STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–17.
Published online in International Academic Press (www.IAPress.org)

A combined Cox-Aalen model for HIV/AIDS progression using data from a clinical trial in South Africa

Claris Shoko [1,*], Delson Chikobvu [2], Pascal Obong Bessong [3], Gorata Duduzile Manyeagae [1]

[1] *Department of Statistics, University of Botswana, Botswana*
[2] *Department of Mathematical Statistics and Actuarial Sciences, University of the Free State, South Africa*
[3] *HIV/AIDS and Global Research Programme, School of Mathematics and Natural Sciences, University of Venda, South Africa*

Abstract **Background and Objective:** Over thirty-nine million individuals globally are living with HIV, and during the past three decades, HIV/AIDS has claimed millions of lives globally. The study's primary goal was to determine the variables that affect the course of HIV/AIDS in individuals receiving antiretroviral therapy, including body mass index, age, gender, TB before treatment, response to treatment, WHO stage baseline, and CD4 baseline. **Materials and Methods:** The research employed a cohort analysis of surveillance data from the HIV Wellness Clinic in the province of Limpopo, South Africa, for 318 HIV-positive patients on antiretroviral therapy (ART). The effects of covariates were evaluated in this study using Aalen's additive regression model, the Cox regression model and the Cox-Aalen ensemble model, in comparison. **Results:** Results from the Cox-Aalen model, complemented by findings from Aalen's additive and Cox proportional hazards, show a substantial association between HIV/AIDS progression for patients undergoing treatment therapy and CD4 baseline, gender, age, and incidence of TB during treatment. All models show that the CD4 baseline is highly significant in comparison to the other covariates in the models, with a very small $p - value < 0.0001$. Male patients had greater rates of immune system deterioration than female patients. Patients with baseline CD4+ cell counts below 350 cells/mm^3 have greater rates of immunological recovery than those with baseline counts over 350 cells/mm^3. **Conclusion:** This study has reinforced the need for an integrated HIV/TB care so that these two diseases are concurrently managed for the benefit of the infected individual.

Keywords Aalen's additive regression model, Cox's hazard regression model, Cox-Aalen model, antiretroviral therapy, treatment adherence, immune deterioration, cohort analysis.

AMS 2010 subject classifications 62N01, 62P10

DOI: 10.19139/soic-2310-5070-2628

1. Introduction

The HIV/AIDS disease progression is characterised by continuous destruction of the CD4+ cells leading to immunosuppression, neoplasm, wasting, or low CD4+ T-cell count that defines AIDS [1, 2]. Some clinical markers, such as the CD4+ cell count and the RNA viral load, help in providing information on HIV/AIDS disease progression. A normal CD4+ cell count varies from individual to individual, and it is usually between 500 and 1,400 cells per mm^3. CD4+ cell count values below 500 are usually an indication of immune suppression and vulnerability to opportunistic infections [3].

The Joint United Nations Programme on HIV/AIDS (UNAIDS) estimates show that approximately 39 million people globally are living with HIV, and 40.4 million have died of AIDS-related illnesses since the beginning of the epidemic by the end of 2022. Although sub-Saharan Africa constitutes a small fraction of the world population,

*Correspondence to: Claris Shoko (shokoc@ub.ac.bw). Department of Statistics, University of Botswana, 4775 Old Notwaane Rd, Gaborone, Botswana.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

approximately 67% of the HIV/AIDS cases have been reported in sub-Saharan Africa, with approximately 23.4 million HIV positive cases by 2022 [4]. South Africa was leading with 7.8 million, followed by Nigeria with 3 million HIV-positive individuals. Between 2001 and 2011, South Africa, the country with the largest number of HIV infections, reduced new HIV infections by 41%. South Africa has one of the largest HIV treatment programmes in the world [5]. Therefore, it is important to identify factors associated with survival from HIV/AIDS for patients who are on antiretroviral therapy.

South Africa has one of the highest dual burdens of HIV and tuberculosis (TB) in the world, with HIV as a critical driver of the TB epidemic and increased mortality. While the country has made significant efforts and progress in reducing the burden of TB, the ongoing challenge remains the management of TB/HIV co-infection. South Africa is one of 30 nations that account for 87% of global TB cases and is on the World Health Organisation's (WHO) list of 14 countries with the highest burden of TB, HIV-associated TB, and multi-drug resistant TB (MDR-TB).

Survival analysis is particularly important for accurately identifying risk factors affecting survival rates [6]. Researchers commonly use survival analysis techniques, such as Cox proportional hazards regression models, to identify factors that predict mortality. These models help in understanding the hazard ratios (HR) associated with different variables, providing estimates of the increased risk of death for specific patient groups. In 2016, [7] analysed the survival of patients receiving antiretroviral therapy. In their analysis, they concluded that prolonged survival times for patients with HIV/AIDS under treatment, and hence improved survival probability. They emphasised the need for constant follow-up, regular monitoring of CD4+ cell count, as well as timely and early antiretroviral intake to achieve better survival chances. In 2019, Mangal et al. [8] carried out a study on people living with HIV/AIDS using piecewise constant exponential models. The study shows that early treatment contributes to improved survival probabilities, with the greatest benefit obtained in women and younger age groups. In 2016, Kowaliska et al. [9] used the Cox proportional hazard models to identify factors related to the first combination antiretroviral therapy (cART) modification in three age groups. Their results reveal significantly longer time on cART for the older age groups than their younger counterparts.

However, relatively few studies on HIV/AIDS progression have been carried out using both the Cox proportional hazards model and the Aalen additive hazards model. While the Cox proportional hazards model is frequently used to analyse survival in censored data, [17] contended that when the proportional assumption is not met, bias may result. Then, the additive model turns into an alternative option. The Aalen additive model has proved to be advantageous over the Cox proportional hazards model for HIV modelling primarily because it allows covariate effects to vary over time and directly models absolute risk differences, providing different and potentially more informative insights than the Cox model's assumption of constant multiplicative hazard ratios. This flexibility is crucial in HIV modelling, where the impact of factors like CD4 count or treatment adherence can change significantly as a patient's disease progresses. The additive model also offers simpler, closed-form solutions for parameter estimation and can provide a clearer understanding of "excess risk".

Using both Aalen's additive model with a Cox proportional hazards (PH) model for HIV modelling offers advantages by providing a comprehensive view of covariate effects over time, with the additive model highlighting changes in absolute risk and the Cox PH model explaining relative risk. The additive model can reveal time-varying covariate effects that the Cox PH model might obscure or fail to detect due to its assumption of constant hazard ratios [11], thus offering a richer, more complete understanding of HIV progression. Thus, in this study, both the Cox proportional hazards model and Aalen's additive model are used. In addition, the Cox-Aalen model is presented, a hybrid model that combines both the multiplicative and the additive covariate effect. For this cohort, an investigation of whether CD4 baseline (CD4BL), age, gender, TB before enrollment (TBB4), developed TB during treatment (DTB), adverse reaction to treatment (Reaction), WHO stage baseline (WSBL), body mass index progression (BMI), and viral load baseline (VLBL) are associated with HIV/AIDS progression for patients under treatment therapy is carried out.

1.1. Contributions

The study employs the Cox-Aalen model, a model that combines both Aalen's additive regression model and the Cox proportional hazards model, providing a robust comparative analysis of their applicability to HIV/AIDS progression. The research addresses critical factors influencing HIV/AIDS progression, such as CD4 baseline,

gender, and TB co-infection, which are pivotal for clinical decision-making. The use of a well-documented cohort (n=318) from a real-world setting enhances the study's external validity.

1.1.1. Key findings

- CD4 Baseline Dominance: Both the Cox proportional hazards model and Aalen's Additive model consistently highlight low CD4 baseline (< 350 cells/mm^3) as a strong predictor of immune deterioration, aligning with global ART initiation guidelines. The Cox-Aalen model further identifies the additive effect of the continuous CD4 cell baseline on HIV/AIDS progression.
- Gender Disparity: Male gender is associated with higher progression risk, potentially reflecting adherence or biological factors.
- TB Coinfection: Incident TB during treatment is significantly linked to progression, underscoring the need for integrated HIV/TB care–a critical public health message
- Findings underscore the importance of early TB detection and gender-specific treatment strategies, aligning with global HIV management guidelines.

1.2. Organization

The rest of the paper is combined as follows: Section 2 presents a detailed methodology for the Cox, Aalen and the Cox-Aalen models. This is followed by Section 3 where the key findings from the analysis are presented. Section 4 discusses the key findings and compares them with the literature. Finally, Section 5 concludes the findings.

2. Materials and Methods

Ethical considerations: The data collection procedures used in this study were approved by the Research Ethics Committee of the University of Venda, SA (ref. no. SMNS/13/ MBY/01/0625), following the 1964 Declaration of Helsinki and its subsequent amendments. Additionally, permission to access health facilities was obtained from the Limpopo Provincial Department of Health, SA, and collaborating health facilities. Informed consent was obtained from study participants before their involvement, and data obtained were stripped of personal identifiers to ensure the anonymity and confidentiality of the participants.

2.1. Study population

The study used a cohort analysis of the surveillance data for 318 HIV-infected patients under anti-retroviral therapy (ART) from the HIV Wellness Clinic in the Limpopo Province of South Africa. From these individuals, there were 227 females and 91 males. The survey was conducted between 2005 and 2009, and follow-up was done every 6 months. The age range of the patients at enrolment was from 2 to 77 years, with a mean age of 36.47 years. The age group 31-40 years had the highest percentage of individuals, with approximately 31.4%. Before the commencement of treatment, the body mass index (BMI) was calculated, and the mean BMI was 19.17 at enrolment. World Health Organisation (WHO) stage baseline and CD4 baseline were measured. The CD4 cell count at enrolment ranged from 1 cell/mm3 to 518 cells/mm3. 70% of the individuals had a CD4 cell count below 200, and 30% had a CD4 count between 200 and 750. There were no individuals with a CD4 count above 750. Some of the patients were enrolled at the clinic with TB as the initial marker of HIV. During treatment, CD4 count was noted, and any form of adverse reaction to treatment was noted, as well as any development of a TB co-infection.

2.2. Variable coding

For this analysis, the variables were coded as follows: Gender: 1 - Male, 0 - Female; prior ART TB (PATB): 1 - yes, 0 - no; incidence TB on ART (ITB): 1 - yes, 0 - no; Virologic failure (VF): 1 - yes, 0 - no. The CD4BL are as follows: 0: CD4 ≥ 350; 1: 500 CD4 < 350. The continuous baseline CD4BL is also considered in the analysis to complement the categorical CD4BL. Analysis was done using the R statistical software since it can handle Aalen's additive regression analysis.

2.3. Cox proportional hazards model versus Aalen's regression model in modelling HIV progression

The proportional hazards regression model, known as the Cox regression model, has been the popular choice for right-censored survival data [13, 14]. Its popularity is based on the fact that it is simple to fit and that the results are easy to explain. However, Cox's regression model has its limitations. Firstly, the validity of the analysis relies heavily on the proportional hazards assumptions. Secondly, it cannot include time-varying covariate effects since the regression coefficients are assumed to be constants [15]. The alternative is Aalen's regression model. Aalen's model postulates a different relationship for the hazard and covariates than does the Cox model [13, 16]. The coefficients in Aalen's model are functions that vary over time without any specific shape or reliance on the parameter functions [17]. Events that happen to one person are therefore presumed to be independent of those of other people. Because of this, it is non-parametric [18] as opposed to Cox's proportional hazards model, which is semi-parametric. Thus, Aalen's model is ideally suited to investigate potential temporal variations in the variables' impacts.

It is argued [19] that it is not appropriate to use Aalen's additive hazards model for all data sets because the calculation of cumulative regression functions $\beta(t)$ is restricted to a time interval, and Z, the matrix of covariates, is of full rank. This means that $Z'Z$ should be invertible, which is not always the case. However, with the Cox regression model, this is not a problem. Table 1 gives a summary of the differences between the two models.

Table 1. A summary of the differences between the Cox proportional hazards model and Aalen's additive model

	Cox proportional Hazards model	**Aalen's Additive model**
Assumption:	Assumes covariate effects are multiplicative on the hazard rate and that these effects are constant over time (the proportional hazards assumption).	Assumes covariates affect the hazard rate in an additive way, on an absolute scale.
Effect of Covariates:	Estimates time-invariant regression coefficients.	Estimates time-varying coefficients, allowing the effect of covariates to change over time.
Interpretation:	Focuses on hazard ratios (relative risks), indicating how the hazard changes multiplicatively for an individual with a given covariate value compared to a baseline.	Focuses on excess risks or the absolute difference in the hazard rate per unit of covariate change.
Model Type:	Semi-parametric.	Non-parametric, making it more flexible for exploring how effects change over time.
Application:	A widely used standard for survival analysis, but its validity hinges on the often-restrictive proportional hazards assumption.	Useful for understanding when and how covariate effects change, offering new insights into prognostic factors over time.

2.3.1. Cox proportional hazards regression model Cox proportional hazards model is specified as $h(t|\bar{Z}) = h_0(t)exp\{\bar{\beta}'_r\bar{Z}\}$ where $h_0(t)$ is a baseline hazard, which may vary arbitrarily over time, and $\bar{Z}' = (z_1, z_2, \ldots, z_r)$ is the covariate vector, $\bar{\beta} = (\beta_1, \beta_2, \ldots, \beta_r)$ is a vector of covariate coefficients and is assumed constant. The baseline hazard is treated to be non-parametric. The null hypothesis to test for the significance of a given covariate z_j in the Cox model is given by Equation 1:

$$H_j : \beta_j(t) = 0.. \tag{1}$$

Where $\beta_j(t)$ is the regression coefficient corresponding to the j^{th} covariate [20].

If two individuals with covariate values, Z_a and Z_b being compared, the ratio of their hazard rates at any time point simplifies to Equation 4 below:

$$\frac{h(t|Z)}{h(t|Z*)} = \frac{h_0(t)exp\{\sum_{k=1}^{r}\beta_k Z_{ak}}{h_0(t)exp\{\sum_{k=1}^{r}\beta_k Z_{bk}\}} \quad (2)$$

$$= \frac{exp\{\sum_{k=1}^{r}\beta_k Z_{ak}}{exp\{\sum_{k=1}^{r}\beta_k Z_{bk}\}} \quad (3)$$

$$= exp\{\sum_{k=1}^{r}\beta_k(Z_{ak} - Z_{bk})\}. \quad (4)$$

Where $h_0(t)$ is the baseline hazard rate, β_k is the regression coefficient for the k^{th} covariate. This ratio is considered constant throughout the study period.

A sample of size n yields data with F distinct failure times denoted by $t_1 < t_2 < \cdots < t_F$ and $n - F$ censored times. $t_{(i)}$ is an ordered event while T_j is the follow-up time for the j^{th} individual. The set of indices of subjects at risk at time t_i is denoted by $R_i = R(t_i)$. The covariate for the individual who had an event at time t_i is given by $Z_{(i)}$ to be distinguished from covariate Z_j for individual j according to the previous study [15]. This allows the proportion for all the individuals at risk at time t_i such that $\{j \in R_i\} = \{j|T_j \geq t_i\}$. The partial likelihood for estimating the parameter β is given by Equation 5:

$$L(\beta) = \Pi_{i=1}^{F}\frac{exp(\beta' Z_{(i)}}{\sum_{j\in R_i} exp(\beta' Z_j)} \quad (5)$$

Where i denotes the subscript of the subject who dies at any time $t_{(i)}$ [21].

2.3.2. Aalen's additive hazard model Aalen's additive model is a non-parametric survival model that specifies the hazard rate as a sum of a baseline hazard and a linear function of covariates. Aalen's additive regression model is constructed as follows: Suppose a possibility of a censored lifetime of a number of HIV patients is observed. Let λ_i denote the hazard rate of patient i, n be the number of HIV-infected patients, and r be the number of predictors in the analysis. The additive model, is constructed by taking into account the column vector $\lambda(t)$ of hazard rates $\lambda_i(t), i = 1, 2, \ldots, n$, is given by $\lambda(t) = Y(t)\alpha(t)$, where the $n \times (r + 1)$ matrix $Y(t)$ is built as follows: If the patient is a member of the risk set at time t, them the i^{th} row of the $Y(t)$ is the vector $\bar{z}^i(t) = (1, z_1^i(t), z_2^i(t), \ldots, z_r^i(t))'$ where $z_j^i(t), j = 1, 2, \ldots, r$, are time-dependent covariates values. The corresponding row of $Y(t)$ contains only zeros if at time t the individual i is at risk. The vector $\alpha(t) = (\alpha_0(t), \alpha_1(t), \ldots, \alpha_r(t))'$ contains the regression information where α_0 is a baseline hazard rate, while the remaining components are regression functions which quantify the impact of the corresponding covariates. Over time, their functions are free to change.

One primary question is "Do specific covariates have any influence on the distribution of lifetimes?". This corresponds with the hypothesis; $H_j : \alpha_j(t) = 0$, where j corresponds to the j^{th} covariate in the analysis. This corresponds to the null hypothesis in the Cox model. However, the alternative hypothesis for the Cox model $H_j : \alpha_j(t) \neq 0$ is valid for all t, whereas the alternative hypothesis in Aalen's model $H_j : \alpha_j(t) \neq 0$ is valid for some t. For Aalen's additive model, cumulative regression functions or coefficients are estimated as:

$$A_j(t) = \int_0^t \alpha_j(u)du. \quad (6)$$

$\alpha_j(t)$ is estimated from the slope of $A_j(t)$, which is often done using kernel smoothing techniques

2.4. Model formulation

The HIV data used in this study have $n = 318$ patients and $j = 9$ variables. The covariates are: Baseline CD4 (CD4BL), age, sex, pre-ART TB (PATB), incident TB (ITB), adverse reaction to treatment defined as virologic failure (VF) or lactic acidosis (LA) or peripheral neuropathy (PN), WHO stage baseline (WSBL) and viral load baseline (VLBL). The effects of these covariates on the progression of HIV were analysed using the "timereg" package in R, and the results are discussed in the following paragraphs.

The process of identifying the model that best fits the data is performed using backward elimination. This involved starting by considering a model that included the full set of covariates, and then removing non-significant covariates step by step. The effects of the removal of these variables are noted by checking the change in p-values of the global test statistics, and also checking the changes in R^2 for the fitted models. The analysis is done for both Aalen's and Cox's regression, and the comparison of these two models is based on the p-values of each of the covariates as well as the p-values for the model and other statistics. The identification of the better model is done in four steps.

The first step presents both Aalen's and Cox's models with all the variables included. The Cox regression model fitted is:

$$\begin{aligned} h(t|\bar{Z} = h_0(t)exp\{&\beta_1 \times Gender + \beta_2 \times WHOSBL + \beta_3 \times Age \\ &+ \beta_4 \times CD4BL + \beta_5 \times LA + \beta_6 \times PN + \beta_7 \times ITB + \beta_8 \times PATB\} \end{aligned} \tag{7}$$

and the Aalen's additive model is given by:

$$\begin{aligned} \lambda_i(t) = &\alpha_0(t) + \alpha_1(t) \times Gender_i + \alpha_2(t) \times WHOSBL_i + \alpha_3(t) \times Age_i \\ &+ \alpha_4(t) \times CD4BL_i + \alpha_5(t) \times VF_i + \alpha_6(t) \times PN_i + \alpha_7(t) \times ITB_i + \alpha_8(t) \times PATB_i) \end{aligned} \tag{8}$$

where $\alpha_0(t)$ is the baseline hazard rate and $\alpha_i(t)$ is the time-dependent regression function that represents the additive effect (risk) of the ith covariate on the hazard.

However, continuous elimination of non-significant variables is done to ensure that the model remains with only those variables that explain the progression of HIV/AIDS better. The final theoretical models being estimated in the final step are:

$$h(t|\bar{Z} = h_0(t)exp\{\beta_1 \times Gender + \beta_2 \times Age + \beta_3 \times CD4BL + \beta_4 \times ITB\} \tag{9}$$

and;

$$\lambda_i(t)|z_i = \alpha_0(t) + \alpha_1(t) \times Gender_i + \alpha_2(t) \times Age_i + \alpha_3(t) \times CD4BL_i + \alpha_4(t) \times ITB_i \tag{10}$$

where $\alpha_r(t)$ for $r = 0, 1, 2, \ldots, 4$ is the regression information, and $\beta_r(t)$ for $r = 1, 2, \ldots, 4$ vector of covariate coefficients for Aalen's additive model and Cox regression model, respectively.

2.4.1. Confidence Interval for Aalen's Additive model The confidence interval for the cumulative function $\alpha_k(t)$ is used to test the hypotheses: $H_{01} : \beta_k(t) = 0$ and $H_{02} : \beta_k(t) = bt$ for all $t \in [0, \tau]$ and τ being the largest observed time. For the k^{th} cumulative function, H_{01} is based on the supremum test statistic. The Kolmogorov-Smirnov test statistic is used to test the hypothesis of a time-dependent effect of the k^{th} covariate.

2.5. Cox-Aaalen Hazards Regression model

The Cox-Aalen regression model was proposed by Scheike and Zhang [22], is a more flexible model than the other additive-multiplicative models, which combines Aalen's additive regression model and the Cox proportional hazards model discussed earlier. This model is based on an additive structure on the basis of the multiplication model. In the Cox-Aalen model, the covariates are partitioned into two parts; some covariate effects work additively on the intensity, and other covariate effects act multiplicatively. Combining these two models helps to overcome the limitations of a single model approach that struggles with mixed effects timelines. Thus, the Cox-Aalen model is more comprehensive when dealing with covariates that behave differently over time, capturing both relative risk (from the Cox component) and absolute effects over time (from the Aalen component). This allows for a richer and more nuanced understanding of how different covariates influence survival over time. It is a more flexible and potentially useful model which is defined by, Equation 11 below.

$$\lambda(t)|x_i = Y(t) \left[X(t)^T \alpha(t)\right] exp\left(Z(t)^T \beta\right) \tag{11}$$

where $Y(t)$ is the risk indicator; $(X(t), Z(t))$ is a $(p+q\times 1)$ vector of covariates, $\alpha(t)$ is a $(p\times 1)$ vector of time varying regression coefficients and β is a $(q\times 1)$ vector of relative risk regression coefficients. This model allows some covariate effects to be additive, nonparametric and time varying $(X(t))$ and other covariates $(Z(t))$ to have constant multiplicative effects

2.6. Goodness of Fit

In order to avoid misrepresentation of the effects of covariates, it is important to check how well the fitted models fit the data. For the Aalen additive regression part of the model, we used the martingale residuals. The martingale residual gives a precise indication of the point of the problem in the fitted model as a function of time [23, 24]. At a given time t, the martingale residual for individual i is given by Equation 12:

$$\hat{M}_i(t) = N_i(t) - \hat{H}\left[t|\mathbf{Z}_i(t)\right], i = 1, 2, 3, \ldots, n, \tag{12}$$

where $N_i(t)$ represent the actual number of events and $\hat{H}\left[t|\mathbf{Z}_i(t)\right]$, is the predicted number of events. The martingale residuals are the sum based on the levels of the covariate, and a sum near zero when plotted against time represents model fitness. P-values can also be used to quantify the departure of the residuals from zero.

3. Results

In this section, the results on the fitted Aalen's additive model and the Cox proportional hazards model are fitted on the data.. Estimates of the coefficients $\alpha_r(t)$ for $r = 0, 1, 2, \ldots, 8$, and β_r for $r = 1, 2, \ldots, 8$ for Aalen's additive model and Cox regression model, respectively, are presented. For each model, a step-wise approach to the selection of significant covariates is used. The approach starts with a full model containing all covariates, and at each stage, some covariates are removed based on the p-values at 5% level until the best fitting model is obtained. The Aalen's additive model and the Cox proportional hazards model are presented separately, and at the end, a comparison of the findings from each model is discussed.

3.1. Cox Proportional Hazards model

The Cox proportional hazards model is applied to the data. The first step starts with a full model that contains all covariates (the full model). A step-wise approach is used to eliminate the variables that do not satisfy the proportional hazards assumption until the best model is obtained (the reduced model). This is done by examining the Schoenfeld residuals for the variables in a fitted model. Covariates with p-values below 0.05 are removed from the model because they do not satisfy the proportional hazards assumption. The results are presented in Table 2. Proportionality of covariates is satisfied by five covariates: the HIV patient's age, having TB prior to ART initiation (PATB), virologic failure (VF), TB incident (ITB), and CD4 baseline (CD4BL). The global test shows no evidence of departure from the standard Cox model (p-value=0.270).

The plot of the scaled Schoenfeld residuals is a useful diagnostic tool. A non-zero slope is an indication of a violation of the proportional hazard assumption. The plots of scaled Schoenfeld residuals are presented to test if the proportional hazards assumption is satisfied for all the covariates in the final/reduced Cox proportional hazards model. The proportional hazards assumption is satisfied for all the variables: TB prior to ART, peripheral neuropathy, virologic failure, incident TB, and CD4 baseline. The graphs of the scaled Schoenfeld residuals are presented in Figure 1

In Table 3 below presents the estimations of regression coefficients from the reduced Cox model together with the corresponding p-values for each covariate. The relative risk, represented by the exp(β_r) is interpreted as a multiplicative effect on the hazard function. The risk of having a CD4 cell count above 350 among patients who had TB before ART is 0.62146 (p-value=0.0242) times lower than that of patients who did not have TB prior to ART. This means a reduction by 38% in the risk of progressing to a better CD4 cell count level as a result of having TB prior to ART. The risk among patients with TB incidence is increased by approximately 7% (p-value=0.04857). This could be because these patients are on follow-up, and as a result, any TB incident was quickly diagnosed and

Table 2. Test for proportional hazards assumption for the Cox Proportional Hazards model

	Outcome from the full model:			**Outcome from the reduced model:**		
Covariates	chisq	df	p-value	chisq	df	p-value
Age	3.48e+00	1	0.06197	2.92e+00	1	0.088
CD4B	1.84e+01	1	1.8e-05			
factor(Gender)	3.92e+00	1	0.04768			
factor(PATB)	1.60e+00	1	0.20525	1.38e+00	1	0.241
factor(PN)	2.40e+00	1	0.12120			
factor(VF)	2.27e-08	1	0.99988	4.51e-08	1	1.000
factor(LA)	4.71e+00	1	0.02999			
factor(VLBL)	4.70e+00	1	0.03013			
factor(ITB)	7.33e-06	1	0.99784	1.72e-01	1	0.678
factor(CD4BL)	7.77e-01	1	0.37808	6.57e-01	1	0.418
GLOBAL	3.07e+01	10	0.00065	7.59e+00	6	0.270

treated. For patients who started treatment with a CD4BL below 350, the risk of progressing to a CD4BL above 350 is 0.105 times lower than that of patients who started treatment with a CD4BL above 350. This calls for the need for early diagnosis of HIV and treatment, that is, the treat and test model. For every unit increase in age, the risk of progressing to a better CD4 cell count level is reduced by about 2% (p-value=0.0316). Based on the likelihood ratio, Wald, and the score tests, the overall test of the fitted model is statistically very highly significant (p-value< 0.0001).

Table 3. Estimation of regression coefficient, relative risk, standard error, z and p-value from the reduced Cox model

	β_r	exp(β_r)	se(coef)	z-value	p-value
Age	-0.02216	0.97809	0.01230	-1.802	0.0316*
factor(PATB)1	-0.47568	0.62146	0.21106	-2.254	0.0242 *
factor(ITB)1	0.07072	1.07328	0.12974	0.545	0.04857*
factor(CD4BL)1	-2.25497	0.10488	0.24059	-9.373	$< 2e-16$ ***
Concordance= 0.62 (se = 0.021)					
Likelihood ratio test= 71.27 on 6 df, p=2e-13***					
Wald test = 87.56 on 6 df, $p =<$ 2e-16***					
Score (logrank) test = 132.4 on 6 df, $p =<$ 2e-16***					

3.2. Cox-Aalen's Additive model

In this subsection, we fit the Cox-Aalen model to assess the additive effects of all the covariates. The results are presented in two steps: **Step 1:** The first output presented in Table 4 below tests the additive effect of all the variables. This step is analogous to an Aalen's additive model. From this test, covariates: gender (1-Male, 0-Female), virologic failure (VF: 1-Yes, 0-No), CD4B (continuous CD4 baseline variable), and CD4BL (0-CD4BL≥ 350, 1-CD4BL< 350) turned out to be statistically significant at the 5% level of significance, thus satisfying the additive (time-varying assumption). The other variables: VLBL, PATB, LA, incident TB (ITB) and PN do not satisfy the additive assumption.

Step 2: To fit the Cox-Aalen model, the covariates are partitioned into two parts to accommodate all covariate effects, those that work additively (Aalen's model) on the intensity and those that work multiplicatively (Cox

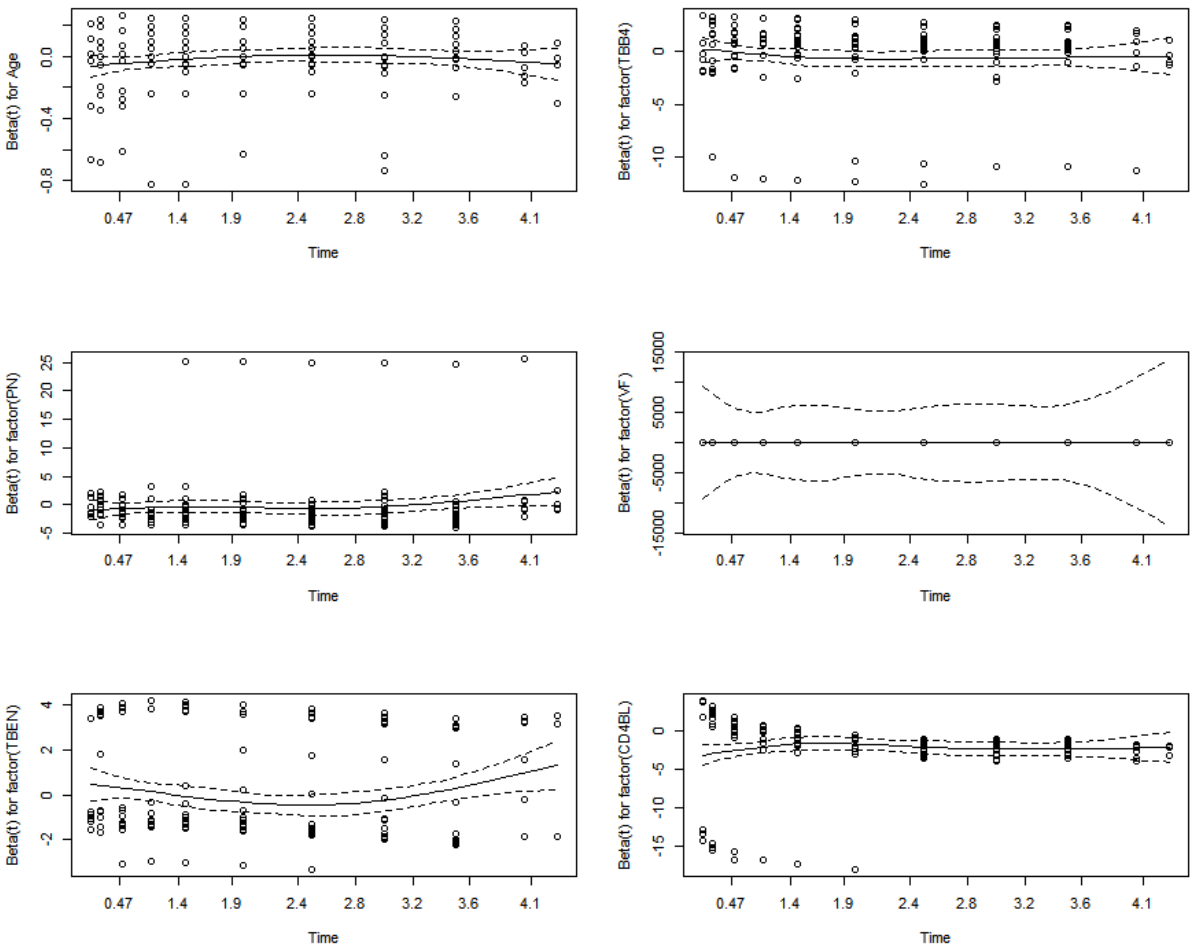

Figure 1. Smoothed scaled Schoenfeld residual plots with 95% pointwise confidence intervals for dfor the reduced Cox proportional Hazards model

model). This results in a combined Cox-Aalen model. In Table 5 and Table 6, we present results from the Cox-Aaalen model for the additive part and the multiplicative part of the model, respectively. For the additive part of the model, a non-parametric significance test is performed using the supremum test and the test for time invariance is performed using two tests: the Kolmogorov-Smirnov test and the Cramer von Mises test. The Cox-Aalen model clearly separates the multiplicative part from the additive part as presented in the Cox and Aalen models, respectively.

Results from the multiplicative part of the Cox-Aalen model are given in Table 6. The multiplicative part of the model suggested that the relative risk of patient age is 1.0179, that of incidence TB is 1.1286, and the relative risk of pre-ART TB is 1.1526. These results are similar to the Cox model. However, peripheral neuropathy has neither an additive effect nor a multiplicative effect. These results now show that, for every unit increase in age, the risk of progressing to a higher CD4 cell count level increases by 2%. The risk increases by 12.8% for patients with TB incidence. However, the risk for patients with TB prior to ART is now increased compared to the estimated risk when the Cox proportional model is used.

Table 4. Test for nonparametric terms for Aalen's additive model: Part of First output

Test for non-significant effects		
	Supremum-test of significance	p-value $H_0 : \alpha(t) = 0$
(Intercept)	6.50	0.000***
Age	3.08	0.047 *
factor(VLBL)1	1.63	0.649
CD4B	7.08	0.000***
factor(Gender)1	5.10	0.000 ***
factor(PATB)1	1.32	0.910
factor(VF)1	5.87	0.000 ***
factor(LA)1	2.72	0.109
factor(ITB)1	2.68	0.109
factor(CD4BL)1	5.71	0.000 ***
factor(PN)1	2.24	0.291

Table 5. Test for nonparametric terms for the Cox-Aalen's additive model: Final output –Part A

Test for non-significant effects		
	Supremum-test of significance	p-value $H_0 : \alpha(t) = 0$
(Intercept)	13.00	0.000 ***
factor(VLBL)1	2.99	0.050*
CD4B	7.00	0.000***
factor(Gender)1	5.04	0.000***
factor(VF)1	5.87	0.000***
factor(LA)1	3.58	0.008**
factor(CD4BL)1	5.77	0.000***
Test for time invariant effects		
	Kolmogorov-Smirnov test	p-value H_0:constant effect
(Intercept)	1.360	0.000***
factor(VLBL)1	0.241	0.033 *
CD4B	0.00223	0.000***
factor(Gender)1	0.232	0.000***
factor(VF)1	0.254	0.000***
factor(LA)1	0.217	0.006**
factor(CD4BL)1	2.020	0.000***
	Cramer von Mises test	p-value H_0:constant effect
(Intercept)	1.3100	0.004**
factor(VLBL)1	0.0493	0.025*
CD4B	6.05e-06	0.000***
factor(Gender)1	0.0579	0.000***
factor(VF)1	0.0519	0.000***
factor(LA)1	0.0361	0.002**
factor(CD4BL)1	3.2500	0.000***

Table 6. Estimation of regression coefficient, relative risk, standard error, z and p-value for multiplicative part of Cox-Aalen model: Final output-Part B

	β_r	exp(β_r)	SE	Robust SE	z	P-val	lower2.5%	upper97.5%
const(Age)	0.0147	1.0179	0.0026	0.00194	7.57	0.000***	0.0096	0.0198
const(factor(PATB))1	0.1420	1.1526	0.0662	0.04950	2.88	0.00399**	0.0123	0.2720
const(factor(PN))1	-0.0336	0.9670	0.0831	0.07450	-0.45	0.652	-0.1960	0.1290
const(factor(ITB))1	0.1210	1.1286	0.0514	0.04190	2.90	0.00377 **	0.0203	0.2220

From the fitted multiplicative part of the Cox-Aalen model (Figure 2, the plots of the cumulative hazards for CD4 baseline, TB incidence, and prior-ART TB are presented. The steeper slope for the females (gender=0) compared to males (gender=1) suggests that immune recovery (CD4 above 350) is expected sooner for females than males. For patients who did not have TB prior to ART, immune recovery is also expected sooner compared to patients with TB prior to ART.

In order to evaluate the goodness of fit of the covariates included in the multiplicative part of the Cox-Aalen model, it is considered a cumulative score process. To assess the influence of the significant covariates on the risk of progressing to a CD4 cell count higher than 350, the graph of the cumulative regression coefficients against the survival time was plotted, showing the 95% pointwise confidence intervals (Figure 3. The slope of the plot informs about the behaviour of the covariate on the hazard of CD4 above 350 over time (immune recovery). The confidence intervals for the intercept, gender, virologic failure, and CD4 baseline (ungrouped) do not include the zero line, hence they are significant. The plot in the top right corner is the estimated baseline cumulative hazards function (intercept). With its positive slope, the plot suggests that when there are no covariates, there is an overall significant and increasing hazard of immune recovery. The plot on the top right corner represents the additive effect of the CD4 baseline (a continuous variable. The positive slope indicates that an increase in the CD4 cell count at baseline increases the risk of immune recovery. Row 2 and column 1 show the plot for gender, and it decreased steadily with a negative slope throughout the follow-up time. This plot suggests that males have a decreased risk of immune recovery compared to their female counterparts. On the second row and second column is the plot for virologic failure, and its negative slope suggests that virologic failure reduces the risk of immune recovery. The bottom right corner is the slope for CD4 baseline with a negative slope, which indicates that starting ART with a CD4 baseline below 350 reduces the risk of immune recovery.

The additive effect of age of an HIV patient was assessed by dividing it into four equal intervals. The estimated cumulative functions for each quadrant are presented in Figure 4 below.

The fitness of the Aalen additive model was examined for the continuous variables, age and CD4 baseline, using a series of cumulative martingale residuals with 100 random simulations under the null. These variables are divided into quartiles to identify the most influential categories, and the results are presented in Figures 5. However, Martinussen and Scheike recommended using P-values to ascertain the extent of the departure from the null that is observed when a large size of residuals is calculated [25]. The p-values for the test for cumulative MG-residuals are presented in Table 7. The p-values for the observed cumulative residuals for the CD4 baseline category (336, 499] and the age category (20, 26.5] are 0.26 and 0.08, respectively. This confirms that the CD4 baseline category (336, 499] and the age category (20, 26.5] fit the data well (Figure 5). The observed cumulative residuals for the other categories have P-values that favour the alternative hypothesis that the residuals are significantly different from zero. Thus, leading to unacceptable fits as shown in the figures below.

4. Discussion

The Cox model assumes constant proportional hazards. Aalen's additive model is a non-parametric model that allows time-varying covariate effects in which covariates are modelled as additive risks to a baseline hazard and are allowed to vary over time. Combining these two models helps to overcome the limitations of a single model.

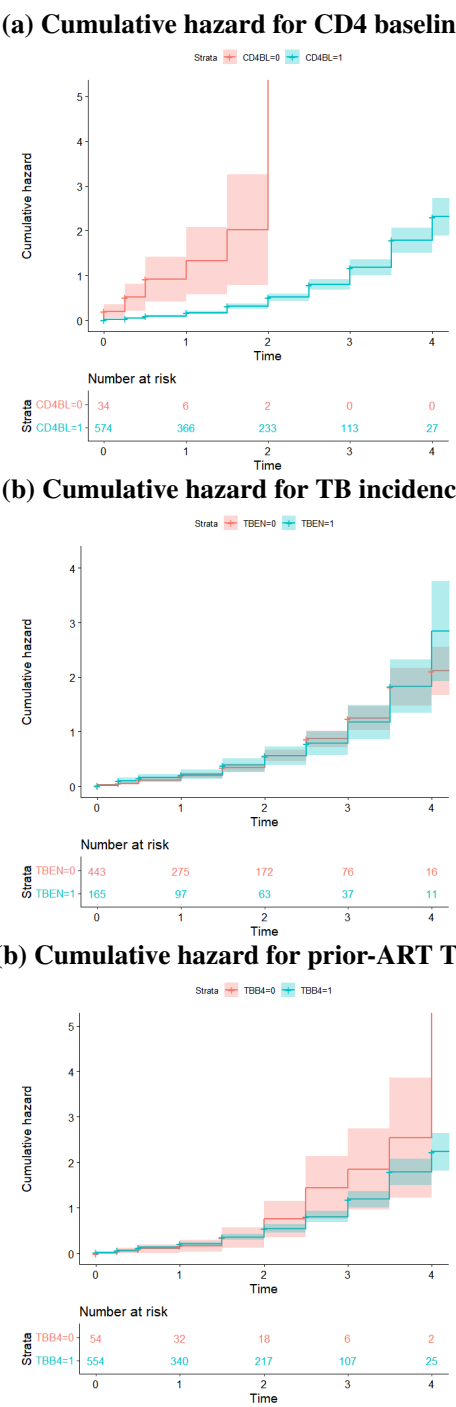

Figure 2. Plots of Cumulative hazards for the covariates in the multiplicative part of the Cox-Aalen Model

The Cox-Aalen model is more comprehensive in capturing both relative risk (Cox component) and the absolute effects over time (Aalen component). By combining the Cox proportional hazards model and Aalen's additive model, this study analyses the risk of immune recovery using data from a retrospective cohort study from South Africa.

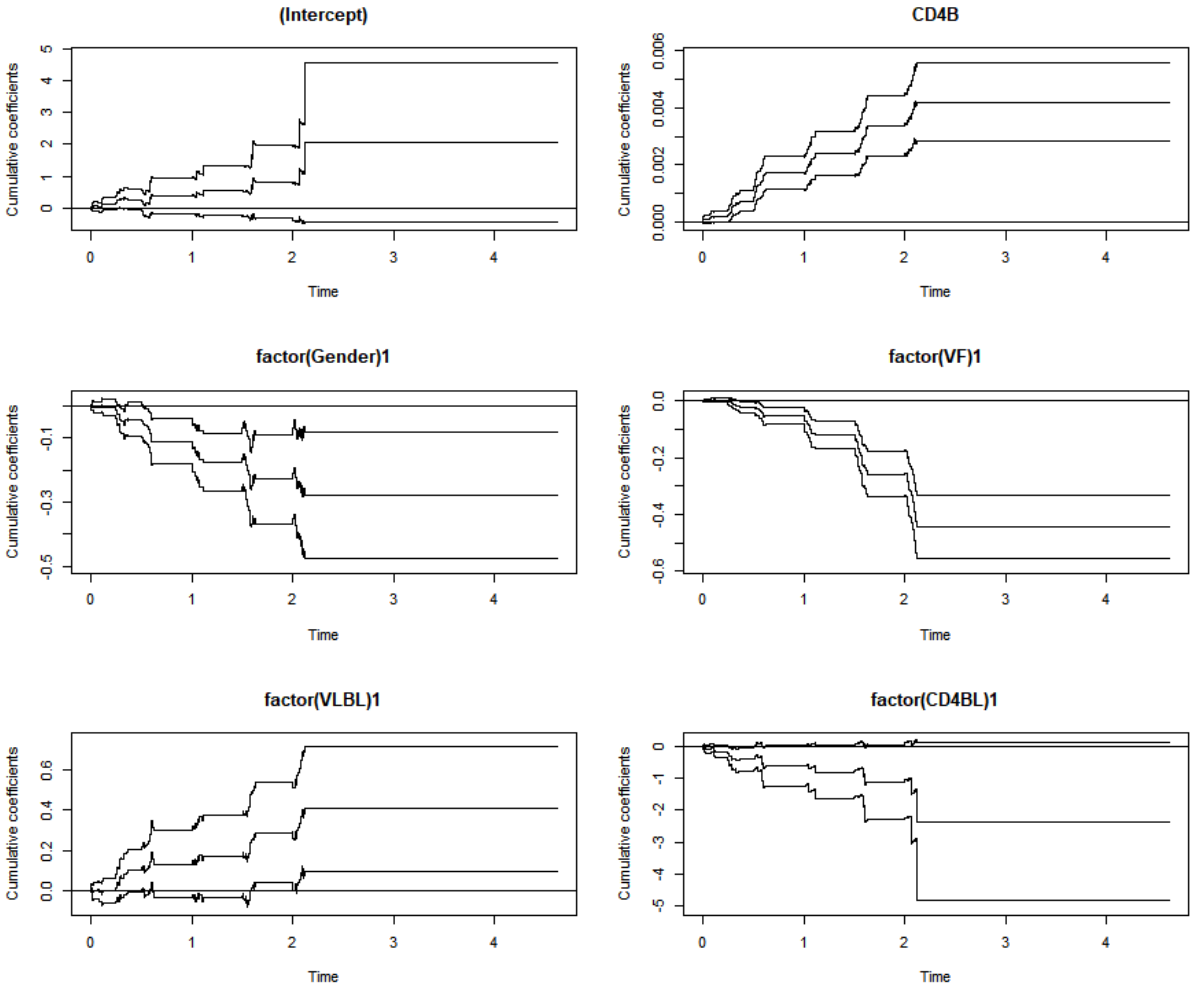

Figure 3. EEstimated cumulative regression functions with 95% pointwise confidence intervals based on Aalen's additive model

Table 7. Test for Cumulative MG-residuals for additive effect of the continuous variables, age and CD4 baseline part of Cox-Aalen model: Final output-Part B

	sup$\hat{\beta}(t)$	p-value $H_0 : \beta(t) = 0$
factor(cut(CD4B, 4))(9.35,173]	68.166	0.00
factor(cut(CD4B, 4))(173,336]	47.285	0.00
factor(cut(CD4B, 4))(336,499]	3.550	0.26
factor(cut(CD4B, 4))(499,663]	8.999	0.00
factor(cut(Age, 4))(14,20.2]	12.110	0.01
factor(cut(Age, 4))(20.2,26.5]	5.878	0.08
factor(cut(Age, 4))(26.5,32.8]	52.239	0.00
factor(cut(Age, 4))(32.8,39]	57.773	0.00

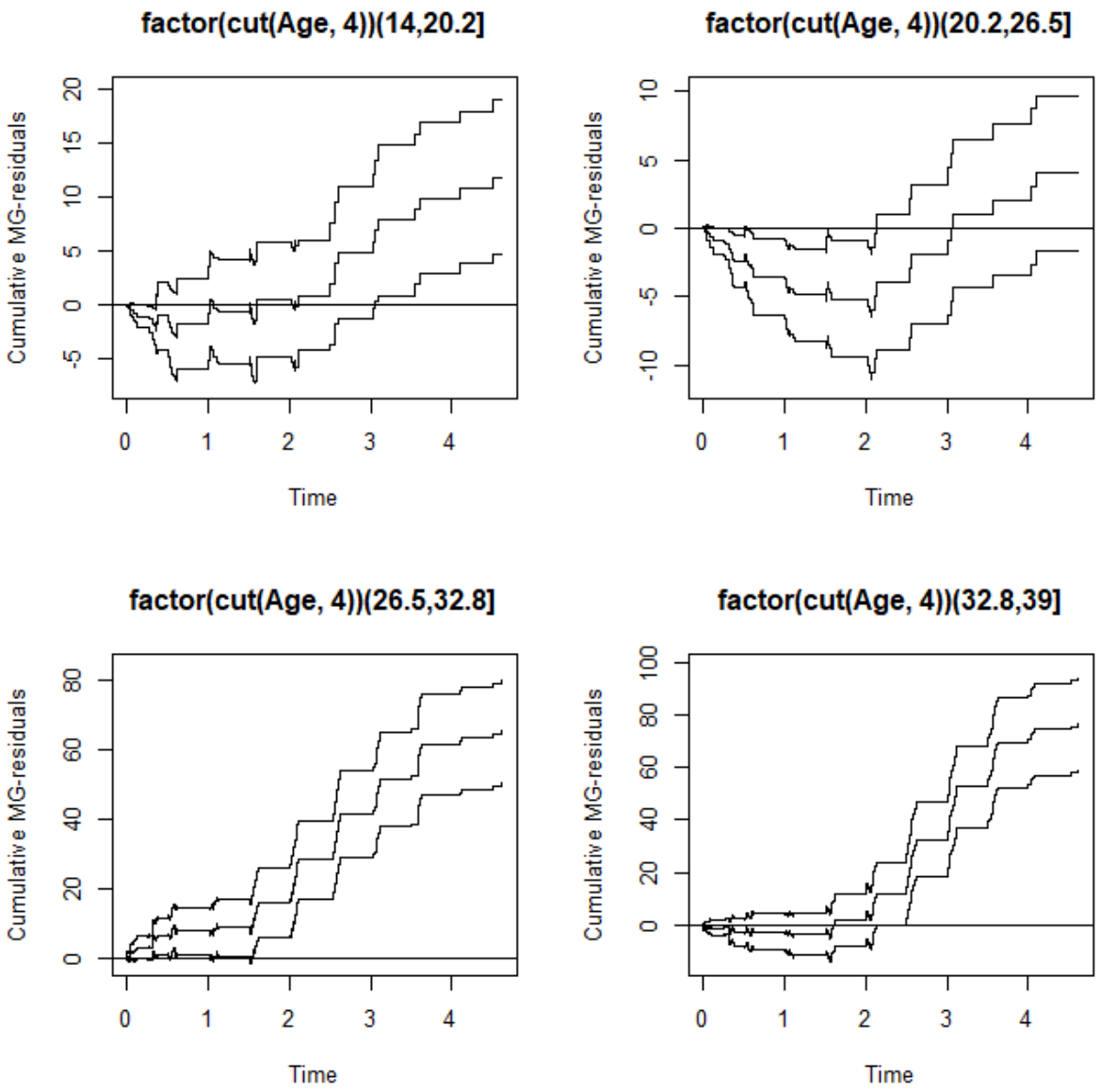

Figure 4. Cumulative residuals with 95% pointwise confidence intervals for 4 age groups with 100 simulations based on Aalen's additive model

The results from the analysis single Cox proportional hazards function suggest that age of patient, having TB prior to ART, virologic failure, TB incidence post ART commencement, and CD4 baseline satisfy the proportional hazards assumptions based on the Schoenfeld residuals test. However, from the final reduced Cox model, virologic failure was dropped. Goodness of fit for the Cox proportional hazards model was based on the likelihood ratio, Wald and Score tests, which show a highly significant fit of the model. On the other hand, the Aalen's additive model suggests that gender, virologic failure, CD4 baseline, and age had a statistically significant time-varying effect based on the results from the supremum test. Combining the two models (Cox-Aalen model) improves the significance of the covariates, both from the additive and the multiplicative classes, with more covariates being

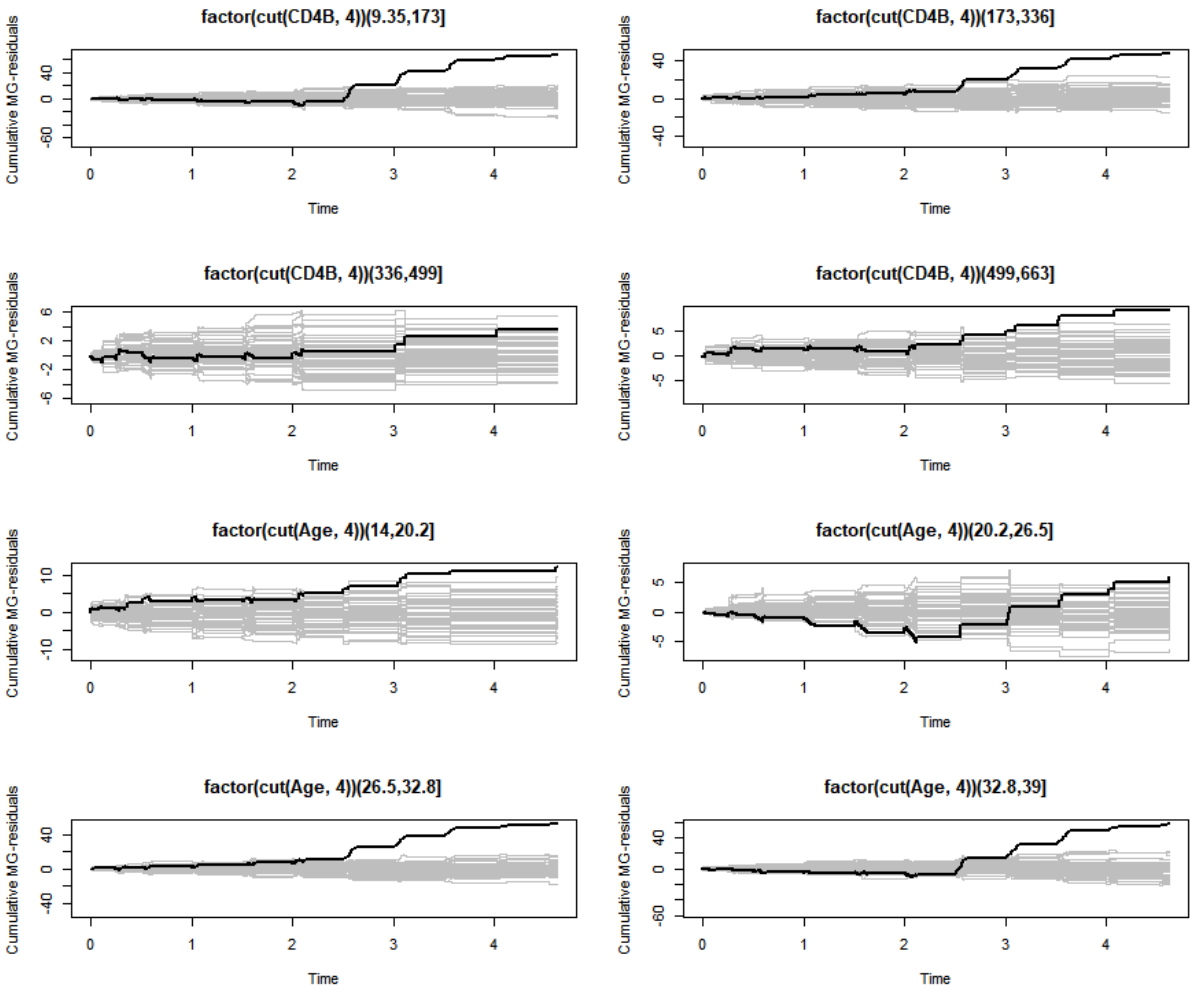

Figure 5. Score test for constant multiplicative effect for patient CD4 baseline and Age in the additive part of the Cox-Aalen model with 100 randomly chosen score processes under the null of a constant multiplicative effect.

added to the additive part of the model. Thus, if a single model were used, some covariates could have been dropped, yet their contribution is significant if properly classified. This makes the Cox-Aalen model a powerful tool for survival analysis.

A comparison of the Cox model and Aalen's additive model is based on the p-values for the selected covariates. The results show that p-values in Aalen's additive model are fairly lower than the p-values in the Cox model. This could be because non-parametric models are generally less powerful in detecting significant effects compared to parametric models [17]. The overall significance of Aalen's model is higher than the overall significance of the Cox model. However, it is not possible to base our conclusion on this because Aalen's model uses a χ^2 test and the Cox model uses a likelihood ratio test. However, besides these differences, the Cox model and Aalen's additive model give almost similar results as far as the selection of covariates to remain in the final model is concerned. The signs of the coefficients of covariates for both models are also the same. Hence, Aalen's additive model can be regarded as complementary to the Cox model. These findings agree with the results from studies on survival from breast cancer [15, 16].

The results from both Aalen's additive hazard model and Cox regression model show that the covariates, TB prior ART, adverse reaction to treatment (reaction), WHO stage baseline (WSBL), BMI progression, and viral load baseline (VLBL) are not significantly associated with HIV/AIDS progression for patients under treatment therapy. The other variables, CD4 baseline (CD4BL), gender, and developed TB during treatment, contribute significantly to HIV/AIDS progression. These findings are in corroboration with the findings from multistate modelling by Shoko and Chikobvu [26]. Although the variable age did not have any significant effect on the event of interest, it was not removed from the model because of its epidemiological importance.

Results from this study show that males have poor adherence to treatment compared to their female counterparts. There are several potential reasons for male vulnerability as highlighted by the literature [27, 31, 32, 33]. These include stigma, delayed ART initiation and biological factors. A study carried out in 2014 shows that males have higher patient attrition and mortality compared to females, and this may be attributed in part to late presentation for HIV treatment and care [27]. Women are biologically more vulnerable to acquiring HIV infection than males [33], and on average, women may exhibit slightly better immunological response to ART [36]. Mwamba et al. [31] identified one of the factors influencing delayed ART initiation being patients with a history of a gap in HIV. Males' delayed reporting to health facilities could be a result of differences in social responsibilities, which give women more entry points to HIV services, for instance, during pregnancy 30. Perceived social stigma around HIV continues to be a challenge for treatment initiation [31]. Patients who develop TB whilst on treatment, if diagnosed and treated, are likely to have a better HIV treatment outcome than their counterparts. Another study also confirms that early recognition and appropriate management of these consequences can reinforce the successful integration of therapy in HIV-infected patients with TB [28]. Commencing HIV/AIDS treatment when the CD4 cell count is below 350 cells/mm^3 results in better treatment adherence than starting treatment when CD4 cell counts are above 350 cells/mm^3. This finding is supported by the work from Sabin et al., which also recommended commencement of ART when the CD4 cell count is below 350 cells/mm^3 and to improve treatment adherence [29, 30]. Initially, the WHO recommended ART initiation at a CD4+ cell count of 200 cells/μl or lower, but recent updates to these guidelines increased the recommended threshold for ART initiation to 250, 350, and 500cells/μl [33]. The second recommendation by WHO, which was based on clinical trial results, confirms the efficacy of the ARV drug tenofovir for use as PrEP to prevent people from acquiring HIV in a wide variety of settings and populations, especially the most vulnerable populations [33]. However, for future studies on HIV/AIDS, the use of Markov multistate models is recommended because they allow monitoring of disease progression through multiple mutually exclusive states.

5. Conclusion

In this study, both the Cox proportional hazards model and Aalen's additive model were used on HIV data. The Cox model and Aalen's model both yield similar results with regard to the selection of covariates that contribute significantly to HIV progression. These covariates are gender, CD4 baseline, and the development of TB. Males have an increased risk of immune deterioration than females. Patients who start ART with lower CD4 cell counts have smaller risks of immune deterioration. Developing TB on ART and having it monitored increases the rates of immune recovery. TB Coinfection: Development of TB during treatment is significantly linked to progression, underscoring the need for integrated HIV/TB care—a critical public health message.

5.1. Limitation

Cohort Homogeneity: The sample is limited to Limpopo Province, South Africa. Results may not generalise to other regions with different ART access, TB prevalence, or viral subtypes. Short FollowUp: Follow-up every 6 months (2005–2009) may miss long-term ART effects. Recent data (e.g., post2010 ART guidelines) could provide updated insights.

Significance statement

This study presents a comprehensive analysis of the determinants of HIV progression in patients receiving antiretroviral therapy by combining Aalen's additive model and the Cox proportional hazard model. The study findings reinforce the need for early initiation of treatment for improved treatment outcomes and constant TB diagnostics to reduce mortality rates among HIV-co-infected individuals. This study suggests that in order to relay accurate information to the intended audience, at least two models can be used so that these models complement each other. The purpose of this study is to model the progression of HIV among a TB co-infected cohort.

Acknowledgment

We are grateful for the cooperation of the study participants in data collection. POB's research was supported by the South African Medical Research Council (RCDI) through funding received from the South African National Treasury; the South African National Research Foundation (GUN109312, GUN86037), and the University of Venda. The views expressed are solely the responsibility of the authors and do not necessarily represent the official views of the South African Medical Research Council, the National Research Foundation, or the University of Venda.

Author's Contribution:

Claris Shoko devised the initial idea and drafted the first manuscript. Delson Chikobvu finalised and proofread the article. Gorata D. Manyeagae wrote the conclusion, edited and proofread the article. Claris Shoko and Delson Chikobvu contributed to the analysis and interpretation of the data. Pascal O. Bessong collected the data used in the current study and aided with both revisions and proofreading of the final manuscript. All authors participated in critically revising the manuscript drafts and approved the final version.

REFERENCES

1. A.A. Okoye and L.J. Picker, *CD4(+) T-cell depletion in HIV infection: mechanisms of immunological failure*, 1 depletion in HIV infection: mechanisms of immunological failure. Immunol Rev, vol. 254, no.1, pp. 54-64, 2013. doi:10.1111/imr.12066..
2. G. D. Biase, A. D. Girolamo, J. Janssen, S. Iacobelli, N. Tinari, and R. Manca, *A Stochastic Model for the HIV/AIDS Dynamic Evolution*, Mathematics Problems in Engineering, vol. 2007, no. 1, 065636, 2007, https://doi.org/10.1155/2007/65636.
3. R. Ying, R. M. Granich, S. Gupta, B. G. Williams, *CD4 Cell Count: Declining Value for Antiretroviral Therapy Eligibility*, Clin Infect Dis, vol. 62, no. 8, pp. 1022-1028, 2016, doi:10.1093/cid/civ1224.
4. E. Moyo, P. Moyo, G. Murewanhema, M. Mhango, I. Chitungo, T. Dzinamarira, *Key populations and Sub-Saharan Africa's HIV response*, Front Public Health, vol. 16, 11:1079990, 2023. doi:10.3389/fpubh.2023.1079990.
5. J. E. Bennett, R. Dolin, MD, M. J. Blaser, , *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases, 8th Edition*, Elsevier Saunders, Philadelphia, vol. 1, pp. 2584-2589, 2015.
6. Q. Cai, L. Yang, Y. Ling, W. Pan, Q. Zhong, C. Wang, and X. Pan, *Survival prediction models for people living with HIV based on four machine learning models*, Scientific Reports, vol. 15, no. 1, pp. 1-11, 2025, https://doi.org/10.1038/s41598-025-16479-3.
7. G. Zhang, Y. Gong, Q. Wang, L. Deng,L.et al., *Outcomes and factors associated with survival of patients with HIV/AIDS initiating antiretroviral treatment in Liangshan Prefecture, southwest of China. A retrospective cohort study from 2005 to 2013*, Medicine, vol 95, no. 27(e3969), 2016, http://dx.doi.org/10.1097/MD.0000000000003969
8. T. D. Mangal, M. V. Meireles, A. R. P. Pascom et al., *Determinants of survival of people living with HIV/AIDS on antiretroviral therapy in Brazil 2006–2015*, BMC Infect Dis, vol. 19, 206, 2019 https://doi.org/10.1186/s12879-019-3844-3.
9. J. D. Kowalska, J. Kubicka, E. Siwak et al., *Factors associated with the first antiretroviral therapy modification in older HIV-1 positive patients*, AIDS Res Ther, vol. 13, no. 2, 2016. https://doi.org/10.1186/s12981-015-0084-5
10. E. Basar, *Aalen's additive, Cox proportional hazards and the Cox-Aalen Model: Application to kidney transplant data*, Sains Malaysian , vol 46, no. 3, pp. 469–476, 2017, http://dx.doi.org/10.17576/jsm-2017-4603-15.
11. X. Xie, H. D. Strickler, and X. Xue, *Additive Hazard Regression Models: An Application to the Natural History of Human Papillomavirus*, Computational and Mathematical Methods in Medicine, 2013, 796270. https://doi.org/10.1155/2013/796270

12. A. Zierle-Ghosh, A. Jan, *Physiology, Body Mass Index*, StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK535456/
13. D. Y. Lin, and Z. Ying, *Semiparametric analysis of the additive risk model*, Biometrika 81, 61–71, 1994.
14. L. Sun, et al., 2006. *Modeling the Subdistribution of a Competing Risk* Statistica Sinica, vol. 16, pp. 1367–1385.
15. A. Abadi, S. Saadat, P. Yavari, C. Bajdik, and P. Jalili, *Comparison of Aalens Additive and Cox Proportional Hazards Models for Breast Cancer Survival: Analysis of Population-Based Data from British Columbia, Canada*, Asian Pacific J Cancer Prev, vol. 12, pp. 3113-3116, 2011.
16. O.O. Aalen, *A model for nonparametric regression analysis of counting processes*, Lecture Notes in Statistics, vol. 2, pp.1-25, 1980, New York: Springer.
17. E. Basar *Additive, Cox Proportional Hazards and The Cox-Aalen Model: Application to Kidney Transplant Data*, Sains Malaysiana, vol. 46, no. 3, pp. 469–476, 2017. http://dx.doi.org/10.17576/jsm-2017-4603-15Aalens
18. Aalen, O.O., 1989. *A linear regression model for the analysis of life times*, Statist. Med., vol. 8, pp. 907-925.
19. F. Lefebvre, R. Giorgi, *A strategy for optimal fitting of multiplicative and additive hazards regression models*, BMC Med Res Methodol, vol. 21, 100, 2021, https://doi.org/10.1186/s12874-021-01273-2
20. Devarajan, K, Ebrahimi, N., 2009. *Testing for Covariate Effect in the Cox Proportional Hazards Regression Model*, Commun Stat Theory Methods, vol. 38, no. 14, pp. 2333–2347. doi:10.1080/03610920802536958
21. J. Fan, H. Lin, and Y. Zhou, 2006. *Local Partial-Likelihood Estimation for Lifetime Data*, The Annals of Statistics, vol. 34, no. 1, pp. 290–325, 2006. DOI:10.1214/009053605000000796
22. T. H. Scheike, and M. J. Zhang, *An additive-multiplicative Cox-Aalen model*, Scand. J. Statist. vol. 28, pp. 75-88, 2002.
23. J. P. Klein, M. L. Moeschberger, *Survival Analysis: Techniques for Censored and Truncated Data*, 2nd ed. Springer; 2003
24. H. L. Cao, *A Comparison Between the Additive and Multiplicative Risk Models*, 2005
25. T. Martinussen, T. H. Scheike. *Dynamic Regression Models for Survival Data*, Springer Science and Business Media, 2007.
26. C. Shoko, D. Chikobvu and P.O Bessong, *A Markov model for the effects of virologic failure on HIV/AIDS progression in TB co-infected patients receiving antiretroviral therapy in a rural clinic in northern South Africa*, South African Medical Journal, vol. 110, no. 4, pp. 313-319, 2020, https://doi.org/10.7196/SAMJ.2020.v110i4.13934.
27. K. C. Takarinda, A. D. Harries, R.. W. Shiraishid, A. T. Mutasa, A. Abdul-Quaderd, and O. Mugurungi, *Gender-related differences in outcomes and attrition on antiretroviral treatment among an HIV-infected patient cohort in Zimbabwe: 2007–2010*, Int J Infect Dis., vol. 30, pp. 98—105, 2015. doi:10.1016/j.ijid2014.11.009
28. W. Manosuthi, S. Wiboonchutikul and S. Sungkanuparph, *Integrated therapy for HIV and tuberculosis*, AIDS Res Ther., vol. 13, no. 22, 2016. https://doi.org/10.1186/s12981-016-0106-y
29. C. A. Sabin, A. N. Phillips, 2009. *Should HIV therapy be started at a CD4 cell count above 350 cells/µl in asymptomatic HIV-1-infected patients?* Current Opinion in Infectious Diseases, vol. 22, no. 2, pp.191–197, 2009, https://www.who.int/hiv/events/artprevention/sabine.pdf.
30. C. Shoko, D. Chikobvu, P. O. Bessong, *Effects of Antiretroviral Therapy on CD4+ Cell Count, HIV Viral Load and Death in a South African Cohort: A Modelling Study*, Pakistan Journal of Biological Sciences, vol. 23, no. 4, pp. 542–551. DOI:10.3923/pjbs.2020.542.551.PMID:32363840.
31. C. Mwamba, L. K. Beres, S. M. Topp, N. Mukamba, S. Simbeza, K. Sikombe, ..., and C.Bolton Moore, *'I need time to start antiretroviral therapy': understanding reasons for delayed ART initiation among people diagnosed with HIV in Lusaka, Zambia'*, Annals of Medicine, vol. 54, no. 1, pp. 830–836, 2022. https://doi.org/10.1080/07853890.2022.2051069
32. T. Cassidy, M. Cornell, B. Makeleni, C. R. Horsburgh, L. T. Duran, V. de Azevedo, ... and M. P. Fox, *Attrition from care among men initiating ART in male-only clinics compared with men in general primary healthcare clinics in Khayelitsha, South Africa: a matched propensity score analysis*, AIDS and Behavior, vol. 27, no. 1, pp. 358-369, 2023.
33. T. Girum, A. Wasie, K. Lentiro, et al. *Gender disparity in epidemiological trend of HIV/AIDS infection and treatment in Ethiopia*, Arch Public Health. vol. 76, no. 51, 2018. https://doi.org/10.1186/s13690-018-0299-8
34. A. Kuznik, G. Iliyasu, A. G. Habib, B. M. Musa, A. Kambugu, and M. Lamorde, *Initiation of antiretroviral therapy based on the 2015 WHO guidelines*, Aids, vol. 30, no. 18, pp. 2865-2873, 2016.
35. F. Mosha, *Gender Differences in Human Immunodeficiency Virus (HIV) Disease Progression and Treatment Outcomes*, IntechOpen, 2021. doi:10.5772/intechopen.92898
36. K. Taylor-Smith, H. Tweya, A. Harries, E. Schoutene, and A. Jahn, *Gender differences in retention and survival on antiretroviral therapy of HIV-1 infected adults in Malawi*, Malawi Medical Journal, vol. 22, no. 2, 2010, https://doi.org/10.4314/mmj.v22i2.58794