Journal of Communicable Diseases
Volume 51, Issue 1 - 2019, Pg. No. 22-28
Peer Reviewed & Open Access Journal

Research Article

Statistical Analysis of Risk factors of Malaria re-infection among Outpatients in DR, Congo: A Comparison Approach of AFT and Cox PH Models

Ruffin Mutambayi, Adeboye Azeez*, James Ndege, YongSon Qin

Department of Statistics, University of Fort Hare, P. Bag-X1314, Alice-5700, South Africa.
DOI: https://doi.org/10.24321/0019.5138.201904

I N F O

Corresponding Author:
Adeboye Azeez, Department of Statistics, University of Fort Hare, P. Bag-X1314, Alice-5700, South Africa.
E-mail Id:
azizadeboye@gmail.com
Orcid Id:
https://orcid.org/0000-0001-9427-7374
How to cite this article:
Mutambayi R, Azeez A, Ndege J et al. Statistical Analysis of Risk factors of Malaria re-infection among Outpatients in DR, Congo: A Comparison Approach of AFT and Cox PH Models. *J Commun Dis* 2019; 51(1): 22-28.

Date of Submission: 2019-02-03
Date of Acceptance: 2019-03-05

A B S T R A C T

Background: Malaria is an infectious disease caused by a Plasmodium parasite and is one of the highest causes of mortality globally. This study aims to determine models to detect the effect of risk factors on malaria re-infection of patients survival.

Methods: The study includes 109 malaria outpatients in Lubumbashi Congo Hospital, who had re-infection status after six months follow-up. The survival status of the re-infected patients was based on the effect of various factors. The best model was selected through Akaike Information Criterion (AIC) and Cox-Snell Residuals using SAS and R packages.

Results: The results from the analysis showed that Gamma model (AIC=147.092) was better in the analysis compared to accelerated failure-time models.

Conclusion: Although, many researchers prefer proportional hazard model in analysing a survival data but accelerated failure-time model is a good alternative method as they do not require proportionality of hazards as key assumption.

Keywords: Accelerated FT Models, Cox PH, Cox-Snell Residuals, Gamma Distribution, Malaria

Introduction

Malaria is one of the leading causes of death in some African countries. It is a vector-borne parasitic infection causing the most deaths worldwide and among the highest causes of mortality in the Democratic Republic of Congo.[1,2,3,4] Globally, malaria is the greatest prevalent parasitic sickness and accounts for more than half a million death annually.[5] A bite from an infected Anopheles mosquito cause malaria.[6] Despite having many antimalarial treatments, many treatment methods, awareness and prevention reports on the reduction of malaria death in many countries, it still remains a public health challenge in many of these countries.[7] Almost half of the entire world population is at risk of malaria and more than 200 million cases are reported with an estimated of 5 million malaria death cases annually and ninety percent of these deaths occur

in Africa, 7% from South-East, Asia and 2% from Eastern Mediterranean Regions.[8]

In the endemic countries like the DR Congo, malaria accounts for at least 40% of public health challenge with nearly 25 to 55% of hospitalization.[9,8] The increase of re-infection cases is a great challenge and contributor to death of malaria patients, which necessitate from the effect different determinants or factors that could be linked to the rise of patients death.

Various statistical methods have been identified to evaluate the impact factors of survival status on patients' infection of a particular disease including parametric, semi-parametric and non-parametric survival models such as Proportional Hazard (PH) model and Accelerated Failure-Time (AFT) model. In a parametric proportional hazard model, Weibull, exponential or/and Gompertz models can be used to evaluate the effect of covariates on hazard function. While in non-parametric hazard model, the Cox proportional hazard model can be used to detect the effect of covariates on hazard function. However, accelerated failure-time model is used to assess the effect of covariates on the logarithm of survival time such as generalized Gamma AFT models, Log-logistic AFT models, Log-normal AFT models, Weibull AFT models and exponential AFT models. Both Weibull and exponential models are common to parametric PH models and accelerated failure-time models.

In analysing a survival data using the Cox regression model, there is no need for probability assumption for the survival times. As a result, the hazard function does not have a restriction to a specific functional form but a flexible model with widespread acceptability. However, if the assumption of a proportional hazard model is not valid or unacceptable, the inference of such assumption will be incorrect and Cox model estimates obtained will have improper fitting of the model.[10,11] Hence, AFT models may be relevant and used instead. Since AFT models having a parametric distribution for the survival times, they tend to have accurate statistical inference and proper model fitting.[12]

Many studies have used Cox proportional hazard model to identified factors affecting the survival of patients for a particular disease but none of these studies have tested the assumptions of PH models if violated, to provide an alternative model in analysis survival data particularly in malaria re-infection study.[13] However, this study is comparing and identifying the best model between Cox-PH and an alternative AFT model for patients with malaria relapse treatment. The Cox-Snell Residuals and Akaike Information Criterion (AIC) will be used to verify the goodness-of-fit of the survival models.

Materials and Methods

The study consists of 109 patients with malaria re-infection were identified based on the daily screening for re-infection Malaria. The data collected (at screening, inclusion into the study and daily until Malaria free, death following inclusion), from Poly-clinic Kiubo, Katumba III, Katanga Province, DR, Congo from February 2016 to July 2017 included the following: age, type of malaria, environment, type of house, toilet type and the survival time of the patient for having malaria after being treated. The effects of environmental, sanitation, household characteristics, preventive and demographical variables on the re-infection of patients were evaluated and compare among various models considered into this study using the time to re-infection as dependent variable. Akaike Information Criterion (AIC) and Cox-Snell Residuals were used to compare and identify the best model for the survival times. The Cox-Snell residuals plots aim to evaluate the goodness-of-fit of both Cox PH and AFT models. The short residuals deviation from the straight line through the origin with a slope of 1 is a better and suitable survival model.[14, 15] Visual error might be associated with Cox-Snell residuals Plot, hence for better decision, AIC concept are advisable to reach a better and precise decision. AIC model selection is used to measure the model goodness of fit and a smaller value of AIC indicating a better model selection.[16, 17] The formula bellow was applied in the determination of AIC's related to the models investigated in this research:

$$AIC = -2 \times \log(\ell) + 2(n)$$

where n is the number of model parameters and ℓ is the model likelihood function.[18,19]

A smaller AIC-value relate to a more powerful model in identifying the risk factors.[20,21] Furthermore, to access the comparison of the variables' variances, standardized variability is used in the model and is calculated as $\sigma\sigma^2 = \frac{se(\hat{\beta})}{\hat{\beta}}$, which was used to standardize the variance of parameters estimated, where $se(\hat{\beta})$ is the standard error of parameter and $\hat{\beta}$ is the coefficient of parameter in the survival model. SAS 9.3 and R software was used for all analyses and the significance level was set at 5%.

Statistical Analysis

The demographic variables are reported as frequency and percentage unless otherwise stated. Categorical variables were analysed by Fisher exact test and all interval variables by t-test, where suitable in application. All analyses were performed using SAS 9.3 and R software. Binary data like sex were coded in a simple indicator way of 0 and 1. Time to re-infection of Malaria were censored after the patients' test indicate the presence of Malaria at 21 days was analysed initially using Kaplan-Meier, Cox model and AFT model for comparison. Model selection to choose best performing model was accomplished using Akaike and Bayesian Information Criteria (AIC and BIC) respectively, and the model comparison was done using standardized

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.201904

variability, Relative risk and Hazard ratio based on the partial likelihood estimates criteria. Basically, Cox model fit was evaluated through Kaplan-Meier estimates, residual plots and assumption test of hazard function. The re-infection time was also modelled by parametric survival analysis using Accelerated Failure Time (AFT) model, where we parameterised the log time, the covariate vector and the corresponding coefficient to have a specified distribution for gamma, exponential, Weibull, log-normal and log-logistic respectively.

Results

The majority of patients were male (61.2%) patients; 81.6 percent of the patients were between the ages of 18-59 years; 34.7 percent were single; 42.9 percent did not complete secondary school education and 67.9% of the patients were living in a brick house type. More than two-thirds of the total patients (89.9%) relapsed from the uncomplicated falciparum malaria after 6 months follow-up from the day the treatment completed to the time of re-infection of malaria. The median survival time for re-infection was 3.01 months with a mean survival time for re-infection of 2.81 and a survival rate of re-infection of 0.1560. Meanwhile, 69.7% were re-infected through water from an unprotected source; 57.1% were re-infected through dirty facilities, 58% of patients were re-infected through no information about malaria, and 61.2% of patients were re-infected through the use of mosquito's spray (Table 1).

In Figure 1, Cox-Snell residuals plots approach was used to fit the malaria data using AFT and Cox PH models. Generally, the AFT models seem to fit the data better. In AFT models, the Cox-Snell residuals plot for the Gamma model showed a better fit for the re-infection data (Figure 1). Akaike information criterion confirms these results with the lowest value related to Gamma distribution model (AIC= 147.092) is better than the rest of models used in this study (Table 2).

Moreover, the risk factors of malaria re-infection were identified by using AFT and Cox PH models to analysis the data through the standardized variances, Hazard Ratio (HR) and Relative Risk (RR) for all the covariates. The results were summarized in Table 2 shows that only the exponential and Gamma model were statistically significant based on partial likelihood estimation from AFT models and the rest show insignificance. Although, the hazard rate in Cox PH model is virtually the same as the results of AFT models and Gamma model has better results while the exponential has the least results according to Akaike information criterion based on AFT models.

Results of Cox PH model show that the usage of mosquito's spray and awareness about malaria identified as risk factors of re-infection among the malaria patients (P-value < 0.05). the Cox PH models results and the analysis of Gamma AFT model revealed that gender, source of drinking water, maintenance of the toilet was identified as risk factors of malaria re-infection of among the patients (P-value <0.05). The exponential model results revealed that none of the covariates considered were significant as risk factors of malaria re-infection among malaria patients (P-value >0.05). In this study, treatment, age of patient, marital status, level of education, type of housing, type of toilet, dumping space, pit toilet, stagnant water, cultivating in the compound, usage of mosquito's net and net treated with chemical did not have any significant effect on the re-infection malaria.

Table 1.Malaria re-infected patients' demographic characteristics

Variables	Frequency (%)
Gender	
Male	60 (61.2)
Female	38 (38.8)
Age Group	
18-35 years	41 (41.8)
36-59 years	39 (39.8)
60 years and above	19 (19.4)
Marital Status	
Single	34 (34.7)
Married	49 (50.0)
Divorced	6 (6.1)
Widowed/ Widower	9 (9.2)
Level of Education	
Completed primary school	18 (18.4)
Primary school not completed	10 (10.2)
Completed secondary school	13 (13.3)
Secondary school not completed	42 (42.9)
Completed College/ University	7 (7.1)
College education not completed	8 (8.2)
Surrounding Area	
Clean place	42 (42.9)
Dirty place	56 (57.1)
Source of Drinking Water	
Unprotected dug well	68 (69.7)
Protected dug well	30 (30.3)

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.201904

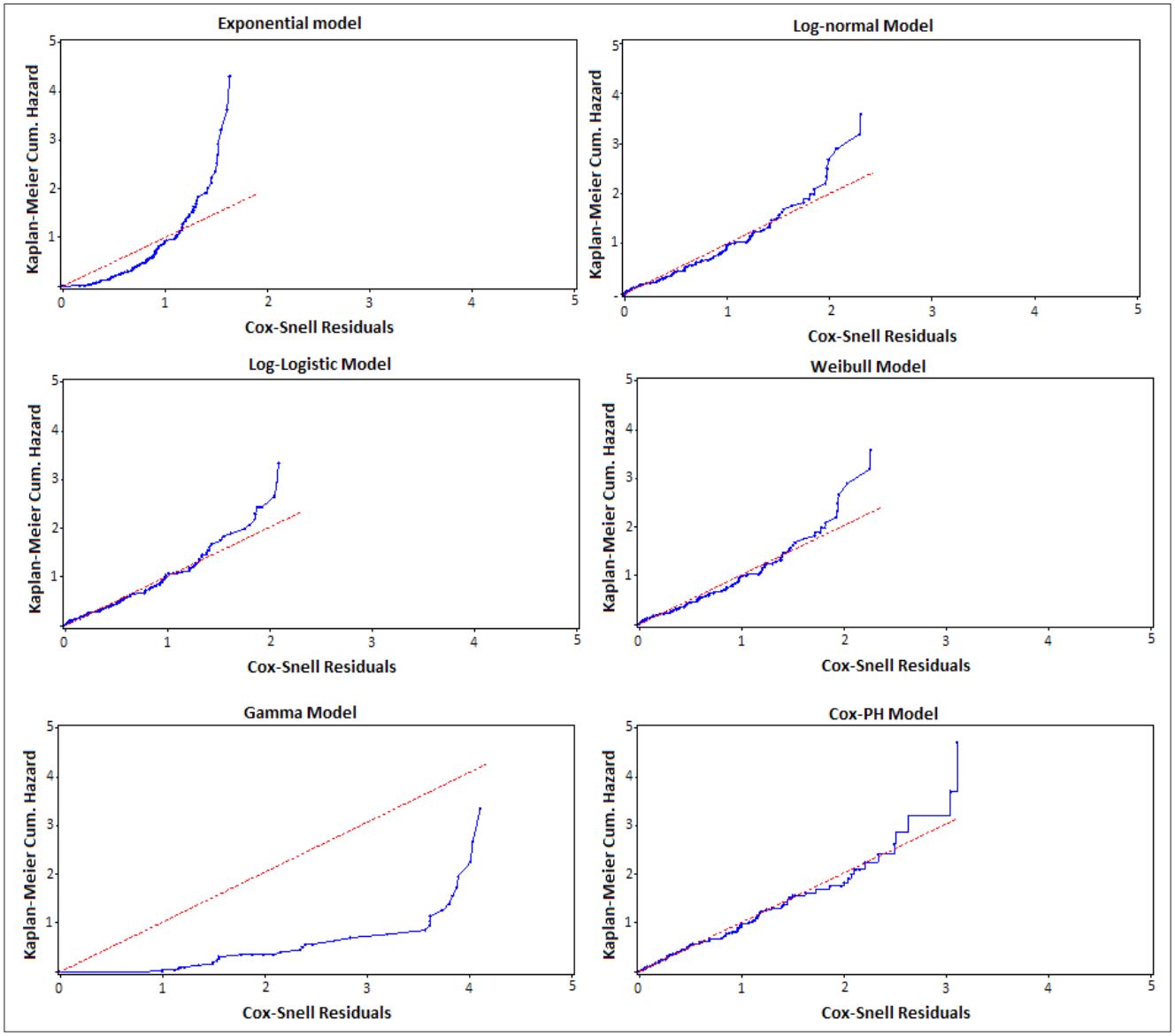

Figure 1.The Cox-Snell residuals plots in the considered AFT and Cox PH models

Table 2.Contrast of outcome of threat determinants between AFT and Cox PH

Risk factors	Exponential RR (SV)	Weibull RR (SV)	Log-logistic RR (SV)	Log-normal RR (SV)	Gamma RR (SV)	Cox HR (SV)
Treatment	0.963 (0.2193)	0.977 (0.0835)	0.986 (0.1022)	0.963 (0.1027)	0.998 (0.0526)	1.052 (0.23)
Gender	0.823 (0.2380)	0.893 (0.0896)	0.830 (0.1078)	0.827 (0.1126)	0.908 (0.0477)	1.326 (0.25)
Age	0.891 (0.1505)	0.952 (0.0562)	0.912 (0.0702)	0.877 (0.0730)	1.006 (0.0371)	1.108 (0.16)
Marital Status	0.932 (0.1402)	0.956 (0.0558)	0.923 (0.0627)	0.904 (0.0622)	1.021 (0.0262)	1.112 (0.16)
Level of education	0.992 (0.0660)	0.994 (0.0236)	1.005 (0.0311)	1.002 (0.0330)	0.978 (0.0156)	1.029 (0.07)
Type of dwelling	0.768 (0.2684)	0.852 (0.1040)	0.915 (0.1216)	0.917 (0.1191)	0.924 (0.0688)	1.734 (0.30)
Source of drinking water	0.873 (0.7408)	0.977 (0.2778)	1.126 (0.3915)	1.018 (0.3602)	0.770 (0.1129)	0.851 (0.78)
Type of toilet	0.815 (0.4175)	0.861 (0.1486)	0.941 (0.2155)	0.945 (0.2105)	0.865 (0.0897)	1.468 (0.42)

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.201904

Maintained toilet	0.542 (0.8951)	1.886 (0.3395)	1.787 (0.3843)	1.920 (0.4380)	2.022 (0.0864)	0.179 (0.92)
Dumping sites	1.348 (0.7021)	1.693 (0.3733)	1.571 (0.2808)	1.249 (0.2633)	1.263 (1.6370)	0.203 (1.08)
Pit toilet covered	0.480 (1.2027)	0.421 (0.4974)	0.406 (0.5605)	0.415 (0.5580)	0.283 (1.6409)	14.158 (1.43)
Stagnant water	2.492 (1.2567)	1.809 (0.5340)	1.494 (0.5000)	1.874 (0.5773)	2.995 (1.6459)	0.348 (1.48)
Cleaning the yard	1.394 (0.5412)	1.352 (0.2709)	1.348 (0.2294)	1.381 (0.2228)	2.265 (1.6285)	0.401 (0.74)
Mosquito's spray	0.833 (0.2319)	0.880 (0.0844)	0.971 (0.1075)	0.964 (0.1127)	0.787 (0.0382)	1.578 (0.24)
Net	0.992 (0.3863)	0.972 (0.1427)	0.976 (0.1870)	1.021 (0.1858)	0.893 (0.0868)	1.070 (0.40)
Net treated with chemical	0.979 (0.2427)	1.005 (0.0955)	1.088 (0.1091)	1.085 (0.1127)	0.992 (0.0493)	1.029 (0.27)
Information about malaria	0.242 (1.1684)	0.298 (0.4269)	0.336 (0.4900)	0.310 (0.5761)	0.260 (0.1433)	18.311 (1.23)
AIC	278.697**	183.946	199.491	201.07	147.092*	780.037

SV: Standardized Variability, RR: Relative Risk, HR: Hazard Ratio, **Based on Partial Likelihood.

Discussion

Many studies have used Cox PH model to investigate the effect of different covariates on survival of patients in a particular disease rather than AFT models. Among the AFT models, Gamma was the best choice, which demonstrated to be a good time-varying covariate effect[22,23] and similarly noted in a previous study[24] that the number of variables for non-nested multivariable models while exponential was the least choice to model malaria re-infection. The lack of proportionality of hazards assumption may lead to an unreliable and biased model, therefore, accelerated failure-time models such as Exponential, Gompertz, Log-logistic, Log-normal, generalized Gamma or Weibull can stand as an alternative and better choice in such circumstances. As accelerated failure-time models assess a statistical distribution for survival time and they do not require proportionality of hazards, they are suitable alternatives to Cox PH model.

In this study, the results of Cox PH models and AFT models were compared to identify the risk factors of malaria re-infection among the follow-up patients. The analysis of Cox-Snell residuals (Figure 1) showed that AFT models fit the data better compare to Cox PH model. Moreover, based on Akaike information criterion in Table 1, the analysis of models showed that Gamma model was the best alternative for Cox proportional hazard model. There was a significant difference between Gamma AFT model and proportional hazard model in identifying factors related to the survival of patients with malaria re-infection. In one hand the analyses of AFT and Cox PH models showed that the usage of mosquito's spray, awareness and information about malaria, gender, source of drinking water, maintenance of the toilet were covariates risk factors on the re-infection of malaria's patients (P-value < 0.05). These results are consistent with other studies in this area of research.[25-28] However, the study revealed that, age of patient, marital status, level of education, type of dwelling, type of toilet, dumping space, pit toilet, stagnant water, cultivating in the compound, usage of mosquito's net and net treated with chemical did not have any significant effect on the re-infection of malaria patients from the AFT models. Based on AIC & Cox-Snell residual, Gamma model is the best parametric alternative for Cox proportional hazard model. In some studies, however, Weibull has been considered as the good model but Gamma is a specific case of Weibull.

Conclusion

Though, Cox proportional hazard model has been used by many studies in medical and survival researches and the results of accelerated failure-time models are often be more valid and less bias since AFT models have better fit due for the survival time and have no need for assumptions. Accelerated failure-time models can sometime be a substitute of the Cox proportional hazard model when PH assumptions fail. Furthermore, accelerated failure-time models produce more efficient parameters than Cox PH base on asymptotic results. Note that a sample's size reduction may affect the relative efficiency of parameters which may benefit AFT models. When experimental findings are adequate, AFT models can be used into the form of the baseline hazard.

Ethical Considerations

Ethical research consideration was ensured during the study, with ethical clearance being obtained from the ethics'

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.201904

committee of Fort Hare University. Ethical considerations including informed consent, anonymity, confidentiality, protection from harm, provision of privacy, objectivity and honesty in reporting findings and procedures were also applied.

Acknowledgement

The authors thankfully appreciate the financial support for this piece of work provided by NRF (85307) and Fort Hare University through Govan Mbeki Center of Research (GMRDC) and, enunciate also that there is no disagreement of interests.

Conflict of Interest: None

References

1. Mapping Malaria Risk in Africa (MARA), *Low-end Information Tool (LITe),* 2002, V3.0.0 Build 7. http://www.mara.org.za/lite/information.htm.
2. Parham PE, Waldock J, Christophides GK et al. Climate, environmental and socio-economic change: weighing up the balance in vector-borne disease transmission. *Philos Trans R Soc Lond B Biol Sci* 2015; 370(1665). pii: 20130551.
3. Van Herp M, Parque V, Rackley E et al. Mortality, violence and lack of access to healthcare in the Democratic Republic of Congo. *Disasters* 2003; 27(2): 141-153.
4. World Health Organization Global Health Observatory. Available from: http://www.who.int/gho/database/en/.
5. Murray CJ, Rosenfeld LC, Lim SS et al. Global malaria mortality between 1980 and 2010: a systematic analysis. *Lancet* 2012; 379(9814): 413-431.
6. Prudêncio M, Rodriguez A, Mota MM. Review The silent path to thousands of merozoites: the Plasmodium liver stage. *Nat Rev Microbiol* 2006; 4(11): 849-856.
7. O'Meara WP, Mangeni JN, Steketee R et al. Changes in the burden of malaria in sub-Saharan Africa. *Lancet Infect Dis* 2010; 10: 545-555.
8. World Health Organization. World malaria report, 2014. WHO Press, Geneva, Switzerland. Available from: http://www.who.int/malaria/media/world-malaria-report-2014/en/.
9. Bourtzis K, Lees RS, Hendrichs J et al. More than one rabbit out of the hat: radiation, transgenic and symbiont-based approaches for sustainable management of mosquito and tsetse fly populations. *Acta Trop* 2016; 157: 115-130.
10. Attallah O. The risk of re-intervention after endovascular aortic aneurysm repair. PhD thesis, Aston University. 2016.
11. Zare A, Hosseini M, Mahmoodi M et al. A comparison between accelerated failure-time and cox proportional hazard models in analyzing the survival of gastric cancer patients. *Iranian Journal of Public Health* 2015; 44(8): 1095-102.
12. Montaseri M, Charati JY, Espahbodi F. Application of parametric models to a survival analysis of hemodialysis patients. *Nephro-urology Monthly* 2016; 8(6): e28738.
13. Floege J, Gillespie I, Kronenberg F et al. Development and validation of a predictive mortality risk score from a European hemodialysis cohort. *Kidney Int* 2015; 87: 996-1008.
14. Collett D. Modelling survival data in medical research. 2nd ed. CRC Press, 2003.
15. Gong Q, Fang L. Comparison of different parametric proportional hazards models for interval-censored data: A simulation study. *Contemp Clin Trials* 2013; 36(1): 276-283.
16. Gauthier F, Allard M, Hétu B. Permafrost and periglacial processes. *Ice Wall Growth and Decay: Meteorological Analysis and Modelling* 2015; 26(1): 84-102.
17. Mazerolle MJ. Appendix 1: making sense out of Akaike's information criterion (AIC): its use and interpretation in model selection and inference from ecological data. 2004: 174-190. Available from: https://pdfs.semanticscholar.org/a696/9a3b5720162eaa75deec3a607a9746dae95e.pdf.
18. Akaike H. *Information theory as an extension of the maximum likelihood principle*. Petrov BN, Csaki F, (Eds.), Second International Symposium on Information Theory, Akademiai Kiado, Budapest. 1973: 267-281.
19. Burnham KP, Anderson DR. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2nd ed., Springer-Verlag. 2002.
20. Vrieze SI. Model selection and psychological theory: a discussion of the differences between the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). *Psychological Methods* 2012; 17(2): 228-243.
21. Zare A, Mahmoodi M, Mohammad K et al. Comparison between parametric and proportional hazard models in modelling transition rates of a multi-state model: application in patients with gastric cancer undergoing surgery at the Iran Cancer Institute. *Asian Pac J Cancer Prev* 2013; 14(11): 6751-5.
22. Kalbfleisch JD, Prentice RL. Rank regression and the accelerated failure time model. In Statistical Analysis of Failure Time Data, 2nd edn (eds J. D. Kalbfleisch & R. L. Prentice), 2002; 218-246. Hoboken: JD Wiley & Sons.
23. Petersen T. Fitting parametric survival models with time dependent covariates. *Applied Statistics* 1986; 35: 281-288.
24. Dudley RA, Harrell FE Jr, Smith LR *et al.* Comparison of analytic models for estimating the effect of clinical factors on the cost of coronary artery bypass graft surgery. *Journal of Clinical Epidemiology* 1993; 46: 261-271.

25. Fraser-Hunt N, Lyimo E. Insecticide-treated nets and treatment service: A trial using public and private sector channels in rural United Republic of Tanzania. *Bulletin of the World Health Organization* 1998; 76(6): 607-615.
26. Heggenhougen K, Hackethal V, Vivek P. The behavioral and social aspects of malaria and its control: An introduction and annotated bibliography. UNCP/ World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR), Geneva.
27. Liehl P, Meireles P, Albuquerque IS et al. Innate immunity induced by plasmodium liver infection inhibits malaria reinfections. *Infect Immun* 2015; 83(3): 1172-1180.
28. Wellems TE, Hayton K, Fairhurst RM. The impact of malaria parasitism: from corpuscles to communities. *The Journal of Clinical Investigation* 2009; 119(9): 2496-505.

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.201904