Cure Fraction Model for the Estimation of Long-term Survivors of HIV/AIDS Patients under Antiretroviral Therapy

MK Varshney\textsuperscript{1}, Gurprit Grover\textsuperscript{2}, Vajala Ravi\textsuperscript{3}, Arpan Kumar Thakur\textsuperscript{4}

\textsuperscript{1}Department of Statistics, Hindu College, University of Delhi, Delhi-110007, India.\textsuperscript{2}\textsuperscript{4}Department of Statistics, University of Delhi, Delhi-110007, India.\textsuperscript{3}Department of Statistics, Lady Shri Ram College, University of Delhi, Delhi-110024, India.

\textbf{Keywords:} AIDS, ART, Cure rate model, CD4, MCMC

The primary aim of this research is to estimate the proportion of long-term survivors among HIV/AIDS patients receiving Antiretroviral therapy (ART). A cure fraction model has been used to accomplish the same. Although, cure fraction models are extensively being used in oncology for modeling survival time data with long-term survivors, but there are minimal research that considers this model in HIV/AIDS set up. Here, we have defined survival time under the purview of CD4 cell counts. CD4 is considered to be disease marker for HIV/AIDS patients. Bayesian Analysis of the various mixture and non-mixture cure fraction models under exponential, generalized exponential, Raleigh, Weibull, Exponentiated Weibull distributions are exemplified using a real data set. Effect of prognostic factors like baseline CD4, age, sex, medication on cure fraction are studied. The MLE’s are obtained using Gibbs sampling techniques with MCMC method in Open BUGS package. Convergence diagnostic (like trace plots, density plots & MC errors) are used to detect in any unexpected anomalies in MCMC output. The DIC (Deviance information criterion) has been utilized to compare the efficiency of different models. A real-life data set from the ART center of RML hospital, Delhi, India are taken for this study.

\textbf{Introduction}

In the contemporary world of medical science, there are plethora of research and development going on to enhance the lifespan of the patients suffering from all kinds of diseases including deadly diseases like cancer and HIV/AIDS.

As a result of this many patients’ suffering from certain type of cancer get cured permanently. So, the population of patients becomes heterogeneous in nature, i.e. a large proportion of population who responded positively to the treatment set to be free of any sign and symptoms of the disease are treated as cured or long-term survivors or immunes. And other remaining proportion of the population on whom treatment has no significant impact, i.e. who develop a recurrence of the disease are considered as susceptible or un-cured.

Of late it has been observed that universally used survival models like proportional hazard (PH) are based on unstated assumption that every subject or individual is susceptible to the event of interest.\textsuperscript{1} Nevertheless, this assumption may not be valid in present time because now a large proportion of subjects get cured, or at-least survives a sufficiently long period of time.

\textbf{Corresponding Author:} Arpan Kumar Thakur, Department of Statistics, University of Delhi, Delhi- 110007, India.
\textbf{E-mail Id:} arpankmr3@gmail.com
\textbf{Orcid Id:} https://orcid.org/0000-0001-7052-8351
\textbf{How to cite this article:} Varshney MK, Grover G, Ravi V et al. Cure Fraction Model for the Estimation of Long-term Survivors of HIV/AIDS Patients under Antiretroviral Therapy. \textit{J Commun Dis} 2018; 50(3): 1-10.
Boag JW propounded cure rate model in terms of mixture models that encompasses an element representing the proportion of long-term survivors (immunes) in the population and a distribution representing the survival experiences of the susceptible (called latency distribution).

Cure models can more adequately describe survival trends when there is a non-negligible proportion of patients alive at the end of follow-up with a plateau at the end of the survival curve. In this case, the “cured” proportion from cure models provides an estimate of the proportion of patients who will not fail during follow-up, which may be a clinically relevant value.

Cure models can be a useful alternative to the standard Cox proportional hazards models for data with heavy censoring, for several reasons. First, the assumption of proportional hazards can fail when survival curves have plateaus at their tails. Second, survival plots with long plateaus may indicate heterogeneity within a patient population that can be useful to describe the data explicitly. Cure models allow us to investigate which covariates are associated with either short-term or long-term effects. For example, cure models can allow us to evaluate whether a new therapy is associated with an increase or decrease in the probability of being a long-term survivor or an improvement or detriment in survival for those who are not long-term survivors. Cure fraction models basically focus on the proportion of patients who survive for a long period of time following disease. Additionally, these models focus on the probability of survival of un-cured patients up to a given point in time.

There are plenty of applications of this model in the medical field. For example, Struthers used a cure rate model to study the incubation time of AIDS for HIV-I positive patients and results were better as compared to standard survival models. Tournoud M and Ecochard R studied this model for analyzing an HIV-I mother-to-child transition (MTCT) data set and for nosocomial urinary tract infections data set. They elaborately studied to model the delay from contamination to the event, to take in to account successive and multiple exposure to the infectious agent, to identify which part of infection is due to each exposure occasion. Jolayemi ET used this model to study tuberculosis (TB) data with HIV co-infection. The cure rate model can also be used in reliability.

Research Motivation and Expected Outcome of Our Work

Many a places in statistical fraternity it is mentioned time to time that cure model is an underutilized statistical tool in spite of the fact that it is well developed in statistical literature. Nevertheless, there are arguably plenty of publications in oncology using cure models but when we searched on popular and quality databases we could found countably quite a few research publications on cure models in the HIV/AIDS, namely Struthers CA et al., Tournoud M. So, the ultimate aim of this work is to build a cure model using HIV/AIDS survival data under Bayesian set up to enhance the applicability of cure models and elucidate the same to the non-statistician like public health researcher, psychiatrist, educationist, criminologist and reliability engineers.

Remaining part of the paper is systematized as follows. In the immediate section of “Material and Methods”, data set is described as a material and we briefly delineate types of cure models, specifically standard cure rate model and promotion time cure rate model. In Section 3 we have provided results based on our data set, and ultimately discussion and conclusion have been accorded in section.

Material and Methods

Acquired Immune Deficiency Syndrome (AIDS) is a leading pandemic in terms of incidence and prevalence, Anti-Retroviral Therapy (ART) is considered to be the best treatment for HIV/AIDS.

As per WHO guidelines decision on initiation of ART must be based on two tests viral load test and CD4 count test. Viral load test counts the number of HIV virus in the blood while CD4 test count the number of T cells that gives immunity against infections but in India, due to economic constraints, NACO suggest ART centers across the country to initiate treatment based on CD4 count only.

Due to improved ART facilities, there is a noticeable improvement in the lifespan of patients suffering from the disease. The impact of predominant disease marker CD4 on patient’s survival is studied by many researchers including Grover G et al. and Williams BG et al. As a result of the improvement in treatment facilities, survival time data from HIV/AIDS patients became heterogeneous i.e. data consists of two kinds of subjects, those who survived for the longer duration of time and those who succumb to death.

CD4 count is closely related with patient’s survival, i.e. patients who have high CD4 count probably would survive longer than those who have low CD4 count. Similarly, increasing CD4 count signals longer survival and decreasing CD4 count indicate deteriorating health and survival. To extract this intricate relation, we have defined our event of interest slight differently, unlike in the widely used survival model here we define survival time with respect to CD4 count. The time (in days) at which CD4 count starts decreasing is taken as survival time and corresponding occurrence point as an event of interest. List of prognostic factors that are affecting survival time are the age at enrolment, baseline CD4 count, sex, treatment combination.

Data are retrospectively collected from Dr. Ram Manohar
Lohia Hospital, New Delhi of patients who were on ART during the period April 2004 to December 2014. Inclusion criteria are, the patient should be adult, should have a baseline CD4 count, and have periodic CD4 count available.

On applying complete-case analysis approach we are left with a sample of 545 patients only, of which 463 survival times of our time censored at the termination of study period and CD4 count of remaining 82 patients start decreasing before the termination of study even though they are also on ART. This signifies that 84.95% of patients got benefited with ART and they are surviving for a longer duration of time whilst remaining 15.05% patient’s health deteriorates or improved for the very short duration of time. Figure 1 gives Kaplan-Meier estimate of the survival function for the data set; it can be observed that there is a plateau in the right near 0.86 which suggest that cure fraction model is a reasonable alternative to the widely used survival models.

**Standard Cure Rate Model**

Let C be the probability of an HIV/AIDS patient being a long-term survivor and (1 - C) be the probability of a patient being susceptible. Then, the overall population survival function at any time t can be written as:

\[ S(t) = C + (1 - C) \times S_u(t) \]  \hspace{1cm} (1)

where, \( S(t) \) is the survival function of the overall population which may be assumed to follow some lifetime distribution and here we use exponential, gamma, Weibull, Rayleigh, generalized exponential, and exponentiated Weibull distribution to estimate cure fraction \( C \).

Probability density function \( f(t) \) of the overall population may be written as:

\[ f(t) = (1 - C) \times f_u(t) \]  \hspace{1cm} (2)

where \( f_u(t) \) is the probability density function of the susceptible population.

Now let \((t_i, \delta_i)\) be the observed data of size \( n \), where \( t_i \) is the survival time of the \( i^{th} \) patient and \( \delta_i \) is censoring indicator variable which is defined as follows: \( \delta_i = 0 \) for censored observation and \( \delta_i = 1 \) for uncensored observation \((i = 1, 2, \ldots, n)\).

Thereupon, the individual patient’s contribution in the likelihood is given by:

\[ L_i = [f(t_i)]^{\delta_i} \times [S(t_i)]^{(1-\delta_i)} \]  \hspace{1cm} (3)

So, complete data likelihood is given by:

\[ L = \prod_{i=1}^{n} L_i = \prod_{i=1}^{n} [(1 - C) f_u(t_i)]^{\delta_i} \times [C + (1 - C) S_u(t_i)]^{1-\delta_i} \]  \hspace{1cm} (4)

Upon maximizing the complete data likelihood in equation (4) in WinBUGS software package using Gibbs sampling approach we can get the estimated values of the parameters. Different models resulting from equation (1) are compared using DIC values and presented in Table 2. Farewell VT\textsuperscript{12} used mixture cure rate model to estimate long-term survivors in toxicological experiment, Peng and Dear\textsuperscript{13} developed non-parametric cure rate model, Sy and Taylor\textsuperscript{14} studied estimation procedures in Cox-proportional hazard rate model, Tsodikov et al.\textsuperscript{15} proposed an alternative to mixture cure model. Maller and Zhou\textsuperscript{16} advanced testing procedures for censored cure fraction model, Yamaguchi\textsuperscript{17} combined accelerated failure time model and cure fraction model to study employment data of Japan, Chen et al.\textsuperscript{18} provided a Bayesian counterpart of cure fraction model that is now most widely used in the literature. Bivariate cure rate models are derived by Chatterjee N\textsuperscript{19}, Exponentiated cure rate model by Kannan N\textsuperscript{20}, Generalized Gompertz cure rate model by Swain PK\textsuperscript{21}. Similarly, multivariate cure rate models are derived by Chen MH.\textsuperscript{22}

**Promotion Time Cure Rate Model**

To accentuate the biological phenomena that after the initiation of ART a patient might go through biological changes like an opportunistic infection that may result in reduction of CD4 count which consequently leads to a shorter lifetime. Survival time may also get shrink due to negligence at patient’s side. To accomplish the same a non-mixture cure rate model is applied which is defined as:

\[ S(t) = \exp[(ln C) \times F_u(t)] \]  \hspace{1cm} (5)

where ‘C’ is the probability of an HIV/AIDS patient being a long-term survivor and \( F_u(t) \) is the distribution function for susceptible (uncured) patients, in our case we have taken the same six distributions as in mixture cure rate model. Since hazard function \( h(t) = f(t)/S(t) \), it can be written as \(- (ln p) f_u(t)\). The model given in Eq. (5) is also called as bounded cumulative hazard (BCH) model, in parametric setup, there is no difference whether model (1) or (5) is used as a basis for estimation of \( C \), but for non-parametric, there is difference, Tsodikov.\textsuperscript{15}

For the model (5), \( i^{th} \) \( (i = 1, 2, \ldots, n) \) individual’s contribution to the likelihood function is given by:

\[ L_i = [h(t_i)]^{\delta_i} \times [S(t_i)]^{(1-\delta_i)} \times \exp[(ln C) \times F_u(t_i)] \]  \hspace{1cm} (6)

So, complete data likelihood is given by:

\[ L = \prod_{i=1}^{n} L_i = \prod_{i=1}^{n} [(- (ln p) f_u(t_i))]^{\delta_i} \times \exp[(ln C) \times F_u(t_i)] \]

All the models are compared using DIC values and the corresponding estimate of cure fraction are reported in Table 3.

**Results**

In order to analyze the data, set we first plotted the Kaplan-Meier survival curve to check the suitability of the cure
fraction model. As the shape of overall survival curve is like plateau which connotes that there is scope for cure rate model. Figure 1 attains flatness/stagnancy at level 0.86 after approximately 3100 days (8.5 years) that means that 86% of the HIV patients are surviving at-least up to 8.5 years, in terms of CD4 level we can say that 86% of the patients are having their CD4 level non-decreasing at-least up to 8.5 years after the initiation of the ART. It is known that CD4 cell count is most important determinant for the survival of HIV/AIDS patients so, 86% of non-decreasing CD4 cell count indicates that 86% of HIV-I patients survives at-least 8 years, which is comparable to study reported by Grover G et al.  

To see whether there is any effect of gender of the patient on the proportion of long-term survivors we have plotted K-M survival curve with respect to sex (male or female). The plot is given in Figure 2 that can be interpreted as about 94% of the females have their non-decreasing CD4 level and it attained at 7.25 (approx.) and about 85 % of the males have their non-decreasing CD4 level, and flatness attained at 8.2 years. If we fix a cut-off year say seven years then at that level about 90% males and 95% of females are having non-decreasing CD4 cell count. So, at any given time larger proportion of females are having non-decreasing CD4 cell count.
Among the 545 filtered HIV/AIDS patients 333 patients were given Virocomb-N combination medicine and remaining 212 were given virolans-30, Tenolam + Efravinez-600 etc. We have categorized our data set in two groups, one who were prescribed virocomb-N medicine and ‘others’ group consists of all patients who have taken medicine other than virocomb-N. Based on the aforementioned categorization we have plotted K-M survival curve given in Figure 3 from which it is evident that Virocomb-N treatment combination is giving better results as compared to others treatment. Since virocomb-N treatment is making more long-term survivors than the others treatment. Additionally, observe that between 1 to 2-year period others treatment is giving a greater number of long-term survivors.

For observing the effect of baseline CD4 count on prospective progression/retrogression of the CD4 count, on grouping baseline CD4 count in two parts, one for baseline CD4 count between 0-200 cells/µL and one for CD4 count more than 200 cells/µL. Survival curve with respect to two group of baseline CD4 count is given in Figure 4, it is observed that patients who have high CD4 cell count (200+) have more chance of surviving for a longer duration of time. Similarly, for patients who have low CD4 cell count (0-200) are more prone to death due to HIV infection.⁹

All the four figures drawn above reinforced our surmise to go for the cure fraction model and that we call our stimulus or motivation. So, from next paragraph onwards our whole results will linger around the cure fraction model, parameter estimation and model comparison.

To study the mixture and non-mixture models under Bayesian setup, a uniform U (0, 1) prior distribution for cure fraction ‘C’ is assumed, for scale and shape parameters Gamma (1, 1) priors are assumed. Additionally, prior independence among C and all other parameters are presumed.
Different models (viz. Exponential, gamma, Weibull, Rayleigh, generalized exponential, exponentiated Weibull distribution) without considering cure fraction ‘C’ (in this sense it is standard survival models) along with their DIC values are presented in Table 1.

Table 1. Parameter estimates considering standard models

| Model     | Mean    | Std. dev  | Error   | Median  | CI             | DIC  |
|-----------|---------|-----------|---------|---------|----------------|------|
| Exponential |        |           |         |         |                |      |
| $\lambda$ | 0.01565 | 0.002001  | 2.12E-05| 0.01558 | (0.01202, 0.01986)| 612.5 |
| Gamma     |         |           |         |         |                |      |
| $\gamma$  | 3.935   | 1.235     | 0.20356 | 4.693   | (1.057, 4.497)  | 498.6 |
| Rayleigh  |         |           |         |         |                |      |
| $\alpha$  | 0.00207 | 2.71E-04  | 3.07E-06| 0.002056| (0.001577, 0.002646)| 665.2 |
| $\gamma$  |         |           |         |         |                |      |
| Weibull   |         |           |         |         |                |      |
| $\alpha$  | 0.02103 | 0.01098   | 8.81E-04| 0.01961 | (0.01191, 0.0333) | 613.8 |
| $\gamma$  | 0.8773  | 0.1252    | 0.007548| 0.8806  | (0.6558, 1.097)  | 665.2 |
| General exponential |     |           |         |         |                |      |
| $\alpha$  | 0.01256 | 0.005722  | 3.94E-04| 0.01264 | (0.00001781, 0.02364) | 613.9 |
| $\beta$   | 0.2706  | 0.1074    | 0.009969| 0.2598  | (0.205, 0.3873)  |      |
| $\gamma$  | 0.8985  | 0.1708    | 0.01712 | 0.9367  | (0.5314, 1.149)  | 508.5 |

Table 2. Parameter estimates under mixture cure models

| Model     | Mean    | Std. dev  | Error   | Median  | CI             | DIC  |
|-----------|---------|-----------|---------|---------|----------------|------|
| Exponential |        |           |         |         |                |      |
| $\lambda$ | 0.118   | 0.05822   | 0.003773| 0.1165  | (0.01688, 0.2319)| 604.9 |
| C         | 0.7423  | 0.1889    | 0.01579 | 0.8092  | (0.05824, 0.8795)|      |
| Gamma     |         |           |         |         |                |      |
| $\gamma$  | 4.322   | 1.351     | 0.2396  | 4.929   | (1.07, 4.987)   | 366.9 |
| C         | 0.2747  | 0.3173    | 0.0567  | 0.2207  | (0.005497, 0.9125)|      |
| Rayleigh  |         |           |         |         |                |      |
| $\alpha$  | 0.02218 | 0.03493   | 0.002727| 0.002638| (0.001767, 0.06923)|      |
| $\gamma$  | 0.3717  | 0.3843    | 0.03846 | 0.1136  | (0.005885, 0.047)|      |
| Weibull   |         |           |         |         |                |      |
| $\alpha$  | 0.1216  | 0.05228   | 0.003677| 0.1227  | (0.02619, 0.2268)| 603.8 |
| $\gamma$  | 1.111   | 0.189     | 0.01244 | 1.107   | (0.7716, 1.461) |      |
| C         | 0.7803  | 0.1631    | 0.01292 | 0.8414  | (0.2509, 0.8945)|      |
| General exponential |     |           |         |         |                |      |
| $\alpha$  | 0.844   | 0.3989    | 0.06939 | 0.5895  | (0.45, 1.561)   | 606.9 |
| $\lambda$ | 0.2379  | 0.4248    | 0.07254 | 0.137   | (0.06216, 2.218)|      |
| C         | 0.8455  | 0.07227   | 0.01176 | 0.8596  | (0.5806, 0.8928)|      |
| Exp. Weibull |       |           |         |         |                |      |
| $\alpha$  | 0.0259  | 0.1221    | 0.01206 | 5.73E-05| (0.00001402, 0.565)| 318.3 |
| $\beta$   | 0.01662 | 0.04388   | 0.004337| 0.006603| (0.001304, 0.221)|      |
| $\gamma$  | 3.422   | 1.01      | 0.1014  | 3.776   | (0.6107, 4.21)  |      |
| C         | 0.8568  | 0.01558   | 0.001447| 0.8517  | (0.84, 0.9003)  |      |
For all the Bayesian analysis we took first 15,000 samples as ‘Burn-in’ sample for the elimination of the effect of the initial values, after this we simulated 100,000 Gibbs samples retaining every 100th sample that amounts to give a final sample of size 1000. All the MCMC estimates of parameters are based on 1000 sample retained finally. Convergence was monitored using a trace plot.

From Table 1, we can see that exponentiated Weibull distribution is giving least DIC value (479), so this model can be considered to be the best model among the six fitted models.

Now, we fit the same set of models under cure fraction ‘C’ (Table 2 for mixture models and Table 3 for non-mixture models). Based on DIC values we conclude that for our dataset performance of mixture and non-mixture models are almost identical.

Since exponentiated Weibull mixture model is giving least DIC value, therefore, it is selected as the best model with estimated cure fraction 88% so, for checking the impact of prognostic factors on the cure proportion ‘C’ (Table 4) we deliberately took the same selected model and for this, regression equations would be:

\[
\begin{align*}
\alpha &= \beta_0 \exp (\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 - \xi)
\end{align*}
\]

\[
\ln (C/(1 - C)) = \alpha_0 + \alpha_1 (X_1 - \xi) + \alpha_2 (X_2 - \xi) + \alpha_3 (X_3 - \xi) + \alpha_4 (X_4 - \xi)
\]

Gamma prior distribution is used for \(\beta_0\) and \(\alpha_0\) and normal prior \(N(0, 100)\) distribution for all other regression parameters. Simulated samples are generated for the joint posterior distribution by Markov Chain Monte Carlo (MCMC) using Gibbs sampling. Samples are generated to get the posterior values of the parameters.

### Table 3. Parameter estimates under non-mixture cure models

| Model        | Mean  | Std. dev | Error | Median | CI            | DIC  |
|--------------|-------|----------|-------|--------|---------------|------|
| Exponential  |       |          |       |        |               |      |
| \(\lambda\) | 0.1267| 0.1467   | 0.0133| 0.107  | (0.01929, 0.2577) | 605.5|
| \(C\)       | 0.7764| 0.1032   | 0.008107| 0.8067| (0.4412, 0.8799)  |      |
| Gamma        |       |          |       |        |               |      |
| \(y\)       | 4.741 | 1.892    | 0.3363| 5.587  | (1.144, 6.844)  | 374.3|
| \(C\)       | 0.2215| 0.3727   | 0.06625| 0.01573| (0.0005058, 0.9154) |      |
| Rayleigh     |       |          |       |        |               |      |
| \(a\)       | 0.05443| 0.03152 | 0.001877| 0.05237| (0.03403, 0.07363) | 625.7|
| \(C\)       | 0.8773| 0.04285 | 0.003029| 0.8815| (0.8429, 0.9101)  |      |
| Weibull      |       |          |       |        |               |      |
| \(a\)       | 0.1131| 0.04661 | 0.002662| 0.1114| (0.02975, 0.2117) | 605.2|
| \(y\)       | 1.151 | 0.2006   | 0.01346| 1.157  | (0.7919, 1.483)  |      |
| \(C\)       | 0.8048| 0.1261   | 0.009291| 0.8443| (0.4162, 0.8949)  |      |
| Gen. exponential | | | | | | |
| \(a\)       | 3.214 | 0.6779   | 0.03923| 3.167  | (2.008, 4.673)  | 605.9|
| \(\lambda\) | 0.7677| 0.07149  | 0.004502| 0.7655| (0.6366, 0.9021)  |      |
| \(C\)       | 0.02319| 0.04631| 0.004126| 0.01402| (0.0005797, 0.077727) |      |
| Exp. Weibull |       |          |       |        |               |      |
| \(a\)       | 0.00476| 0.0278  | 0.001992| 5.23E-05| (0.00003794, 0.038222) | 328.7|
| \(\beta\)   | 0.03728| 0.1274  | 0.01247| 0.001696| (0.00003315, 0.5306) |      |
| \(y\)       | 3.303 | 0.9443   | 0.09479| 3.844  | (0.7785, 3.884)  |      |
| \(C\)       | 0.8858| 0.07424 | 0.007199| 0.9066| (0.61, 0.9165)   |      |
Monte Carlo (MC) error measures the variation of the mean of the parameter of interest due to simulation. Since, in our all cases MC errors are low as compared with corresponding estimated posterior standard deviations, which suggest that the posterior mean was estimated with high precision. By increasing the number of iterations we can decrease the MC error, at the same time posterior standard deviation will be relatively stable. Boxplot of cure fraction C in Figure 5 is slightly different from the usual one. Limits of each box show the posterior quartile, while bar the posterior mean (by default it is mean, but we can change it to median also).

Table 4. Parameter estimates under non-mixture cure model for exponentiated Weibull distribution with covariates

| Parameter | Mean  | Std. dev. | MC_error | Median  | CI                  |
|-----------|-------|-----------|----------|---------|---------------------|
| $a_0$     | 0.9982| 0.9994    | 0.003176 | 0.9991  | (-0.9675, 2.953)   |
| $a_1$     | -4.08E-04 | 0.1005  | 3.06E-04 | -4.90E-04 | (-0.1969, 0.1967) |
| $a_2$     | -1.40E-05 | 0.1001  | 3.33E-04 | -1.14E-04 | (-0.1958, 0.1972) |
| $a_3$     | -7.83E-04 | 0.1003  | 3.17E-04 | -0.00111 | (-0.1976, 0.1947) |
| $a_4$     | -1.19E-04 | 0.09973 | 3.02E-04 | -2.45E-04 | (-0.1944, 0.1957) |
| $b_0$     | 1.003 | 1.001     | 0.003255 | 0.6951  | (0.02525, 3.702)   |
| $b_1$     | -2.26E-04 | 0.09993 | 3.24E-04 | -1.50E-04 | (-0.196, 0.195)   |
| $b_2$     | 3.37E-04 | 0.09994 | 3.03E-04 | 4.41E-04 | (-0.1962, 0.1959) |
| $b_3$     | -2.81E-04 | 0.1001  | 3.09E-04 | -1.48E-04 | (-0.1971, 0.1942) |
| $b_4$     | -5.29E-04 | 0.1001  | 3.25E-04 | -8.01E-04 | (-0.1978, 0.1954) |
| $\beta$   | 0.999 | 1         | 0.003277 | 0.6914  | (0.02555, 3.698)   |
| $\gamma$  | 1     | 1.001    | 0.003081 | 0.693   | (0.02554, 3.702)   |

Figure 6. (a) Posterior density plot; (b) trace plot; (c) history plot

Checking Convergence

The convergence refers to whether the algorithm has its equilibrium (target) distribution or not. If the algorithm converges, then the generated samples come from the precisely assumed (target) distribution. Therefore, monitoring the convergence is a must for getting results from the posterior distribution of interest. There are several ways of monitoring the convergence first and foremost is to observe MC error, for our dataset small error indicates we have
high precision consequently the convergence. Another way is to draw the trace plot: the plot of iterations against generated values. Since our trace plot [Figure 6(b)] is free from strong periodicities and all generated values are within a parallel zone so, we assume a convergence for our model parameters. Density plot [Figure 6(a)] produces an approximate visual Kernel estimate of the posterior density or probability function. History plot [Figure 6(c)] draws a full trace plot of all stored values Ntzoufras I. In all the model algorithms convergence was achieved smoothly.

Discussion and Conclusion

The foremost impetus of this research was to estimate the proportion of long-term survivors for the HIV/AIDS data set. The stimulus was the fact that in HIV/AIDS there could have the existence of cure fraction and obviously the covariate effect on cure fraction too. As an alternative to the standard survival model, cure rate model has been employed which is gaining accreditation from many researchers of diverse fields. It has been shown that traditional survival distributions (without considering cure fraction) do not perform well, as cure fraction model do. So, survival models that consider cure fraction can greatly be used for analyzing HIV/AIDS life-time data. Additionally, here we have used Bayesian inference for parameter estimation and this technique is supplemented by various modern software packages like OpenBUGS, JAGS (Just Another Gibbs Sampler) etc. that only need survival time distribution and some prior distribution of parameters under investigation, the latter one can be taken from experience of medical practitioner regarding proportion of long-term survivors that will immensely improve the precision. This article strives to extend the use of cure rate model from oncology to HIV/AIDS. Point of departure was not merely limited to the application of this model to HIV/AIDS data set in fact we have elicited the information provided by the CD4 count in defining our event of interest and consequently the survival time. Moreover, we have provided the aftermath from strong periodicities and all generated values are

Conflict of Interest: None

References

1. Cox DR. Regression models and life-tables (with discussion). Journal of the Royal Statistical Society, Series B (Methodological) 1972; 34: 187-220.
2. Boag JW. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. Journal of the Royal Statistical Society, Series B (Methodological) 1949; 11(1): 15-53.
3. Taweab F, Ibrahim NA. Cure rate models: a review of recent progress with a study of change-point cure models when cured is partially known. J. Appl. Sci. 2014; 14: 609-16.
4. Struthers CA, Farewell VT. A mixture model for time to AIDS data with left truncation and an uncertain origin. Biometrika 1989; 76(4): 814-7.
5. Tournoud M, Ecochard R. Application of the promotion time cure model with time-changing exposure to the study of HIV/AIDS and other infectious diseases. Statistics in Medicine 2007; 26(5): 1008-21.
6. Jolayemi ET. Modeling of a cure rate model for TB with HIV co-infection. The Pacific Journal of Science and Technology 2017; 18(1): 288-99.
7. Nelson WB. Applied life data analysis. John Wiley & Sons, 2005.
8. Grover G, Vajila R, Swain PK. On the assessment of various factors effecting the improvement in CD4 count of aids patients undergoing antiretroviral therapy using generalized Poisson regression. Journal of Applied Statistics 2015; 42(6): 1291-305.
9. Williams BG, Korenromp EL, Gouws E, et al. HIV infection, antiretroviral therapy, and CD4+ cell count distributions in African populations. The Journal of Infectious Diseases 2006; 194(10): 1450-8.
10. Burton A, Altman DG. Missing covariate data within cancer prognostic studies: a review of current reporting and proposed guidelines. British Journal of Cancer 2004; 91(1): 4.
11. Berkson J, Gage RP. Survival curve for cancer patients following treatment. Journal of the American Statistical Association 1952; 47(259): 501-15.
12. Farewell VT. The use of mixture models for the analysis of survival data with long-term survivors. Biometrics 1982; 38(4): 1041-6.
13. Peng Y, Dear KB. A nonparametric mixture model for cure rate estimation. Biometrics 2000; 56(1): 237-43.
14. Sy JP, Taylor JM. Estimation in a Cox proportional hazards cure model. Biometrics 2000; 56(1): 227-36.
15. Tsodikov AD, Ibrahim JG, Yakovlev AY. Estimating cure rates from survival data: an alternative to two-component mixture models. Journal of the American Statistical Association 2003; 98(464): 1063-78.
16. Maller RA, Zhou S. Testing for the presence of immune or cured individuals in censored survival
data. *Biometrics* 1995; 51(4): 1197-205.

17. Yamaguchi K. Accelerated failure-time regression models with a regression model of surviving fraction: an application to the analysis of “permanent employment” in Japan. *Journal of the American Statistical Association* 1992; 87(418): 284-92.

18. Chen MH, Ibrahim JG, Sinha D. A new Bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association* 1999; 94(447): 909-19.

19. Chatterjee N, Shih J. A bivariate cure-mixture approach for modeling familial association in diseases. *Biometrics* 2001; 57(3): 779-86.

20. Kannan N, Kundu D, Nair P, Tripathi RC. The generalized exponential cure rate model with covariates. *Journal of Applied Statistics* 2010; 37(10): 1625-36.

21. Swain PK, Grover G, Goel K. Mixture and Non-Mixture Cure Fraction Models Based on Generalized Gompertz Distribution under Bayesian Approach. *Tatra Mountains Mathematical Publications* 2016; 66(1): 121-35.

22. Chen MH, Ibrahim JG, Sinha D. Bayesian inference for multivariate survival data with a cure fraction. *Journal of Multivariate Analysis* 2002; 80(1): 101-26.

23. 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.

24. Achcar JA, Coelho-Barros EA, Mazucheli J. Cure fraction models using mixture and non-mixture models. *Tatra Mountains Mathematical Publications* 2012; 51(1): 1-9.

25. Ntzoufras I. *Bayesian modeling using WinBUGS*. John Wiley & Sons, 2011.

*Date of Submission: 2018-08-22*

*Date of Acceptance: 2018-09-10*