On the Estimation of Expected Survival Time of AIDS Patients Undergoing Antiretroviral Therapy Using Censored Generalized Poisson Regression Model

Antiretroviral Tedavisi Alan AIDS Hastalarının Beklenen Sağkalım Süresinin Sansürlü Genelleştirilmiş Poisson Regresyon Modeli Kullanılarak Tahmini

Gurprit GROVER,[a]
Vajala RAVI[a]

[a]Department of Statistics,
University of Delhi, Delhi, INDIA

Geliş Tarihi/*Received:* 16.09.2013
Kabul Tarihi/*Accepted:* 25.10.2013

Yazışma Adresi/*Correspondence:*
Vajala RAVI
Lady Shri Ram College
University of Delhi,
Department of Statistics,
Lajpat Nagar-IV
New Delhi - 110024, INDIA
vravi.vr@gmail.com

ABSTRACT Objective: The progression of HIV infection depends not only on the combination of drugs and adherence to the Anti Retroviral therapy (ART), but also on various other factors. The survival of an HIV patient depends on interaction of these factors with the therapy. **Material and Methods:** The parametric approach with and without covariates have been used to analyze the survival data of HIV/AIDS patients. When covariates are not considered, survival distributions are fitted and best one chosen. The inclusion of covariates is analyzed using Censored Generalized Poisson Regression Model (CGPR). **Results:** For a retrospective right censored data of 1689 HIV/AIDS patients undergoing ART at the Ram Manohar Lohia Hospital, New Delhi, India, appropriate survival distributions in the absence of covariates are fitted. Using the Akaike Information Criterion, Gamma distribution is found to be the distribution of best fit. Assuming Gamma survival distribution with right censored survival time data, the mean survival time of AIDS patients using the method of maximum likelihood is found to be 11.14 years. Also, CGPR model for estimating the survival time while accounting for the impact of cofactors is used for the first time on the AIDS dataset. **Conclusion:** The advantage of using a CGPR model over the gamma right censored distribution is that the mean survival time of the AIDS patient undergoing ART can be estimated in the presence of significant prognostic factors. Using this model, age, gender, smoking status, alcoholism, WHO staging, improvement in CD4 count, opportunistic infections, number of visits and weight at initiation of ART are identified as significant prognostic factors effecting the survival time. On averaging the predicted survival times based on CGPR model, the mean survival time of AIDS patients on ART is found to be 12.12 years.

Key Words: Survival distribution; gamma and weibul distribution; censoring; generalized poisson regression model

ÖZET Amaç: HIV enfeksiyonunun ilerlemesi sadece ilaçların kombinasyonuna ve Anti Retroviral tedavisine(ART) bağlı değil, bunların dışında diğer faktörlere de bağlıdır. Bir HIV hastasının sağkalımı bu faktörlerin tedavi ile etkileşimine bağlıdır. **Gereç ve Yöntemler:** HIV/AIDS hastalarının sağkalım süresi ortak değişkenlerin varlığı ve yokluğu durumunda parametric yaklaşım kullanılarak analiz edilmiştir. Ortak değişkenlerin sansürlendiği durumda, sağ kalım dağılımları uydurulmuş ve en iyisi seçilmiştir. Ortak değişkenlerin dahil edilmesi Sansürlü Genelleştirilmiş Poisson Regresyon Modeli (Censored Generalized Poisson Regression Model-CGPR) kullanılarak analiz edilmiştir. **Bulgular:** Yeni Delhide bulunan Ram Manohar Lohia hastanesinde ART alan 1689 HIV/AIDS hastasının geriye dönük sağdan sansürlü verisi için ortak değişkenlerin varlığında uygun sağkalım dağılımları uydurulmuştur. Akaike bilgi kriteri kullanılarak, Gamma dağılımı en uygun dağılım olarak bulunmuştur. Gamma sağkalım dağılımı sağdan sansürlü sağkalım verisi ile göz önüne alındığında, en çok olabilirlik yöntemi kullanılarak AIDS hastalarının ortalama sağkalım süresi 11,14 yıl olarak bulunmuştur. Ek olarak CGPR modeli, kofaktörlerin etkisi dikkate alındığında sağkalım süresinin tahmini için AIDS veri seti üzerinde ilk kez kullanılmıştır. **Sonuç:** CGPR modelinin sağdan sansürlü Gamma dağılımı üzerinde kullanılmasının avantajı, ART alan AIDS hastalarının ortalama sağkalım süresinin prognostik faktörlerin varlığı durumunda tahmin edilebilmesidir. Bu model kullanılarak yaş, cinsiyet, sigara içme durumu, alkolizm, WHO evrelemesi, CD4 sayısındaki ilerleme, fırsatçı enfeksiyonlar, muayenelerin sayısı ve ART'ın başlama ağırlığı sağkalım süresini etkileyen anlamlı prognostic faktörler olarak tanımlanmıştır. CGPR modeli üzerinden hesaplanan tahmini sağkalım sürelerinin ortalaması alındığında, ART alan AIDS hastalarının ortalama sağkalım süresi 12,12 yıl olarak bulunmuştur.

Anahtar Kelimeler: Sağkalım dağılımı; gamma ve weibul dağılımı; sannsürleme; genelleştirilmiş poisson regresyon modeli

Turkiye Klinikleri J Biostat 2014;6(1):24-32

The advent of highly active antiretroviral therapy (HAART) has drastically improved the survivability of patients with HIV infection by achieving more sustained elevations in CD4 lymphocyte counts and effective suppression of HIV replication. Although the cure for it doesn't exist till date,[1] multiple drug formulations have taken care of compliance problems and therefore HIV infection is now manageable as a chronic disease in patients who have access to medication and who achieve durable virologic suppression.[2]

Several studies have been conducted to model the progression of HIV infection and to identify the prognostic factors affecting them. This progression is not only dependent on combination of drugs and adherence to the therapy, but also on various biological and socio-demographic factors that tend to have varying influences on the degree of infection. These factors range from age, co-infections (infection other than HIV), ethnicity, geographic location, genetics, infection route (how the disease was transmitted), nutrition, pregnancy, stress, and whether or not the patient smokes or uses recreational drugs can affect the rate at which an HIV patient develops AIDS.[3-5]

Bakanda et al.[6] used Kaplan-Meier curves and Weibul analysis to identify the positive correlation of age with survival time of HIV patients on ART. Authors[7,8] identified body mass index, WHO stages, and baseline CD4 count as significant predictors of mortality in the individuals undergoing ART in Malawi. Various studies[8,9] identified Gender as a significant factor effecting mortality in AIDS patients. Kitahata et al.[10] concluded on the basis of their studies that deferred ART therapy group was more susceptible to mortality as compared to the early therapy group. Researchers[11] identified age, intravenous drug users and stage of AIDS as significant markers for the survival of patients while gender was insignificant. Authors[12-14] also identified Tuberculosis, Diarrhea, Liver Cirrhosis, as comorbidities associated with HIV infection.

Poisson Regression analysis is a useful tool for the analysis of count data.[15] It derives its name from the Poisson distribution which is a mathematical distribution often used to describe the probability of occurrence of count data. When survival time of patients is recorded in days, the data becomes discrete thus providing us an opportunity to explore the use of Poisson distribution. The effect of various covariates on survival time of AIDS patients may be evaluated using this model. Poisson regression is an important alternative to partial-likelihood based analysis of proportional hazards model and to parametric analyses of such models. It creates proportional hazards models which is very often used for survival analysis. In fact, Poisson models usually replace Cox model, which cannot be easily applied to aggregated data. Terza[16] extended the Poisson regression model to censored count data with constant censoring threshold. Caudill and Mixon[17] considered the case of variable threshold. One of the earliest uses of Poisson regression for analyzing the survival data were seen when Frome[18] found lung cancer death rate to be closely related to the Age, dose rate (cigarettes per day) and the duration of smoking. Dickman et al.[19] used Poisson regression to estimate and model the net survival of cancer patients.

The basic underlying assumption for Poisson model is that sample mean equals sample variance. However, count data often exhibit substantial variations where the sample variance is either smaller or larger than the sample mean and it is classified as under- or over-dispersion, respectively. Several models have been proposed to overcome these limitations.[15,17,20,21] However, such models may not be generalized for handling both kinds of dispersion. In such an event, the generalized Poisson regression (GPR) model[22-24] is one of the few that can accommodate both under- and over-dispersion.

In a non-censored count data, the relationship between the sample mean and sample variance is a good measure of the amount of underlying dispersion. However, for a censored count data, it is very unlikely that the relationship between the mean and the variance is known. For such data the observed mean will be less than the true mean and same thing holds for variance. If the observed mean is less than the observed variance, the true mean could be less than or more than the true variance. Thus, in a censored count data, one may not know

the type of dispersion. In such situations, an extension of the GPR model for censored data is given by Famoye and Wang.[25]

Cox regression, a semi-parametric model, has been widely used for analyzing the survival time of AIDS patients by including covariates in the model.[26-29] This is a preferred model because fewer assumptions are needed to predict the prognostic factors associated with survival. Parametric models, however, are known to be more accurate than non-parametric methods when using survival models to make projections about the risk of death and future trends in mortality.[30,31] In this article, we propose to fit appropriate probability distributions to model the survival time of AIDS patients undergoing ART and obtain the maximum likelihood estimate of the mean survival time. On the other hand, we assume that the mean survival time is dependent on covariates like improvement in CD4 count, age, gender, opportunistic infection etc. Survival time being in days is discrete in nature and hence it's relation with covariates would be verified using Censored generalized Poisson regression (CGPR) model. The individual impact of these prognostic factors on the survival time would be calculated and mean survival time and dispersion parameter of every individual would be predicted using this model. Using the mean and variance of fitted values of the survival time, an estimate of the overall mean and variance of the survival time will be calculated. Even though, many authors have used different procedures and models.[11,32-37] to estimate the survival time of AIDS patients undergoing ART, generalized Poisson regression has never been used in this context. To the best of our knowledge, this is the first study where CGPR is being used to model the survival data of AIDS patients and that too, on Indian population.

MATERIAL AND METHODS

The data from cohort survival studies typically consist of information whether or not the event of interest occurred, the event or censoring time t, and a vector of possibly time dependent covariates $\boldsymbol{X}$ for each cohort member. Since interest centers on hazard rates, it is natural and useful for the purpose of analysis or summarization to reorganize such data into an event-time table defined by a cross classification over a set of time intervals and covariate categories.

Let $\boldsymbol{T_i}$ denote the survival time (in days) for the i^{th} AIDS patient ($i = 1, 2 \ldots n$) undergoing ART. The following three approaches are used to estimate the mean survival time:

2.1. PARAMETRIC APPROACH (WITHOUT COVARIATES)

Here, we assume that the survival times of AIDS patients follow some specific statistical distribution. Even though the list of such distributions is large, we consider 4 important survival distributions namely Exponential, Gamma, Weibul and Lognormal distribution respectively. Exponential distribution depicts a constant hazard rate and has been widely used to model lifetime distribution. Gamma and Weibul distributions are the generalizations of exponential distribution. The positively skewed distributions where the average values are low, variances are high and the values are not negative, generally accord with lognormal distribution. A general feature of these distributions is that they all belong to the exponential family and have been most often found to fit all kinds of survival data. For choosing the most appropriate distribution, the criteria of log-likelihood and Akaike information criterion (AIC)[38] is used.

$$\mathrm{AIC} = -2\log L + k \qquad (1)$$

where L denotes the likelihood function of the model evaluated at maximum likelihood estimates and k is the total number of parameters in the model. The models which have a higher log-likelihood or a lower AIC value are considered to be the best. Also, the estimates of parameters of survival distribution are obtained by the method of maximum likelihood and hence mean survival time is calculated.[39]

2.2. PARAMETRIC APPROACH (WITH COVARIATES)

The survival time of HIV/AIDS patient depends on various prognostic factors and therefore it is imperative to estimate the survival with reference to these covariates. Section 2.1 identifies the distribution that best fits the survival data of AIDS patients. In

order to estimate the mean survival time of AIDS patients in the presence of covariates, we have considered the Generalized Poisson regression model.

Let $\boldsymbol{T_i}$ (i = 1, 2 ... n) denote the survival time (in days) for the i^{th} patient which is affected by k explanatory variables $\mathbf{X} = \{X_{i1}, X_{i2}, .. X_{ik}\}$. We assume that $\boldsymbol{T_i}$ follows generalized Poisson distribution with probability function

$$P[T_i = t_i] = p_i = \frac{1}{t_i!}\frac{\lambda_i}{(1+a\lambda_i)}\left(\frac{\lambda_i(1+a\lambda_i)}{1+a\lambda_i}\right)^{t_i-1} \exp\left\{\frac{-\lambda_i(1+at_i)}{1+a\lambda_i}\right\},\ t_i = 0, 1, \ldots \quad (2)$$

with

$$E[\boldsymbol{T_i} \mid \mathbf{X}] = \lambda_i = e^{X'\beta} \text{ and } V[\boldsymbol{T_i} \mid \mathbf{X}] = \lambda_i(1+a\lambda_i)^2 \quad (3)$$

where $\boldsymbol{\beta}$ is the k-dimensional vector of regression parameters and the a measures the type of dispersion in the data. It is easily seen that the parameter $a = 0$ indicates the presence of equi-dispersion in count data and that the probability function in equation (1) reduces to the Poisson regression model, while $a > 0$ is over-dispersion and $a < 0$ is under-dispersion in the GPR model.

Since the data on the survival time of AIDS patients ($\boldsymbol{T_i}$) is censored therefore,

$$T_i = \begin{cases} \geq t_i & \text{if } i^{th} \text{ observation is censored} \\ t_i & \text{otherwise} \end{cases}$$

For obtaining the distribution of sample data under censored observations, we define an indicator variable d_i as

$$d_i = \begin{cases} 1 & \text{if } i^{th} \text{ observation is censored} \\ 0 & \text{otherwise} \end{cases}$$

The likelihood function of CGPR model[25] is given by,

$$L(a, \beta; \boldsymbol{T_i}) = \prod_{i=1}^{n}\left[P(T_i = t_i)\right]^{1-d_i}\left[P(T_i \geq t_i)\right]^{d_i} \quad (4)$$

$$= \prod_{i=1}^{n}[p_i]^{1-d_i}\left[1-\sum_{j=0}^{t_i-1}p_j\right]^{d_i} \quad (5)$$

$$\text{Log } L(a, \beta; \boldsymbol{T_i}) = \sum_{i=1}^{n}(1-d_i)\log[p_i] + \sum_{i=1}^{n} d_i \log\left[1-\sum_{j=0}^{t_i-1}p_j\right] \quad (6)$$

Differentiating equation (6) with respect to a and β, we get

$$\frac{\partial LogL(a,\beta)}{\partial \beta} = \sum_{i=1}^{n}\left\{(1-d_i)\left(\frac{t_i-\lambda_i}{(1+a\lambda_i)^2}\right)x_i - d_i\frac{\sum_{j=0}^{t_i-1}\partial p_i / \partial \beta}{1-\sum_{j=0}^{t_i-1}p_j}\right\} = 0 \text{ and} \quad (7)$$

$$\frac{\partial LogL(a,\beta)}{\partial a} = \sum_{i=1}^{n}\left\{(1-d_i)\left(\frac{-t_i\lambda_i}{(1+a\lambda_i)} + \frac{t_i(t_i-1)}{(1+at_i)} - \frac{\lambda_i(t_i-\lambda_i)}{(1+a\lambda_i)^2}\right) - d_i\frac{\sum_{j=0}^{t_i-1}\partial p_j / \partial a}{1-\sum_{j=0}^{t_i-1}p_j}\right\} = 0 \quad (8)$$

$$\text{where } \frac{\partial p_j}{\partial \beta} = p_j\frac{j-\lambda_i}{(1+a\lambda_i)^2}x_i \text{ and } \frac{\partial p_j}{\partial a} = p_j\left(\frac{-j\lambda_i}{(1+a\lambda_i)} + \frac{j(j-1)}{(1+aj)} - \frac{\lambda_i(j-\lambda_i)}{(1+a\lambda_i)^2}\right)$$

Equation (7) and (8) being non-linear in a and β may be solved by any iterative algorithm. The variances of the estimators of a and β can be obtained by solving the inverse of Fisher's information matrix given by

$$I(\beta, a) = \begin{bmatrix} \frac{-\partial^2 LogL(a,\beta)}{\partial\beta\partial\beta'} & \frac{-\partial^2 LogL(a,\beta)}{\partial\beta\partial a} \\ \frac{-\partial^2 LogL(a,\beta)}{\partial a\partial\beta} & \frac{-\partial^2 LogL(a,\beta)}{\partial a^2} \end{bmatrix} \quad (9)$$

The analysis of the CGPR model was done on Statistical Analysis Software (SAS 9.1). The requisite macro program for the analysis is given in by Chow& Steenhard.[40] Also, R software, Statistical Package for Social Sciences (SPSS 15.0) were used to fit Exponential, Weibul and Lognormal distributions while Statgraphics Centurion 15.1.0.2 was used to fit Gamma distribution for censored data.

DATA DESCRIPTION

The retrospective data of 1689 patients is obtained from the ART centre of Ram Manohar Lohia Hospital, New Delhi, India. This is the largest ART centre in Northern India and caters to a large number of patients coming from across various states in India. The patients coming for the ART treatment are followed up till 16th November, 2010 which is the threshold for right censoring. A record for every patient is maintained in terms of date of subsequent visits, CD4 counts, WHO stage etc. The improvement in CD4 count is obtained as the difference between last known CD4 count and the initial value. Patients who were lost to follow-up after just one visit were excluded from the study as their improvement in CD4 count could not be determined.

RESULTS

Four distributions, namely Exponential, Gamma, Weibul and Lognormal are fitted to the survival data of HIV/AIDS patients undergoing ART. The

TABLE 1: Distributions fitted to survival time of AIDS patients.

Distribution	Estimates of parameters	Log likelihood	AIC
Exponential *(θ)*	0.0002	-1996.8	3995.673
Gamma *(α, β)*	(8.132, 0.002)	-1789.3	3497.285
Weibul *(k, λ)*	(11.913, 2.29)	-1883.8	3771.507
Lognormal *(μ, σ²)*	(7.79, 1.09)	-1894.7	3793.453

results are given in Table 1 and the probability plots are made in Figure 1. Using the criterion of log-likelihood and AIC, Gamma distribution is found to be the best fitted distribution. The maximum likelihood estimates of the shape and scale parameters of the gamma distribution are obtained as $\hat{\alpha}$ = 8.132 and $\hat{\beta}$ = 0.002. Hence the mean survival time of AIDS patients undergoing ART is $\frac{\hat{\alpha}}{\hat{\beta}}$ = 4066 days = 11.14 years.

The descriptive summary of all the predictors expected to be of interest is presented in Table 2. The mean age of patients is 32.95 ± 10.48 years. Majority of the patients are males (70.1%) followed by females (29.2%) and Eunuchs (0.7%) respectively. Smoking, alcoholic and Drug habits are recorded in 31.7%, 35.7% and 1.2% of the patients respectively. The mean number of visits for patients is 20.02 ± 11.33 while the mean weight at the

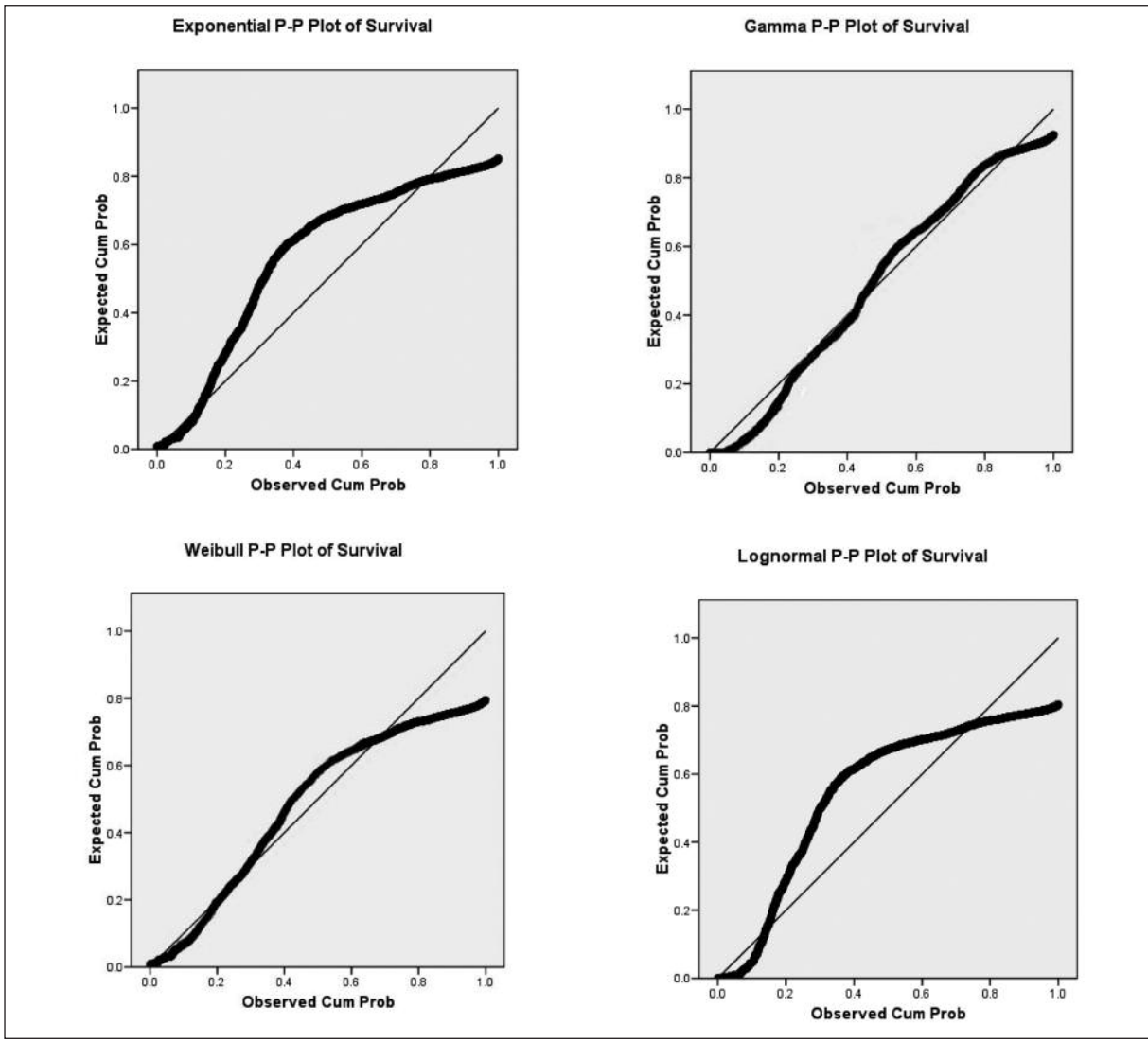

FIGURE 1: Probability plots to evaluate the fit of various Survival distributions.

TABLE 2: Effect of various covariates on survival time (in days) of AIDS patients using Censored Generalized Poisson distribution.

		Censored Generalized Poisson Model		
Variables	**Frequency(%)/ mean ± sd**	**Estimate**	**Std. error**	**p-value**
Intercept	-	2.5185	0.0605	0.000
Age	32.95 ± 10.48	-0.0007	0.0003	0.002
Gender				
Male	1184 (70.1%)	Ref.	-	-
Female	493 (29.2%)	0.0133	0.0055	0.007
Eunuch	12 (0.7%)	-0.2447	0.1345	0.034
Smoker	535 (31.7%)	-0.0472	0.0228	0.019
Alcohol	603 (35.7%)	-0.0593	0.0330	0.036
Drug	21 (1.2%)	-0.0064	0.1048	0.476
WHO Stage				
Stage 1	152 (9%)	Ref.	-	-
Stage 2	434 (25.7%)	-0.0263	0.0415	0.263
Stage 3	857 (50.7%)	-0.0447	0.0205	0.015
Stage 4	246 (14.6%)	-0.0654	0.0191	0.000
Change in CD4 count	149 (-15, 293)*	0.0006	0.00004	0.000
Opportunistic Infection				
None	929 (55%)	Ref.	-	-
Diarrhea	85 (5%)	-0.1598	0.0516	0.001
Tuberculosis	312 (18.5%)	-0.1843	0.0320	0.000
Others	363 (21.5%)	-0.1113	0.0287	0.000
No. of visits	20.02 ± 11.34	0.0559	0.0011	0.000
Weight (in Kg) at initiation of ART	47.09 ± 12.86	0.0083	0.0010	0.000
Dispersion Parameter	-	0.0579	0.0131	< 0.001
Log likelihood			-6805.14	
AIC			12644.28	

* Median (Quartile1, Quartile3)

initiation of ART is 47.09 ± 12.86 kg. The average improvement in the CD4 count of patients is recorded as 182.59 ± 262.51 cells/mm^3. Majority (50.7%) of patients are diagnosed as WHO stage 3. Tuberculosis or other infections are found to be the most predominant opportunistic infections. In order to evaluate the effect of each of these covariates on the survival time of AIDS patient undergoing ART, CGPR model is fitted and the results are shown in Table 2.

Firstly, we note that the estimate of dispersion parameter, = 0.058 is highly significant (p-value< 0.001) which justifies the use of CGPR over other count data models. We note that all the covariates included in the analysis have come out to be significant in explaining the survival time of AIDS patients undergoing ART. Age (in years) is negatively correlated with survival time. In fact, a unit increase in the age will cause the expected logarithm of survival time to decrease by 0.00078 days (p-value = 0.002). With respect to gender, the survival time (in days) of Eunuchs is the lowest. In comparison to males, females have 1.33% significant higher survival and eunuchs have 24.46% lower survival (p-values < 0.05). Smokers have 4.72% (p-value = 0.019) lower survival over non-smokers and Alcoholics have 5.93% (p-value = 0.036) lower survival as compared to non-alcoholics. However, drug users have no significant impact on the survival time of AIDS patient. The expected logarithm of survival time for WHO stage 1 and 2 patients is not significantly different. However, patients of

stage 3 and 4 have 4.46% and 6.54% lower survival in comparison to stage 1. Also, as the improvement in CD4 count increases by 1 unit, the logarithm of survival time increases by 0.00057 days. With reference to no opportunistic infection, Diarrhea patients have 15.98% lower survival. Similarly, Tuberculosis and other infection patients have 18.43% and 11.12% lower survival as compared to those without any infection. As the number of visits increase by unity, the expected logarithm of survival time increases by 0.056 days. A 1 kg increase in the weight of patient at the initiation of ART, increases the survival time in days by 0.82%.

The predicted values of expected logarithm survival time (in days) were calculated for all the patients. The mean of these values is found to be 8.395 which gives the estimated survival time of AIDS patients undergoing ART as = 4425 days i.e., 12.12 years.

DISCUSSION AND CONCLUSION

Gamma distribution is perhaps the most widely used distribution in the analysis of survival data. This distribution has been extensively used for studying various aspects of HIV/AIDS disease like prevalence,[41] incubation period distribution of AIDS,[42,43] life expectancy of patients who are newly diagnosed of HIV infection[37] and survival times of AIDS patients. In this paper, we have attempted to fit four distributions namely, Exponential, Gamma, Weibul and Lognormal to estimate the survival time of an AIDS patient after the initiation of ART. Of these four distributions, Gamma distribution provided the best fit and using this the mean survival time of AIDS patients after initiation of ART is estimated as 11.14 years. Authors[33,37,44] in past have estimated this survival time as anywhere between 1.18 years to 15 years. The variation in these estimates of survival times is not only due to the use of divergent statistical procedures but also because of advent of newer therapies. With the introduction of more potent new-generation first-line HAART regimens and advances in medical care for opportunistic infections and malignancies,[45,46] survival in HIV-positive patients is likely to see improvements.

On the other hand, the success of the ART also depends on the interaction of these antiretroviral drugs with the various physical and biological characteristics associated with the AIDS patient. Many authors have tried to assess the impact of these factors on the survivability of AIDS patients using different statistical techniques. However, to the best of our knowledge, no author has used CGPR model for analyzing these associations. When the survival time is measured in days, it's discrete nature allows the use as dependent variable in the CGPR model. Also for censored data, there is no means of verifying the over/under dispersion assumption and hence CGPR model gives us a platform for analyzing data without worrying too much about the assumptions. Consistent with literature,[11,44,47] age has a negative correlation with survival time of AIDS patient. A younger person undergoing ART is more likely to survive longer as compared to an older person i.e., old age is associated with high risk of disease progression. Also the survival time of females is significantly higher than that of males, while eunuchs record the lowest survival and this agrees with other findings.[48,49] Smokers and alcoholics have significantly lower survival time while drug users have no significant decrease in the survival time. WHO stages have a negative correlation with the survival time. While there is no significant difference in the survival time for patients in stages 1 and 2, it decreases for patients in stages 3 and 4. Also, the improvement in CD4 count is positively correlated with the survival time of patients which corroborate with results of other researchers.[44,48,50] CD4 count is an important marker of the progression of HIV infection and hence an improvement in it will certainly have positive effect on the survival time.

There is increasing evidence that opportunistic infection may affect subsequent HIV disease, perhaps by increasing HIV replication during the period of acute disease or by increasing cytokines that in turn impact on disease progression. HIV-mediated immunosuppression changes host control of the infectious agent resulting in disease and the disease process in turn activates HIV, hastening the rate of immunosuppression.[43] In our study, the patients

who had no opportunistic infection had the highest survival time. Amongst those who suffered such infections at some point or other in the study period, the lowest survival was seen in Diarrhea closely followed by Tuberculosis and other infections. Similar to the findings of authors,[11,47,51] our study established a positive correlation between weight and survival time. According to recent studies, ART regimens require 70-90% adherence in order to be effective.[52] However, sustaining adherence to antiretroviral therapy (ART) over the long term requires accurate and consistent monitoring. In our study, the a crude method to identify adherence to the therapy is captured through the number of visits made to the ART centre. As these visits increase, the survival time of AIDS patients also increases and this is in agreement with other studies.[52,53]

The advantage of using a CGPR model over the gamma right censored distribution is that the mean survival time of the AIDS patient undergoing ART can be estimated in the presence of significant prognostic factors. The impact of each of these factors on the survival time can be assessed and these results could be kept in mind while administering ART on the AIDS patients. For a new patient reporting to ART centre as well as existing patients, the characteristics may be included into the variables of CGPR model and hence the length and pattern of survival could be suitably predicted. Most of the statistical procedures till date, employ the Cox-proportional hazards for assessing the impact of prognostic factors on the survival time which is treated as a continuous variable. Such models depend heavily on the proportionality assumption and fail in most of the practical situations. Even though alternatives models have been developed to get over these assumptions, these models still consider survival time in a continuous sense. In such a scenario, the CGPR model with survival time taken as discrete variable provides a viable alternative which depends very little on the assumptions on model. Also, this model easily overrides the over/under dispersion criteria which cannot be exactly identified and verified for a censored data set. This study could be further improved by including the interactive effects amongst various prognostic factors in the model. Such a model would definitely yield a more complex equation with difficulty in attaining convergence of estimates, but will also provide more precise estimates. Also, in the parametric approach, we have considered only four popular lifetime distributions all of which, belong to exponential family. However, the results of such approach may be improved by considering a more general class of distributions which encompass not only, the ones considered in this article, but also cover a wider range of distributions. The limitations of our study arise from the fact that the analysis was done for data obtained from only one ART centre. Therefore a strict generalization of our findings at national level would require at most care and further conformity of multi-centre analysis.

Acknowledgements

This work was supported by grants to first author under R&D program of University of Delhi, India. We thank the reviewers for giving their valuable suggestions on this research article.

REFERENCES

1. Richman DD, Margolis DM, Delaney M, Greene WC, Hazuda D, Pomerantz RJ. The challenge of finding a cure for HIV infection. Science 2009;323(5919):1304-7.
2. Palella FJ Jr, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. N Engl J Med 1998;338(13): 853-60.
3. Leserman J, Jackson ED, Petitto JM, Golden RN, Silva SG, Perkins DO, et al. Progression to AIDS: the effects of stress, depressive symptoms, and social support. Psychosom Med 1999;61(3):397-406.
4. Weisser M, Rudin C, Battegay M, Pfluger D, Kully C, Egger M. Does pregnancy influence the course of HIV infection? Evidence from two large Swiss cohort studies. J Acquir Immune Defic Syndr Hum Retrovirol 1998;17(5): 404-10.
5. Rodriguez RA, Mendelson M, O'Hare AM, Hsu LC, Schoenfeld P. Determinants of survival among HIV-infected chronic dialysis patients. J Am Soc Nephrol 2003;14(5): 1307-13.
6. Bakanda C, Birungi J, Mwesigwa R, Ford N, Cooper CL, Au-Yeung C, et al. Association of aging and survival in a large HIV-infected cohort on antiretroviral therapy. AIDS 2011; 25(5):701-5.

7. Zachariah R, Fitzgerald M, Massaquoi M, Pasulani O, Arrnould L, Makombe S, et al. Risk factors for high early mortality in patients on antiretroviral treatment in a rural district of Malawi. AIDS 2006;20(18):2355-60.
8. Chen SC, Yu JK, Harries AD, Bong CN, Kololа-Dzimadzi R, Tok TS, et al. Increased mortality of male adults with AIDS related to poor compliance to antiretroviral therapy in Malawi. Trop Med Int Health 2008;13(4):513-9.
9. Zachariah R, Harries K, Moses M, Manzi M, Line A, Mwagomba B. et al. Very early mortality in patients starting antiretroviral treatment at primary health centres in rural Malawi. Trop Med Int Health 2009;14(7):713-21.
10. Kitahata MM, Gange SJ, Abraham AG, Merriman B, Saag MS, Justice AC, et al; NA-ACCORD Investigators. Effect of early versus deferred antiretroviral therapy for HIV on survival. N Engl J Med 2009;360(18):1815-26.
11. Grover G, Das R, Swain PK, Deka B. On the estimation of survival of HIV/AIDS patients on antiretroviral therapy using NPMLE method: an application to interval censored data. American Journal of Mathematics and Statistics 2013; 3(4):213-9.
12. Corbett EL, Watt CJ, Walker N, Maher D, Williams BG, Raviglione MC, et al. The growing burden of tuberculosis global trends and interactions with the HIV epidemic. Arch Intern Med 2003;163(9):1009-21.
13. Moore D, Liechty C, Ekwaru P, Were W, Mwima G, Solberg P, et al. Prevalence, incidence and mortality associated with tuberculosis in HIV-infected patients initiating antiretroviral therapy in rural Uganda. AIDS 2007;21(6):713-9.
14. Bica I, McGovern B, Dhar R, Stone D, McGowan K, Scheib R, et al. Increasing mortality due to end-stage liver disease in patients with human immunodeficiency virus infection. Clin Infect Dis (2001);32(3):492-7.
15. Cameron AC, Trivedi PK. Regression Analysis of Count Data (Econometric Society Monographs) (Book 30). 1sted. Cambridge (UK): Cambridge University Press; 1998. p.436.
16. Terza JV. A tobit-type estimator for the censored Poisson regression model. Economics Letter 1985;18(4):361-5.
17. Caudill SB, Mixon FG. Modeling household fertility decisions: estimation and testing censored regression models for count data. Empirical Economics 1995;20(2):183-96.
18. Frome EL. The analysis of rates using Poisson regression models. Biometrics 1983; 39(3):665-674.
19. Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. Stat Med 2004;23(1):51-64.
20. Winkelmann R, Zimmermann KF. Recent developments in count data modelling: theory and application. Journal of Economic Survey 1995;9(1):1-24.
21. Karlis D. A general EM approach for maximum likelihood estimation in mixed Poisson regression models. Statistical Modeling 2001; 1(4):305-18.
22. Consul PC. Generalized Poisson Distributions: Properties and Applications. 1sted. New York: Marcel Dekker Incorporated; 1989. p.302.
23. Consul PC, Famoye F. Generalized Poisson regression model. Comm Statist Theory Meth 1992;21(1):81-109.
24. Famoye F. Restricted generalized Poisson regression model. Comm Statist Theory Meth 1993;22(5):1335-54.
25. Famoye F, Wang W. Censored generalized Poisson regression model. Computational Statistics & Data Analysis 2004;46(3):547-60.
26. Carvour ML. Patterns and predictors of survival following an HIV/AIDS-related neurologic diagnosis-dissertation. University of Iowa, 2012.
27. Hambisa MT, Ali A, Dessie Y. Determinants of mortality among HIV positives after initiating antiretroviral therapy in Western Ethiopia: a hospital-based retrospective cohort study. ISRN AIDS 2013;2013:491601.
28. Malta M, Bastos FI, da Silva CM, Pereira GF, Lucena FF, Fonseca MG, et al. Differential survival benefit of universal HAART access in Brazil: a nation-wide comparison of injecting drug users versus men who have sex with men. J Acquir Immune Defic Syndr 2009; 52(5):629-35.
29. de Pinho AM, Santoro-Lopes G, Harrison LH, Schechter M. Chemoprophylaxis for tuberculosis and survival of HIV infected patients in Brazil. AIDS 2001;15(16):2129-35.
30. May M, Sterne J, Egger M. Parametric survival models may be more accurate than Kaplan-Meier estimates. BMJ 2003;326(7393): 822.
31. Schneider MF, Gange SJ, Williams CM, Anastos K, Greenblatt RM, Kingsley L, et al. Patterns of the hazard of death after AIDS through the evolution of antiretroviral therapy: 1984-2004. AIDS 2005;19(17):2009-18.
32. Sun J. A non-parametric test for interval censored failure time data with application to AIDS studies. Stat Med 1996;15(13):1387-95.
33. Grover G, Banerjee T. Estimation of survival times of HIV-1 infected children for doubly censored data. Electronic Journal of Applied Statistical analysis 2011;4(2):155-63.
34. Grover G, Shakeri N. Non parametric estimation of survival function of HIV+ patients with doubly censored data. J Commun Dis 2007; 39(1):7-12.
35. Nakhaee F, Law M. Parametric modelling of survival following HIV and AIDS in the era of highly active antiretroviral therapy: data from Australia. East Mediterr Health J 2011; 17(3):231-7.
36. King JT, Justice AC, Roberts MS, Chang CC, Fusco JS; Collaboration in HIV Outcomes Research-US Program Team. Long-term HIV/AIDS survival estimation in the highly active antiretroviral therapy era. Med Decis Making 2003;23(1):9-20.
37. Fang CT, Chang YY, Hsu HM, Twu SJ, Chen KT, Lin CC, et al. Life expectancy of patients with newly-diagnosed HIV infection in the era of highly active antiretroviral therapy. QJM 2007;100(2):97-105.
38. Lindsey JK, Jones B. Choosing among generalized linear models applied to medical data. Statist Med 1998;17(1):59-68.
39. Kaplan EL, Meier P. Non-parametric estimation from incomplete observations. Journal of the American Statistical Association 1958;53:457-81.
40. Chow NT, Steenhard D. A flexible count data regression model using SAS® PROC NLMIXED. SAS Global Forum: Statistics and Data Analysis 2009;250:1-14.
41. Joint United Nations Programme on HIV/AIDS., Nations Unies. Division de la population, ONUSIDA., UN. Population Division. The Demographic Impact of HIV/AIDS, Report on the Technical Meeting. New York: UN; 1999. p. 77.
42. Phillips AN, Sabin CA, Elford J, Bofill M, Janossy G, Lee CA. Use of CD4 lymphocyte count to predict long term survival free of AIDS after HIV infection. BMJ 1994;309(6950):309-13.
43. Osmond DH. Epidemiology of Disease Progression in HIV. HIV In Site Knowledge Base Chapter 1998.
44. Saah AJ, Hoover DR, He Y, Kingsley LA, Phair JP. Factors influencing survival after AIDS: report from the Multicenter AIDS Cohort Study (MACS). J Acquir Immune Defic Syndr 1994;7(3):287-95.
45. Lampe FC, Gatell JM, Staszewski S, Johnson MA, Pradier C, Gill MJ, et al. Changes over time in risk of initial virological failure of combination antiretroviral therapy: a multicohort analysis, 1996 to 2002. Arch Intern Med 2006;166(5):521-8.
46. Crum NF, Riffenburgh RH, Wegner S, Agan BK, Tasker SA, Spooner KM, et al. Comparisons of causes of death and mortality rates among HIV-infected persons: analysis of the pre-, early, and late HAART (highly active antiretroviral therapy) eras. J Acquir Immune Defic Syndr 2006;41(2):194-200.
47. Bachani D, Garg R, Rewari BB, Hegg L, Rajasekaran S, Deshpande A, et al. Two year treatment outcomes of patients enrolled in India's national first-line antiretroviral therapy programme. Nat Med J India 2012;23(1):7-12.
48. Ghate M, Deshpande S, Tripathy S, Godbole S, Nene M, Thakar M, et al. Mortality in HIV infected individuals in Pune, India. Indian J Med Res 2011;133:414-20.
49. Donnelly CA, Bartley LM, Ghani AC, Le Fevre AM, Kwong GP, Cowling BJ, et al. Gender differences in HIV-1 RNA viral loads. HIV Med 2005;6(3):170-8.
50. May M, Sterne JA, Sabin C, Costagliola D, Justice AC, Thiébaut R, et al. Prognosis of HIV-1 infected patients up to 5 years after initiation of HAART: collaborative analysis of prospective studies. AIDS 2007;21(9):1185-97.
51. Reda AA, Biadgilign S, Deribew A, Gebre B, Deribe K. Predictors of change in CD4 lymphocyte count and weight among HIV infected patients on anti-retroviral treatment in Ethiopia: a retrospective longitudinal study. PLoS ONE 2013;8(4):e58595.
52. Reda AA, Biadgilign S. Determinants of adherence to antiretroviral therapy among HIV-infected patients in Africa. AIDS Research and Treatment 2012;2012:574656.
53. Nachega JB, Mills EJ, Schechter M. Antiretroviral therapy adherence and retention in care in middle-income and low-income countries: current status of knowledge and research priorities. Curr Opin HIV AIDS 2010;5(1):70-7.