JURNAL BIOMETRIKA DAN KEPENDUDUKAN
(Journal of Biometrics and Population)

EVALUATING CLUSTER EFFECTS IN MALARIA SURVIVAL ANALYSIS WITH A SIMULATED EXTENDED COX MODEL

Peter Enesi Omaku[1], *Joseph Odunayo Braimah[2,3], Fabio Mathias Correa[2]

[1]Department of Mathematics and Statistics, Federal Polytechnic Nasarawa, Nasarawa State, Nigeria
[2]Department of Mathematical Statistics and Actuarial Sciences, University of the Free State, 205 Nelson Mandela Drive, Park West, Bloemfontein, South Africa
[3]Department of Mathematics and Statistics, Ambrose Alli University, Ekpoma, Edo State, Nigeria
***Corresponding author:** Joseph Odunayo Braimah ; **Email :** braimahjosephodunayo@aauekpoma.edu.ng

Published by *Fakultas Kesehatan Masyarakat Universitas Airlangga*

Keywords:
Time to event,
Risk factors,
Clusters,
Epidemiology,
Malaria

ABSTRACT

Malaria remains a significant global health challenge, particularly in tropical regions. Accurate analysis of patient survival data is essential for understanding disease progression and evaluating the effectiveness of interventions. However, traditional survival analysis often overlooks clustering effects from factors like location, healthcare or family relationship. This study examines how unshared heterogeneity in treatment regimens and reporting time affect malaria patient survival analysis. A simulated dataset, following a Weibull distribution for typical malaria treatment duration (3-7days) was generated to assess the extended Cox model's ability to handle clustering. Three cluster sizes (20, 10, 5 observations) and varying total clusters (25, 50, 100) were used to mimic a 500-patient malaria dataset from Keffi General Hospital, Nigeria, considering shared treatment similarities within clusters. Cluster effects were introduced through a normally distributed random variable. Model 2, with 10 observations per cluster, performed best based on constant hazard, low AIC, and BIC. This suggests that 50 clusters of 10 observations each effectively capture the malaria data's underlying structure. The analysis of simulated covariates revealed that male patients had 15% higher risk of death compared to females. Additionally, younger patients (0-5years), patients with blood types A, B, or AB (particularly type A), and those with increasing body temperatures were identified as high-risk groups. This study underscores the importance of considering clustering effects in analyzing malaria time-to-event data, especially for clustered datasets; a sample size of 500, divided into 50 clusters of 10 patients each, seems optimal for analyzing real-world malaria datasets using the extended Cox model.

Kata Kunci:
Waktu terjadinya,
Faktor Risiko,
Klaster,
Epidemiologi,
Malaria

ABSTRAK

Malaria masih menjadi tantangan kesehatan global yang signifikan, khususnya di wilayah tropis. Analisis yang akurat terhadap data kelangsungan hidup pasien sangat penting untuk memahami perkembangan penyakit dan mengevaluasi efektivitas intervensi. Namun, analisis kelangsungan hidup tradisional sering mengabaikan dampak pengelompokan dari faktor-faktor seperti lokasi, layanan kesehatan, atau hubungan keluarga. Studi ini mengkaji bagaimana heterogenitas yang tidak terbagi dalam rejimen pengobatan dan waktu pelaporan mempengaruhi analisis kelangsungan hidup pasien malaria. Kumpulan data simulasi, mengikuti distribusi Weibull untuk durasi pengobatan malaria pada umumnya (3-7 hari) dihasilkan untuk menilai kemampuan model Cox yang diperluas dalam menangani pengelompokan. Tiga ukuran cluster (20, 10, 5 observasi) dan total cluster yang bervariasi (25, 50, 100) digunakan untuk meniru kumpulan data 500 pasien malaria di Rumah Sakit Umum Keffi, Nigeria, dengan mempertimbangkan kesamaan pengobatan dalam cluster. Efek cluster diperkenalkan melalui variabel acak yang terdistribusi normal. Model 2, dengan 10 observasi per cluster, memiliki kinerja terbaik berdasarkan bahaya konstan, AIC rendah, dan BIC. Hal ini menunjukkan bahwa 50 cluster yang masing-masing terdiri dari 10 observasi secara efektif menangkap struktur dasar data malaria. Analisis kovariat yang disimulasikan menunjukkan bahwa pasien laki-laki memiliki risiko kematian 15% lebih tinggi dibandingkan pasien perempuan. Selain itu, pasien yang lebih muda (0-5 tahun), pasien dengan golongan darah A, B, atau AB (terutama tipe A), dan mereka yang suhu tubuhnya meningkat diidentifikasi sebagai kelompok risiko tinggi. Studi ini menggarisbawahi pentingnya mempertimbangkan efek pengelompokan dalam menganalisis data time-to-event malaria, terutama untuk kumpulan data yang mengelompok; ukuran sampel sebesar 500, dibagi menjadi 50 kelompok yang masing-masing terdiri dari 10 pasien lebih optimal untuk menganalisis kumpulan data malaria (real-world) menggunakan model Cox yang diperluas.

Received in 11 June 2024 **; Reviewed** in 01 November 2024 **; Accepted** in 02 December 2024 **; p-ISSN** 2302–707X - **e-ISSN** 2540–8828 ; **DOI:** https://doi.org/10.20473/jbk.v13i2.2024.211-218 **; Cite this as :** Omaku PE, Braimah JO, Correa FM. Evaluating Cluster Effects in Malaria Survival Analysis with A Simulated Extended Cox Model. J Biometrika dan Kependud [Internet]. 2024;13(2):211–8. Available from: https://doi.org/10.20473/jbk.v13i2.2024.211-218

INTRODUCTION

Malaria remains a serious public health problem and is endemic in Nigeria. Africa accounted for 94% (213 million cases) of all malaria cases and 94% (386,000 deaths) of all malaria deaths worldwide in 2019, according to data published by the World Health Organization (WHO) [(1)]. Nigeria accounted for 27% (61.8 million cases) of the global malaria burden and 23% (94,070 deaths) of global malaria deaths in 2019, despite the fact that malaria is preventable, treatable and curable. Also, Nigeria is one of six countries accounting for 51% of all malaria cases worldwide in 2019 (1). The most vulnerable groups in Nigeria are children under five and pregnant women. In Nigeria, malaria accounts for 30% of hospital admissions and 60% of outpatient visits. About 11% of maternal deaths and 30% of under-five deaths are caused by malaria (2). This paper will attempt to find the prevalence of malaria disease for a wider range of age categories, with children under the age of 5 years in one category. Malaria exacerbates the country's already fragile health system and places a heavy socio-economic burden on the country, reducing Gross Domestic Product (GDP) by 40% annually and resulting in direct and indirect medical costs of approximately 480 billion Naira (3).

Time of survival or failure time is a measure of how long it takes for an event to occur from a given starting point. The event is often referred to as a "failure event," although in the context of survival analysis, "failure" is a general term that does not always imply physical failure. Failure times can include things like waiting to accept a job offer, paying off a house loan, returning to criminal activity after being found guilty of a previous crime, being punished, reformed and released from prison, and moving from single to married life (4). Survival time can refer to the duration of illness remission in medicine, the time between diagnosis and death, or the time between onset and recovery. Survival data cannot be analyzed using standard statistical approaches because the underlying distribution is rarely normal (4).

Standard linear regression methods cannot be used to analyze the relationship between survival time and certain biological, socioeconomic and demographic parameters that may affect a patient's survival status because of the concept of censoring (5). If an individual's failure event occurs at an unknown time, that individual's survival time is said to be censored. This may be the case because some study participants were still alive at the time of data collection. More so, a person's survival time was unclear at the time of analysis if they were lost to follow-up; all we know about people who are censored is that it takes longer for them to fail than for them to be censored (6). Time to failure is often the response (dependent) variable in survival data, and the hazard function is often used to model survival data. The median is used as a measure of center position rather than the arithmetic mean, which can be influenced by extreme values, because survival time is often skewed (non-normal).

An important area of interest in the analysis of survival data is how prognostic factors affect the hazard function. Semi-parametric models such as the Cox proportional model and parametric models such as the Weibull model are often used (6). The shape of the baseline risks is one of the main factors influencing the choice of technique for modeling survival data. Parametric models can provide some insight into the shape of the baseline hazard if the empirical data are adequate (7). Likewise, it is possible to extrapolate survival functions, which, although theoretical, may be useful in some applications (8). Most often in medical research, the assumptions underlying parametric models may not apply to the data set because the true hazard is either unknown or complex. Say, for example, a parametric Weibull model may produce estimates that are biased and inefficient when used to analyze data from a population that do not fit a Weibull survival distribution. The Cox model provides a robust alternative and performs better than parametric Weibull model analysis in this scenario (9).

Nigeria as a developing country, data collected from health facilities for survival analysis often lack relevant variables to efficiently describe the effects of risk factors, a research challenge that can be addressed by simulating some characteristics, one of which is the effect of clustering. When observations are grouped according to a common feature or

characteristic - for example, patients belonging to the same hospital, family or region - this is referred to as clustering (10). Observations that are clustered may not be completely independent of each other because they have similar characteristics or exposures (11). Other authors suggested that if people in the same cluster have comparable survival times or event risks, this could lead to within-cluster correlation, because it can affect the assumptions and conclusions of the study, clustering is an important concept in survival analysis (12). Under the proportional hazards assumption, the value of the hazard ratio is assumed to be constant over time. The hazard ratio is an influence that can be observed when comparing two things with different conditions. Imperatively, proportional hazards assumption should be evaluated because time can often lead to changes in the hazard ratio. If these assumptions are not met, an additional technique - the extended Cox regression proposed in this paper from the Cox proportional hazard model - is required to estimate the probability of the resistance test (13).

This paper considered time clustering, where patients reporting to hospital at around the same time appear to have similar symptoms, to be on the same treatment regimen, or to be from the same household. Basically, clusters based on time periods can capture changes in malaria incidence and treatment over time, as well as the impact of interventions. The aim of this paper was to assess the Cox (ignoring clusters) and several classes of extended Cox models (with cluster specific frailty or random effect) on a partially simulated malaria data set obtained from Keffi General Hospital, time was generated from the Weibull distribution for between 3 and 7 days when one is expected to complete malaria treatment and three conditions of clustering effects that follow the normal distribution, other risk factors include; death indicator (status), age, malaria type, sex, weight, temperature and blood group which also adds as a new investigation in the context of studying this endemic diseases using the extended Cox model. This paper further assesses the impact of this disease by considering some of the risk factors using survival analysis.

METHODS

Simulation Recipe

A partial simulation was performed on the malaria data collected from the Keffi General Hospital registry, Nasarawa State, Nigeria. The Head Officer of the Registry Department of Keffi General Hospital registry, Nasarawa State, Nigeria provided informed verbal consent for the research to be conducted and the findings to be published. Failure times followed a Weibull distribution when testing the dataset for goodness of fit, and were simulated for between 3 and 7 days when treatment for malaria is expected to be completed. These times were obtained for shape and scale parameters assumed to be 2 and 4 respectively. The number of observations generated per cluster are 20, 10 and 5 using 25, 50, 100 clusters, respectively, resulting in a sample size of 500 in each situation of the total number of inpatients and outpatients treated for Malaria between January, 2022- April, 2024, this is done to under study the unshared heterogeneity in treatment regimes and reporting time amongst patients who visited the facility within the study period. The cluster effect (v) follows a normal distribution with mean zero and variance 0.5, i.e. $v \sim N(0,0.5)$.

The reminder of the data frame consists of some categorical variables: Status; coded "1" for patients who died of malaria disease and zero for those who recovered or are lost to follow-up, malaria type coded "1" for patients with severe malaria and "0" for those not severe, sex was coded "1" for male patients and "0" for female patients, blood groups O, A, B, AB were coded 0,1,2,3, respectively. Metric covariates were age, weight and temperature. Age was categorized as 1 for 1 to 5 years, 2 for $\leq$18 years and 0 for >18 years (reference category). The summary statistics for data set on malaria are displayed in Table 1.

The Cox Model

The hazard rate of the model is represented as a product, that is;

$$\lambda_i(t, X) = \lambda_0(t) \exp.\left(\sum_{j=1}^{p} \beta_j X_j\right) = \lambda \exp(X'\gamma) \quad (1)$$

where the covariates X_j are measured at study entry (t =0). A key attribute of the model is that the hazard ratio $\frac{\lambda(t,X=x)}{\lambda(t,X=0)} = exp(\beta x)$ depends on the covariates $x_1, \ldots, x_p$ but not on time t. The baseline hazard rate is unnamed and with the exponentially associated function, covariates $x = (x_1, \ldots, x_p)$ act in a multiplicatively manner on the hazard rates, γ is a vector of regression effects which is only estimable through partial likelihood estimation procedure (14).

Model Specification (Extended Cox Model)

$$\lambda_i = \lambda_0(t) \exp.(\gamma_1 type + \gamma_2 sex + \gamma_{3i} agegroup + \gamma_{4i} blood\ group + \beta_1 temperature + \beta_2 weight + v_i) \quad (2)$$

where $\lambda_0(t)$is the unnamed baseline hazard, γ_i are the categorical covariates, β_i are metrical in nature and v_i is the random effect associated with the i^{th} cluster, which allows for modeling the unobserved heterogeneity within clusters and adjusting for the potential correlation.

Proportional Hazard Assumption (PHA)

Testing the statement that the proportionate hazard condition is valid, the following hypothesis stated as cited in (15,16) is made.

H_0: $\delta_1 = \delta_2 = \cdots = \delta_p = 0$ (Statement is valid)

H_1: $at\ least\ one\ pair\ of\ the\ \delta_i's\ is\ not\ equal\ to\ zero$ (Statement violated)

Decision rule: Reject H_0 if $p - value \leq \alpha$.

Schoenfeld residual measures are used to examine the violation of the constant hazard assumption. The technique is classically computed at each failure time under the comparative hazard postulate, and is typically not defined for non-informative observations (17). The general significance test is called the global test. The criteria used to select the best modeling strategy for the malaria data were; p-value, i.e, the higher the value above the threshold of 0.05 the better, Akiake Information Criteria (AIC) and Bayesian Information Criteria (BIC); in both cases, the smaller the values, the better the model.

Models

$$\lambda_{ICE} = \lambda_0(t) \exp.(\gamma_1 type + \gamma_2 sex + \gamma_{3i} agegroup + \gamma_{4i} blood\ group + \beta_1 temperature + \beta_2 weight) \quad (3)$$

$$\lambda_{WCE_i} = \lambda_0(t) \exp.(\gamma_1 type + \gamma_2 sex + \gamma_{3i} agegroup + \gamma_{4i} blood\ group + \beta_1 temperature + \beta_2 weight + v_i) \quad (4)$$

$WCE_i = 1, 2, 3\ for\ 25, 50, 100$ Clusters respectively

where ICE = Ignoring Cluster and WCE = with cluster effects for 25, 50 and 100 cluster numbers.
Model 1 = model with 20 patients within a cluster, with total clusters of 25
Model 2 = model with 10 patients within a cluster, with total clusters of 50
Model 3= model with 5 patients within a cluster, with total clusters of 100

RESULTS

The study made use of statistical software (R4.3.) to generate the results as presented in Table 1 and 2.

Table 1: Descriptive Statistics for Categorical and Metrical covariates

Variables	Covariate level	Condition	
		Alive	Dead
Malaria Type	0	222 (93.3%)	16 (6.7%)
	1	241(90.8%)	21(9.2%)
Sex	0	258 (92.7 %)	21 (7.5 %)
	1	202(91.4%)	19(8.6%)
Age group	0	125(94.7%)	7(5.3%)
	1	184(88.9%)	23(11.1%)
	2	151(93.8%)	10 (6.2 %)
Blood group	0	116(98.3%)	2(1.7%)
	1	120(88.9 %)	15(11.1%)
	2	150(91.5%)	14 (8.5%)
	3	74(89.2%)	9(10.8%)
Temperature	Min/Max	32/39	32/39
Weight	Min/Max	3/80	7/78

Table 2: Table of hazard ratios (with reference category "0" as in table 1) and model selection criteria

Covariates (Reference "0")	ICE	Model 1	Model 2	Model 3	Model2 coeff.
Malaria type	1.006804	1.006230	1.040947	0.99486	0.040131
Sex	1.063870	1.063555	1.155101	1.06236	0.144187
Age group1	3.637784	3.641018	3.736388	3.73486	1.318119
Age group 2	1.213944	1.215251	1.185672	1.22542	0.170310
Blood group1	5.033967	5.030504	5.726482	5.04158	1.745101
Blood group2	3.385701	3.385575	3.809125	3.36940	1.337400
Blood group3	2.994272	2.994760	2.990308	2.97338	1.095376
Temperature	2.627788	2.628384	2.649645	2.63650	0.974426
Weight	0.989088	0.989106	0.995148	0.98963	-0.004864
p value for PHA	**0.0490**	**0.0477**	**0.0553**	**0.0729**	
AIC	**410.6053**	**412.6052**	**406.3768**	**412.555**	
BIC	**425.8052**	**429.494**	**423.2656**	**429.4438**	

DISCUSSIONS

The constant hazard assumption holds for model 2 and model 3, while it fails for ICE and model 1 as can be seen in Table 1. Apparently, it can be inferred that in trying to solve the problem of non-constant hazard as suggested by a research findings in 2020 (13) in Cox ICE, the model only got worse in model 1, which may be due to misspecification of the cluster numbers - too many subjects within the clusters. When the models were further evaluated using the AIC and BIC values, it was seen that model 2 outperformed the others with the lowest values for both selection criteria which reflects the submission of a case study in 2024 (12) on Cluster Analysis Integration Model with Survival Analysis for Late Payment of House Ownership Loan. Model 3 may appear to be better in terms of proportional hazard, but failed to capture the model effects within clusters with too few members.

The best model (model 2) fitted is shown below:

$$\lambda_{WCE_{=50}} = \lambda_0(t)\exp.(0.040131.type + 0.144187.sex + 1.318119.agegroup1 + 0.170310.agegroup2 + 1.745101.bloodgroup1 + 1.337400.bloodgroup2 + 1.095376.bloodgroup3 + \gamma_{4i}blood\ group + 0.974426.temperature - 0.004864.weight + 0.404537.v_i)$$

Malaria type has an increasing effect of 0.040131 and a hazard ratio of 1.041 for patients with the effect of cluster-specific frailty, suggesting that patients with severe malaria have a slightly 4% higher risk of dying from malaria compared to non-severe cases. Sex has an increasing effect of 0.144 and the hazard ratio of 1.155, for patients with the effect of cluster-specific frailty, suggesting that male patients are 15% at risk of death from malaria compared to their female counterparts.

Age group 1 has an increasing effect of 1.32 and the hazard ratio of 3.74 for patients with the effect of cluster specific frailty, suggesting that patients within the age group 0-5 years are 3.74 times more likely to die from malaria than those within the age group over 18 years. Age group 2 has an increasing effect of 0.17 and the hazard ratio is 1.19, for patients with the effect of cluster specific frailty, suggesting that patients in the age group 6-19 years are 1.19 times or 19% more at risk of dying from malaria than those in the age group over 19 years, this is consistent with the findings of national malaria elimination program organized by the National Population Commission, National Bureau of Statistics and ICF International on Nigeria Malaria Indicator Survey held in Abuja, Nigeria (2).

Blood group 1 has an increasing effect of 1.75 and a hazard ratio of 5.73 for patients with the effect of cluster specific frailty, suggesting that patients with blood group A are 5.73 times more likely to die from malaria than those with blood group O. Blood group 2 has an increasing effect of 1.34 and a hazard ratio of 3.81 for patients with cluster specific frailty, suggesting that patients with blood group B are 3.81 times more likely to die from malaria than those with blood group O. Blood group 3 has an increasing effect of 1.0954 and the hazard ratio of 2.99, for patients with the effect of cluster specific frailty, suggesting that patients with blood group "AB" are 2.99 times more at risk of dying from malaria than those with blood group "O". Although empirical study comparing the hazard or risk posed by specific blood group as considered in this study are rarely seen in literature. However, the assessment of blood groups in this study mirrors the laboratory research of Bertrand et al. (18) which suggested that the severity of malaria is partly determined by the presence of blood group A. Temperature has an increasing effect of 0.97 and a hazard rate of 2.65 for patients with the effect of cluster-specific frailty, suggesting that for a unit change in patient temperature, the baseline is associated with 2.65 times the risk of death from the disease. Weight has a decreasing effect of -0.0049 and a hazard rate of 0.995 for patients with the effect of cluster-specific frailty, suggesting that for a unit change in patient weight, the baseline as a result of this slight reduction in weight is associated with 0.995 times the risk of death from the diseases. This finding aligns with Oldenburg et al. (19), who reported weight gain after days of treatment.

Advantages of the Study

This study enhances the relevance of the research to the local context by leveraging data from General Hospital Keffi. This improved accuracy leads to more precise estimates of malaria risk factors and a deeper understanding of the factors influencing malaria survival. Additionally, the inclusion of blood group as a covariate provides valuable insights into the varying degrees of vulnerability among different blood groups. The study ensures correct resource allocation to the hospital facility by appropriately clustering datasets. Ultimately, this research contributes to the development of more targeted and effective treatment and control strategies for similar settings and populations.

Limitations of the Study

Potential inaccuracies in record-keeping and data collection may have introduced some degree of error into the study. Additionally, the reliance on simulated data may limit the extent to which the findings can be directly applied to real-world scenarios.

CONCLUSIONS AND SUGGESTIONS

Conclusions

The study reveals that not only is it problematic to fit Cox regression to the data structure in most cases due to non-constant hazards, but misspecifying the number of cluster random effects and the correct number of homogeneous sets within each group in the

hazard structure also has a detrimental effect. Here, the malaria data is best represented by 10 patients with similar conditions and treatment regimens being treated by their doctors at a given time among the 50 blocks of clusters in a sample of 500 patients in the health facility. Patients with the covariate set; age group 1; (0-5 years), blood groups A, B and AB, especially those of group "A," coupled with an increasing temperature were seen to be the most vulnerable and at great risk of death from malaria - which is consistent with the submissions of (1, 2 and 3). Owing to modeling; it is clearly observed that data cleaning by making up for the inadequacy of data structures obtained from record rooms in this part of the world is paramount to effectively capture the effects of the model.

Suggestions

Adequate time should be allowed to collect all relevant information from patients to allow effective measurement of risk trends over time, which may help to advise patients appropriately and review treatment strategy over time. Due to model effects, preventive and proactive measures should be taken to reduce the number of deaths from this endemic disease, especially for children under five years of age. Some of these could be: the use of mosquito nets, promoting and maintaining a clean and healthy environment, seeking medical attention in case of rising temperature, and knowing the blood group to have a good idea of the risk of malaria disease if the child is exposed.

ACKNOWLEDGMENT

The authors are grateful to the Keffi General Hospital, particularly the registry, for providing the data used in this study. Dr Braimah, Joseph Odunayo is grateful to the University of the Free State, Bloemfontein for the opportunity to conduct this study as a Postdoctoral Researcher.

REFERENCES

1. World Health Organization (WHO). World Malaria Report 2020: 20 Years of Global Progress and Challenges [Internet]. Geneva: World Health Organization; 2020. Available from: https://www.who.int/publications/i/item/9789240015791
2. National Malaria Elimination Programme (NMEP), National Population Commission (NPopC), Bureau N, of Statistics (NBS), ICF International. Nigeria Malaria Indicator Survey 2015 [Internet]. Abuja, Nigeria, and Rockville, Maryland, USA: NMEP, NPopC, and ICF International; 2016. Available from: https://dhsprogram.com/pubs/pdf/mis20/mis20.pdf
3. National Population Commission (NPC) [Nigeria], ICF. Nigeria Demographic and Health Survey 2018 [Internet]. Abuja, Nigeria, and Rockville, Maryland, USA: NPC and ICF; 2019. Available from: https://dhsprogram.com/pubs/pdf/FR359/FR359.pdf
4. Chilot D, Mondelaers A, Alem AZ, Asres MS, Yimer MA, Toni AT, et al. Pooled Prevalence and Risk Factors of Malaria Among Children Aged 6-59 Months in 13 Sub-saharan African Countries: a Multilevel Analysis Using Recent Malaria Indicator Surveys. PLoS One [Internet]. 2023;18(5):e0285265. Available from: https://doi.org/10.1371/journal.pone.0285265
5. Maji A, Mohan A, Choudhari TP, Shailendra SK. Hazard-based Overtaking Duration Model for Mixed Traffic. Transp Lett [Internet]. 2023;16(7):715–724. Available from: https://doi.org/10.1080/19427867.2023.2236409
6. Cox DR. Regression Models and Life-Tables. J R Stat Soc Ser B [Internet]. 1972 Jan 1;34(2):187–202. Available from: https://doi.org/10.1111/j.2517-6161.1972.tb00899.x
7. Mbona SV, Mwambi H, Ramroop S, Chifurira R. Survival Analysis of Patients with Multidrug-Resistant Tuberculosis in Kwazulu-Natal, South Africa : A Comparison of Cox Regression and Parametric Models. Int J Sci Math Technol Learn [Internet].

2024;31(1):571–81. Available from: https://hdl.handle.net/10321/5467

8. Dai R, Ma J, Wu M, Mai Y, He W. A Flexible Ensemble Learning Method for Survival Extrapolation. Ther Innov Regul Sci [Internet]. 2023;57(3):580–8. Available from: https://doi.org/10.1007/s43441-022-00490-1
9. Zuo W, Li Y. A New Stochastic Restricted Liu Estimator for the Logistic Regression Model. Open J Stat [Internet]. 2018;8(1):25–37. Available from: https://doi.org/10.4236/ojs.2018.81003
10. Liang J, Chen S, Kochunov P, Hong LE, Chen C. Integrative Data Analysis Where Partial Covariates Have Complex Non-linear Effects by Using Summary Information from an External Data. Am Stat [Internet]. 2024;2:1–22. Available from: https://doi.org/10.48550/arXiv.2303.03497
11. Rubio FJ, Drikvandi R. MEGH: A parametric class of general hazard models for clustered survival data. Stat Methods Med Res [Internet]. 2022;31(8):1603–1616. Available from: https://doi.org/10.1177/09622802221102620
12. Zahra A, A. R. Fernandes A, B. Astuti A. Cluster Analysis Integration Model with Survival Analysis for Late Payment of House Ownership Loan (Case Study: House Ownership Loan Bank X Customer). J Stat Appl Probab [Internet]. 2024;13(1):15–26. Available from: https://dx.doi.org/10.18576/jsap/130102
13. Jong VMT de, Moons KGM, Riley RD, Smith CT, Marson AG, Eijkemans MJC, et al. Individual Participant Data Meta-analysis of Intervention Studies with Time-to-event Outcomes: A review of the Methodology and an Applied Example. Res Synth Methods [Internet]. 2020;11(2):148–168. Available from: https://doi.org/10.1002/jrsm.1384
14. O.I. A, A.A. A. First Birth Interval: Cox Regression Model with Time Varying Covariates. Covenant J Phys Life Sci [Internet]. 2018 Jun 20;0(0 SE-Articles). Available from: https://journals.covenantuniversity.edu.ng/index.php/cjpls/article/view/930
15. Zhang Z, Reinikainen J, Adeleke KA, Pieterse ME, Groothuis-Oudshoorn CGM. Time-Varying Covariates and Coefficients in Cox regression Models. Ann Transl Med [Internet]. 2018;6(7):121. Available from: https://doi.org/10.21037/atm.2018.02.12
16. Turkson AJ, Addor JA, Ayiah-Mensah F. The Cox Proportional Hazard Regression Model Vis-à-Vis ITN-Factor Impact on Mortality Due to Malaria. Open J Stat [Internet]. 2021;11:931–62. Available from: https://doi.org/10.4236/ojs.2021.116055
17. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika [Internet]. 1982 Apr 1;69(1):239–41. Available from: https://doi.org/10.1093/biomet/69.1.239
18. Lell B, May J, Schmidt-Ott RJ, Lehman LG, Luckner D, Greve B, et al. The Role of Red Blood Cell Polymorphisms in Resistance and Susceptibility to Malaria. Clin Infect Dis [Internet]. 1999 Apr 1;28(4):794–9. Available from: https://doi.org/10.1086/515193
19. Oldenburg CE, Guerin PJ, Berthé F, Grais RF, Isanaka S. Malaria and Nutritional Status Among Children With Severe Acute Malnutrition in Niger: A Prospective Cohort Study. Clin Infect Dis [Internet]. 2018 Sep 14;67(7):1027–34. Available from: https://doi.org/10.1093/cid/ciy207