The International Journal of Science, Mathematics and Technology Learning
ISSN: 2327-7971 (Print), ISSN: 2327-915X (Online)
Volume 31, Issue 1, 2024
https://www.ijsmtl-cgrn.org

The Performances of the Cox, Andersen-Gill, Prentice-Williams-Peterson-Total-Time and Wei-Lin-Weissfeld-Total-Time models in Identifying Risk Factors in Patients with Recurrent Diseases: A kidney infections example

Sizwe Vincent Mbona*, Department of Statistics, Faculty of Applied Science, Durban University of Technology, Durban, South Africa.
Tsepang Patrick Mzamane, Department of Statistics, Faculty of Applied Science, Durban University of Technology, Durban, South Africa.
Anisha Ananth, Department of Statistics, Faculty of Applied Science, Durban University of Technology, Durban, South Africa.
Bonginkosi Duncan Ndlovu, Department of Statistics, Faculty of Applied Science, Durban University of Technology, Durban, South Africa.

Received: *06/21/2024;* ***Accepted:*** *07/01/2024;* ***Published:*** *07/19/2024*

Abstract: In many longitudinal studies, when subjects are followed over a period of time, recurrent event frequently occur. However, some analysis focusses only on time to the first event, ignoring the subsequent events. The main objective of this paper was to compare the extended standard Cox models, such as Andersen-Gill (AG), Prentice-Williams-Peterson total time (PWP-TT), PWP-Gap time model, Wei-Lin-Weissfeld total time (WLW-TT), and Cox frailty model, to identify risk factors associated with kidney re-infection. Empirical evaluation and comparison of these different models were performed. The better model was assessed based on the goodness of fit criteria (AIC, BIC and likelihood ratio test). Kidney data that was downloaded from the R statistical software using the command data("kidney") was used to perform analyses in this study. The PWP-TT model had lower standard errors, AIC and BIC values compared to other models, therefore fitted data better and was used to interpret results. The results showed that 81% (HR = 0.19; 95% CI: 0.09-0.39) of the female patients were less likely to experience kidney reinfection than male patients. The risk of recurrent kidney infection was significantly high (HR = 2.32; 95% CI: 1.25-4.29) to patients having an Acute Neptiritis (AN) disease compared to patients with other diseases. While the prevalence of kidney infection remains the public health problem, intervention strategies and awareness campaigned are needed to in order to minimize risk factors behind the recurrent of the disease.

Keywords**:** Andersen-Gill model; Prentice-Williams-Peterson-total-time model; Wei-Lin-Weissfeld-total-time model; Cox frailty model; recurrent disease.

Introduction

Recurrent disease, characterized by repeated alternations between acute relapse and long remission, can be a feature of both common diseases, like ear infections, and serious chronic diseases, such as human immunodeficiency virus (HIV) infection or multiple sclerosis [1]. Subsequent re-occurrences of a disease are influenced by previous occurrences and hence the

correlations among the re-occurrences should be considered when modeling recurrent disease data. For example, the Cox proportional hazards (PH) model for any one re-occurrence ignores that diseases re-occurrences may be correlated [2-4]. The consequence of this is that if there are correlations among disease re-occurrences, then the standard errors of the model parameter estimates are underestimated or biased downwards. Due to the independence assumption of the Cox PH model, it is only appropriate for modelling the time to the first event (disease occurrence) [2,5], which is an inefficient use of the data because data from the later events are discarded.

Models such as “Variance-corrected Cox based models” and “Frailty/random effects” models have been developed that consider correlations among the re-occurrences of diseases [4,6]. In this paper we wish to compare the performances of such models in identifying risk factors associated with recurrent of kidney infections. The current study uses the kidney real data from McGilchrist and Aisbett [7]. In particular, we consider the following extensions of the original Cox PH model and then compare their perfomances:

- Andersen-Gill (AG) model [8];
- Prentice-Williams-Peterson-total-time (PWP-TT) and gap-time (PWP-GT) model [9];
- Wei-Lin-Weissfeld-total-time (WLW-TT) model [10]; and
- Cox frailty model [7].

These models can be separated into either conditional or marginal models. The AG and PWP-TT models are conditional since that a patient at risk is determined from the previous disease re-occurrences. That is, a patient who have had the disease is at risk of its re-occurrence. The at-risk group of patients in the marginal WLW-TT models is not determined from previous re-occurrences of the disease, i.e. the model assumption is that re-occurrences are independent. To adjust for correlations among re-occurrences, a robust sandwich estimator is used to estimate variances of the model parameter estimates [11]. The first three models can be generalized to a Cox frailty model by adding a patient specific random effect to account for the dependencies among the re-occurrences of the disease within a patient. In addition, marginal means/rates model [12-15], which model the mean number of events or the rate of event occurrence, have been considered.

The performances of the above models (AG, PWP-TT, PWP-GT, WLW-TT and Cox frailty) are expected to be different due to their different underlying assumptions [6]. Furthermore, they model correlations among disease re-occurrences differently [5]. The researchers are interest to know which of these models (discussed above) best describe the data of recurrent kidney infections in patients. Although this data is more than a decade old, the researchers believe it is still informative about knowing the risk factors in patients that are associated with recurrent kidney infections.

Materials and Methods

This study uses kidney data from McGilchrist and Aisbett [7], which can be retrieved from the R software package using the command data("kidney"). Altogether, there were 77 cases of data from 38 patients of which one case has missing data. McGilchrist and Aisbett [7] define recurrent times to infection at the point of insertion of the catheter for kidney patients using portable dialysis equipment. These times are right censored in the dataset due to either the catheter being removed for other reasons or the final recurrence times are being censored. Furthermore, each patient was followed for a predetermined number of recurrence times, some of which became censored [7].

Table 1 contains a subset of the dataset and the variables in the dataset. The dataset consists of variables time=recurrence times; status (0=censored; 1=recurrence time); age (years), sex (1=male, 2=female) and disease type (0=Glomerulo Neptiritis (GN), 1=Acute Neptiritis (AN); 2=Polycyatic Kidney Disease (PKD); 3=other). The frail variable is the frailty prediction.

Table 1: A subset of the recurrent kidney infection data from McGilchrist and Aisbett [7]

ID	time	status	Age (years)	sex	Disease type	frail
1	8	1	28	1	Other	2.3
1	16	1	28	1	Other	2.3
2	23	1	48	2	GN	1.9
2	13	0	48	2	GN	1.9
3	22	1	32	1	Other	1.2
3	28	1	32	1	Other	1.2

Standard Cox PH model

The standard Cox PH model for the survival data specifies the hazard of the ith individual as:

$$\lambda_i(t) = \lambda_0(t)\exp(\beta x_i) \quad (1)$$

where $\lambda_0(t)$ is an unspecified baseline hazard function, β is the vector of regression coefficients and x_i is the vector of covariates of the ith individual. The extended Cox models (AG, PWP-TT, PWP-GT, WLW-TT and Cox frailty), model the recurrent time to event outcomes within a subject comprehensively than the standard Cox PH model.

Andersen-Gill (AG) model

The AG model uses the counting process structure of data inputs. The counting process model of AG generalizes the Cox PH model, which is formulated in terms of increments in the number of events along the time line [8]. The AG model assumes that the correlation between event times for a subject can be explained by past events, which implies that the time increments between events are conditionally uncorrelated, given the covariates. It is a suitable model when correlations among events for each individual are induced by measured covariates [16]. The counting process style of data input of the AG model is represented as a series of observations with recurrent times given as $(t_0, t_1], (t_1, t_2], \dots, (t_{n-1}, t_n]$. Each recurrent event for the ith individual; $i = 0,1,2, \dots, n$; is assumed to follow the PH model. The AG model ignores

the order of the events leaving each subject to be at risk for any event as long as they are still under observation at the time of the event. This further means that a subject could be at risk for a subsequent event without having experienced the prior event. The hazard function is given as:

$$\lambda_i(t) = \lambda_0(t)exp\{\beta_k x_i(t)\} \quad (2)$$

Under this model, the risk of recurrent event for an individual follows the Cox PH model assumption, but the number of recurrent events is not taken into consideration.

Prentice, Williams and Peterson (PWP) model

The PWP model analyses ordered multiple events by stratification, based on the prior number of events during the follow-up period [9]. The PWP counting process model is similar to the AG counting process model but stratified by events. All participants are at risk for the first stratum, but only those with an event in the previous stratum are at risk for the successive one [17].

Besides using the total time (TT) as in the AG model, the PWP model can also be usually defined in terms of gap time (GT), which is the time since the previous event. When using a gap or waiting-time scale, the time index is reset to zero after each recurrence of the event, with assumption of a renewal process. Gaps between events are often useful with infrequent events, when a renewal occurs after an event or when the interest lies on prediction of a next event. Hence, two stratified PWP models can be fitted: (1) PWP-TT, which evaluates the effect of a covariate for the kth event since the entry time in the study; (2) PWP-GT, which evaluates the effect of a covariate for the kth event since the time from the previous event. Unlike the AG model, the effect of covariates may vary from event to event in the stratified PWP models. Therefore, the PWP models might be preferable to the AG model when the effects of covariates are different in subsequent events, which is likely to be the case for diseases such as kidney infections in patients. The baseline hazards vary from event to event, the hazard function for the kth event for the ith individual with the PH form is written as

$$\lambda_{ik}(t) = \lambda_0(t)e^{\beta_k x_i(t)} \quad (3)$$

Both PWP approaches are conditional models as an individual is not considered in the risk set for the kth event until experiencing the $(k-1)$th event. The PWP-GT model describes an intensity process from the occurrence of an immediately preceding event, with the gap time defined as $(t - t_{k-1})$.

$$\lambda_{ik}(t) = \lambda_0(t - t_{k-1})e^{\beta_k x_i(t)} \quad (4)$$

where $\lambda_{0k}(t)$ represents the event-specific baseline hazard for the kth event over time. AG model and both PWP models are adjusted by estimating the sandwich type estimators and hence they are known as variance corrected models [19].

Wei, Lin and Weissfeld (WLW) model

The WLW (marginal means/rates) model is an alternative model for analysing recurrent events, which can be interpreted in terms of the mean number of events when there are no time-dependent covariates [14,15,20-22]. This approach does not specify dependence structures among recurrent event times within a subject. However, since the marginal means/rates model considers all recurrent events of the same subject as a single counting process and does not require time-varying covariates to reflect the past history of the process, this model is more flexible and parsimonious than AG model [15]. If no time-dependent covariates are included in the AG model to account for all the influence of the prior events on future recurrences, point

estimates from the means/rates model and the AG model are the same. Nevertheless, the covariance matrix estimate for the regression coefficients for the marginal means/rates model uses score residuals in the middle of the sandwich estimate, which corrects for the dependency structure. This approach can be of interest in many medical applications when the dependence structure is complex and unknown, especially when it cannot be characterized by including time-varying covariates, as in the AG model.

The hazard function for the kth event for the ith individual is

$$\lambda_{ik}(t) = \lambda_0(t)e^{\beta_k x_i(t)} \quad (5)$$

Unlike the AG model, this model allows a separate underlying hazard for each event. When an event is zero, it means that an individual is no longer at risk after the last given event [20,23].

Cox frailty model

The frailty model also known as the random effects approach, is an extension of the Cox PH model, in which, the hazard function depends on an unmeasured random variable [23,24]. The Cox frailty model introduces a random covariate into the model that induces dependence among the recurrent event times [25]. The idea is that the random effect describes excess risk or frailty for distinct individuals, considering unmeasured heterogeneity that cannot be explained by observed covariates alone. The most commonly used frailty model is a shared frailty model with random effects assumed to follow a gamma distribution with mean equal to one and unknown variance [23]. The model assumes that the recurrent event times are independent conditional on the covariates and random effects. When there is heterogeneous susceptibility to the risk of recurrent events, the frailty model can be applied. For instance, when evaluating recurrent infections at the point of catheter insertion in dialysis patients, the study population can be considered as a mixture of individuals with different hazards, but the characteristics for differences between individuals are not captured by the measured covariates. In such applications, frailty models can be a possible choice.

The hazard function $\lambda_{ij}(t)$ for the recurrent time of the kth event in the ith individual $(j = 1, 2, \dots, k_i; i = 1, 2, \dots, n)$ conditional on the frailty z_i, follows the PH form and its given by:

$$\lambda_{ij}(t) = \lambda_{0k}(t)z_i e^{\beta_k x_i(t)}, \quad t > 0 \quad (6)$$

where, $\lambda_{0k}(t)$ is the common baseline hazard function, x_i is a vector of observable covariates and β is a vector of unknown regression coefficients. Frailty z_i is the unobserved (random) common risk factors shared by all subjects in cluster 'i' and is assumed to be independent and identically distributed (i.i.d) random variable with unit mean and unknown variance (θ) [24,26]. The Frailty effects occur when the observed sources of variation in the observed or unobserved explanatory variables fail to account for the true difference in risk. That is, when there are other important but omitted variables presented, the effect of omitted variable can be captured by frailty.

In order to compare the performance of the models, researchers of this study used the Akaike information criteria (AIC) and Bayesian information criteria (BIC). The smallest AIC and BIC values suggest a better model.

Results

The results in Tables 2 and 3 contain the demographic information of the patients in the kidney infection data and other descriptive statistics of the data. The median follow-up time was 39.5 months (range = 560 months). Half of the patients (50%) were between 41-55 years old, followed by those who were 56 years and above (21.1%), 26-40 years (15.8%), and 10-25 years (13.2%) (Table 2). The researchers further observed that more females (73.7%) participated in the study than males (26.3%). The results also show that 34.2% of the patients participated in the study had other types of diseases that were not mentioned, followed by those who had AN (31.6%), GN (23.7%), and PKD (10.5%). The results presented in Table 3 show the percentage number (by variables) of patients who experienced the first event and percentage of those who experienced the second event.

Table 2: Demographic information of the patients in the kidney infection data.

Variables	Baseline (n=38)
	n (%)
Median follow-up time (months)	39.5 (Range = 560)
Median age in years (SD)	45 (14.8)
Age group	
10 to 25	5 (13.2)
26 to 40	6 (15.8)
41 to 55	19 (50.0)
56 and above	8 (21.1)
Sex	
Male	10 (26.3)
Female	28 (73.7)
Disease	
GN	9 (23.7)
AN	12 (31.6)
PKD	4 (10.5)
Other	13 (34.2)

- SD= standard deviation

Table 3: Demographic information of the patients in the kidney in data by kidney infection recurrence.

Variable	Occurrence 1	Occurrence 2
	n (%)	n (%)
Age group		
10 to 25	5 (15.6)	4 (15.4)
26 to 40	6 (18.7)	4 (15.4)
41 to 55	14 (43.8)	12 (46.1)
56 and above	7 (21.9)	6 (23.1)
Sex		
Male	10 (31.3)	8 (30.8)

Female	22 (68.7)	18 (69.2)
Disease		
GN	8 (25.0)	6 (23.1)
AN	9 (28.1)	9 (34.6)
PKD	3 (9.4)	3 (11.5)
Other	12 (37.5)	8 (30.8)

The cumulative hazard plot in Figure 1 showed that both male and female patients have different estimated cumulative hazard over time. Male patients had higher cumulative hazard than female patients at the beginning of a follow-up study. This means that males were experiencing kidney infection faster than females from the beginning of a follow-up period until month 400.

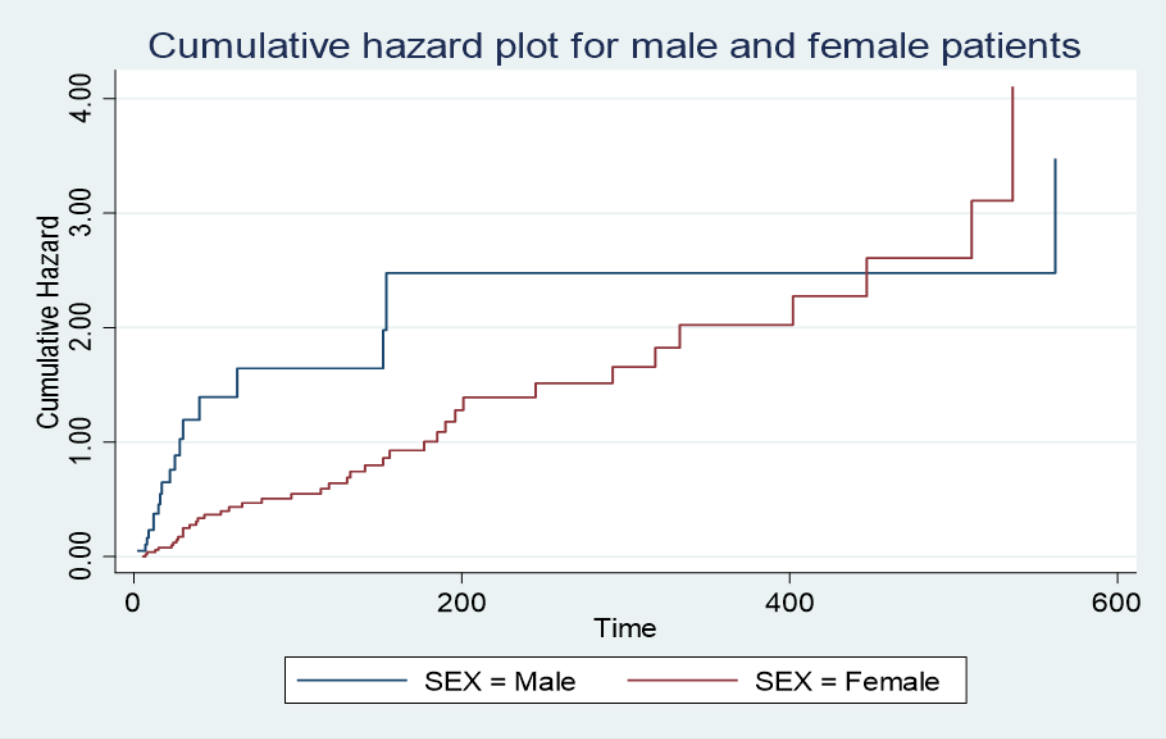

Figure 1: Cumulative hazard plot for kidney infection recurrence over a time of follow-up for male and female patients.

The researchers present the results for AG, PWP-TT, PWP-GT, WLW-TT and Cox frailty models with common effects. The researchers present the hazard ratios (HR) and corresponding 95% confidence intervals (CIs) for the risk factors for kidney infection recurrences. The likelihood ratio test (LRT) result for testing the regression parameter estimates, standard errors (S.E), and p-value of the models are displayed in Table 4. The parameter estimates obtained were consistent in sign for all the models. However, the standard error was relatively smaller for the PWP-TT model when compared to the other models. The LRT value for the AG model was 18.73 with a p-value = 0.009. For the PWP-TT, PWP-GT, WLW-TT and Cox frailty models, the LRT values were 20.04, 21.67, 34.42 and 37.72, respectively. It was observed that the PWP-TT model had the lowest AIC and BIC values of 263.5477 and 277.9708, respectively, suggesting that the PWP-TT model was better fitted to

the kidney dataset than the other models. The HR, standard error, and p-value for the PWP-TT model displayed in Table 4 show that sex and AN were important factors for recurrent of kidney infection. However, the variables, age had no significant effect on the recurrent of kidney infection in patients.

The estimated acceleration factor or risk of recurrent of kidney infection for the female patients was 0.19, with a 95% confidence interval (CI) of (0.09 – 0.39). After adjusting for other covariates, a patient with AN disease had an increased risk of recurrent of kidney infection by 2.32 times compared to other diseases, holding the effects of other covariates constant.

Table 4: Risk factors for kidney infection recurrent event data using extended standard Cox models.

Variable	AG model			PWP-TT model			PWP-GT model			WLW-TT model			Cox frailty model		
	HR (95% CI)	S.E	P value	HR (95% CI)	S.E	P value	HR (95% CI)	S.E	P value	HR (95% CI)	S.E	P value	HR (95% CI)	S.E	P value
Age group															
10 to 25 (Ref)															
26 to 40	0.81 (0.33-1.99)	0.48	0.64	1.02 (0.35-2.93)	0.48	0.98	1.13 (0.43-294)	0.49	0.81	0.93 (0.27-3.16)	0.50	0.90	1.00 (0.69-1.46)	0.51	0.79
41 to 55	0.63 (0.34-1.19)	0.48	0.15	0.65 (0.29-1.44)	0.43	0.29	0.75 (0.35-1.60)	0.50	0.45	0.78 (0.30-2.00)	0.51	0.60	0.69 (0.39-1.95)	0.48	0.32
56 and above	1.08 (0.50-2.37)	0.53	0.84	1.07 (0.43-2.65)	0.52	0.89	1.33 (0.58-3.03)	0.55	0.50	1.22 (0.43-3.45)	0.56	0.71	1.06 (0.74-2.63)	0.55	0.83
Sex															
Male (Ref)															
Female	0.22 (0.11-0.44)	0.37	<0.01*	0.19 (0.09-0.39)	0.35	<0.01*	0.18 (0.09-0.37)	0.38	<0.01*	0.11 (0.05-0.26)	0.41	<0.01*	0.26 (0.11-0.60)	0.43	<0.01*
Disease															
Other (Ref)															
GN	1.43 (0.81-2.51)	0.47	0.21	1.78 (0.90-3.54)	0.46	0.10	1.71 (0.86-3.44)	0.47	0.13	1.76 (0.79-3.90)	0.48	0.16	1.62 (0.76-2.74)	0.52	0.35
AN	1.87 (1.09-3.19)	0.45	0.02*	2.32 (1.25-4.29)	0.44	0.01*	1.71 (0.96-3.05)	0.46	0.07	2.96 (1.53-5.73)	0.47	<0.01*	1.96 (1.33-3.87)	0.44	0.04*
PKD	0.40 (0.11-1.51)	0.63	0.18	0.41 (0.10-1.68)	0.60	0.21	0.31 (0.07-1.44)	0.67	0.14	0.29 (0.06-1.32)	0.66	0.11	0.36 (0.18-1.45)	0.61	0.18
AIC	330.8817			263.5477			291.9754			287.5066			326.4762		
BIC	345.3048			277.9708			306.3985			301.9297			344.0794		
Likelihood ratio test	18.73 on 7 df, p-value = 0.009*			20.04 on 7 df, p-value = 0.005			21.67 on 7 df, p-value = 0.003			34.42 on 7 df, p-value < 0.001			37.72 on 7 df, p-value < 0.001		
Concordance	0.697 (se = 0.033)			0.717 (se = 0.039)			0.684 (se = 0.047)			0.752 (se = 0.038)			0.809 (se = 0.031)		
Wald test	25.42 on 7 df, p-value < 0.001			28.31 on 7 df, p-value < 0.001			29.20 on 7 df, p-value < 0.001			36.65 on 7 df, p-value < 0.001					
Score (log rank) test	20.81 on 7 df, p-value = 0.004			20.77 on 7 df, p-value = 0.004			24.79 on 7 df, p-value < 0.001			41.03 on 7 df, p-value < 0.001					
Robust	15.05, p-value = 0.040			17.04, p-value = 0.02			14.25, p-value = 0.05			16.45, p-value = 0.02					
Variance of random effect													0.295		
DF for terms													0.60		

*S.E = standard error; CI = confidence interval; HR = hazard ratio

Discussion

The increased prevalence of kidney failure and early stages of kidney disease, and the high costs and poor outcomes of treatment constitute a worldwide public health threat. The main responsibility for developing guidelines for chronic kidney disease has now been assumed by Kidney Disease Improving Global Outcomes (KDIGO)—a global non-profit foundation dedicated to improving the care and outcomes of patients with kidney disease worldwide. According to Eknoyan et al. [27] and Levin et al. [28], KDIGO guidelines rate the strength of recommendations and evidence with rigorous and well accepted methods.

In this paper, the researchers discussed known approaches under independent censoring assumption for analysis of recurrent event data. Parameter estimates obtained from various models were considerably different due to the disparity in their properties and assumptions. The AG mode assumes that the time increments between events are conditionally uncorrelated given the covariates. However, omission of an important covariate could induce dependence. In such case, the standard errors would be underestimated, causing inflation of type I error. The PWP models (TT or GT) are also indicated when there is interest in estimating effects for each event separately. The PWP models assume that the individuals can only be at risk for a given event after he/she experienced the previous event. On the other hand, the WLW-TT model also called the means/rates marginal model is useful when the interest is in modelling the expected number of events or the rate of event occurrence, conditional on

covariates. These models are also useful in many applications where there are multiple types of events and it is of interest to simultaneously describe marginal aspects of them.

The frailty models are indicated when a subject-specific random effect can explain the unmeasured heterogeneity that cannot be explained by covariates alone. There is a strong evidence in the literature that if frailty is present but ignored, then the covariate effects will be underestimated [29,30]. A debate about using frailty models is regarding the amount of information, such as number of events, number of subjects and the distribution of events/subjects required to produce stable estimates. When random effects are large, a smaller number of events seems to be adequate, otherwise a larger number of events would be necessary [23].

Based on the findings for AIC, BIC and likelihood ratio test, the PWP-TT model was the better model that fit current recurrent event data. The results agree with earlier researchers [4,31,32]. The negative value of regression coefficient of covariate “sex” indicate that the female patients had a lower risk of kidney infection than male patients. AN disease was found to be the most significant risk factor associated with recurrence of kidney infection in patients. The current finding is consistent with other studies [33].

Limitations

There are various risk factors for the causes of recurrent of kidney infection in patients during treatment period. In this study, we included five risk factors; therefore, there was incomplete information for important demographic, as well as clinical, variables; thus, the study was limited to only those variables described in the methods section. The study used small sample size. Hence, the study was subject to selection bias. Despite these limitations, this study is very important for the follow-up of kidney patients by health professionals during the intensive phase and will contribute in decision making to reducing recurrence during kidney treatment.

Conclusion

To improve outcomes for kidney disease, new treatments will need to be translated into clinical practice and public health. Recommendations for prevention are needed including improvements in surveillance, screening, education, and awareness, which are directed at three target populations: people with or at increased risk of kidney disease; providers, hospitals, clinical laboratories; and the general public.

Declarations

Acknowledgements
Declared none.

Authors’ contributions
All the authors made contribution in this paper. SVM planned the study and wrote the initial draft of the paper and did the analysis. All other authors revised, edited the paper and approved the final work of this manuscript.

Financial disclosure
None.

Potential conflict of interests

The authors declare no conflict of interest, financial or otherwise.

Ethics approval and consent to participate
Not applicable

Consent for publication
Not applicable

Reference

1. Zhang, Wenjing, "Understanding Recurrent Disease: A Dynamical Systems Approach" (2014). Electronic Thesis and Dissertation Repository. 2265. https://ir.lib.uwo.ca/etd/2265
2. Cox DR. Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological). 1972 Jan;34(2):187-202.
3. Anker SD, McMurray JJ. Time to move on from 'time-to-first': should all events be included in the analysis of clinical trials?. European Heart Journal. 2012 Nov 1;33(22):2764-5.
4. Thenmozhi M, Jeyaseelan V, Jeyaseelan L, Isaac R, Vedantam R. Survival analysis in longitudinal studies for recurrent events: Applications and challenges. Clinical Epidemiology and Global Health. 2019 Jun 1;7(2):253-60.
5. Amorim LD, Cai J. Modelling recurrent events: a tutorial for analysis in epidemiology. International journal of epidemiology. 2015 Feb 1;44(1):324-33.
6. Ullah S, Gabbett TJ, Finch CF. Statistical modelling for recurrent events: an application to sports injuries. British journal of sports medicine. 2014 Sep 1;48(17):1287-93.
7. McGilchrist CA, Aisbett CW. Regression with frailty in survival analysis. Biometrics. 1991 Jun 1:461-6.
8. Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. The annals of statistics. 1982 Dec 1:1100-20.
9. Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika. 1981 Aug 1;68(2):373-9.
10. Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. Journal of the American statistical association. 1989 Dec 1;84(408):1065-73.
11. Kim JI, Lin FC, Fine JP. On the self-triggering Cox model for recurrent event data. Statistics in Medicine. 2019 Sep 30;38(22):4240-52.
12. Pepe MS, Cai J. Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. Journal of the American Statistical Association. 1993 Sep 1;88(423):811-20.
13. Lawless JF, Nadeau C. Some simple robust methods for the analysis of recurrent events. Technometrics. 1995 May 1;37(2):158-68.
14. Lawless JF. The analysis of recurrent events for multiple subjects. Journal of the Royal Statistical Society: Series C (Applied Statistics). 1995 Dec;44(4):487-98.
15. Lin DY, Wei LJ, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2000;62(4):711-30.

16. Moulton LH, Dibley MJ. Multivariate time-to-event models for studies of recurrent childhood diseases. International journal of epidemiology. 1997 Dec 1;26(6):1334-9.
17. Pandeya N, Purdie DM, Green A, Williams G. Repeated occurrence of basal cell carcinoma of the skin and multifailure survival analysis: follow-up data from the Nambour Skin Cancer Prevention Trial. American journal of epidemiology. 2005 Apr 15;161(8):748-54.
18. Kleinbaum D, Klein M. Survival Analysis: A Self-Learning Text: Springer Science. NY, NY. 2005.
19. Clayton D. Some approaches to the analysis of recurrent event data. Statistical methods in medical research. 1994 Oct;3(3):244-62.
20. Cook RJ, Lawless JF. The statistical analysis of recurrent events. New York: Springer; 2007 Apr 9.
21. Cai J, Schaubel DE. Analysis of recurrent event data. Handbook of statistics. 2003 Jan 1;23:603-23.
22. Amorim LD, Cai J, Zeng D, Barreto ML. Regression splines in the time-dependent coefficient rates model for recurrent event data. Statistics in medicine. 2008 Dec 10;27(28):5890-906.
23. Therneau TM, Grambsch PM. The cox model. InModeling survival data: extending the Cox model 2000 (pp. 39-77). Springer, New York, NY.
24. Hougaard P, Hougaard P. Analysis of multivariate survival data. New York: Springer; 2000 Aug.
25. Kelly PJ, Lim LL. Survival analysis for recurrent event data: an application to childhood infectious diseases. Statistics in medicine. 2000 Jan 15;19(1):13-33.
26. Duchateau L, Janssen P. The frailty model. New York: Springer Verlag; 2008.
27. Eknoyan G, Lameire N, Barsoum R, Eckardt KU, Levin A, Levin N, Locatelli F, Macleod A, Vanholder R, Walker R, Wang H. The burden of kidney disease: improving global outcomes. Kidney international. 2004 Oct 1;66(4):1310-4.
28. Levin NW, Kotanko P, Eckardt KU, Kasiske BL, Chazot C, Cheung AK, Redon J, Wheeler DC, Zoccali C, London GM. Blood pressure in chronic kidney disease stage 5D—report from a Kidney Disease: Improving Global Outcomes controversies conference. Kidney international. 2010 Feb 2;77(4):273-84.
29. Pickles A, Crouchley R. A comparison of frailty models for multivariate survival data. Statistics in Medicine. 1995 Jul 15;14(13):1447-61.
30. Hougaard P. Frailty models for survival data. Lifetime data analysis. 1995 Sep;1(3):255-73.
31. Lintu MK, Kamath A. Performance of recurrent event models on defect proneness data. Annals of Operations Research. 2020 Dec 7:1-0.
32. Joe Turkson A, Simpson T, Awuah Addor J. An Introspective Overview of the Dynamics of Recurrent Events Data Analysis. 2021.
33. Hanagal DD. Analysis of kidney infection data using correlated Inverse Gaussian frailty model. Statistics and Applications. 2020;18(1):1-9.