

# Analysis of Recurrent Events with Associated Informative Censoring: Application to HIV Data

Jonathan Ejoku<sup>1</sup>, Collins Odhiambo<sup>1,\*</sup> and Linda Chaba<sup>1</sup>

<sup>1</sup>Strathmore Institute of Mathematical Sciences, Strathmore University, Madaraka Estate, Ole Sangale Road, P.O. Box 59857, 00200, City Square, Nairobi, Kenya

**Abstract:** In this study, we adapt a Cox-based model for recurrent events; the Prentice, Williams and Peterson Total-Time (PWP-TT) that has largely, been used under the assumption of non-informative censoring and evaluate it under an informative censoring setting. Empirical evaluation was undertaken with the aid of the semi-parametric framework for recurrent events suggested by Huang [1] and implemented in R Studio software. For validation we used data from a typical HIV care setting in Kenya. Of the three models under consideration; the standard Cox Model had gender hazard ratio (HR) of 0.66 (p-value=0.165), Andersen-Gill had HR 0.46 (with borderline p-value=0.054) and extended PWP TT had HR 0.22 (p-value=0.006). The PWP-TT model performed better as compared to other models under informative setting. In terms of risk factors under informative setting, LTFU due to stigma; gender [base=Male] had HR 0.544 (p-value =0.002), age [base is < 37] had HR 0.772 (p-value=0.008), ART regimen [base= First line] had HR 0.518 (p-value= 0.233) and differentiated care model (Base=not on DCM) had HR 0.77(p-value=0.036). In conclusion, in spite of the multiple interventions designed to address incidences of LTFU among HIV patients, within-person cases of LTFU are usually common and recurrent in nature, with the present likelihood of a person getting LTFU influenced by previous occurrences and therefore informative censoring should be checked.

**Keywords:** Recurrent events, Loss to follow-up, HIV, Prentice, Williams and Peterson Gap-Time, Informative censoring.

## 1. BACKGROUND

Recurrent events occur in a variety of disciplines/areas of life such as recurrent opportunistic infections in HIV patients, episodes of asthmatic attacks. Underlying processes that generate data from these events are called “recurrent event processes and the data they provide are called “recurrent event data” [3]. There is a wide range of research on the analysis of this data including the analysis of recurrence of sports injuries, multiple episodes of childhood diseases [4] and hospitalizations for chronic kidney disease [5].

A key characteristic of recurrent events is that observations per individual are usually not independent and in many cases are correlated, with current event incidences influenced by previous incidences. A number of approaches that have been proposed to analyze recurrent events data, including both parametric and non-parametric methods such as the use of the Poisson and Negative binomial models [3]. Both of these however accommodate only time-independent covariates, with the Poisson further pre-supposing a constant events rate per individual.

While these and a number of other models have been suggested, the models most predominant in recurrent event literature for ordered outcomes are the Anderson and Gill [6], Prentice, Williams, and Peterson total time (PWP-TT) and gap time (PWP-GT) models

[7], Wei, Lin and Weissfeld (WLW) [8]; and Lee, Wei and Amato (LWA). These models are essentially extensions of the extensively used semi-parametric Cox Proportional Hazards Model, predominantly applied under the assumption of non-informative censoring.

To the best of our knowledge, none of these models have been applied to researches with an underlying assumption of informative censoring. Subsequently, we hope this paper will significantly bridge the gap by evaluating the common recurrent event based models in situations where informative censoring exists.

Informative censoring in HIV related-studies exist in situations, where there are informed drop-outs; such as sicker persons or those in either WHO staging III or IV being withdrawn from the study or persons getting lost to the study due to stigma-related factors. Ghosh *et al.*, [9] note that informative censoring occurs in situations where the censoring times “depend on the observed or non-observed recurrent times”. An illustration of this in the HIV setting, is the informative drop-outs such as the withdrawal of sicker patients from a study way before the study ends. Ghosh *et al.*, [9] specify two possible sources of informative censoring in practical settings:- the voluntary withdrawal of subjects for reasons related to the event process as well death.

The exclusion of censoring assumption from different studies may generally lead to biased estimates. Castelli *et al.* [10] in adapting the Inverse Probability of Censoring Weighted [IPCW] to study the survival times of asthma patients while including informative censoring [patients felt they were okay and

\*Address correspondence to this author at the Strathmore Institute of Mathematical Sciences, Strathmore University, Madaraka Estate, Ole Sangale Road, P.O. Box 59857, 00200, City Square, Nairobi, Kenya; Mob: +254 703 034 000/200/300 or [+254] [0] 730-734000/200/300; Fax: +254 20 6007498; E-mail: codhiambo@strathmore.edu

did not need to consult a doctor] show that information coming from censoring process improves the survival estimate. Another approach is that by Ghosh *et al.*, [9] where they introduce a semi-parametric approach for recurrent events in the presence of dependent censoring and apply it ALIVE cohort study. A comparison with the accelerated failure times model as proposed by Lin *et al.* [11], which assumes non-informative censoring found that their proposed method yielded “a much larger estimate for the effect of baseline HIV status on hospitalizations than the method proposed by Lin”.

There is a body of literature to show that drop-outs in HIV programs are informative. Berheto *et al.*, [12] in assessing the predictors of LTFU in Patients Living with HIV/AIDS after Initiation of Antiretroviral Therapy found that sicker patients [baseline CD4 < 200 cells/mm<sup>3</sup> – HR 1.7, 95% Cis 1.3-2.2, regimen substitution – HR 5.2; 95% Cis 3.6-7.3] and receiving non-isoniazid (INH) prophylaxis [HR 3.7; 95% CIs 2.3-6.2] as accelerators for LTFU. According to Assemie *et al.*, [13], being on WHO clinical stage IV as well as receiving isoniazid preventive therapy were significant predictors of LTFU. On the premise of this, informative censoring is thus incorporated in this study.

**2. METHOD AND SETTING**

Let the rate of occurrence of recurrent events in a given time interval i.e. (0, Γ<sub>o</sub>), where Γ<sub>o</sub> > 0 is set with the information of possible epochs of recurrences observable up to Γ<sub>o</sub>. Suppose N(t) is the number of recurrent events that occur on or before t, t > 0. The functional rate of a recurrent event at t, t ∈ [0, Γ<sub>o</sub>], is expressed as

$$\lambda(t) = \lim_{\Delta \rightarrow 0^+} \frac{\Pr[N[t + \Delta] - N[t] > 0]}{\Delta}$$

The rate is considered theoretically different from the intensity function. We define the functional rate as

the occurrence rate of recurrent events and is not conditional to the event history. The intensity function on the other hand is the occurrence rate and is conditional on the event history.

Let the cumulative rate function be described as

$$\Lambda(t) = \int_0^t \lambda[\mu] du.$$

Also, suppose Y be the censoring time for observation of the recurrent event process at termination. The interest remains the occurrence rate in the time interval (0, Γ<sub>o</sub>), and therefore, recurrent event data beyond Γ<sub>o</sub>, is not important. We define τ = min [Y, Γ<sub>o</sub>] as the new censoring time used in the proposed models.

Consider PWP-TT defined as λ<sub>ik</sub>(t) = λ<sub>ok</sub>(t) exp{X<sub>ik</sub>β} with the following features events are ordered and handled by stratification, everyone is at risk for the first stratum, but only those who had an event in the previous stratum are at risk for the successive one. The model estimates both overall and event-specific effects and uses robust standard errors to account for correlation.

**2.1. Review of the Prentice, Williams, and Peterson [PWP] Model**

The PWP model as proposed by Prentice, Williams, and Peterson [7] is a ‘conditional model’ where the individuals are at risk for an event if and only if they were at risk for a previous event. To achieve this, each event occurrence is put into a different stratum with all participants at risk in the first stratum. Under this model, only participants that experienced the previous event would then be at risk for the next event [14]. When time since entry is of interest, this model condenses to total time model (PWP-TT), while when time since last event is of interest the model becomes a gap time model

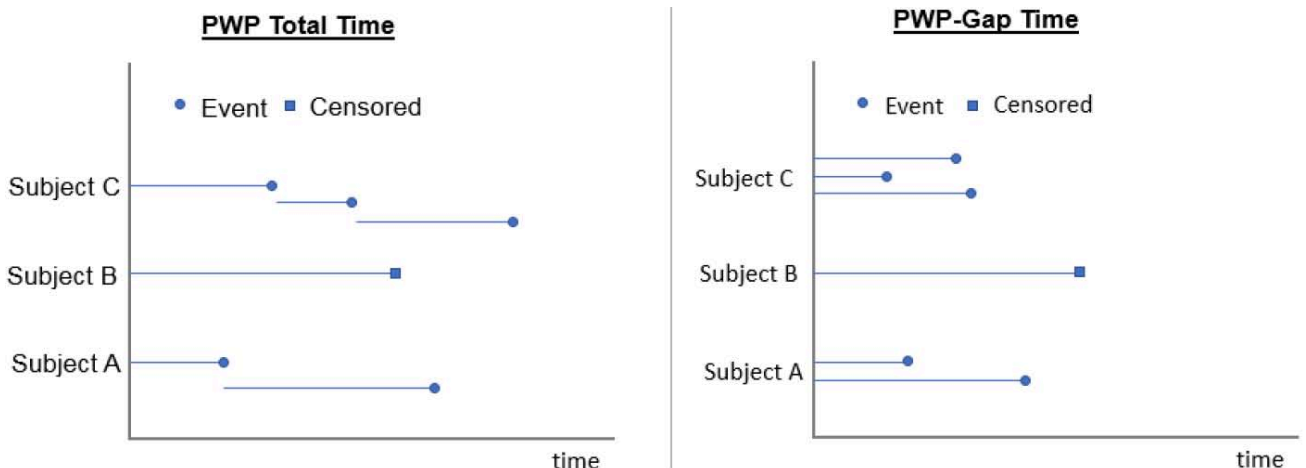


Figure 1: Representation of the PWP Models.

commonly abbreviated as PWP-GT. See Figure 1. The key difference between this model and the Andersen-Gill (AG) is in terms of the effects of the covariates in different strata. In the AG model, effects of covariates are constant across all strata, while this varies in different strata for the PWP [14].

**2.2. Mathematical Representation of PWP Model**

To include informative censoring, we built on generic the approach proposed by Huang *et al.* [1] where they jointly model the recurrent event process and failure times. The key to this approach is to model this relationship via a subject specific latent variable,  $Z_i$ , that models the association between the intensity of the recurrent event and the hazard of the failure time. This approach is able to account for only time-independent predictors while also allowing for informative censoring.

As relayed by Huang *et al.*, the intensity function is

$$\lambda_i(t) = Z_i \lambda_o(t) \exp\{x\alpha\}.$$

And Hazard of the failure time, D, is given by:

$$h(t) = Z_i h_o(t) \exp\{x\beta\}.$$

Where:

$\lambda_o$  and  $h_o$  are baseline intensity and hazard functions respectively.

$\alpha$  and  $\beta$  are coefficients.

$X_i[t]$  is the p-dimensional set of time-dependent covariates

Based on the proposed intensity function by Huang, the PWP models are modified as below. A mathematical representation of each of the models is presented below, as well as the individual model specification with informative censoring incorporated (Table 1). Essentially, each model is multiplied by the unobserved frailty,  $Z_i$ .

Where:-

PWP-GT is the Prentice, Williams, and Peterson Gap-Time Model,

PWP-TT is the Prentice, Williams, and Peterson Total-Time Model

And: -

$M_i(t)$  is the at risk indicator for the j-th event and ith person at time t. This is 1 when at risk for event j, and zero when not at risk for event j

$\lambda_o$  is the baseline intensity function

$X_i[t]$  is the p-dimensional set of time-dependent covariates

$\gamma$  is a q x 1 dimensional parameter.

**2.3. Evaluation of the PWP-GT model under informative censoring**

**2.3.1. Parameter Estimation**

The key to this approach is to use a subject specific latent variable to model the association between the recurrent event and hazard of the failure times. Additionally, no assumptions are made on the distributions of censoring times and latent variables. Specifically, we estimate the cumulative hazard and intensity functions. For estimation, we use the approach relayed by Huang *et al.* as it remains consistent with the PWP model with the at risk indicator treated as a nuisance parameter.

The estimation by Huang is briefly provided below. Starting with notation:-

Let  $\mathbf{N}[t]$  which represents events occurring at or before some time  $T_o$ ,  $\mathbf{D}$  be failure time, say LTFU due to stigma for this study,  $\mathbf{C}$  failure time due to other reasons than D,  $\mathbf{x}$  is a vector of 1xp covariates.

Let  $\mathbf{Y}$  be the point at which the observation of recurrent events ceases, such that,  $Y = \min[C, D, T_o]$ .

We introduce a non-negative latent variable z, such that given  $X=x$ , and  $Z=z$ , the **intensity function** is given by

$$\lambda_i(t) = Z_i \lambda_o(t) \exp\{x\alpha\},$$

with  $E[Z|x] = E[z]$

And Hazard function by

$$h(t) = Z_i h_o(t) \exp\{x\beta\}.$$

Implicitly, a large/small z increases or decreases the intensity and hazard respectively. The rate function is defined as  $\mu_z \lambda_o(t) \exp [x\alpha]$

**Table 1: Mathematical Representation of the Intensity Functions of PWP Models with Informative Censoring**

Name of The Model	Recurrent Model Mathematical presentation	Form under Informative/Non-informative censoring
PWP-GT	$M_{ij}(t)\lambda_{oj}[t - t_{j-1}] \exp [x'\beta]$	$Z_i M_{ij}(t)\lambda_{oj}[t - t_j - 1] \exp [x'\beta]$
PWP-TT	$M_{ij}(t)\lambda_{oj}[t] \exp [x'\beta]$	$Z_i M_{ij}(t)\lambda_{oj}[t] \exp [x'\beta]$

Assuming that for individual  $i$ , the collection  $\{x_i, z_i, m_i(t_i, \dots, 0, t_{im_i}), Y_i\}$  are iid, then the density function is given by

$$\frac{Z_i M_i \lambda_o(t) \exp[\alpha x_i]}{Z_i M_i \Lambda_o(t) \exp[\alpha x_i]}$$

However we assume that the density does not depend on  $Z_i$ ,  $M_i$ , nor  $x$ , reducing the density to

$$\frac{\lambda_o(t)}{\Lambda_o(t)}$$

Since we assume iid, the conditional likelihood is generated for  $n$  subjects, assuming  $m_i$  events per individual as

$$\prod_{i=1}^n \prod_{j=1}^{m_i} \frac{\lambda_o[t_{ij}]}{\Lambda_o(y_i)}$$

Huang notes that the rate estimator for  $\Lambda_o(y_i)$  can be given by  $\prod[1 - \frac{d}{R}]$  [1]

where  $d$  is the number of patients experiencing event at time  $t$ , and  $R$  is total persons at risk. To estimate the hazard  $\alpha$ , a class of estimators given below is solved.

$$n^{-1} \sum w_i x_i' [m_i \Lambda_o(Y)^{-1} - \exp(\alpha x_i)] = 0.$$

Huang proposes the estimation of  $\alpha$  by replacing  $\Lambda_o[Y]^{-1}$  by 1.

Asymptotic properties of these estimators are extensively studied by Huang and bear no repeating in this article.

### 2.3.2. Empirical Evaluation of Data

Patient-level data from four facilities in central Kenya, collected between 2013 and 2018 was used. The recurrent event of interest were the incidences of loss to follow-up. Censoring for a drop-out was considered informative, based on the reason provided for drop-out, a person cited 'stigma' or adverse drug reaction. Additionally, time independent covariates were incorporated in the analysis. Data preparation was prepared to reflect the PWP-GT outline format as laid out by Therneau [15]. Primarily, we sought to establish if the incorporation of informative censoring improves estimation of the hazard and rate estimates. While this study acknowledges the possibility of competing risks, they were not incorporated in this study.

### 2.3.3. Computing Environment

The models were assessed under the Rstudio computing environment. To account for informative

censoring, the 'method' function within 'reReg function' in the reReg was set to "cox.HW". Fitting the models was achieved by using the cluster and strata functions in the base survival package required by the reReg package. The Hmisc package was used to provide basic descriptive statistics.

## 3. APPLICATION TO HIV DATA

### 3.1. Data Description

Determination of the incidence of due to LTFU was computed using the cumulative baseline hazard technique with start of ART as time 0 and LTFU at the time a particular patient failed to return to CCC clinic for 38 weeks since the scheduled appointment date. Patients who did not experience the event [LTFU], were right censored at the last clinic visit. Other patients exit such as mortality and transfer out were considered as non-informative censoring as they relate to LTFU, as mortality and transfer out that were known to specific clinics and had happened within 38 days of the last clinical appointment were managed in a competing risk framework. After determination of LTFU, several consulted efforts to retrace clients who are lost are implemented. The efforts include calling back clients and visiting them at their homes with the aim of returning them back to clinic. Patients who return back to care are exposed to recurrent of the event [LTFU]. Among the sample of LTFU clients, we also collected other parameters include Patient ID: This is the patient ID which may be repeated due to recurrence of LTFU, time to LTFU: the event of interest is LTFU in this setting which is recurrent, status; for episodes of LTFU vs no episode; a particular individual can experience several LTFU episodes, event (Informative Censoring); the patient is no longer observed-Informative censoring was defined by the patient being stopped from highly active antiretroviral therapy (HAART) due to either drug reaction/ stigma, Differentiated HIV care for patients who are either under differentiated care model or not, age of the patient, and gender of the patient ARV regimen line as defined by WHO. All PLHIV adults on HAART who enrolled at four facilities in central Kenya in January 2013 to December 2018 were considered for analysis. Those who had at-least one follow-up visit were eligible to be included. Children below 15 years and confirmed pregnant mothers were excluded. We also excluded PLHIV who had unknown ART start date, unknown outcome, and transferred in with incomplete base-line information. Data was pulled from point of care electronic medical database, consolidated in MS excel, and exported to R Studio for further analysis. The main event variable in the analysis was time to LTFU (in months) with other exits

treated as a competing event. LTFU was defined as patients not taking ART refill for a period of 38 days or missed clinical appointment over the same duration.

#### 4. DISCUSSION

##### 4.1. Recurrent Events Model

In this study, we extended PWP-TT model to cover informative censoring when making inferences on recurrent events under HIV retention setting. The proposed methodology intrinsically extends the existing method in literature on recurrent events under typical HIV resource limited setting. A well-defined truncation time  $T$ , with  $Z_i$  overall i.e.  $Z_i M_{ij}(t) \lambda_{oj}(t - t_j - 1) \exp [X_i(t)' \beta]$  has been assumed. Here, we have concentrated on the LTFU as recurrent events and fitted real data. It can also be developed further to other regression/Cox methods. The concept can be extended to other approaches by modeling  $\lambda_{oj}[t]$  and  $\lambda_{oj}[t - t_j - 1]$  with proportional hazards models. The main purpose of this work was to provide an overview, applicable statistical techniques when analyzing recurrent event data under informative censoring in HIV retention setting. The typical real data used in this work is from a routine well established HIV care clinic. The longitudinal approach to analyze recurrent-event data applied here can also be applicable to other observational cohort studies. Because the technical approach employed here are extensions of Cox proportional hazards regression, explicit issues that affect model modification can also be handled in the same manner as the classic applied techniques.

##### 4.2. Results

Majority of the patients were female 37 (64%). The average age was 36.4 years (SD=6.28). There were 256 incidences of patients lost to follow up, albeit there were no noted incidences among seven patients. Table

2 shows a breakdown of incidence of LTFU as well as those resulting from the informative terminal event, stigmatization. These are disaggregated by gender and age [binary covariate set at population mean]. Overall, 6 in 10 patients [61%] experienced 1 to 6 LTFU incidences. The median time to LTFU was 8.6 months [range: 3.0-82.9 months].

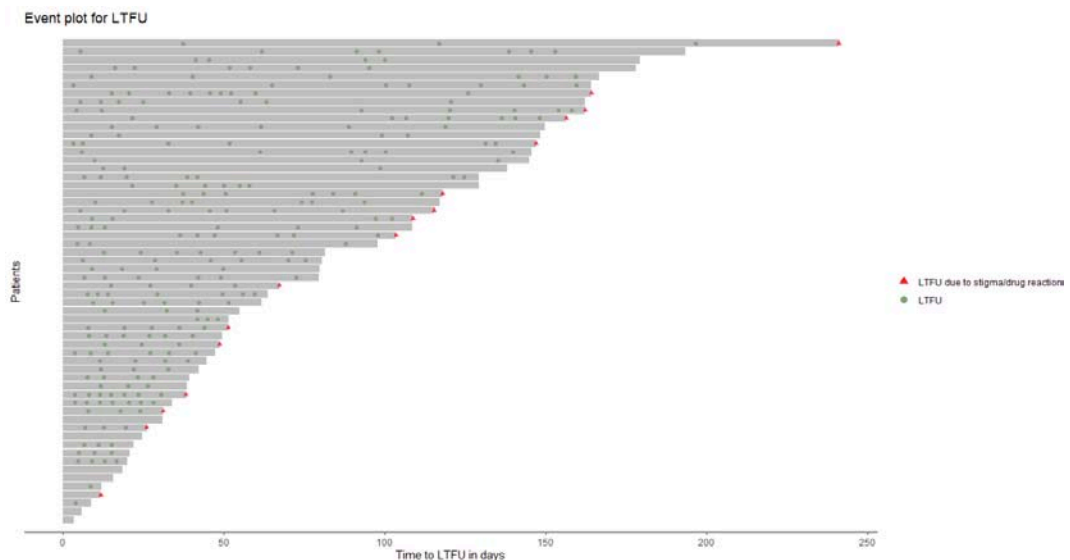
Patients were either censored at the end of the study or informative drop-out due to stigma or drug reaction. Sixteen instances of informative censoring were reported. Comparison of crude proportions of LTFU incidence due to stigma revealed a higher incidence among patients above 37 years. However, there was no significant difference by gender. The timing of the LTFU incidences and corresponding informative censoring arising from either stigma or drug reaction are shown in Figure 2.

Additionally, cumulative sample mean plots are provided to provide the cumulative number of LTFU occurrences during the study period. The number of incidences is shown to increase consistently across the monitoring window.

To assess the performance of the proposed PWP-TT model under joint modelling of both the failure and event time, we apply it to the patient level data described in section 3.1. For this analysis, sex [1 for male, 2 for female], age [1 for  $\leq 37$ , 2 for  $>37$  years], Regimen [1 for first line, 2 for second line], and differentiated care [1 for on differentiated care, 0 for not on differentiated care] were used. The event was LTFU (1 for LTFU occurrence, 0 for no occurrence), while a composite censoring variable was defined (1 if due to stigma, 0 for other exits). Specifically, we investigate the effect of these covariates on rate of LTFU and risk of LTFU arising from stigma. Parameter estimates, standard errors (SE), and corresponding p-values are shown in Table 3. Standard errors were estimated by resampling 100 times from patient data. Additionally,

**Table 2: A distribution of the Total Number of LTFU Incidences Occurring among 58 Patients**

Subgroup	Total Patients	LTFU due to stigma	Number of LTFU incidences since ART start									
			0	1	2	3	4	5	6	7	8	9
<b>Male</b>	21	8	2	1	0	5	3	0	7	3	-	0
<b>%</b>	36%	50.0%	10%	5%	0%	24%	14%	0%	33%	14%		0%
<b>Female</b>	37	8	5	1	0	8	5	1	6	10		1
<b>%</b>	64%	50.0%	14%	3%	0%	22%	14%	3%	16%	27%		3%
<b>Age <math>\leq 37</math></b>	30	6	3	2	0	6	3	0	5	10		1
<b>%</b>	52%	37.5%	10%	7%	0%	20%	10%	0%	17%	33%		3%
<b>Age <math>&gt; 37</math></b>	28	10	4	0	0	7	5	1	8	3		0
<b>%</b>	48%	62.5%	14%	0%	0%	25%	18%	4%	29%	11%		0%



**Figure 2:** Recurrent history of LTFU per patient. Red triangles represent informative censoring due to stigma/drug reaction. The green circles are incidences of LTFU.

**Table 2: Estimates for Risk Factors with Standard Errors, and p-Values**

Risk factor	Estimate	SE	z value	p-value	
<b>LTFU Recurrence</b>					
Gender [base=Male]	0.043	0.131	0.323	0.746	
Age [base is < 37]	-0.014	0.087	-0.166	0.868	
Regimen [base= First line]	0.067	0.162	0.412	0.681	
Differentiated care [Base=not on Differentiated care]	-0.356	0.164	-2.166	0.03	*
<b>LTFU due to stigma</b>					
Gender [base=Male]	-1.694	0.544	-3.113	0.002	**
Age [base is < 37]	2.034	0.772	2.634	0.008	**
Regimen [base= First line]	0.618	0.518	1.192	0.233	
Differentiated care [Base=not on Differentiated care]	-1.615	0.77	-2.097	0.036	*

Codes: 0 \*\*\*\* 0.001 \*\*\* 0.01 \*\* 0.05 \* 0.1 . 1.

results from the AG and Cox model are also computed (output not included) and compared.

From Table 3, it can be noted that patients on differentiated care had lower likelihoods for both LTFU recurrence (30% lower) as well as for the incidence of LTFU arising from stigma (80% less likely). These findings are unsurprising given that one well known benefit of differentiated care models is to fight stigma. Gender and age are also significant predictors of LTFU due to stigma. The instantaneous risk for LTFU from stigma is lower for females [82% less risk], but higher for persons over the age of 37 years.

For the Cox model, none of the covariates was significant for risk of LTFU from stigma. On the other hand, the AG model also indicated a 73% risk of LTFU from stigma for persons on a differentiated care program, as well as a 23% reduced recurrence of LTFU. The effect and direction of the other covariates

for the AG are also comparable with those for the PWP-TT. These results suggest Cox regression for recurrent events may not be suitable in observational studies.

As an alternative to cox regression models, one may consider using the Generalized Estimating Equations (GEE), which may be employed in instances of longitudinal and correlated data, especially if the responses are binary. One key advantage of this approach is its robustness in giving consistent results even with a mis-specification of the correlation structure. However one key limitation is that responses need to be correlated, which is not always the case with recurrent survival data.

On the other hand, the analysis of recurrent-events based only on the first event time is not ideal when examining the effect of risk factors. This is underlined by Ullah *et al.*, [16] who wrote several of articles on

recurrent injuries with a use of statistical methods, which only account for the first event thus excluding key information from subsequent injuries. Odhiambo *et al.*, [17] applies a Block-Borges-Savits (BBS) minimal repair model to HIV retention data to recurrent event data, with subsequent use of smooth tests to assess fit.

Several approaches have been proposed in literature to account for censoring that rises in survival analysis setting.

The major assumption in AG model is that the inter-event time increments are conditionally uncorrelated with given covariates. It also assumes the same baseline hazard for all persons, which may not be the case for LTFU in an HIV setting, given that intensive adherence counseling may alter the subsequent likelihood of an individual getting LTFU. On the other hand, it is best suited to cases of independent increments across observation units, and is also the easiest extension to the Cox model to replicate. On the other hand, either of the PWP models (GT or TT) adjusts for varying baseline hazards across observation units which may be efficient in the case of LTFU, where due to adherence counseling, the baseline hazard may change.

Generally, the choice of recurrent event data analysis technique is determined by several factors, i.e. events; relationship between events; varying effects across recurrences; the medical/biological process; and independence/dependence structure. Usually the stratified models, as PWP [total or gap times] or multi-state models, are useful whenever there are relatively few recurrent-events per individual and the risk of recurrences.

A recurrent events model will ideally help to provide insights into the program/disease structure and process. Hence, it is critical to consider the censoring mechanism and perform analysis that enhances comprehension of the risk factors.

### LTFU

In a typical HIV care clinic, the risk posed by instances of LTFU is undesired and has the potential of undoing antiretroviral treatment benefits. Specifically, patients' retention in HIV care is critical to ensuring better health outcomes especially in reduced viral load suppression, mitigating mortality and averting possible drug resistance caused by non-adherence.

In spite of the multiple interventions in place to address incidences of LTFU among HIV patients like enhanced adherence counselling, within-person cases of LTFU are usually common and recurrent in nature,

with the present likelihood of a person getting lost to follow-up influenced by previous occurrences.

### ABBREVIATION

AD	= Anderson and Gill
AIDS	= Acquired Immunodeficiency Syndrome
ART	= Antiretroviral Therapy
BBS	= Block-Borges-Savits
GEE	= Generalized Estimating Equations
HAART	= Highly Active Antiretroviral Therapy
CCC	= Comprehensive Care Center
DCM	= Differentiated Care Model
HIV	= Human Immunodeficiency Virus
HR	= Hazard Ratio
LTFU	= Lost to follow-up
LWA	= Lee, Wei and Amato
PWP-GT	= Prentice, Williams, and Peterson gap time
PLHIV	= Persons Living with HIV
PWP-TT	= Prentice, Williams, and Peterson total time
SD	= Standard Deviation
SE	= Standard Error
WLW	= Wei, Lin and Weissfeld and
INH	= Isonicotinylhydrazide
WHO	= World Health Organization

### REFERENCES

- [1] Huang C-Y, Wang M-C. Joint Modeling and Estimation for Recurrent Event Processes and Failure Time Data. *J Am Stat Assoc* 2004; 99(468): 1153-65. <https://doi.org/10.1198/01621450400001033>
- [2] Chiou SH [Steven], Huang C-Y. reReg: Recurrent Event Regression [Internet]. 2018 [cited 2019 Jul 23]. Available from: <https://CRAN.R-project.org/package=reReg>
- [3] Lawless JF. Regression Methods for Poisson Process Data. *J Am Stat Assoc* 1987; 82(399): 808-15. <https://doi.org/10.1080/01621459.1987.10478502>
- [4] Kelly PJ, Lim LL-Y. Survival analysis for recurrent event data: an application to childhood infectious diseases. *Stat Med* 2000; 19(1): 13-33. [https://doi.org/10.1002/\(SICI\)1097-0258\(20000115\)19:1<13::AID-SIM279>3.0.CO;2-5](https://doi.org/10.1002/(SICI)1097-0258(20000115)19:1<13::AID-SIM279>3.0.CO;2-5)
- [5] Yang W, Jepson C, Xie D, Roy JA, Shou H, Hsu JY, *et al.* Statistical Methods for Recurrent Event Analysis in Cohort Studies of CKD. *Clin J Am Soc Nephrol* 2017; 12(12): 2066-73. <https://doi.org/10.2215/CJN.12841216>

- [6] Andersen PK, Gill RD. Cox's Regression Model for Counting Processes: A Large Sample Study. *Ann Stat* 1982; 10(4): 1100-20.  
<https://doi.org/10.1214/aos/1176345976>
- [7] Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. *Biometrika* 1981; 68(2): 373-9.  
<https://doi.org/10.1093/biomet/68.2.373>
- [8] Wei LJ, Lin DY, Weissfeld L. Regression Analysis of Multivariate Incomplete Failure Time Data by Modeling Marginal Distributions. *J Am Stat Assoc* 1989; 84(408): 1065-73.  
<https://doi.org/10.1080/01621459.1989.10478873>
- [9] Ghosh D, Lin DY. Semiparametric Analysis of Recurrent Events Data in the Presence of Dependent Censoring. *Biometrics* 2003; 59(4): 877-85.  
<https://doi.org/10.1111/j.0006-341X.2003.00102.x>
- [10] Christel Castelli, Philippe Saint-Pierre, Jean-Pierre Daures. Informative censoring in survival analysis and application to asthma | Christel Castelli | Request PDF [Internet]. ResearchGate. 2006 [cited 2019 Jul 22]. Available from: [https://www.researchgate.net/publication/262413091\\_Informative\\_censoring\\_in\\_survival\\_analysis\\_and\\_application\\_to\\_asthma/citations](https://www.researchgate.net/publication/262413091_Informative_censoring_in_survival_analysis_and_application_to_asthma/citations)
- [11] Lin DY, Wei LJ, Ying Z. Accelerated Failure Time Models for Counting Processes. *Biometrika* 1998; 85(3): 605-18.  
<https://doi.org/10.1093/biomet/85.3.605>
- [12] Berheto TM, Haile DB, Mohammed S. Predictors of Loss to follow-up in Patients Living with HIV/AIDS after Initiation of Antiretroviral Therapy. *North Am J Med Sci* 2014; 6(9): 453-9.  
<https://doi.org/10.4103/1947-2714.141636>
- [13] Assemie MA, Muchie KF, Ayele TA. Incidence and predictors of loss to follow up among HIV-infected adults at Pawi General Hospital, northwest Ethiopia: competing risk regression model. *BMC Res Notes* [Internet]. 2018 May 10 [cited 2019 Jul 4]; 11.  
<https://doi.org/10.1186/s13104-018-3407-5>
- [14] Amorim LD, Cai J. Modelling recurrent events: a tutorial for analysis in epidemiology. *Int J Epidemiol* 2015; 44(1): 324-33.  
<https://doi.org/10.1093/ije/dyu222>
- [15] Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. Springer Science & Business Media 2000; p. 372.  
<https://doi.org/10.1007/978-1-4757-3294-8>
- [16] Ullah S, Gabbett TJ, Finch CF. Statistical modelling for recurrent events: an application to sports injuries. *Br J Sports Med* 2014; 48(17): 1287-93.  
<https://doi.org/10.1136/bjsports-2011-090803>
- [17] Odhiambo C, Odhiambo J, Omolo B. A Smooth Test of Goodness-of-Fit for the Baseline Hazard Function for Time-to-First Occurrence in Recurrent Events: An Application to HIV Retention Data. *Int J Stat Med Res* 2017; 6(3): 104-113.  
<https://doi.org/10.6000/1929-6029.2017.06.03.2>

Received on 26-02-2020

Accepted on 19-03-2020

Published on 29-03-2020

<https://doi.org/10.6000/1929-6029.2020.09.03>© 2020 Ejoku *et al.*; Licensee Lifescience Global.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.