



Research Article

Estimating Disability Adjusted Life Years using Survival Models in HIV/ AIDS Risk Groups

Gurprit Grover¹, Sangeeta Chakravarty², Sanya Aggarwal¹, Vishal Deo³

¹Department of Statistics, Faculty of Mathematical Sciences, University of Delhi, Delhi, India.

²Institute of Economic Growth, Delhi University Enclave, Delhi, India.

³Department of Statistics, Ramjas College, University of Delhi, Delhi, India.

DOI: <https://doi.org/10.24321/0019.5138.202172>

I N F O

Corresponding Author:

Sanya Aggarwal, Department of Statistics, Faculty of Mathematical Sciences, University of Delhi, Delhi, India.

E-mail Id:

sanya.aggarwal.stats@gmail.com

Orcid Id:

<https://orcid.org/0000-0003-1177-0708>

How to cite this article:

Grover G, Chakravarty S, Aggarwal S, Deo V. Estimating Disability Adjusted Life Years using Survival Models in HIV/ AIDS Risk Groups. J Commun Dis. 2021;53(4):36-47.

Date of Submission: 2021-09-26

Date of Acceptance: 2021-12-31

A B S T R A C T

Introduction: Advances in human immunodeficiency virus (HIV) treatment have led to greater survival rates and have brought about a shift in the burden of disease from mortality to morbidity. The main purpose of this study is to estimate the Disability Adjusted Life Years (DALYs) of HIV infected patients associated with different modes of transmission.

Methods: Non-parametric Kaplan-Meier estimate has been utilised to develop survival function, and the mean residual life model has been utilised to estimate the life expectancy of patients alive at the end of the study. The impact of factors such as age, sex, hepatitis B and syphilis on life expectancy has also been assessed by fitting a proportional mean residual life model. DALYs have been calculated based on the results of both models.

Results: Retrospective time to event data of HIV patients undergoing Antiretroviral Therapy (ART) in Dr Ram Manohar Lohia Hospital, New Delhi, India has been utilised to illustrate the modelling technique. The study suggests that in total, 42300.15 DALYs were lost which includes 39765.10 years of life lost due to premature death and 2535.05 years of life lived with disability. When the covariates were taken into consideration, 47592.14 DALYs were found to have been lost with an average of 17.64 DALYs lost per patient.

Conclusion: Our results suggest that the high-risk groups such as homosexuals and parent to child transmission are a major cause of concern, which are in accordance with the existing national policies. Also, we would suggest that gender-based and age-based policies should be incorporated to reduce the burden of disease.

Keywords: DALY, HIV/ AIDS, Mode of Transmission, Survival Analysis, Kaplan Meier Model, Mean Residual Life



Introduction

India has come a long way in addressing the human immunodeficiency virus (HIV) epidemic. After attaining its peak in the year 2000, the country has seen a decline in the overall trend of the epidemic.¹ The improved survivorship of HIV patients has led to an increase in life expectancy, resulting in a higher prevalence of people living with HIV (PLHIV). As of 2020, the total number of PLHIV in India is estimated at 23.19 lakh (18.33 lakh-29.78 lakh).² Children accounted for 3.5%, and 44.3% of total infections were among females in 2020 as compared to 3.4% CLHIV in 2019 and women constituted around 44% of total estimated adult PLHIV.^{2,3} However, the heterosexual route of transmission remains a predominant mode of acquiring infection accounting for almost 837 of every 1000 new infections reported nationally in 2019-20.⁴ Further, while around 4-6% of the cases have been reported to have acquired the infection via infected syringes and needles, 2-3% have been infected through the homosexual/ bisexual route.⁴

Acquisition and transmission of HIV infection are also closely associated with sexually transmitted infections (STIs) which presents a huge burden of disease and reproductive health problems.^{5,6} Syphilis is a sexually transmitted complex disease with increasing incidence all over the world, especially among men who have sex with men (MSM).⁷ HIV, hepatitis C, and hepatitis B (HBV) are among the blood-borne viruses transmissible from parent to child, through sexual intercourse as well as blood or blood products. HBV and HIV co-infection is a cause of concern in India in the management of HIV infected individuals, due to their common routes of transmission.⁸ Early detection and treatment of STIs reduces the chances of HIV infection by 40%. Therefore, assessing disease burden in terms of mortality or morbidity among HIV infected cases with respect to the route of transmission and commonly prevalent co-infections may provide a deeper insight into the current challenges posed by the epidemic.

Estimation of the burden of disease is fundamental in the allocation of resources, policy formulation and implementation. Disability Adjusted Life Years (DALYs) is one of the popular measures of burden of disease. It is a summary measure of population health that combines years of life lost due to disability (YLD) and years of life lost (YLL) due to premature mortality in a single index.⁹ DALYs are mainly calculated on population-based data utilising different approaches, from incidence, prevalence, and hybrid-based perspectives.¹⁰

Despite the advantages of being a uniform measure of population health, estimation of DALYs can be perplexing for censored data. In observational studies of time to event data, complete data is rarely available, with censored observations at the end of the study. The analysis can

be quite challenging due to induced censoring. While calculating DALY for such type of data, a major obstacle is to estimate the duration of illness and years of life lost for the cases alive after the end of study. The WHO guide to Cost effectiveness 2003 defines four methods of calculating life expectancy for estimating DALY.¹¹ Standard expected YLL (SEYLL) uses a global life expectancy curve based on the longest observed life expectancies rather than using local data.¹¹

Furthermore, the follow-up data provides detailed risk factor information which is not present in population-level data. The estimation becomes complicated when multiple conditions co-exist in the individual. The YLD measure indicates the morbidity level, measuring health conditions in terms of severity as represented by the Disability Weight (DW).^{12,13} Global burden of disease (GBD) studies have ignored the co-morbidities in the estimation of DW, which might result in overestimation of the overall disease burden. At present, the three most frequently used approaches in accounting for DW for multiple morbidities are additive, multiplicative and maximum limit methods.¹⁴ We developed a method to incorporate the effect of coinfections in the estimation of the remaining life expectancy of censored patients.

Various studies have been conducted to estimate DALYs corresponding to different diseases in India. Grover & Aggarwal¹⁵ used standard life expectancy for the calculation of DALY among HIV infected patients. Dandona et al.¹⁶ reported DALYs saved in the prevention of HIV infection among high-risk groups in Andhra Pradesh, India. Chu et al. estimated the expected life years lost in cancer patients using Monte Carlo method from the life table of the general population.¹⁷ Struijk et al.¹⁸ developed a methodology for DALY based on cohort data and compared different methods of calculation of life expectancy. Haagsma et al. compared the three comorbidity adjustment approaches in patients with injuries and common diseases with non-trivial health impacts suggesting that the three models do not fit well in the case of mild to moderate pre-existing disease.¹⁴

There is a lack of studies related to the methods of incorporating co-morbidities in the model and use of survival models in DALY estimations. This article thus attempts to introduce a novel approach in estimating DALYs based on retrospective follow-up data of HIV infected patients based on MoT by using survival analysis models in the presence of other factors. The idea is to estimate the remaining life expectancy and expected life years lost due to death among the censored group using mean residual life models. Also, to avoid overestimation of disability weights, the co-infections such as syphilis and HBV are incorporated in estimating remaining life expectancy using PMRL model. Sensitivity Analysis is then conducted by changing the

assumptions of DALY parameter to determine robustness of the results. To the best of our knowledge, this is the exclusive study using PMRL method for estimation of life expectancy and comparing results of DALYs for HIV/ AIDS patients based on MoT. In the next section, a description of the models utilised are described. In Section 3, the results from the implementation of the models have been presented. Section 4 contains a detailed discussion on the results and is followed by a conclusion about the results.

Material and Method

Study Population

The study uses retrospective longitudinal data on 3015 HIV/ AIDS patients who were enrolled for Antiretroviral Therapy (ART) therapy at the Dr Ram Manohar Lohia Hospital, Delhi during the period of 2004-2014 and were followed up till 2015. After eliminating 236 patients who were lost to follow up and 81 patients whose date of admission or date of final entry or death were not reported, a total of 2698 patients were found eligible for the study. No personal identification details of the patients were available in the ART register. No ethical approval was required for this study as no personal identification details were available except the registration number, which has not been mentioned in the study.

By the end of the study period, 438 patients had died and 2260 were alive. A vital response variable considered for the study is time to death which is defined as the time between the date of entry into the cohort and the date of death, or the date of last visit, or the end of the study. Other patient-specific variables considered were age at the date of entry in the study, sex, MoT, co-infections such as HBV and syphilis, and CD4 cell count on the first visit. Wherever the data on co-infection is missing, the values were imputed using k-nearest neighbour (kNN) imputation method. Data were managed in Microsoft Excel and the analysis was conducted using R software.

Disability Adjusted Life Years (DALY)

DALY represents one year of healthy life lost. It is an indicator of the burden of disease in a population, which is the measurement of the gap between current health status and an ideal disability and disease-free life.^{9,19} In practice, DALYs for a specific cause is calculated as the sum of years of life lost due to premature mortality (YLLs) from that cause and the years of healthy life lost due to disability (YLDs) for people living in states of less than good health resulting from the specific cause.¹² Mathematically, the mortality component, YLLs of DALY is calculated as²⁰:

$$YLL = \frac{KCe^{ra}}{(r + \beta)^2} [e^{-(r+\beta)(L+a)}\{- (r + \beta)(L + a) - 1\} - e^{-(r+\beta)a}\{- (r + \beta)a - 1\}] + \frac{1-K}{r}(1 - e^{-rL}) \quad (1)$$

Where k = age weighting modulation factor; C = constant; r = discount rate; a = age of death; β = parameter from the age weighting function; and L = standard expectation of life

at age a. The formula for YLDs differs only in the addition of D (the disability weight) as below²⁰:

$$YLD = D \left\{ \frac{KCe^{ra}}{(r + \beta)^2} [e^{-(r+\beta)(L+a)}\{- (r + \beta)(L + a) - 1\} - e^{-(r+\beta)a}\{- (r + \beta)a - 1\}] + \frac{1-K}{r}(1 - e^{-rL}) \right\} \quad (2)$$

Where k = age weighting modulation factor; C = constant; r = discount rate; a = age of onset of disability; β = parameter from the age weighting function; L = duration of disability; and D = disability weight. Disability weight measures the severity on the scale of zero (perfect health) to one (worst possible health state or death). The disability weight reflects the proportional reduction in the good health due to disability or disease. The disability weights are generally taken from published Global burden of disease (GBD) studies which compile the data from all over the world for global usage.²⁰ In the estimation of standard expectation of life at age of death, the standardised life expectancy table is utilised. The maximum life span is 82.5 years for females and 80 years for males and is taken from the country with the highest life expectancy in the world, Japan.¹⁹

Kaplan-Meier Estimator

Let T be the random variable representing survival time that has a distribution function, F(t). The survival function S(t) at time t is defined to be the probability that the survival time is greater than t, where $S(t) = P(T > t) = 1 - F(t)$. Let T_i be the time when the event occurs for the ith patient and C_i be the corresponding censoring time. When T_i is subject to right censoring, X_i is the observed time which is minimum of (T_i, C_i) , i.e., X_i is equal to T_i if the event was observed and T_i if it is censored. Finally, define $\delta_i = I(T_i \leq C_i)$ where $I(\cdot)$ is an indicator function and takes the value 0 for right censored event time and value 1 when $T_i \leq C_i$. Let W_i be the covariates of p-dimensions for the ith patient. Thus, the observed data for the ith patient is $(X_i, \delta_i, T_i, W_i)$.

The Kaplan-Meier estimator is the most commonly used survival analysis method. It is a non-parametric method used to estimate the survival probabilities of the occurrence of an event at a certain point in time.²¹ It is a piecewise constant and an empirical survival function for censored data. The Kaplan-Meier estimate is also known as “product limit estimate”. Few assumptions have been made in the study. Firstly, it is assumed that subjects that are being censored have equal probabilities of surviving like those who remain in the study. Secondly, the survival probabilities are assumed to be the same for patients that entered early and later in the study. Finally, it is assumed that the events took place at the specified time. The survival probability²² at any particular time is thus obtained as:

$$S(t_i) = \frac{(\text{No. of patients living at the start}) - (\text{No. of patients who died})}{(\text{No. of subjects living at the start})} \quad (3)$$

The cumulative survival function, $S(t_i)$, at time t_i is shown in the following equation:

$$S(t_i) = \prod_{i=1}^j \left(1 - \frac{d_i}{n_i}\right) \quad (4)$$

Where $S(t_i)$ is the probability of being alive at time t_i ; n_i is the number of patients alive before t_i ; d_i is the number of patients who died at time t_i ; and $t_0 = 0$, $S(0) = 1$, $i = 1, 2, \dots, j$. It should be noted that the data is noisy when there are few observations particularly when the events are rare, since it is a piecewise constant.

Mean Residual Life (MRL) and Proportional Mean Residual Life (PMRL) Model

In survival studies, mean residual life is the expected additional lifetime given the patient has survived until time t . More specifically, assuming X is a continuous random variable having distribution function $F(t)$, the mean residual life function after a fixed time point t is defined as follows:^{21,23}

$$m(t) = E(X - t | X > t) \quad (5)$$

$$= \frac{1}{S(t)} \int_t^{\infty} (x - t) dy(x), t \geq 0 \quad (6)$$

where $S(t)$ is the survival function, $m(t)$ is the MRL function. Writing and employing Tonelli's theorem yields an equivalent formula:

$$m(t) = \frac{1}{S(t)} \int_t^{\infty} \int_t^{\infty} du dy(x) = \frac{1}{S(t)} \int_t^{\infty} \int_u^{\infty} dy(x) du \quad (7)$$

$$m(t) = \frac{\int_t^{\infty} S(y) dy}{S(t)} \quad (8)$$

which exists for all t if and only if $\int_0^{\infty} m^{-1}(y) dy = \infty$. The necessary and sufficient conditions for the existence of mean residual life function of continuous non-negative random variable is summarised as:²³

- $m(t) \geq 0$ for all $t \geq 0$, and continuous,
- $m(t) + t$ is non-decreasing in t ,
- If there exists a w such that $m(w) = 0$, then $m(t) = 0$ for all $t \geq w$; otherwise, $\int_0^{\infty} m^{-1}(y) dy = \infty$.

The MRL function can be inverted to the survival function by using inversion formula:²³⁻²⁵

$$S(t) = \frac{m(0)}{m(t)} \exp \left\{ - \int_0^t \frac{dy}{m(y)} \right\} \quad (9)$$

MRL is often referred to as life expectancy or expectancies of life in the study of human population. To study the association between T and covariate vector W , the following proportional mean residual life (PMRL) model is incorporated.^{23,26}

$$m(t; \theta) = \theta m_0(t) \quad (10)$$

Where $\theta = \exp(\beta' W)$, β is an unknown $p \times 1$ vector of regression parameters. Here, $m(t; \theta)$ is the MRL function of the survival time T corresponding to a $p \times 1$ vector covariate W , $m_0(t)$ is some unknown baseline MRL function. Applying the inversion formula twice on equation (9) and using equation (10), we have:

$$S_1(t) = \frac{m_1(0; \theta)}{m_0(t; \theta)} \exp \left\{ - \int_0^t \frac{dy}{m_1(y; \theta)} \right\} \quad (11)$$

$$= \frac{m_1(0)}{m_0(t)} \exp \left\{ - \frac{1}{\theta} \int_0^t \frac{dy}{m_1(y)} \right\} \quad (12)$$

$$= S_0(t) \left(\frac{\int_0^{\infty} S_0(y) dy}{\mu_0} \right)^{\frac{1}{\theta-1}} \quad (13)$$

Where $\mu_0 = m_0(0)$ is the mean distribution of the origin.

Taking the first derivative of equation (12) gives the pdf as:

$$f_1(t; \theta) = - \frac{dS_1(t; \theta)}{dt} = \frac{\mu_0}{m_0^2(t)} \left\{ m_0'(t) + \frac{1}{\theta} \right\} \exp \left\{ - \frac{1}{\theta} \int_0^t \frac{dy}{m_1(y)} \right\} \quad (14)$$

The maximum likelihood estimation method under the PMRL model can be applied to estimate the parameters for the baseline distribution and the proportionality parameter θ . The log-likelihood function is given as,²³

$$l(\theta; x_i) = \sum_{i=1}^n [\delta_i \log\{f_1(x_i; \theta)\} + (1 - \delta_i) \log\{S_1(x_i; \theta)\}] \quad (15)$$

The positive value of the parameter estimate would indicate an increase of the MRL function and the value would imply the magnitude of the increase within an interval of interest. On the other hand, a negative value of the parameter estimate would indicate a decrease in the MRL function and hence a negative effect of the covariate of interest. The analysis is conducted in R software utilising the function developed by Jeong.²³

Missing Data Imputation

K-nearest neighbour (kNN) algorithm is an instance-based learning method to impute missing values. This method is based on the principle of approximating real-valued or discrete-valued target functions by considering its k closest cases from the training set with known values in the attributes according to a given distance metric. The distance computation for defining the nearest neighbours is based on the extension of proposed Gower distance, which can handle mixed-type variables.^{27,28} Considering two input variables, i and j , the distance between the two cases is calculated as the weighted mean of each variable. The distance can be calculated by:

$$d_{i,j} = \frac{\sum_{k=1}^p w_k \delta_{i,j,k}}{\sum_{k=1}^p w_k} \quad (16)$$

Where w_k is the weight and $\delta_{i,j,k}$ is the distance between the two cases on its k th attribute, such that,

$$\delta_{i,j,k} = \begin{cases} 1 & \text{if } s_k \text{ is missing in } i \text{ or } j, \\ \delta_{i,j,0} & \text{if } s_k \text{ is a discrete variable,} \\ \delta_{i,j,N} & \text{if } s_k \text{ is a continuous variable} \end{cases} \quad (17)$$

With s_k as the k th attribute. In equation (17) it is considered that all $\delta_{i,j,k}$ are in $[0,1]$. For continuous type of variable, the absolute distance divided by the total range is used:

$$\delta_{i,j,k} = \frac{|x_{i,k} - x_{j,k}|}{r_k} \quad (18)$$

Where $x_{i,k}$ is the value of the k th variable of the i th observation and r_k is the range of the k th variable. For nominal and binary variables, 0/1 distance is used:

$$\delta_{i,j,k} = \begin{cases} 0 & \text{if } x_{i,k} = x_{j,k}, \\ 1 & \text{if } x_{i,k} \neq x_{j,k}. \end{cases} \quad (19)$$

Given an incomplete vector x , the missing value in x is estimated by kNN imputation method by finding the set of k closest neighbours in the missing feature to be imputed. Once the k nearest neighbours is found, the missing value is estimated depending on the type of data. The imputed value is given by the mean (continuous data) or the mode

(categorical data) of the set of k nearest neighbours.^{28,29}

Sensitivity Analysis

Sometimes the results obtained in a study can vary significantly upon the introduction of even small changes in the assumptions and values of certain parameters, leading to a situation where the entire nature of the results may alter. Sensitivity analysis is a methodological approach to deal with such uncertainties in the choice of parameters to assess the confidence in the chosen course of action and ascertain the value of collecting additional information to better inform the decision.³⁰ Univariate sensitivity analysis is conducted on the disability weight parameter of DALY to estimate the change in result from the base value. This is done to increase confidence in the results of the study. The disability weight parameter is adjusted on its range's endpoints to see the variation in the results.

Methodological Structure

Our analysis first focuses on estimating baseline life expectancies for censored population using equations (4) and (8) by incorporating Kaplan-Meier survival estimator. These estimates are then utilised in equation (10) for calculating PMRL for each censored case, in the presence of the covariates age, sex, syphilis, and HBV. These results are further utilised to estimate the remaining life expectancy at

death for determining YLL and duration of illness to estimate YLD. The estimates of YLLs and YLDs, and hence, those of DALYs, are obtained both with and without considering the effects of covariates. Finally, sensitivity analysis is conducted on DALY estimates for reporting the variation in results. For the purpose of this study, eunuchs are combined with males for the estimation of life expectancy as the standard life expectancy table is available for males and females only.

Results

Table 1 represents the summary data of all the predictors of interest in the study. The majority of the patients were in the age group of 25-30 years with a mean of 34.26 ± 10.95 . The minimum age was 1 year and the maximum was 82 years. 65% of the patients were males and 34% were females with eunuchs constituting only 1% in the study. By the end of the study, 84% of the patients were alive and 16% had died. The self-reported MoT of HIV in these patients was majorly heterosexual activities (87%), followed by infected syringes and needles (5%), mother to child (4%), blood and blood products (2%), and homosexual activities (2%). About 62% of the patients entered the study at stage IV of HIV/ AIDS, with only 3% of the patients enrolled at stage I. Looking at the addictions of the patients in the study, about 21% were addicted to smoking while 26% of the registered patients were addicted to alcohol.

Table 1. Descriptive Summary of Patients

Variables	Categories	Frequency	(%)	Mean ± SD	Range
Age (years)	< 15	132	5	34.05 ± 10.95	(1,82)
	15-25	321	12		
	25-35	1192	44		
	35-45	698	26		
	45-55	262	10		
	≥ 55	93	3		
Gender	Female	921	34		
	Male	1756	65		
Status	Eunuch	21	1		
	Alive	2261	84		
Addictions	Dead	437	16		
	Smoking	572	21		
Stage	Alcohol	700	26		
	I	75	3		
	II	142	5		
	III	815	30		
	IV	1666	62		
Mode of Transmission (MoT)	Blood transfusion	55	2		
	Heterosexual	2338	87		
	Homosexual	46	2		

	Infected syringe and needle	145	5		
	Parent to child	114	4		
Syphilis	Yes	76	3		
	No	2473	92		
	Unknown	149	5		
Hepatitis B	Yes	67	3		
	No	2445	90		
	Unknown	186	7		

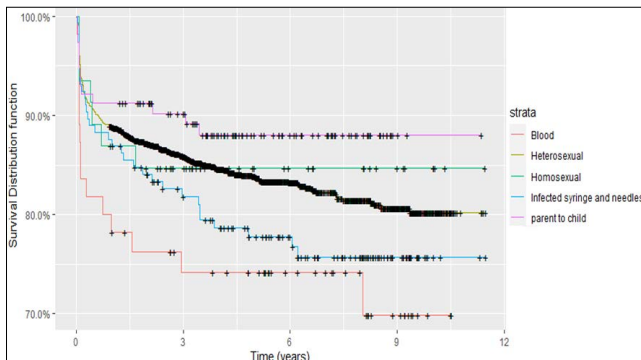


Figure 1. Kaplan-Meier Survival Curves of Different Groups of HIV Patients based on MoT

Table 2. Input Parameters for Estimation of DALY and Summary of Co-infections after KNN Imputation

Variables		Values	Uncertainty Bounds
HIV/ AIDS receiving ART		0.053 ³¹	0.034-0.079
C		0.1658 ²⁰	
b		0.04 ²⁰	
k		0 ²⁰	
r		0.03 ²⁰	
KNN imputed results		Frequency	Percentage (%)
Syphilis	Yes	78	3
	No	2620	97
HBV	Yes	68	3
	No	2630	97

At a quick glance at the Kaplan-Meier survival curve (Figure 1) estimates on the basis of HIV MoT groups, it is evident that the parent to child transmission group has the highest survivability as compared to the other groups. The survival in the infected syringe and needle group is significantly higher than blood transfusion group. Those infected through the homosexual route seems to have better survival than those who had experienced heterosexual transmission. Table 2 lists the values of input parameters utilised in the

estimation of DALYs and the prevalence of co-infections after kNN imputation. The disability weight was taken to be 0.053, as suggested by the GBD 2010 study.³⁰ We used the base case recommended by Murray and Lopez for the calculation of DALYs with $C = 0.1658$, $r = 0.03$, $K = 0$ and $b = 0.04$.¹⁸ After imputing the missing values on co-infection through kNN method, it is recorded that there were 68 HIV patients with syphilis coinfection and 78 with HBV (Table 2).

The results of YLDs, YLLs, and DALYs for all MoT groups of HIV patients, obtained without including the covariates in the analysis, have been presented in Table 3. The results show that when covariates are not considered, the overall DALYs lost are 42300.15 with 2535.05 healthy years lost due to disability (YLDs) and 39765.10 healthy years lost due to premature death (YLL). This means that on average, 15.68 DALYs are lost per patient. Among all MoT groups, maximum average DALYs are lost due to parent to child transmission (17.99 DALY per patient). Transmission through blood transfusion results in an average loss of 16.84 DALYs per patient and heterosexual mode of transmission leads to an average loss of 15.50 DALYs per patient. Transmissions due to infected syringes or needles and homosexual behaviour cost 16.33 and 15.72 years of DALYs per patient, respectively.

Figure 2 shows the survival curve for age groups after discretisation of the variable for further analysis. A significant difference between the two age groups, above 40 years and below 40 years, can be observed. For the HIV patients above 40 years of age, the probability of surviving about 10 years (x-axis in years) is slightly below 75%, while for those below 40 years of age is substantially above 75%. Similar inference can be drawn from the risk table. This cut-off point value of age (i.e., 40) is utilised for the estimation of remaining life expectancy using proportional mean residual life.

The impact of covariates on life expectancy is estimated using PMRL model. The PMRL model for subject i is specified as:

$$m(t|\theta) = m_0(t) \exp(\beta_{1i}Age_i + \beta_{2i}Sex_i + \beta_{3i}Syphilis_i + \beta_{4i}Hepatitis\ B_i) \quad (20)$$

Where, $m_0(t)$ is some unknown baseline MRL function. The covariates considered are coded as: age (below 40 =

0, above 40 = 1); sex (female = 0, male = 1); syphilis (no = 0, yes = 1) and hepatitis B (no = 0, yes = 1). Table 4 lists the estimate coefficients (β) of the covariates present in the PMRL model. The estimated coefficients of all the variables are negative, indicating a decrease in the MRL function when the factors take value, keeping all other factors constant. This suggests that males, above 40 years of age, with syphilis and HBV co-infection, have reduced residual life.

The values of YLDs, YLLs and DALYs calculated after including the covariates in the PMRL model are reported in Table 5. Overall, 47592.14 DALYs are lost with 2277.20 healthy years lost due to disability (YLDs) and 45314.94 healthy years lost due to premature death (YLL). On average, 17.64 DALYs are lost per patient. Among the different MoT, maximum per patient loss of DALYs is associated with the parent to child transmission (21.23 DALYs per patient). Average per patient loss of DALYs associated with transmission through blood transfusion and heterosexual activities are 17.84 DALYs and 17.40 DALYs, respectively. Transmissions due to infected syringes or needles and homosexual behaviour cost 18.25 and 18.82 years of DALYs per patient, respectively. The results of DALYs with and without covariates, for the five MoT groups, are compared in Figure 3. It is clear from the comparison that after taking into account the effects of covariates, there is a consistent increase in the estimates of DALYs lost by patients across all MoT groups. This suggests

that the covariate age, sex, and presence of co-infections play a significant role in assessing the burden of disease. Further, it can be inferred that the patient with one or more of the three attributes - older age (above 40 years), being male (or a eunuch) and having co-infections (syphilis and/or HBV) are expected to experience an increased burden of disease.

The gender-based distribution of the DALY values obtained from both models is presented in Table 6. Parent to child transmission has led to considerable DALYs lost among both males and females. Among the sexual transmission groups, females have acquired infection only through heterosexual behaviour and have lost maximum years of life to the disease. It can be observed that sexual behaviour is the mode of transmission of HIV among eunuchs with 18.65 DALYs lost per patient in the presence of covariates as compared to 15.59 DALYs per patient lost in the base case scenario. The results suggest that in the absence of covariates, females lost higher DALYs (16.72 DALY per patient) as compared to males and eunuchs (15.14 and 15.21 DALYs per patient, respectively). After taking into consideration other factors, it is evident that homosexual activities and infection through infected syringes and needles have higher DALYs lost. This means that co-infection along with gender are major contributors in increasing the burden of disease in these groups as compared to the others.

Table 3. YLDs, YLLs and DALY Estimates (Analysis without Covariates)

MOT Groups	No. of Patients	YLD	YLL	DALY	DALY/ Patient
Blood	55	47.32	878.67	925.99	16.84
Heterosexual	2338	2166.36	34065.45	36231.82	15.50
Homosexual	46	45.34	677.90	723.24	15.72
Infected syringes and needles	145	131.48	2236.41	2367.90	16.33
Parent to child	114	144.55	1906.66	2051.21	17.99
Total	2698	2535.05	39765.10	42300.15	15.68

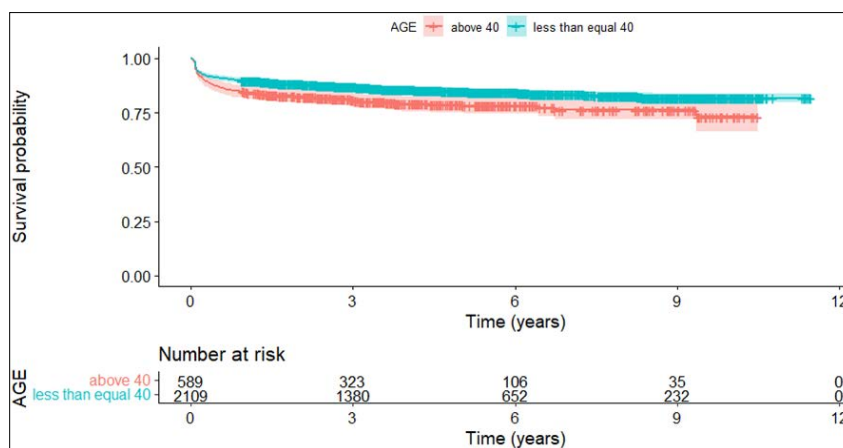


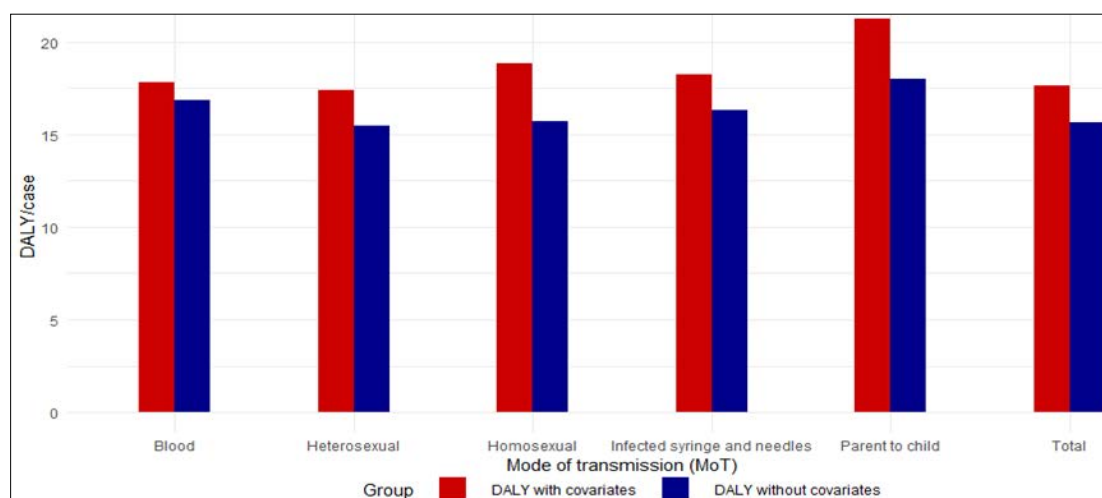
Figure 2. Kaplan-Meier Survival Curve based on Age Cut Point

Table 4. Estimated Covariates for Proportional Mean Residual Life Models

Covariates	Age	Sex	Syphilis	Hepatitis B
β	-0.16434	-0.2901	-0.07337	-0.1985

Table 5. DALYs for HIV Patients with Covariates

MOT Groups	No. of Patients	YLD	YLL	DALY	DALY/ Patient
Blood	55	44.95	936.31	981.26	17.84
Heterosexual	2338	1940.25	38738.97	40679.22	17.40
Homosexual	46	38.88	826.67	865.55	18.82
Infected syringes and needles	145	119.31	2526.94	2646.25	18.25
Parent to child	114	133.82	2286.04	2419.86	21.23
Total	2698	2277.20	45314.94	47592.14	17.64

**Figure 3. DALY Estimates (With and Without Covariates)****Table 6. DALY's Distribution based on Gender**

MoT Groups	Without Covariates			With Covariates		
	Male	Female	Eunuch	Male	Female	Eunuch
Blood	15.57	17.62	-	18.04	17.72	-
Heterosexual	14.94	16.57	12.92	17.73	16.78	16.01
Homosexual	15.81	-	15.59	18.93	-	18.65
Infected syringes and needles	15.86	17.32	-	18.61	17.50	-
Parent to child	17.70	18.63	-	22.37	18.74	-
Total	15.14	16.72	15.21	18.01	16.93	18.27

The results of sensitivity analysis on DALYs by adjusting disability weights are presented through Figure 4. The effect is observed by changing the DALY assumption of disability weight from 0.053 to 0.034 and 0.079.

The baseline value is at level 0 and x-axis represents the percentage change in the DALY estimates from the base case scenario. In the absence of covariates, it is evident that

the results vary between the range of -4% to 3% (Figure 4). Similarly, the percentage change in DALY estimates lies within the range -3% to 2% (Figure 4). It is observed that in both cases, DALY estimates of parent to child transmission group varied the most. The results of DALY estimates (with covariates) showed lesser variation as compared to the other group.

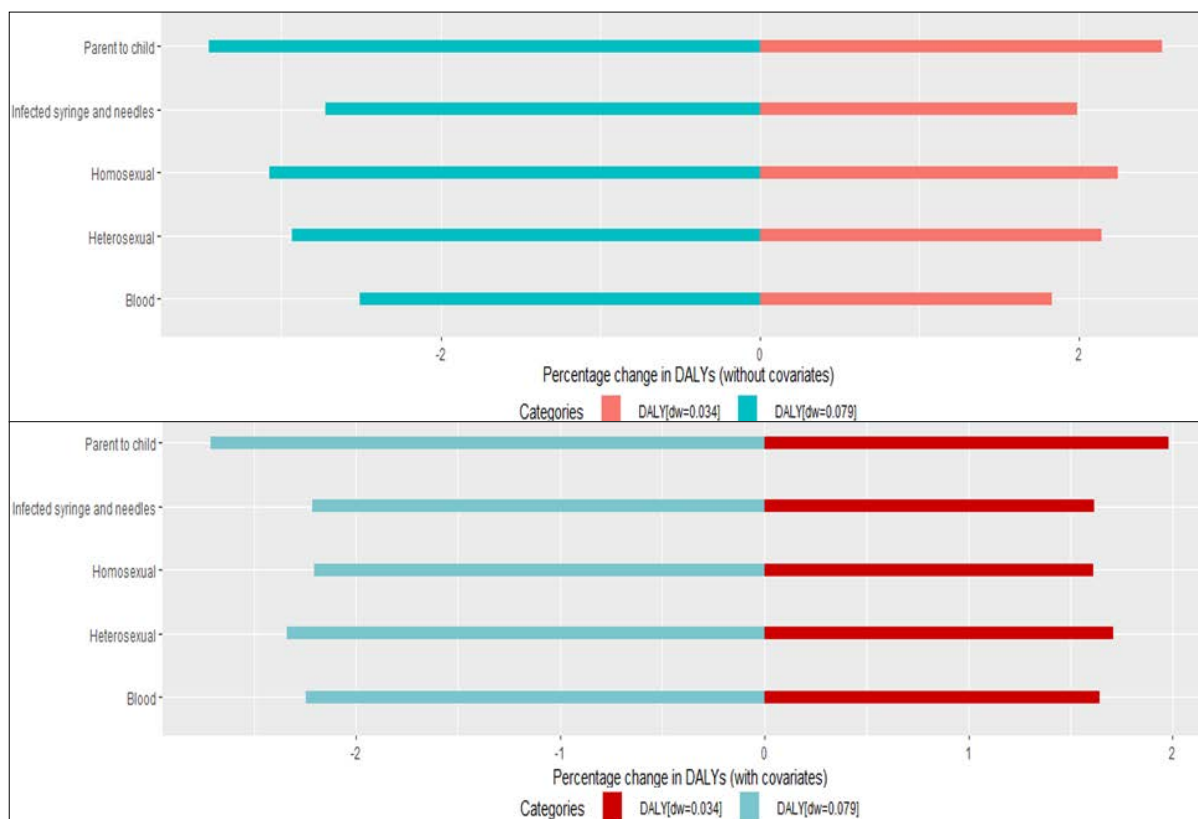


Figure 4. Univariate Sensitivity Analysis of DALY Estimates

Discussion

The study provides various significant findings. Analysis of self-reported MoT in the study suggests that heterosexual activity is the primary mode of transmission of infection followed by infected syringes and needles. Similar findings have been reported by National AIDS Control Organisation (NACO), reporting about 84%-86% of infections via heterosexual route.⁴ The findings of the study also suggests that 2.89% of the HIV patients in the study had syphilis and 2.29% had Hepatitis B co-infection which is comparable with findings of previous studies.³²⁻³⁴ The study shows that only 4% of the patients reported had acquired the infection through parent to child transmission of the virus. Although the overall survivorship curve shows the maximum probability of surviving in the parent to child transmission group, the findings from DALYs lost suggest that this group accounts for the highest number of healthy years of life lost to the disease. Based on gender distribution, females seem to have a higher burden of disease as compared to males in this group. The findings of our study are in line with other published works^{4,35} which report the prevalence of parent to child transmission among HIV in India. But the studies have not reported the burden of disease in terms of DALY, thus results are not comparable. Thus, this study provides the evidence of burden of disease in terms of DALYs to substantiate the pre-existing NACO policy of Prevention of Parent to Child Transmission (PPTCT) programme which

focuses on eliminating mother to child transmission of infection.³⁶

Baseline results on DALYs obtained in the absence of covariates suggest that transmission through blood transfusion and infected syringes and needles are significant contributors to DALYs lost. Infection through sexual behaviour (heterosexual/ homosexual) have similar DALYs. Further, the findings indicate that more DALYs are lost among females as compared to males and eunuchs, which is consistent with the results of other studies.^{37,38} These results are alarming as most of the females are in the reproductive (15-49) age group and are also at risk for parent to child transmission of the virus. Infected syringes and needles are becoming a major cause of concern of HIV transmission among males and females. In this study, eunuchs have reported having acquired the infection through sexual behaviours with a higher burden on homosexual route. Since the data on eunuchs is very small, the results may not be the representative of the true population. The results are in accordance with the existing policies of NACO focusing on high-risk groups in limiting the transmission of disease.^{3,4}

It is evident that when covariates such as sex, age, syphilis, and HBV are incorporated, a significant increase in DALYs lost is observed among all transmission groups. Homosexual activities and infection through infected syringes and needles become a significant contributor to DALYs lost, followed by heterosexual activities and blood transmission

groups when adjusted for covariates. A minor increase in DALYs lost is observed in females when covariates are included. The major impact of the inclusion of covariates is observed in homosexual males and eunuchs with increased DALYs as compared to their heterosexual counterparts. A significant increase of about 17% is observed in DALYs lost in eunuchs in the presence of covariates. The result is in accordance with the study by Solomon et al.,³⁹ which reports a high prevalence of HIV among homosexual and syphilis infected cases in India. The study is not directly comparable but provides a similar conclusion based on the prevalence of disease.

The study provides a novel approach in estimating DALYs by utilising the MRL model for estimating remaining life expectancies and compiling DALYs lost based on HIV infection transmission groups. There are various limitations of the study. Firstly, the study only considers right censoring. Patients who dropped out in the middle were not considered in the study. Another limitation is the combining of eunuchs and males in one variable during the estimation of YLLs and covariates for ease of calculation. Other factors such as education level and occupation of infected patients are not utilised in the study which could have provided more insight on MoT. Also, this study is based on only one ART centre in Delhi, India. Therefore, generalisation of our results at the national level should be done with utmost care.

Conclusion

Since the detection of the first HIV case in the year 1986, India has come a long way in addressing the epidemic. The improved survivorship of HIV patients has led to an increase in life expectancy, but at the same time, there is a problem of the significant burden of disease as a result of limited quality of life.

This study proposed a new methodology of estimating DALYs for HIV infected patients based on MoT using the mean residual life model to evaluate life expectancy of the censored group. The results suggest that parent to child transmission and homosexual behaviour have a higher burden of disease in terms of DALYs lost. The rising number of cases acquiring infection through infected syringes and needles have already become a cause of concern with significant DALYs lost. Other factors such as sex, age, and co-infection of syphilis and HBV play a crucial role in reducing life expectancy and increasing disabled years both in terms of mortality and morbidity.

The study will help policymakers and researchers to use these estimates in understanding the burden of HIV from the perspective of transmission groups. Healthcare services and planners must meet the challenge of providing utmost care to HIV/ AIDS patients and educating them regarding

different modes of transmission of the infection. In this way, the burden of disease can be curtailed, if not completely eliminated, by improving the quality of life of the infected population and reducing the risks of further transmission.

Source of Funding: None

Conflict of Interest: None

References

1. National AIDS Control Organization [Internet]. HIV facts & figures; [cited 2021 Jun 1]. Available from: <http://naco.gov.in/hiv-facts-figures>
2. National AIDS Control Organisation & ICMR-National Institute of Medical Statistics. India HIV Estimates 2020: Technical Brief. New Delhi: NACO, Ministry of Health and Family Welfare, Government of India; 2021.
3. National AIDS Control Organization & ICMR-National Institute of Medical Statistics. India HIV Estimates 2019: Report. New Delhi: NACO, Ministry of Health and Family Welfare, Government of India; 2020.
4. National AIDS Control Organization. Sankalak: Status of National AIDS Response. 2nd ed. New Delhi: NACO, Ministry of Health and Family Welfare, Government of India; 2020.
5. National AIDS Control Organization. Operational guidelines for programme managers and service providers for strengthening STI/RTI services. NACO, Ministry of Health and Family Welfare, Government of India; 2011.
6. Cohen MS, Council OD, Chen JS. Sexually transmitted infections and HIV in the era of antiretroviral treatment and prevention: the biologic basis for epidemiologic synergy. *J Int AIDS Soc.* 2019;22:e25355. [PubMed] [Google Scholar]
7. Madi D, Achappa B, Vernakar P, Bhaskaran U, Ramapuram JT, Mahalingam S. Prevalence of Syphilis among People living with HIV (PLHIV) attending a tertiary care hospital in coastal south India. *Asian J Med Sci.* 2015;6(6):77-79. [Google Scholar]
8. Koli S, Girish Kumar CP, Selvaraj V, Prabu R, Chandrasekar C, Valan AS, Suria Kumar J, Raja K. Profile and prevalence of HBV among HIV affected individuals attending the largest public HIV care center in India. *Virusdisease.* 2016 Sep;27(3):215-9. [PubMed] [Google Scholar]
9. Murray CJL, Lopez AD. *The Global Burden of Disease.* Cambridge, MA: Harvard School of Public Health; 1996.
10. Schroeder SA. Incidence, prevalence, and hybrid approaches to calculating disability-adjusted life years. *Popul Health Metr.* 2012 Sep;10(1):1-7. [PubMed] [Google Scholar]
11. Tan-Torres Edejer T, Baltussen RM, Adam TA, Hutubessy R, Acharya A, Evans DB, Murray CJ. Making choices

- in health: WHO guide to cost-effectiveness analysis. World Health Organisation, Geneva (Switzerland); 2003. [Google Scholar]
12. Murray CJ, Lopez AD. Quantifying disability: data, methods and results. *Bull World Health Organ.* 1994;72(3):481-94. [PubMed] [Google Scholar]
 13. Hilderink HB, Plasmans MH, Snijders BE, Boshuizen HC, Poos MJ, van Gool CH. Accounting for multimorbidity can affect the estimation of the Burden of Disease: a comparison of approaches. *Arch Public Health.* 2016 Aug;74:37. [PubMed] [Google Scholar]
 14. Haagsma JA, van Beeck EF, Polinder S, Toet H, Panneman M, Bonsel GJ. The effect of comorbidity on health-related quality of life for injury patients in the first year following injury: comparison of three comorbidity adjustment approaches. *Popul Health Metr.* 2011 Apr;9:10. [PubMed] [Google Scholar]
 15. Grover G, Aggarwal S. A study comparing cost-effectiveness of combination therapy for preventing opportunistic infections among human immunodeficiency virus-infected adults on antiretroviral therapy. *Value Health Reg Issues.* 2021 May;24:107-13. [PubMed] [Google Scholar]
 16. Dandona L, Kumar SG, Kumar GA, Dandona R. Cost-effectiveness of HIV prevention interventions in Andhra Pradesh state of India. *BMC Health Serv Res.* 2010 May 10;10:117. [PubMed] [Google Scholar]
 17. Chu PC, Wang JD, Hwang JS, Chang YY. Estimation of life expectancy and the expected years of life lost in patients with major cancers: extrapolation of survival curves under high-censored rates. *Value Health.* 2008 Dec;11(7):1102-9. [PubMed] [Google Scholar]
 18. Struijk EA, May AM, Beulens JW, de Wit GA, Boer JM, Onland-Moret NC, van der Schouw YT, Bueno-de-Mesquita HB, Hoekstra J, Peeters PH. Development of methodology for disability-adjusted life years (DALYs) calculation based on real-life data. *PLoS One.* 2013 Sep;8(9):e74294. [PubMed] [Google Scholar]
 19. Donev D, Zaletel-Kragelj L, Bjegovic V, Burazeri G. Measuring the burden of disease: disability adjusted life year (DALY). In: Zaletel-Kragelj L, Božikov J, editors. *Methods and tools in Public Health.* 2010:715. [Google Scholar]
 20. Murray CJ, Lopez AD. *Global health statistics: a compendium of incidence, prevalence and mortality estimates for over 200 conditions (No. 2).* Harvard University Press; 1996. [Google Scholar]
 21. Gupta RC, Bradley DM. Representing the mean residual life in terms of the failure rate. *Math Comput Model Dyn Syst.* 2003;37(12-13):1271-80. [Google Scholar]
 22. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc.* 1958;53(282):457-81. [Google Scholar]
 23. Jeong JH. *Statistical inference on residual life.* Springer; 2014. [Google Scholar]
 24. Gumbel EI. Eine Darstellung der Sterbetafel. *Biometrika;* 1924:283-96. German. [Google Scholar]
 25. Cox DR. *Renewal Theory.* London: Methuen; 1962.
 26. Oakes D, Dasu T. A note on residual life. *Biometrika.* 1990;77(2):409-10. [Google Scholar]
 27. Gower JC. A general coefficient of similarity and some of its properties. *Biometrics;* 1971. p. 857-71. [Google Scholar]
 28. García-Laencina PJ, Abreu PH, Abreu MH, Afonso N. Missing data imputation on the 5-year survival prediction of breast cancer patients with unknown discrete values. *Comput Biol Med.* 2015 Apr;59:125-33. [PubMed] [Google Scholar]
 29. Batista GE, Monard MC. An analysis of four missing data treatment methods for supervised learning. *Appl Artif Intell.* 2003;17(5-6):519-33. [Google Scholar]
 30. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD; ISPOR-SMDM Modeling Good Research Practices Task Force. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. *Value Health.* 2012 Sep-Oct;15(6):835-42. [PubMed] [Google Scholar]
 31. Institute for Health Metrics and Evaluation (IHME). *Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2010 (GBD 2010) Disability Weights.* Seattle, United States of America; 2012.
 32. Agrawal A, Goyal A, Goyal S, Kumari S, Singh PK. Seroprevalence of hepatitis B and syphilis co-infection in human immunodeficiency virus-positive antiretroviral therapy attendees and human immunodeficiency virus-negative sexually transmitted infection attendees. *Indian J Sex Transm Dis AIDS.* 2016;37(1):94. [Google Scholar]
 33. Desikan P, Khan Z. Prevalence of hepatitis B and hepatitis C virus co-infection in India: a systematic review and meta-analysis. *Indian J Med Microbiol.* 2017 Jul-Sep;35(3):332-9. [PubMed] [Google Scholar]
 34. Chakravarty R, Pal A. Insights into human immunodeficiency virus-hepatitis B virus co-infection in India. *World J Virol.* 2015;4(3):255-64. [Google Scholar]
 35. Darak S, Panditrao M, Parchure R, Kulkarni V, Kulkarni S, Janssen F. Systematic review of public health research on prevention of mother-to-child transmission of HIV in India with focus on provision and utilization of cascade of PMTCT services. *BMC Public Health.* 2012 May;12:320. [PubMed] [Google Scholar]
 36. National AIDS Control Organization [Internet]. Prevention of parent to child transmission (PPTCT); [cited 2021 Jun 1]. Available from: <http://naco.gov.in/>

- prevention-parent-child-transmissionpptct
37. Hawkes S, Santhya KG. Diverse realities: sexually transmitted infections and HIV in India. *Sex Transm Infect.* 2002 Apr;78 Suppl 1(Suppl 1):i31-9. [PubMed] [Google Scholar]
 38. World Bank. World Development Report 1993: Investing in Health. New York: Oxford University Press for the World Bank; 1993. [Google Scholar]
 39. Solomon SS, Mehta SH, Srikrishnan AK, Vasudevan CK, Mcfall AM, Balakrishnan P, Anand S, Nandagopal P, Ogburn EL, Laeyendecker O, Lucas GM, Solomon S, Celentano DD. High HIV prevalence and incidence among MSM across 12 cities in India. *AIDS.* 2015 Mar;29(6):723-31. [PubMed] [Google Scholar]