Epidemiology, Health
and
Population
(Statistical Applications)

Proceedings are sponsored by the Grant from
Council for Scientific and Industrial Research (CSIR)

Edited by
Anil Kumar

B.R. Publishing Corporation
[A Division of BRPC (India) Ltd.]
Delhi-110035

2002,143-155

Proc. of The 18th Annual Conf. of
Indian Soc. for Medical Statistics

13

Methods for Projection of HIV/AIDS Epidemic

P. Venkatesan

Summary

There has been substantial interest worldwide in understanding the current state of HIV/AIDS epidemic and prediction of its future path throughout the pandemic. The various approaches that have been developed to provide the description of HIV spread and modeling the transmission processes for short-term and long-term projections will be reviewed in this paper. Some of the recent developments in combining forecasting techniques with surveillance data-basis for more valid projections will be discussed.

1. Introduction

The extensive worldwide spread of Human Immunodeficiency Virus (HIV) started in the mid-to late 1970s. In less than two decades-—during the first of which it was unknown and unsuspected - HIV became the first modern pandemic. Today,in 2000 there are signs that the pandemic of HIV may be out of control - that is, a pandemic whose broad course through and among societies has yet to be influenced in any substantial way by policies and programmes mounted against it. The main concern is that, what we see at present is undoubtedly just the tip of the iceberg. If the transmission of HIV ceased tomorrow, there would still be considerable number of AIDS cases appearing for many years to come since the disease has a long incubation period. It is estimated that there are about 33 million virus carriers in the world at present time [15], many of these people are completely unaware that they are infected and hence may unwillingly pass the virus to the others. Further cogent reasons for the concern are that AIDS tends to attack people in the prime of life, is invariably fatal, with death often occurring in a most impleasant manner and above all the disease is preventable.

The Acquired Immune Deficiency disease (AIDS) and its related syndromes have also changed virtually every aspect of medicine and society at large. AIDS has affected basic science. clinical practice and social perspective. First and foremost, we are required to accept the reality that even at the end of 20th century, it is still possible for a previously unknown pathogen to initiate a pandemic that spared no country. Secondly, we were made to realize that the distinction, those who perform basic research and those who engage in clinical care have become somewhat blurred. AIDS challenges the usual compartmentalization of medical disciplines. To specialize in AIDS require a depth of knowledge in almost every field of social and medical sciences [1]. For those who view AIDS from a public health policy perspective, there must be a growing recognition that AIDS and poverty are strongly related; and for those with a medical orientation or clinical agendas in AIDS will have multifaceted civil implications as to the rights and expectations of individuals who may already be facing oppression from society at large.

The global pandemic of HIV infection comprises many different epidemics, each with its own dynamics, and influenced by many factors–e.g., time of introduction, population density and cultural and social issues. Spread of the epidemic had varied considerably between developed and developing countries, depending on the culture as well as other social and behavioural patterns. Incidence rates have been the highest in developing countries where hetero-sexual transmission is most common. Present work reviews the uses and limitations of various methods for projection of HIV/AIDS epidemic and the potential usefulness of some of the recent statistical models. This work also discusses the future direction of research needed for valid and reliable projections. In India, Myanmar and Thailand, the volatility of the pandemic is most striking, Because of the density of the population, there may be more HIV infected individuals in Asia by the year 2000 than in any other region of the world. With continued escalation of HIV transmission in Asia and Sub-Saharan Africa the estimate is that nearly 90% of all HIV infected people will reside in these developing countries by the year 2000 [12].

2. The Current Global Situation of the HIV/AIDS Pandemic

Though AIDS was not recognized as a new disease until the early 1980s, it has now reached every country. Worldwide the Joint United Nations Programme on HIV/AIDS [15] estimates that more than 33 inillion people are currently living with HIV/AIDS (Table 13.1). The developing world bears the brunt of the epidemic, accounting for over 90% of the global infections so far. The highest prevalence of HIV infection is found in Africa where one in three adults in some cities are HIV positive. However,the virus is now spreading explosively in Asia which already holds over 6 million adults living with HIV even though the virus gained a foothold in the region as recently as the mid-1980s. World-wide 5-6 million people are estimated to be newly infected with HIV in 1999 alone and 12.7 million adult deaths due to AIDS till recently (Table 13.2).

As on 30. June 1997, 1,644,183 cumalative AIDS cases in the adults and children have been reported since the onset of the pandemic [17]. This represents an increase of 18% from the 1,396,640 cases reported till June 30. 1996. The distribution of the reported AIDS cases by continent is provided in Table 13.1.From this, we see that about half of the AIDS cases were from America followed by about one-third of the cases from Africa. Allowing for under recognition, incomplete reporting and reporting delays, and based on the available data on HIV infections around the world, the USAIDS/WHO) working group on Global HIV/AIDS and sexually Transmitted Diseases has estimated that approximately 16.3 million AIDS deaths in the adults and children have occurred worldwide since the pandemic began, i.e., an estimated 12.7 million deaths in adults and 3.6 million deaths in children.

The number of persons living with HIV/AIDS is more indicative of the current and future course of the pandemic than the cumulative number of AIDS cases.

Table 13.1: Global Summary of HIV/AIDS Epidemic (Dec.1999)

(in Millions)

	Totol	*Adults*	*Women*	*Children*
Newly Infected in 1999	5.6	5.0	2.3	0.6
Living with HIV/AIDS	33.6	32.4	14.8	1.2
AIDS Deaths in 1999	2.6	2.1	1.1	0.5
Total no. AIDS Deaths (From the beginning of the epidemic)	16.3	12.7	6.2	3.6

Source: (WHO/UNAIDS-2000).

From Table 13.2 we see that about two-third of the HIV/AIDS are from sub-Saharan Africa which contains only 8.8% of the world population. The other three geographical areas of affinity with excessive number of cases compared to their population are- Latin America (4% of the world HIV cases, 7.8% of the world population), North America 2.7% of the HIV world cases, 5.4% of the world population and Carribeans (1.0% world HIV cases, 0.6% of the world population). The South and South East Asia accounts for 18% of the world HIV cases. Compared to the total population in this area, this may be considered as very small, but

Table 13.2: Regional HIV/AIDS Statistics and Features (Dec. 1999), (In Millions)

Region	*Total HIV/AIDS*	*Newly Infected*	*Adult Prevalence rate (%)*	*Adult Women(%)*
Sub-Saharan	23.3(69%)	3.8	8.0	55
North-Africa/ Middle East	0.22(1%)	0.019	0.13	20
south/ South East Asia	6.00(18%)	1.3	0.69	30
East,Asia/Pacific	0.53	0.12	0.068	15
Latin America	1.3(4%)	0.15	0.57	20
Caribbean	0.36(1%)	0.057	1.96	35
Eastern Europe/ Central Asia	0.36	0.095	0.14	20
Western Europe	0.52	0.030	0.25	20
North America	0.92	0.044	0.56	20
Australia/N.Z.	0.012	0.0005	0.10	10
Total	33.6	5.6	1.1	46

the trend over the years, which was less than 1% before 1990 is now the level of 18%, indicates the emerging situation. The cumulative number of AIDS cases in India and prevalence rates for different risk groups are given in Table 13.3.

The pattern and rate of spread of the HIV are reported to vary greatly both within and between different countries [4]. The World Health Organization Global Programme on AIDS initially categorized the HIV epidemic in different countries according to the estimated time at which the virus entered the country, the particular risk groups (homosexuals, IDUs and heterosexuals) and first affected (i.e., Patterns I, II and III) [16]. However, the initial patterns have been altered over time and much heterogeneity in

Classification of countries/regions by prevalence

Category	*Adult HIV Prevalence (%)*
Highest	≥ 8.0
Very High	2.0-8.0
High	0.5-2.0
Intermediate	0.125-0.50
Low	0.031-0.125
Very Low	0.008-0.031
Lowest	0.002-0.008
Entry	< 0.002

Table 13.3: HIV/AIDS in India

Cumulative number of AIDS cases (upto June-2000)	*6487*
Prevalence (Adult rate)	0.82%
Prevalence among CSW	30-65%
Prevalence among STD	23.4%.
Incidence among CSW	26.1/100Py
Incidence among Men	9.4/100Py
Incidence among NSCW	8.4/100Py
Dominant mode of transmission	Hetero-Sexual

spread is apparent even within countries initially designated to a particular pattern. The notion that the epidemic has attained an epidemic equilibrium in some communities has led some researches to describe the epidemic as consisting of three major phases, representing an early silent phase of slow spread, a rapid epidemic phase and an endemic plateau in sero-prevalence. This is far too simplistic a view as in reality the epidemic is formed from a series of waves of infection representing transmission within and between different risk groups. A number of studies have indicated that the overall epidemic may even be multipeaked if mixing between the different groups is limited.

A wider range of factors determines the overall shape of the epidemic. These include the presence of other sexually transmitted diseases (STD) that act as co-factors in the transmission, poverty promoting commercial sexual activity, the supply and acceptability of barrier contraceptives, social disruption created by armed conflict, pattern of mixing between sexual activity classes, and different geographical locations and the manner sexual activity changes with age in the two sexes and pattern of sexual contact between age classes [4]. The relative importance of each of those and other factors in promoting HIV transmission varies widely in different societies. Such heterogeneity makes prediction of the pattern of the epidemic in the coming decades extremely difficult. However, the WHO gives the global projections for different countries. The projections are based on the 'epimodel' developed by the WHO Global Programme on AIDS [16]. The model used is essentially a gamma function for the prevalence of infection with time. The assumption upon which the models were built and projections were made raises many doubts as to the validity of the projections.

3. Modelling the Spread of HIV/AIDS

Mathematical models are being constructed to generate insights into how different factors combine to shape the epidemic. One or few factors can be allowed to vary in the model structure while the rest remain constant in order to pinpoint the influence of each factor on the shape of the epidemic. More generally they can be used to determine what needs to be measured to improve understanding to gain a quantitative impression of the potential demographic impact of AIDS in defined locations and most importantly to generate insights into the potential effectiveness of different intervention strategies either used in isolation or in combination [13].

Many research groups in the world have produced models of varying degrees of complexity, which purport to describe the spread of the HIV/AIDS epidemic, and each team tends to feel that their own model is best under particular circumstances. In this paper, we have gone along a rather different path. Instead of concentrating on a particular model, we have investigated the potential usefulness of a number of different methods. In particular we will address the task of estimating the number of HIV/AIDS cases.

The model developed in the early stage of the epidemic can be split into three categories: statistical projections, adhoc methods and mathematical models. Since the data required for input into sophisticated models is strictly limited, the emphasis had been on the use of simple models to provide estimates of magnitude.

(i) Statistical Projections

These models use the reported AIDS cases data for short-term projections of AIDS cases. Such projections have been made by several groups who have applied statistical extrapolation techniques such as growth curves, power index curves, log linear models on doubling times to the observed temporal curve of reported AIDS cases [7,8,10]. The models assume that after adjustment for inherent reporting delays and incomplete reporting, past trends in the reported cases will continue for the next few years in a pattern similar to that already observed. Such methods are easy to apply but they consider no biological or epidemiological data other than reported AIDS cases.

(ii) Adhoc Methods

This category includes methods, which do not have a direct statistical or modeling connotation. The spread of AIDS in the developed countries like USA is of a different magnitude to that reported in developing countries. It would seem possible to attempt to gain insight from the American experience using Lag time. This approach works on the member of AIDS cases per million populations and the lag in the incidence of AIDS can be calculated for any country. Applying this incidence the cumulative total of

AIDS cases can be estimated. Making the lag time not constant made the procedure a little more accurate.

It is a sad fact that most of the persons who will be diagnosed as having AIDS in the intervening year are already infected with the virus several years back. Unfortunately, it is not possible to estimate the date of infection quite precisely. However, there is a set of data from which it is possible to estimate the date of infection quite precisely [11]. These data pertains to AIDS patients in USA who received a single transfusion of infected blood. Madly et.al. (1987) found that a Weibull distribution produced a good fit to the distribution of incubation times. The two parameters Weibull distribution has the density function;

$$f(n) = pq^{p} X^{p-1} \exp(-qx)^{p}, \quad X > 0$$

Using maximum likelihood techniques, the fitted parameter values were p = 2.396, q = 0.1077 for adult cases in the range 5-59 years, giving a mean and median 8.23 and 7.97 years respectively. Suppose it assumed that the doubling time of HIV infection is similar to that of AIDS and one infection was introduced in India during early eighties, the cumulative number of HIV cases for each year can be estimated and hence the incidence of new HIV cases occurring in a year. Using the percentage points of the cumulative Weibull distribution, we can estimate the number of AIDS for future that stem from each of the earlier year HIV cases.

(iii) Mathematical Models

One of the advantage of mathematical models is that it is possible to interrogate, the system on the familiar “what if?” basis. Another advantage is that it is possible to incorporate whatever knowledge exists relating to factors such as transmission mechanism, incubation period etc., in mathematical models.

The heterosexual transmission accounts for more than 80% of all HIV infections and therefore it is reasonable to attempt to model cases resulting from this particular transmission characteristic. The group may be divided into three subgroups as follows: (a) Susceptibles (b) Infections and (c) AIDS.

Griffith and Whealer [6] gave a schematic representation. The model can be translated into 4 non-linear differential equations and can be solved using numerical techniques. The model can also be used to extrapolate backwards to find the first infection in a country. A good review of mathematical models for hetero-sexual communities was given by many authors [2,4,14].

(iv) WHO Model (Chin and Lwanga, 1991)

The WHO model was developed for estimation and short-term projections of AIDS cases. This model uses a HIV point prevalence estimate in combination with the estimated year of HIV transmission become widespread and HIV infection curve during the epidemic period. The data are then used to calculate annual cohorts of HIV infected adults. AIDS cases are then determined by multiplying each annual cohort by progression rates from infection to clinical AIDS.

The year of extensive spread of HIV infection is defined as the year when at least 1% HIV sero-prevalence in high risk groups. The model assumes 1980 ± 2 years for North America, Sub-Saharan Africa and Western European Countries. For South and South East Asian countries it is assumed as 1988. Similarly HIV point prevalence was estimated with the available published and unpublished databases most of them are not accurate. The shape of the HIV incidence curve is assumed to be a sigmoid curve and a simple gamma function, $tp^{-1} e^{-t} /(p-l)^{1}$ is used to describe the HIV incidence curve at time t. p derives the shape of the curve and a value of p = 5 is used for projections. Also, the progression rates from HIV to AIDS are assumed to follow a Weibull function.

(v) Age and Sex Structural Model

This model [9] basically combines forecasting and surveillance activities to estimate the HIV prevalence trends, behavioural interventions and demographic and economic impacts based on comprehensive approach using age and sex structural surveillance data. The deterministic models for long-term projections and estimation models for short-term projections are evaluated. The age and sex structural model is an estimation model based on a HIV infection curve, by time and age, which is applied to the population structural by age. The data inputs required are, estimate of HIV prevalence, date of onset of infection, distribution of population by age and sexually transmitted disease by age incidence. The estimates of HIV prevalence and of date of epidemic onset used are similar to that of epimodel. For demographic data, the actual data of Cole-Demney model life table are used. The natural history data is taken as Weibull curve and the age profile of risk of HIV infection in a susceptible population is assumed to be Pearson type-I curve. The model has much strength when compared to the previous approaches and one of the widely used methods. However, it has the limitation of other models, as no accurate estimates are available for source of the parameters.

(vi) Recent Developments

Cumulative Incidence and Survival Methods

This method treats HIV prevalence at any given age for a stable endemic population as the cumulative incidence of new infections at each preceding age adjusted for mortality. A model for age-specific incidence is fitted using maximum likelihood technique.

Constant Prevalence Method

The method calculates the incidence of new infections within a time interval as the difference between observed prevalence levels at two successive age intervals after adjusting mortality. A good description of the two approaches is given with application to Uganda data [5].

(vii) Future Work Needed

Models, which characterize the disease transmission process and mixing between infections and susceptibles, taking biological and behavioural inputs, are to be developed to get reliable estimates of the magnitude of the epidemic. Models to provide qualitative descriptions of HIV spread and long-term projections are also the need.

References

[1] Broader S., Merigan Jr. T.C., Bolognese D. *Text Book of AIDS Medicine.* Williams and Wilkins Pub. 1994.

[2] Brookmeyer R. AIDS, Epidemics and Statistics. *Biometrics.* 1996, **52;** 781-796.

[3] Chinn J., Lwanga S.K. Estimation and Projection of Adult AIDS Cases: A simple epidemiological model. *Bulletin of World Health Organization* 1994, **69** : 399-406.

[4] Garnett G.P., Anderson R.M. Factors controlling the spread of HIV in heterosexual communities in Developing countries: Patterns of mixing between different age and sexual. Activity classes. *Phil. Trans. R Soc. Lond.* 1993, **B342:** 137-59.

[5] Gregson S., Donnelly C.A., and Porker C.G., Anderson. Demographic approaches to the estimating of incidence of HIV infection among adults from age specific prevalence data in stable endemic conditions. *AIDS.* 1996, 10: 1689-97.

[6] Griffths J.D. and Wheeler K.A. Modelling the spread of AIDS. In: *Statistics in Medicine.* Ed. F. Dunstom and J. Rickles. Oxford Pub. 1991.

[7] Healy M.R.J. and Tillett H.E. Short-term extrapolation of the AIDS Epidemic. *J. Roy. Statist. Soc.(A).* 1988, **151:** 50-65.

[8] Isham V. Mathematical modelling of the transmission dynamics of HIV infection and AIDS: A Review, *J.Roy. Statist. Soc., (A).* 1988, **151:** 5-30.

[9] Low B.D. and Stoneburner R.L. An Age and Sex Structural HIV Epidemiological Model: Features and Applications. *Bulletin of World Health Organization,* 1997, **75:** 213-21.

[10] Morgan W.M. and Curran J.W. AIDS: Current and future trends. *Public Health Reports.* 1986, **101:** 459-65.

[11] Peterman T.A., Stoneburner R.L., Allen J.R., Jaffe H.W. and Guran J.W. Risk of HIV Transmission from Heterosexual Adults with Transfusion Associated Infection. *J. Amer. Medical Assoc.,* 1988, **59:** 55-63.

[12] Quinn T.C. Global Burden of the HIV Pandemic. The *LANCET.* 1996, **348:** 99-106.

[13] Rowley J.T., Anderson R.M., Ng W.T. Reducing the Spread of HIV Injection in Sub-Saharan Africa: Some Demographic and Economic Implications. *AIDS.* 1990, **4:** 47-56.

[14] Schwager S.J. Statistical and Mathematical Modeling: A Review. Lecture Notes in Biomathematices No. 83. Springer-Verlag 1989.

[l5] UNAIDS/WHO. Global Summary of HIV/AIDS Epidemic. Geneva 2000.

[16] WHO Global Programme on AIDS. Current and Future Dimensions of *HIV/AIDS Pandemic*-ACapsule. Geneva. 1991.

[17] WHO. *The Current Global Situation of HIV/AIDS.* Geneva. 1996.