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

Collins Odhiambo[1], John Odhiambo[1] and Bernard Omolo[2,*]

[1]*Strathmore Institute of Mathematical Sciences, Strathmore University, Ole Sangale Road, Nairobi, Kenya*

[2]*Division of Mathematics and Computer Science, University of South Carolina-Upstate, 800 University Way, Spartanburg, South Carolina, USA*

Abstract: Motivated by HIV retention, we present an application of the smooth test of goodness-of-fit under right-censoring to time to first occurrence of a recurrent event. The smooth test applied here is an extension of Neyman's smooth test to a class of hazard functions for the initial distribution of a recurrent failure-time event. We estimate the baseline hazard function of time-to-first loss to follow-up, using a Block, Borges and Savits (BBS) minimal repair model of the data (n = 2,987,72% censored). Simulations were conducted at various percentages of censoring to assess the performance of the smooth test. Results show that the smooth test performed well under right-censoring.

Keywords: BBS model, Hazard function, Loss to follow-up, Neyman's smooth test, Recurrent events, Retention in HIV care.

1. BACKGROUND

Recurrent events refer to failure-time events where individuals or entities experience repeated occurrences of the same type of event of interest. The events are frequently encountered in many biomedical settings, where clinically important events occur repeatedly over the course of follow-up. See [1-6] for specific and more comprehensive examples of recurrent events. In many applications, the gap-time, is the main outcome of interest. Several survival models have been proposed to handle gap-time in recurrent event (e.g. the Wei-Lin-Weissfeld total time (WLW-TT) marginal model, the Prentice-Williams-Petersen gap-time (PWP-GT) conditional model, the Block-Borges-Savits (BBS) minimal repair model, the recurrence rate models, the semiconditional models, etc.) [6-12].

In this paper, we focus on the application of the BBS model [13] to the time to first loss to follow-up for patients in HIV treatment. The BBS model is more appropriate in our scenario, due its flexibility and can be adopted in many areas of applications [10, 14]. In the BBS model, a system is put on test at time 0. When the system fails at time, say *t*, it undergoes a perfect repair with probability $\rho(t)$ or an imperfect repair with probability $q(t) = 1 - \rho(t)$. A perfect repair reverts the system's age to 0, whereas a minimal repair leaves the system's age the same as at the age at failure. After repair, the testing process is continued, and at each subsequent failure, either a perfect repair is performed with probability $\rho(t)$ or an imperfect repair is performed with probability $q(t)$, where t is the age of the system. The assumption here is that repairs take negligible time. Although the BBS model is primarily utilized in the reliability and operations research settings, it is also applicable to other areas since it admits as special cases some of the models commonly encountered in practice [11] (e.g. by taking $\rho(t) = 0$ we obtain a nonhomogeneous Poisson process as a special case of the model that is commonly used in biomedical settings).

This work is motivated by the problem of retention of patients in HIV care. During "retention", patients are known to be alive and receiving highly active antiretroviral therapy (HAART) by the end of a follow-up period. An "exit" from an ART program is defined as a discontinuation of ART for any reason, including death, loss to follow-up (LTFU), stopping ARV medications while remaining in care and transfer to another ART facility. ART treatment discontinuation and poor adherence can lead to drug resistance and sub-optimal benefits in an HIV treatment programs. Patients with clinical AIDS who discontinue ART will likely die within a relatively short time [15]. High rates of attrition from treatment programs thus pose a serious challenge to program implementers and constitute an inefficient use of scarce treatment resources in sub-African countries.

A patient who experiences LTFU default's treatment. S/he can be traced back to ART program and either initiated ART afresh (perfect repair) or continued on the same regimen (imperfect repair).

*Address correspondence to this author at the Division of Mathematics & Computer Science, University of South Carolina-Upstate, 800 University Way, Spartanburg, South Carolina, USA; Tel: +1 864-503-5362; Fax: +1 864-503-5930; E-mail: bomolo@uscupstate.edu

E-ISSN: 1929-6029/17

Long-term retention of patients in ART treatment programs is critical to achieving long term benefits of ART. Most large-scale ART treatment providers, particularly in sub-Saharan Africa, face the challenge of LTFU. Attrition from ART treatment programs is generally divided into four categories: death, active on ART, transfer-out and loss to follow-up (LTFU)) [15].

Here, we fit the BBS model and test the baseline hazard of the model using a smooth tests of goodness-of-fit. The smooth test of goodness-of-fit problem considered here is an extension of Neyman's smooth test [16] to recurrent events models. Assuming that the LTFU data accrual follows a BBS model, the test involves testing the null hypothesis that the initial hazard function of the lifetime is $H_0 : \lambda(\cdot) \in \{\lambda_0(\cdot;\varepsilon) : \varepsilon \in \Xi\}$ versus the alternative hypothesis $H_1 : \lambda(\cdot) \notin \{\lambda_0(\cdot;\varepsilon) : \varepsilon \in \Xi\}$ where the functional form of $\lambda_0(\cdot;\varepsilon)$ is known, except for the $p \times 1$ vector ε. The parameter ε is a nuisance parameter in this testing problem and Ξ is a $p \times r$ vector space [11]. The smooth tests are score tests and have been shown by [17, 18] to be powerful against a wide range of alternatives. Furthermore, smooth tests of goodness-of-fit are considered to be a compromise between omnibus tests and directional tests. This work presents the first attempt to fit HIV retention data to a recurrent event model and then assessing the fit using a smooth test of goodness-of-fit. The motivating premise is that of a practical clinical setup where the primary purpose of the inference is risk of LTFU, tracking defaulters in ART treatment, and prediction of the risk of LTFU. This is common in clinical setting where patients in an ART are expected to remain in a treatment program with the goal of suppressing viral load. Further, in such situations, particularly in sub-Saharan Africa, the challenge has been how to model the hazard rate function more efficiently.

The formulation of the smooth test presented here is through hazard functions, which permits the derivation to obtain omnibus as well as directional tests [18]. Since the smooth tests are score tests, they are endowed with asymptotic optimality properties. A major contribution of this paper is the application of the BBS model to a primary HIV retention dataset and the assessment of the model using the smooth test of goodness-of-fit. In modelling the LTFU data, we utilize stochastic formulation, which allows a direct estimation of nuisance parameters [19]. By fitting LTFU data to a parametric family of baseline hazard functions, we are able to make more efficient inferences about the risk.

In section 2, we revisit the general framework of BBS model and development of a smooth test of goodness-of-fit for the initial baseline function. Monte Carlo simulation results pertaining to the finite sample size properties of the tests are presented in section 3. In section 4, we present model fitting and analysis of results including motivation for fitting LTFU data. We discuss overall results in section 5. Finally, we provide the concluding remarks and limitations of the study in section 6.

2. GENERAL FRAMEWORK

Let $\{\omega_0 = 0, \omega_1, \omega_2, \cdots\}$ be a sequence of failure times generated under a minimal repair model, with ω_i being continuously distributed with probability density function f and hazard function λ. Let $\omega_0 = 0 < \omega_1 < \omega_2 < \cdots$ be successive failure times of a component, and $U_1, U_2, \cdots$ is a sequence of i.i.d Uniform $[0,1]$ random variables which are independent of the failure times. The sequence $(\omega_0 = 0, \omega_1, \omega_2, \cdots, \omega_\nu)$, where $\nu = \inf\{k \in \{1,2,\cdots\} : U_k < p(\omega_k)\}$, is an epoch of the BBS model. Since a perfect repair restores a component to as good-as-new state, it suffices to observe a component only until the time of its first repair, a situation that is naturally similar to time to first loss to follow-up (LTFU) in a typical HIV clinical setting. See [1, 4, 10, 12-14, 20-24] for other applications of BBS models.

2.1. Smooth Test of Goodness-of-Fit

Our interest here is to test the null hypothesis $H_0 : \lambda(\cdot) = \lambda_0(\cdot)$, where $\lambda_0(\cdot)$ is a completely specified hazard function. The smooth test of goodness-of-fit is derived by nesting the hypothesised hazard function $\lambda_0(\cdot)$ to a larger parametric family of hazard functions to get

$$\beth_k = \{\lambda_k(\cdot;\boldsymbol{\theta}) = \lambda_0(\cdot)\exp\{\boldsymbol{\theta}^{\mathrm{T}}\Psi(\cdot)\} : \boldsymbol{\theta} \in R^k\}, \tag{1}$$

where k is some fixed positive integer and $\Psi(\cdot)$ is a $k \times 1$ vector of locally bounded predictor process [12, 18-20, 23]. Note that, the null hypothesis $H_0 : \lambda(\cdot) = \lambda_0(\cdot)$ can be rewritten as $H_0^* : \boldsymbol{\theta} = \mathbf{0}$.

The score processes associated with $\boldsymbol{\theta}$ have been derived by [12, 18-20, 23] to yield

$$\mathbf{U}_{\boldsymbol{\theta}}(t;\boldsymbol{\theta}) = \sum_{j=1}^{n} \int_0^t \Psi(s) dM_j(s;\boldsymbol{\theta}), \tag{2}$$

where $M_j(s;\boldsymbol{\theta}) = N_j(s) - A_j(s;\boldsymbol{\theta}), j = 1,2,\cdots,n$, $N_j(s)$ is the counting process defined on a filtered probability space $(\Omega, F = (\Im_t : t \geq 0), P)$ and $A_j(s;\boldsymbol{\theta})$ is the

$F-$compensator of $N_j(s)$. The asymptotic α-level smooth test of goodness-of-fit for $H_0^*:\theta=0$ (or equivalently $H_0:\lambda(\cdot)=\lambda_0(\cdot)$) can be generated by the test statistic (see [19])

$$S(t)=\frac{1}{n}\mathbf{U}_{\boldsymbol{\theta}}(t;\boldsymbol{\theta})^T\Sigma^{-1}(\cdot)\mathbf{U}_{\boldsymbol{\theta}}(t;\boldsymbol{\theta}), \quad (3)$$

where $\Sigma^{-1}(\cdot)$ is the generalized inverse of $\Sigma(\cdot)$. We reject H_0 whenever $S(t)\geq\chi^2_{k^*;\alpha}$, where $\chi^2_{k^*;\alpha}$ is the $(1-\alpha)100^{th}$ percentile of the chi-square distribution with degree of freedom $k^*=\text{rank}[\Sigma(t)]$. For a comprehensive coverage of the choice of the processes $\boldsymbol{\Psi}(\cdot)$, polynomial specification (k), smoothing process of $\boldsymbol{\Psi}(\cdot)$, achieved power of $S(t)$ and asymptotic distribution of $S(t)$, see [3, 12, 17-20, 23].

3. SIMULATION

We conducted Monte Carlo simulations to investigate the performance of the smooth tests. Similar tests have been examined in extensive simulation studies [18-20]. The goal of simulations was to compare the empirical significance levels of the tests with the specified nominal asymptotic levels as the sample size and the degree of censoring are varied. These comparisons indicate which tests qualify as good omnibus tests and which tests have good control of Type I error among a wide range of alternatives. Simulations were also helpful in determining appropriate values of the smoothing parameter k (i.e. $k=\{1,2,3,4\}$). Since we are fitting the BBS model with the probability of perfect repair $\rho(\cdot)=1$, we considered simulations for $n=\{20,50,100,200,500,1000\}$.

In the simulations, the initial failure-time variables were generated according to the exponential distribution with mean $\theta=8$, the Weibull distribution with shape parameter γ and scale parameter β and the Gamma ($\varsigma=3,\alpha=4$) distribution. The failure-time variables $(T_1,T_2,\cdots T_n)$ were generated using the chosen alternative, and the censoring variables $(C_1,C_2,\cdots,C_m)$ were distributed according to the exponential distribution with mean 1. By utilizing the resulting randomly censored data $\{(Z,i),i=1,\cdots,n\}$, the null hypothesis was tested according to the different 5% asymptotic level tests. The percentage out of the replicates that a test rejected H_0 was then

Table 1: Empirical Control of the Type I Error Rate Under $H_0:\lambda(\cdot)$ is Distributed as Exponential ($\theta=8$) at $\alpha=0.05$. The Failure Times under the Null Hypothesis were Generated According to a BBS Model

	Empirical Type I error rate (70% Censoring)					
Sample size, n	**n=20**	**n=50**	**n=100**	**n=200**	**n=500**	**n=1,000**
ψ_1	0.0452	0.0406	0.0422	0.0394	0.0323	0.0401
ψ_2	0.0418	0.0455	0.0497	0.0453	0.0498	0.0428
ψ_3	0.0560	0.0485	0.0443	0.0490	0.0545	0.0563
ψ_4	0.0568	0.0491	0.0553	0.0598	0.0654	0.0570
Censoring %	**Empirical Type I error rate (50% Censoring)**					
Sample size, n	**n=20**	**n=50**	**n=100**	**n=200**	**n=500**	**n=1,000**
ψ_1	0.0366	0.041	0.0333	0.0394	0.0303	0.0371
ψ_2	0.0422	0.0459	0.0373	0.0453	0.0363	0.034
ψ_3	0.0457	0.049	0.0398	0.05	0.0401	0.0482
ψ_4	0.0564	0.0496	0.0403	0.0528	0.0509	0.0491
Censoring %	**Empirical Type I error rate (20% Censoring)**					
Sample size, n	**n=20**	**n=50**	**n=100**	**n=200**	**n=500**	**n=1,000**
ψ_1	0.0311	0.0322	0.0334	0.0304	0.0423	0.0309
ψ_2	0.0384	0.0385	0.0398	0.0376	0.0486	0.0394
ψ_3	0.043	0.0424	0.0438	0.0422	0.0525	0.0447
ψ_4	0.044	0.0432	0.0446	0.0431	0.0533	0.0558

Table 2: Empirical Control of the Type I Error Rate under $H_0 : \lambda(\cdot)$ is Distributed as Weibull ($\beta = 6$, $\gamma = 10$) at $\alpha = 0.05$. The Failure Times under the Null Hypothesis were Generated According to a BBS Model

	Empirical Type I error rate (70% Censoring)					
Sample size, n	**n=20**	**n=50**	**n=100**	**n=200**	**n=500**	**n=1,000**
ψ_1	0.0352	0.0406	0.0422	0.0494	0.0423	0.0371
ψ_2	0.0431	0.0411	0.0401	0.0411	0.0411	0.0431
ψ_3	0.0444	0.0423	0.0677	0.069	0.0663	0.0471
ψ_4	0.0798	0.0919	0.081	0.0788	0.0703	0.0998
	Empirical Type I error rate (50% Censoring)					
Sample size, n	**n=20**	**n=50**	**n=100**	**n=200**	**n=500**	**n=1,000**
ψ_1	0.0466	0.041	0.0333	0.0494	0.0403	0.0301
ψ_2	0.0471	0.0466	0.0411	0.0494	0.0455	0.0411
ψ_3	0.0516	0.0511	0.0594	0.0503	0.0571	0.0479
ψ_4	0.0533	0.0666	0.0811	0.0777	0.0603	0.0578
	Empirical Type I error rate (20% Censoring)					
Sample size, n	**n=20**	**n=50**	**n=100**	**n=200**	**n=500**	**n=1,000**
ψ_1	0.0311	0.0322	0.0334	0.0304	0.0323	0.0309
ψ_2	0.0471	0.0466	0.0371	0.0394	0.0403	0.0371
ψ_3	0.0499	0.0511	0.0454	0.0422	0.0471	0.0455
ψ_4	0.0521	0.0596	0.0511	0.0484	0.0509	0.0588

Table 3: Empirical Control of the Type I Error Rate under $H_0 : \lambda(\cdot)$ is Distributed as Gamma($\varsigma = 3, \alpha = 4$) at $\alpha = 0.05$. The Failure Times under the Null Hypothesis were Generated According to a BBS Model

	Empirical Type I error rate (70% Censoring)					
Sample size, n	**n=20**	**n=50**	**n=100**	**n=200**	**n=500**	**n=1,000**
ψ_1	0.0444	0.0354	0.0423	0.0394	0.0323	0.0371
ψ_2	0.0497	0.0396	0.0474	0.0453	0.0398	0.0428
ψ_3	0.0531	0.0423	0.0506	0.049	0.0445	0.0563
ψ_4	0.0537	0.0529	0.0512	0.0598	0.0854	0.077
	Empirical Type I error rate (50% Censoring)					
Sample size, n	**n=20**	**n=50**	**n=100**	**n=200**	**n=500**	**n=1,000**
ψ_1	0.0366	0.031	0.0333	0.0394	0.0403	0.0371
ψ_2	0.0422	0.0459	0.0373	0.0453	0.0463	0.044
ψ_3	0.0557	0.049	0.0498	0.059	0.0601	0.0482
ψ_4	0.0964	0.0696	0.0503	0.0598	0.0909	0.0991
	Empirical Type I error rate (20% Censoring)					
Sample size, n	**n=20**	**n=50**	**n=100**	**n=200**	**n=500**	**n=1,000**
ψ_1	0.0331	0.0322	0.0355	0.0304	0.0323	0.0309
ψ_2	0.044	0.0385	0.0422	0.0476	0.041	0.0394
ψ_3	0.0599	0.0624	0.0663	0.0722	0.0564	0.0447
ψ_4	0.0691	0.1122	0.1072	0.0931	0.0875	0.1058

Table 4: Empirical Control of the Type I Error Rate under $H_0:\lambda(\cdot)$ is Distributed as Weibull ($\beta=5,\gamma=15$) at $\alpha=0.05$. The Failure Times under the Null Hypothesis were Generated According to a BBS Model

	Empirical Type I error rate (70% Censoring)					
Sample size, n	**n=20**	**n=50**	**n=100**	**n=200**	**n=500**	**n=1,000**
ψ_1	0.0415	0.0331	0.0395	0.0361	0.0382	0.044
ψ_2	0.0464	0.037	0.0442	0.0417	0.0452	0.0493
ψ_3	0.0496	0.0395	0.0472	0.0551	0.0695	0.0526
ψ_4	0.0502	0.04	0.0478	0.0559	0.0704	0.0533
	Empirical Type I error rate (50% Censoring)					
Sample size, n	**n=20**	**n=50**	**n=100**	**n=200**	**n=500**	**n=1,000**
ψ_1	0.0435	0.0383	0.0311	0.0361	0.037	0.0333
ψ_2	0.0487	0.0429	0.0348	0.0417	0.0426	0.0397
ψ_3	0.052	0.0458	0.0472	0.0551	0.0561	0.0437
ψ_4	0.0527	0.0464	0.0677	0.0759	0.1069	0.0946
	Empirical Type I error rate (20% Censoring)					
Sample size, n	**n=20**	**n=50**	**n=100**	**n=200**	**n=500**	**n=1,000**
ψ_1	0.0333	0.0388	0.0318	0.0364	0.0375	0.0362
ψ_2	0.0397	0.0446	0.0381	0.0432	0.0456	0.0442
ψ_3	0.0437	0.0583	0.0619	0.0574	0.0607	0.0791
ψ_4	0.0846	0.089	0.1028	0.0683	0.0817	0.0802

calculated. The bootstrapping procedure was applied 1,000 times to each generated dataset to obtain the significance level of the test. Within the context of model selection, size estimates were based on the proportion of replications that indicate acceptable fit, with a larger number of replications resulting in smaller CIs (higher power, more accuracy) around the estimates. The data-generating process was performed using the *SimSurv* function of the *prodlim* package from *R*.

To investigate the performance of the different tests, we evaluated the exponential, Weibull and Gamma null hypothesis against their generalised alternatives for the initial distribution of failure ages. Hence, given the values of different parameters, values for each alternative were generated. Simulations were done for the 5% asymptotic level tests. We also performed simulations for values of $\beta=6$, $\gamma=10$ and $\beta=5,\gamma=15$ for the Weibull initial distribution and found the results to be consistent with those presented by [18-20, 25]. Examining the performance of the directional tests, we again noticed that we are able to achieve required significance levels. Tables **1**, **2**, **3** and **4** summarize the percentage rejection of the tests for the exponential alternatives, Gamma alternatives and Weibull alternatives. For the exponential alternative, the ψ_1,ψ_2 and ψ_3, have the highest power under H_0. The directional tests based on $\psi_i, i=1,2,3,4$, are sensitive for the Gamma-type distributions. The fact that the ψ_1 test is powerful for this alternative is expected because this test was derived against such alternatives, whereas the observation that the ψ_4 test is not powerful for this alternative is also expected because the normalized total-time-on-test statistic is invariant to changes in scale [19]. Against the Weibull-type, Gamma-type and exponential alternatives, the ψ_1 test performed best, followed by the ψ_2 test.

4. FITTING HIV RETENTION DATA

4.1. Motivation for Analysis of HIV Retention Data

HIV/AIDS has consistently been a major challenge in Kenya. The national prevalence is currently estimated to be 6% and there are at least 1.6 million Kenyans living with HIV (PLHIV) with at least 800,000 of PLHIV on ART [26]. In practice, the quality of the ART services is measured against the rate of retention of PLHIV on ART. With the advent of the United Nation

AIDS (UNAIDS) programme on HIV/AIDS targets in 2013, the focus have turned to interventions that quicken elimination of HIV/AIDS at the global, regional, country, province, district and city levels [27]. The strategy popularly known as $90-90-90$ targets that by 2020, 90% of people living with HIV know their HIV status, 90% of people who know their status are receiving ART treatment and 90% of people on HIV treatment have a suppressed viral load so that their immune system remains stronger and the likelihood of their infection being passed to others is greatly reduced. This strategy is currently being implemented in Kenya and this paper focuses on statistical innovation that hinges on one of the pillars: the third 90% i.e. viral suppression. Viral suppression is achievable by retaining patients on ART for long. There are potential benefits whenever a PLHIV's viral load is reduced to an undetectable level (i.e. people with undetectable viral loads are generally healthier than those people with higher levels of virus in their blood and are also less likely to transmit HIV to their sexual partners [28]). High retention rate in ART treatment plays a crucial role in maintaining viral load suppression [28-33]. See [34] for excellent summary on the relevance of LTFU and the rationale for fitting LTFU data to a parametric distribution.

4.2. Data Description

Data comprised all patients who were initiated ART at two government hospitals in Kenya. Patients under observation were enrolled between 1st of October 2011 and 31st December 2014. The event of interest was time to first LTFU. Data was collected routinely whenever patients came for clinical check-up or drug refill. The time between ART initiation to first LTFU was given in months. Time to first LTFU was defined as missing routine clinical appointment within 48 hours from the scheduled appointment date. Out of those initiated on ART, 854 patients experienced LFTU while the rest were right-censored. See [34] for details about the data.

4.3. Application to HIV Retention

4.3.1. Modelling LTFU

Let $\omega_1,\omega_2,\cdots,\omega_n$ be the gap-time between started on ART to LTFU for PLHIV. The date of start of ART for each patient is independent and the start time of observation is set at 0. The gap time is the period between start of ART until time *t*, when the patient experiences LTFU (failure) during the observation period. Patients who experience other exits (i.e. transfer-out or death) are censored. Patients who remain active on ART at the end of the observation period are also censored. A patient who experiences LTFU (failure) can still be recovered back (repaired) to ART treatment through a mechanism called defaulter tracing. Such patients are still exposed to the risk of LTFU even after recovery and so the event can recur several times. The setting is routine hospitals visits, hence the risk for the second, third and subsequent episodes of LTFU is the same as the first. Let $S_1,S_2,\cdots$ be the successive event times (LTFU) for the A process, and let $T_1,T_2,\cdots$ be the successive event times of the B process. To obtain a realization of a BBS (Λ,p) model, events in the A process correspond to the imperfectly repaired failures, whereas events in the B process are associated with the perfectly repaired failures. The equivalence is then seen by noting that the process $A+B$ is a nonhomogeneous Poisson process with intensity function $\int_0^{(\cdot)}\lambda(s)ds$, and, given that at time t an event has occurred, the probability that it is from the B process is $p(t)$.

4.3.2. Model Fitting

Here, we begin by checking the difference in residual behaviour and result to be detected in the hypothesis testing.

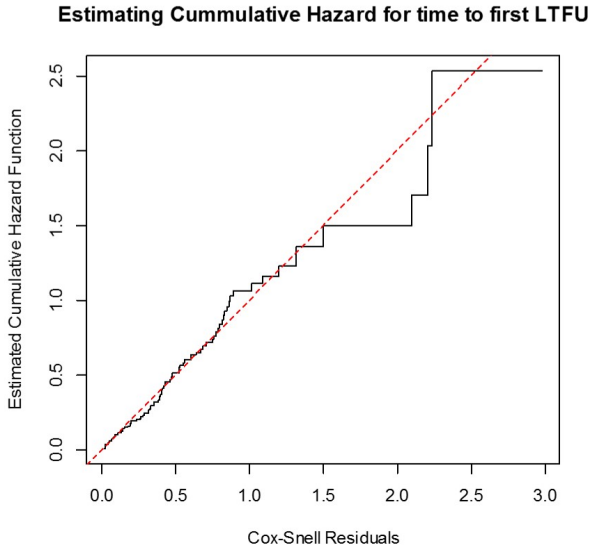

Figure 1: Comparing baseline hazard for time to first occurrence of LTFU.

The residuals graphically showed moderate signs of a different behaviour for violated models when the sample size is small and censoring is present. This indicates that for small samples with a higher degree of censoring the residuals could be sensitive for model violations. The advantage of using smooth test in the next subsection is that, the test is not affected by sample size, number of covariates and the level of censoring. We now check the fit for the distribution of the initial distribution.

Table 5: Assessing the Initial Distribution. Only the Weibull Distribution Fails to Reject the H_0 : Initial Distribution is Weibull, at $\alpha < 0.05$

Distribution	Scale	Chi-square value	*p*-value
Weibull	2.11	4.13	0.213
Exponential	1	8.29	0.016
Logistic	3.73	21.13	2.6e-05
Rayleigh	0.5	9.37	0.0093
Extreme Value	3.97	16.25	3e-04
Gaussian	7.45	29.83	3.3e-07
Student t	4.89	21.58	2.1e-05
Log-normal	0.287	30.39	6.3e-05

Table 6: Smooth Tests of Up to Order 4 for the BBS Model with Initial Distribution Weibull ($\beta = 6, \gamma = 30$) against an Initial Distribution of Generalized Weibull Family

Ψ	Test statistic	*p*-value
ψ_1	2.211	1.0
ψ_2	3.334	0.811
ψ_3	4.121	0.565
ψ_4	5.011	0.323

5. DISCUSSION

We extended the application of the BBS model to HIV retention data by setting loss to follow-up data to represent recurrent event scenario. This was motivated by LTFU data comprising of typical records of patients with repeated LTFU and repeated time-to-failure measurements of multiple patients. First, we reviewed the BBS model and generally discussed its features and applicability. The BBS model is used to estimate the baseline hazard function. Furthermore, applying the BBS model to fit LTFU is more straightforward, particularly when the analysis involves several recurrent observations. Smooth tests for assessing model fit for BBS model were presented. The application of the tests to assess overall fit of the BBS model have also been revisited. The BBS model is often used in reliability studies. In this paper, we demonstrate that the model can be extended to cover other scenarios particularly in public health. The BBS model is flexible and a typical HIV retention data can be fitted to the model. This finding emphasizes the point that BBS model and time-to-event analyses can be used to model LTFU. An analysis of the time to first event also shows the flexibility of the model. The application to the LTFU data is special in that we have attempted to show varied applications of the BBS Model. The procedures for estimating the parameters of a general and flexible class of the models for recurrent events have been revisited and its properties examined through simulation studies. Some data sets in the biomedical and reliability or engineering settings can be reanalyzed using BBS models.

The importance of HIV retention and adherence is also reflected in the 2011 General Assembly Political Declaration on HIV/AIDS, which emphasizes the need to address factors that limit treatment uptake and contribute to poor adherence and calls for the mobilization and capacity building of communities to support treatment scale-up and patient retention as well as programmes that support improved treatment adherence [35]. A focus on the client side was also underscored by UNAIDS in their 2011–2015 strategy, which stated that the demand side of treatment – the factors that make people enrol for treatment and adhere to it – has not received enough attention [36, 37]. One of the main challenges to the response to HIV treatment is insufficient adherence to treatment. Suboptimal viral suppression as a result of LTFU may yield a higher risk of developing drug resistance, as well as the transmission of HIV. We consider the problem of testing that the baseline hazard function $\lambda(\cdot)$ of the time to first LTFU equals some specified baseline hazard function $\lambda_0(\cdot)$. The goodness-of-fit procedures were derived as score tests obtained by nesting the null hypothesis in a larger family of hazard rate functions and has been studied by [18-20]. The resulting smooth test of goodness-of-fit procedures are also related to model validation procedures that utilize generalized residuals and, consequently, through the asymptotic results, the appropriate adjustments needed to properly use procedures based on generalized residuals can be obtained [18]. Several classes of goodness-of-fit tests, both omnibus and directional, can also be generated. Because the smooth tests are

viewed as score tests, they posses optimality properties. Through a simulation study, the acceptability of the asymptotic approximations were ascertained for the BBS model, and the powers under the null hypothesis of the different tests were obtained for a wide range of alternatives.

6. CONCLUSION

The smooth test of goodness-of-fit for the distribution of the initial failure-times of a BBS model was considered. An application of the intensity-based smooth goodness-of-fit tests developed in [12, 17-20] was employed. By varying the smoothing process $\Psi(\cdot)$, a generalization to the BBS model to the tests was obtained. The results of the Monte Carlo simulations shows the tests are powerful directional tests. Thus, the smooth test has the potential of being applicable in more complex dynamic models in survival analysis, reliability, and other settings, where the specification of the model is through hazard functions. Retention in ART appeared to be poor over time whereas loss to follow-up (LTFU) is common in resource-limited settings. Recognition of the importance of adherence to ART to the success of HIV care has gained widespread acceptance over the past decade. Today, acknowledgment of the importance of engagement in HIV care is necessary to maintain the health of HIV-infected populations.

Some of the limitations of this study include the issue of what smoothing degree to use, (k), and how to improve the asymptotic approximations for small to moderate sample sizes and in the presence of a high degree of censoring. Furthermore, there is a need to compare the smooth test, in terms of efficiency, with other procedures such as generalizations of the Kolmogorov-Smirnov and Cramer-von Mises types of tests in the settings considered in this paper.

LIST OF ABBREVIATIONS

AIDS = Acquired Immune Deficiency Syndrome

HIV = Human Immunodeficiency Virus

LTFU = Lost To Follow-Up

ART = Antiretroviral Therapy

BBS = Block, Borges and Savits

PLHIV = People Living with HIV

UNAIDS = United Nations AIDS

CPH = Cox Proportional Hazard

A-G = Andersen-Gill

WLW-TT = Wei-Lin-Weissfeld total time

PWP-GT = Prentice-Williams-Petersen gap time

CCC = Comprehensive Care Center

WHO = World Health Organization

IQR = Inter-quartile range

SD = Standard deviation

HAART = Highly Active Antiretroviral Therapy

GOF = Goodness-of-fit

ACKNOWLEDGEMENTS

This work was partially supported by a grant from the Simons Foundation (# 282714 to BO).

APPENDIX

The BBS Model

Generally, the BBS model has two parameters [12]: a lifetime distribution function F, which we have assumed to be continuous, and a function $p[0,\infty)\to[0,1]$. Thus, we write the model as BBS (F,p). Because there is a one-to-one correspondence between F and its hazard function Λ, given by

$$\Lambda(t)=\int_0^t \frac{dF(\omega)}{1-F(\omega)}=-\log(1-F(t)) \tag{4}$$

and

$$F(t)=1-\exp\{-\Lambda(t)\}, \tag{5}$$

we can therefore specify a BBS model and rewrite it as BBS (Λ,p) [12]. Under a model of minimal repair, the sequence $\{\omega_j\}_{j=1}^{\infty}$ is a Markov process, and the conditional survivor function of ω_j given $\omega_0,\omega_1,\cdots,\omega_{j-1}$ is

$$\bar{S}(t\mid\omega_{j-1})=\frac{\bar{S}(t)}{\bar{S}(\omega_{j-1})},t\geq\omega_{j-1},j\geq 1,$$

where $S=1-F$ is the survivor function. Let $U_1,U_2,\cdots$ be a sequence of identically distributed and independent standard uniform variables, which are independent of the ω_j's. Let $\nu=\min\{k\in\{1,2,\cdots\}:U_k\leq p(\omega_k)\}$. An epoch in the BBS (Λ,p) model is the sequence $\omega_1,\omega_1,\cdots,\omega_\nu$. Because the system's effective age is restored to 0

after a perfect repair is performed, it suffices to observe the system until the ν^{th} failure, which occurs at time ω_ν. Hence we focus on the feature of an epoch of a BBS(Λ, p) model. The probability mass function, $f_\nu(k)$, for a BBS(Λ, p) is given by (see [12])

$$f_\nu(k) = \frac{1}{(k-1)!}\int_0^\infty \exp\{-\Lambda(\omega)\} \times [\Lambda^*(\omega)]^{k-1} p(\omega)\lambda(\omega)d\omega, k = 1,2,\cdots, \quad (6)$$

where $\Lambda^*(\omega) = \int_0^\omega \lambda^*(s)ds$ and $\lambda^*(s) = [1 - p(s)]\lambda(s)$.

Smooth test of goodness-of-Fit for BBS Model

We consider the stochastic formulation developed by Hollander *et al.* [23].

Let $\mathbf{N}^* = \{(\mathbf{N_1^*}(t), \mathbf{N_2^*}(t), \cdots, \mathbf{N_n^*}(t))\}$ be a multivariate counting process defined by $\mathbf{N_j^*}(t) = \sum_{k=1}^\infty I(\omega_{jk} \le t), j = 1,2,\cdots$ and filtration $\mathbf{F}^* = \{\mathfrak{F}_t^* : t \in \mathfrak{F}\}$ defined by $\mathfrak{F}_t^* = \mathfrak{F}_\circ \vee \bigwedge_{j=1}^n \mathfrak{F}_{jt}^*$, where

$$\mathfrak{F}_{jt}^* = \sigma\{\{N_j^*(s) : s \le t\} U \{U_{jk} : k \ge 1\}\}, \quad (7)$$

with $\mathfrak{F}_\circ$ containing all null sets of $\mathfrak{F}$ [19]. Suppose n independent BBS epochs $\{\omega_{jk} : 1 \le j \le n, 1 \le k \le \nu_j\}$ associated with n units, are observed, where the j^{th} unit has time-dependent covariate process $\mathbf{X}_j(\cdot)$ [20]. Define the stochastic processes $\mathbf{N} = \{N_1(t), N_2(t), \cdots, N_n(t) : i \in \mathfrak{F}\}$ with $N_j(t) = N_j^*(t \wedge \omega_{j\nu_j}), j = 1,2,\cdots$ and the corresponding filtration $\mathbf{F} = \{\mathfrak{F}_t : t \in \mathfrak{F}\}$ is given by $\mathfrak{F}_t = \vee_{j=1}^n \mathfrak{F}_{j(t\wedge\omega_{j\nu_j})}^*$. The compensator $\mathbf{F}$ of $\mathbf{N}$ is given as $\mathbf{A} = \{A_1(t), A_2(t), \cdots, A_n(t) : t \in \mathfrak{F}\}$, with

$$A_j(t) = \sum_0^t Y_j(s)\lambda(s)ds, j = 1,2,\cdots,n, \quad (8)$$

where $Y_j(s) = I\{\omega_{j\nu_j} \ge s\}$, and $\lambda(\cdot)$ is the baseline hazard function.

REFERENCES

[1] Aalen OO, Husebye E. Statistical Analysis of Repeated Events Forming Renewal Processes. Statistics in Medicine. 1991; 10(8): 1227-1240. https://doi.org/10.1002/sim.4780100806

[2] Andersen PK, Borgan O, Gill RD, Keiding N. Statistical models based on counting processes. Springer Science & Business Media 2012.

[3] Cook RJ, Lawless JF. Analysis of repeated events. Statistical Methods in Medical Research 2002; 11(2): 141-166. https://doi.org/10.1191/0962280202sm278ra

[4] Nelson WB. Recurrent events data analysis for product repairs, disease recurrences, and other applications. SIAM 2003. https://doi.org/10.1137/1.9780898718454

[5] Cook RJ, Lawless J. The statistical analysis of recurrent events. Springer Science & Business Media 2007.

[6] Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika 1981; pp. 373-379. https://doi.org/10.1093/biomet/68.2.373

[7] Cox DR, Oakes D. Analysis of survival data. vol. 21. CRC Press 1984.

[8] Huang CY, Luo X, Follmann DA. A model checking method for the proportional hazards model with recurrent gap time data. Biostatistics 2011; 12(3): 535-547. https://doi.org/10.1093/biostatistics/kxq071

[9] Jahn-Eimermacher A, Ingel K, Ozga AK, Preussler S, Binder H. Simulating recurrent event data with hazard functions defined on a total time scale. BMC Medical Research Methodology 2015; 15(1): 16. https://doi.org/10.1186/s12874-015-0005-2

[10] Baxter LA, Kijima M, Tortorella M. A Point Process Model for the Reliability of a Maintained System Subject to General Repair. Stochastic Models 1996; 12(1): 12-11. https://doi.org/10.1080/15326349608807372

[11] Agustin MZN, Peña EA. A basis Approach to Goodness-of-fit Testing in Recurrent Event Models. Journal of Statistical Planning and Inference 2005; 133(2): 285-303. https://doi.org/10.1016/j.jspi.2004.03.022

[12] Agustin MZN, Pena EA. Order Statistic Properties, Random Generation, and Goodness-of-fit Testing for a Minimal Repair Model. Journal of the American Statistical Association 1999; 94(445): 266-272. https://doi.org/10.1080/01621459.1999.10473842

[13] Block HW, Borges WS, Savits TH. Age-dependent Minimal Repair. Journal of Applied Probability 1985; pp. 370-385. https://doi.org/10.1017/S0021900200037839

[14] Brown M, Proschan F. Imperfect Repair. Journal of Applied Probability 1983; pp. 851-859. https://doi.org/10.1017/S0021900200024153

[15] Sydney Rosen CJG Matthew P Fox. Patient Retention in Antiretroviral Therapy Programs in Sub-Saharan Africa: A Systematic Review. PLoS Medicine 2007; 4(10): e298. https://doi.org/10.1371/journal.pmed.0040298

[16] Neyman J. Smooth test for Goodness of fit. Scandinavian Actuarial Journal 1937; 1937(3-4): 149-199. https://doi.org/10.1080/03461238.1937.10404821

[17] Pena EA. Smooth goodness-of-fit tests for the baseline hazard in Cox's proportional hazards model. Journal of the American Statistical Association 1998; 93(442): 673-692. https://doi.org/10.2307/2670118

[18] Pena EA. Smooth Goodness-Of-Fit Tests for Composite Hypothesis in Hazard Based Models. The Annals of Statistics 1998; 26(5): 1935-1971. https://doi.org/10.1214/aos/1024691364

[19] Agustin MZN, Peña EA. Goodness-of-fit of the Distribution of Time-to-first occurrence in Recurrent Event Models. Lifetime Data Analysis 2001; 7(3): 289-306. https://doi.org/10.1023/A:1011692525705

[20] Agustin MZN, Peña EA. A basis approach to goodness-of-fit testing in recurrent event models. Journal of Statistical Planning and Inference 2005; 133(2): 285-303. https://doi.org/10.1016/j.jspi.2004.03.022

[21] Dorado CA. On a General Repair Model for Repairable Systems 1995.

[22] Dorado C, Hollander M, Sethuraman J. Nonparametric Estimation for a General Repair Model. The Annals of Statistics 1997; pp. 1140-1160.

[23] Hollander M, Presnell B, Sethuraman J. Nonparametric Methods for Imperfect Repair Models. The Annals of Statistics 1992; pp. 879-896. https://doi.org/10.1214/aos/1176348661

[24] Kijima M. Some Results for Repairable Systems with General Repair. Journal of Applied Probability 1989; pp. 89-102. https://doi.org/10.1017/S0021900200041826

[25] Odhiambo, Collins J, Odhiambo, Omolo B. Validation of the Smooth Test of Goodness-of-Fit for Proportional Hazards in Cancer Survival Studies. International Journal of Statistics in Medical Research 2017; 6(2): 49-67.

[26] Kenya A. Response progress report 2014: Progress towards zero. Nairobi, Kenya: Government Printer 2014.

[27] UNAIDS. 90-90-90: An ambitious treatment target to help end the AIDS epidemic. UNAIDS Geneva 2014.

[28] Rachlis B, Cole DC, van Lettow M, Escobar M, Muula AS, Ahmad F, et al. Follow-Up Visit Patterns in an Antiretroviral Therapy (ART) Programme in Zomba, Malawi. PloS One 2014; 9(7): e101875. https://doi.org/10.1017/S0021900200041826

[29] Megerso A, Garoma S, Tolosa Eticha TW, Daba S, Tarekegn M, Habtamu Z. Predictors of Loss to Follow-up in Antiretroviral Treatment for Adult Patients in the Oromia region, Ethiopia. HIV/AIDS (Auckland, NZ) 2016; 8: 83. https://doi.org/10.2147/hiv.s98137

[30] Rasschaert F, Koole O, Zachariah R, Lynen L, Manzi M, Van Damme W. Short and Long term Retention in Antiretroviral Care in Health Facilities in Rural Malawi and Zimbabwe. BMC Health Services Research 2012; 12(1): 444. https://doi.org/10.1186/1472-6963-12-444

[31] Gwynn RC, Fawzy A, Viho I, Wu Y, Abrams EJ, Nash D. Risk factors for loss to follow-up prior to ART initiation among patients enrolling in HIV care with CD4+ cell count 200 cells/μL in the multi-country MTCT-Plus Initiative. BMC health Services Research 2015; 15(1): 1. https://doi.org/10.1186/s12913-015-0898-9

[32] Ramadhani HO, Thielman NM, Landman KZ, Ndosi EM, Gao F, Kirchherr JL, *et al.* Predictors of Incomplete adherence, virologic failure, and antiviral drug resistance among HIV-infected adults receiving Antiretroviral Therapy in Tanzania. Clinical Infectious Diseases 2007; 45(11): 1492-1498. https://doi.org/10.1086/522991

[33] Haddow L, Edwards S, Sinka K, Mercey D. Patients Lost to Follow up: Experience of an HIV clinic. Sexually Transmitted Infections 2003; 79(4): 349-350. https://doi.org/10.1136/sti.79.4.349-b

[34] Odhiambo, Collins J, Odhiambo, Omolo B. A Smooth Test of Goodness-of-Fit for the Weibull Distribution: An Application to an HIV Retention Data. International Journal of Statistics in Medical Research 2017; 6(1): 68-78.

[35] Assembly UG. Political Declaration on HIV and AIDS: Intensifying our Efforts to Eliminate HIV and AIDS. 65th Session New York: UN General Assembly (UNGASS) 2011.

[36] Programme JUN, HIV/AIDS o. Getting to Zero: 2011-2015 strategy. UNAIDS, Geneva 2010.

[37] Stricker SM, Fox KA, Baggaley R, Negussie E, de Pee S, Grede N, *et al.* Retention in care and adherence to ART are critical elements of HIV care interventions. AIDS and Behavior 2014; 18(5): 465-475. https://doi.org/10.1007/s10461-013-0598-6

Received on 21-04-2017 Accepted on 02-05-2017 Published on 04-08-2017

https://doi.org/10.6000/1929-6029.2017.06.03.2