

Statistical Modelling of HIV/AIDS in Nepal: A Necessary Enquiry

Sathian B¹

¹Assistant Professor, Department of Community Medicine, Manipal College of Medical Sciences, Pokhara, Nepal

Editorial

Corresponding Author:

Dr. Brijesh Sathian, Assistant Professor, Department of Community Medicine, Manipal College of Medical Sciences, Pokhara, Nepal.

Email: drsathian@gmail.com

The current situation of HIV in Nepal is different from when the first case was diagnosed in 1988. The current number of positive cases reported till 2009 are 14787, out of which 13005 are receiving HIV care^{1,2}. There are gaps and challenges to be addressed in the fight against HIV and AIDS.

HIV/AIDS surveillance methods evolve over time, so data from the same source may not be directly comparable year to year. The type of data available and the lag-time in availability, may pose challenges to assessing recent impact. Epidemiological measures of HIV/AIDS are numerous and each has important and distinct definitions. Much of the data are estimates only. This is true globally and in all countries, even the United States, due to the lag-time between HIV infection and the development of AIDS, the fact that many do not know their status, stigma which leads to under reporting, and surveillance systems that may not be complete. Attention should be given to ranges given around any estimate, as well as any notes that may accompany data, since these may provide important information that can help in your interpretation. Rates/percents, not just numbers, are important.

In 2001, the UNAIDS/WHO introduced a new version of its epidemiological model for all developing countries. The model known as Estimation and Projection Package (EPP) makes use of the surveillance data from each HIV sentinel site in estimating HIV/AIDS prevalence rates. The Spectrum Program also utilizes the information on the birth rate,

death rate and the output of EPP for the country in calculating the estimated number of people living with HIV, number of new infections, number of AIDS cases, number of AIDS deaths, number of orphans, etc. The extent to which this data source is representative of the entire adult population of Nepal is doubtful. This is because not all pregnant women attended antenatal clinics and not all women of adult age were pregnant at the time of the surveys. The selection procedure of the survey sites also systematically excludes private clinics where many births occur. The Spectrum Program of UNAIDS and WHO makes use of vital rates which were obtained from a poor vital registration system. Therefore, national estimates based on these surveys rely on a fraction of women who attended the selected antenatal clinics.

Focusing on the epidemiology of HIV/AIDS in Nepal, some studies have been published, but nothing has been done in the area of modelling the course of the infection overtime. Infectious disease data has two features that distinguish it from other data. It is highly dependent on the chosen treatment modalities and the process of infection is only partially observable³. A consequence of these features is that the analysis of data is usually most effective when it is based on a model that describes aspects of the infection process Becker and Britton (1999). Again, an understanding of the magnitude and trajectory of the HIV/AIDS epidemic, as well as the uncertainty around the parameters is critically important both for planning and evaluating control strategies and for preparing for vaccine efficacy trials⁴. Mathematical models can become very useful tools in this area. Therefore modelling is an integral part of statistical work in HIV/AIDS research.

Modelling exercises are additionally aimed at making use of the available data (no matter how little) to provide information about the trend inherent in the course of the epidemic. Since, in Nepal, data on HIV/AIDS are scanty, a better insight can be provided if analysis of the data is based on estimate of statistical models whose assumption are realistic and with parameters defined to capture the

situations peculiar to the locality. One such statistical models is the back-calculation method that was first proposed by Brookmeyer and Gail (1986) for estimating infection distribution and for providing short-term projection of future AIDS case^{5,6}.

Back-calculation is a method for estimating past infection rates from AIDS incidence data. The method has been useful for obtaining short term projections of AIDS incidence and estimating HIV prevalence. Back-calculation requires reliable counts of the numbers of AIDS cases diagnosed over time and a reliable estimate of the incubation period distribution. Early references on back-calculation are Brookmeyer and Gail^{5,7}.

The basic idea is to use AIDS incidence data together with an estimate of the incubation period distribution to reconstruct the numbers of individuals who must have been previously infected in order to give rise to the observed pattern of AIDS incidence. Then, the incubation distribution is applied to these estimated numbers of previously infected individuals to project AIDS incidence.

The fundamental relation between the expected cumulative number of AIDS cases diagnosed by calendar time t , $A(t)$, the infection rate $g(s)$ at calendar time s , and the incubation period distribution $F(t)$, is given by the convolution equation
Back Calculation formula is

$$A(t) = \int_0^t g(s) \times F(t-s) ds \rightarrow (1)$$

$A(t)$ = the cumulative HIV cases at time t

$g(S)$ = infection rate at time s

$F(t-s)$ = the weibull distribution function for the incubation of the people who lives after the time s to t .

The convolution equation (1) is justified by noting that in order for an individual to be diagnosed by calendar time t , he or she must have been infected at some prior time s and then have an incubation period less than $t - s$. Essentially, the back-calculation methodology uses equation (1) together with knowledge of $A(t)$ (obtained from registries of AIDS cases) and $F(t)$ (obtained from epidemiological studies) to glean information about previous infection rates $g(s)$.

Talking about the models for incubation time period of HIV, it is the random time between the HIV infection and the onset of clinical AIDS symptoms. The probability distribution of this non negative random variable is known as HIV incubation period distribution. Back calculation methods are very sensitive to incubation time distribution so we have to find out a best fitted model for the incubation time of HIV. Some most used models are Weibull and Gamma models, Log logistic and Log normal models, Mixture and Staging Model, Change point and Contact Models⁸⁻¹⁷.

The limitations of studies which consist only of AIDS cases, such as the transfusion study, called attention to the importance of following cohorts of infected individuals to monitor their immune systems and to determine progression rates to AIDS. One rapid and convenient study design involves assembling a cohort of individuals who are already HIV seropositive but whose dates of seroconversion are typically unknown. These are called prevalent cohort studies. A number of statistical problems arise in the interpretation and analysis of prevalent cohorts.

Given the observed limitations in the current estimation procedures in Nepal, it is better to examine other modelling approaches for HIV/AIDS epidemic (Sathian B et al, 2011) that could be more suitable and give more accurate estimates for the Nepal epidemic scenario².

Conflict of Interests

The author has no conflict of interest arising from the study.

References

1. Fact Sheet: HIV and AIDS Epidemic Update of Nepal, National Center for AIDS and STD Control, Ministry of Health and Population, November 2009.
2. Sathian B, Sreedharan J, Mittal A, Baboo NS, Chandrasekharan N, Devkota S, Abhilash ES, Rajesh E, Dixit SB. Statistical Modelling and Forecasting of Reported HIV Cases in Nepal. Nepal Journal of Epidemiology 2011;1(3): 107-11.
3. De Angelis, D., Day, N. E. and Gill, O. N. (Acquired Immune Deficiency Syndrome projections in England and Wales; interplay of Methodology and data. J. R. Statist. Soc. A, 1998; 161(2): 167 - 76.
4. Salomon J. A. and Murray J. L. Modeling HIV/AIDS epidemics in Sub-Saharan Africa using Seroprevalence data from antenatal clinics. Bulletin of WHO. 2001; 79(7): 596 – 606.
5. Brookmeyer & Gail. Minimum size of the acquired immunodeficiency syndrome (AIDS) epidemics in the United States. Lancet, 1986. 2 (8519): 1320-22.
6. Tan W.Y. Stochastic Modelling of AIDS Epidemiology and HIV Pathogenesis. World Scientific publishing Co. 2000; 199-232.
7. Brookmeyer, R and Damiano, A. Statistical methods for short-term projection of AIDS incidence. Statist. Med. 1989; 8:23-34.
8. Medley GF, Anderson RM, Cox DR and Billard L. Incubation period of AIDS in patients infected via blood transfusion. Nature 1987;328:719-21.
9. Brookmeyer R and Goedert JJ. Censoring in an epidemic with an Application to haemophilia-associated AIDS. Biometrics 1989;45:325-35.
10. Boldson JL, Jensen JL, Sogarrd J and Sorensen M. On the incubation time distribution and the Danish AIDS data. J Royal Stat Soc, series A 1988;151:42-3.

11. Lawless J and Sun J. A comprehensive backcalculation framework for the estimation and prediction of AIDS cases. In: "AIDS Epidemiology: Methodological Issues", (Eds. Jewell, N.P., Dietz, K. and Farewell, V.T.), 1992; 81-104.
12. Swaminathan S, Ramachandran R, Baskaran C et al. Risk of development of Tuberculosis in HIV infected patients. *Inter J Tuber Lung Dis* 2000; 4:832-44.
13. Venkatesan P. Methods of projection of HIV/AIDS epidemic; In *Epidemiology, Health and Population*, (Eds; Anil Kumar), 2002; 143-55.
14. Anbupalam T, Ravanan R and Venkatesan P. Backcalculation of HIV/AIDS in Tamilnadu; In *Bio statistical Aspects of Health and Epidemiology* (Eds; Pandey, C.M., Pradeep Mishra and Uttam Singh), Deptt of Biostatistics, Sanjay Gandhi Postgraduate Institute of Medical Research, Lucknow, India. 2002; 232-43.
15. Ravanan R. Statistical modeling of HIV/AIDS epidemic: A backcalculation approach. PhD thesis, University of Madras, 2004.
16. Auger I, Thomas P, Gruttola VD, Morse D, Moore D, Williams R, Truman B and Lawrence CE. Incubation periods for paediatric AIDS patients. *Nature* 1988; 336: 575-77.
17. Brookmeyer R and Liao J. Statistical modeling of the AIDS spread for forecasting health care need. *Biometrics* 1990; 46: 1151-63.

Additional Notes: Artist's depiction of HIV-1 Virus, Image Attribution:
<http://www.flickr.com/photos/phylomon/4639953725/>
Cover design and journal layout by NJE Associate Managing Editor, Dr. Nishida Chandrasekharan.