A Joint Model Approach for Evaluating the Efficacy of HAART Treatment for AIDS patients in Taiwan

Yi-Kuan Tseng[1], Ying-Hen Hsieh[2], and Shiing-Jer Twu[3]

[1]Graduate Institute of Statistics, National Central University, Chungli, Taiwan

[2]Department of Applied Mathematics, National Chung Hsing University, Taichung, Taiwan

[3]National Health Research Institute, Taiwan

Running title: Joint Model for Evaluating HAART

Word counts: 2430

Correspondence to: Dr. Ying-Hen Hsieh, Department of Applied Mathematics, National Chung Hsing University, Taichung, Taiwan 401; Tel: 886-4-22853949, Fax: 886-4-22853949, Email: hsieh@amath.nchu.edu.tw

Abstract

In recent years, benefits of Highly Active Antiretroviral Therapy (HAART) conferred on the morbidity and mortality of HIV/AIDS patients have been observed in numerous studies, mostly in industrialized countries where both AIDS incidence and the mortality resulting from AIDS have been declining. In this study, we make use of a cohort data of 138 AIDS patients in Taiwan, where the annual number of AIDS incidence is still increasing, to investigate the impact of HAART on survival time from onset of AIDS to death, by attempting to identify the association between patterns of CD4 counts and survival times for the HAART and non-HAART groups. Complications in data used, such as CD4 counts measured intermittently at different time points for different patients, measurement errors, or measurements unavailable after the survival time, could result in biased inferences for this study. Therefore we make use of a joint modeling approach including graphic techniques to model survival time and the time-dependent CD4 count simultaneously, by which we are able to derive unbiased and efficient estimators. The statistical inferences and graphic techniques reveal that both HAART and CD4 count are positively associated with survival time. The p-values for the impact of HAART and CD4 count are less than 0.00001 and 0.013, respectively. The mean survival time for non-HAART and HAART group are 2.31 years [95%CI: (1.72, 3.83)] and 5.82 years [95%CI: (4.34, 8.96)], respectively; and the respective median survival time are 1.88 years [95%CI: (1.60, 2.65)] and 6.51 years [95%CI: (5.54, 9.31)]. Moreover, increasing log CD4 count by one unit will reduce hazard rate by 38.9% and the relative risk of non-HAART group to HAART group is 7.76. Our findings are consistent with previous statistical studies, indicating that HAART is helpful in improving the survival time of AIDS patients and CD4 count is positively related to lifetime. The results derived through joint modeling approach are more accurate and reliable due to the advantages offered by our procedure.

Introduction

Since the advent of Highly Active Antiretroviral Therapy (HAART) in 1996, substantial decrease in mortality and slower progression to AIDS had been observed in many studies (e.g., [1-4]); while other studies have shown HAART to be unable to completely and durably suppress HIV replication and sometimes even with adverse drug reaction (e.g., [5]). Currently, HAART regimens with 3 or more antiretroviral agents, usually from at least 2 different classes, are the standard treatment for HIV-infected persons receiving antiretroviral therapy [6]. A systematic review of 54 randomized, controlled trials consisting of 16,684 HIV-infected patients with limited or no antiretroviral experience concludes that 3-drug therapy was more effective than 2-drug therapy [7]. Moreover, observational studies have also indicated that HAART can result in sustained (up to 4 to 5 years) improvements in CD4 cell counts and viral loads [8-10]. However, all these studies have been done in industrialized countries where both AIDS incidence and the mortality resulting from AIDS have been declining in recent years.

In Taiwan, the annual number of AIDS incidence is still increasing, from 16 in 1991, to 136 in 1997, and to 259 in 2004 [11]. Beginning in April 1997, the Department of Health of Taiwan adopted a policy to provide all HIV-infected citizens with free access to HAART through the National Health Insurance program, with the timing of the initiation of HAART and the regimens based on guidelines recommended in the United States [12]. The combination of an effective surveillance system and a policy to provide free HAART to all HIV-infected citizens made it possible to study the effect of widespread use of HAART in Taiwan under the condition that the yearly incidence of HIV infection is still increasing. Several studies have been carried out which focused on the effects of HAART in HIV incidence [13] and the quality of life [14-15] in Taiwan. More recently, Li et al. [16] constructed the trajectory of the changing viral load with treatment time. They used longitudinal viral load data of

HIV-infected patients in Taiwan with complete viral load observations, of which 208 are in the HAART group and 164 in the non-HAART group, to apply a nonparametric mixed-effects model. This study concluded that, at the beginning of treatment, non-HAART group has strong antiviral activity, which is lacking with HAART group. However, during the course of the treatment, the superiority of non-HAART treatment lessens, and this therapy ultimately fails, whereas HAART maintains a constant effect throughout the treatment.

However, the impact of HAART on the level of CD4 counts in AIDS patients and the mortality of AIDS patients in Taiwan has never before been reported. The primary aim of this study focuses on the evaluation of the impact of HAART on the survival time from onset of AIDS to death of AIDS patients by attempting to identify the association between patterns of CD4 counts and survival times for the HAART and non-HAART groups using a joint model approach [17], which derives estimates through joint likelihood function of the Cox model and linear mixed effects models. The joint model approach, which efficiently utilizes information of survival time and time-dependent covariates at same time, yields consistent and (semiparametrically) efficient estimates (e.g., [18]). For lucid and insightful reviews on joint modeling, the readers are referred to [19, 20]. The joint modeling approach has also been used for a clinical trial comparing risperidone with a placebo for the treatment of schizophrenia in [21, 22], however their estimating procedures differ from ours in the way one maximizes the joint likelihood function.

The rest of this article is organized in the following manner. We describe the data used for our study in Section 2. Section 3 gives the proposed statistical procedure, namely, the joint model. In Section 4 we give the results obtained by applying our procedure to the data described in Section 2. Finally, in Section 5, we discuss the advantages of our method as well as limitations in applications.

2. Taiwan AIDS cohort data

The HIV/AIDS cohort data of 1054 HIV-infected patients in Taiwan, collected from 1990 to January of 2003, was used in this study. All patients were advised to return every three or four months for medical treatment and measurement (especially CD4 count). Although this original data set contains more than 200 AIDS patients, only patients developed AIDS were included in our study. Moreover, since the original data set was not from a randomized clinical trial (almost all patients in the data set who developed AIDS after 1996 were treated with HAART), we may encounter various biases when analyzing the original data. As an alternative, a historical comparison may be applicable here. Consequently, after deleting those AIDS patients who switched between the two treatment groups and discarding a few AIDS patients who had onset after 1996 but only received non-HAART treatment, 138 AIDS patients were then selected. The fact that patients who had onset of AIDS after 1996 are more likely to receive HAART could be attributable to its being freely available and to treatment efficacy, the use of historical control group enables us to handle such dilemmas.

Of the selected 138 patients, there are 101 patients in non-HAART group who had onset of AIDS before 1996, when HAART treatment was first introduced in Taiwan, and received non-HAART (single or double drugs) at all the time. The other 37 patients developed AIDS after 1996 and received HAART after the onset time. Therefore, this group of 101 patients is the historical control group when deriving the effect of HAART treatment for AIDS patients.

3. Statistical Method

3.1. Graphic method

As a complement to formal statistical methodology, graphic technique is a helpful tool to capture the features of longitudinal CD4 count and survival time simultaneously.

Three graphic methods are employed in this study to gain insights into the data. The event history graph was developed by Dubin et al. [23]. The graph contains the Kaplan-Meier estimator for the right-censored data and a simultaneous display of the patterns of time dependent covariates and survival time for each subject in the sample. The second method is a 3-D display of spline smoothing surface for survival time, measurement time, and log CD4 count (see perspective plot in [24]). With the 3-D graphs, the relationship between survival time and log CD4 count can be simply grasped. Another method uses LOWESS smoothing technique on log CD4 count to specify the behavior of CD4 count in different treatment groups (see Lowess plot in [24]).

3.2. Joint method

The recent developed joint model approach [17] is used to model survival time and the time-dependent CD4 count simultaneously. Let $X(t)$ denote the time dependent CD4 count and $e(t)$ denote the measurement errors. As $X(t)$ may be disturbed by $e(t)$, the CD4 count actually observed is denoted by $W(t)= X(t) + e(t)$, where $X(t)$ cannot be directly observed. The relationship between survival time and time dependent CD4 count is then through the Cox proportional hazards model

$$\lambda\{t \mid \overline{X}(t), Z\} = \lambda_0(t)\exp\{\beta X(t) + \gamma Z\},$$

where Z is the treatment group with the non-HAART group coded 0 and HAART group coded 1. Moreover, $\overline{X}(t) = \{X(s) : 0 \le s < t\}$ denotes the covariate history up to time t, (β, γ) is the regression coefficients, and $\lambda_0(t)$ is the unspecified baseline hazard rate function. The patterns of the time dependent CD4 count $X(t)$ can be well described by linear mixed-effects models. As we can see from the results in the next section, a cubic random coefficient model is suitable to describe $X(t)$, which is defined as

$$X(t) = b_0 + b_1 t + b_2 t^2 + b_3 t^3,$$

where (b_0, b_1, b_2, b_3) follow a 4-dimesional multivariate normal distribution. Note that the regression coefficient β reflects the relationship between survival time and CD4 count, and γ reveals the effect of HAART on survival time. Regression coefficients being significantly less than 0 indicate that the corresponding covariates tend to be associated with increased survival time, while regression coefficients significantly greater than 0 indicate that the covariates are associated with shortened survival time and increased mortality. The estimates of coefficients (β, γ) are derived by maximizing joint likelihood function [17] through Monte Carlo EM algorithm and bootstrap procedures [20, 25]. The CD4 count was transformed logarithmically to yield a better fit for $X(t)$. Furthermore, the proportionality assumption was assessed by scaled Martingale residuals to ensure the validity of Cox model.

4. Results

The event history graph is presented in Fig 1(a)-1(c) for all patients, the non-HAART group, and the HAART group, respectively. Each colored bar denotes the CD4 history of each patient. Four different colors denote different levels of log CD4 counts. The red color indicates a high log CD4 count of greater 2.7, the orange color indicates a medium-high log CD4 count of between 2.0 and 2.7; the yellow color indicates a medium-low log CD4 count of between 1.2 and 2.0; and the green color indicates a low log CD4 count of less than 1.2. The Kaplan-Meier curve is the upper boundary of the union of all bars. Fig 1(a), with the CD4 counts of all patients, may not indicate any special pattern. However, if we plot the event history graph separately for the two treatment group, then certain interesting trends do appear. In Fig 1(b), the CD4 counts of non-HAART group eventually decreased to a medium-low level or low level, while in Fig 1(c) most of patients in HAART group kept a medium-high or high level CD4

counts all during the time period with some patients even having increased CD4 counts for some time interval. In addition, Fig 1(c) has a much higher survival probability than 1(b) when comparing with the Kaplan-Meier estimators. The event history graphs may reveal the possibly positive association among CD4 count, HAART, and survival time of the patient.

Fig 2(a)-2(c) show the 3-D smoothing surface for all patients, non-HAART group and HAART group, respectively. Fig 2(a) indicates the trends that CD4 counts eventually decrease, and the patients with longer lifetime maintain higher level CD4 count for a longer period. To obtain more information, we need to examine the 3-D smoothing surface for each group. In Fig 2(b), patients in non-HAART group with shorter lifetimes tend to have rapid CD4 decreasing rates, while patients in the same group but with longer lifetime have slow CD4 decreasing rate. It is intuitively reasonable that the longer lifetimes the patients have, the slower CD4 decreasing rate would be expected. In Fig 2(c), patients in HAART group tend to have increasing CD4 counts at early measurement time. However, in the long run the CD4 counts are still decreasing, albeit with decreasing rate which is slower than that of non-HAART group. Furthermore, the survival time in HAART group is longer than that in non-HAART group. Again, these figures may suggest possibly positive associations among CD4 count, HAART, and survival time.

To observe the trends of CD4 counts in both groups clearly, we plot the profile of CD4 for each group and use LOWESS smoothing technique to obtain the trends. Fig 3(a) and Fig 3(b) show the trends of CD4 count in both treatment groups. The overall trend of CD4 counts in non-HAART group decreases quickly at the first two years and stays stable thereafter. However, the overall trend of CD4 counts in HAART group gradually increases to attain a peak but then decreases slowly to a stable level. Hence our findings show that, using graph methods, HAART seem to confer a benefit in both survival and CD4 counts of patients when comparing with patients in

non-HAART group.

For statistical analysis, the scaled Martingale residuals do not show violation to the proportionality assumption, and hence Cox model is valid to be used for this data. The statistical analyses via joint model reveal two facts. First, the pattern of CD4 counts of individual can be well described by a cubic random coefficient model. The evidence can be seen in Fig 4 which shows that cubic functions fit very well to the four randomly selected patients' CD4 counts. Consequently, the profile of CD4 counts of each individual can be approximated by a cubic function with coefficients different. If we assume the four coefficients of each patient are random and come from the same population then this is cubic random coefficient model. Second, the regression coefficients $(\beta, \gamma) = (-0.389, -2.049)$ are both significantly less than 0 with p-value (0.013, <0.00001). This indicate that a larger CD4 count or delivery of HAART predict a longer survival time. With adjusted by CD4 counts, the survival functions for the two treatment groups are presented in Fig 5 which shows patients in HAART group (in red curve) have higher survival probability. Moreover the mean survival time for non-HAART and HAART group are 2.31 years [95%CI: (1.72, 3.83)] and 5.82 years [95%CI: (4.34, 8.96)] , respectively, and the respective median survival time are 1.88 years [95%CI: (1.60, 2.65)] and 6.51 years [95%CI: (5.54, 9.31)]. To be more specific, increasing log CD4 count by one unit will reduce hazard rate by 38.9% and the relative risk of non-HAART group to HAART group is 7.76.

5. Conclusions and Discussions.

The results obtained in this work by employing graphic methods are consistent with those found by statistical analyses. That is, both agreeing that HAART is helpful to lengthening lifetime for AIDS patients, and that CD4 count is positively related to lifetime.

Of the HIV/AIDS cohort data in Taiwan, we only used the AIDS patient data. In

theory, we could also explore the improvement in morbidity, i.e., the delay in onset of AIDS in HIV-infected patients, due to the fact that the Taiwan cohort data includes many patients who have received HAART for several years before onset of AIDS. However, this direction of study is complicated by the fact that the infection time of the each patient is unknown, as opposed to our present study of AIDS mortality where the onset time of every AIDS patient in the data is known.

When performing partial likelihood for Cox proportional hazards model, it is necessary to have complete knowledge of the covariate history for all individuals. Therefore, biases occur when time-dependent covariates (i.e. CD4 counts) are measured intermittently at different time points for different patients (possibly with measurement errors) or when measurements are not available after the survival time. Joint modeling approach offers a solution to these predicaments when Cox proportional hazards model is used. Moreover, the estimators derived by joint modeling approach have some nice properties (consistency, asymptotic normality and efficiency) under a large number of repeated measures and small measurement errors (see [18, 20]). Consequently, one needs to be cautious when using joint modeling approach if there are very few repeated CD4 measurements obtained for each patient or if the measurement errors of CD4 counts are large. In these cases, a much more complicated functional analysis [26] could be used instead.

Acknowledgement. This research is supported by National Science Council (NSC) of Taiwan under grant 95-125-M005-003 (YHH) and 95-2118-M008-003(YKT). This article was written partially while YHH visited the School of Mathematics at University of New South Wales, Sydney, Australia, funded by Taiwan CDC grant (DOH95-DC-1407)

References

1. Egger M, Hirschel B, Francioli P, et al. Impact of new antiretroviral combination therapies in HIV infected patients in Switzerland: prospective multicentre study. Swiss HIV Cohort Study. *BMJ* 1997; **315**(7117):1194-9.
2. Brodt HR, Kamps BS, Gute P, Knupp B, Staszewski S, Helm EB. Changing incidence of AIDS-defining illnesses in the era of antiretroviral combination therapy. *AIDS* 1997; **11**:1731-8.
3. Palella FJ Jr, Delaney KM, Moorman AC, 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.
4. Forrest DM, Seminari E, Hogg RS, et al. The incidence and spectrum of AIDS defining illnesses in persons treated with antiretroviral drugs. *Clin Infect Dis* 1998; **27**:1379-85.
5. Lucas CM, Chaisson RE, Moore RD. Highly active antiretroviral therapy in a large urban clinic: risk factors for virologic failure and adverse drug reactions. *Ann Intern Med* 1999; **131**:81-87.
6. *Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents.* Bethesda, MD: U.S. Department of Health and Human Services. 7 April 2005. Accessed at http://www.aidsinfo.nih.gov/guidelines/adult/aa_040705.pdf on 16 May 2005.
7. Jordan R, Gold L, Cummins C, Hyde C. Systematic review and meta-analysis of evidence for increasing numbers of drugs in antiretroviral combination therapy. *BMJ* 2002; **324**:757.
8. Gulick RM, Meibohm A, Havlir D, et al. Six-year follow-up of HIV-1-infected adults in a clinical trial of antiretroviral therapy with indinavir, zidovudine, and lamivudine. *AIDS* 2003; **17**:2345-9.
9. Kaufmann GR, Perrin L, Pantaleo G, et al. CD4 T-lymphocyte recovery in

individuals with advanced HIV-1 infection receiving potent antiretroviral therapy for 4 years: the Swiss HIV Cohort Study. *Arch Intern Med.* 2003; **163**:2187-95.

10. Garcia F, De Lazzari E, Plana M, et al. Long-term CD4+ T-cell response to highly active antiretroviral therapy according to baseline CD4+ T-cell count. *J Acquir Immune Defic Syndr.* 2004; **36**:702-713.
11. Department of Health, Executive Yuan, Taiwan. *Health and vital statistics: General health statistics.* Taipei, Taiwan: Department of Health, Executive Yuan, 2004.
12. Centers for Disease Control and Prevention. Report of the NIH panel to define principles of therapy of HIV infection and guidelines for the use of antiretroviral agents in HIV-infected adults and adolescents. *MMWR* 1998; **47**(RR-5):1–41.
13. Fang CT, Hsu HM, Twu SJ, et al. Decreased HIV transmission after a policy of providing free access to highly active antiretroviral therapy in Taiwan. *J Infect Dis.* 2004; **190**(5):879-85.
14. Fang CT, Hsiung PC, Yu CF, Chen MY, Wang JD. Validation of the World Health Organization quality of life instrument in patients with HIV infection. *Qual. Life Res*. 2002; **11**:753-762.
15. Yen CF, Tsai JJ, Lu PL, et al. Quality of life and its correlates in HIV/AIDS male outpatients receiving highly active antiretroviral therapy in Taiwan. Psychiatry *Clin. Neurosci*. 2004; **58**(5): 501-6.
16. Li C S, Liang H, Hsieh YH, Twu SJ. Comparison Of Viral Trajectories In Aids Studies By Using Nonparametric Mixed-Effects Models. *J. Modern Applied Stat. Methods* 2003; **2**(2):443-450.
17. Wulfsohn MS, Tsiatis AA. A joint model for survival and longitudinal data measured with error. *Biometrics* 1997; **53**:330-9.
18. Zeng D, Cai J. Asymptotic results for maximum likelihood estimators in joint analysis of repeated measurements and survival time. *The Annals of Statistics* 2005; **33**(5):2132-63.

19. Tsiatis AA, Davidian M. Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica* 2004; **14**:809-34.
20. Hsieh F, Tseng YK, Wang JL. Joint Modeling of Survival Time and Longitudinal Data: Likelihood Approach Revisit. *Biometrics* 2006; accepted. DOI:10.1111/j.1541-0420.2006.00570.x.
21. Henderson R, Diggle P, Dobson P. Joint modelling of longitudinal measurements and event time data. *Biostatistics* 2000; **4**:465-80.
22. Xu J, Zeger SL. Joint analysis of longitudinal data comprising repeated measures and times to events. *Appl. Statist.* 2001; **50**(3):375-87.
23. Dubin J, Mueller HG, Wang JL. Event history graphs for censored survival data. *Statistics in Medicine* 2001; **20**:2951-64.
24. Everitt B, Rabe-Hesketh S. *Analyzing Medical Data Using S-PLUS* 2001; Springer-Verlag: New York.
25. Tseng YK, Hsieh F, Wang JL. Joint modeling of accelerated failure time and longitudinal data. *Biometrika* 2005; **92**:587-603.
26. Yao F, Müller HG, Wang JL. Functional data analysis for sparse longitudinal data. *Journal of American Statistical Association* 2005;**100**:577-90.

Figure Legends

Fig 1(a). Event history graph for all AIDS patients. The red color indicates a high log CD4 count of greater than 2.7, the orange color indicates a medium-high log CD4 count of between 2.0 and 2.7, the yellow color indicates a medium-low log CD4 count of between 1.2 and 2.0, and the green color indicates a low log CD4 count of less than 1.2.

Fig 1(b). Event history graph for AIDS patients in non-HAART group.

Fig 1(c). Event history graph for AIDS patients in HAART group.

Fig 2(a). 3-D surface using smoothing spline for all patients.

Fig 2(b). 3-D surface using smoothing spline for non-HAART group.

Fig 2(c). 3-D surface using smoothing spline for HAART group.

Fig 3 (a). The profile of log CD4 counts of the patients in non-HAART group. Each blue curve denotes CD4 history of a patient. The red curve is the LOWESS smoothing trend.

Fig 3 (b). The profile of log CD4 counts of the patients in HAART group. Each blue curve denotes CD4 history of a patient. The red curve is the LOWESS smoothing trend.

Fig 4. Individual profile of log CD4 counts fitted by cubic function. (a) patient ID 267 fitted by $2.66 + 0.95\,x - 0.73\,x^2 + 0.12\,x^3$; (b) patient ID 450 fitted by $3.26 - 0.46\,x - 0.11\,x^2 + 0.03\,x^3$; (c) patient ID 519 fitted by $0.62 + 1.56\,x - 0.43\,x^2 + 0.043\,x^3$; (d) patient ID 600 fitted by $2.12 - 0.20\,x + 0.08\,x^2 - 0.12\,x^3$.

Fig 5. Survival function for non-HAART group (in blue) and HAART group (in red).

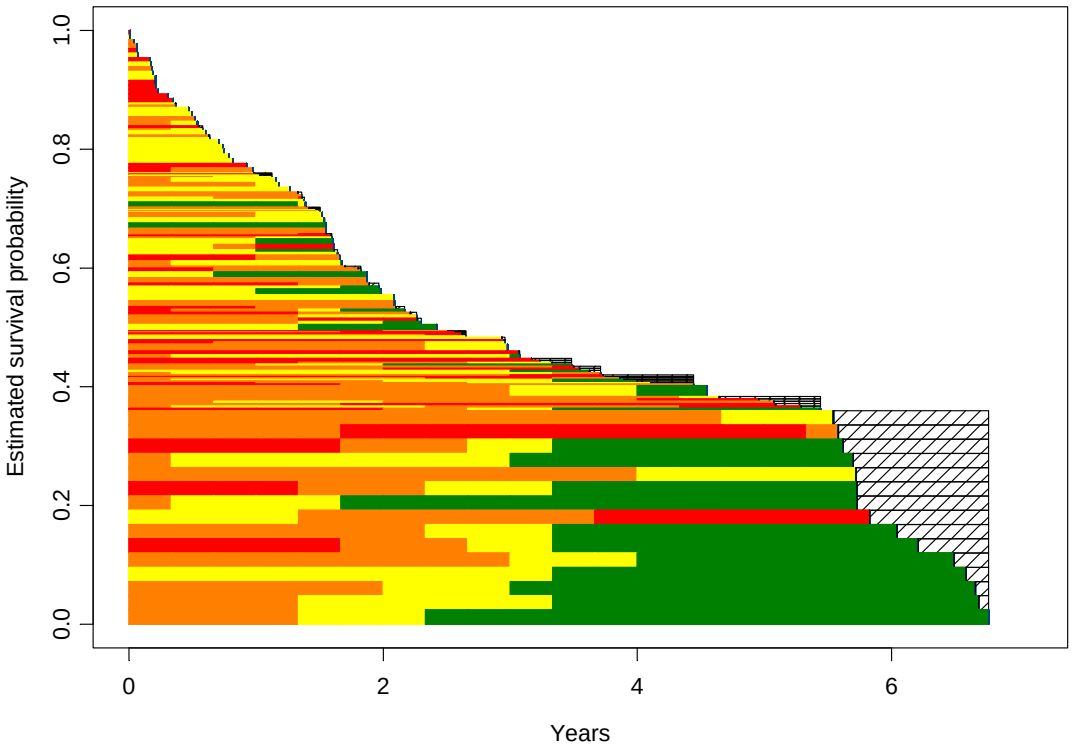

Fig 1(a). Event history graph for all AIDS patients. The red color indicates a high log CD4 count of greater than 2.7, the orange color indicates a medium-high log CD4 count of between 2.0 and 2.7, the yellow color indicates a medium-low log CD4 count of between 1.2 and 2.0, and the green color indicates a low log CD4 count of less than 1.2.

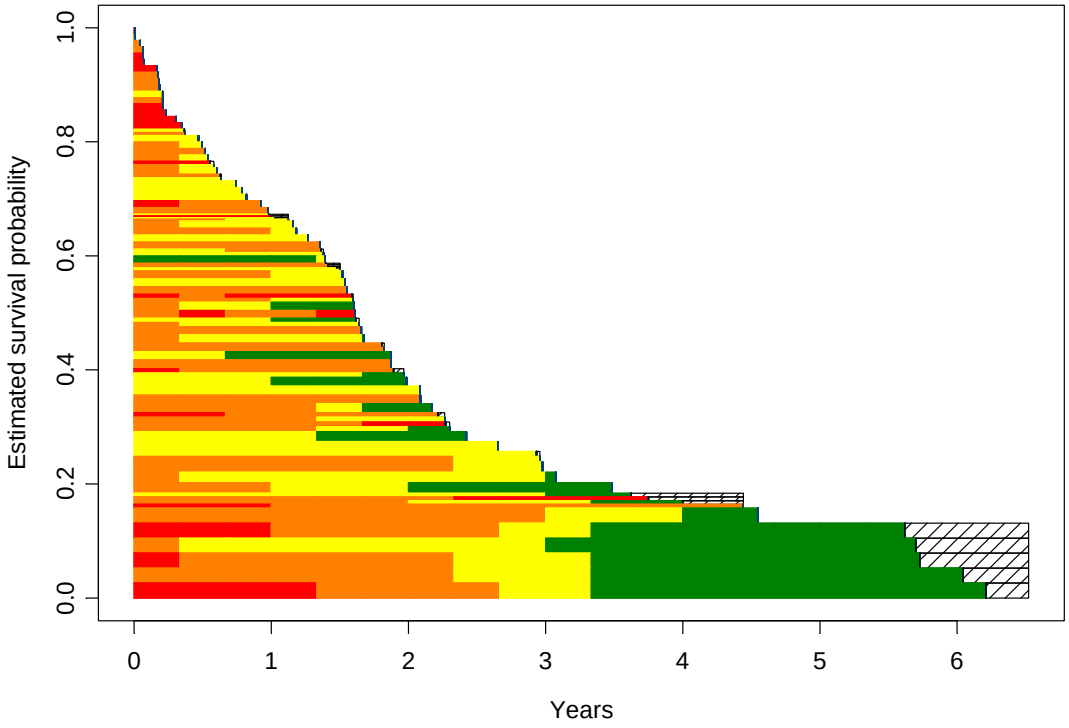

Fig 1(b). Event history graph for AIDS patients in non-HAART group.

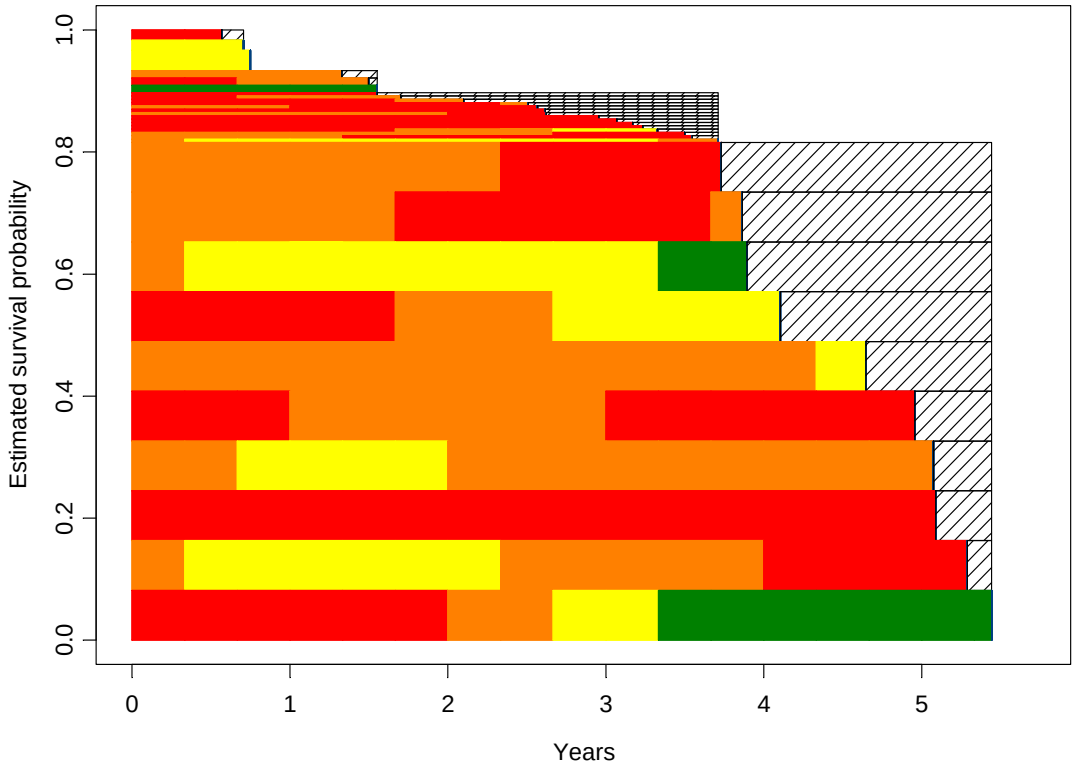

Fig 1(c). Event history graph for AIDS patients in HAART group.

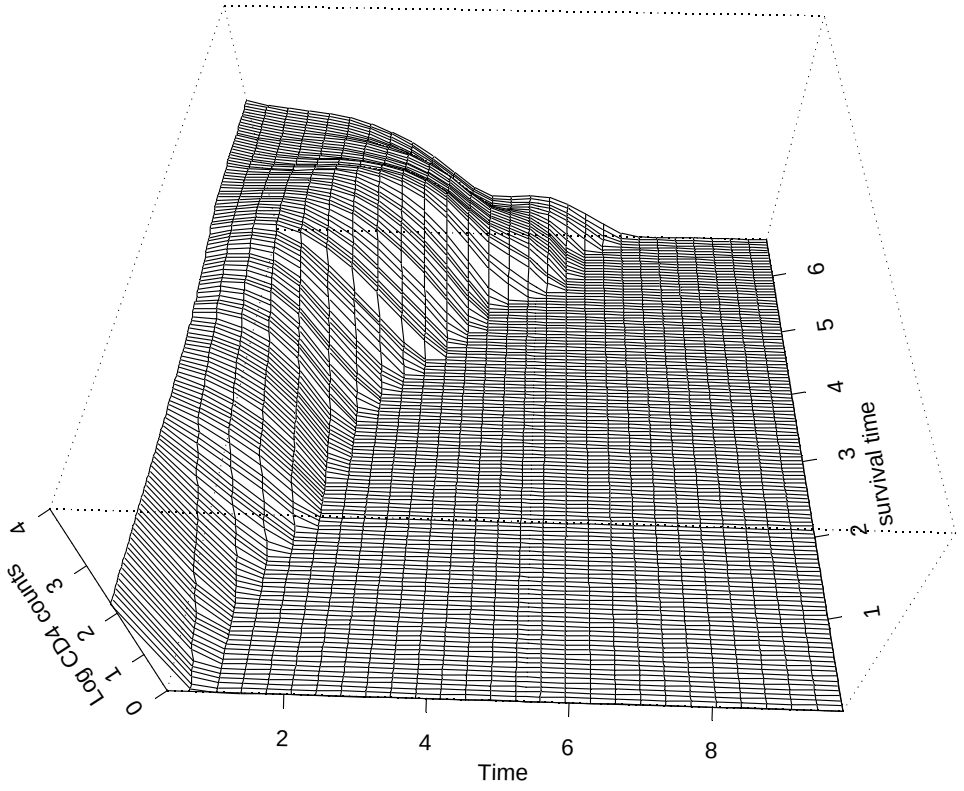

Fig 2(a). 3-D surface using smoothing spline for all patients.

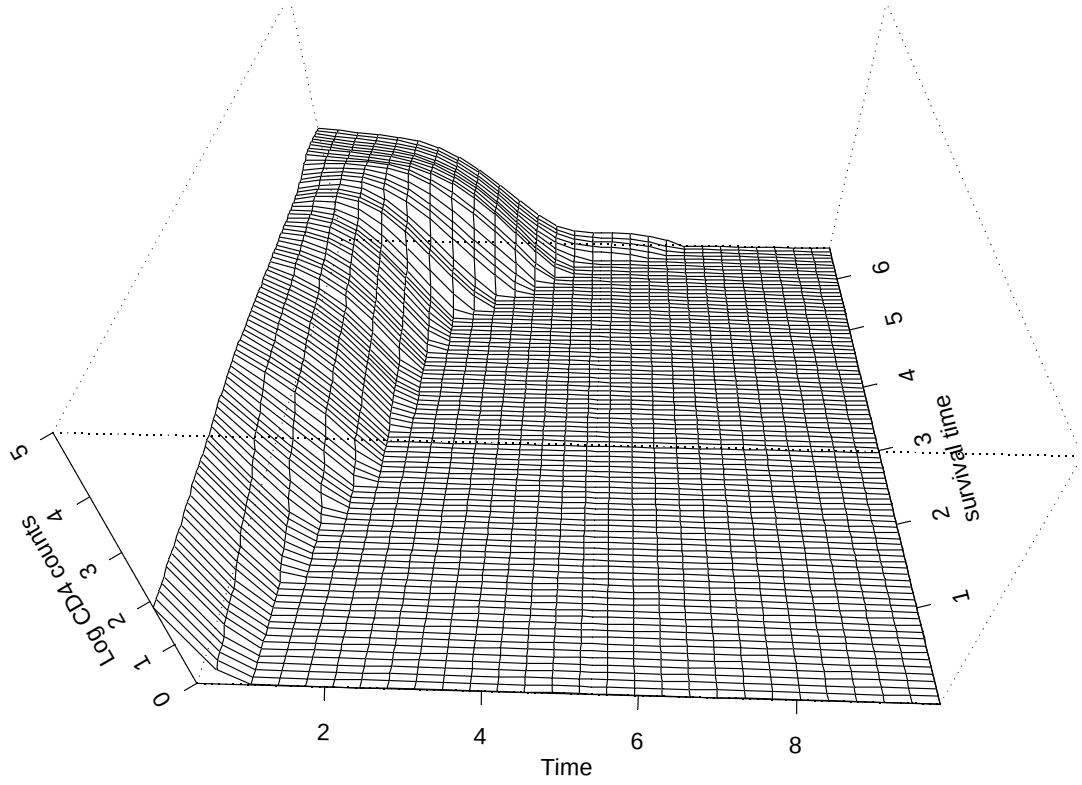

Fig 2(b). 3-D surface using smoothing spline for non-HAART group.

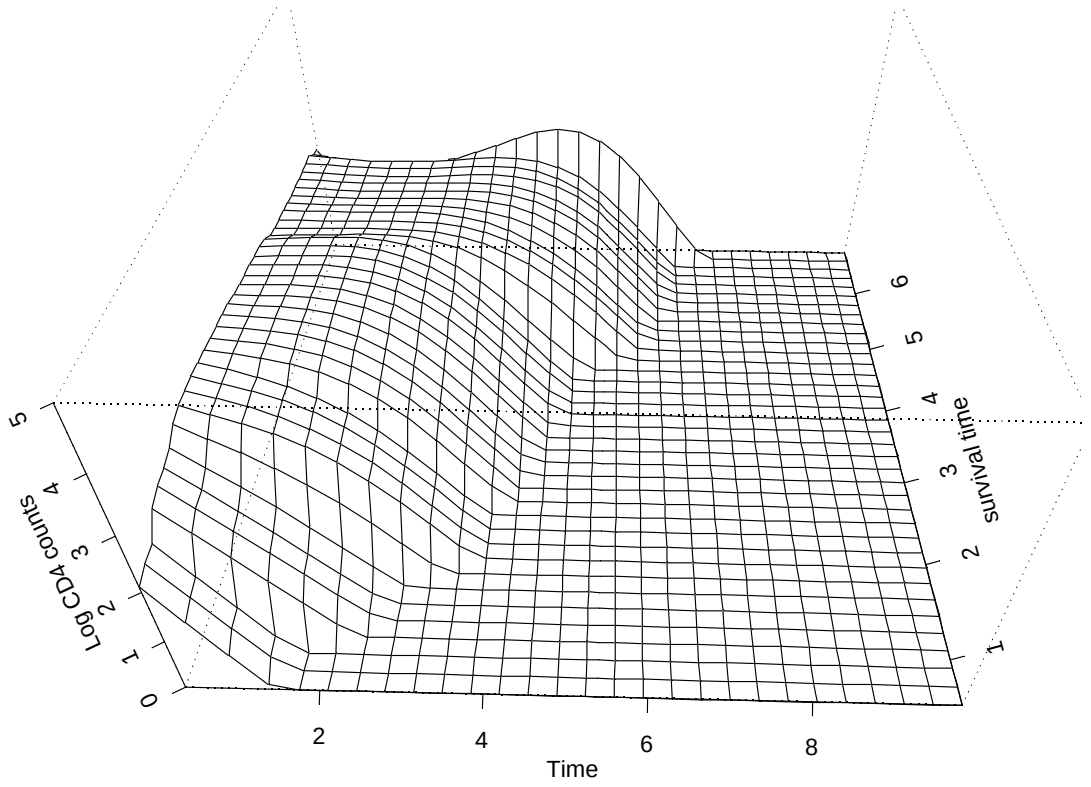

Fig 2(c). 3-D surface using smoothing spline for HAART group.

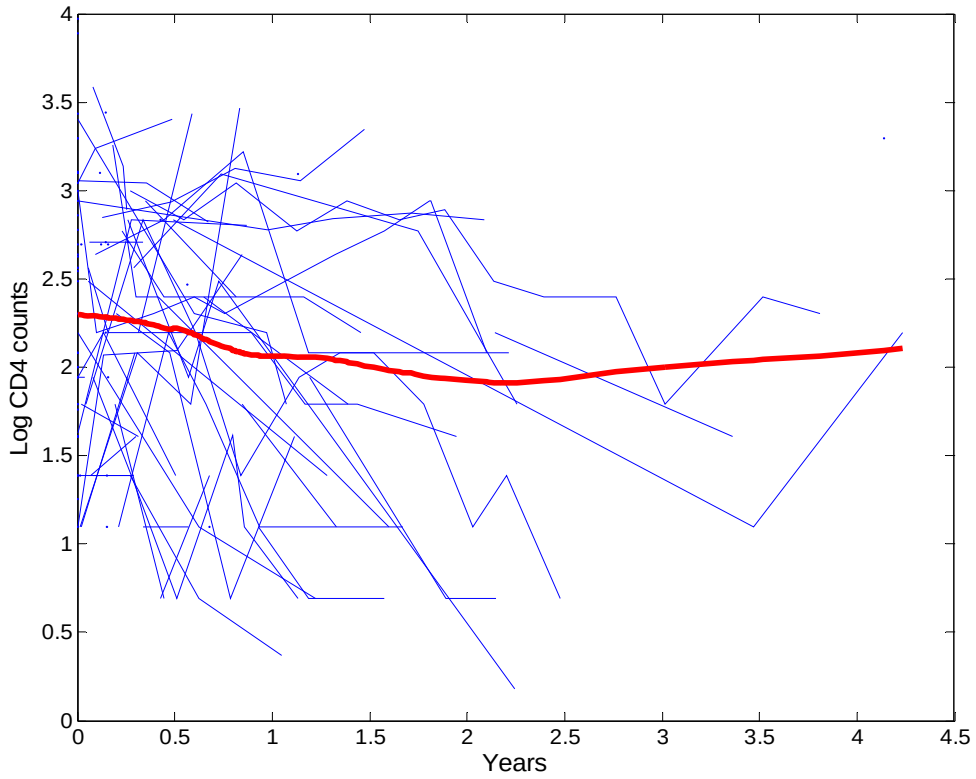

Fig 3 (a). The profile of log CD4 counts of the patients in non-HAART group. Each blue curve denotes CD4 history of a patient. The red curve is the LOWESS smoothing trend.

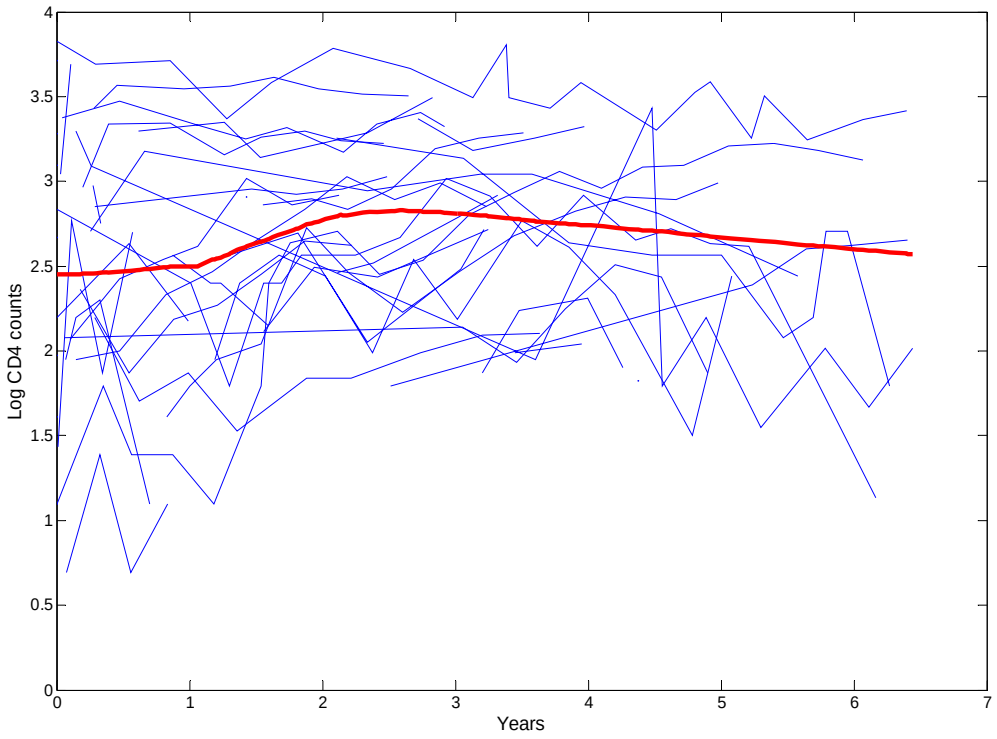

Fig 3 (b). The profile of log CD4 counts of the patients in HAART group. Each blue curve denotes CD4 history of a patient. The red curve is the LOWESS smoothing trend.

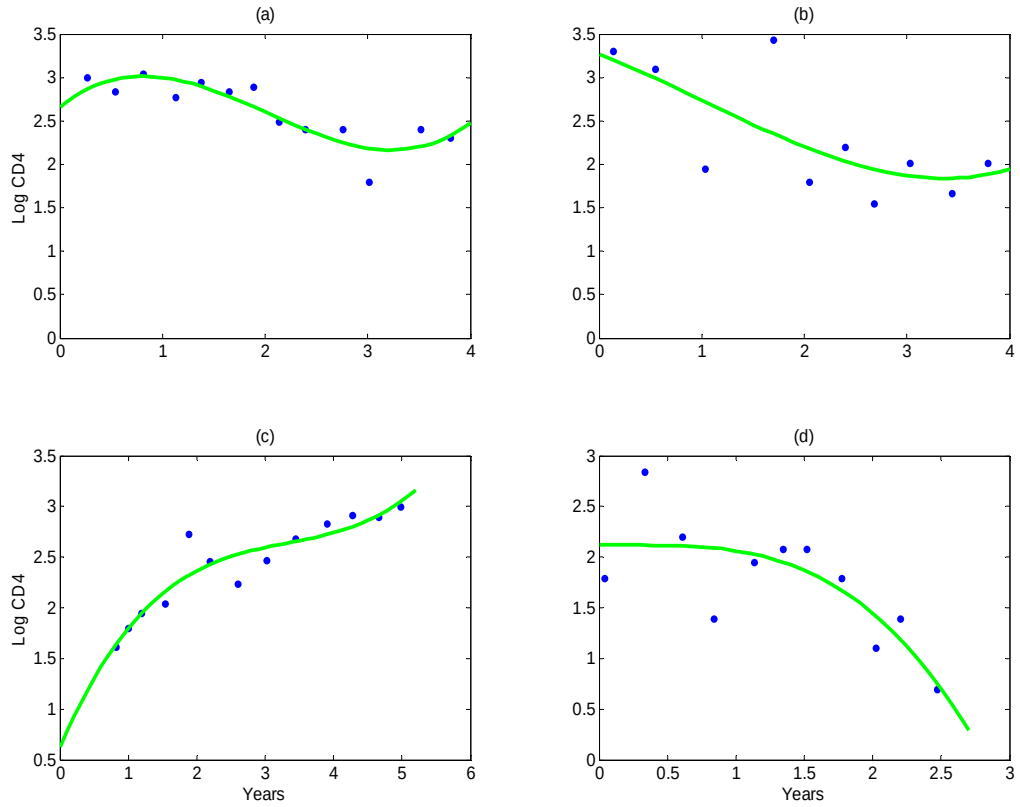

Fig 4. Individual profile of log CD4 counts fitted by cubic function. (a) patient ID 267 fitted by $2.66 + 0.95\,x - 0.73\,x^2 + 0.12\,x^3$ (b) patient ID 450 fitted by $3.26 - 0.46\,x - 0.11\,x^2 + 0.03\,x^3$ (c) patient ID 519 fitted by $0.62 + 1.56\,x - 0.43\,x^2 + 0.043\,x^3$ (d) patient ID 600 fitted by $2.12 - 0.20\,x + 0.08\,x^2 - 0.12\,x^3$

.

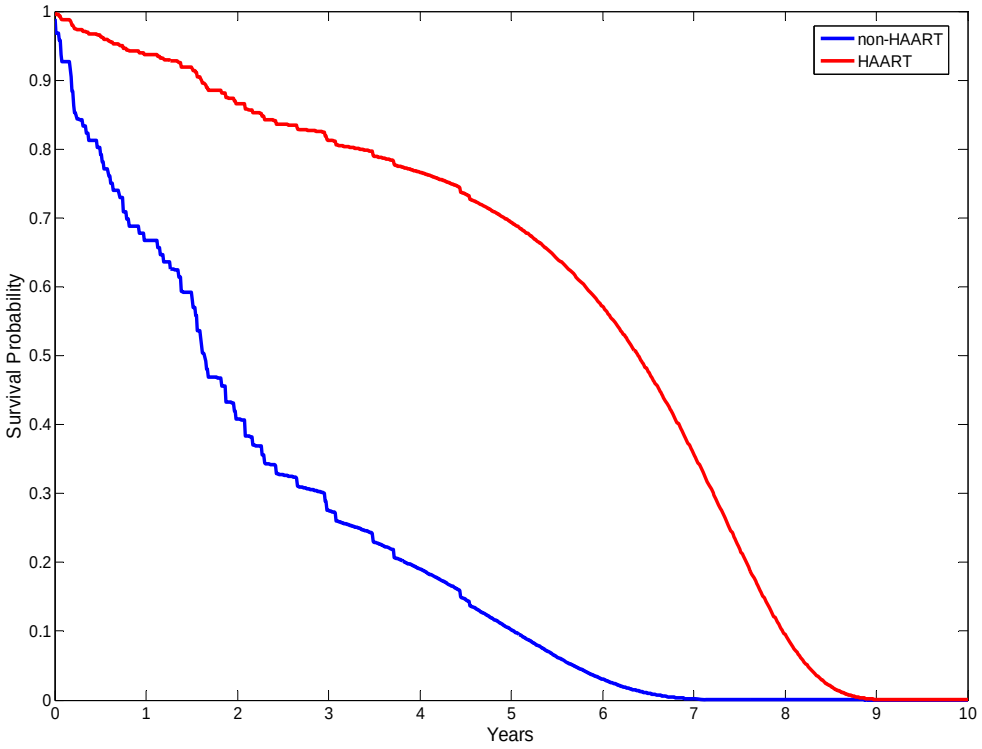

Fig 5. Survival function for non-HAART group (in blue) and HAART group (in red).