A Probability Model for Estimating the Force of Transmission of HIV Infection and Its Application

Mariamma Thomas[1], Damodar Sahu[1,*], Yujwal Raj[2], Arvind Pandey[1]

[1]National Institute of Medical Statistics, Ansari Nagar, New Delhi, 110029, India
[2]Department of AIDS Control, National AIDS Control Organization, New Delhi, 110001, India

Abstract　Information on force of transmission of HIV infection is crucial for better understanding of the epidemic. It is useful for better programme planning and evaluation. The current study aims at fitting probabilistic models for estimating the force of transmission of HIV infection due to different risk behaviour, heterosexual/homosexual/IDUs. The force of transmission between the risk groups bridging the infection is obtained as the joint probability function of the growth probabilities of HIV positivity over time among them. The growth probabilities of HIV positivity is estimated by fitting appropriate probability distributions to observed data from HIV Sentinel Surveillance (HSS) in Andhra Pradesh, India. The joint probabilities are estimated assuming that the risk behavior among the groups is random and the growth of infection within each group is independent in a region. In the study population, the average force of transmission is estimated to be 3.4755E-09 among general population, 1.5508E-03 among MSMs and 4.8455E-03 among IDUs, and 2.132E-03 in the total population. The results are subsequently used to estimate the HIV burden and compared with the estimates derived by other approaches.

Keywords　Force of transmission, Distribution of incubation period, high risk behavior groups (HRGs), Men having sex with men (MSM), Injecting drug users (IDUs), Female sex workers (FSWs)

1. Introduction

Information on the force of transmission of any infectious disease is important information for understanding and controlling the epidemic. In case of HIV/AIDS epidemic, the forces of transmission among the risk groups that bridge the infection are useful in understanding the risk behavior. Further, the knowledge of the force of transmission over time enables direct estimation of HIV/AIDS burden and other related parameters such as incidence, ART needs and AIDS related mortality.

In the past, a number of models have been propounded to describe the complex phenomena of the human immunodeficiency virus/acquired immuno-deficiency syndrome (HIV/AIDS) [1, 2, 3]. These models under various assumptions illustrate possible epidemic process due to varying parametric conditions. The models on distribution of incubation period and its use for back projections [4, 5, 6] and regression analysis for understanding the process of AIDS spread and disease progression [7, 8] etc., are a few models to mention. In this article, an attempt has been made to estimate the force of transmission of HIV infection in a population by estimating the growth probability of HIV positivity rates observed among different risk groups over time and deriving the interaction probabilities between risk groups that transmit the infection.

The article aims at developing a simple probability model to estimate the growth probabilities of HIV positivity in different risk groups over time. It also emphasizes the application of models for estimating the force of transmission of HIV infection among heterosexual high risk behavior groups and within homosexuals and injecting drug users (IDUs). It is assumed that the risk behavior between heterosexual groups and within homosexuals and IDUs are random and the growth of infection among each risk group is independent of each other in any region. The new HIV infections over time are estimated by applying the aggregate force of transmission to the adult (age above 15 years) population. The number of new infections developing to AIDS each year is estimated by applying the Weibull incubation time probabilities to develop AIDS since infection. In this process we get an estimate of new infections in the form of number developed to AIDS and that remaining as HIV in the following years. The multiple decrement life table framework is used to derive the annual cumulative number of cases developing AIDS and that remaining as HIV. The exercise has been done for one state, Andhra Pradesh in India.

2. Data and Methodology

2.1. Data

The HIV infection is transmitted between the high risk behaviour population, i.e., (i) female sex workers (FSWs) and their clients (regular /casual) (ii) intravenous drug users (IDUs) and (iii) men having sex with men (MSMs). The path diagram explaining the transmission dynamics is presented in Figure 1.
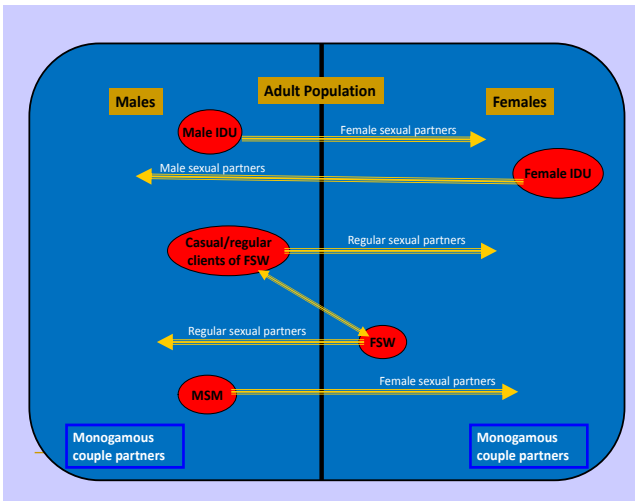

Figure 1.   HIV transmission dynamics

In India, the National AIDS Control Organization (NACO) tracks the epidemic among different risk groups through the HIV Sentinel Surveillance (HSS) since 1998. Therein the epidemic among FSWs, MSMs and IDUs are monitored through targeted intervention sites and the clients are tracked in STD clinics of health facilities. The general population excluding above risk groups is monitored through HIV testing of antenatal care attendees in the health care facilities for adult population aged 15-49. The above surveillance data on HIV positivity rates for the period 1998 - 2011 [8, 9] have been used to estimate the growth probabilities of HIV infection over time within each risk groups.

2.2. Methodology

Fitting of an appropriate probability distribution over time of the HIV positivity rate for each risk group is expected to provide growth probabilities of HIV infection among them. The force of transmission of HIV infection over time between two risk groups bridging the infection is thought of a function of the growth probabilities (joint probability function) of infection among the two groups.

We consider the following three models to study growth probabilities of HIV infection in different risk groups:

(i) Weibull Distribution

The general form of Weibull distribution [10, 11] with three parameters is given by,

$$f(x) = \frac{\beta}{\alpha} \left(\frac{x - \gamma}{\alpha} \right)^{\beta - 1} e^{-\left(\frac{x - \gamma}{\alpha} \right)^{\beta}}$$

where β = shape parameter; α = scale parameter and γ = location parameter. Since the HIV positivity rate before the start of the epidemic was zero, the location parameter in the above model could be set at zero. If the location parameter γ is set to zero as the time when HIV infection initiated in the population is considered zero. Then the two parameter Weibull distribution becomes:

$$f(x) = \frac{\beta}{\alpha^{\beta}} x^{\beta - 1} e^{-(x/\alpha)^{\beta}} \tag{1}$$

(ii) Gamma distribution

The general form of Gamma distribution with three parameters is given by

$$f(x) = \frac{\left(\frac{x - \gamma}{\alpha} \right)^{\beta - 1} exp\left(-\frac{x - \gamma}{\alpha} \right)}{\alpha \Gamma(\beta)}$$

where β = shape parameter; α = scale parameter and γ = location parameter. As explained above, if the location parameter is set to zero, the two parameter Gamma distribution function becomes:

$$f(x) = \frac{x^{\beta - 1} e^{-x/\beta}}{\alpha^{\beta} \Gamma(\beta)} \tag{2}$$

(iii) Lognormal distribution

The probability distribution function of the Lognormal distribution is given by

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} exp\left\{ -\frac{1}{2\sigma^2} (\log(x) - \mu)^2 \right\} \tag{3}$$

where, μ and σ are the population parameters of the model.

2.2.1. Estimation of Parameters and Fitting of Models

Setting the initial year of infection in India as 1985 (in view of the fact that the first case of HIV positive in India was detected in 1986), the parameters of the models are estimated by the method of moments. With these estimates of parameters, the growth probability curves for each risk group are plotted using the aforesaid three models as presented in Figure 2.

The best epidemiological curve has been considered for further analysis. The shape of the curve is the criterion for deciding the best epidemiological fit. Epidemiologically, the HIV infection grows slowly in the initial time reaches the peak and then starts declining.

The Gamma and Lognormal curves are epidemiologically acceptable for all groups except for ANC. Among the two models, the Lognormal curve appeared to be more appropriate for HRG population and thereby selected for deriving the force of transmission. In case of ANC group, the Weibull model provided epidemiologically plausible growth curve and is therefore chosen for further analysis.
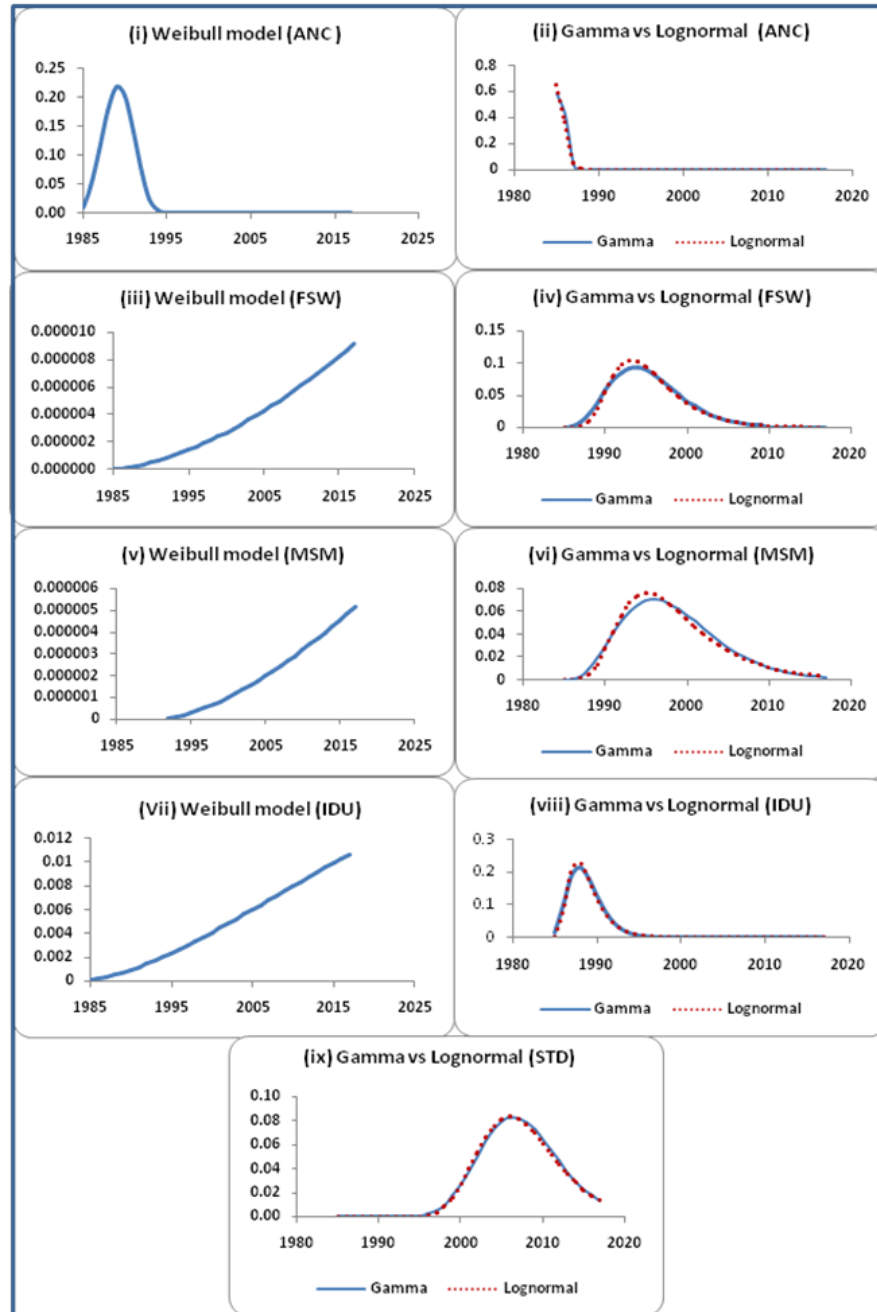
Figure 2. Growth probability curves of HIV infection among different risk groups by three models

2.2.2. Estimation of Transmission Probabilities

The transmission probabilities of HIV infection over time between risk groups that bridge the population has been estimated as the joint probabilities of the growth probabilities among the respective risk groups over time under the following assumptions.

(i) The HIV infection is transmitted through high risk behaviour – sexual/ intravenous drug use.

(ii) The groups are mixing randomly in the total population.

(iii) Growth of infection among each risk group in a region is independent.

(iv) The HIV epidemic curves gradually rise to the peak, stays stable for some time and slowly declines.

The transmission probabilities due to heterosexual risk behaviour have been derived as the product of the growth probabilities of infection among STI patients, FSWs and ANC women.

Similarly, under the assumption that the risk behavior among MSMs and the needle exchange among IDUs are random, the transmission probabilities due to respective risk behaviour are derived as the product of the growth probabilities of infection within them over time. The three transmission probabilities are thereafter aggregated to obtain the force of transmission in the total population over time as presented in Figure 3.
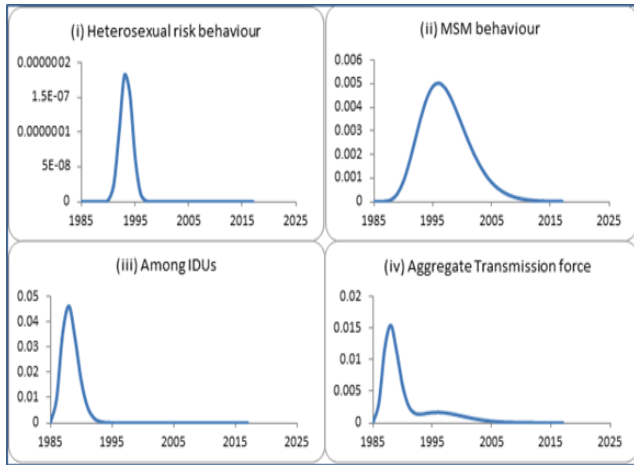
Figure 3. Force of transmission of HIV infection in Andhra Pradesh due to different risk behaviour

2.2.3. Estimation of HIV Burden

The estimated force of transmission over time on the adult population is then used to derive the annual number of new infections. Other epidemiological parameters of HIV burden, like number of people living with HIV (PLHIV), need for antiretroviral treatment (ART) and number of HIV related deaths are also derived as illustrated fellows.

Estimation of New Infections

The number of persons contracting HIV infection in each year from initial year of the epidemic t_0 to t_n, a specified future time is estimated by applying the force of transmission λ_0 to λ_n to adult population (age 15 and above) exposed to risk. The population at risk, R_t at time t is taken from the Spectrum population projection, the data used for HIV estimation in India [12]. The initial year t_0 is considered to be 1985 as the first case of HIV infection in India was detected in 1986. The steps for estimation of new infections are as follow:

Let

P_t = $P_0, P_1, P_2, \ldots, P_n$ the adult population at time t, $t = 0$ represents 1985.

λ_t = $\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_n$ the force of transmission of HIV infection at time t.

P_{tr} = $P_t - P_{t-1}$ the new population entering to risk at time t.

D_t = $D_0, D_1, D_2, \ldots, D_n$ the proportion of adult deaths at time t.

R_t = $R_{t-1} - N_t + P_{tr}$ the population exposed to risk at time t; where $R_0 = P_0$ and N_t is as given below.

N_t = $(R_{t-1} - R_{t-1} \times D_{i-1}) \times \lambda_{i-1}$ the new infections in the beginning of time t out of R_{i-1} exposed to risk.

Estimation of PLHIV and ART needs

An HIV infected individual needs treatment when the disease progress to AIDS, i.e., when the CD4 counts reduce to 350 or less [13]. In order to estimate the ART needs the Weibull model for the incubation time distribution (time to develop AIDS) is applied to the estimated new infections over time. The median duration (8 years) and the shape

(2.415) of the incubation period distribution is taken as the same as that considered by the spectrum model used globally for HIV estimation. The scale parameter is estimated using moment method. The incubation time distribution for progression to need for treatment since initial infection is presented in Figure 4.
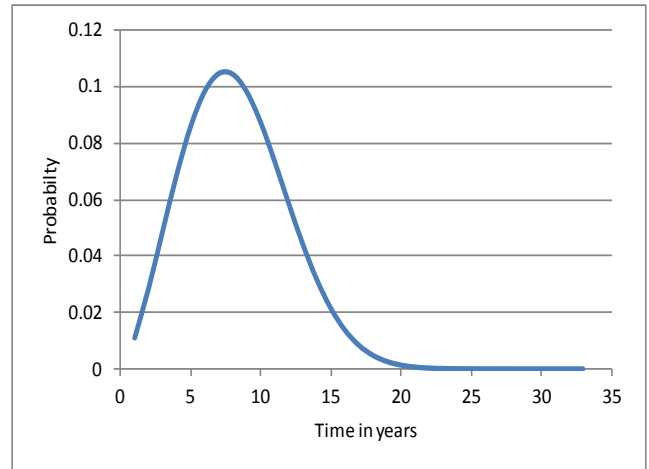

Figure 4. Weibull probability curve for time to ART needs

The estimated incubation times are applied to the new infections over time in the multiple decrement life table format to obtain the number needing treatment and remaining as HIV positives in consecutive years as the steps given below:

Let

π_i = the proportion developing AIDS in the ith year of infection, then

A_{12} = $(N_1 - N_1 \times D_1) \times \pi_1$ is the number progressing to AIDS and needing treatment in the second year out of N_1 infected in the first year (t_1).

H_{12} = $N_1 - A_{12}$ is the number remaining as HIV +ve in the second year out of N_1 infected in the first year (t_1).

A_{13} = $(H_{12} - H_{12} \times D_2) \times \pi_2$ is the number progressing to AIDS and needing treatment in the third year out of N_1 infected in the first year (t_1).

H_{13} = $H_{12} - A_{13}$ is the number remaining as HIV +ve in the third year out of N_1 infected in the first year (t_1).

(Note: N_t and D_t are the same parameters specified earlier)

Similarly for each N_i's, H_{ij} and A_{ij} are calculated for each year. Summation over j in the ith year for H_{ij} and A_{ij} (H_i and A_i) is the number living as HIV positives and needing treatment respectively in i^{th} year. The sum ($H_i + A_i$) is the number of PLFIV in the respective year.

Estimation of AIDS related deaths

Deaths due to AIDS are estimated from A_i's by applying the survival rates derived from the Weibull model. The median time (3 years) and the shape (2.415) of the Weibull model for time to death from need for ART in the absence of treatment is taken as the same as that considered by the spectrum model used globally for HIV estimation. The Weibull survival probability curve for AIDS cases is presented in Figure 5.
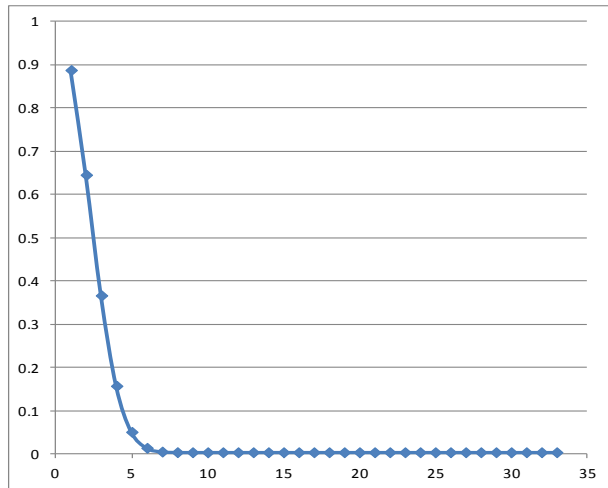
Figure 5. Weibull survival probabilities of AIDS cases

The number of AIDS deaths is calculated as below:
Let

l_i = the survival rate of AIDS cases in the ith year after progression to AIDS.

L_{22} = A_2, number living with AIDS in the second year of the epidemic, i.e., initial AIDS cases.

L_{23} = $L_{22} \times l_1$ is the number surviving in the first year after AIDS out of L_{22}

M_{22} = $L_{22} - L_{23}$ is the number of AIDS death in the first year of AIDS

Similarly, L_{ij}'s and M_{ij}'s are estimated for each A_i's. Summation over j in the ith year for L_{ij}'s and M_{ij}'s is the number living with AIDS and AIDS deaths respectively in i^{th} year. $B_i = H_i + L_i$ is the HIV burden in the ith year.

3. Results and Discussion

The model based estimates of the force of transmission of HIV infection among different risk groups and at the aggregate level has been derived in the state of Andhra Pradesh, India. The force of transmission is estimated as the joint probability of the growth probabilities of HIV positivity in the risk groups that bridge the infection. The growth probabilities of HIV infection in different risk groups are estimated from a probability model fitted on respective HIV positivity data. Three probability models (Weibull, Gamma and Lognormal) have been fitted over the five risk groups viz, STI patients, FSWs, MSMs, IDUs and ANC women. While the Lognormal curve has been selected for all four high risk groups, viz. STI patients, FSW, MSM and IDU, the Weibull model is found epidemiologically plausible for the ANC group.

The growth curves depict that the epidemic started rising in the early stage of the epidemic (1985-87) in the risk groups FSW, MSM, IDU, and ANC. However, in the STI group it started rising only around 1995. The curve reached the peak only in 2006 for STI group for others the peak was during 1989-95 (ANC-89, IDU-90, FSW-93, and MSM-95).

While the curve for the force of HIV transmission among IDUs and MSMs started rising as early as in 1986 and 1989, it started rising in the general population due to heterosexual risk only in 1990.

The average force of transmission in the general population due to heterosexual risk behaviour in the state is estimated to be 3.4755E-09 and the same among MSMs and IDUs are 1.5508E-03 and 4.8455E-03 respectively. The average force of transmission in the population of Andhra Pradesh is estimated to be at 2.132E-03.

The force of transmission due to heterosexual risk in the population stated above includes the transmission among high risk behaviour population and their partners who are otherwise at low risk. In order to understand the programmatic relevance of heterosexual high risk behaviour, this parameter is further disaggregated to obtain a rough estimate of the force of transmission between FSWs and their clients. The mean force of transmission due to heterosexual high risk behaviour is estimated at 3.2072E-04. However, this is only a rough estimate as the STI patents may include homosexuals too.

The aggregate force of transmission in the total population over time is further used to estimate the HIV burden. The results for the years 2007-11 are compared with the Spectrum estimates [14] and are presented in Table 1. The trends of all estimates of impact indicators (new HIV infections, adult HIV prevalence, PLHIV, ART needs and AIDS deaths) by both the methods are declining. However, the current method estimated smaller values for new infections, ART needs and AIDS deaths. Whereas the estimates for PLHIV and adult HIV prevalence are larger by the current method compared to Spectrum estimates. However, the Spectrum estimates include HIV positive children also. The current estimates are only for population above 15 years which accounts for a portion of the difference in all estimates.

The major difference between the two methods is that the Spectrum model estimates the PLHIV first by multiplying the observed HIV positivity from small samples to the larger populations and the incidence is estimated by back calculation. That is, new infections are dependent on the estimated PLHIV. In the current methodology all parameters are estimated as byproducts of the new infections based on the estimated transmission probabilities and other biological and population parameters. The PLHIV in the current method is obtained by cumulating the survivors of the new infections each year.

In the methodology the distribution of incubation period is applied to new infections in a multiple decrement framework where certain percentage of the cases are not developing AIDS even after 20 years and hence AIDS death is reduced. This is one of the reasons for larger number of PLHIVs.

Table 1. HIV/AIDS related estimates in Andhra Pradesh

		2007		2008		2009		2010		2011	
		Current Model	Spectrum	Current Model	Spectrum	Current Model	Spectrum	Current Model	Spectrum	Current Model	Spectrum
New HIV infection (15+)	Estimate	8384	22063	6112	19588	4432	18548	3201	17465	2304	16603
	Minimum Maximum	-	12209 33510	-	8899 31319	-	6904 31141	-	5947 30975	-	5251 30739
Adult HIV Prevalence (15-49)	Estimate	2.11	0.97	1.99	0.91	1.90	0.85	180	0.8	1.72	0.75
	Minimum Maximum	-	0.86 1.08	-	0.8 1.01	-	0.75 0.96	-	0.7 0.92	-	0.65 0.88
PLHIV*	Estimate	1094976	489063	1047136	467994	1004179	449103	965864	432643	931914	419180
	Minimum Maximum	-	431991 551360	-	413656 526351	-	395801 505012	-	379083 493906	-	361450 488232
Need for ART (Adult 15+)	Estimate	34081	124639	28839	128390	24136	160436	19954	164792	16281	170005
	Minimum Maximum		107164 144674		112523 147589		142418 180462		147234 183772		152825 188106
AIDS deaths	Estimate	44797	46427	38449	42337	32765	38919	27666	35276	23109	31347
	Minimum Maximum		4003 53932		36292 49315		33209 45333		29575 41481		25669 37138

*Spectrum estimates are including children whereas the current estimates are only for adults

4. Conclusions

The article has attempted to evolve a basic probabilistic model in order to estimate the force of transmission of HIV infection between heterosexual high risk behaviour groups and within homosexuals and IDUs. Growth probabilities of HIV infection over time among different risk groups are estimated in the process. The transmission probability function is estimated as joint probability function of the growth probabilities of the risk groups that bridge the population.

Transmission probabilities due to various risk behaviours are important useful information from programmatic point of view. This could be an indicator of the epidemic control or otherwise. It also indicates the impact of intervention programmes and further needs for initiation of new programmes or modification of existing ones.

As an application the aggregate force of transmission over time has further been used to estimate the annual new HIV infections. The number of new infections progressing to AIDS is estimated applying the Weibull incubation probability distribution and using the multiple decrement process. The deaths among estimated AIDS cases are also derived applying the Weibull survival probabilities for AIDS persons in the absence of treatment. Though the numbers are not tallying with the Spectrum results, the trend is declining in both the methods. Major reason for differences in the results is that the current method estimated only the population above 15 years of age whereas the Spectrum results include children. Others include the methodological differences mentioned in the results section.

Improving the estimation process by adding, age-sex components, fertility among HIV positives and ART of various groups could improve the result. Since the basis of HIV estimation in this exercise is the force of transmission of HIV infection in the population, the methodology eliminates the use of population sizes of HRGs - an uncertain input in the Global methods.

REFERENCES

[1] R. Brookmeyer and Mitchell, H. G., 1994, *AIDS Epidemiology: A Quantitative Approach*, Oxford University Press, London.

[2] M. Thomas, Sudhendu Biswas, Rehana Sultana and Arvind Pandey, 2006, Stochastic models to study the dynamics of HIV/AIDS: An analytical review, In A. Pandey (Ed.):

Biostatistical Aspects of Health and Population, 106-132; Hindustan publishing Corporation (India).

[3] S.M. Berman. 1990, A stochastic model for the distribution of HIV latency time based on T4 counts, *Biometrika*, 77(4), 733-741.

[4] M. Alvaro and Jane XU. 1996, Models for the incubation of AIDS and variations according to age and period, *Statistics in Medicine,* 15, 2459-2473.

[5] M. Thomas. 2002, EM Algorithm approaches to estimate HIV intensities and incubation period from reported AIDS cases, *Epidemiology, Health and Population* (*Statistical Applications*), BRPC (India) Ltd. edited by Anil Kumar.

[6] W.Y. Tan and Hsu, H., 1989, Some stochastic models of AIDS spread, *Statistics in Medicine*, 8: 121-136.

[7] M. Thomas. and A. Pandey, Ed., 2002, A Study on the Spread and Progression of HIV/AIDS in India'; *Biostatistical Aspects of Health and Epidemiology*, (Eds) C.M. Pandey, U. Singh & P. Mishra, Prakashan Kendra, Lucknow, pp. 228-231.

[8] *Annual Sentinel Surveillance for HIV infection in India Country Report 2004*. National Institute of Health and Family Welfare and National AIDS Control Organization, Ministry of Health and Family Welfare.

[9] *Annual HIV Sentinel Surveillance Country Report 2008-09*. National Institute of Health and Family Welfare and National AIDS Control Organization, Ministry of Health and Family Welfare.

[10] Christian Walck, 2007, *Handbook on Statistical Distributions for Experimentalists*, University of Stockholm.

[11] Mohammad A. Al-Fawzan, 2000, Methods for Estimating the Parameters of the Weibull distribution, *interstat. Statjournals. net/YEAR/ 2000/articles/0010001.PDF*.

[12] Technical Report India: HIV Estimates 2012, National Institute of Medical Statistics and National AIDS Control Organization.

[13] *Consolidated Guidelines on the Use of Antiretroviral drugs for treating and preventing HIV infection – Recommendations for a public health approach*, World Health Organization, June 2013.

[14] Stover J, Johnson P, Hallett T, Marston M, Becquet R, Timaeus IM. The Spectrum projection package: improvements in estimating incidence by age and sex, mother-to-child transmission, HIV progression in children and double orphans. *Sex Transm Infect* 2010; 86 *Suppl 2*: ii16–21.