MODELING OF AIDS INCIDENCE AND THE RESPONSE OF TRANSMISSION RATES TO INCREASE PREVENTION EFFORT: A CASE STUDY OF THE THAI PROVINCE OF NAKHON PATHOM

Sirilak Suwanwongse¹, Saowaluk Chasreechai¹, Yongwimon Lenbury¹ and Supat Kataunyuthita²

¹Department of Mathematics, Faculty of Science, Mahidol University, Bangkok; ²Office of Provincial Chief Medical Officer, Ministry of Public Health, Nakhon Pathom, Thailand

Abstract. AIDS is a serious public health problem. Our projections of the likely incidence of AIDS are of vital importance to the assessment of future healthcare needs. This paper considers an epidemic model of the population dynamics of AIDS, which has been adjusted to take into account the changes in the transmission rate in response to changes in risk behaviors and increased AIDS awareness due to public health policy, AIDS campaigns, and other means of disease prevention. The model, adjusted for reporting delays and for the variable incubation period of the disease, has been applied to AIDS incidence data gathered in Nakhon Pathom, Thailand. Using the least-squares criterion, we solved for the appropriate values of the parameters which gave the best fit of the model to the observation data. The model was found to be capable of generating short-term projections, and offers an explanation for the decline in the number of cases that is evident in more recent data.

INTRODUCTION

According to the report on the status and trends of the world's HIV/AIDS epidemics by the MAP (Monitoring the AIDS Pandemic) network (Lamtey and Tarantola, 1998), HIV/ AIDS epidemics continue to spread unevenly. Whereas in most developed countries, the number of annual AIDS cases continues to decline, the status of HIV epidemics in most other areas of the world remains uncertain because of inadequate data on the prevalence of HIV-risk behaviors (Lamtey and Tarantola, 1998).

Since the HIV pandemic is a comparatively recent phenomenon in many Asia-Pacific countries, the care and support of people with HIVrelated illness is noticeably lacking. Apart from Australia and Japan, Thailand is one of the few Asia-Pacific countries that have responded relatively early to the need for a healthcare workforce that is adequately prepared to care for the multitude of patients with HIV-related illness (Lamtey and Tarantola, 1998).

In Thailand, Nakhon Pathom had the third highest provincial rate of increase of AIDS infection by the year 1995, according to an official report of the Department of Epidemiology, Ministry of Public Health (1995). Only Bangkok and Chiang Rai had higher rates. Nakhon Pathom has a population of 734,723 and covers an area of 2168.33 km², while the average annual income per head is approximately US\$ 2,000. Nakhon Pathom is considered an unlikely candidate to have been ranked third on the list of provinces with the highest rate of increase of reported AIDS incidences in the country because its average annual income is higher while its population density is lower than many other provinces.

In this paper, we consider an epidemic

Correspondence: Yongwimon Lenbury, Department of Mathematics, Faculty of Science, Mahidol University, Rama 6 Road, Bangkok 10400, Thailand. Fax: ++66 (0) 2644-5419 E-mail: scylb@mahidol.ac.th

model of the population dynamics of AIDS that was proposed by Pickering *et al* (1986); we have modified the model in order to take into account the effect of changes in sexual behavior and increased disease awareness, which might be attributable to the increased government spending on prevention efforts, AIDS campaigns, and intervention programs.

Studying the data of the reported cases of AIDS that were diagnosed in Nakhon Pathom during the 1989-2000 period, we saw a decline in HIV-positive cases in the more recent reports (1999 and 2000): this could be attributed to reporting delays (Amaral *et al*, 2000) or the time lag from diagnosis to reporting to the Nakhon Pathom Office of the Provincial Chief Medical Officer, which could be up to 4 months or more. These time lags are often seen and cause considerable difficulty for modelers of epidemics (Pickering *et al*, 1988). The decline seen during recent months could be artificial or the result of increased prevention efforts.

We have therefore investigated such possibilities by incorporating both of the above factors in our model, by using time delays and allowing the rate of transmission to vary in response to government spending on AIDSrelated campaigns, which is in turn assumed to be influenced by the number of AIDS cases that were diagnosed in the past few years. The resulting model can explain, to a certain extent, how a combination of both factors may contribute to the leveling off followed by the drop that is observed in the data of AIDS incidence in Nakhon Pathom.

Modeling the incidence of AIDS

Cases of AIDS diagnosed in Thailand were reported on a voluntary basis to the Ministry of Public Health until 1989, when the reporting of the incidence of AIDS was made compulsory for all healthcare units. Data about HIV/AIDS patients are now submitted on a standard report form that is issued by the Ministry of Public Health. By the end of 1996, the estimated incidence of all cases of AIDS in Thailand had risen to 102,000, ranking Thailand thirteenth in the world (Lamtey and

Tarantola, 1998).

We note that while considering the cases by the onset of symptoms might be preferable, we feel that it would be problematic, partly because of the many and varied manifestations of the syndrome. We therefore chose to work with the data on the number of cases detected with the virus, including both those with confirmed seropositivity and those with confirmed viremia.

Table 1 shows the observation data on the number of cases diagnosed with AIDS and the total number of detected cases within each sixmonth period between the years 1989-2000; the data were provided by the Nakhon Pathom Office of the Provincial Chief Medical Officer. From this data, we found that from January 1989 to June 2000, as many as 2,483 males and 708 females were diagnosed with AIDS.

By studying the frequency distributions of reported incidences of AIDS among males, females, and the total number of cases every 6 months (Fig 1), where the observation data are fitted by third order polynomials, we find

Table 1					
Observation data on the number of AIDS					
cases diagnosed in Nakhon Pathom,					
counted on a biannual basis during					
1989-2000.					

Month	June		December		Total
Year	Males	Females	Males	Females	
1989	2	0	1	0	3
1990	2	1	2	0	5
1991	4	0	1	0	5
1992	3	0	18	2	23
1993	20	4	40	6	70
1994	64	3	61	5	133
1995	76	19	101	22	218
1996	176	14	237	50	477
1997	261	69	206	69	605
1998	286	93	214	73	666
1999	224	92	253	84	653
2000	231	102	165	86	584
Total					3,442



- Fig 1–Frequency distribution of males, females, and the total number of cases diagnosed with AIDS, December 1992-June 2000. Observation data is plotted every 6 months and fitted with third-order polynomials, showing similar distribution patterns among all three data sets. The observation data for males is indicated by ●, that for females by ■, and that for the total number of diagnosed cases by ▲.
- estimating curve for males.
- ----- estimating curve for females.
- ----- estimating curve for total number of cases.

that the distributions are quite similar. This is in line with the hypothesis that the transmission of AIDS in Thailand is primarily among heterosexuals, while up until 1998 the disease had been predominantly homosexually transmitted in many other countries in the world (Lamtey and Tarantola, 1998). Moreover, the time lag between diagnosis and reporting which has been noted earlier makes it more reasonable to consider the yearly reported data instead of the biannual ones. We shall therefore derive a model for D_t, the total number of AIDS cases diagnosed in the year t, expecting that the distribution of D_t is similarly reflected by males and females.

In the past, there have been essentially three different approaches to the study of HIV epidemics: the deterministic modeling approach, the stochastic modeling approach, and the statistical modeling approach (Tan and Ye, 2000). The deterministic and stochastic approaches construct the probability of infection by taking into account the dynamics and epidemiology of an HIV epidemic. On the other hand, the statistical modeling approach attempts to estimate the probability of infection from data without considering the dynamics and epidemiology of the disease. Also, the probabilities are stochastic variables in the stochastic approach, whereas they are deterministic functions of time t in the deterministic and statistical approaches (Tan and Ye, 2000).

We have adopted the deterministic approach and assumed that the rate of transmission, in the dynamic model, is dependent on the contact rate between infectious and uninfected individuals (Pickering *et al*, 1986). We have also assumed that transmission across province line is relatively small. Thus, the model may be written as follows:

$$D_{t} = k_{1}R_{t}\left(\sum_{i=t-\delta+\sigma}^{t-\delta+\mu}D_{i}\right)\left(1-k_{2}\sum_{i=1}^{t-1}D_{i}\right)$$
(1)

where the term in the first parentheses represents the infectious individuals who are assumed to transmit the disease agent starting δ - μ years after exposure until σ years before diagnosis, as proposed by Pickering *et al* (1988). Thus, the parameter δ is the length of time from exposure to diagnosis with AIDS.

The term in the second parentheses represents the susceptible individuals with

$$\mathbf{k}_2 = \frac{1}{\mathbf{N} (1-\mathbf{v})}$$

where N is the population size and v is the proportion of exposed individuals that will never develop the severest form of the disease and be diagnosed with AIDS. Thus k_2 is the inverse of the portion N (1-v) of the population that will develop the symptoms and be diagnosed with the disease if exposed. That is, we normalized the number of susceptible individuals

$$N(1-\nu) - \sum_{i=1}^{t-1} D_i = N(1-\nu) \left(1 - \frac{1}{N(1-\nu)} \sum_{i=1}^{t-1} D_i \right)$$

giving rise to the term in the second parentheses of equation (1), where the factor N (1- ν) was then absorbed into the coefficient k_1 and the number $\frac{1}{N(1-\nu)}$ was set equal to k_2 .

The term R_{t} is assumed to take the form

$$R_t = \frac{k_3}{a + bS_{t-\alpha}}$$
(2)

which accounts for the decline in the transmission rate due to increased awareness and changes in risk behaviors resulting from prevention efforts and campaigns budgeted for by government's responding to the increase in the observed cases diagnosed with AIDS each year. The term $S_{t-\alpha}$ represents government spending in the year t- α which takes α years to have an effect on the number of cases detected D_t in year t. Thus, α is the average amount of time required for the prevention efforts to bear fruit, taking into account the incubation period before the infected individuals are diagnosed or show symptoms and seek medical advice.

We, in turn, assumed that government spending on AIDS programs depends on the number of cases detected ρ years earlier, the time lag being the result of late reporting, late reaction and budget planning in response to the increase in AIDS cases. Thus, the term S_t is taken to have the form

$$S_t = k_4 + k_5 D_{t-\rho}^r \tag{3}$$

Substituting (3) into (2), and letting $\omega = \alpha + \rho$, $\kappa_1 = \frac{k_1 k_3}{a + b k_4}$, $\kappa_2 = k_2$, and $\kappa_3 = \frac{b k_5}{a + b k_4}$,

the model in equation (1) becomes

$$D_{t} = \frac{\kappa_{1}}{1 + \kappa_{3} D_{t-\omega}^{r}} \left(\sum_{i=t-\delta+\sigma}^{t-\delta+\mu} D_{i} \right) \left(1 - \kappa_{2} \sum_{i=1}^{t-1} D_{i} \right)$$
(4)

The time lag ω , therefore, accounts for the reporting delays, the amount of time required for the government to react, as well as the time required before the influence of the campaigning efforts and interventions is felt and reflected in the reduction of the transmission rate.

Parameter estimation

Utilizing the data on the total number of reported cases diagnosed each year during the years 1990-1999, the appropriate values of the parameters κ_1 , κ_2 , κ_3 , δ , σ , μ , ω , and ν were found with the lowest sum-of-squares (ss).

Here, N is equal to 734,723, the total population of Nakhon Pathom in December, 1999. We discovered that the model was not very sensitive to the changes in the value of κ_2 , as we obtained best fitting curves for different v which were indistinguishable. We therefore set v = 0.64, so that κ_2 = 0.0000038, following the guideline adopted by Pickering *et al* (1986).

In Fig 2, the outputs of the model in (4) using different r are plotted against observation data. Setting $\delta = 5$, $\delta - \mu = 1$, $\sigma = 0$, and $\omega = 5$, we numerically solved for κ_1 and κ_3 which give the lowest sum-of-squares for each value of r. We found that the lowest sum-of-squares ss = 43,838 when r = 2, ss = 9,529 when r = 1, ss = 2,310 when r = $\frac{2}{3}$, ss = 2,833 when $\frac{1}{2}$, and ss = 23,596 when r = $\frac{1}{3}$. Thus, for this set of parameters, r = $\frac{2}{3}$, κ_1 = 1.43696, and κ_3 = 0.13637 appear to give the best fitting model.

Fig 3 shows the output of the model using $r = \frac{2}{3}$, $\delta - \mu = 1$, $\sigma = 0$ and $\omega = \delta$ for different values of δ . Again, we numerically solved for κ_1 and κ_3 which give the lowest sum-of-squares for each value of δ . We found that ss = 18,998 when $\delta = 3$, ss = 16,786 when $\delta = 4$, ss = 2,310 when $\delta = 5$, and ss = 31,108 when $\delta = 6$. This seems to indicate that the most appropriate value of δ is 5, in which case $\kappa_1 = 1.43696$ and $\kappa_1 = 0.13637$.

Fig 4 shows the output of the model using $r = \frac{2}{3}$, $\delta = 5$, $\delta - \mu = 1$, and $\omega = 5$ for different values of σ between 0 and 4. We discovered that the lowest sum-of-squares occurred when



Fig 2–Output of the model of AIDS incidence given in equation (4) for different values of *r*. Here, $\delta = 5$, $\delta - \mu = 1$, $\sigma = 0$, and $\kappa_2 = 0.0000038$. The observation data is indicated by \bullet .

 $\sigma = 0$, in which case κ_1 and κ_3 are as given above. We observed in this figure that while the curve with $\sigma = 0$ gives the best fit to the 10 data points which we used during the years 1990-1999, the predicted value for the year 2000 (which was not used in the curve fitting) is not as close to the observed data as that predicted by the model with $\sigma = 2$. This may have been due to late reporting, which resulted in the number of cases reported by December 2000 being less than the actual number of diagnosed cases. This led us to believe that the curve with $\sigma = 0$ may have been the better fit and the predicted value for the year 2000 may have been more accurate than that predicted by the model with $\sigma = 2$.

Now, $\delta = 5$ suggests that it takes on the average 5 years from exposure to diagnosis in the case of Nakhon Pathom, and infectious individuals appear to transmit the disease agent starting 1 year after exposure ($\delta - \mu = 1$) until right before diagnosis ($\sigma = 0$), a situation which



- Fig 3–Output of the model of AIDS incidence given in equation (4) for different values of δ . Here, $r = \frac{2}{3}$, $\delta \mu = 1$, $\sigma = 0$, and $\kappa_2 = 0.0000038$. The observation data is indicated by \bullet .
- $\begin{array}{c} \square \square \square \square : \delta = 3, \kappa_1 = 1.913706, \kappa_3 = 0.067300, \text{ss} = 18,998\\ \textcircled{math\square} \textcircled{math\square} \textcircled{math\square} \textcircled{math\square} \sub{math\square} \emph{math\square} \emph{mat$



Fig 4–Output of the model of AIDS incidence given in equation (4) for different values of σ . Here, $r = \frac{2}{3}$, $\delta = 5$, $\delta - \mu = 1$, $\sigma = 0$, and $\kappa_2 = 0.0000038$. The observation data is indicated by \bullet .

 is to be expected.

Modeling the impact of AIDS incidence on government spending

We found that $r = \frac{2}{3}$, $\kappa_3 = 0.13637$, and $\omega = 5$ gave the best fitting model, where $\delta = 5$, $\delta - \mu = 1$, $\sigma = 0$, $\kappa_1 = 1.43696$ and $\kappa_2 = 0.0000038$. We then used the criterion of the least sum-of-squares to numerically solve for the parameters k_4 , k_5 , and ρ in the following model for the government spending S_t which is allocated for AIDS prevention efforts and campaigning,

$$S_{t} = k_{4} + k_{5} D_{t-\rho}^{2/3}$$
(5)

in order to determine how budget allocations respond to the reports of AIDS incidence $D_{t-\rho}$ with a time lag of ρ .

The data on government budgets for prevention programs and AIDS campaigns during the years 1998-2001 were provided by the Nakhon Pathom Office of the Provincial Chief Medical Officer. Prior to 1998, accurate data was not available since there was no systematic accounting for budget allocations. Increased awareness has since resulted in more careful collection of data, categorizing spending into different headings, according to the relevant programs.

After setting a value for ρ between 1 and 5, we numerically solved for k_4 and k_5 which give the lowest sum-of-squares. In Fig 5, S_t is plotted against $D_{t,\rho}$, and the estimating curves with $\rho = 2$, 3, 4, and 5 are shown against the observation data in Figs 5a - 5d respectively.

Table 2 presents the lowest sum-of-squares values for the model given by (5), and the corresponding coefficients of determination for different ρ . We found that ss was lowest when $\rho = 3$, for which we found that $k_4 = 15,880.78$ and $k_5 = 862.94$. This value of ρ suggests that it takes about 3 years for the government to respond to the increase in the reported AIDS cases diagnosed each year, to plan to increase its spending, and eventually obtain an increased annual budget in the cause of AIDS.

Now, $\omega = 5$ in the best fitting model given by (4). This could be interpreted as saying that, following an increase in government spending



Fig 5–Output of the model for allocated annual budget for AIDS prevention campaigns in response to the reported AIDS incidence for different values of ρ . The observation data is indicated by \bullet . In (a), $\rho = 2$; (b), $\rho = 3$; (c), $\rho = 4$; (d), $\rho = 5$.

Table 2
The lowest-sum-squares values of the
model in equation (5) for different values of
ρ between 1 and 5, and the corresponding
coefficients of determination R^2 , k_4 , and k_5 .

ρ	ss x 10 ⁻⁸	\mathbb{R}^2	k_4	k ₅
1	7.26	0.05518	-22,137.92	1,241.08
2	0.61	0.92054	-88,045.37	2,207.51
3	0.56	0.93712	15,880.78	862.94
4	2.99	0.63185	39,938.98	591.42
5	3.52	0.58216	46,994.84	620.20

on AIDS prevention, 2 years are required before the resultant changes in high-risk behavior produce a decrease in the number of new infections, which are detected some time after the exposure to the virus.

DISCUSSION

The primary assumptions of this model are that AIDS can be modeled as an infectious disease and the transmission rates of the agents responsible for AIDS respond to changes in sexual behavior or other high-risk behaviors, such as needle and syringe sharing among drug users.

Although Lamtey and Tarantola (1998) reported that the understanding of HIV epidemics and their determinants in the Asia-Pacific region has improved substantially, as witnessed by the implementation of comprehensive surveillance systems for HIV prevalence as well as by assessments of sexual and other risk behaviors in many countries, a number of countries in the region remain unable to assess the incidence of HIV infection and related activities and monitor the impact of interventions (Lamtey and Tarantola, 1998).

One of the important factors that prevents a timely response to the need for immediate action has been the economic crisis experienced in the Asia-Pacific region in recent years. This has resulted in delays and cut-backs in government spending. The impact of such actions can be projected by the model considered in this paper with the use of different values of r, ρ , and ω . As can be seen in Fig 2, when r is given lower value (representing a budget cut-back), a rise in AIDS cases is produced. More specifically, a sharper drop in the detected cases D, can be effected if the allocated budget responds as the square (r = 2 in Fig)2) of the number of cases reported 3 years earlier, given that the delay from diagnosis to actual provision of funding is fixed. Furthermore, the model with a smaller time lag δ , corresponding to a quicker response to the observed incidence of AIDS, projects fewer cases in the year 2001 than those projected by the model with a bigger δ , which corresponds to a slower response. This is shown in Fig 3, where the curves are projected until the year 2001.

On a cautionary note, our model assumes that cross-provincial infection is small, when in actual fact infected individuals from outside the province can easily come in contact with a member of the province who may then become infected. However, since we have found from our regression analysis, that the model is relatively insensitive to the value of k,, we feel that our restriction is still within the bounds of reasonable assumption. Moreover, our model discounts the effect of nationwide intervention schemes, which may also have increased awareness and changed risk behaviors among the population of Nakhon Pathom. However, to take into account such effects, more reliable and comprehensive data is needed, which is not readily obtainable at this time. This stresses the need for greater efforts in accurate data collection and analysis.

Although there are such questions concerning the reliability of the data, partly due to underreporting because not all AIDS cases are recognized as such, the model discussed in this paper still provides a useful tool for the projection of AIDS incidence in response to interventions that should be an integral part of policy development. The model is a deterministic one, which is based upon the understanding of the dynamics and epidemiology, rather than relying on mathematical functions that fit existing data. It should therefore be capable of forcasting the incidence of the disease over an extended period of time, if the model is progressively updated with additional data when they become available.

Earlier studies involving epidemic models have concluded that more intensive studies are needed to deal with the concerns about the reliability and validity of the data gathered in the past. (Anderson *et al*, 1987; May and Anderson, 1987; Bailey, 1997; Griffiths *et al*, 2000; Lenbury *et al*, 2000). The model could be improved in several ways: we could study the reporting delay distribution in the region, as suggested by Amaral *et al* (2000), and then incorporate a variable time lag instead of the constant averaged time lag that we used in this paper.

For countries to have timely and effective responses, high-level political commitment is crucial to containing or avoiding a serious HIV problem. Unfortunately, as stated in the report on the world's HIV/AIDS epidemics (Lamtey and Tarantola, 1998), policy-makers and the public as a whole often operate on false assumption about sexual and other risk behaviors in society. This in parts results from a reluctance to discuss sexual and drug issues, due to either faith, social, or sexual biases. Lack of data to contradict false assumptions is also an important contributing factor to the neglect of prevention needs.

On the basis of this observation, it is essential that the necessary behavioral and epidemiological data is collected, analyzed, understood, and clearly presented to both the policy-makers and the public. Specific studies to demonstrate the effectiveness of prevention programs on behavioral changes that leads to case reduction can mobilize increased political will and commitment to HIV prevention. Above all, such epidemiological models provide a conceptual framework into which accumulating data can be integrated, leading to both a better understanding of the transmission dynamics of the infection, and projections for the future.

ACKNOWLEDGEMENTS

This research project was supported by a grant from the Thailand Research Fund (Grant Contract Number RTA/02/2542).

REFERENCES

- Amaral JA, Rosário MB, Paixão MT. Data and projections of HIV and AIDS in Portugal. J Appl Stat 2000; 27: 269-79.
- Anderson RM, Medley GF, Blythe SP, Johnson AM. Is it possible to predict the minimum size of the acquired immunodeficiency syndrome (AIDS) epidemic in the United Kingdom? *Lancet* 1987; 1:1073-5.
- Bailey NTJ. A revised assessment of the HIV/AIDS incubation period assuming a very short early period of high infectivity and using only San Francisco public health data on prevalence and incidence. *Stat Med* 1997; 16: 2447-58.
- Department of Epidemiology, Ministry of Public Health, Thailand. Official report released annually: AIDS epidemics in each of the provinces in Thailand, 1995.
- Griffiths J, Lowrie D, Williams J. An age-structured model for the AIDS epidemic. *Eur J Operat Res* 2000; 124: 1-14.
- Lamtey P, Tarantola D. The status and trends of the HIV/AIDS epidemics in the World: final report 1998. Boston, USA: World AIDS Conference on the Status and Trends of HIV/AIDS Epidemics in the World; June 28-July 3, 1998.
- Lenbury Y, Ouncharoen R, Tumrasvin N. Higher dimensional separation principle for the analysis of relaxation oscillations in nonlinear systems: application to a model of HIV infection. *IMA J Math Appl Med Biol* 2000; 17: 243-61.
- May RM, Anderson RM. Transmission dynamics of HIV infection. *Nature* 1987; 326: 137-42.
- Pickering J, Wiley JA, Padian NS, Lieb LE, Echenberg DF, Walker J. Modelling the incidence of acquired immunodeficiency syndrome (AIDS) in San Francisco, Los Angeles, and New York. *Math Model* 1986; 7: 661-88.
- Pickering J, Wiley JA, Lieb LE, Walke, J, Rutherford GW. Modelling the incidence of AIDS in New York, Los Angeles, and San Francisco. Statistical analysis and mathematical modelling of AIDS. Oxford: Oxford Press, 1988.
- Tan WY, Ye Z. Estimation of HIV infection and incubation via state space models. *Math Biosci* 2000; 167: 31-50.