ORIGINAL ARTICLE

Backcalculation of HIV Infection Rates in Malaysia

H C Ong, PhD, K L Soo, B.Sc.

School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Minden, Penang

Summary

It has been almost two decades ago since the first AIDS case was reported in Malaysia. It has also been approximately eight years ago when the method of backcalculation was used to estimate the past HIV infection rate from the AIDS incidence data and an estimate of the incubation period distribution. This method is used because it makes use of the Malaysian AIDS incidence which is fairly reliable and reflects the trend of the epidemic as compared to the HIV infection rate recorded. The latest results generated show a slowdown in the increase of the number of estimated infected HIV+ cases in the late 1990s and this trend is supported by a slowdown in the increase of the number of AIDS cases recorded.

Key Words: Backcalculation, AIDS modelling, HIV infection

Introduction

Methods for predicting future trends in the incidence of AIDS are based on one of three methods. These methods are similar in that they all fit some function of calendar time to the incidence of AIDS data but they differ in the degree to which the mechanisms that generate the data are incorporated into the model.

At one extreme, the method fits a function of calendar time such as a polynomial or other mathematically convenient curves to the AIDS incidence curve and then extrapolates into the future (Healy and Tillet[1]). Although this method is easily implemented, it has the danger of extrapolating a fitted model outside the range of the observed data. The prevalence of HIV infection cannot be estimated with this approach. This method ignores what is known about the epidemiology of the disease and it cannot incorporate information that one might have on changing patterns of transmission.

The second method models the full dynamics of transmission of the epidemic in the population providing much insight into the qualitative evolution of the epidemic and identifying the key variables that determine the short and medium term forecast of the number of cases (Arca, Perucci and Spadea[2]). Unfortunately, the deterministic models proposed for general epidemics are complicated, and the stochastic models are even more complicated.

Furthermore, predictions based on such models are particularly sensitive to unknown parameters such as the long incubation distribution from infection to the development of AIDS, the frequency and pattern of sexual activity and behavioural changes which change with time and also the proportion of infected people who eventually develop AIDS with allowance for emigration, immigration and death. Such models are complicated and contain many unknown parameters (Heisterkamp *et al*[3]). Moreover, the parameters involved are not easily accessible basically due to the sensitivities involved with the epidemic and the social stigma associated with the disease.

The third approach, to be used in this study, which is intermediate between these two approaches, is the back projection method. (Brookmeyer[4]; Bacchetti, Segal and Jewell[5]; Day, Gore and De Angelis[6]). Backcalculation is a method of estimating past HIV

This article was accepted: 25 November 2006
Corresponding Author: Ong Hong Choon, School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Pulau Pinang

infection rates from the AIDS incidence data, and an estimate of the incubation period distribution. The method requires reliable counts of the number of AIDS cases diagnosed over time and a reliable estimate of the incubation period distribution. The method is popular because it makes use of the AIDS incidence data which represent the most readily available information on the AIDS epidemic as most national AIDS surveillance data systems record only AIDS cases. The incubation period distribution can then be applied to the estimated past HIV infection rates to project future AIDS incidence.

The main objective in this paper is to determine the underlying trend in the HIV infection rates in Malaysia as compared to the number of HIV positive cases recorded. The method of backcalculation is used because it makes use of the AIDS incidence data which reflects better the trend of the epidemic. The number of HIV+ cases, on the other hand, is dependent on the test made and is unreliable as a trend. For example, a steep rise in the number of HIV+ cases may be due to the mandatory testing of all intravenous drug users in drug rehabilitation centres and increase in detection through aggressive case finding. The backcalculation method provides a simple conceptual framework for relating the incubation distribution with the AIDS incidence data and the infection rate. The weaknesses and advantages of using the backcalculation method is discussed in Ong, Quah and Low[7].

Materials and Methods

The basic convolution equation in backcalculation relates the number of new cases of AIDS in time t to $t + dt$ (designated $Z(t)$)and the number of new HIV infections at each time s since the start of the epidemic (s=0) through the incubation period distribution $f(u)$, where u is the time spent between the initial infection and the eventual diagnosis of AIDS. The basic convolution equation is given as:

$$Z(t) = \int_0^t g(s)f(t-s)\, ds. \qquad (1)$$

From the above equation, for an individual to be diagnosed as an AIDS case by calendar time t, he or she must have been infected at some prior time s, and then have an incubation period less than $t-s$. In other words, the backcalculation method uses the above equation together with knowledge of $Z(t)$ (obtained from the AIDS cases registries) and $f(t)$ (obtained from the epidemiological studies) to give information on past infection rates $g(s)$. If $f(t)$ is known, the above relationship could be inverted to express $g(s)$ for all $0 \le s \le t$ as a function of $Z(t)$. In general, a family of values for $g(s)$, $0 \le s \le t$ can be constructed which are consistent with a realization of $Z(s)$, $0 \le s \le t$.

Let $z_1, z_2, \ldots, z_n$ be the number of AIDS cases diagnosed in the calendar time interval $[t_{i-1}, t_i)$, $i = 1, 2, \ldots, n$. It is assumed that individuals become infected according to a point process. Then the expected number of AIDS cases occurring during the time interval $[t_{i-1}, t_i)$ is given by (Brookmeyer and Gail[8])

$$E(Z_i) = \int_0^{t_i} g(s).\{F(t_i-s) - F(t_{i-1}-s\}\, ds. \qquad (2)$$

where $F(t)$ is defined to be 0 for $t \le 0$. By convention, we shall define calendar time 0 to be the start of the epidemic (that is $g(s) = 0$ prior to that time) and thus $t_0 = 0$.

Backcalculation essentially involves the deconvolution of equation (1) or (2) above to get the infection rate. The deconvolution involves inverting the relationship in equation (1) or (2) to express the infection rate g as a function of Z using an estimate of the incubation period distribution f. The equation is deconvoluted by assuming a Fourier transform of $Z(t)$ of the form. $\tilde{Z}(s) = \int_{-\infty} e^{ist} Z(t)dt$. Similarly $\tilde{g}(s)$ and $\tilde{f}(s)$ are the Fourier transforms of $g(s)$ and $f(s)$. It then follows from equation (1) that $\tilde{Z}(s) = \tilde{g}(s)\tilde{f}(s)$ and, therefore, $g(t)$ can be obtained by taking the inverse transform of $\tilde{Z}(s)/\tilde{f}(s)$.

Results

The definitive diagnostic method for diseases indicative of AIDS used in Malaysia is taken from the Ministry of Health's publication *"Plan of Action for Prevention and Control of AIDS"* released in May 1988. This definition is similar to the CDC's 1987 definition and the WHO definition and has been in used in Malaysia since 1988.

Ong, Quah and Low[7] applied the backcalculation method on the Malaysian data until August 1996 and obtained the estimated number of cumulative HIV infection as shown in Figure 1.

An obvious result from Figure 1 is that the HIV/AIDS epidemic in Malaysia is in their early stage which is evident in its rapid exponential increase in the number of infected cases until August 1996. A backcalculation program in Fortran from Bacchetti, Segal and Jewell[5] is used again on the Malaysian data in this paper. The program is based on the incubation period distribution from Brookmeyer[4]. The basis for using the incubation period distribution from Brookmeyer[4], which is based mainly on homosexuals, on the Malaysian data which are mainly intravenous drug users is because the incubation period distribution of the two cohorts are similar (Mariotto *et al.*[9]).

The two curves in Figure 2 are obviously different as the recorded number of cumulative HIV+ cases in Malaysia does not reflect the true picture of the epidemic. The number of AIDS cases projected backward shows an exponential rise in the number of HIV+ cases in the early 1990s and a sub-exponential growth in the late 1990s. The almost plateau or slow growth in the estimated cumulative HIV+ cases in the early 2000s is reflected by a slowdown in the reported number of AIDS cases which could be due to delayed reporting or better drug treatment. One of the ways of dealing with the problem of delay reporting is by overlooking or truncating the most recently reported data as was done by Brookmeyer and Gail[7].

Discussion

There is a slowdown in the increase of the number of estimated infected HIV+ cases in the late 1990s as can be seen from Figure 2. This trend is supported by a slowdown in the increase of the number of AIDS cases (Table I) which gives a clear picture of the trend of the epidemic. Also, there is a narrowing of the gap between the number of recorded cumulative and estimated HIV+ cases. This is probably due to an increasing awareness among the population (especially the high risk group) towards the epidemic and also due the efforts by the Malaysian government and non-government organizations to promote this awareness. This slowdown in the increase is also similar to the trend in developed countries like the US and UK where the AIDS epidemic had begun earlier. (Bacchetti, Segal and Jewell[5]; Chariotti *et al.*[11])

The introduction of life-prolonging drugs, used in the Highly Active Antiretroviral Therapy (HAART) will definitely increase the incubation period distribution f in equation (1) but it was only in recent years that it has been made available to the general population. As such it does not affect our result that there was a sub-exponential growth in the AIDS epidemic in the late 1990s.

Table I: Cumulative HIV+ and AIDS cases recorded from the Ministry of Health, Malaysia (http://dph.gov.my/aids/[10]) and the corresponding estimated cumulative HIV+ cases from the backcalculation method in this study.

Year	Cumulative HIV+ cases recorded	AIDS Cases reported	Estimated cumulative HIV+ Cases
1986	3	1	1.62
1987	5	0	402.26
1988	14	2	1036.68
1989	214	2	1828.84
1990	992	18	2897.21
1991	2786	60	4502.37
1992	5298	73	7208.14
1993	7805	71	12064.01
1994	11198	105	20214.95
1995	15396	233	30897.54
1996	19993	347	40536.07
1997	23917	568	46741.35
1998	28541	875	50055.38
1999	33233	1200	51718.45
2000	38340	1168	52542.59
2001	44278	1302	52950.32
2002	51256	1193	53152.01
2003	58012	1076	61839.43
2004 (until September)	61486	661	61854.00

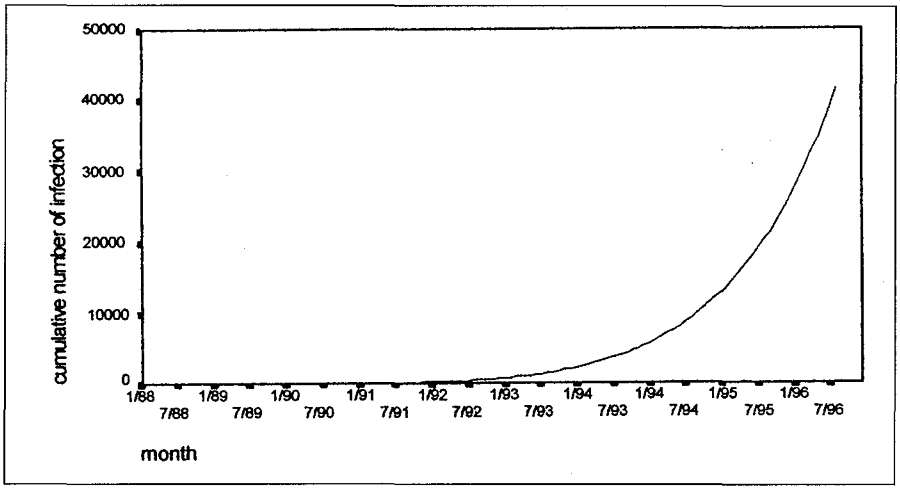

Fig 1: Estimated cumulative number of HIV infection applied to the Malaysian data till August 1996 (Ong, Quah and Low[7])

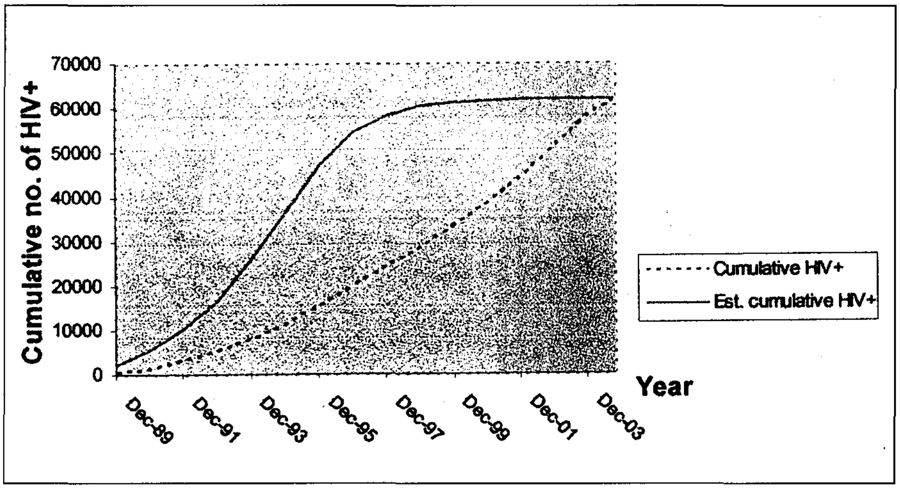

Fig 2: A comparison between the recorded and estimated cumulative number of HIV infections in Malaysia

References

1. Healy MJR, Tillett HE. Short-term extrapolation of the AIDS epidemic. Journal of the Royal Statistical Society A 1988; 151: 50-61.
2. Arca M, Perucci CA, Spadea T. The epidemic dynamics of HIV-1 in Italy: Modelling the interaction between intravenous drug users and heterosexual population. Statistics in Medicine 1992; 11: 1657-84.
3. Heisterkamp SH, De Haan BJ, Jager JC, Van Druten JAM, Hendricks JCM. Short and medium term projections of the AIDS/HIV epidemic by dynamic model with an application to the risk group of homo/bisexual men in Amsterdam. Statistics in Medicine 1992; 11: 1425-41.
4. Brookmeyer R. Reconstruction and future trends of the AIDS epidemic in the United States. Science 1991; 253: 37-42.
5. Bacchetti P, Segal MR, Jewell NP. Backcalculation of HIV infection rates. Statistical Science 1993; 8: 82-119.
6. Day NE, Gore SM, De Angelis D. Acquired immune deficiency syndrome predictions for England and Wales (1992-1997): Sensitivity analysis, information, decision. Journal of Royal Statistical Society A 1995; 158: 505-24.
7. Ong HC, Quah SH, Low HC. An application of the backcalculation method to estimate past HIV infection rates in Malaysia. The Medical Journal of Malaysia 1998; 53: 385-91.
8. Brookmeyer R, Gail MH. AIDS Epidemiology: A Quantitative Approach. New York: Oxford University Press 1994.
9. Mariotto AB, Mariotti S, Pezzotti P, Rezza G, Verdecchia A. Estimation of the acquired immunodeficiency syndrome incubation period in intravenous drug users: A comparison with male homosexuals. American Journal of Epidemiology 1992; 135: 428-37.
10. Ministry of Health, Malaysia (2005) [Online] [Accessed 20th December 2005] http://dph.gov.my/aids/.
11. Chiarotti F, Palombi M, Nicola S, Ghirardini A, Prospero L. Effects of different parametric estimates of seroconversion time on analysis of progression to AIDS among Italian HIV positive haemophiliacs. Statistics in Medicine 1992; 11: 591-601.