a 95% prediction interval (883,725, 1,161,735). The annual percentage of increase will continue to decline to 4.1% in 2015.

Estimated number of death from people living with HIV and AIDS has shown a steady linear increase from 31,347 in 2005 to 46,716 in 2012. The annual percentage of increase has declined from 7% in 2005 to 4.4% in 2012. The model projection for combined HIV and AIDS death in 2015 is 51,229, with a 95% prediction interval (45,332, 59,829). The annual percentage of increase will continue to decline to 2.5% in 2015.

Based on our estimations, among the HIV positive population, the percentage of people who have been diagnosed has risen from 25% in 2005 to 54.4% in 2012, and is projected to continue increase to 65.6% in 2015. Percentage of people under treatment among the diagnosed population has risen from 16.1% in 2005 to 35.2% in 2012, and is projected to continue rise to 53.3% in 2015.

Our validated model was used to make projections on potential interventions for the period from 2015-2020. We examine the impacts of continuing scale-up in HIV diagnosis and treatment enrollment in three different scenarios:

- Scenario 1: Scaling up of both HIV diagnosis and treatment enrollment after 2015 as pre-2015.
- Scenario 2: Scaling-up of HIV diagnosis as pre-2015 and HIV treatment enrollment is kept at the 2015 level.
- Scenario 3: Scaling-up of HIV diagnosis as pre-2015 and HIV treatment enrollment is kept at the 2013 level.

Model projections for the four scenarios are shown graphically in Figure 4.5.

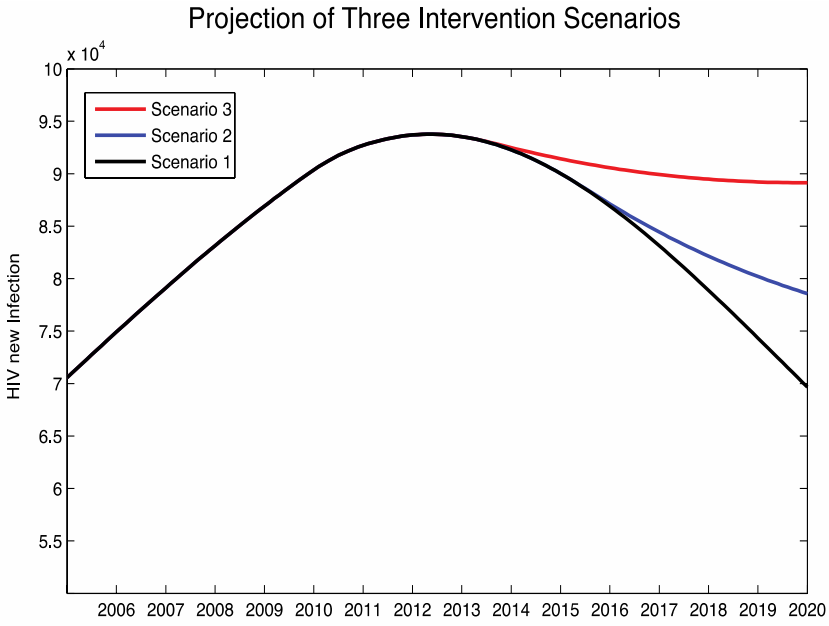


Figure 4.5. Projection of new HIV infections under three intervention scenarios.

Under the best scenario (Scenario 1), if we maintain the linearly increasing trend in HIV diagnosis and treatment enrollment through 2020, the proportion of people who

diagnosed among people living with HIV will rise to 80%, and the proportion of people receiving treatment among diagnosed population will reach 75% by the end of 2020. The annual new HIV infections by the end of 2020 will be reduced by 22.8% from its 2010 level, and by 25.7% from its highest level in 2012.

Our model projections show that to achieve the turning point of the HIV epidemics around 2013 and maintain the downward trend in annual new infections, it is important to keep the momentum in diagnosis and treatment during the past 5 years well beyond 2015. If the diagnosis and treatment efforts are only maintained at the 2013 level without further scaling up, the downward trend after 2014 will quickly reverse and the annual new infections will level off.

Our validated model was used to calculate the control reproduction number R_C: the actual average number of secondary cases per primary case observed in a population for an infectious disease in the presence of control measures [77]. Unlike the basic reproduction number R_0 which is measured at the beginning of an epidemic, the control reproduction number takes in the effect of intervention measures and varies as the epidemic progresses in time. Since our model parameters are time dependent, the control reproduction number is more suitable an indicator of the strength of the transmission dynamics than R_0. In our model, effects of the national HIV/AIDS programs between 2005 and 2015 were incorporated into the time dependent diagnosis rate $\alpha(t)$ and ART enrolment rate $\gamma(t)$. The impact of the response programs during the time period 2005-2015 will be measured by the time-varying control reproduction number $R_c = \frac{\beta_I}{\alpha(t)+d_I} + \frac{\beta_D \alpha(t)(\rho+d_T)}{(\alpha(t)+d_I)(\gamma(t)d_T+\rho d_D+d_D d_T)} + \frac{\beta_T \alpha(t)\gamma}{(\alpha(t)+d_I)(\gamma(t)d_T+\rho d_D+d_D d_T)}$, which is calculated as the spectral radius of the next generation matrix FV^{-1}, where

$$F=\begin{pmatrix} \beta_I & \beta_D & \beta_T \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad V=\begin{pmatrix} \alpha(t)+d_I & 0 & 0 \\ -\alpha & \gamma(t)+d_D & -\rho \\ 0 & -\gamma(t) & \rho+d_T \end{pmatrix}.$$

In Figure 4.6, we have shown plots of distribution of R_C at year 2015 (Figure 4.6(a)), and at year 2020 for the three intervention scenarios (Figure 4.6 (b)-(d)). The distribution was produced by sampling model parameters in their confidence intervals.

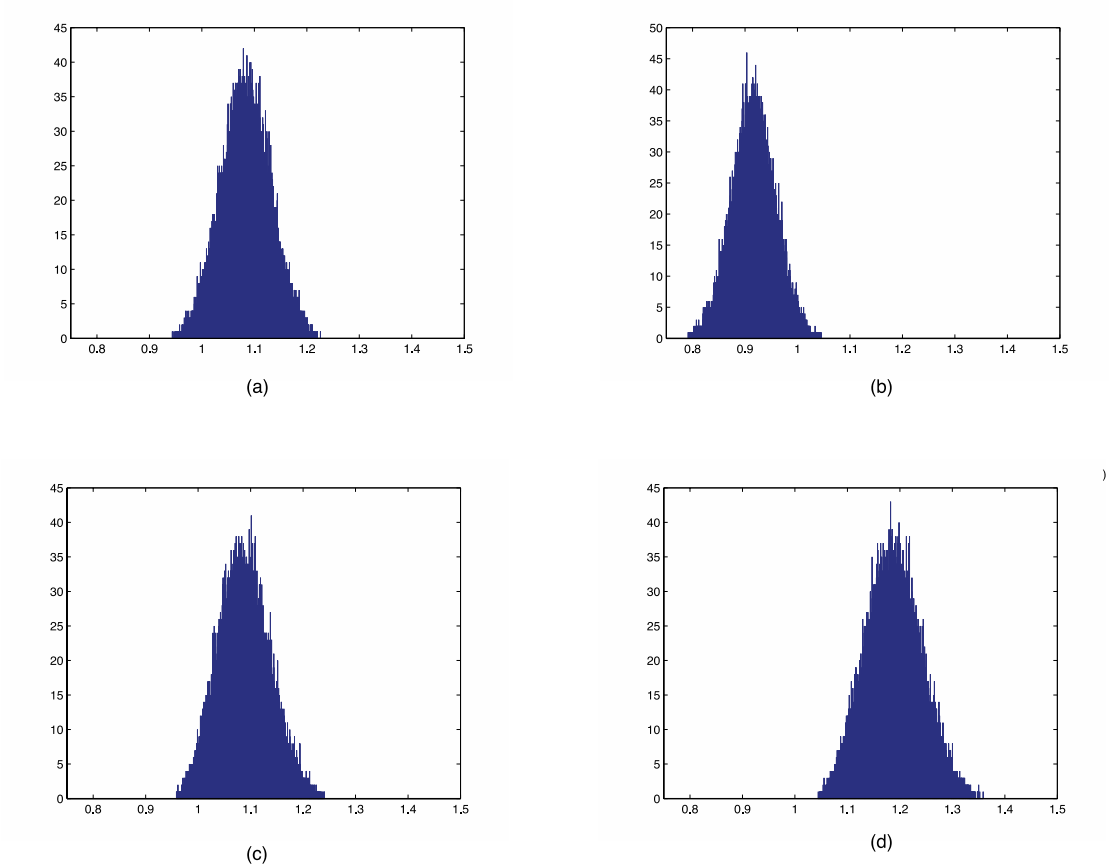


Figure 4.6. Distribution of the control reproduction number calculated from our model when parameter values are sampled from their confidence intervals.

Figure 4.6(a) shows the mode of R_C at 1.1, indicating the epidemic is slowing down. Figure 4.6(b) shows the mode of R_C in 2020 at 0.9 and the entire distribution lies below threshold value 1. This is another indication that continuing the scaling up in diagnosis and treatment through 2020 can ensure sustained decline in new HIV infections and bring the epidemic under control. In contrast, Figure 4.6 (c) and (d) show modes of R_C remain above 1 under intervention scenarios 2 and 3, indicating that ceasing the scaling up in diagnosis and treatment before 2020 will not achieve control of the epidemic.

We also examined the impact of varying α and γ independently on the values of reproduction number R_C. The results are shown in a contour plot in Figure 4.7.

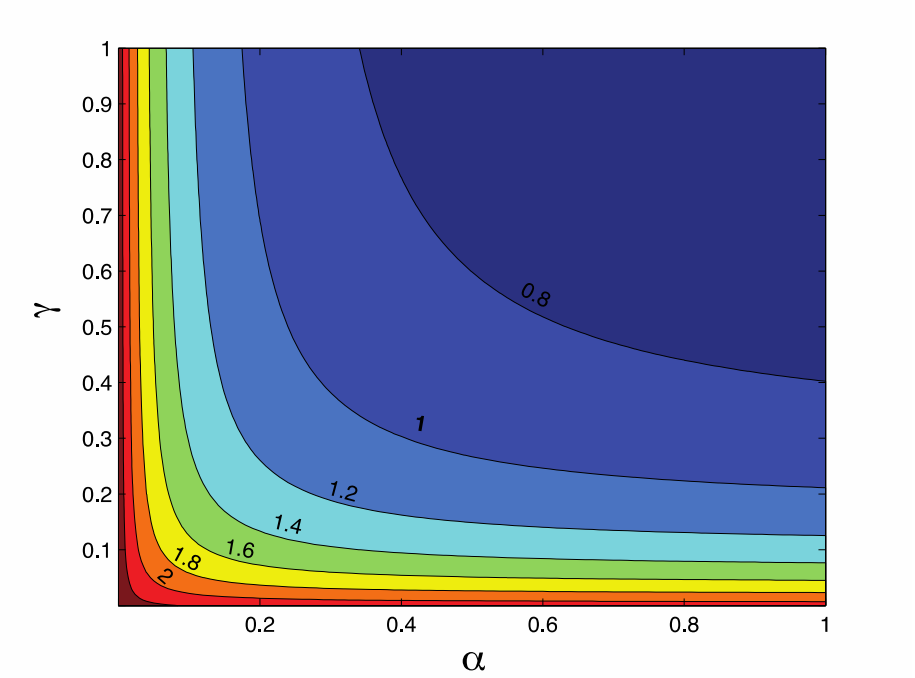


Figure 4.7. Dependence of the control reproduction number on parameters α and γ.

We can see that increasing both α and γ is an effective approach to reduce R_C to below 1. In 2012, the values of α and γ were estimated as both at 0.2 and were projected to reach 0.26 and 0.31 by 2015, a 30% and 55% increase in three years, respectively. If the momentum of increase in diagnosis and treatment is kept until 2020, it is reasonable to expect that values of α and γ will both reach beyond 0.4, doubling the level of 2012. At such a level, the value of R_C will be below 1, indicating control of the epidemic.

Discussion

China has drastically scaled up its efforts in HIV testing and treatment since 2005. These are reflected by the increases in total number of people tested, number of sentinel testing sites, and number of treatment centers. These efforts are partly responsible for the ever-increasing number of new HIV/AIDS diagnosis and new treatments in the country. To

account for the scaling up in testing and treatment, we model the new diagnosis and new treatment enrollment by $\alpha(t)I(t)$ and $\gamma(t)D(t)$, respectively, and allow both parameters α and γ to be time dependent. The surveillance and treatment data show that function $\gamma(t)$ is a piece-wise linear function with a higher slope after 2010 (Figure 4.8).

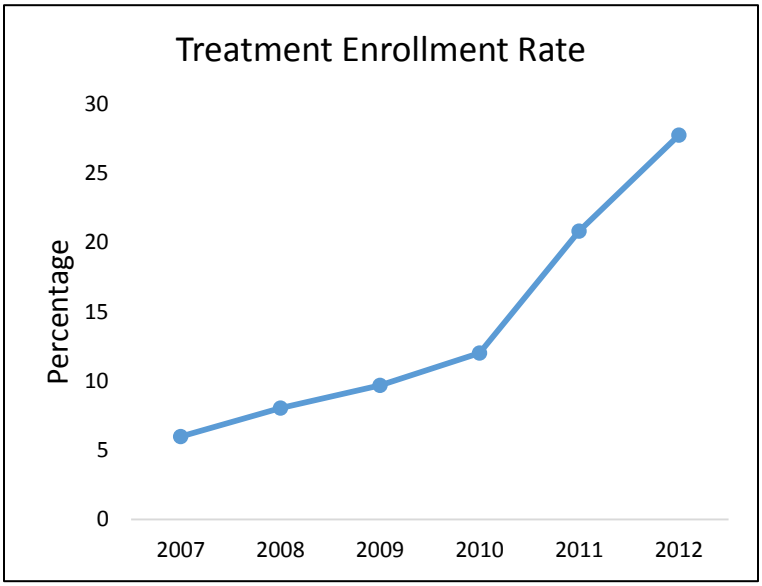


Figure 4.8. Annual treatment enrollment rate during the period 2007-2012.

The model informs us that $\alpha(t) =$ annual HIV new diagnosis$/I(t)$. If we assume HIV positive patients are uniformly distributed among the total population, annual HIV new diagnosis$/I(t)$ can be approximated by HIV test rate. Therefore function $\alpha(t)$ can be approximated by annual HIV test rate. The HIV testing data shows that the HIV test rate from 2007 to 2012 can be simulated as a linear function (Figure 4.9), and as a result we fix the function $\alpha(t)$ as a linear function after 2005. These adjustment for function $\alpha(t)$ allowed us to estimate the increase in HIV diagnosis due to the intrinsic HIV transmission dynamics and avoided overestimation of the transmission coefficient β_I (Figure 4.10).

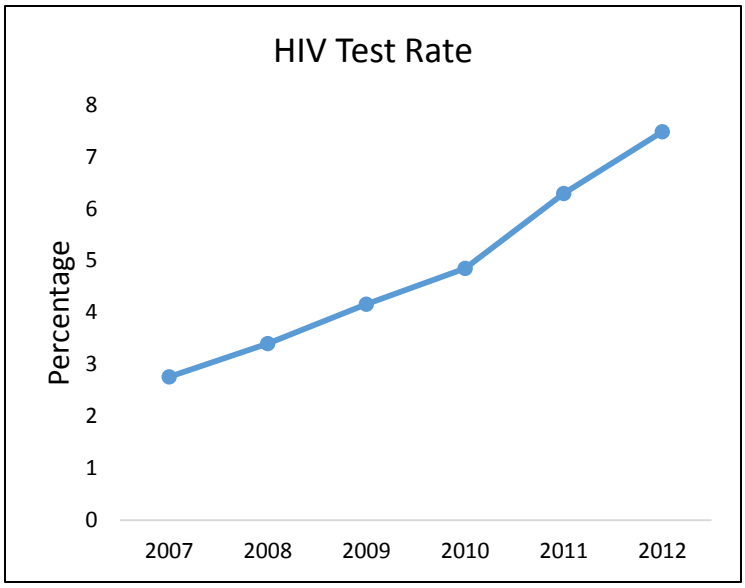


Figure 4.9. Annual HIV test rate during the period 2007-2012.

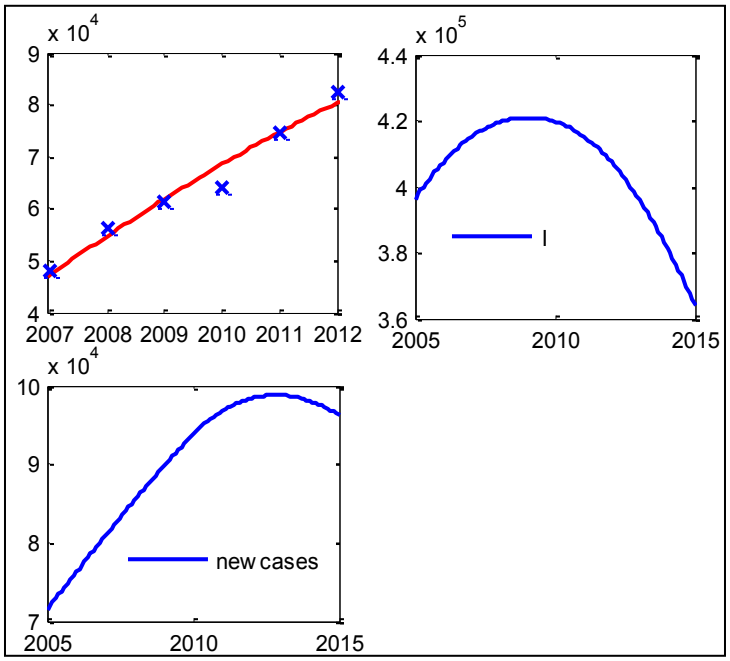


Figure 4.10. Model fitting results, the trends for I and annual HIV new infections when $\alpha(t)$ is a linear function

If we overestimate the strength for HIV testing, such as assuming $\alpha(t)$ is a quadratic function, although model fitting can be as well, we would underestimate the intrinsic HIV transmission (Figure 4.11).

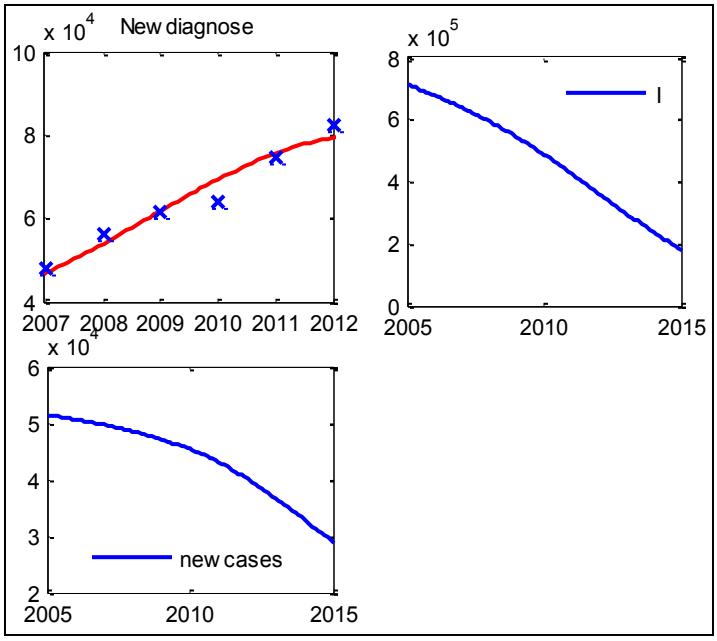


Figure 4.11. Model fitting results, the trends for I and annual HIV new infections when $\alpha(t)$ is a quadratic function

If we underestimate the strength for HIV testing, such as assuming $\alpha(t)$ is a constant, although model fitting can be as well, we would overestimate the intrinsic HIV transmission (Figure 4.12).

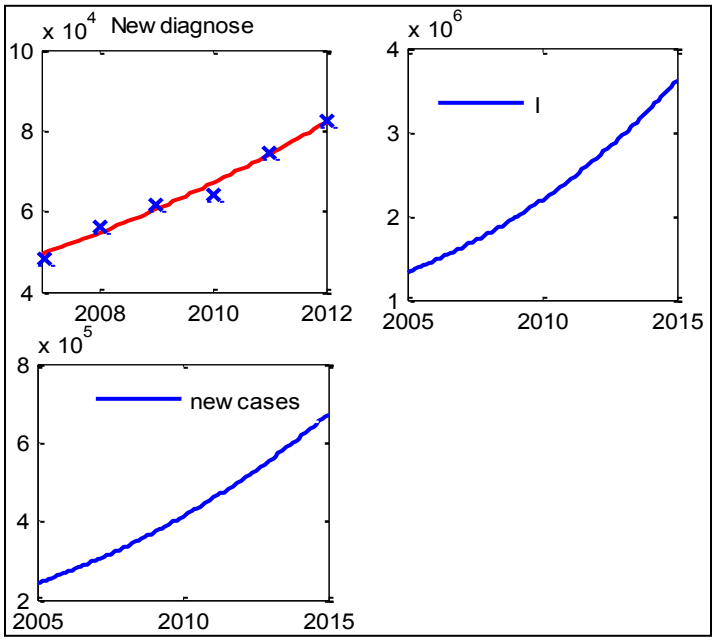


Figure 4.12. Model fitting results, the trends for I and annual HIV new infections when $\alpha(t)$ is a constant

Therefore the proper selection for the form of $\alpha(t)$ is crucial for our estimation and prediction. All the three different forms of $\alpha(t)$ can lead to good model fitting, which means that observable model outputs under these three assumptions are almost the same, while model unobservable outputs can be totally different. For our single group SIDT model, the selection for the form of $\alpha(t)$ depends on the trend for HIV testing rate.

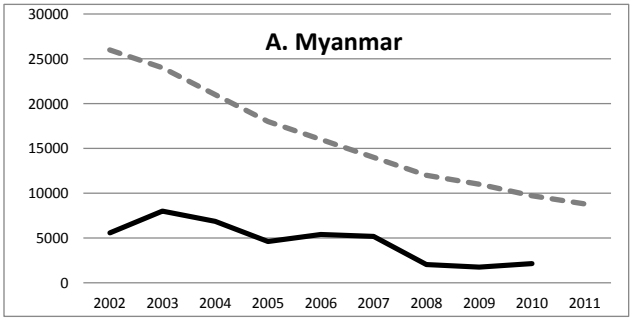
Unlike the national estimate showing a steady decline of new HIV infections in China, the mathematical modeling work showed a steady increase. The model results estimate 29,125 and 44,735 more new HIV infections than the national estimate in 2007 and 2011 respectively. Additionally, the model results estimate 15,759 and 16,732 more AIDS deaths than the national estimate in 2007 and 2011 respectively. The differences in the total number of PLHIV are mixed, there are 75,330 less and 16,723 more PLHIV estimated by the model than the national estimate in 2007 and 2011 respectively (Table 4.3). These differences show that there may be an underestimation of both the new HIV infections and HIV/AIDS deaths by the National HIV Epidemic Estimation Working Group. Further evidence for a possible underestimation can be seen in the irregularity in

the trends of China's HIV/AIDS surveillance reporting data and estimate data, compared with the same set of data of Myanmar, Thailand and the United States (Figure 4.13). Firstly, all the datasets of the other countries show the same increasing or decreasing trend in both reporting data and estimation data of new HIV infections. Secondly, their estimated numbers are always larger than the reported numbers. The ratio of estimated verses reported is 1.1 for the USA, 1.6 for Thailand and above 2 for Myanmar. It is understood that a country with more resources and skills should have a smaller ratio than a country with less resources and skills. The Chinese such dataset showed an unusual reversed ratio of less than one, with 0.7 in 2009 and 0.52 in 2011 (Figure 4.13).

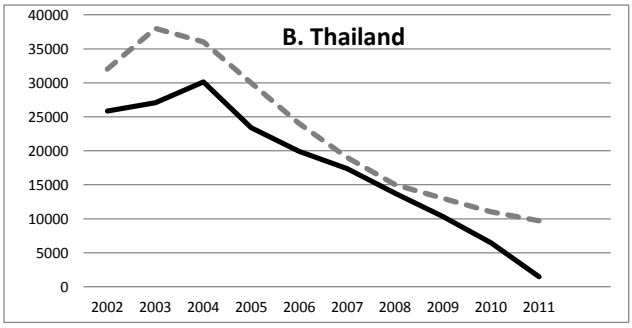
Year	2007+	2007*	Difference	2009+	2009*	Difference	2011+	2011*	Difference
New infections	79,125	50,000	29,125	86,912	48,000	38,912	92,860	48,000	44,860
PLHIV	624,670	700,000	-75,330	716,414	740,000	-23,586	812,326	780,000	32,326
HIV/AIDS death	35,759	20,000	15,759	40,294	26,000	14,294	44,749	28,000	16,749

Notes: + model estimates; * NCAIDS national HIV/AIDS estimates.

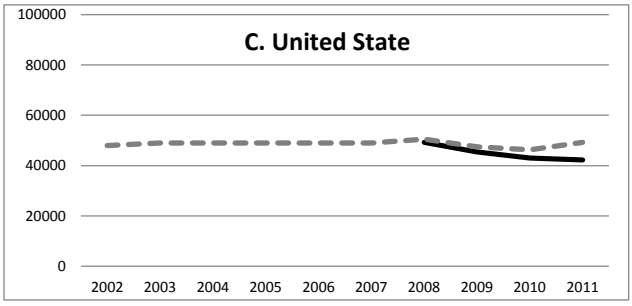
Table 4.3: Comparisions of the HIV/AIDS estimation results of both national HIV/AIDS estimation and our model estimation



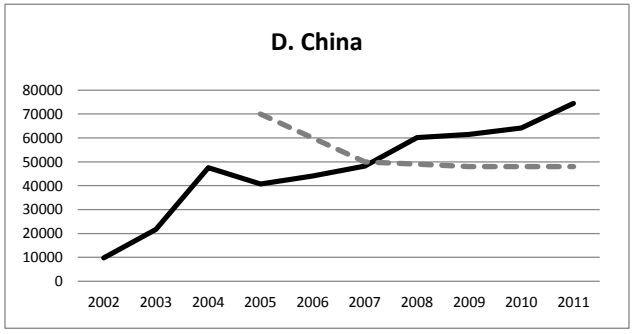
Myanmar	HIV/AIDS reporting data	Estimated new infections data	Ratio*
2002	5567	26000	4.67
2005	4601	18000	3.91
2007	5179	14000	2.70
2009	1735	11000	6.34
2010	2144	9700	4.52



Thailand	HIV/AIDS reporting data	Estimated new infections data	Ratio*
2002	25854	32000	1.24
2005	23401	30000	1.28
2007	17351	19000	1.09
2009	10301	13000	1.26
2010	6443	11000	1.70



United States	HIV/AIDS reporting data	Estimated new infections data	Ratio*
2008	49226	50501	1.03
2009	45470	47408	1.04
2010	43051	46268	1.07
2011	42181	49273	1.17



China	HIV/AIDS reporting data	New infections estimated data	Ratio*
2005	40711	70000	1.72
2007	48161	50000	1.04
2009	61470	48000	0.70
2011	74517	48000	0.52

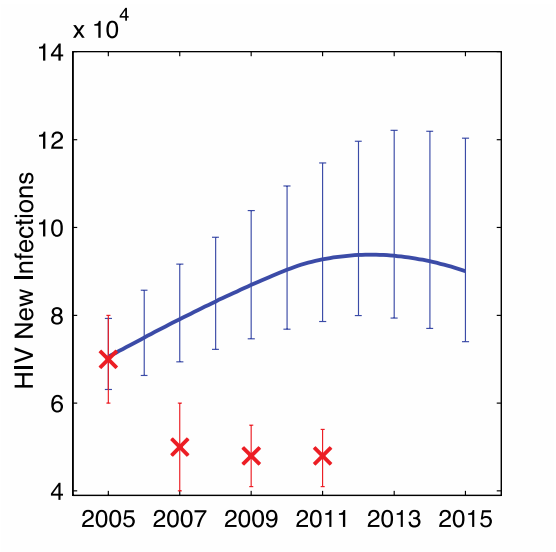
HIV new infections estimated data

* Ratio= new infections estimated data/ HIV/AIDS reporting data

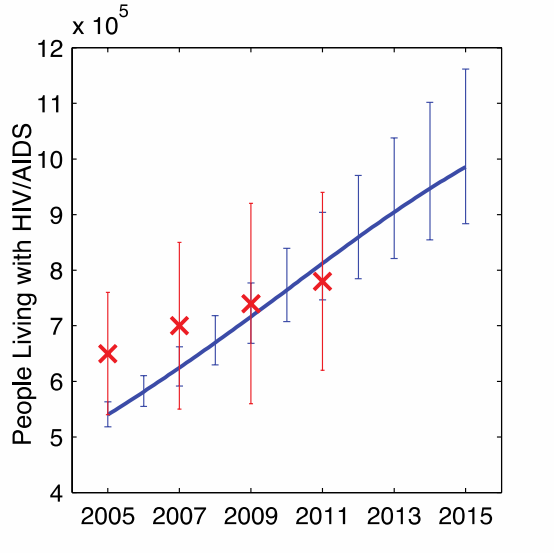
Source：UNAIDS. (2012) AIDS info 1.1

Figure 4.13: Relationship between HIV/AIDS reporting data and estimated new infections data

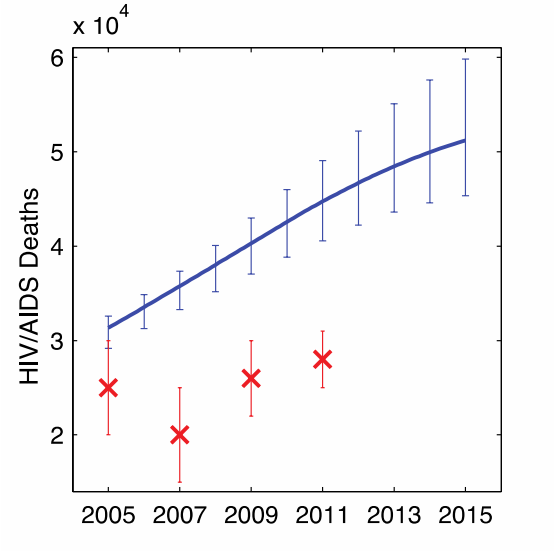
An adapted Workbook method was used in the national HIV/AIDS estimation exercises in China in 2005, 2007, 2009, and 2011, by the National HIV Epidemic Estimation Working Group. Results of the estimation exercises from the Ministry of Health (MOH) reports are plotted in Figure 4.14. National estimation of HIV prevalence and population sizes of risk groups were obtained from estimations at the prefecture level in 2005, and from estimations at the county level since 2007. Data on HIV prevalence rates came from sentinel surveillance data, special epidemiological surveys, mass screenings of target populations, and literature searches in scientific journals [78, 79]. Data from the 12 months prior to the estimation exercise were used. When recent data were not available, adjustments were made to previous data as considered necessary by local CDC officials [79]. Missing HIV prevalence and risk-group population size data for specific geographic areas were imputed using data from areas with similar geographic and socio-economic characteristics. EPP was used to generate a national HIV prevalence curve based on prevalence curves from lower level regions and different subpopulations. EPP output was entered into Spectrum to calculate the annual number of new HIV infections and HIV/AIDS [79]. The quality of the national estimation results can be greatly influenced by the quality of data collection at the regional level. Data collection from the county level required mobilization and training of large number of survey workers, and enormous amount of time and resources. When survey data were tallied at each of the county, prefecture, and provincial level, approval from local government was required before the data was submitted to the next level. The process is time consuming and costly, and political interferences could influence the quality of data at the local level. With China's great expanse and heterogeneity of geographical regions, estimations applied to unsurveyed local regions can lead to variability in the quality of local data.



(a) Annual new HIV infections



(b) People living with HIV/AIDS (PLHIV).



(c) Deaths among PLHIV

Figure 4.14: Model estimates from 2005 to 2015. Lines are baseline model estimates, crosses are estimates in the MOH reports in year 2005, 2007, 2009 and 2011, and vertical bars indicte 95% confidence intervals.

Laboratory-based algorithms are another approach to the estimation of HIV incidence. Recent advancement in laboratory diagnostic of HIV allowed development of immunological assays for the classification of recent HIV infections (within 4 to 12 months prior to diagnosis) among the newly diagnosed HIV cases. Statistical methods such as stratified extrapolation method can be applied to new infections data to produce estimation of HIV incidence. The BED HIV capture enzyme immunoassay was used in the estimation of HIV incidence in the United States since 2008 [80, 81]. A CD4-staged back-calculation model using the new HIV diagnosis and CD4 count information was used to estimate HIV incidence and prevalence in the UK [82]. In China, diagnostic techniques have been developed and implemented for BED and CD4 analysis on serum specimen of newly diagnosed cases. It is possible to adopt the laboratory-based methods for HIV estimations in conjunction with the Workbook Method.

Mathematical model has been used to estimate HIV incidence and prevalence. Mathematical models are developed to reconstruct the processes of transmission of HIV, diagnosis, treatment and occurrence of AIDS and death for populations under a study. Models are validated through fitting to multiple data sources. A validated model can produce estimations for HIV epidemics in the past, project incidence and prevalence of the epidemic into the future, and to provide quantitative evaluations and cost-benefits analysis of potential prevention and intervention measures. Another advantage of the model-based estimation is that models can be fitted to a variety of data sources. In our model study, only the HIV/AIDS surveillance and treatment data from the national direct-reporting surveillance database and treatment database are used in model fitting. The independence of our study from the prevalence surveys and national estimation from Workbook Method allowed us to avoid any inaccuracy in data collection associated with the Workbook Method. Our model accounted for the increasing rates of HIV testing and treatment in recent years in China, in order to produce more accurate estimation on the size and mean transmission rate of undiagnosed population. We only present our

estimations for the national population in this report. Estimations for high-risk groups at the national level are underway, and the results will be presented in a separate report. Model-based approaches are also applicable to high-risk subpopulations and populations and at the local level.

The objectives for the control of HIV/AIDS in the 12th FiveYear Plan period (2011-15) are to keep the total number of people living with HIV/AIDS (PLHIV) within 1.2 million, reduce the number of newly infected people by 25 percent and the number of HIV/AIDS deaths by 30 percent. Our model analysis showed that the control target for PLHIV can be realistically achieved, but it is difficult to achieve the reduction targets in new HIV infections and HIV/AIDS deaths by 2015. Based on our model projections, if China continues to increase HIV testing and treatment and continues to improve treatment effectiveness to reduce mortality among people receiving treatment, a turning point for the HIV incidence can be achieved by the end of 2013, and the reduction targets in new infections and deaths can be achieved by 2020. If the efforts in HIV testing and treatment work hand in hand with well-implemented public awareness campaigns, targeted education programs and other preventive measures to reduce HIV transmission, it is highly likely that these reduction targets can be reached during the next Five-Year Plan period (2016-2020).

4.3 Project 2: **The HIV dataset for two remote townships in Sichuan Province, China.**

A major challenge to the effective control of the HIV epidemics in China is the heterogeneity of populations, economic conditions, social structures and local HIV transmission dynamics across its vast and diverse geographic regions. At the national level, it was estimated that about 740,000 people are living with HIV (PLHIV) in 2009, and additional 105,000 people are surviving AIDS (CD4 count < 200 /mm^3) [83]. While all of the country's 31 provinces and autonomous regions have reported HIV cases, the top six high-prevalence provinces (Yunnan, Guangxi, Henan, Sichuan, Xinjiang, and

Guangdong) in 2009 have accounted for 77.1% of the cumulative HIV reports in the country, and for over 60% of the national number of people living with HIV (PLHIV) [83]. In certain remote rural townships, the HIV prevalence among adults can be 200 times higher than the national average [84].

Liangshan Yi Autonomous Prefecture in the Sichuan Province is among the regions with the highest HIV prevalence rate, accounting for 56.4% of cumulative reported cases of HIV/AIDS in the province at the end of 2010. The severe HIV epidemic in Liangshan was attributed to the prefecture's location along one of the major drug trafficking routes to northwest and central China from the "Golden Triangle," one of the world's largest illicit heroin production and distribution centers, and to the large number of injection drug users from the remote areas in the northeastern part of the prefecture. A large migrant population of farmers, who go to other regions and provinces to seek work and travel between home and their work places, accounts for 15% of the prefecture's 4.73 million population in 2011. National HIV surveillance data has shown that migrant workers from Liangshan have contributed to the spread of HIV to 30 other provinces. To effectively control the HIV epidemics in Liangshan, the local government has partnered with the Chinese Center for Disease Control and Prevention, and Sichuan Center for Disease Control and Prevention to start in 2005 comprehensive HIV intervention and control programs in two remote rural townships, Jiudu and Muer, with a combined population of about 10,000. The interventions included expanded methadone clinics to help people off injection drugs, and expanded HIV testing and ART treatment coverage. Experiences and lessons learned from these localized programs will help to improve the HIV control programs in larger high-prevalence regions.

Among the intervention measures carried out in the Jiudu and Muer townships were two population-wide HIV screening programs in 2008 and 2010, aimed to understand the baseline for the HIV epidemic. The 2008 screening program tested close to 50% of the population in both townships for HIV. Test data showed a HIV prevalence rate of 18.32%

among adults aged between 15 and 49 [84]. In 2010, with the support from the National 11-5 Major Research Project on the Prevention and Control of HIV, Viral Hepatitis and Other Major Infectious Diseases, a population-wide physical examination was conducted in Jiudu and Muer. Baseline individual health records, including HIV status, were established for over 99% of the population in the two townships. Concurrent with the increase of HIV screening, enrollment into ART treatment programs was also greatly expanded. In 2005, there was only one HIV patient receiving ART treatment in Jiudu and Muer, and the number has risen to 166 by the end of 2010. HIV data collected in the Jiudu and Muer townships were categorized and analyzed by Dr. Ping Li in her doctoral dissertation [84].

The objective of our mathematical modeling study is to assess the impact of HIV interventions implemented at the Jiudu and Muer townships between 2005 and 2010, with specific emphasis on the two population-wide screening programs in 2008 and 2010, on the temporal trend of the HIV epidemic in the two townships. Our mathematical model for the HIV transmission dynamics was calibrated and validated using the surveillance and treatment data for the period 2005-2010 collected in the two townships [84]. The validated model produces estimations on the number of new HIV infections, people living with HIV (PLHIV), and HIV/AIDS related death for 2005-2010. The impacts of two screening programs were evaluated in terms of HIV cases averted and life years saved. Our results can inform policy on the control of HIV epidemic in Liangshan and other high prevalence regions in China.

Data source: Aggregated yearly HIV/AIDS surveillance data from 2005 to 2010 was obtained from the Surveillance Database of the China Center for Disease Control and Prevention. Aggregated data on ART treatment for the same period was obtained from the Treatment Database of the China Center for Disease Control and Prevention. Population data for the two townships was obtained from the demographic database for the Liangshan prefecture.

Parameter estimation and fitting: Same as the whole nation**,** Demographic parameters Λ and d_S are estimated by fitting the equation of the total population $N' = \Lambda - d_S N$ to the population data for the two townships from 2005 to 2009. Values of parameters d_D, d_T, ρ, and the form of function $\gamma(t)$ are estimated directly from the surveillance and treatment data. Since the HIV testing data is not available for these two townships, we assume that function $\alpha(t)$ has the same form as that for the whole nation when there is no screening interventions. From 2005 to 2007, $\alpha(t)$ is assumed to be a linearly increasing function. The two screening interventions in 2008 and 2010, which tested around half and 99% of the total population in these two townships, make us to fix $\alpha(t)$ at 0.5 in 2008 and 0.99 in 2010. Values of other parameters, including transmission coefficients $\beta_I, \beta_D, \beta_T$, death rate d_I and the initial population size I_0 for the undiagnosed compartment at the beginning of our fitting (end of 2005), cannot be estimated directly from surveillance data and need to be obtained through model fitting.

Our Matrix Decomposition method was applied to detect and resolve non-identifiability issue in the model when fitted to the two townships data. Same as dealing with the HIV dataset for the whole nation, there are two dependencies among model parameters $\beta_I, \beta_D, \beta_T, d_I, I(2005)$. We fixed the value of d_I using survival data in the literature for newly infected treatment naive people [67, 68], $\beta_D/\beta_I = 0.75$ [69, 70], and $\beta_T/\beta_I = 0.1$ [71-76].

The Nonlinear Least Squares method was applied to find the point estimates for model parameters β_I, I_0, and the coefficients for $\alpha(t)$, which minimize the sum of squared error between model output and the available surveillance and treatment data. More specifically, the following data are used in our modeling fitting:

- Annual number of new diagnosis of HIV and AIDS combined from 2005 to 2009;
- Annual number of death due to HIV and AIDS among diagnosed from 2005 to 2009;
- Annual treatment enrollment from 2005 to 2009;

- Annual number of death among diagnosed people and people in treatment from 2005 to 2009.
- Annual number of treatment failure and loss to follow-up from 2005 to 2009
- The number of total population from 2005 to 2009.

The model output has the same form as we discussed in the whole nation project.

Bayesian-based MCMC method [29, 30] was applied to calculate the 95% confidence intervals for model parameters β_I , d_I, d_D, d_T, ρ and I_0. Each time we calculate the 95% confidence interval for one parameter fixing all remaining parameters at the point estimates we obtained.

Model Validation: Validation of our model was carried out using three independent approaches. First, the goodness of fit ratio R^2=0.97 indicates an excellent fit between our model output and the data. Second, leave-one-out cross-validation method was applied to further verify the consistency of the model fit [35]. The population-wide physical examination at the two townships in 2010 provided the HIV prevalence data for over 99% of the population. The availability of such good prevalence data is rare and provides a "gold standard" for model validation. The prevalence data in 2010 was not used in our model fitting and reserved for model validation. As shown in Figure 4.15, this prevalence data point is within the 95% prediction interval of our model for the total number of people living with HIV (PLHIV) in 2010.

Parameter	Description	Best-fit value[1]	95% CI[2]	Source
β_I	Transmission rate for compartment I	0.214	[0.184, 0.237]	Fitting
β_D	Ratio of transmission rates for D and I	0.1605	[0.138, 0.17775]	[69,70]
β_T	Ratio of transmission rates for T and I	0.0214	[0.0184, 0.0237]	[71-76]
d_S	Death rate for compartment S	0.007*	[0.00503, 0.0089]	Fitting
d_I	Death rate for compartment I	0.063*	[0.045, 0.083]	[67-68]
d_D	Death rate for compartment D	0.0427*	[0.032, 0.055]	Data
d_T	Death rate for compartment T	0.0652*	[0.051, 0.077]	Data
ρ	Treatment drop-out rate	0.0169*	[0.0108, 0.025]	Data
I_0	Undiagnosed HIV positive population in 2005	813	[651, 934]	Fitting
Λ	Influx of susceptible	276	[199, 354]	Fitting
$\alpha(t)$	Time-dependent diagnosis rate	$0.0602t - 120.585^*, \quad 2005 < t \leq 2007$ $0.5, \quad 2007 < t \leq 2008$ $0.0602t - 120.585, 2008 < t \leq 2009$ $1, \quad 2009 < t \leq 2010$ $0.0602t - 120.585, \quad t > 2010$		Fitting
$\gamma(t)$	Time-dependent treatment enrollment rate	0.01886t-37.807*,		Data
R^2	Goodness of Fitting	0.97		Fitting

Table 4.4. Model parameters and their best-fit values together with 95% confidence intervals

Notes: * The unites for these parameters are 1/person/year

Results

The point estimates for all model parameter together with the corresponding confidence intervals are shown in table 4.4.

Three key assessment indicators for the HIV/AIDS dynamics: annual number of new HIV infections, annual number of HIV/AIDS related deaths, and the total number of people living with HIV (both diagnosed and undiagnosed), were estimated from 2005 to 2010 using the validated model.

The estimated annual number of new HIV infections rose from 166 in 2005 to 219 in 2010. In the meantime, the year-over-year rate of increase declined from 10.84% in 2006 to 0.92% in 2010, indicating that the momentum of rise of the HIV epidemics in the two townships has slowed down significantly during the 5-year period.

The estimated total number of people living with HIV and AIDS, including both diagnosed and undiagnosed people, increased from 931 in 2005 to 1615 in 2010. The year-over-year rate of increase declined from in 13.43% 2006 to 9.57% in 2010. The estimated HIV prevalence rate among the whole population at the end of 2010 was 14.57%.

The estimated total number of deaths among people living with HIV and AIDS increased from 53 in 2005 to 78 in 2010. The year-over-year rate of increase declined from 11.32% in 2006 to 5.4% in 2010.

Our model was used to further predict the three key indicators for the period 2011-2015, under the assumption that the treatment enrollment rate would continue to increase linearly following its pre-2010 trend, and that the diagnosis rate would maintain the same trend during the period from 2005 to 2007 before the population-wide screening.

Our model predicted that continuous scale-up of HIV testing and ART treatment will reverse the temporal trend of the HIV epidemics in the Jiudu and Muer townships; the predicted annual number of new HIV infections would decline starting from 2013 (Figure 4.15). The predicted number of new HIV infections in 2015 is 212, lower than its 2010 level. The predicted value for PLHIV in 2015 is 2208, with a HIV prevalence rate at 18.98%. The model prediction for the number of combined HIV and AIDS deaths in 2015 is 116. These predicted values would be updated using more recent HIV data between 2011-15.

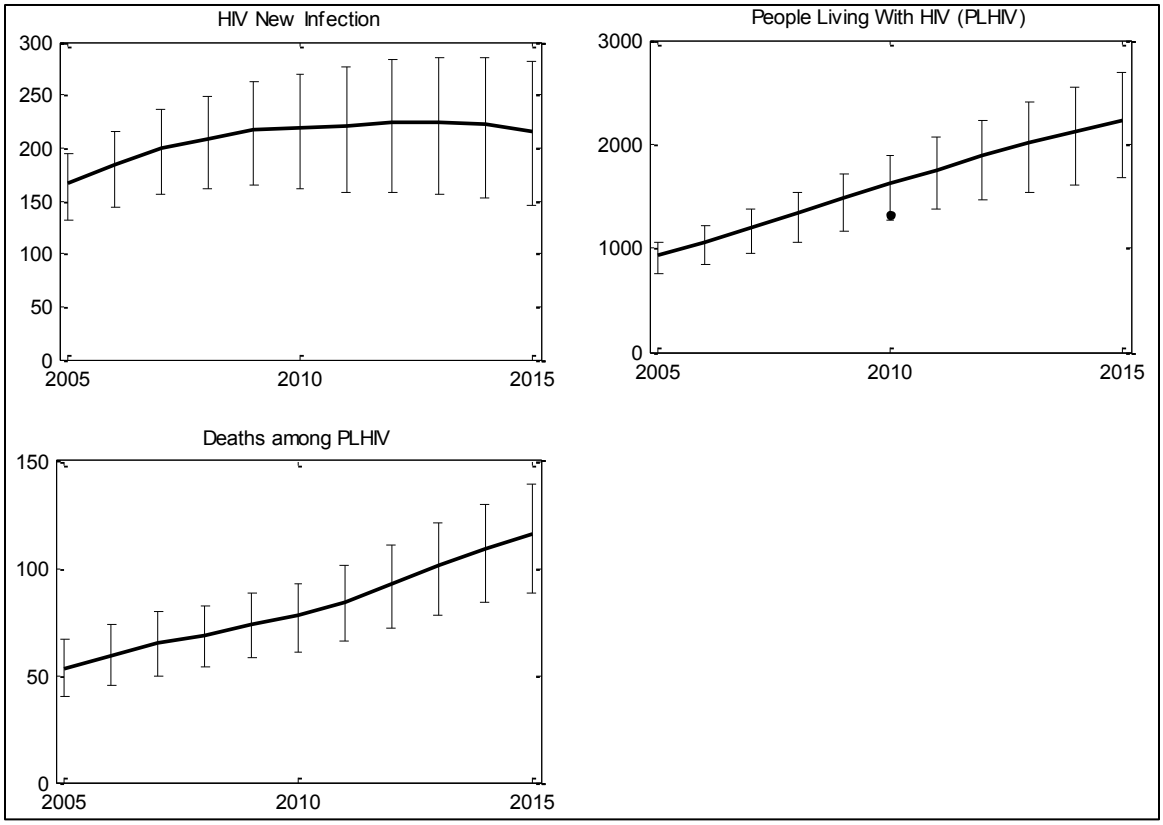


Figure 4.15 Model estimations from 2005 to 2015. Lines are baseline model estimations and vertical bars indicate 95% confidence intervals

We used our model to estimate that the two population-wide screening programs in 2008 and 2010 alone have saved 134 life years, and averted 123 new HIV infections from the beginning of 2008 to the end of 2015. The number of averted new HIV infections account for 6.6% of the new HIV infections that would have occurred without the two screening interventions. Figure 4.16 shows the cumulative number and percentage of HIV new infections averted from 2008 to 2015.

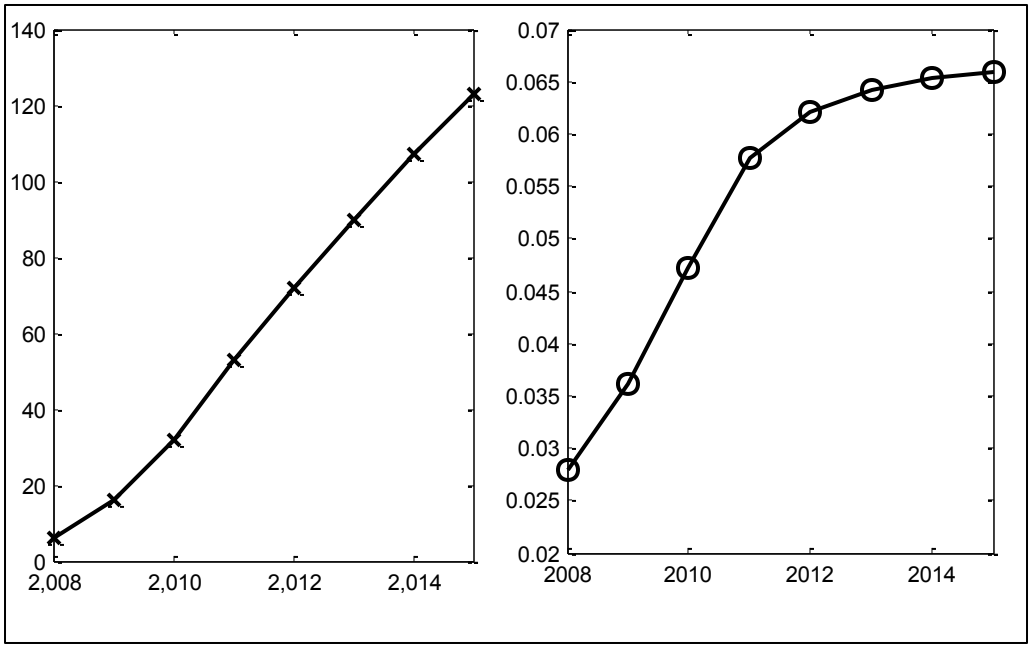


Figure 4.16 Accumulative HIV new infections averted number (left) and percentage (right) due to the screening interventions

Discussion

Mathematical models are becoming a useful tool for estimating HIV/AIDS incidence and prevalence and for predicting their temporal trends. Mathematical models can be constructed to describe HIV epidemic in a population, incorporating dynamic processes of HIV transmission, testing, treatment, as well as deaths. Mathematical models can be calibrated to a specific population by fitting them to epidemiological and surveillance data of the population. Model validation is of paramount importance for model predictions to be reliable. Validated mathematical models can be used to assess retrospectively the impacts of past intervention programs, predict temporal trends of HIV incidence and prevalence in the future, to project the effects of potential interventions, and to provide evidence to inform policies on HIV control and prevention.

In our study, we constructed and calibrated a mathematical model for the HIV epidemic among the populations in the two remote and rural townships, Jiudu and Muer, in the Liangshan Prefecture of the Sichuan Province during 2005-2010. The model was

validated by a good agreement of its predicted value of the 2010 HIV prevalence with the true prevalence data based on a HIV screening of 99% of the population in 2010. Our model estimations have shown that, while the HIV epidemic in Jiudu and Muer were rising during the period of 2005-2010, the intervention programs have significantly slowed down the rising momentum. Our model further predicted that continuing the scale-up of HIV intervention programs, especially HIV testing and ART treatment as prevention of HIV transmission, would revert the temporal trend of the HIV epidemic in the two townships, and the annual new HIV infections would decline starting as early as 2013. These predictions will be updated when more recent data becomes available.

For the period from 2008 to 2015, our model projected that the two population-wide screening programs in 2008 and 2010 have averted 123 HIV new infections and saved 134 life-years from the prevented HIV deaths. Considering the majority of people who live with HIV and AIDS in the two townships were male farmers of age from 25 to 49, and most of them were the only laborer and bread maker for their families, the social and economic impact of the screening programs are significant on the local population.

There are several limitations in our study due to the limited number of data points. The heterogeneity among the people living with HIV and AIDS was not considered in our model. The transmission coefficients β_I, β_D and β_T in our model were averaged among people in different high-risk groups. Incorporating risk groups will make our model more realistic, while calibration of a more complex model will require a larger number of time points in the data than what were available to us. In our parameter estimation, the death rate d_T for the population under treatment was larger than the death rate d_D of people who were diagnosed of the HIV but not in treatment programs. This was likely because most patients receiving treatment in 2005 were in the late stage of HIV infection and suffered a higher fatality rate. As the HIV testing and treatment being continuously scaled up, we expect that a larger proportion of the people receiving treatment will be in the early stage of infection, and the fatality rate will drop. It will be

more reasonable to assume that the death rate d_T is time-dependent and decreases in time. This would be possible when a longer time series of data become available.

Our study demonstrated that mathematical models could be an effective research tool for HIV epidemiology. When integrated with epidemiological and surveillance data, models can produce estimations that are not readily available from standard epidemiological studies, including but not limited to the number of new HIV infections, the size of the hidden (undiagnosed) HIV positive populations, and death among the hidden HIV positive populations. These estimations can provide health authorities with reliable assessments for the true burden of the HIV epidemic and evidence for planning of HIV control and interventions.

4.4 Project 3: The HIV dataset for Guangxi province, China

Since China adopted an open-door economic policy in 1978, with its fast economic developments and profound social changes, the country has witnessed a rapid reemergence of drug trafficking and abuse, commercial sex, and sexually transmitted diseases (STDs), as well as the emergence of an HIV epidemic [85]. The first reported case of an HIV-infected injecting drug user (IDU) in China was reported in 1989 in the southwestern province of Yunnan, along the border with Myanmar. By 2002, HIV infections among IDUs had been reported in all 31 mainland provinces, and 71% of the infections in China were attributed to injection drug use by government estimation [86]. The high HIV prevalence among drug users in Southwest China is largely due to its close proximity to the world's major heroin producing area, known as the Golden Triangle. The majority of heroin in China's market is brought from Myanmar into Yunnan or from Vietnam into Guangxi, then transported to the southwestern and western provinces of Sichuan, Guizhou, and Xinjiang [87-90]. China is no exception to the evolution of the HIV epidemic occurring in other Asian countries, which began as a drug-driven epidemic and is shifting to one driven predominantly by sexual contacts [85]. Since sexual contacts are the key mode of transmission of HIV from high-risk groups to the general population,

the HIV epidemic in China has entered a new phase. The government has shown strong leadership responding to HIV/AIDS. Under the national comprehensive HIV and AIDS response policy, the central government implemented the "Four Frees and One Care" program in 2003, which provides free voluntary counseling and testing, free antiretroviral therapy, free prevention of mother to child transmission (PMTCT), free education to AIDS orphans, and financial assistance and social support to HIV and AIDS patients [91]. Furthermore, harm reduction programs such as knowledge education, HIV testing and counseling, promotion of condom use, methadone maintenance therapy and needle exchange have been rapidly scaled-up and implemented at the national and local levels since 2004 [92-96].

Guangxi Zhuang Autonomous Region is located in southern China bordering Vietnam and has a population close to 50 million. It is located along the major heroin trafficking route linking Guangxi with Yunnan bordering Myanmar and eastern-ward to Guangdong and Hong Kong. Its HIV epidemic was fueled primarily by injection drug use at the beginning of the epidemic [95-99]. According to the Guangxi Center for Disease Control and Prevention (Guangxi CDC), HIV prevalence among the IDU population ranged between 11% and 60% in 2003, based on data collected from sentinel sites [100]. Guangxi has been among the top three provinces that reported the highest number of HIV and AIDS cases in China since 2005 [100]. While the HIV infection amongst injecting drug users accounted for 69% of reported HIV cases in 2003, there has been a constant increase of HIV infections through sexual contacts since then, and sexual contacts have accounted for 66% of the reported HIV cases by 2009 in Guangxi [100]. In response to this severe HIV epidemic, the government of Guangxi launched a 5-year Guangxi AIDS Conquering Project (GACP) in 2010. The GACP has doubled the investment in financial and human resources that were required for implementing the national "Four Frees and One Care" program, and further expanded HIV testing and antiretroviral therapies. At the completion of the 5-year program, assessment of its impact on Guangxi's HIV

transmission and epidemic is important for informing evidence-based decision making and future HIV control policies.

The HIV epidemiology and treatment data are collected, tabulated and reported vertically by hospitals and the CDC networks in each province, autonomous region and municipality of China. The provincial CDC provides coordination and quality control for the data collection. Data from all provinces, autonomous regions and municipalities are reported to the Chinese Center for Disease Control and Prevention (China CDC) to form the national HIV epidemiology and antiretroviral treatment databases, which are analyzed at the national level for China's HIV epidemic trends and for formulating the national control policy [101, 102]. The aim of our study is to understand the Guangxi HIV trends in reference to the national context, by analyzing Guangxi CDC datasets from 2005 to 2014 using epidemic surveillance and mathematical modelling tools. Results of our study provided an assessment of the performance and impacts of GACP. The study results further provided scientific evidences to inform future provincial AIDS control policies for Guangxi Zhuang Autonomous Region.

Data source: HIV epidemic reports, surveillance data and ART treatment data were collected by the Guangxi Center for Disease Control and Prevention from 2005 to 2013. The datasets included the provincial HIV/AIDs case reports, HIV testing, HIV sentinel surveillance, HIV testing for marriage, pregnancy and child delivery, antiretroviral treatment, etc. Population data was collected from the Guangxi Zhuang Autonomous Region Bureau of Statistics.

Parameter estimation and fitting: Same as the whole nation**,** Demographic parameters Λ and d_S are estimated by fitting the equation of the total population $N' = \Lambda - d_S N$ to the population data for the two townships from 2005 to 2013. Values of parameters d_D, d_T, ρ, and the form of function $\gamma(t)$ are estimated directly from the surveillance and treatment data. $\alpha(t)$ was estimated from HIV testing data of Guangxi. Values of other parameters, including transmission coefficients $\beta_I, \beta_D, \beta_T$, death rate d_I and the initial population size

I_0 for the undiagnosed compartment at the beginning of our fitting (end of 2005), cannot be estimated directly from surveillance data and need to be obtained through model fitting.

Our Matrix Decomposition method was applied to detect and resolve non-identifiability issue in the model when fitted to the two townships data. Same as dealing with the HIV dataset for the whole nation, there are two dependencies among model parameters $\beta_I, \beta_D, \beta_T, d_I, I(2005)$. We fixed the value of d_I using survival data in the literature for newly infected treatment naive people [67, 68], $\beta_D/\beta_I = 0.75$ [69, 70], and $\beta_T/\beta_I = 0.1$ [71-76].

The Nonlinear Least Squares method was applied to find the point estimates for model parameters β_I and I_0, which minimize the sum of squared error between model output and the available surveillance and treatment data. More specifically, the following data are used in our modeling fitting:

- Annual number of new diagnosis of HIV and AIDS combined from 2005 to 2013;
- Annual number of death due to HIV and AIDS among diagnosed from 2005 to 2013;
- Annual treatment enrollment from 2005 to 2013;
- Annual number of death among diagnosed people and people in treatment from 2005 to 2013.
- Annual number of treatment failure and loss to follow-up from 2005 to 2013
- The number of total population from 2005 to 2013.

The model output has the same form as we discussed in the whole nation project.

Bayesian-based MCMC method [29, 30] was applied to calculate the 95% confidence intervals for model parameters $\beta_I, d_I, d_D, d_T, \rho$ and I_0. Each time we calculate the 95% confidence interval for one parameter fixing all remaining parameters at the point estimates we obtained.

Model Validation: First, the goodness of fit ratio R^2=0.97 indicates an excellent fit

between our model output and the data. Second, leave-one-out cross-validation method was applied to further verify the consistency of the model fit.

Results

The point estimates for all model parameter together with the corresponding confidence intervals are shown in table 4.5.

Parameter	Description	Best-fit value[1]	95% CI[2]	Source
β_I	Transmission rate for I	0.1733	[0.1512, 0.1921]	Fitting
β_D	Ratio of transmission rate in D and that in I	0.13	[0.1186 0.1431]	[69,70]
β_T	Ratio of transmission rate in T and that in I	0.01733	[0.01534, 0.01912]	[71-76]
d_S	Death rate for S	0.007*	[0.0059 0.00805]	Fitting
d_I	Death rate for I	0.063*	[0.04501, 0.0803]	[67, 68]
d_D	Death rate for D	0.115*	[0.09873, 0.1597]	Data
d_T	Death rate for T	0.0321*	[0.02318, 0.0407]	Data
ρ	Treatment drop-out rate	0.0769*	[0.05898, 0.09576]	Data
I_0	Undiagnosed HIV positive population in 2006	47940	[44401, 50905]	Fitting
Λ	Influx of susceptible	820738	[750143 882361]	Fitting
$\alpha(t)$	Time-dependent diagnosis rate	$0.03345 * (t - 2006) + 0.155^*, \quad t \leq 2009$ $0.0543 * (t - 2009) + 0.25, \quad t > 2009$		Data
$\gamma(t)$	Time-dependent treatment enrollment rate	$0.004195 * t - 8.2738^*, \quad t \leq 2009$ $0.059966 * t - 120.3177, \quad t > 2009$		Data
R^2	Goodness of Fitting	0.96		Fitting

Table 4.5. Model parameters and their best-fit values together with 95% confidence intervals.

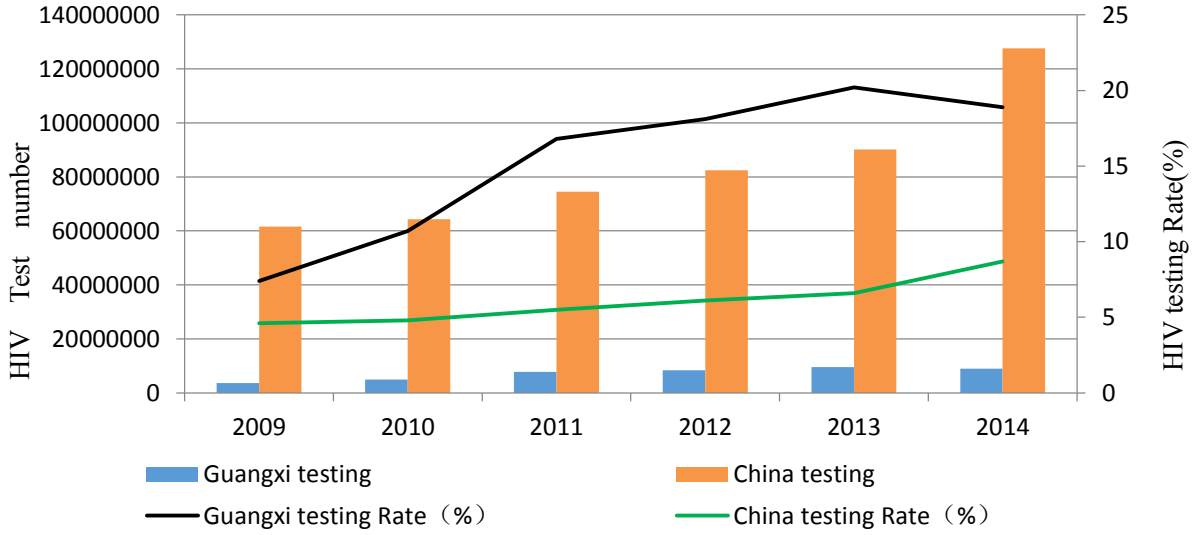
Notes: * The unites for these parameters are 1/person/year

HIV testing, ART and case reporting

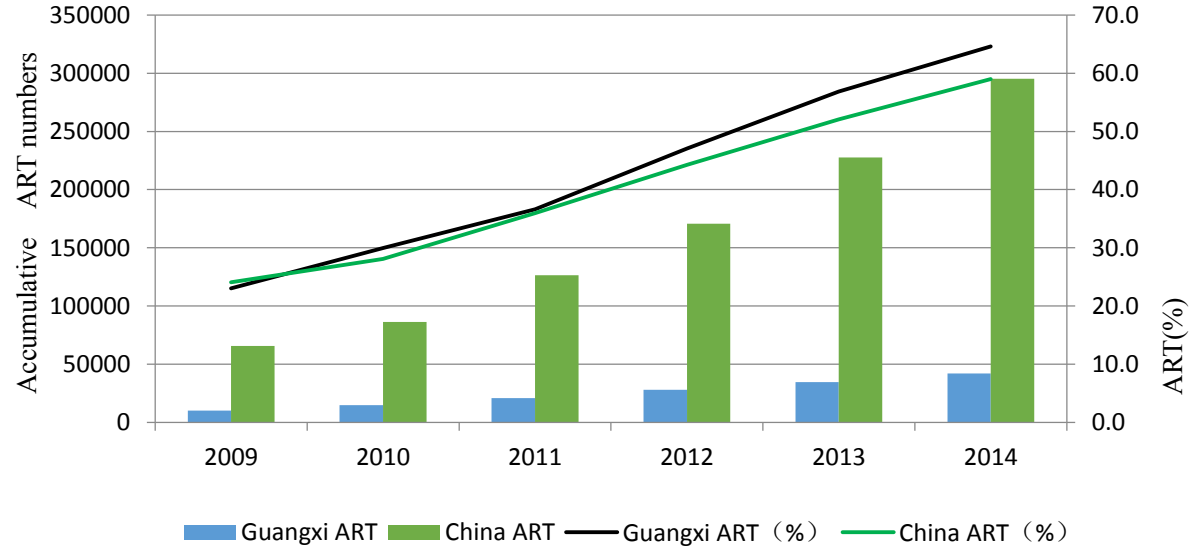
Under GACP and national Four Free One Care policy, the AIDS surveillance, intervention and ART activities have been vigorously scaled up. The person-time for HIV testing in Guangxi has jumped from 3.6 million in 2009 before the GACP to 7.8 million in 2011, and further to more than 9.5 million in 2013 (Figure 4.17a, and Table 4.6a). The

ART treatment for PLWHA has increased from 10035 persons in 2009 before the GACP to 41859 persons in 2014 at the end of the GACP (Figure 4.17b, and Table 4.6b). As the result of the scaling up of HIV testing and expanded ART treatment as well as other intensified intervention measures during the GACP campaign, Guangxi's HIV/AIDS case reports peaked in 2011 and then started to decline since 2012 for 3 consecutive years, even as its HIV testing numbers continuously to rise to its peak in 2013, almost tripling the testing numbers before the GACP campaign in 2009 (Figure 4.17c, and Table 4.6c).

a)



b)



(c)

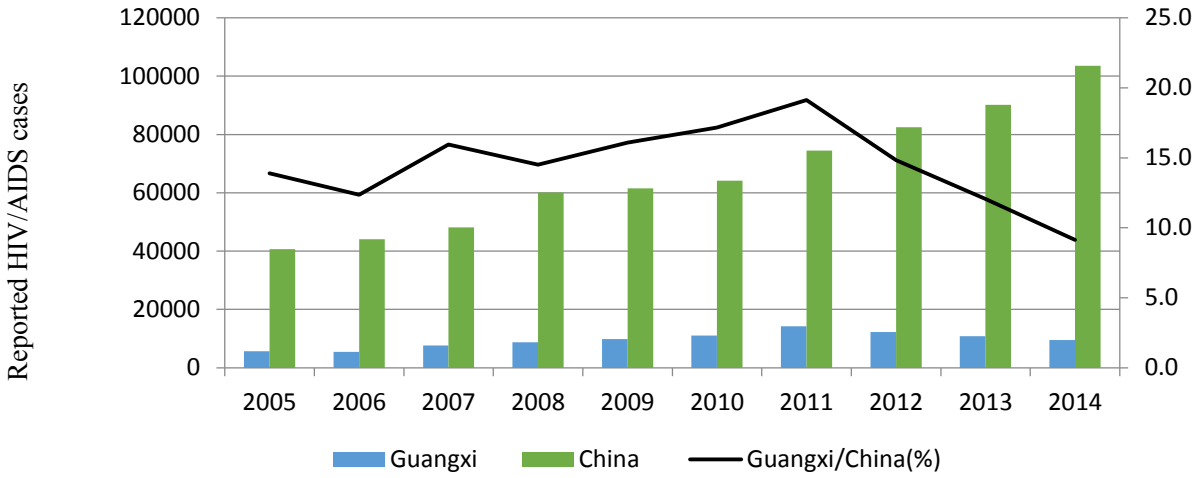


Figure 4.17：The HIV testing, ART and HIV/AIDS reports in Guangxi and China

a) The HIV testing rate in Guangxi and China.

	HIV Tests		Population (millions)		testing Rate (%)	
Year	China Average	Guangxi	China Average	Guangxi	China	Guangxi
2009	61626000	3602517	1334.5	48.56	4.6	7.4
2010	64338000	4924400	1340.91	46.1	4.8	10.7
2011	74517000	7798630	1347.35	46.45	5.5	16.8
2012	82434000	8471933	1354.04	46.82	6.1	18.1
2013	90119000	9540710	1360.72	47.19	6.6	20.2
2014	127560180	8976867	1367.82	47.54	8.7	18.9
Average rate (10-14)					5.3	14.1

b）ART coverage rate in Guangxi and China.

	Current ART Number		Accumulated PLWHA		ART Rate (%)	
Year	China	Guangxi	China	Guangxi	China	Guangxi
2009	65481	10035	271871	43568	24.1	23.0
2010	86122	14830	306732	49528	28.1	29.9
2011	126448	20874	351709	57019	36.0	36.6
2012	170655	27893	385871	59240	44.2	47.1
2013	227489	34496	436817	60687	52.1	56.8
2014	292538	41859	500679	64792	34.8	36.8
Average rate (10-14)					45.7	48.0

c) HIV/AIDS reported cases in Guangxi and China

Reported HIV/AIDS cases	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**
Guangxi	5652	5450	7685	8715	9894	11007	14250	12229	10877	9460
China total	40711	44070	48161	60081	61470	64108	74517	82434	90119	103501
Guangxi/China(%)	13.9	12.4	16.0	14.5	16.1	17.2	19.1	14.8	12.1	9.1

Table 4.6 The HIV testing, ART and HIV/AIDS reports in Guangxi and China

The control reproduction number R_c

The control reproduction number R_c measures the expected number of secondary cases produced from a single infection source in a susceptible population in the presence of control measures [77, 103]. Unlike the basic reproduction number R_0, which is typically calculated at the beginning of an epidemic and will not vary with time, R_c takes into account the effects of control measures and may vary as the epidemic progresses. In our model, effects of the national "Four Frees and One Care" program between 2005 and 2009 and the Guangxi 5-year GACP after 2010 were incorporated into the time dependent diagnosis rate $\alpha(t)$ and ART enrolment rate $\gamma(t)$. The impact of the response

programs during the two time periods 2005-2009 and 2010-2014 will be measured by the time-varying control reproduction number R_c.

Sensitivity analysis for R_c

Sensitivity of R_c to variations of model parameters was measured by local sensitivity analysis using the one-at-a-time method [32, 33]. Results of the sensitivity analysis at the years 2005, 2010 and 2014 are shown in Figure 4.18b, c and d. In 2005, the most sensitive parameters for R_c are the transmission coefficient β_I and death rate d_I of the undiagnosed population in compartment I. With the scale-up in HIV testing and ART coverage, more HIV positive people were diagnosed. Accordingly, in 2010, the most sensitive parameters for R_c are the transmission coefficient β_D and death rate d_D of the diagnosed population in compartment D. In 2014, as the result of further scale-up of ART treatment under the 5-year GACP, the death rate d_T of the population under ART becomes highly sensitive for R_c (Figure 4.18).

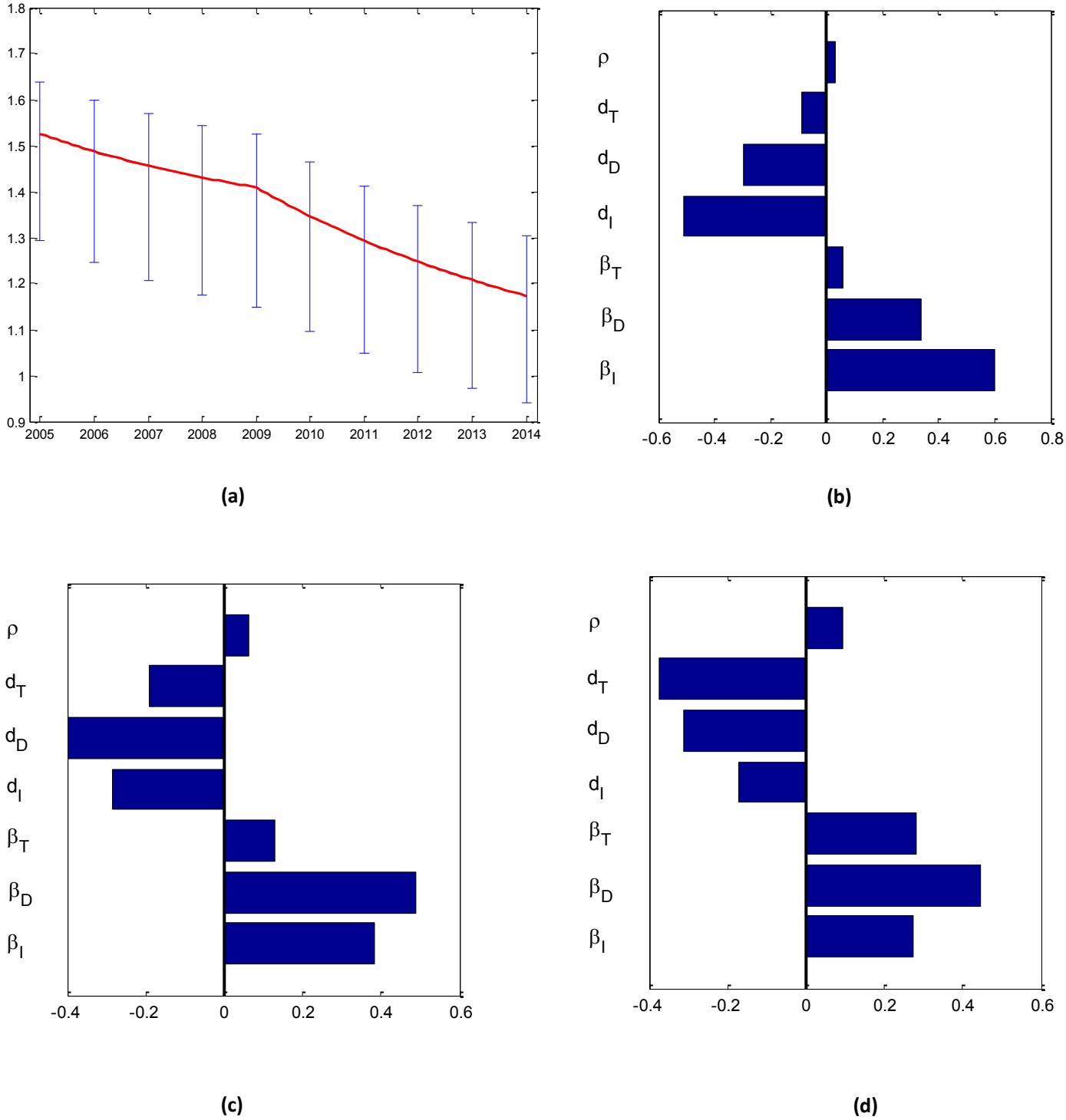


Figure 4.18: Estimated annual value of R_c and its sensitivity analysis

Uncertainty analysis for R_c

We used the Latin Hypercube Sampling (LHS) method [31] to compute the 95% confidence intervals for R_c from 2005 to 2014, when all parameters are sampled from their 95% confidence intervals given in Table 4.5 We generated 20,000 samples using LHS, and results of the uncertainty analysis are given in Figure 4.18.

Comparative study of alternative intervention scenarios.

Our calibrated and validated model was used to carry out comparative studies of several hypothetical options under which the 5-year GACP could have been implemented. The baseline case was the actual implementation of the 5-year GACP from 2010 to 2014, in combination with the national HIV/AIDS response program in Guangxi. In this case, the scale-up in the levels of HIV testing and ART treatment was modeled by the time dependent diagnosis rate $\alpha(t)$ and ART enrollment rate $\gamma(t)$. Note that the slopes of $\alpha(t)$ and $\gamma(t)$ are greater during the 5 years of GACP than those during the period 2005-2009 when only the national response program was implemented. For comparisons, we have modified assumptions on $\alpha(t)$ and $\gamma(t)$ during the period 2010-2014 according to the following three alternative scenarios:

- Scenario 1: The trend of HIV testing and ART treatment continues that of the national program of 2005-2009;
- Scenario 2: The trend of HIV testing continues that of the national program, and the trend of ART treatment follows the baseline;
- Scenario 3: The trend of HIV testing follows the baseline, and the trend of ART treatment continues that of the national program.

Control production numbers R_c for the three alternative scenarios were calculated and compared to that of the baseline case.

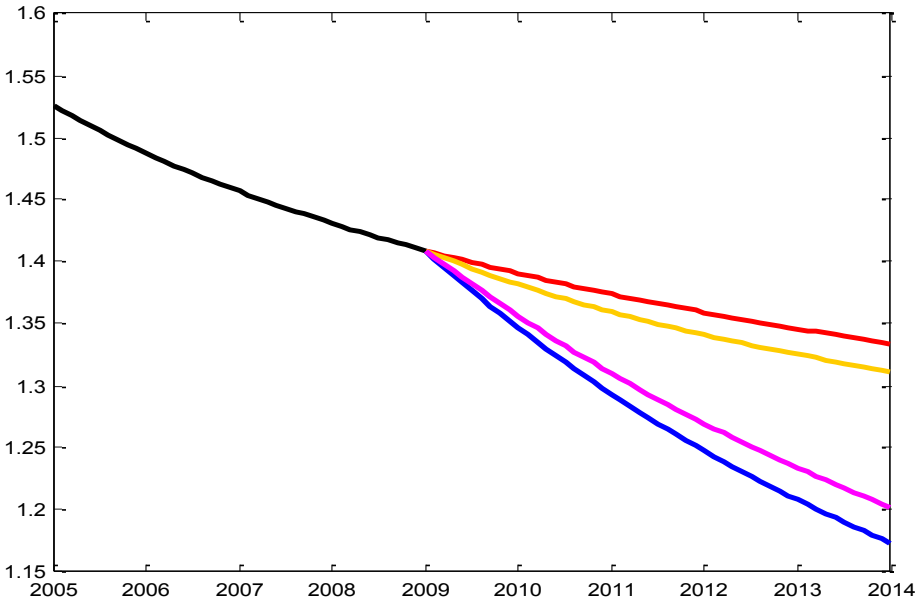


Figure 4.19: The estimations of the AIDS control effectiveness under various situations (blue:baseline, red:Scenario 1, pink:Scenario 2, Blue:Scenario 3)

Figure 4.19 shows that combining the national control program with GACP has produced the most reduction in R_c, from 1.526 in 2005 to 1.174 in 2014. In the alternative Scenario 1, R_c has a gradual decline from 2005 to 2014, since the scale-ups in testing and ART coverage were assumed to be at the national level throughout the period. It is interesting to compare Scenarios 2 and 3: while both have resulted in a substantial reduction in R_c either by scaling up to the GACP level of the ART (Scenario 2) or the HIV testing (Scenario 4), a greater reduction is achieved by scaling up the ART than scaling up the HIV testing. This is in part because people under ART treatment have much smaller transmission coefficients than those who are not.

HIV transmission is mainly through sexual contacts in the world, and in China since 2009. The HIV epidemic in Guangxi has been driven by sexual transmission since 2007. As shown in Figure 4.20 and Table 4.7, after the implementation of GACP, report on sexually transmitted diseases in Guangxi has shown a continuous and significant drop.

The gonorrhea case reports in Guangxi showed a two-digit annual decline (10.5 - 15.8%) from 7822 cases in 2010 to 4477 cases in 2014. The syphilis case reports in Guangxi first experienced an increase and then a continuous sharp decrease (25.5 - 50%) from 41673 cases in 2010 to 11416 cases in 2014. The STD reports are generally expected to be in line with the HIV/AIDS reports in the same population, and case reports for both HIV and STD have shown a significant decline in Guangxi (Figures 4.20).

a) Guangxi and China syphilis report

	2010	2011	2012	2013	2014
Guangxi	41673	46699	34742	15330	11416
China	399565	447525	465713	464292	471312

b) Guangxi and China gonorrhea report

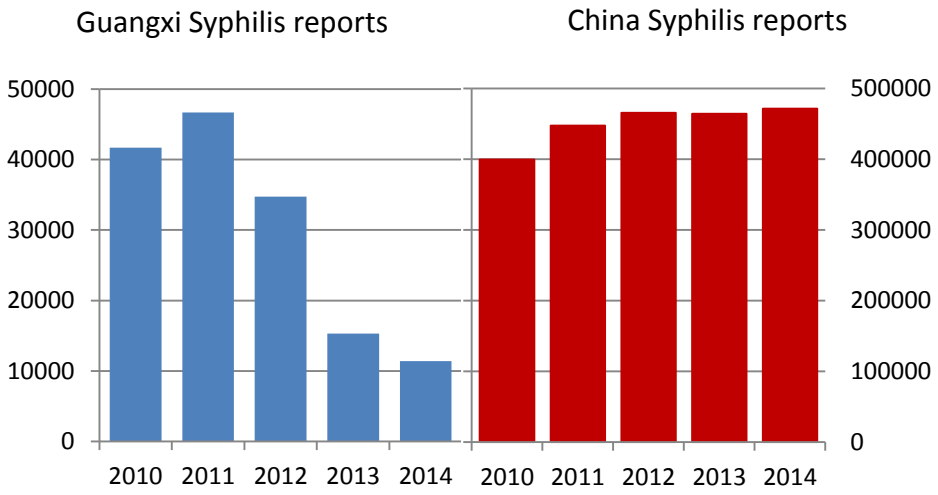
	2010	2011	2012	2013	2014
Guangxi	7822	6997	5891	5273	4477
China	109299	103728	96890	104245	99482

c) Year on year comparison on syphilis and gonorrhea report for Guangxi and China.

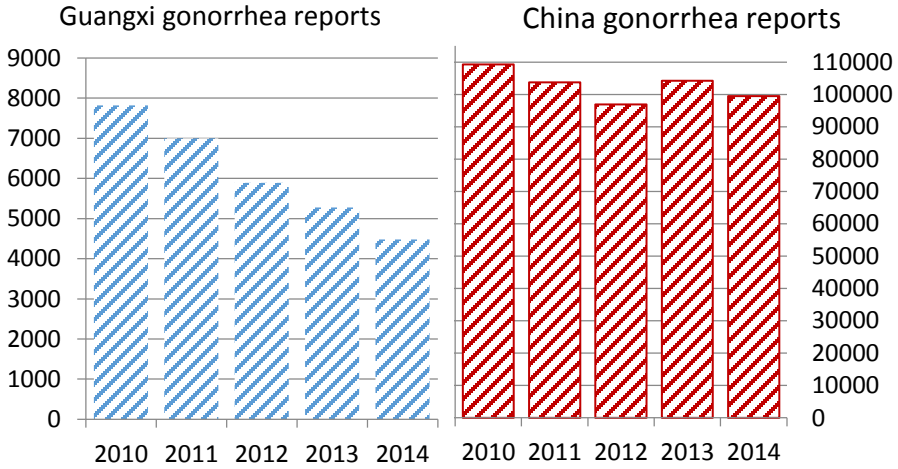
	2010	2011	2012	2013	2014
Guangxi Syphilis	24.1	12.1	-25.6	-55.9	-25.5
Guangxi Gonorrhea	-12.9	-10.5	-15.8	-10.5	-15.1
China Syphilis	17.7	12	4.1	-0.3	1.5
China Gonorrhea	-11.6	-5.1	-6.6	7.6	-4.6

Table 4.7. Syphilis and Gonorrhea Reports in Guangxi and China

(a)


Guangxi Syphilis reports
China Syphilis reports
50000
40000
30000
20000
10000
0
500000
400000
300000
200000
100000
0
2010 2011 2012 2013 2014
2010 2011 2012 2013 2014

(b)


Guangxi gonorrhea reports
China gonorrhea reports
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
110000
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
2010 2011 2012 2013 2014
2010 2011 2012 2013 2014

(c)

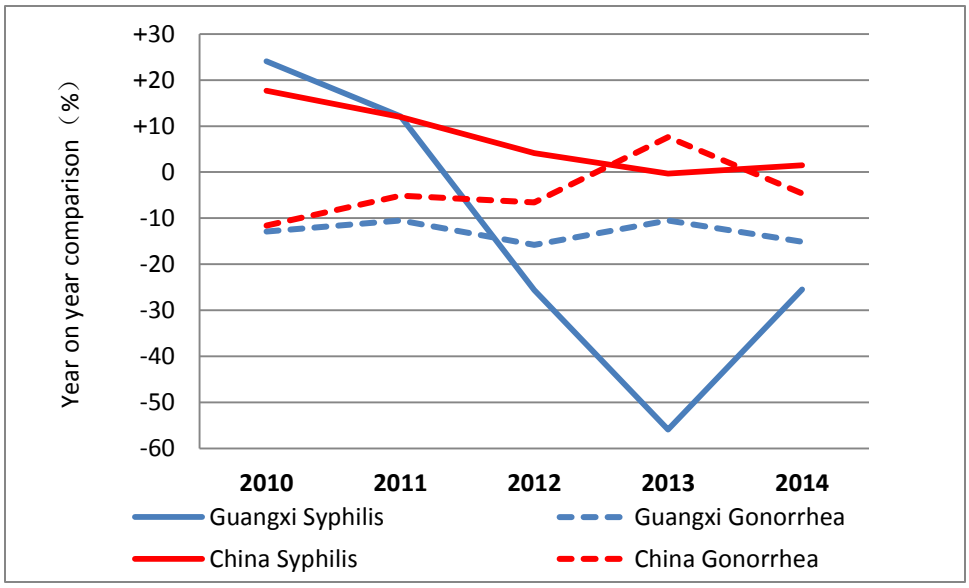


Figure 4.20: Syphilis and Gonorrhea Reports in Guangxi and China

Discussion

Facing the challenge of a rapidly expanding AIDS epidemic, the Guangxi government put AIDS control as a top priority of the government agenda. They formulated and implemented a five-year Guangxi AIDS Conquering Project (GACP) in 2009 with unprecedented efforts. Governments at all levels in Guangxi invested a total of more than 2.25 billion Yuan (RMB) to AIDS control programs and opened 1482 permanent government positions for AIDS control offices and public health agencies. With these efforts, Guangxi ranked number one in local governmental financial commitment and human resources investments to AIDS control programs, among China's 31 provinces, autonomous regions and municipalities. Guangxi is also among the few provinces in China that passed provincial AIDS control legislation to ensure governmental and societal commitment, as well as to safeguard local AIDS control programs by law.

Such efforts have resulted in the most rapid scaling-up of HIV testing and antiviral treatment campaigns in Guangxi during the past 5 years in comparison to other provinces in China and the national average (Figure 4.17and Table 4.6). Guangxi's average HIV

testing strength during the implementation of GACP was 14.1 per 100 population, 2.67 fold higher than that of the national average. The Guangxi ART coverage rate has also caught up rapidly and was kept at a higher than national average throughout the GACP period. The unprecedented scale-up in testing and treatment in Guangxi has resulted in a significant decrease in the reported HIV/AIDS cases from 14250 in 2011 to 9460 in 2014, while the reported cases in China increased from 74517 to 103501 during the same period (Figure 4.17c, and 4.6c). The significant drop of the reported HIV/AIDS cases in Guangxi in three consecutive years, concurrent with the continuous scale-up in the number of HIV tests, suggests that the local HIV incidence may be declining.

We used a mathematical model to assess quantitatively the impact of the GACP on the HIV transmission dynamics in Guangxi. Based on our analysis of Guangxi's HIV surveillance and treatment data from 2005 to 2009, the control reproduction number R_c was reduced annually by 1.92% from 2005 to 2009. Our results demonstrate the efficacy of China's national comprehensive AIDS response policy and the "Four Frees and One Care" program in Guangxi during the period. Meanwhile, a greater annual reduction of 3.34% in R_c was estimated for the period from 2010 to 2014, which provides strong evidence in support of the Guangxi government's decision to implement the 5-year GACP. The overall reduction of 16.7% in the value of R_c during Guangxi's 5-year GACP demonstrates the effectiveness of the campaign.

Our mathematical model describes the HIV transmission in Guangxi by following the flow of the HIV positive population, from being infected, to being tested and diagnosed, and to being enrolled into ART treatment. The model was set up differently from the standard SIR formulation for ease of comparison to the diagnosis and treatment data. The model was calibrated and validated using the HIV surveillance and ART treatment data from the Guangxi CDC. To reflect China's efforts in scaling up HIV testing and treatment since 2005 [104], we have used time-dependent rates for HIV diagnosis and ART enrollment from 2005 to 2014. By comparing the differences in

diagnosis rates and ART enrollment rates during the two periods 2005-2009 and 2010-2014, we were able to compare the impacts of the national HIV and AIDS response program implemented in Guangxi during 2005-2009 to that of Guangxi's 5-year GACP during 2010-2014, in terms of the control reproduction number.

The control reproduction number R_c is an important measure for assessing the burden of HIV transmission in a population and evaluating the effectiveness of the implemented AIDS control program. A value of R_c below 1 is an indication that the epidemic is under control, as long as the current control and prevention efforts are maintained. If the value of R_c is above 1, then further scaling-up of the testing and ART coverage will be needed, or other prevention programs need to be further expanded to complement the test-and-treat programs. Based on our study, the value of R_c for the Guangxi population showed a moderate annual decline of 1.92% from 2005 to 2009 under the national HIV and AIDS response program. A sharp annual decline of 3.34% in the value of R_c is achieved from 2010 to 2014 under the Guangxi GACP implementation. The impact of reducing the value of R_c from 1.409 at the end of 2009 to 1.174 in 2014 by the GACP indicated that the expanded HIV testing and ART coverage have significantly slowed down the HIV transmission in Guangxi. While the program in Guangxi is shown to be successful, the value of R_c is still above 1 at the end of 2014. Further scaling-up of HIV testing and ART, complemented by an expansion of harm reduction preventions - increasing the rate of condom use, reducing high-risk behaviors among IDUs and unprotected sex with clients among FSWs - as part of the "Four Frees and One Care" national policy [105, 106] will be necessary beyond 2014 to bring the epidemic fully under control.

Results of sensitivity analysis on our model results show that, with more HIV patients being diagnosed and treated, the most sensitive parameters for the value of R_c will gradually shift from β_I to β_D. This implies that reducing the transmission among people who are diagnosed with HIV but not in ART programs by enhancing the

harm reduction measures and HIV education campaigns can be highly effective in reducing the overall transmissibility of HIV, as the HIV testing and ART treatment are further scaled up.

Several alternative scenarios for the scaling-up of HIV testing and ART coverage were studied retrospectively using our calibrated model. The results show that, if the 5-year GACP only scaled up the ART coverage after 2009 while maintained level of HIV testing at that of the national response program, the value of R_c would be reduced to 1.2 in 2014. If the HIV testing was scaled up while the ART coverage was maintained at the national program level, the value of R_c would only be reduced to 1.31 in 2014. This suggests that scaling-up of ART coverage will be more effective in controlling the HIV epidemic, as people under ART treatment have a greatly reduced transmissibility.

Our study has limitations. Heterogeneity among the HIV positive population is not considered in our model. Transmission rates β_I, β_D, and β_T are estimated as the average over all high-risk groups and all age groups. Further studies with improved models that take the heterogeneity into account are needed to produce more accurate assessments of the HIV epidemic. Expanded ART treatment is expected to improve the survival of people living with HIV and AIDS. The overall fatality rates among people under treatment are expected to decline with the scale-up of ART. Due to limitations of the available data, we have used a constant death rate for the population under ART throughout the 10-year study period, which may result in a slight over-estimation of the value of R_c towards the end of the study period.

Heterosexual transmission accounted for over 80% and 90% of reported HIV/AIDS cases in Guangxi in 2011 and 2013, respectively. This indicates that the HIV epidemic in Guangxi was predominantly driven by heterosexual transmissions. For this reason, the Guangxi STD database can be a good reference in the evaluation of the local HIV epidemic. The national STD reports showed a rise of syphilis from 399,565 cases in 2010 to 471,312 cases in 2014, while the gonorrhea reports experienced a modest decline from

109,299 in 2010 to 99,482 in 2014 (Figure 4.20, and Table 4.7). This is in contrast to the rapid decline in both syphilis and gonorrhea reports observed in Guangxi (Figure 4.20, and Table 4.7). The STD epidemic reports are in good agreement with the HIV/AIDS reports; both showing an increasing trend nationally, and both showing a decline in Guangxi (Figures 4.17 and 4.20). This again indicates that the HIV incidence in Guangxi may be declining.

Effectiveness of testing and ART treatment as a HIV prevention strategy has not previously been assessed retrospectively in a real-world setting of developing countries [106-108]. Our study is the first to use mathematical models to analyze and assess retrospectively the impact of HIV testing and treatment as prevention strategies in a major Chinese HIV epidemic province with a population of 50 million. This study can provide real-time data and evidence in support of the test-and-treat HIV prevention strategy in the setting of a developing country with a concentrated HIV epidemic.

Results of our modelling study and HIV and STD surveillance data analysis strongly support the public health policy of expanded HIV testing and ART as a mean to prevent HIV transmission in the setting of developing countries. Our results also show that, to achieve full control of the HIV epidemic, it is necessary to combine national strategy and local AIDS control programs. The national strategy can provide a political framework and policy environment for the local AIDS control programs. Local initiatives such as the GACP in Guangxi can provide targeted interventions tailored to the local situation with needed resources. Our study further shows that the control reproduction number R_c needs to be further reduced to stop the HIV transmission in Guangxi. To achieve this ultimate goal, the test and treat strategy needs to be complemented by social behavior interventions, such as various harm reduction programs for the traditional high-risk groups, as well as the education programs for the general population. This again requires combining a national strategy with local initiatives to fight AIDS. It is very encouraging to see that the Guangxi government is initiating the Phase II of GACP (2015-2019) and

the Chinese government is making the nation's Thirteenth Five-Year Plan (2016-2020) to fight AIDS.

In this chapter, we built a single group SIDT model to study different HIV datasets from China CDC. The single group SIDT model describes the HIV process of susceptible-infected-diagnosed-treated. For each dataset, it consists of HIV diagnosis related data and HIV treatment related data. Non-identifiability existed when we fitted the model to different HIV datasets, which were detected and resolved using the method we discussed in Chapter 3. After that, model parameters were uniquely estimated based on different HIV datasets.

In the first project, we studied the HIV dataset for China. After model parameters were uniquely estimated based on this dataset, the HIV epidemic in China was estimated and predicted. The results were compared to the estimates published by the National HIV Epidemic Estimation Working Group

In the second project, we studied HIV datasets for two remote townships in Sichuan province of China. Same as the first project, the HIV epidemic in these two townships was estimated and predicted based on uniquely estimated model parameters. This project provided a validation for our model, since a population-wide HIV screening being conducted in the two remote townships in 2010, about 99% of the residents and mobile population being tested for HIV, which provides a rare data for the true HIV prevalence, and a gold standard for model validation.

In the third project, we studied HIV dataset for Guangxi province in China. After model parameters were uniquely estimated based on the dataset, the effect of Guangxi AIDS Conquering Project (GACP) were estimated and predicted based on the calculation of control reproduction number R_c.

Chapter 5

The Multi-group SIDT Model for Assessment of HIV Datasets from China CDC

In this chapter, we will construct a multi-group SIDT model to analyze HIV datasets. The multi-group SIDT model will consider HIV/AIDS process susceptible-infected-diagnosed-treated in general population and high risk groups. Our method discussed in Chapter 3 will be applied to the model to detect and resolve non-identifiability. After model parameters are uniquely estimated, HIV epidemics in general population and each high risk group will be estimated and predicted.

Compared to the single group SIDT model, the multi-group model will enable us to understand the true burden of HIV epidemic in each high-risk group, to design more effective HIV interventions, and to investigate HIV testing intensity in each high risk group.

The challenge for the analysis of multi-group model lies in the non-identifiability issue, since we have more parameters. We applied our Matrix Decomposition Method for resolving non-identifiability to the multi-group SIDT model, the method clearly informed us how many parameters we need to fix before the model fitting process and what options we have for the fixing procedure. Our Matrix Decomposition method was proved to be useful for analyzing more complicated models.

5.1 Project 1: The HIV dataset for the two remote townships in Sichuan province, China

As we discussed above, these two townships are among the regions in China where HIV transmission is very severe. In 2008, the Chinese government tested close to half of the total population in these two townships for HIV, 409 HIV cases were reported, and 48.17% of them were injection drug users (IDU). In 2010, the population-wide physical examination established baseline individual health records, including HIV status, for over

99% of the population in the two townships, 449 HIV cases were reported, and 29.62% of them were injection drug users (IDU). Injection drug users play an import role in HIV transmission in these two townships. We aim to investigate HIV transmission among injection drug users, HIV cross-transmission between IDU and general population in these two townships, and to identify targeting effective interventions, which will bring HIV transmission in these two townships fully under control. Based on the above motivation, we built the following two-group SIDT model. The diagram for the two-group SIDT model:

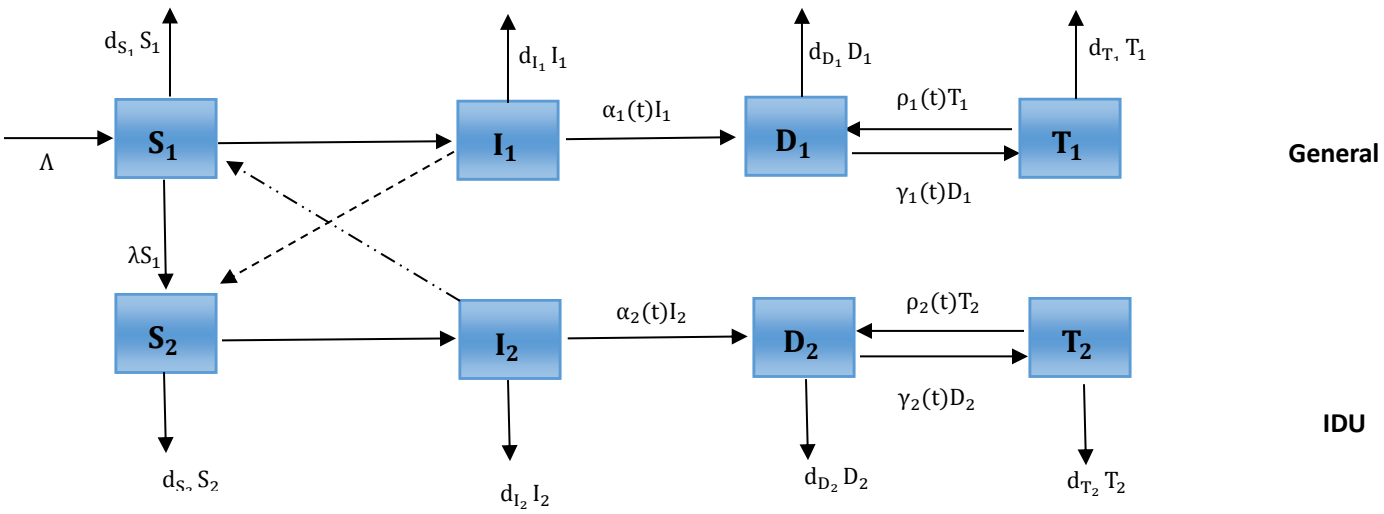


Figure 5.1: Transfer diagram for the two-group HIV transmission model

Based on the transfer diagram in Figure 5.1, the model is described by the following system of nonlinear differential equations

$$\dot{S}_1(t) = \Lambda - \frac{S_1(t)}{N_1(t)}\left(\beta_{11}I_1(t) + \beta_{1D1}D_1(t) + \beta_{1T1}T_1(t) + \beta_{12}I_2(t) + \beta_{1D2}D_2 + \beta_{1T2}T_2(t)\right) - d_{S_1}S_1(t)$$

$$\dot{I}_1(t) = \frac{S_1(t)}{N_1(t)}\left(\beta_{11}I_1(t) + \beta_{1D1}D_1(t) + \beta_{1T1}T_1(t) + \beta_{12}I_2(t) + \beta_{1D2}D_2 + \beta_{1T2}T_2(t)\right) - \alpha_1(t)I_1(t) - d_{I_1}I_1(t)$$

$$\dot{D}_1(t) = \alpha_1(t)I_1(t) + \rho_1T_1(t) - \gamma_1(t)D_1(t) - d_{D_1}D_1(t)$$

$$\dot{T}_1(t) = \gamma_1(t)D_1(t) - \rho_1T_1(t) - d_{T_1}T_1(t)$$

$$\dot{S}_2(t) = \lambda S_1 - \frac{S_2(t)}{N_2(t)}\left(\beta_{21}I_1(t) + \beta_{2D1}D_1(t) + \beta_{2T1}T_1(t) + \beta_{22}I_2(t) + \beta_{2D2}D_2 + \beta_{2T2}T_2(t)\right) - d_{S_2}S_2(t)$$

$$\dot{I}_2(t) = \frac{S_2(t)}{N_2(t)}\left(\beta_{21}I_1(t) + \beta_{2D1}D_1(t) + \beta_{2T1}T_1(t) + \beta_{22}I_2(t) + \beta_{2D2}D_2 + \beta_{2T2}T_2(t)\right) - \alpha_2(t)I_2(t) - d_{I_2}I_2(t)$$

$$\dot{D}_2(t) = \alpha_2(t)I_2(t) + \rho_2T_2(t) - \gamma_2(t)D_2(t) - d_{D_2}D_2(t)$$

$$\dot{T}_2(t) = \gamma_2(t)D_2(t) - \rho_2T_2(t) - d_{T_2}T_2(t).$$

We split the whole population into two groups, the general population and IDU, denoting them as group 1 and group 2 respectively. At a time t, the number of susceptible people in each group is denoted by $S_i(t), i = 1,2$, the number of HIV positive people who are not diagnosed in each group is denoted by $I_i(t), i = 1,2$, the number of diagnosed HIV positive people who are not under treatment in each group is denoted by $D_i(t), i = 1,2$, and the number of diagnosed HIV positive people that are under treatment in each group is denoted by $T_i(t), i = 1,2$. The sum $N_i(t) = S_i(t) + I_i(t) + D_i(t) + T_i(t), i = 1,2$ denotes the total population for each group. The time unit used in the model is per year to align with the available data.

Parameter Λ is the influx of susceptibles into the general group, λ is the transfer rate from general population to IDU. $d_{S_i}, d_{I_i}, d_{D_i}$ and $d_{T_i}, i = 1,2$, are death rates specific to compartments $S_i, I_i, D_i, T_i, i = 1,2$. Parameter $\rho_i, i = 1,2$ is the combined rate for treatment failure and loss to follow-up, $\beta_{ii}, \beta_{iDi}, \beta_{iTi}, i = 1,2$ are the within-group transmission coefficients for compartment $I_i, D_i, T_i, i = 1,2, \beta_{12}, \beta_{1D2}, \beta_{1T2}$ are the cross-group

transmission coefficients for compartment $I_2, D_2, T_2, \beta_{21}, \beta_{2D1}, \beta_{2T1}$ are the cross-group transmission coefficients for compartment I_1, D_1, T_1.

Terms $\alpha_i(t)I_i(t)$ and $\gamma_i(t)D_i(t), i = 1,2$ in the model are the annual number of new reports for HIV and AIDS combined and the number of new treatment enrollments in each group, respectively. With the implementation of the "Four Frees and One Care" program in 2003, the Chinese government has rapidly scaled up HIV testing and ART treatment. This was reflected by an increase in the number of people tested for HIV and in the number of treatment centers. To correctly adjust for the increase in new HIV tests and ART treatments, we used a time-dependent diagnosis rate $\alpha_i(t)$and treatment enrolment rate $\gamma_i(t), i = 1,2$.

Data source: Aggregated yearly HIV/AIDS surveillance data for general population and IDU from 2005 to 2010 was obtained from the Surveillance Database of the China Center for Disease Control and Prevention. Aggregated yearly data on ART treatment for the same period for the two groups was obtained from the Treatment Database of the China Center for Disease Control and Prevention. Yearly population data for the two townships was obtained from the demographic database for the Liangshan prefecture.

Parameter estimation and fitting: Same as the whole nation**,** Demographic parameters Λ and d_{S_1} are estimated by fitting the equation of the total population $N' = \Lambda - d_S N$ to the population data for the two townships from 2005 to 2010. Values of parameters d_{D_i}, d_{T_i}, ρ_i, $i = 1,2$ and the form of function $\gamma_i(t), i = 1,2$ are estimated directly from the surveillance and treatment data for these two groups. Since the HIV testing data is not available for these two townships, $\alpha_i(t), i = 1,2$ is assumed to be a linearly increasing function when there is no screening interventions, which is consistent with single group study. The two screening interventions in 2008 and 2010, which tested around half and 99% of the total population in these two townships, make us fix $\alpha_i(t), i = 1,2$ at value 0.5 in 2008 and 0.99 in 2010. Values of other parameters, including within-group transmission coefficients and cross-group transmission coefficients, the transfer rate λ,

death rate $d_{I_1}, d_{I_2}, d_{S_2}$ and the initial population size I_{01}, I_{02} for the undiagnosed compartment I_1 and I_2 at the beginning of our fitting (end of 2005), cannot be estimated directly from surveillance data and need to be obtained through model fitting.

We applied our method to the two-group SIDT model to resolve non-identifiability issue [37]. We prefixed the relationships among $\beta_{ii}, \beta_{iDi}, \beta_{iTi}, i = 1,2$, as we did for the single group SIDT model $\beta_{iDi} = 0.75\beta_{ii}, \beta_{iTi} = 0.1\beta_{ii}, i = 1,2$, and perform non-identifiability analysis for the remaining parameters [69-76]. The singular value decomposition informed us that among model parameters β_{12}, β_{21}, β_{11}, β_{22}, I_{01}, I_{02}, S_{02}, $d_{I1}, d_{I2}, d_{S2}, \lambda$, and the coefficients for $\alpha_i(t), i = 1,2$ there are five dependencies, and the variance decomposition informed us that all the above parameters were involved in the five dependencies. Therefore we need to fix five parameters' values before model fitting so that the remaining ones can be uniquely estimated. We fix $d_{I1} = 0.063$, the same as d_I we fixed for the single group SIDT model. By referring to the survival data from a cohort study for IDU [109], we fixed the value for d_{I2}, d_{S2}. The estimation for the number I(2005) from the single group SIDT model study inform us of the value of $I_{01} + I_{02}$. We also reasonably assumed β_{21}, the cross-group transmission coefficient from general population to IDU, to be 0.

The Nonlinear Least Squares method [28] was applied to find the point estimates for model parameters β_{12}, β_{11}, $\beta_{22}, I_{02}, S_{02}, \lambda$, and the coefficients for $\alpha_i(t), i = 1,2$, which minimize the sum of squared error between model output and the available surveillance and treatment data. More specifically, the following data are used in our modeling fitting:

- Annual number of new diagnosis of HIV and AIDS combined from 2005 to 2010;
- Annual number of death due to HIV and AIDS among diagnosed from 2005 to 2010;
- Annual treatment enrollment from 2005 to 2010;
- Annual number of death among diagnosed people and people in treatment from 2005 to 2010.

- Annual number of treatment failure and loss to follow-up from 2005 to 2010
- The number of total population from 2005 to 2010.

Results

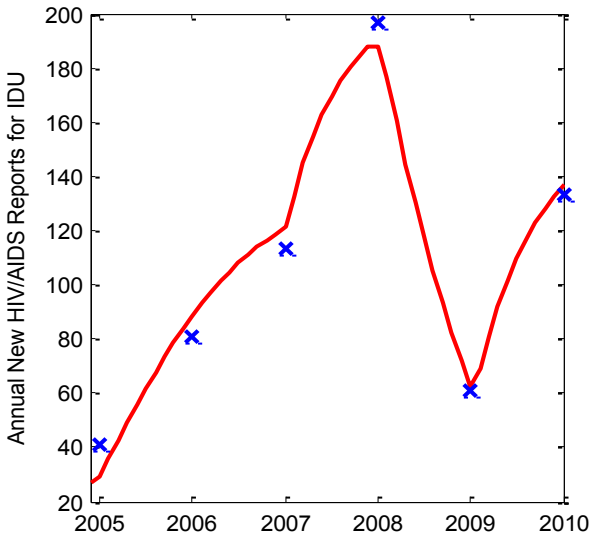
The point estimates for model parameters are shown in table 5.1.

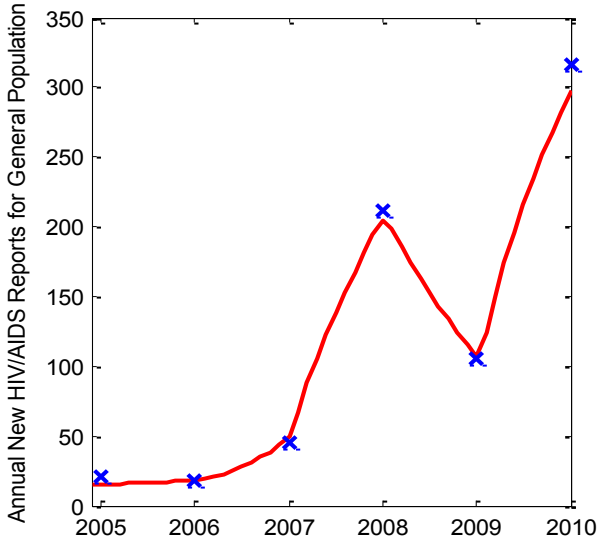
Parameters	Description	Best-fit value	Source
β_{11}	Within-group transmission coefficients for I_1	0.2446	Fitting
β_{1D1}	Within-group transmission coefficients for D_1	0.1835	[69,70]
β_{1T1}	Within-group transmission coefficients for T_1	0.02446	[71,76]
β_{12}	Cross-group transmission coefficients for I_2	0.0266	Fitting
β_{1D2}	Cross-group transmission coefficients for D_2	0.02	[69,70]
β_{1T2}	Cross-group transmission coefficients for T_2	0.00266	[71,76]
β_{22}	Within-group transmission coefficients for I_2	0.0681	Fitting
β_{2D2}	Within-group transmission coefficients for D_2	0.051	[69,70]
β_{2T2}	Within-group transmission coefficients for T_2	0.0068	[71,76]
β_{21}	Cross-group transmission coefficients for I_1	0	Assumption
β_{2D1}	Cross-group transmission coefficients for D_1	0	Assumption
β_{2T1}	Cross-group transmission coefficients for T_1	0	Assumption
d_{S1}	Death rate for S_1	0.007*	Fitting
d_{I1}	Death rate for I_1	0.063*	[67, 68]
d_{D1}	Death rate for D_1	0.0411*	Data
d_{T1}	Death rate for T_1	0.04*	Data
d_{S2}	Death rate for S_2	0.0237*	[109]
d_{I2}	Death rate for I_2	0.117*	[109]
d_{D2}	Death rate for D_2	0.0554*	Data
d_{T2}	Death rate for T_2	0.0427*	Data
ρ_1	Treatment drop-out rate for general Population	0*	Data
ρ_2	Treatment drop-out rate for IDU	0*	Data
I_{01}	Undiagnosed HIV positive population in 2005 for general population	334	Fitting
I_{02}	Undiagnosed HIV positive population in 2005 for IDU	785	Fitting

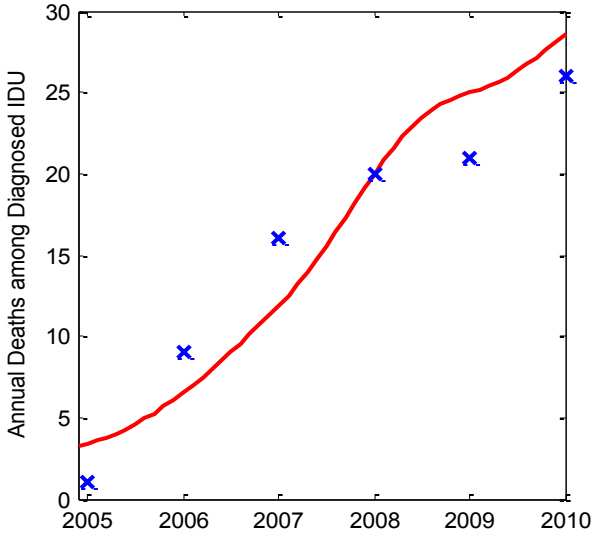
$\alpha_1(t)$	Time-dependent diagnosis rate of undiagnosed for general population	$0.05t - 100.225^*, 2005 \leq t \leq 2007$ $0.5,\ \ 2007 \leq t \leq 2008$ $0.275, 2008 < t \leq 2009$ $1, 2009 < t \leq 2010$	Fitting
$\alpha_2(t)$	Time-dependent diagnosis rate of undiagnosed for IDU	$0.0895t - 178.869^*, 2005 \leq t \leq 2007$ $0.5,\ \ 2007 \leq t \leq 2008$ $0.24, 2008 < t \leq 2009$ $1, 2009 < t \leq 2010$	Fitting
$\gamma_1(t)$	Time-dependent treatment enrollment rate for general population	$0.0039524363t^2 - 15.8637.* t + 15917.8679^*$, 2005-2010	Data
$\gamma_2(t)$	Time-dependent treatment enrollment rate for IDU	$0.007888105t^2 - 31.6311996.* t + 31710.1918^*$, 2005-2010	Data
Λ	The influx of susceptible into the general group	295	Fitting
λ	Transfer proportion from general population to IDU	0.008^*	Fitting
R^2	Goodness of fit ratio		Fitting

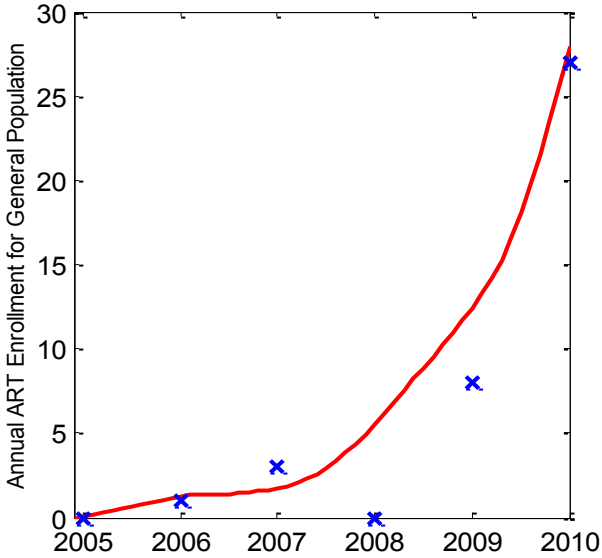
Table 5.1. Model parameters and their best-fit values

Notes: * The unites for these parameters are 1/person/year


Annual New HIV/AIDS Reports for IDU
200
180
160
140
120
100
80
60
40
20
2005
2006
2007
2008
2009
2010


Annual New HIV/AIDS Reports for General Population
350
300
250
200
150
100
50
0
2005
2006
2007
2008
2009
2010


Annual Deaths among Diagnosed IDU
30
25
20
15
10
5
0
2005
2006
2007
2008
2009
2010


Annual ART Enrollment for General Population
30
25
20
15
10
5
0
2005
2006
2007
2008
2009
2010

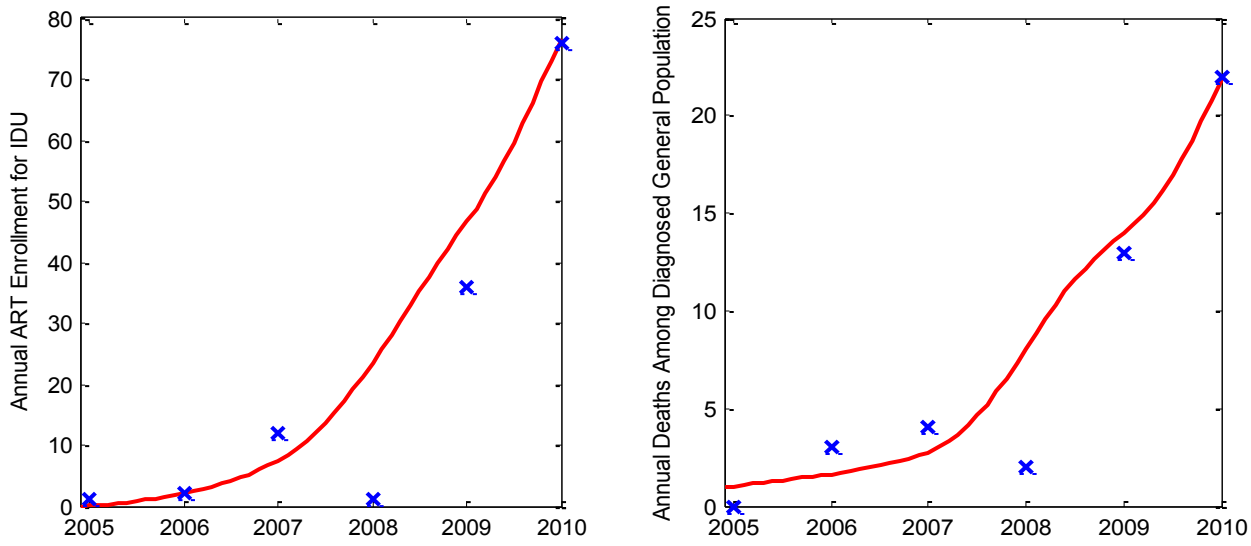


Figure 5.2 Model Fitting Results

The high goodness of fit ratio provided a validation for our two-group SIDT model.

Three key assessment indicators for the HIV/AIDS dynamics: annual number of new HIV infections, annual number of HIV/AIDS related deaths, and the total number of people living with HIV (both diagnosed and undiagnosed), were estimated from 2005 to 2010 for each group using the two-group SIDT model.

The estimated annual number of new HIV infections of general population rose from 98 in 2005 to 162 in 2010. In the meantime, the year-over-year rate of increase declined from 15.31% in 2006 to 5.88% in 2010, indicating that the momentum of rise of the HIV epidemics in general population has slowed down significantly during the 6-year period. The estimated annual number of new HIV infections of IDU declined from 35 in 2005 to 21 in 2010. The estimated total number of people living with HIV and AIDS of general population, including both diagnosed and undiagnosed people, increased from 366 in 2005 to 896 in 2010. The year-over-year rate of increase declined from 24.04% in 2006 to 15.46% in 2010. The estimated total number of people living with HIV and AIDS of IDU, including both diagnosed and undiagnosed people, declined from 866 in 2005 to

670 in 2010. The estimated HIV prevalence rate among the whole population at the end of 2010 was 14.57%.

The estimated total number of deaths among people living with HIV and AIDS of general population increased from 20 in 2005 to 41 in 2010. The year-over-year rate of increase declined from 25.00% in 2006 to 7.89% in 2010. The estimated total number of deaths among people living with HIV and AIDS of IDU declined from 101 in 2005 to 45 in 2010.

Estimations for the three key indicators informed us that HIV transmission in these two township has transferred from high risk group IDU to general population. The HIV interventions implemented from 2005 to 2010, has brought all three indicators down in IDU group and has slowed down the HIV transmission in general population.

Our model was used to further predict the three key indicators for the period 2011-2015, under the assumption that the treatment enrollment rate would continue to increase linearly following its pre-2010 trend, and that the diagnosis rate would maintain the same trend as the period from 2005 to 2007 before the population-wide screening.

Our model predicted that continuous scale-up of HIV testing and ART treatment will reverse the temporal trend of the HIV epidemics in general population, the predicted annual number of new HIV infections would decline starting from 2014. The predicted annual number of new HIV infections in IDU will decline to 5 in 2015. The predicted value for PLHIV of general population, IDU in 2015 is 1512, 567 respectively, with a HIV prevalence rate at 16.11% and 29.06%. The model prediction for the number of combined HIV and AIDS deaths of general population, IDU in 2015 is 65 and 26 respectively (table 5.2).

Year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
HIV new infections	98	114	130	142	155	162	172	182	186	185	178
PLHIV	366	453	552	659	774	896	1019	1150	1276	1400	1512
HIV/AIDS deaths	20	25	31	34	38	41	46	52	57	62	65

Table 5.2(a) Estimations of the three indicators for the General Population

Year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
HIV new infections	35	34	31	28	25	21	18	14	11	8	5
PLHIV	866	809	761	724	695	670	650	628	607	586	567
HIV/AIDS deaths	101	91	79	65	55	45	39	35	32	29	26

Table 5.2(b) Estimations of the three indicators for IDU

Compared to single group SIDT model, besides the estimations for three indicators, the two-group SIDT model enables us to estimate the total population for high risk group IDU (table 5.3). The estimation showed that IDU accounted for around 20% of the total population in these two townships, which is one of the reasons why HIV transmission is severe in this region.

Year	2005	2006	2007	2008	2009	2010
Population(IDU)	2118	2060	2014	1983	1962	1951
Total Population	10347	10523	10648	10806	10836	10902

Table 5.3 The Estimation for IDU Population

We used our model to estimate that the two population-wide screening programs in 2008 and 2010 have saved 114 life years, and averted 101 new HIV infections from the beginning of 2008 to the end of 2015 for general population, and 76 life years, 6 new HIV infections for IDU. The number of averted new HIV infections account for 6.97%,

4.55% of the new HIV infections that would have occurred without the two screening interventions for general population and IDU respectively. Figure 5.3 shows the cumulative number and percentage of HIV new infections averted from 2008 to 2015 for general population and IDU (Figure 5.3).

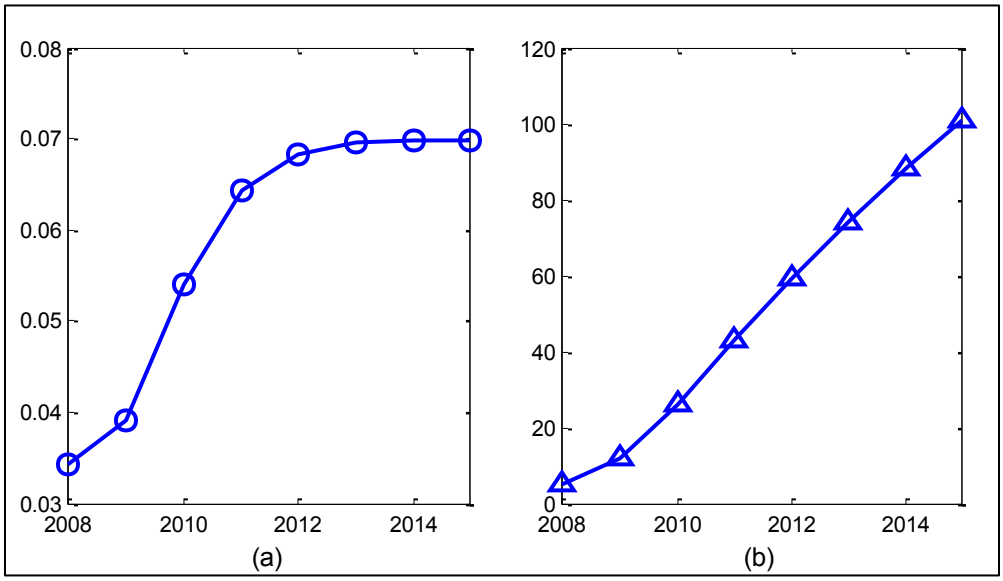


Figure 5.3 (a): Cumulative number and percentage of HIV new infections averted from 2008 to 2015 for the general population due to the screening interventions

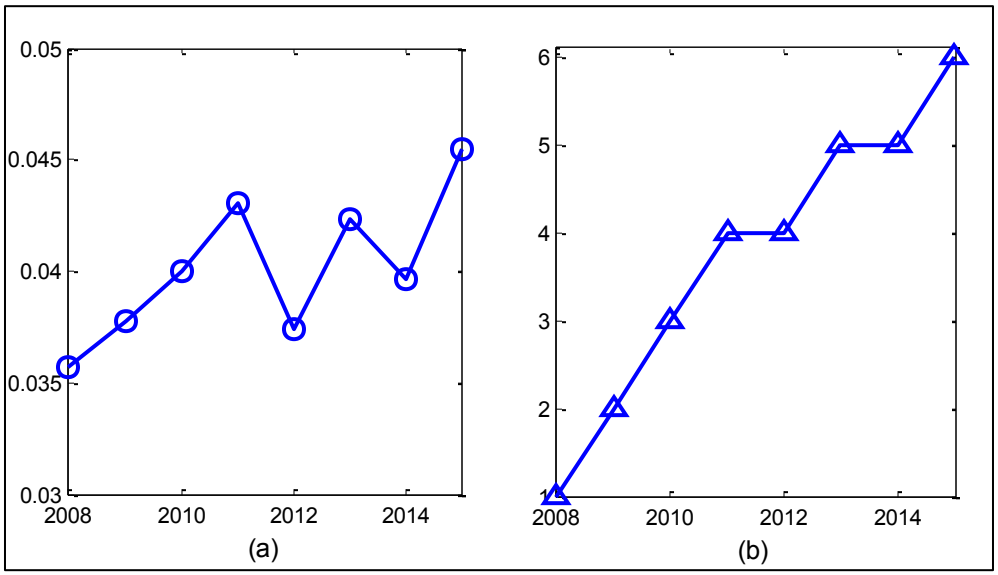


Figure 5.3 (b): Cumulative number and percentage of HIV new infections averted from 2008 to 2015 for IDU due to the screening interventions

Discussion

At first, we assume $\alpha_i(t), i = 1,2$ are identical functions, linearly increasing from 2005 to 2007, fixed at 0.5, 1 at 2008 and 2010 respectively. Under this assumption, the model fitting is as following:

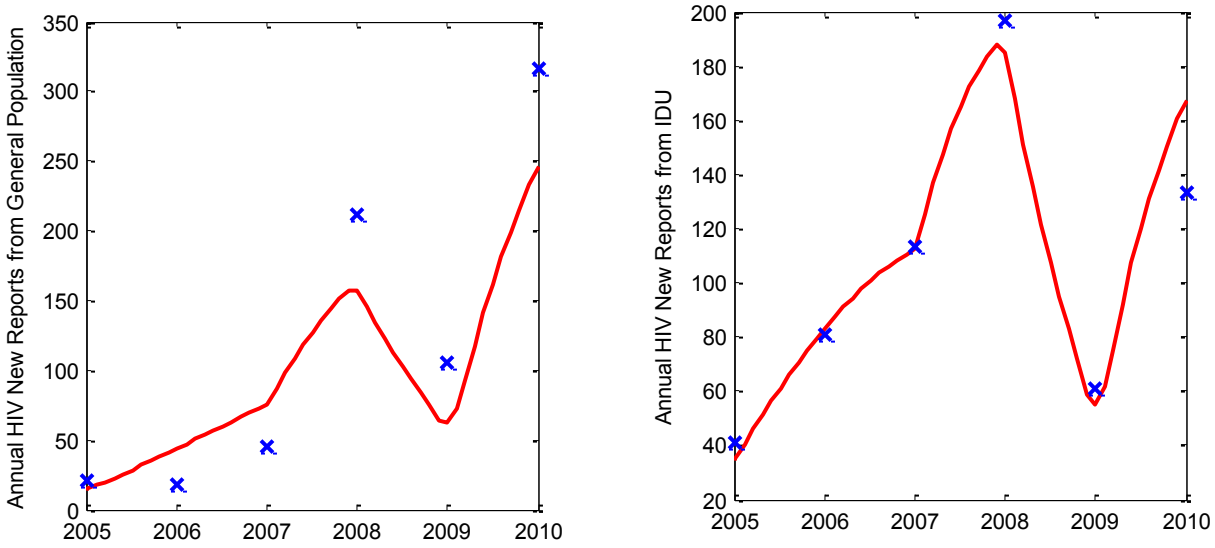


Figure 5.4 Model fitting under the assumption that $\alpha_i(t), i = 1,2$ are identical

Since $\alpha_i(t)I_i(t), i = 1,2$ account for the annual HIV new diagnosis in general population and IDU, if we assume $\alpha_i(t), i = 1,2$ are the same, all the difference of increase in HIV new diagnosis for these two groups is distributed to the difference of increase of $I_i(t), i = 1,2$. Therefore, the faster increase of annual HIV new diagnosis for IDU indicates a faster rise for $I_2(t)$ than $I_1(t)$, resulting in overestimated transmission coefficients, and further an overestimated number of annual HIV new reports in 2010. Flat increase for annual HIV new diagnosis in general population leads to underestimated transmission coefficients, and further an underestimated annual HIV new diagnosis in 2010. The fitting graphs inform us that it is more reasonable to assume that both $\alpha_1(t)$ and $\alpha_2(t)$ are linear increasing functions, but with different coefficients.

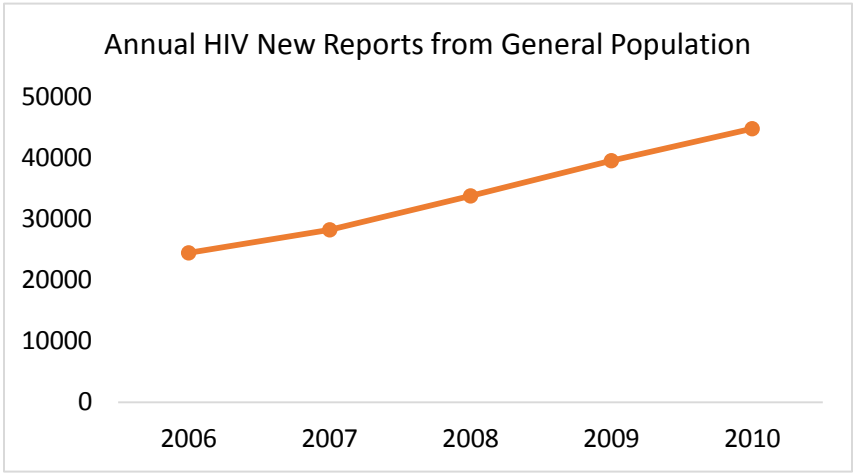
The estimation and prediction for these two townships from two-group SIDT model were consistent with that from single group SIDT model except the estimates of HIV/AIDS related deaths. The estimated HIV/AIDS related deaths from single group model were 53 at 2005, 78 at 2010, while those from two-group model were 121 at 2005, and 78 at 2010. In the analysis for single group model, we fix the value for parameter d_I at 0.063, which was obtained from the survival data for HIV patients not receiving treatment, combining all HIV high risk groups and the general population. Therefore, we will underestimate the death rate d_I in a region where IDU dominates the HIV transmission, since IDU HIV patients have the shortest survival length among all HIV high risk groups and the general population. The analysis for two-group SIDT model informed us that at 2005 IDU dominated in HIV transmission in these two townships, and at 2010 general population accounted for most HIV new infections. Therefore, in the two townships, HIV transmission has transferred from high risk group IDU to the general population. The dominance of IDU in HIV transmission at 2005 explains why estimated HIV/AIDS deaths from the two-group model much higher than that from the single group model.

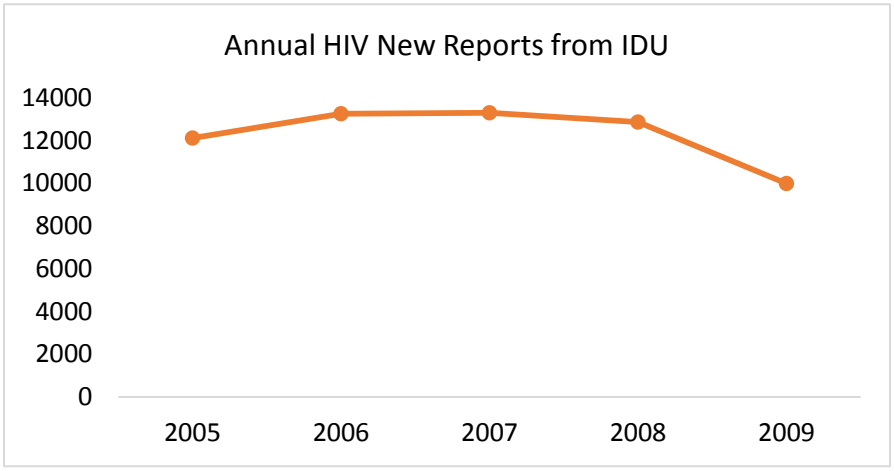
Due to the existence of two screening interventions, there is no non-identifiability issue among the within-group and cross-group transmission coefficients. While in other situations when there is no special HIV interventions, it might be possible for some of the transmission coefficients being non-identifiable.

The estimation and prediction inform us that HIV transmission in IDU for these two townships was under control since all three indicators were declining. Even if we maintain the level of HIV testing and treatment starting from 2009 for IDU, HIV new infections still declined from 35 in 2005 to 21 in 2010 and further to 12 in 2015. However, if the level of HIV testing and treatment for general population was maintained from 2009, HIV new infections would be rising from 2005. Therefore more efforts could be devoted to general population so that HIV transmission in these two townships can be fully under control.

5.2 Project 2: The HIV dataset for China

The surveillance data for high risk groups IDU, men who have sex with men (MSM) and general population showed us various stories about HIV transmission in each group (figure 5.5). The number of annual new HIV reports from MSM has increased by around 30 times from 2005 to 2010, while for IDU that number has been decreasing in the same period. To further understand HIV epidemic transmission situation in each high risk group and the general population, to identify efficient HIV interventions for each group, and to estimate HIV testing intensity in each group, we built the three-group SIDT model.





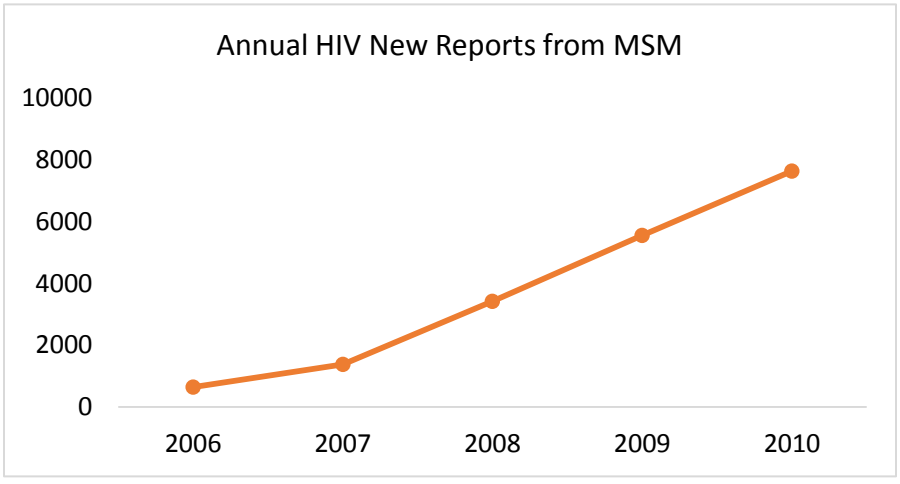


Figure 5.5 Annual number of new HIV reports for the general population, IDU and MSM

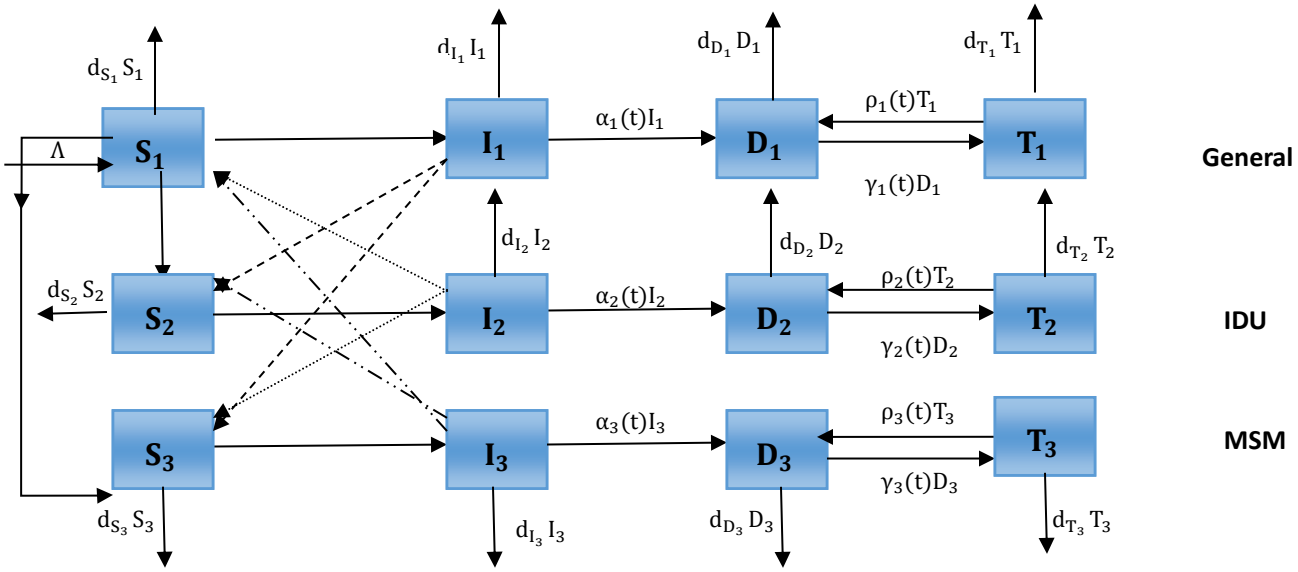


Figure 5.6: Transfer diagram for the three-group HIV transmission model

Based on the transfer diagram in Figure 5.6, the model is described by the following system of nonlinear differential equations

$$\dot{S}_1(t) = \Lambda - \frac{S_1(t)}{N_1(t)}\left(\beta_{11}I_1(t) + \beta_{1D1}D_1(t) + \beta_{1T1}T_1(t) + \beta_{12}I_2(t) + \beta_{1D2}D_2 + \beta_{1T2}T_2(t) + \beta_{13}I_3(t) + \beta_{1D3}D_3 + \beta_{1T3}T_3(t)\right) - d_{S_1}S_1(t)$$

$$\dot{I}_1(t) = \frac{S_1(t)}{N_1(t)}\left(\beta_{11}I_1(t) + \beta_{1D1}D_1(t) + \beta_{1T1}T_1(t) + \beta_{12}I_2(t) + \beta_{1D2}D_2 + \beta_{1T2}T_2(t) + \beta_{13}I_3(t) + \beta_{1D3}D_3 + \beta_{1T3}T_3(t)\right) - \alpha_1(t)I_1(t) - d_{I_1}I_1(t)$$

$$\dot{D}_1(t) = \alpha_1(t)I_1(t) + \rho_1 T_1(t) - \gamma_1(t)D_1(t) - d_{D_1}D_1(t)$$

$$\dot{T}_1(t) = \gamma_1(t)D_1(t) - \rho_1 T_1(t) - d_{T_1}T_1(t)$$

$$\dot{S}_2(t) = \lambda_1 S_1 - \frac{S_2(t)}{N_2(t)}\left(\beta_{21}I_1(t) + \beta_{2D1}D_1(t) + \beta_{2T1}T_1(t) + \beta_{22}I_2(t) + \beta_{2D2}D_2 + \beta_{2T2}T_2(t) + \beta_{23}I_3(t) + \beta_{2D3}D_3(t) + \beta_{2T3}T_3(t)\right) - d_{S_2}S_2(t)$$

$$\dot{I}_2(t) = \frac{S_2(t)}{N_2(t)}\left(\beta_{21}I_1(t) + \beta_{2D1}D_1(t) + \beta_{2T1}T_1(t) + \beta_{22}I_2(t) + \beta_{2D2}D_2 + \beta_{2T2}T_2(t) + \beta_{23}I_3(t) + \beta_{2D3}D_3(t) + \beta_{2T3}T_3(t)\right) - \alpha_2(t)I_2(t) - d_{I_2}I_2(t)$$

$$\dot{D}_2(t) = \alpha_2(t)I_2(t) + \rho_2 T_2(t) - \gamma_2(t)D_2(t) - d_{D_2}D_2(t)$$

$$\dot{T}_2(t) = \gamma_2(t)D_2(t) - \rho_2 T_2(t) - d_{T_2} T_2(t)$$

$$\dot{S}_3(t) = \lambda_2 S_1 - \frac{S_3(t)}{N_3(t)}\big(\beta_{31} I_1(t) + \beta_{3D1} D_1(t) + \beta_{3T1} T_1(t) + \beta_{32} I_2(t) + \beta_{3D2} D_2 + \beta_{3T2} T_2(t) + \beta_{33} I_3(t) + \beta_{3D3} D_3(t) + \beta_{3T3} T_3(t)\big) - d_{S_3} S_3(t)$$

$$\dot{I}_3(t) = \frac{S_3(t)}{N_3(t)}\big(\beta_{31} I_1(t) + \beta_{3D1} D_1(t) + \beta_{3T1} T_1(t) + \beta_{32} I_2(t) + \beta_{3D2} D_2 + \beta_{3T2} T_2(t) + \beta_{33} I_3(t) + \beta_{3D3} D_3(t) + \beta_{3T3} T_3(t)\big) - \alpha_3(t) I_3(t) - d_{I_3} I_3(t)$$

$$\dot{D}_3(t) = \alpha_3(t) I_3(t) + \rho_3 T_3(t) - \gamma_3(t) D_3(t) - d_{D_3} D_3(t)$$

$$\dot{T}_3(t) = \gamma_3(t) D_3(t) - \rho_3 T_3(t) - d_{T_3} T_3(t).$$

We split the whole population into three groups, general population, IDU and MSM, denoting them as group 1, group 2 and group 3 respectively. At a time t, the number of susceptible people in each group is denoted by $S_i(t), i = 1,2,3$, the number of HIV positive people who are not diagnosed in each group is denoted by $I_i(t), i = 1,2,3$, the number of diagnosed HIV positive people who are not under treatment in each group is denoted by $D_i(t), i = 1,2,3$, and the number of diagnosed HIV positive people that are under treatment in each group is denoted by $T_i(t), i = 1,2,3$. The sum $N_i(t) = S_i(t) + I_i(t) + D_i(t) + T_i(t)$ denotes the total population for each group. The time unit used in the model is per year to align with the available data.

Parameter Λ is the influx of susceptibles into the general group, λ_1 is the transfer rate from general population to IDU, and λ_2 is the transfer rate from general population to MSM, $d_{S_I}, d_{I_I}, d_{D_I}$ and $d_{T_I}, i = 1,2,3$, are death rates specific to compartments $S_I, I_I, D_I, T_I, i = 1,2,3$. Parameter $\rho_I, i = 1,2$ is the combined rate for treatment failure and loss to follow-up, $\beta_{ii}, \beta_{iDi}, \beta_{iTi}, i = 1,2$ are the within-group transmission coefficients for compartment $I_I, D_I, T_I, i = 1,2,3, \beta_{12}, \beta_{1D2}, \beta_{1T2}, \beta_{13}, \beta_{1D3}, \beta_{1T3}$ are the cross-group transmission coefficients from compartment $I_2, D_2, T_2, I_3, D_3, T_3$, to compartment $I_1, \beta_{21}, \beta_{2D1}, \beta_{2T1}, \beta_{23}, \beta_{2D3}, \beta_{2T3}$ are the cross-group transmission coefficients from compartment

$I_1, D_1, T_1, I_3, D_3, T_3$, to compartment $I_2, \beta_{31}, \beta_{3D1}, \beta_{3T1}, \beta_{32}, \beta_{3D2}, \beta_{3T2}$ are the cross-group transmission coefficients from compartment $I_1, D_1, T_1, I_2, D_2, T_2$, to compartment I_3.

Terms $\alpha_i(t)I_i(t)$ and $\gamma_i(t)D_i(t), i = 1,2,3$ in the model are the annual number of new reports for HIV and AIDS combined and the number of new treatment enrollments in each group, respectively. With the implementation of the "Four Frees and One Care" program in 2003, the Chinese government has rapidly scaled up HIV testing and ART treatment. This was reflected by an increase in the number of people tested for HIV and in the number of treatment centers. To correctly adjust for the increase in new HIV tests and ART treatments, we used a time-dependent diagnosis rate $\alpha_i(t)$and treatment enrolment rate $\gamma_i(t), i = 1,2,3$.

Data source: Aggregated yearly HIV/AIDS surveillance data for general population, IDU and MSM from 2006 to 2010 was obtained from the Surveillance Database of the China Center for Disease Control and Prevention. Aggregated yearly data on ART treatment for the same period for the three groups was obtained from the Treatment Database of the China Center for Disease Control and Prevention. Yearly population data in the same period on national population was from the National Census. Predictions for annual HIV new infections, PLHIV and annual HIV/AIDS related deaths of the whole nation from 2011 to 2015 were obtained from single group study.

Parameter estimation and non-identifiability analysis: Demographic parameters Λ and d_{S_1} are estimated by fitting the equation of the total population $N' = \Lambda - d_S N$ to the population data for China from 2006 to 2010. Values of parameters $d_{D_i}, d_{T_i}, \rho_i, i = 1,2,3$ and the form of function $\gamma_i(t), i = 1,2,3$ are estimated directly from the surveillance and treatment data for these three groups. Since HIV testing data is not available for each group, and HIV test rate for the whole nation was linearly increasing from 2006 to 2010, we assumed $\alpha_i(t), i = 1,2,3$ being linearly increasing functions. Values of other parameters, including within-group transmission coefficients and cross-group

transmission coefficients, the transfer rate λ_1 and λ_2, death rate $d_{I_1}, d_{I_2}, d_{I_3}, d_{S_2}, d_{S_3}$ and the initial population size $I_{01}, I_{02}, I_{03}, S_{02}, S_{03}$ for the undiagnosed compartment I_1, I_2, I_3 and S_2 S_3 at the beginning of our fitting (end of 2006), cannot be estimated directly from surveillance data and need to be obtained through model fitting.

We applied our Matrix Decomposition method to the three-group SIDT model to resolve non-identifiability issue [37]. We prefixed the relationships among β_{ii}, β_{iDi}, $\beta_{iTi}, i = 1,2$, as we did for the single group SIDT model $\beta_{iDi} = 0.75\beta_{ii}, \beta_{iTi} = 0.1\beta_{ii}, i =$ 1,2, and perform non-identifiability analysis for the remaining parameters [69-76]. The singular value decomposition informed us that among model parameters to be estimated, there are 11 dependencies, and the variance decomposition informed us that all the parameters were involved in the 11 dependencies. We fixed $d_{I_1} = 0.063, d_{I_3} = 0.063, d_{S_3} =$ 0.007 under the assumption that MSM have same death rate as general population, fixed $d_{I_2} = 0.117, d_{S_2} = 0.0237$ as we did in Liangshan two-group SIDT model. For the remaining parameters, we have to fix at least 6 parameters' value so that all other parameters can be uniquely determined.

In each group, there are three transmission coefficients to be estimated, one within-group transmission coefficient and two cross-group transmission coefficient. Non-identifiability analysis revealed that with the available data:

- Annual number of death due to HIV and AIDS among diagnosed from 2006 to 2010;
- Annual number of new diagnosis of HIV and AIDS combined from 2006 to 2010;
- Annual treatment enrollment from 2006 to 2010;
- Annual number of death among diagnosed people and people in treatment from 2006 to 2010.
- Annual number of treatment failure and loss to follow-up from 2006 to 2010
- The number of total population from 2006 to 2010

- The number of annual HIV new infections from 2011 to 2015 obtained from single group study
- The number of PLHIV from 2011 to 2015 obtained from single group study
- The number of HIV/AIDS Deaths from 2011 to 2015 obtained from single group study,

only seven transmission coefficients can be estimated together. We fixed the two cross-group transmission coefficients between IDU and MSM as 0. Among parameters S_{02}, S_{03} λ_1 and λ_2, which determine the size of high risk group IDU and MSM, none of them can be estimated together with the above seven transmission coefficients. Therefore, if we want to estimate transmission coefficients, we have to fix parameters determining the size of high risk groups IDU and MSM. We fixed parameter $S_{03}, \lambda_2, S_{02}, \lambda_1$ by referring to the information that usually MSM accounts for 2%-4% of male population in a region [110], and IDU population in China is around 74,000 during the period from 2006 to 2010 [111]. At last parameters β_{11}, β_{12}, β_{13}, β_{21}, β_{22}, β_{31}, β_{33}, $I_1(2006)$, $I_2(2006)$, $I_3(2006)$ and the coefficients for $\alpha_i(t), i = 1,2,3$ were estimated together by fitting the multi-group SIDT model to data.

Results

The point estimates for model parameters are shown in table 5.4.

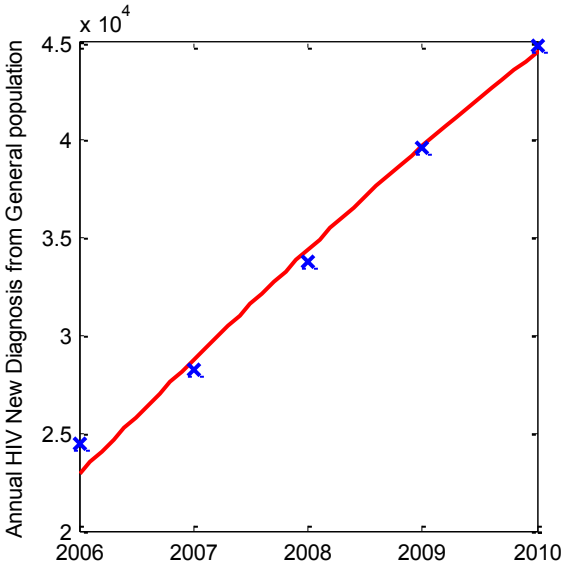
Parameters	Description	Best-fit value	Source
β_{11}	Within-group transmission coefficients for I_1	0.1342	Fitting
β_{1D1}	Within-group transmission coefficients for D_1	0.1007	[69,70]
β_{1T1}	Within-group transmission coefficients for T_1	0.01342	[71-76]
β_{12}	Cross-group transmission coefficients from I_2 to I_1	0.0217	Fitting
β_{1D2}	Cross-group transmission coefficients from D_2 to I_1	0.0163	[69,70]
β_{1T2}	Cross-group transmission coefficients from T_2 to I_1	0.00217	[71-76]
β_{13}	Cross-group transmission coefficients from I_3 to I_1	0.000336	Fitting
β_{1D3}	Cross-group transmission coefficients from D_3 to I_1	0.000252	[69,70]
β_{1T3}	Cross-group transmission coefficients from T_3 to I_1	0.0000336	[71-76]
β_{21}	Cross-group transmission coefficients from I_1 to I_2	0.0000455	Fitting

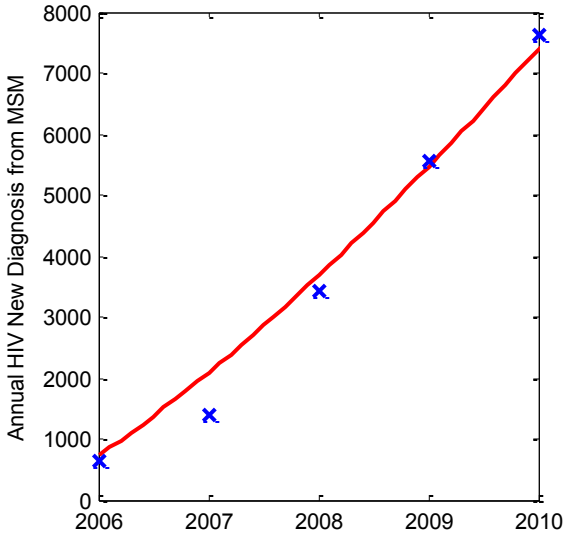
β_{2D1}	Cross-group transmission coefficients for D_1 to I_2	0.0000341	[69,70]
β_{2T1}	Cross-group transmission coefficients for T_1 to I_2	0.00000455	[71-76]
β_{22}	Within-group transmission coefficients for I_2	0.073	Fitting
β_{2D2}	Within-group transmission coefficients for D_2	0.055	[69,70]
β_{2T2}	Within-group transmission coefficients for T_2	0.0073	[71-76]
β_{23}	Cross-group transmission coefficients from I_3 to I_2	0	Assumption
β_{2D3}	Cross-group transmission coefficients for D_3 to I_2	0	Assumption
β_{2T3}	Cross-group transmission coefficients for T_3 to I_2	0	Assumption
β_{31}	Cross-group transmission coefficients from I_1 to I_3	0.000041	Fitting
β_{3D1}	Cross-group transmission coefficients for D_1 to I_3	0.000031	[69,70]
β_{3T1}	Cross-group transmission coefficients for T_1 to I_3	0.0000041	[71-76]
β_{32}	Cross-group transmission coefficients from I_2 to I_3	0	Assumption
β_{3D2}	Cross-group transmission coefficients for D_2 to I_3	0	Assumption
β_{3T2}	Cross-group transmission coefficients for T_2 to I_3	0	Assumption
β_{33}	Within-group transmission coefficients for I_3	0.227	Fitting
β_{3D3}	Within-group transmission coefficients for D_3	0.17	[69,70]
β_{3T3}	Within-group transmission coefficients for T_3	0.0227	[71-76]
d_{S1}	Death rate for S_1	0.007*	Fitting
d_{I1}	Death rate for I_1	0.063*	[67,68]
d_{D1}	Death rate for D_1	0.05*	Data
d_{T1}	Death rate for T_1	0.01285714287.*t.*t- 51.66682862.*t+51906.2185* $t \leq 2010$ 0.036 $t > 2010$	Data
d_{S2}	Death rate for S_2	0.0237*	[109]
d_{I2}	Death rate for I_2	0.117*	[109]
d_{D2}	Death rate for D_2	0.05*	Data
d_{T2}	Death rate for T_2	-0.061t+122.69*, $t \leq 2007$ -0.07.*T+140.7, $2007 \leq t \leq 2010$ 0.035 t>2010	Data
d_{S3}	Death rate for S_3	0.007*	Fitting
d_{I3}	Death rate for I_3	0.063*	[67, 68]
d_{D3}	Death rate for D_3	0.015*	Data

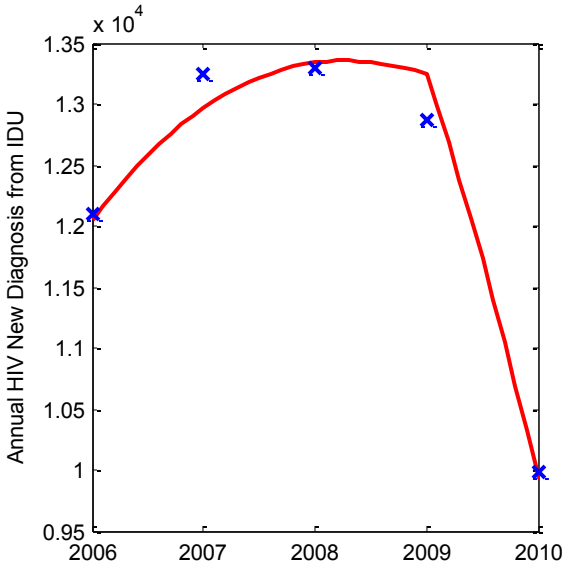
d_{T3}	Death rate for T_3	0.0485*, $t \leq 2007$ 0.051, $2007 \leq t \leq 2009$ 0.025 t>2009	Data
ρ_1	Treatment drop-out rate for general Population	0.03227*	Data
ρ_2	Treatment drop-out rate for IDU	-0.0472844t+95.087*	Data
ρ_3	Treatment drop-out rate for MSM	0.0274*	Data
I_{01}	Undiagnosed HIV positive population in 2006 for general	301710	Fitting
I_{02}	Undiagnosed HIV positive population in 2006 for IDU	147208	Fitting
I_{03}	Undiagnosed HIV positive population in 2006 for MSM	43628	Fitting
S_{02}	Susceptible population in 2006 for IDU	740000	[111]
S_{03}	Susceptible population in 2006 for MSM	19722237	[110]
$\alpha_1(t)$	Time-dependent diagnosis rate of undiagnosed for general population	$0.0172898644t-34.5978$*	Fitting
$\alpha_2(t)$	Time-dependent diagnosis rate of undiagnosed for IDU	$0.01623t - 32.4415$*, $t \leq 2009$ 0.11, $2009 \leq t \leq 2010$ $0.01623t - 32.4415$, $t > 2010$	Fitting
$\alpha_3(t)$	Time-dependent diagnosis rate of undiagnosed for MSM	$0.02596\,t - 52.0479$*	Fitting
$\gamma_1(t)$	Time-dependent treatment enrollment rate for general population	$0.0117367t-23.43$*	Data
$\gamma_2(t)$	Time-dependent treatment enrollment rate for IDU	0.011t -22.03*, $t \leq 2007$, 0.057, $t > 2007$	Data
$\gamma_3(t)$	Time-dependent treatment enrollment rate for MSM	0.12*, $t \leq 2007$, 0.04t -80.18, $t > 2007$	Data
Λ	The influx of susceptible into the general group	15702110	Fitting
λ_1	Transfer from general population to IDU	17539	[111]
λ_2	Transfer from general population to MSM	235332	[110]
R^2	Goodness of fit ratio	0.96	Fitting

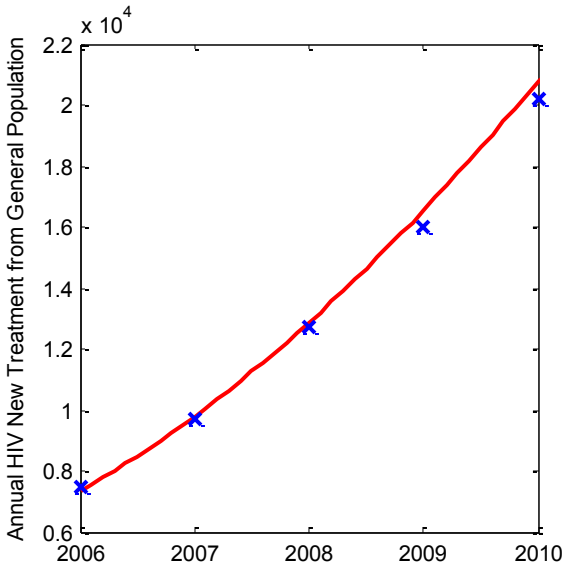
Table 5.4. Model parameters and their best-fit values

Notes: * The unites for these parameters are 1/person/year


x 10^4
Annual HIV New Diagnosis from General population
4.5
4
3.5
3
2.5
2
2006
2007
2008
2009
2010


Annual HIV New Diagnosis from MSM
8000
7000
6000
5000
4000
3000
2000
1000
0
2006
2007
2008
2009
2010


x 10^4
Annual HIV New Diagnosis from IDU
1.35
1.3
1.25
1.2
1.15
1.1
1.05
1
0.95
2006
2007
2008
2009
2010


x 10^4
Annual HIV New Treatment from General Population
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
2006
2007
2008
2009
2010

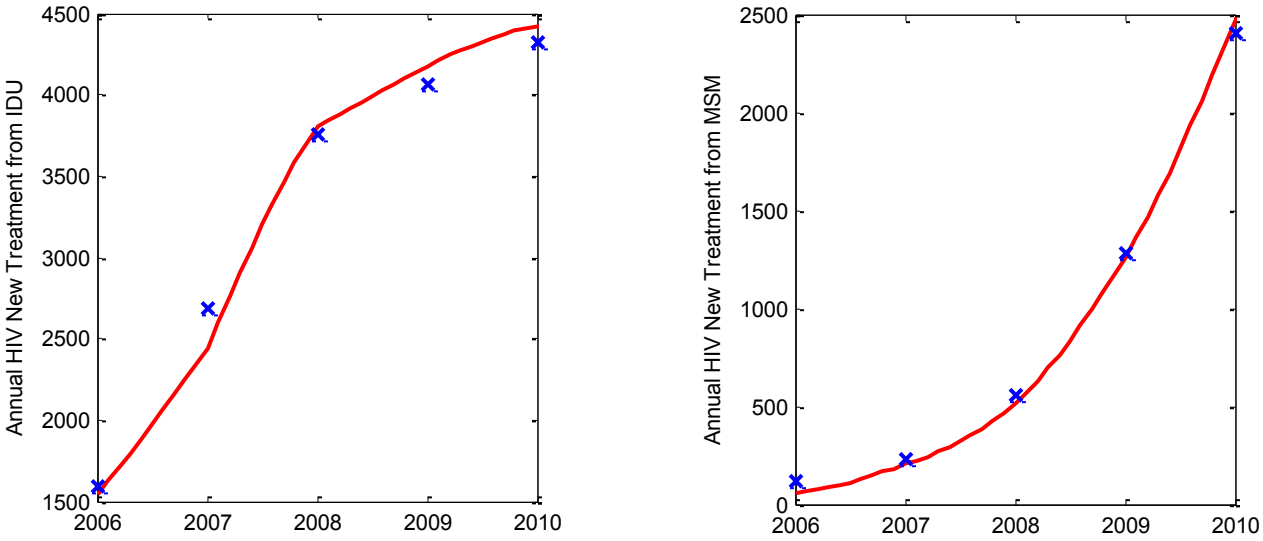


Figure 5.7 Model fitting results

We estimated and predicted HIV new infections in each group based on the assumption that HIV testing and treatment are enhancing from 2011 to 2015 as pre-2010 (table 5.5). Our estimation and prediction showed that annual HIV new infections in MSM kept rising, while annual HIV new infections in general population started to turn down in 2014, and annual HIV new infections in IDU kept declining. Annual HIV new infections for the total population started to decline in 2015, which is consistent with the conclusion of single group SIDT model for China.

Since HIV epidemic in IDU is under control, and the turning point of annual HIV new infections in MSM did not appear, we would like to investigate the effects of the intervention, which maintained the level of HIV testing and treatment in IDU, enhanced that level in MSM and general population, together with harm reduction strategy in MSM (table 5.6). If the harm reduction strategy made MSM within-group transmission coefficient declining by 10%, the turning point of annual HIV new infections in MSM would appear in 2015.

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
General Population	51263	53483	55458	57147	58519	59521	60125	60331	60147	59588
IDU	11154	10502	9858	9240	8679	8151	7645	7187	6776	6407
MSM	8963	10241	11615	13055	14508	15911	17200	18309	19189	19801

Table 5.5 Annual HIV new infections for the general population, IDU and MSM

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
General Population	51263	53483	55458	57147	58519	59523	60132	60347	60175	59630
IDU	11154	10502	9858	9240	8679	8155	7663	7228	6841	6495
MSM	8963	10241	11615	13055	14508	14553	15480	16213	16714	16966

Table 5.6 Annual HIV new infections for the general population, IDU and MSM under a special intervention

Discussion

We started our model fitting based only on surveillance and treatment data for each group from 2006 to 2010. Since we have no HIV testing data for each group, we first assumed that HIV testing rate in each group is the same as that for the whole nation. Model fitting graphs showed that the assumption could be true for general population and MSM , while for IDU other assumption should be tested (Figure 5.8).

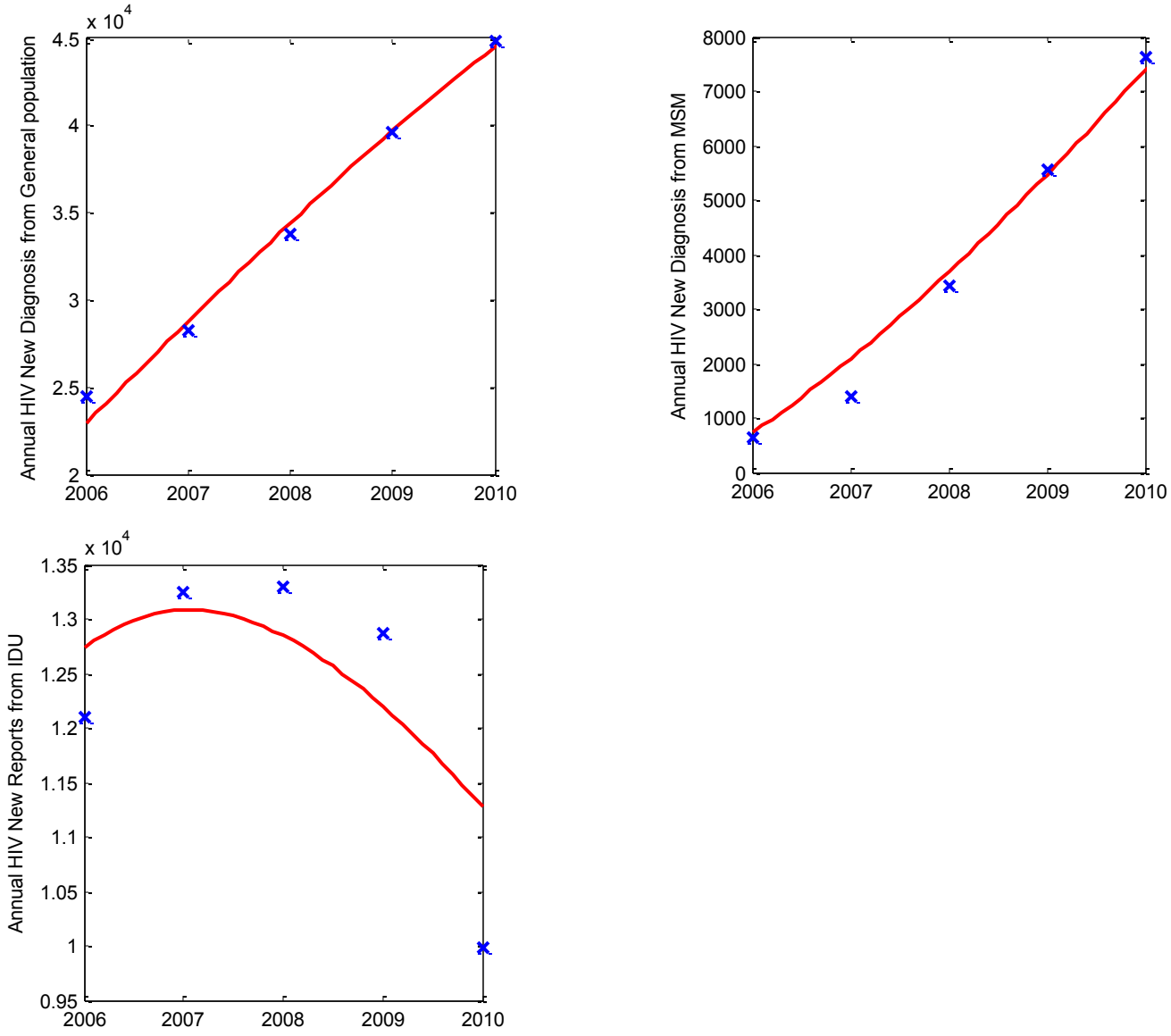


Figure 5.8 Modeling fitting results under the assumption that $\alpha_i(t), i = 1, 2, 3$ are identical

We first tried the assumption that HIV testing rate in IDU followed a quadratic function, which made model fitting worse. We observed that when HIV testing rate in IDU was assumed to be the same as that for the whole nation, the annual new HIV report data of IDU can be fitted well from 2006 to 2009, while the data in 2010 was badly fitted due to the sharp decline in the number of annual HIV new reports from IDU in 2010. There might be some interventions targeting on IDU in 2010, resulting in the sharp

decline in annual HIV reported cases from IDU. We further tried the assumption that from 2006 to 2009, IDU followed the same HIV testing trend as MSM and the general population, while in 2010, it followed different HIV testing rate trend from MSM and the general population. Following this assumption, model fitting becomes better (figure 5.9).

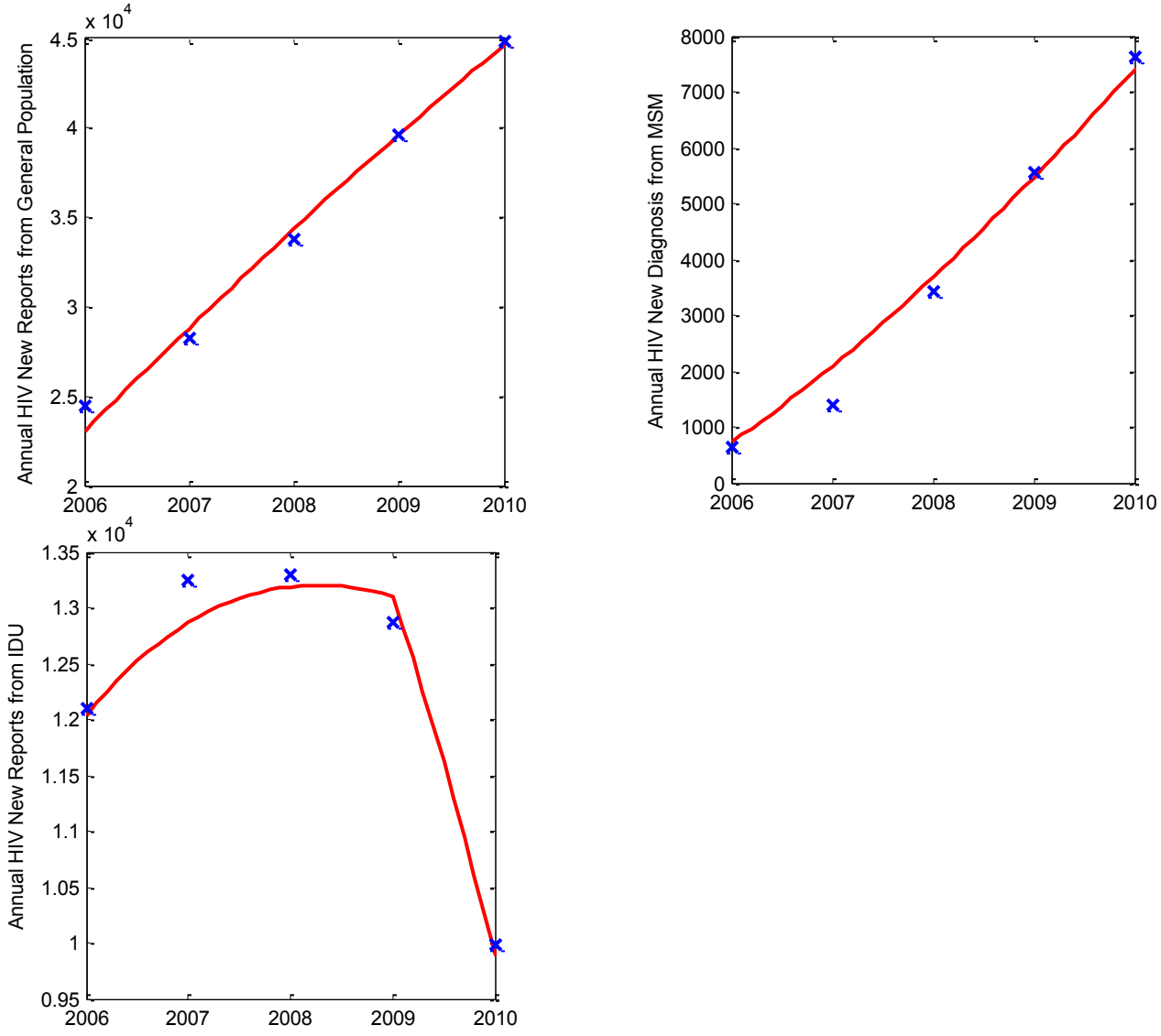


Figure 5.9 Modeling fitting results under the assumption that $\alpha_i(t)$ in IDU is different from that in the other two groups

Based on the above assumption about HIV testing rate in each group, model fitting is well, while model prediction from 2011 to 2015 is not consistent with that from single group SIDT model. Our multi-group model predicted that annual HIV new infections in MSM would increase from 4627 in 2005, to 42897 in 2010 and further to 489654 in 2015. In this case, from 2010 to 2015 annual HIV new infections in MSM would increase by more than 10 times, and in 2015 annual HIV new infections from MSM would account for 70% of the total annual HIV new infections in China.

In single group model study of China, we considered the effect of different forms of $\alpha(t)$ on the trend of HIV new infections. At last, we fixed the form of $\alpha(t)$ by referring to HIV testing rate for the whole nation, which would avoid underestimation or overestimation of annual HIV new infections. While in multi-group model study, we have no information about the HIV testing rate for each group. If we assume HIV testing rate in MSM is the same as the whole nation, the fast increase in the annual number of HIV new diagnosis from MSM would lead to a fast increase in the total number of undiagnosed HIV positive patients, which would further lead to a sharp increase in HIV new infections in MSM. In this case, our prediction for the three indicators would be inconsistent with those from single group study. If we assume HIV testing rate in MSM has the same form as that for the whole nation, a linear function, but with a higher slope, then the fast increase in annual HIV new diagnosis from MSM would be partially attributed to the increase in HIV testing rate, leading to a moderate increase in the number of undiagnosed HIV patients in MSM. In this case, multi-group model predictions for the three indicators were consistent with the predictions from single group study. At last, we used the predictions from the single group model study as data to uniquely fit $\alpha(t)$ in each group.

Fitting results for $\alpha_i(t), i = 1,2,3$ informed us that HIV testing intensity in general population is almost the same as that for the whole nation. HIV testing intensity in IDU increased a little slower than that for the whole nation from 2006 to 2009, which is

consistent with the moderate increase in annual HIV new diagnosis in IDU. HIV testing intensity in MSM in 2006 is lower than that for the whole nation, but it increased faster than that for the whole nation from 2007 to 2010. This estimation would allow us to attribute the sharp increase of annual HIV new diagnosis in MSM partially to the fast increase of HIV testing intensity, which would result in a moderate increase of HIV new infections in MSM.

5.3 Project 3: The HIV dataset for Beijing, China

In 2005, 282 new HIV cases were reported in Beijing, and 19.5% of them came from the high risk group MSM. In 2010, 1138 new HIV cases were reported in Beijing, and 60.2% of them came from MSM. In 2015, 3263 new HIV cases were reported in Beijing, and 75.2% of them came from MSM (Figure 5.10, Table 5.7). With time going, MSM dominated in annual HIV new reports in Beijing.

We built the two-group SIDT model based on HIV surveillance and treatment data from Beijing, which considered the interaction between MSM and the general population. This model will enable us to find out HIV transmission dynamics of MSM and the general population, and effective targeting interventions to control HIV transmission in Beijing.

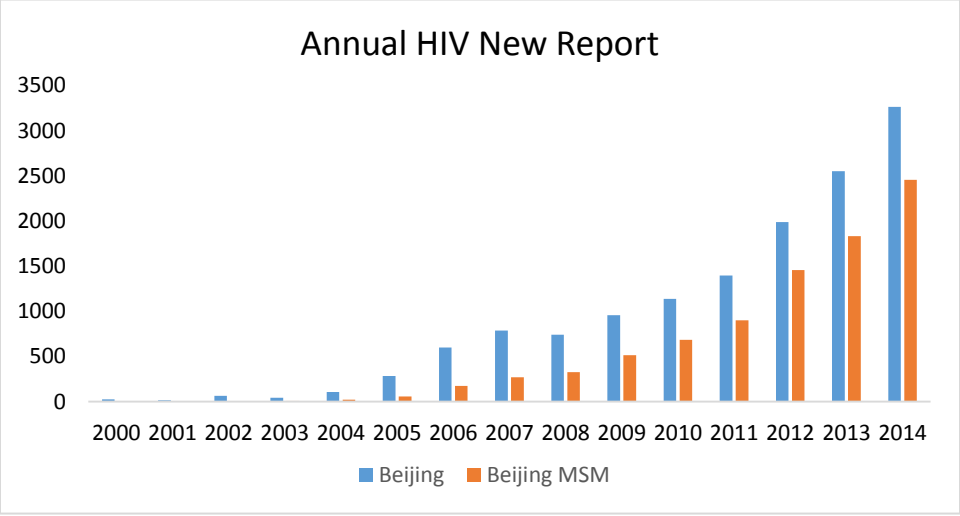


Figure 5.10 Annual HIV new reports in Beijing

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Percentage	0.13	0.23	0.06	0.16	0.19	0.2	0.29	0.34	0.44	0.54	0.60	0.64	0.73	0.72	0.75

Table 5.7 The percentage of HIV new reports from MSM in Beijing

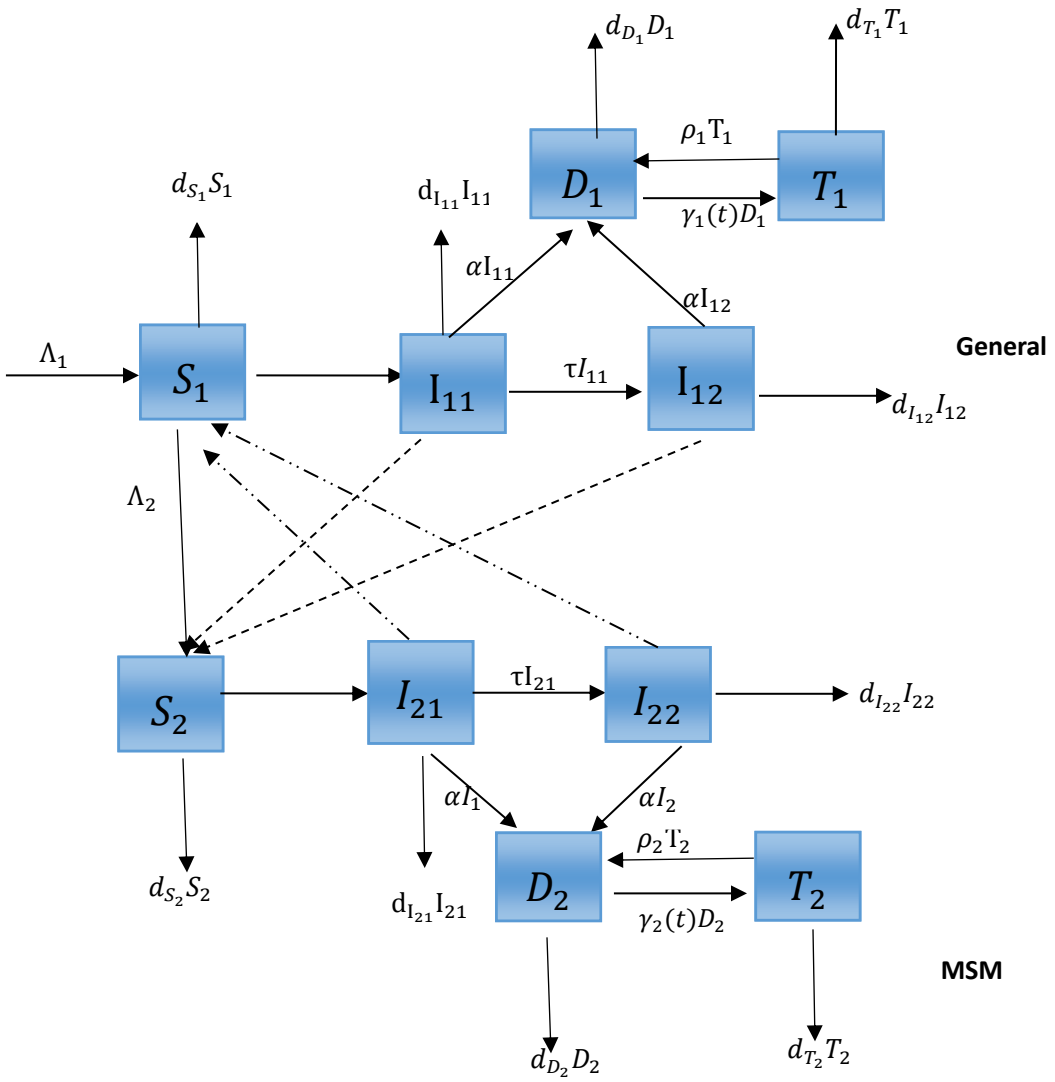


Figure 5.11: Transfer diagram for the two-group HIV transmission model

Based on the transfer diagram in Figure 5.11, the model is described by the following system of nonlinear differential equations.

$$\dot{S}_1(t) = \Lambda_1 - \frac{S_1(t)}{N_1(t)}\Big(\beta_{I_{111}} I_{11}(t) + \beta_{I_{112}} I_{12}(t) + \beta_{1D1} D_1(t) + \beta_{1T1} T_1(t) + \beta_{I_{121}} I_{21}(t) + \beta_{I_{122}} I_{22}(t) + \beta_{1D2} D_2(t) + \beta_{1T2} T_2(t)\Big) - d_{S_1} S_1(t)$$

$$\dot{I}_{11}(t) = \frac{S_1(t)}{N_1(t)}\Big(\beta_{I_{11}} I_{11}(t) + \beta_{I_{12}} I_{12}(t) + \beta_{1D1} D_1(t) + \beta_{1T1} T_1(t) + \beta_{I_{121}} I_{21}(t) + \beta_{I_{122}} I_{22}(t) + \beta_{1D2} D_2(t) + \beta_{1T2} T_2(t)\Big) - \alpha_1(t) I_{11}(t) - \tau I_{11}(t) - d_{I_{11}} I_{11}(t)$$

$$\dot{I}_{12}(t) = \tau I_{11}(t) - \alpha_1(t) I_{12}(t) - d_{I_{12}} I_{12}(t)$$

$$\dot{D}_1(t) = \alpha(t) I_{11}(t) + \alpha(t) I_{12}(t) + \rho_1 T_1(t) - \gamma_1(t) D_1(t) - d_{D_1} D_1(t)$$

$$\dot{T}_1(t) = \gamma_1(t) D_1(t) - \rho_1 T_1(t) - d_{T_1} T_1(t).$$

$$\dot{S}_2(t) = \Lambda_2 S_1(t) - \frac{S_2(t)}{N_2(t)}\Big(\beta_{I_{211}} I_{11}(t) + \beta_{I_{212}} I_{12}(t) + \beta_{2D1} D_1(t) + \beta_{2T1} T_1(t) + \beta_{I_{221}} I_{21}(t) + \beta_{I_{222}} I_{22}(t) + \beta_{2D2} D_2(t) + \beta_{2T2} T_2(t)\Big) - d_{S_2} S_2(t)$$

$$\dot{I}_{21}(t) = \frac{S_2(t)}{N_2(t)}\Big(\beta_{I_{211}} I_{11}(t) + \beta_{I_{212}} I_{12}(t) + \beta_{2D1} D_1(t) + \beta_{2T1} T_1(t) + \beta_{I_{221}} I_{21}(t) + \beta_{I_{222}} I_{22}(t) + \beta_{2D2} D_2(t) + \beta_{2T2} T_2(t)\Big) - \alpha_2(t) I_{21}(t) - \tau I_{21}(t) - d_{I_{21}} I_{21}(t)$$

$$\dot{I}_{22}(t) = \tau I_{21}(t) - \alpha_2(t) I_{22}(t) - d_{I_{22}} I_{22}(t)$$

$$\dot{D}_2(t) = \alpha(t) I_{21}(t) + \alpha(t) I_{22}(t) + \rho_2 T_2(t) - \gamma_2(t) D_2(t) - d_{D_2} D_2(t)$$

$$\dot{T}_2(t) = \gamma_2(t) D_2(t) - \rho_2 T_2(t) - d_{T_2} T_2(t).$$

We split the whole population into two groups, the general population and MSM, denoting them as group 1 and group 2 respectively. At a time t, the number of susceptible people in each group is denoted by $S_i(t), i = 1,2$, the number of HIV positive people who are not diagnosed, infected less than three years and more than three years in each group is denoted by $I_{1i}(t), I_{2i}(t), i = 1,2$, the number of diagnosed HIV positive people who are not under treatment in each group is denoted by $D_i(t), i = 1,2$, and the number of diagnosed HIV positive people that are under treatment in each group is

denoted by $T_i(t), i = 1,2$. The sum $N_i(t) = S_i(t) + I_i(t) + D_i(t) + T_i(t)$ denotes the total population for each group. The time unit used in the model is per year to align with the available data.

Parameter Λ_1 is the influx of susceptibles into the general group, Λ_2 is the transfer rate from general population to MSM. $d_{S_i}, d_{I_{1i}}, d_{I_{2i}}$ d_{D_i} and $d_{T_i}, i = 1,2$, are death rates specific to compartments $S_i, I_{1i}, I_{2i}, D_i, T_i, i = 1,2$. Parameter $\rho_i, i = 1,2$ is the combined rate for treatment failure and loss to follow-up, $\beta_{I_{11i}}, \beta_{I_{22i}}, \beta_{iDi}, \beta_{iTi}, i = 1,2$ are the within-group transmission coefficients for compartment $I_{1i}, I_{2i}, D_i, T_i, i = 1,2, \beta_{I_{12i}}, \beta_{1D2}, \beta_{1T2}$, $i = 1,2$ are the cross-group transmission coefficients for compartment I_{21} $I_{22}, D_2, T_2, \beta_{I_{21i}}, \beta_{2D1}$, β_{2T1} are the cross-group transmission coefficients for compartment I_{11}, I_{12}, D_1, T_1.

Terms $\alpha_i(t)I_i(t)$ and $\gamma_i(t)D_i(t), i = 1,2$ in the model are the annual number of new reports for HIV and AIDS combined and the number of new treatment enrollments in each group, respectively. With the implementation of the "Four Frees and One Care" program in 2003, the Chinese government has rapidly scaled up HIV testing and ART treatment. This was reflected by an increase in the number of people tested for HIV and in the number of treatment centers. To correctly adjust for the increase in new HIV tests and ART treatments, we used a time-dependent diagnosis rate $\alpha_i(t)$and treatment enrolment rate $\gamma_i(t), i = 1,2$.

Data source: Aggregated yearly HIV/AIDS surveillance data for general population and IDU from 2000 to 2014 was obtained from the Surveillance Database of the China Center for Disease Control and Prevention. Aggregated yearly data on ART treatment for the same period for the two groups was obtained from the Treatment Database of the China Center for Disease Control and Prevention. Yearly population data was obtained from the demographic database for Beijing. Predictions for annual HIV new infections, PLHIV and annual HIV/AIDS related deaths from 2015 to 2020 were obtained from single group study

Parameter estimation and fitting: Same as the whole nation**,** Demographic parameters Λ_1 and d_{S_1} are estimated by fitting the equation $N' = \Lambda - d_S N$ to the population data for Beijing from 2000 to 2014. Values of parameters $d_{D_i}, d_{T_i}, \rho_i,\ i = 1,2$ and the form of function $\gamma_i(t), i = 1,2$ are estimated directly from the surveillance and treatment data for Beijing. Since the HIV testing data is not available for these two groups, $\alpha_i(t), i = 1,2$ is assumed to have the same form as that in the single group model, a piece-wise linear function. Values of other parameters, including within-group transmission coefficients and cross-group transmission coefficients, the transfer rate Λ_2, death rate $d_{I_{11}}, d_{I_{12}}, d_{I_{21}}, d_{I_{22}}, d_{S_2}$ and the initial population size $I_{011}, I_{012}, I_{021}, I_{022}, S_{02}$ for the undiagnosed compartment $I_{11}, I_{12}, I_{21}, I_{22}$ and S_2 at the beginning of our fitting (end of 2000), cannot be estimated directly from surveillance data and need to be obtained through model fitting.

We applied our Matrix Decomposition method to the two-group SIDT model to resolve non-identifiability issue [37]. We prefixed the relationships among transmission coefficients by $\beta_{1D1} = 0.75 * \beta_{I_{111}} * I_{11}/(I_{11} + I_{12}) + 0.75 * \beta_{I_{112}} * I_{12}/(I_{11} + I_{12}), \beta_{1T1} = 0.1 * \beta_{I_{111}} * I_{11}/(I_{11} + I_{12}) + 0.1 * \beta_{I_{112}} * I_{12}/(I_{11} + I_{12})), \beta_{2D2} = 0.75 * \beta_{I_{221}} * I_{21}/(I_{21} + I_{22}) + 0.75 * \beta_{I_{222}} * I_{22}/(I_{21} + I_{22}), \beta_{2T2} = 0.1 * \beta_{I_{221}} * I_{21}/(I_{21} + I_{22}) + 0.1 * \beta_{I_{222}} * I_{22}/(I_{21} + I_{22}), \beta_{1D2} = 0.75 * \beta_{I_{121}} * I_{21}/(I_{21} + I_{22}) + 0.75 * \beta_{I_{122}} * I_{22}/(I_{21} + I_{22}), \beta_{1T2} = 0.1 * \beta_{I_{121}} * I_{21}/(I_{21} + I_{22}) + 0.1 * \beta_{I_{122}} * I_{22}/(I_{21} + I_{22}), \beta_{2D1} = 0.75 * \beta_{I_{211}} * I_{11}/(I_{11} + I_{12}) + 0.75 * \beta_{I_{212}} * I_{12}/(I_{11} + I_{12}), \beta_{2T1} = 0.1 * \beta_{I_{211}} * I_{11}/(I_{11} + I_{12}) + 0.1 * \beta_{I_{212}} * I_{12}/(I_{11} + I_{12})$ and performed non-identifiability analysis for the remaining parameters [69-76]. The singular value decomposition informed us that among model parameters $\beta_{I_{111}}, \beta_{I_{112}}, \beta_{I_{121}}, \beta_{I_{122}}, \beta_{I_{221}}, \beta_{I_{222}}, \beta_{I_{211}}, \beta_{I_{212}}, I_{011}, I_{012}, I_{021}, I_{022}, S_{02}, d_{I11}, d_{I12}, d_{I21}, d_{I22}, d_{S2}, \Lambda_2$, and the coefficients of $\alpha_i(t), t = 1,2$ there are 13 dependencies, and the variance decomposition informed us that all the above parameters were involved in the 13 dependencies. Therefore we need to fix 13 parameters' values before model fitting so that the remaining ones can be uniquely estimated. We fixed $d_{I11} = 0.0048$, which is the death rate for general population, $d_{I12} = 0.0811$ by referring to the survival data for HIV

positive people not receiving treatment from a cohort study [68],$d_{I21} = 0.0048$ ($t < 2010$), $d_{I21} = 0.0015$ ($t > 2010$) which is equal to death rate for MSM HIV patients under treatment, $d_{I22} = 0.0811, d_{S2} = 0.0048$. we further assumed that $\beta_{I_{112}} = 0.67\beta_{I_{111}}, \beta_{I_{122}} = 0.67\beta_{I_{121}}, \beta_{I_{222}} = 0.67\beta_{I_{221}}, \beta_{I_{212}} = 0.67\beta_{I_{211}}$ since people that are infected for more than three years have lower viral load than people that are infected less than three years [112]. We also fixed the value for $I_{011} + I_{021}, I_{012} + I_{022}$ by fitting the single group SIDT model to Beijing surveillance data. Parameters Λ_2 and S_{02} were estimated by fitting the population equation $N' = \Lambda - d_S N$ to the MSM population data in Beijing, which accounted for 2%-4% of the total male population [85].

The Nonlinear Least Squares method was applied to find the point estimates for model parameters $\beta_{I_{111}}, \beta_{I_{121}}, \beta_{I_{221}}, \beta_{I_{211}}, I_{021}, I_{022}, \lambda$ and the coefficients for $\alpha_i(t), i = 1,2$, where λ is an arbitrary number between 2% and 4% and used to obtain estimates for $S_{02}$and Λ_2, which minimize the sum of squared error between model output and the available surveillance and treatment data:

- Annual number of new diagnosis of HIV and AIDS combined from 2000 to 2014;
- Annual number of death due to HIV and AIDS among diagnosed for general population from 2000 to 2014;
- Annual treatment enrollment for general population from 2000 to 2014;
- Annual number of death among diagnosed people and people in treatment for general population from 2000 to 2014;
- Annual number of treatment failure and loss to follow-up for general population from 2000 to 2014;
- The number of total population from 2000 to 2014;
- The number of total male population from 2000 to 2014;
- The number of annual HIV new infections from 2015 to 2020 obtained from single group model study

- The number of PLHIV from 2015 to 2020 obtained from single group model study
- The number of HIV/AIDS deaths from 2015 to 2020 obtained from single group model study

Results

Parameters	Description	Best-fit value	Source
β_{111}	Within-group transmission coefficients for I_{11}	0.154	Fitting
β_{112}	Within-group transmission coefficients for I_{12}	0.105	[112]
β_{1D1}	Within-group transmission coefficients for D_1	$0.116I_{11}/(I_{11}+I_{12}) + 0.079I_{12}/(I_{11}+I_{12})$	[69, 70]
β_{1T1}	Within-group transmission coefficients for T_1	$0.015I_{11}/(I_{11}+I_{12}) + 0.01I_{12}/(I_{11}+I_{12})$	[71-76]
β_{121}	Cross-group transmission coefficients for I_{21}	0.0409	Fitting
β_{122}	Cross-group transmission coefficients for I_{22}	0.0274	[112]
β_{1D2}	Cross-group transmission coefficients for D_2	$0.031I_{21}/(I_{21}+I_{22}) + 0.021I_{22}/(I_{21}+I_{22})$	[69, 70]
β_{1T2}	Cross-group transmission coefficients for T_2	$0.004I_{21}/(I_{21}+I_{22}) + 0.003I_{22}/(I_{21}+I_{22})$	[71-76]
β_{221}	Within-group transmission coefficients for I_{21}	0.432	Fitting
β_{222}	Within-group transmission coefficients for I_{22}	0.289	[112]
β_{2D2}	Within-group transmission coefficients for D_2	$0.324I_{21}/(I_{21}+I_{22}) + 0.217I_{22}/(I_{21}+I_{22})$	[69, 70]
β_{2T2}	Within-group transmission coefficients for T_2	$0.0432I_{21}/(I_{21}+I_{22}) + 0.0289I_{22}/(I_{21}+I_{22})$	[71-76]
β_{211}	Cross-group transmission coefficients for I_{11}	0.0237	Fitting
β_{212}	Cross-group transmission coefficients for I_{12}	0.0159	[112]
β_{2D1}	Cross-group transmission coefficients for D_1	$0.0178I_{11}/(I_{11}+I_{12}) + 0.0119I_{12}/(I_{11}+I_{12})$	[69, 70]
β_{2T1}	Cross-group transmission coefficients for T_1	$0.00237I_{11}/(I_{11}+I_{12}) + 0.00159I_{12}/(I_{11}+I_{12})$	[71-76]
d_{S1}	Death rate for S_1	0.0048^*	Fitting
d_{I11}	Death rate for I_{11}	0.0048^*	Fitting
d_{I12}	Death rate for I_{12}	0.0081^*	[68]
d_{D1}	Death rate for D_1	$0.033^*, t \leq 2008$ $0.021, 2008 \leq t \leq 2013$ $0.0106, t > 2013$	Data

d_{T1}	Death rate for T_1	$0.267^*, t \leq 2005$ $0.0316, 2005 \leq t \leq 2010$ $0.0008, t > 2010$	Data
d_{S2}	Death rate for S_2	0.0048^*	Fitting
d_{I21}	Death rate for I_{21}	0.0048^* t<2010 0.0015 t>2010	Fitting
d_{I22}	Death rate for I_{22}	0.0081^*	[68]
d_{D2}	Death rate for D_2	$0.2^*, t \leq 2007$ $-0.008227t + 16.55, 2007 \leq t \leq 2011$ $0.013, 2011 \leq t \leq 2013$ $0.0078, t > 2013$	Data
d_{T2}	Death rate for T_2	$0.011^*, t \leq 2010$ $0.0062, 2010 \leq t \leq 2012$ $0.0015, t > 2012$	Data
ρ_1	Treatment drop-out rate for general Population	$0^*, t \leq 2004$ $-0.0010936t + 2.1948, 2004 \leq t \leq 2007$ $0.06, 2007 \leq t \leq 2009$ $0.031, 2009 \leq t \leq 2011$ $-0.0202t + 40.72, 2011 \leq t \leq 2014$ $0, t > 2014$	Data
ρ_2	Treatment drop-out rate for IDU	$0^*, t \leq 2004$ $-0.0064t + 12.85, 2004 \leq t \leq 2008$ $0.0078t - 15.67, 2008 \leq t \leq 2010$ $0.01, t > 2010$	Data
I_{011}	Undiagnosed HIV positive population in 2000 being infected less than three years for general population	55	Fitting
I_{012}	Undiagnosed HIV positive population in 2000 being infected more than three years for general population	1389	Fitting

I_{021}	Undiagnosed HIV positive population in 2000 being infected less than three years for MSM	20	Fitting
I_{022}	Undiagnosed HIV positive population in 2000 being infected more than three years for MSM	80	Fitting
S_{02}	Susceptible MSM in 2000	260866	Fitting
$\alpha_1(t)$	Time-dependent diagnosis rate of undiagnosed for general population	0.009t-17.9938*, $2000 \le t \le 2005$ 0.0623t-124.846, $t > 2005$	Fitting
$\alpha_2(t)$	Time-dependent diagnosis rate of undiagnosed for MSM	0.0089t-17.7935*, $2000 \le t \le 2005$ 0.06235t-124.946, $t > 2005$	Fitting
$\gamma_1(t)$	Time-dependent treatment enrollment rate for general population	0*, t<2004, $0.13, 2004 \le t \le 2005$ $0.0091637t - 18.33, t > 2005$	Data
$\gamma_2(t)$	Time-dependent treatment enrollment rate for MSM	0*, t<2004, $1, 2004 \le t \le 2005$ $0.02t - 39.89, 2005 \le t \le 2008$ $-0.04733t + 95.294, 2008 \le t \le 2010$	Data
Λ_1	The influx of susceptible into the general group	716991	Fitting
Λ_2	Transfer proportion from general population to MSM	14626	Fitting
R^2	Goodness of fit ratio	0.94	Fitting

Table 5.8. Model parameters and their best-fit values

Notes: * The unites for these parameters are 1/person/year

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
General	179	198	219	241	265	291	319	351	388	432	485	548	612	676	737
MSM	49	75	111	161	229	318	437	590	783	1022	1319	1670	1989	2270	2495

	Year	2015	2016	2017	2018	2019	2020
	General	788	832	868	896	918	935
Scenario 1	MSM	2635	2721	2747	2731	2695	2656
	General	788	837	881	922	960	998
Scenario 2	MSM	2636	2722	2750	2737	2705	2671
	General	784	830	881	938	1000	1068
Scenario 3	MSM	2338	2398	2491	2614	2765	2943
	General	777	810	847	888	934	984
Scenario 4	MSM	2018	1989	1988	2015	2069	2144
	General	782	818	849	877	904	933
Scenario 5	MSM	2321	2295	2232	2152	2075	2014

Table 5.9 Annual HIV new infections under various scenarios

We predicted annual HIV new infections in each group based on various assumed scenarios (Table 5.9). Scenario 1 assumed that HIV testing and treatment are enhancing after 2014 as pre-2014 trend. Our prediction showed that annual HIV new infections in the general population kept rising, while annual HIV new infections in MSM started to turn down in 2019. And annual HIV new infections for the total population started to decline in 2019, which is consistent with the conclusion of the single group SIDT model study for Beijing.

Table 5.9 showed that enhancing HIV testing and treatment can not turn HIV new infections down in the general population. Considering the high cost of enhancing HIV testing and HIV treatment in the general population, we would like to investigate the effect of maintaining the level of HIV testing and treatment in the general population and

enhancing that in MSM (Scenario 2). Comparisons of the effects of Scenario 1 and Scenario 2 informed us that the effect of enhancing HIV testing and HIV treatment in the general population is not so obvious, since only 139, 39 new HIV infections can be averted for the general population and MSM during the period 2015-2020..

Since the effect of enhancing HIV testing and HIV treatment in the general population is not obvious, we would like to investigate the effect of maintaining the level of HIV testing and treatment in the general population, implementing various targeting interventions in MSM. In Scenario 3 and Scenario 4, we assumed that the HIV transmission coefficient in MSM was reduced by 10% and 20% due to the implementation of the harm reduction intervention (knowledge education, promotion of condom use) or the WHO recommended intervention (treating all MSM), while the level of HIV testing and treatment was maintained on the 2014 level in the general population and MSM. In this case, we could avert 547 (Scenario 3) and 3604 (Scenario 4) new HIV infections respectively during 2015-2020 comparing to the effect of enhancing HIV testing and HIV treatment in MSM.

At last we would like to investigate the effect of implementing the harm reduction intervention or the WHO recommended intervention, together with enhancing HIV testing and treatment in MSM (Scenario 5). If the harm reduction intervention or the WHO recommended intervention can reduce the HIV transmission coefficient in MSM by 10% and the HIV testing and treatment in MSM can be enhanced as pre-2014 trend, 3320 new HIV infections can be averted during 2015-2020 comparing to merely enhancing HIV testing and HIV treatment in MSM.

Discussion

HIV surveillance data of Beijing informed us that MSM dominate in HIV transmission in Beijing, and our model estimation revealed that most of HIV transmissions in Beijing come from MSM within-group transmission. Therefore it would be effective of controlling HIV epidemic in Beijing by implementing targeting interventions on MSM.

Our simulations showed that both enhancing HIV testing, treatment and implementing harm reduction interventions or treating susceptible MSM are effective in controlling HIV transmission. Further investigation on the cost of the above two interventions is necessary before advising public health agency for the future decisions.

In this chapter, we studied the multi-group SIDT model, which considered HIV process susceptible-infected-diagnosed-treated in the general population and high risk groups such as IDU, MSM. Our method discussed in Chapter 3 was applied to resolve non-identifiability problem in this model. Although model became more complicated, more model parameters appeared, our method was effective in resolving non-identifiability problem.

In project one, we built a two-group SIDT model, which considered the interaction between IDU and the general population, to study the HIV dataset from two remote townships in Sichuan province of China. After model parameters were uniquely estimated, HIV epidemic in the general population and IDU was estimated and predicted. The estimation and prediction was consistent with the results we obtained in the single group model study. The consistency between our model estimation and the rare prevalence data in 2010 provided a validation for our multi-group SIDT model.

In project two, we built a three-group SIDT model, which considered the interaction among IDU, MSM and the general population, to study HIV dataset for China. To keep multi-group model predictions consistent with that from the single group study, we added model predictions for annual HIV new infections, PLHIV, and HIV/AIDS related deaths during the period 2011-2015 from the single group model study to our dataset. Based on this new dataset, model parameters were uniquely estimated, and HIV epidemic in the general population, IDU and MSM were estimated and predicted. The results were consistent with that from single group study.

In project three, we built a two-group SIDT model, which considered the interaction between MSM and the general population, to study HIV dataset for Beijing. In this model

we split the I group into two parts: I_1, HIV patients that are infected less than three years and I_2, HIV patients that are infected more than three years. To keep model predictions consistent with that from the single group study, we also added model predictions of annual HIV new infections, PLHIV, HIV/AIDS related deaths during the period 2015-2020 from the single group study for Beijing to our dataset. Based on this new dataset, HIV epidemic in MSM and general population in Beijing was estimated and predicted, and the results were consistent with that from the single group study.

In the single group study, there was a dependency between transmission coefficients and the coefficients for $\alpha(t)$, we can fix the coefficients for $\alpha(t)$ to estimate transmission coefficients by referring to HIV testing data. In the multi-group model study, the above dependency existed in each group, while we have no information about HIV testing data for each group, we have to assume $\alpha_i(t)$ to be the same as the whole region. Although model fitting was well, model predictions can be very different from that obtained from the single group study, because the above assumption might result in overestimation or underestimation of HIV epidemic in each group. To keep prediction results consistent with each other, we added new data to our dataset when we performed multi-group model fittings. The new dataset enables $\alpha_i(t)$ in each group being identifiable, therefore we could obtain the information of HIV testing intensity in each group.

Chapter 6

Conclusion

6.1 Summary of results

Parameter estimation is a crucial step for using modeling methods to perform estimation and prediction. The challenge for parameter estimation under modeling framework lies in the existence of non-identifiability, which means that infinite parameter values will have almost the same observable model outputs. Existing methods for resolving non-identifiability will inform us whether there is non-identifiability issue or not, if there is, a ranking for model parameters from the least identifiable one to the most identifiable one is obtained, and we are suggested to fix some least identifiable parameters such that all the remaining ones can be uniquely estimated. It is not always possible to fix some least identifiable parameters, such as transmission coefficients in disease models. In this case, it is necessary for us to study linear dependencies among model parameters, such that we can have more flexibility in fixing model parameters' values.

Our method, the Matrix Decomposition Method has two steps, the singular value decomposition and the variance decomposition. Singular value decomposition will inform us how many dependencies existing among model parameters, which are equal to the number of model parameters we need to fix before model fitting. Variance decomposition will inform us which parameters are involved in each dependency. Therefore, for each dependency, we can choose to fix the value of one parameter which is involved in the dependency, and is possible to be fixed such that the dependency disappears.

All applications in Chapter 4 and Chapter 5 demonstrated the applicability of our method. Our method was first applied to a simple SIDT model which describes the process of HIV patients' infection-diagnosis-treatment to resolve non-identifiability issue. After that, model parameters can be uniquely determined based on various HIV datasets from China. Secondly our method was applied to a multi-group SIDT model, which

considers interactions between HIV high risk groups and the general population. Although the model becomes more complicated with the appearing of more model parameters and more dependencies among parameters, our method is effective in determining the dependencies among model parameters.

Our method is a general approach for resolving non-identifiability issue, and it can be applied to any model to determine dependencies among model parameters.

6.2 Future possibilities

Economic analysis is necessary to be incorporated into our modeling work, since we need to find the most effective intervention to control HIV transmission in each region. This analysis would allow the benefits of various interventions to be qualified, and enable us to directly compare the benefits of different strategies.

Bibliography

[1] S. Aluru. *Haendbook of Computational Molecular Biology (Chapman & All/Crc Computer and Information Science Series)*. Chapman & Hall/CRC; Boca Raton, FL: 2005.

[2] C. Castillo-Chavez, S. Blower, P. van de Driessche, D. Kirschner, A. Yakubu. *Mathematical approaches for emerging and reemerging infectious diseases: models, methods, and theory*. Springer; New York: 2002.

[3] D. A. Lauffenburger, J. L. Linderman. *Receptors: models for binding, trafficking, and signaling*. Oxford University Press; Oxford: 1993.

[4] J. D. Murray. *Mathematical biology I: An introduction*. Springer; New York: 2002.

[5] J. D. Murray. *Mathematical biology II: Spatial models and biomedical applications*. Springer; New York: 2003.

[6] M. A. Nowak, R. M. May. *Virus dynamics: mathematical principles of immunology and virology*. Oxford University Press; Oxford: 2000.

[7] C. Eisenberg, L. Robertson, H. Tien. *Identifiability and estimation of multiple transmission pathways in cholera and waterborne disease*. Journal of Theoretical Biology. 2013;324:84-102.

[8] D. Pauw, K. Steppe, B. Baets. *Identifiability analysis and improvement of a tree water flow and storage model*. Mathematical Biosciences. 2008;211:314-332.

[9] F. Canova, L. Sala. *Back to square one: identification issues in DSGE models. Journal of Monetary Economics*. 2009;56:431-449.

[10] R. Jayasankar, A. Ben-Zvi, B. Huang. *Identifiability and estimability study for a dynamic solid oxide fuel cell model*. Computers & Chemical Engineering. 2009;33:484-492.

[11] S. Srinath, R. Gunawan. *Parameter identifiability of power-law biochemical system models*. Journal of Biotechnology. 2010;149:132-140.

[12] H. Wu, H. Zhu, H. Miao, A. S. Perelson. *Parameter identifiability and estimation of HIV/AIDS dynamic models*. Bulletin of Mathematical Biology. 2008;70:785-799.

[13] N. Metropolis, S. Ulam. *The Monte Carlo Method*. J. Amer. Stat. Assoc. 1949; 44:335–341.

[14] H. Miao, C. Dykes, L. M. Demeter, A. S. Perelson, H. Wu. *Modeling and estimation of kinetic parameters and replication fitness of HIV-1 from Flow-Cytometry-based growth competition experiments*. Bull. Math. Biol. 2008;70:1749–1771.

[15] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, et al. *Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood*. Bioinformatics. 2009;25 (15):1923–1929.

[16] J. A. Jacquez. *Compartmental analysis in biology and medicine*. The University of Michigan Press; Michigan: 1985.

[17] J. Guedj, R. Thiebaut, D. Commenges. *Practical identifiability of HIV dynamics models*. Bull. Math. Biol. 2007;69:2493–2513.

[18] M. Rodriguez-Fernandez, J. A. Egea, J. R. Banga. *Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems*. BMC Bioinformatics. 2006a;7:483–500.

[19] M. Rodriguez-Fernandez, P. Mendes, J. R. Banga. *A hybrid approach for efficient and robust parameter estimation in biochemical pathways*. Biosystems. 2006b; 83:248–265.

[20] D. E. Zak, G. E. Gonye, J. S. Schwaber, F. J. Doyle. *Importance of input perturbations and stochastic gene expression in the reverse engineering of genetic regulatory networks: insights from an identifiability analysis of an in silico network. Genome*. Res. 2003;13:2396–2405.

[21] T. Quaiser, M. Mönnigmann. *Systematic identifiability testing for unambiguous mechanistic modeling - application to JAK-STAT, MAP kinase, and NF-κB signaling pathway models*. BMC Syst. Biol. 2009;3:50–71.

[22] K. Z. Yao, B. M. Shaw, B. Kou, K. B. McAuley, D. W. Bacon. *Modeling ethylene/butane copolymerization with multi-site catalysts: Parameter estimability and experimental design*. Polym. React. Eng. 2003;11:563–588.

[23] K. G. Gadkar, R. Gunawan, F. J. Doyle. *Iterative approach to model identification of biological networks*. BMC Bioinformatics. 2005;6:155–175.

[24] H. Yue, M. Brown, J. Knowles, H. Wang, D. S. Broomhead, D. B. Kell. *Insights into the behaviour of systems biology models from dynamic sensitivity and identifiablity analysis: a case study of an NF-kappaB signalling pathway*. Mol. BioSyst. 2006;2:640–649.

[25] T. Quaiser, W. Marquardt, M. Mönnigmann. *Local identifiability analysis of large signaling pathway models. 2nd Conference Foundation of Systems Biology in Engineering, Proceedings Plenary and Contributed Papers*. Fraunhofer IRB Verlag; Stuttgart, Germany. 2007;465-470.

[26] K. Schittkowski. *Experimental design tools for ordinary and algebraic differential equations*. Ind. Eng. Chem. Res. 2007;46:9137–9147.

[27] S. Vajda, H. Rabitz, E. Walter, Y. Lecourtier. *Qualitative and quantitative identifiability analysis of nonlinear chemical kinetiv-models*. Chem. Eng. Commun. 1989;83:191–219.

[28] D. M. Glover, W. J. Jenkins, S. C. Doney. *Least Squares and regression techniques, goodness of fit and tests, non-linear least squares techniques*. Woods Hole Oceanographic Institute; 2008.

[29] W. R. Gilks, S. Richardson, D. J. Spiegelhalter. *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC; 1996.

[30] O. Ghasemi, M. Lindsey, T. Yang, N. Nguyen, Y. Huang, Y. Jin. *Bayesian parameter estimation for nonlinear modelling of biological pathways*. BMC Syst. Biol. 2011;5(Suppl 3):S9.

[31] S. Blower, H. Dowlatabadi. *Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model as an example*. Int. Stat. Rev. 1994;62:229-43.

[32] D. M. Hamby. *A review of techniques for parameter sensitivity analysis of environmental models*. Environ. Monit. Assess. 1994;32(2):135-154.

[33] C. Daniel. *One-at-a-time plans*. J. Amer. Stat. Assoc. 1973;68:353-360.

[34] S. Marino, I. Bogue, C. Ray, D. Kirschner. *A methodology for performing global uncertainty and sensitivity analysis in systems biology*. Journal of Theoretical Biology. 2008;254:178-196.

[35] R. R. Picard and R. D. Cook. *Cross-validation of regression models*. Journal of the American Statistical Association. 1984;79(387):575-583.

[36] M. C. Eisenberg, S. L. Robertson, J. H. Tien. *Identifiability and estimation of multiple transmission pathways in cholera and waterborne disease*. J. Theo. Bio. 2013;324(7):84-102.

[37] H. Miao, X. Xia, A. S. Perelson, H. Wu. *On identifiability of Nonlinear ODE models and applications in viral dynamics*. SIAM REVIEW. 2011;53(1):3-39.

[38] G. H. Golub, C. Reinsch. *Singular value decomposition and least squares solutions*. Numerische Mathematik. 1970;14(5):403-420.

[39] D. A. Belsley, E. Kuh, R. E. Welsch. *Regression diagnostics: Identifying influential data and sources of collinearity*. John Wiley & Sons; 2005.

[40] P. Yan, F. Zhang, H. Wand. *Using HIV diagnostic data to estimate HIV incidence: method and simulation*. Stat. Comm. Infect. Dis. 2011;3(1):1-28.

[41] N. G. Becker, C. Lewis, Z. Li, A. *Mcdonald. Age-specific Back-projection of HIV diagnosis data*. Statistics in Medicine. 2003;22:2177-2190.

[42] M. J. Sweeting, D. Angelis, O. Aalen. *Bayesian back-calculation using a multi-state model with application to HIV*. Statistics in Medicine. 2005;24:3991-4007

[43] H. I. Hall, R. Song, P. Phodes, J. Prejean, Q. An, L. M. Lee, J. Karon, R. Brookmeyer, E. H. Kaplan, M. T. Mckenna. *Estimation of HIV in the United States*. Journal of the American Medical Association. 2008;300:520-529.

[44] H. Wand, P. Yan, D. Wilson, A. Mcdonald, M. Middleton, J. Kaldor, M. Law. *Increasing HIV transmission through male homosexual and heterosexual contact in Australia: results from an extended back-projection approach*. HIV medicine. 2010;11:395-403.

[45] Q. Yang, D. Boulos, F. Zhang, R. S. Remis, D. Schanzer, C. P. Archibald. *Estimates of the number of prevalent and incident human immunodeficiency virus (HIV) infections in Canada, 2008*. Canadian Journal of Public Health. 2010;101:486-490.

[46] J. Stover, T. Brown, M. Marston. *Updates to the Spectrum/Estimation and Projection Package (EPP) model to estimate HIV trends for adults and children*. Sex. Transm. Infect. 2012;88:i11-i16.

[47] L. Bao, J. A. Salomon, T. Brown, A. E. Raftery, D. R. Hogan. *Modeling national HIV/AIDS epidemics: revised approach in the UNAIDS Estimation and Projection Package 2011*. Sexually Transmitted Infections. 2012;88:i3-i10.

[48] T. Brown, A. E. Raftery, J. A. Salomon, R. F. Baggaley, J. Stover, P. Gerland. *Modeling HIV epidemics in the antiretroviral era: the UNAIDS Estimation and Projection Package 2009*. Sexually Transmitted Infections. 2010;86:ii3-ii10.

[49] J. Stover, P. Johnson, T. Hallett, M. Marston, R. Becquet, I. M. Timaeus. *The spectrum projection package: improvements to estimating incidence by age and sex, mother-to-child transmission, HIV progression in Child and double orphans.* Sexually Transmitted Infections. 2010;86:ii6-ii21.

[50] UNAIDS Reference Group on Estimates, Modeling and Projections. *Improved methods and assumptions for estimation of the HIV/AIDS epidemic and its impact:*

recommendations of the UNAIDS Reference Group on Estimates, Modeling and Projections. AIDS. 2002;16:w1-w14.

[51] A. M. Presanis, D. De Angelis, A. Goubar, O. N. Gill, A. E. Ades. *Bayesian evidence synthesis for a transmission dynamic model for HIV among men who have sex with men*. Biostat. 2011;12: 666–81.

[52] E. Bendvid, M. L. Brandeau, R. Wood, D. K. Owens. *Comparative effectiveness of HIV testing and treatment in highly endemic regions*. Archives of Internal Medicine. 2010;170:1347-1354.

[53] D. Bezemer, F. Wolf, M. C. Boerlijst, A. V. Sighem, Hollingsworth TD, Prins M, Geskus RB, Coutinho RA, Fraser C. *A resurgent HIV-1 epidemic among men who have sex with men in the era of potent antiretroviral therapy*. AIDS. 2008;22:1071-1077.

[54] S. Blower, H. Dowlatabadi. *Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as example*. International Statistical Review. 1994;62:229-243.

[55] W. EL-Sadr, B. Coburn, S. Blower. *Modeling the impact on the HIV epidemic of treating discordant couples with antiretroviral to prevent transmission*. AIDS. 2011;25:2295-2299.

[56] K. L. Zhang, S. J. Ma. *Epidemiology of HIV in China*. BMJ. 2002;324:803-4.

[57] UNAIDS Asia Pacific Intercountry Team. *UNAIDS Task Force on Drug Use and HIV: Drug Use and HIV Vulnerability: Policy Research Study in Asia*. Bangkok: 2000.

[58] Y. Ma, Z. Li, K. X. Zhang, W. Yang, X. Ren, Y. Yang, et al. *Identification of HIV infection among drug users in China [in Chinese]*. Chinese J. Epidemiol. 1990;11: 184-5.

[59] Chinese Center for Disease Control and Prevention: *Analysis of HIV/STD epidemic in 2002*. Beijing, China.

[60] Chinese Center for Disease Control and Prevention: *Analysis of HIV/STD epidemic in 2005*. Beijing, China.

[61] H. Qian, S. H. Vermund, N. Wang. *Risk of HIV/AIDS in China: subpopulations of special importance*. Sex. Transm. Infect. 2005;81(6):442-447.

[62] Chinese Center for Disease Control and Prevention: *Analysis of HIV/STD epidemic in 2007*. Beijing, China.

[63] Chinese Center for Disease Control and Prevention: *Analysis of HIV/STD epidemic in 2011*. Beijing, China.

[64] Z. Jia, L. Wang, R. Chen, D. Li, L. Wang, et al. *Tracking the evolution of HIV/AIDS in China from 1989–2009 to inform future prevention and control efforts*. PLoS ONE. 2011;6:e25671.

[65] Y. Ma, F. Zhang, Y. Zhao, C. Zang, D. Zhao, A. Dou, et al. *Cohort Profile: The Chinese national free antiretroviral treatment cohort*. Int. J. Epidemiol. 2010;39(4):973-9.

[66] Chinese Center for Disease Control and Prevention: *Analysis of HIV/STD epidemic in 2007-2012*. Beijing, China.

[67] J. Todd, J. R. Glynn, M. Marston, T. Lutalo, et al. *Time from HIV seroconversion to death: a collaborative analysis of eight studies in six low and middle-income countries before highly active antiretroviral therapy*. AIDS. 2007;Suppl 6:S55-63.

[68] Collaborative Group on AIDS Incubation and HIV Survival including the CASCADE EU Concerted Action. *Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative re-analysis*. Lancet. 2000;355(9210):1131-7.

[69] G. Marks, N. Crepaz, R. S. Janssen. *Estimating sexual transmission of HIV from persons aware and unaware that they are infected with the virus in the USA*. AIDS 2006;20:1447-1450.

[70] J. Fox, P. J. White, N. Macdonald, et al. *Reductions in HIV transmission risk behaviour following diagnosis of primary HIV infection: a cohort of high-risk men who have sex with men*. HIV Medicine. 2009;10:432-438.

[71] D. Donnell, J. M. Baeten, J. Kiarie, et al. *Heterosexual HIV-1 transmission after initiation of antiretroviral therapy: a prospective cohort analysis*. Lancet. 2010;375(9731):2092-2098.

[72] R. Granich, R. Crowley, M. Vitoria, C. Smyth, et al. *Highly active antiretroviral treatment as prevention of HIV transmission: Review of scientific evidence and update*. Curr. Opin. HIV. AIDS. 2010;5(4):298-304.

[73] D. Donnell. *ART and Risk of Heterosexual HIV-1 Transmission in HIV-1 Serodiscordant African Couples: A Multinational Prospective Study. at 17th Conference on Retroviruses and Opportunistic Infections (CROI)*. San Francisco, USA: 2010.

[74] J. Castilla, J. Del Romero, V. Hernando, B. Marincovich, S. Garcia, C. Rodriguez. *Effectiveness of highly active antiretroviral therapy in reducing heterosexual transmission of HIV*. J. Acquir. Immune. Defic. Syndr. 2005;40:96–101.

[75] R. Bunnell, J. P. Ekwaru, P. Solberg, N. Wamai, W. Bikaako-Kajura, W. Were, A. Coutinho, C. Liechty, E. Madraa, G. Rutherford, et al. *Changes in sexual behavior and risk of HIV transmission after antiretroviral therapy and prevention interventions in rural Uganda*. AIDS. 2006;20:85–92.

[76] T. C. Quinn, M. J. Wawer, N. Sewankambo, et al. *Viral load and heterosexual transmission of human immunodeficiency virus type 1*. N. Engl. J. Med. 2000;342: 921–29.

[77] J. Arino, F. Brauer, P. van den Driessche, J. Watmough, J. Wu. *Simple model for containment of a pandemic*. J. R. Soc. Interface. 2006;3:453-457.

[78] J. Stover, P. Johnson, B. Zaba, M. Zwahlen, F. Dabis, R. E. Ekpini. *The Spectrum projection package: improvements in estimating mortality, ART needs, PMTCT impact and uncertainty bounds*. Sex Transm. Infect. 2010;84 (Suppl I):i24–i30.

[79] N. Wang, L. Wang, Z. Wu, W. Guo, X. Sun, K. Poundstone, Y. Wang, National Expert Group on HIV/AIDS Estimation. *Estimating the number of people living with HIV/AIDS in China: 2003-09*. Int. J. Epidemiol. 2010;39 Suppl 2:ii21-8.

[80] H. I. Hall, R. Song, P. Rhodes, et al. *Estimation of HIV Incidence in the United States*. JAMA. 2008;300(5):520-529.

[81] J. Prejean, R. Song, A. Hernandez, et al. *Estimated HIV incidence in the United States, 2006-2009.* PLoS One. 2011;6(8):e17502.

[82] P. J. Birrell, O. N. Gill, V. C. Delpech, et al. *HIV incidence in men who have sex with men in England and Wales 2001-10: a nationwide population study*. Lancet Infect. Dis. 2013;13(4):313-8.

[83] Ministry of Health, China. *2011 Estimates for the HIV/AIDS Epidemic in China*. 2011. http://www.docin.com/p-399429297.html&key. Accessed Nov 2011.

[84] P. Li. *The establishment of several epidemiological analysis methods and their application in a 11.5th comprehensive prevention and control of HIV/AIDS demonstration area with high HIV/AIDS prevalence - a major infectious diseases special investigation in Butuo County*. PhD Dissertation, Guangxi Medical University and Chinese Center for Disease Control and Prevention. 2013.

[85] H. Qian, S. H. Vermund, N. Wang. *Risk of HIV/AIDS in China: subpopulations of special importance*. Sex. Transm. Infect. 2005;81(6):442-447.

[86] China CDC. *Analysis of HIV/STD epidemic in 2002*. Beijing, China: 2002.

[87] L. Wei, J. Chen, M. Rodolph, et al. *HIV incidence, retention, and changes of high-risk behaviors among rural injection drug users in Guangxi, China*. Subst. Abus. 2006;27(4):53-61.

[88] Y. Ruan, G. Qin, S. Liu, et al. *HIV incidence and factors contributed to retention in a 12-month follow-up study of injection drug users in Sichuan Province, China*. J. Acquir. Immune. Defic. Syndr. 2005;39(4):459-463.

[89] Y. Zhang, H. Shan, J. Trizzino, et al. *HIV incidence, retention rate, and baseline predictors of HIV incidence and retention in a prospective cohort study of injection drug users in Xinjiang, China*. Int. J. Infect. Dis. 2007;11(4):318-323.

[90] L. Yin, G. Qin, H. Z. Qian, et al. *Continued spread of HIV among injecting drug users in southern Sichuan Province, China*. Harm Reduct. J. 2007;4:6.

[91] M. G. Han, Q. F. Chen, Y. Han, et al. *Design and implementation of a China comprehensive AIDS response program (China CARES), 2003–08*. Int. J. Epidemiology. 2010;39:ii47–ii55.

[92] B. Liu, S. G. Sullivan, Z. Wu. *An evaluation of needle exchange programs in China*. AIDS. 2007;21 (suppl 8):S123–S128.

[93] W. Z. Yin, Y. Hao, X. H. Sun, et al. *Scaling up the national methadone maintenance treatment program in China: achievements and challenges*. Int. J. Epidemiology. 2010;39:ii29–ii37.

[94] K. M. Rou, S. G. Sullivan, P. Liu, et al. *Scaling up prevention programs to reduce the sexual transmission of HIV in China*. Int. J. Epidemiology. 2010;39:ii38–ii46.

[95] S. Lai, W. Liu, J. Chen, et al. *Changes in HIV-1 incidence in heroin users in Guangxi Province, China*. J. Acquir. Immune. Defic. Syndr. 2001;26(4):365-70.

[96] X. Yu, J. Chen, Y. Shao, et al. *Emerging HIV infections with distinct subtypes of HIV-1 infection among injection drug users from geographically separate locations in Guangxi Province, China*. J. Acquir. Immune Defic. Syndr. 1999;22(2):180-8.

[97] C. Beyrer, M. H. Razak, K. Lisam, et al. *Overland heroin trafficking routes and HIV-1 spread in south and southeast Asia*. AIDS. 2000;14(1):75-83.

[98] X. Yu, J. Chen, Y. Shao, et al. *Two subtypes of HIV-1 among injection-drug users in southern China*. Lancet. 1998;351(9111):1250.

[99] L. Wei, J. Chen, M. Rodolph, et al. *HIV incidence, retention, and changes of high-risk behaviors among rural injection drug users in Guangxi, China*. Subst. Abus. 2006;27(4):53-61.

[100] Guangxi Public Health Department. *Annual Report on Provincial AIDS/STD Surveillance, 2003 -2010*. Guangxi: 2004 – 2011.

[101] Z. Jia, L. Wang, R. Chen, D. Li, L. Wang, et al. *Tracking the evolution of HIV/AIDS in China from 1989-2009 to inform future prevention and control efforts*. PLoS ONE. 2011;6:e25671.

[102] Y. Ma, F. Zhang, Y. Zhao, C. Zang, et al. *Cohort Profile: The Chinese national free antiretroviral treatment cohort*. Int. J. Epidemiol. 2010;39:973-9.

[103] R. Smith, J. Okano, J. Kahn, E. Bodine, S. Blower. *Evolutionary dynamics of complex networks of HIV drug-resistant strains: the case of San Francisco*. Science. 2010;327:697-701.

[104] A. E. Kurth, C. Celum, J. M. Baeten, S. H. Vermund, J. H. Wasserheit. *Combination HIV prevention: significance, challenges, and opportunities*. Curr. HIV/AIDS Rep. 2011;8(1):62-72.

[105] Y. Ruan, S. Liang, J. Zhu, X. Li, S. Pan, Q. Liu , B. Song, Q. Wang, H. Xing, Y. Shao. *Evaluation of harm reduction programs on seroincidence of HIV, hepatitis B and C, and syphilis among intravenous drug users in southwest China*. Sex. Transm. Dis. 2013;40(4):323-8.

[106] L. Zhang, S. Liang, W. Lu, S. Pan, B. Song, Q. Liu, Y. Xu, H. Dong, H. Xing, Y. Shao, Y. Ruan. *HIV, syphilis, and behavioral risk factors among female sex workers*

before and after implementation of harm reduction programs in a high drug-using area of China. PLoS One. 2014;9(1):e84950

[107] J. D. Shelton. *HIV/AIDS. ARVs as HIV prevention: a tough road to wide impact. Science*. 2011;334(6063):1645-6.

[108] R. Hayes, K. Sabapathy, S. Fidler. *Universal testing and treatment as an HIV prevention strategy: research questions and methods*. Curr. HIV Res. 2011;9(6):429-45.

[109] M. W. Tyndall, K. J. Craib, S. Currie, K. Li, M. V. *O'Shaughnessy, M. T. Schechter. Impact of HIV infection on mortality in a cohort of injection drug users*. J. Acquir. Immune Defic. Syndr. 2001;28:351-357.

[110] Chinese Ministry. *White paper about homosexuality in China*. Beijing: 2004.

[111] Chinese Ministry. *Drug Control Report in China.* Beijing: 2006-2010.

[112] M. J. Chesebro, W. D. Everett. *Understanding the guidelines for treating HIV disease*. Am. Fam. Physician. 1998;57(2):315-322.

.