

Joint Multilevel Model for Analyzing Length of Stay through Competing Endpoints in Dengue Epidemiology

G.H.S. Karunarithna^{*}, M.R. Sooriyarachchi

Department of Statistics, Faculty of Science, University of Colombo, Sri
Lanka

^{*} hasani@stat.cmb.ac.lk

Received: 12th February 2019 / Revised: 27th May 2019 / Accepted: 30th May 2019
©IAppStat-SL2018

ABSTRACT

Dengue is a common mosquito-borne tropical disease caused by a virus. It is a life threatening disease since it sometimes leads to death within a short period of time. Multilevel modeling is a form of statistical modeling when data is at different levels. Due to dengue seriousness and risk being more similar for patients within districts than between districts, there is correlation between patients within districts. Thus district has to be taken as a cluster variable. A frequently encountered response in epidemiological studies is the length of Stay (LOS) of a patient, which measures the time until the event of interest occurs. Complexity arises with the different states/destinations of the time event and competing risk modelling is a more appropriate method for handling such states. The association of platelet count and length of stay of a dengue patient leads to the joint modelling approach for analyzing the dengue patients. Formulation criteria for the joint model with clustered data is to link these models through two sub models that is by using the multilevel multinomial logistic model for the LOS of dengue patients with different destinations and multilevel continuous model for the log platelet count. The linkage between two responses was derived by sharing a common random effect. Factors that have an effect on different destinations of LOS are, time indicators, year, age, classification, rainfall, temperature and humidity, while age, sex, classification, year place treated, rainfall, temp and humidity are associated factors for the log platelet count of dengue patients. Moreover, supremacy of joint modelling was proved by the AIC and BIC values over two separate univariate models.

Keywords: Dengue, Competing risk, Multinomial logistic, Joint modelling.

1. Introduction

Dengue has become endemic in the World Health Organization (WHO) regions of Africa, the Americas, the Eastern Mediterranean, South-East Asia and Western Pacific in recent decades (WHO, 2016). Also dengue is a life threatening disease, since it can lead to death of a patient within a short period of time. Dengue Fever (DF) and Dengue Hemorrhagic Fever (DHF) has become a prominent cause of hospitalization and death. Lately, the occurrence of DF and DHF patients' have been rapidly spreading in Sri Lanka. In 2017, Colombo district showed rapid increment of dengue cases nearly with 34,000 in Sri Lanka (Disease Surveillance, 2018).

Length of stay is an intractive tool for assessing dengue patients as it is a result of patient characteristics, hospital characteristics and social characteristics (Sa, Dismuke, and Guimaares, 2007). Also, dengue patients are influenced by the characteristics of the patients as well as their living surrounding with the spatial variation of the rainfall, temperature, humidity and degree of urbanization. Therefore, patients/individuals are hierachchally structured within their living surrounding such as districts. This means that the patients within the same district are more alike than patients between districts. This leads to multilevel modeling. Clustering effects are arising due to the correlation between the same levels nested within the higher level in such situation. Hence, there are different random effects between levels as a result of this clustering. There is a lot of literature on selecting district as a clustering variable for dengue. Wickremasuriya and Sooriyarachchi (2014) have shown that the seriousness of dengue is more correlated within districts than between districts. Jayanetti and Sooriyarachchi (2015) has indicated that dengue survival is more correlated within districts than between districts. Fernando and Sooriyarachchi(2018) have shown that dengue counts in Sri Lanka are related to the climate and so are correlated within districts as the weather is similar within districts.

Length of stay (LOS) of a patient is influenced by various endpoints/failure event like discharge, dead and transferred. This leads to competing risk modeling. It is used when an individual is under risk of multiple events/different destinations (Prentice, et al., 1978; Haller, 2014).

The shared parameter model is called a joint multilevel/hierachchical or random effect model, which is widely used for joining two dependent outcomes under a clustering nature. Here the random effect is introduced and shared with dependent outcomes to account for the association between these dependent

outcomes. On the other hand, complexity arises when one outcome is a competing risk event with hierarchical nature. This is because it contains multiple events as well as two correlations that is correlation between responses for each cluster and correlation between individuals within the same cluster (Gueorguieva and Agresti, 2001).

The relation between LOS and platelet count of a dengue patient is motivated by the study in past literature in order to evaluate the dengue patients (Pooransingh, Teeluksingh, and Dialsingh, 2016; Karunarathna and Sooriyarachchi, 2017). The joint association was built through the competing risk event response for the LOS of dengue patients and continuous response for platelet count after taking the logarithm of that variable. Although the failure events more often are in continuous form, here, LOS of a dengue patient is recorded in discrete form. As a result, the approach undertaken here is to analyze the LOS via discrete time competing event model. The data was obtained from the Epidemiological unit which had records of dengue patients reported in 2006-2008. In this study districts with high dengue cases (10 districts) were used to aid model convergence. The climate data was obtained from Meteorological Department of Sri Lanka.

Logistic modeling for ordered categories and multinomial logistic modeling for nominal data are the most prominent method for modeling multilevel univariate discrete competing risk models (Agresti, 1984; McCullagh, 1980; Prentice and Gloeckler, 1978; Gerhard, 1995). However joint modeling approaches for competing risk are well known with the continuous time competing risk by considering the cause specific hazard or subdistribution hazard models. Research on joint modeling of longitudinal and continuous competing risk is a field of interest in recent decades for formulating the joint association of competing risk events by sharing the random effect among responses through two marginal sub models of mixed effect models for longitudinal measurement and cause specific hazard for the competing event (Elashoff, Li, and Li, 2007; Elashoff, Li, and Li 2008; Deslandes and Chevret, 2010; Gould, et al., 2015; Gueorguieva, Rosenheck, and Lin, 2012; Williamson, et.al., 2008). The approach undertaken here is multinomial logit modelling which is proposed by Steele, Diamond, & Wang (1996) for modeling discrete time competing risk events since this full multinomial model permits a clear specification of the correlation structure between the different outcomes (Steele, Diamond, & Wang, 1996). The novel method is formulated in here because there was no work found for the joint modeling of the discrete competing risk event by multinomial logit with any other response under the clustering nature in the literature.

The rest of the paper is organised as follows. The methodological aspects which is used to analyze this situation is discussed under the ‘‘Methodology’’ section. The next section discusses the results of the analysis of the dengue application. The last section gives some concluding remarks.

2. Methodology (in 12pt, bold; Below paragraph 6pt, Contents in 11pt)

The methodology to formulate the joint model to clustered data is carried out via two linked sub models of multilevel multinomial logistic model for the LOS of a dengue patient with a different destination and multilevel continuous model for the platelet count. This situation arising in different families of distributions such as multinomial distribution and continuous (normal) distribution. The linkage between the two dependent variables are modeled through the association between random effects.

2.1 Separate Sub Models

Assume that the durations are measured in discrete time intervals indexed by t ($t=1, \dots, T$). For each discrete time interval t for the j^{th} individual in cluster i is symbolized as $Y_{1ij}^{(r)}$ that is the multinomial outcome which denotes whether an event r has occur during the interval t or not for the j^{th} individual in i^{th} cluster. The second response is symbolized as Y_{2ij} (log of platelet count variable) which is a normally distributed outcome in the time interval t for the j^{th} individual in the i^{th} cluster.

2.1.1 Multilevel Multinomial Sub Model

The structural formulation of the model for the discrete competing risk event of $Y_{1ij}^{(r)}$ is given by,

$$Y_{1ij}^{(r)} = \begin{cases} 1; & \text{if event of type } r \text{ is in the interval } t \\ 0; & \text{if no event in } t \end{cases} \quad (1)$$

; where, $r = 1, 2, \dots, S$

Let $h_{ij}^{(r)}$ be the hazard of an event of type r in interval t with corresponding vector containing the probabilities of falling into any particular category. Then, the log odds of experiencing an event of type r relative to an event of types s at time t for the j^{th} individual in the i^{th} cluster is given by,

$$\log \left(\frac{h_{ij}^{(r)}}{h_{ij}^{(s)}} \right) = \beta_0^{(r)} + \sum_{t=1}^t \alpha_t^{(r)} D_{tij}^{(r)} + \beta_1^{(r)} X_{tij}^{(r)} + u_{0j}^{(r)} \quad (2)$$

where, $r = 1, 2, \dots, s - 1$

$$\left(u_{0j}^{(1)}, u_{0j}^{(2)}, \dots, u_{0j}^{(r)}\right) \sim \text{Multivariate Normal}$$

The model can be regarded as a combination of fixed parts and a random part, where, $\beta_0^{(r)}$ and $\beta_1^{(r)}$ are the parameters in the fixed part of the model and $u_{0j}^{(r)}$ are the parameters in the random part. $D_{ij}^{(r)}$ are the indicators for the time interval and $X_{ij}^{(r)}$ are the covariates.

The equation (2) represents the probabilities of being in a particular category at time interval t in the random effects hazards model;

$$h_{tij}^{(r)} = \frac{\exp(\beta_0^{(r)} + \sum_{t=1}^t \alpha_t^{(r)} D_{tij}^{(r)} + \beta_1^{(r)} X_{tij}^{(r)} + u_{0j}^{(r)})}{1 + \sum_{k=1}^{s-1} \exp(\beta_0^{(k)} + \sum_{t=1}^t \alpha_t^{(k)} D_{tij}^{(k)} + \beta_1^{(k)} X_{tij}^{(k)} + u_{0j}^{(k)})} \quad (3)$$

; where, $r = 1, 2, \dots, S - 1$

And

$$h_{tij}^{(s)} = \frac{1}{1 + \sum_{k=1}^{s-1} \exp(\beta_0^{(k)} + \sum_{t=1}^t \alpha_t^{(k)} D_{tij}^{(k)} + \beta_1^{(k)} X_{tij}^{(k)} + u_{0j}^{(k)})} \quad (4)$$

2.1.2 Multilevel Continuous Sub Model

Multilevel continuous sub model can be represented as,

$$Y_{2ij} = \beta_1 + \beta_2' X_{tij} + \varepsilon_{ij} + u_{1j} \quad (5)$$

Here, β_1 is the random intercept, β_2' is the fixed regression coefficient, ε_{ij} is the random error for the individual which is normally distributed with mean 0 and variance σ_ε^2 . Here u_{1j} is a normally distributed random effect for the cluster with zero mean Joint Models

Consider two responses, the competing risk event and a continuous event for the joint model as $\mathbf{Y}_{1tij}^{(r)}$ and \mathbf{Y}_{2ij} denoting the j^{th} individual in the i^{th} cluster for the r^{th} event types of competing event and the continuous outcome, respectively. The joint model is built by describing the joint density $f(\mathbf{Y}_{1tij}^{(r)}, \mathbf{Y}_{2ij})$ of the competing risk event $\mathbf{Y}_{1tij}^{(r)}$ and the continuous vector \mathbf{Y}_{2ij} . Joint models are linked through the random effects.

For simplicity, assume that the random effects ($u_{0j}^{(r)}$) are associated with the i^{th} cluster for the event type r , is the same for each event type. Therefore equation (2) is rewritten as,

$$\log\left(\frac{h_{tij}^{(r)}}{h_{tij}^{(s)}}\right) = \beta_0^{(r)} + \sum_{t=1}^t \alpha_t^{(r)} D_{tij}^{(r)} + \beta_1^{(r)} X_{tij}^{(r)} + u_{0j} \quad (6)$$

; where, $u_{0j} \sim N(0, \sigma^2)$

A joint model is obtained by assuming a distribution for both sets of random effects u_{0j} and u_{1j} jointly. However, in here the authors are interested in considering the shared random intercept model with a scale parameter λ . Therefore, equation (5) has become as,

$$Y_{2ij} = \beta_1 + \beta_2' X_{tij} + \varepsilon_{ij} + \lambda u_{0j} \quad (7)$$

Also, it assumes that, conditional on u_{0j} , the outcome vectors $Y_{1tij}^{(r)}$ and Y_{2ij} are independent, that is the association between two outcome vectors are assumed to be completely captured by the association between the random effects (Ivanova, Molenberghs, and Verbeke, 2015; Elashoff, Li, and Li, 2007).

2.2 Estimation and Inference

Assuming the independence of $Y_{1tij}^{(r)}$ and Y_{2ij} , conditional on u_{0j} , the corresponding likelihood function to the joint model is given by,

$$L(\theta) = \prod_{j=1}^J \prod_{n=1}^N \int f_{1tij}(Y_{1tij} | u_{0j}) f_{2ij}(Y_{2ij} | u_{0j}) f(u_{0j}) du_j \quad (8)$$

Where, θ is the vector containing all parameters in the conditional distribution and normally distributed random effect u_{0j} .

where,

$$f_{1tij}(Y_{1tij} | u_{0j}) = \prod_{t=1}^T \left[\prod_{r=1}^S h_{tij}^{(r)} \right]^{\Delta} * \exp \left\{ - \int_0^T \left[\sum_{r=1}^S h_{tij}^{(r)} \right] dt \right\} \quad (9)$$

; where Δ = censoring indicator

and $h_{tij}^{(r)}$ can be derived from equation (3) to be,

$$f_{2ij}(Y_{2ij}|u_{0j}) = \frac{1}{\sqrt{2\pi\sigma_{\epsilon}^2}} \exp\left\{-\frac{1}{2\sigma_{\epsilon}^2}(Y_{2ij} - \beta_1 - \beta_2'X_{tij} - \lambda u_{0j})^2\right\} \quad (10)$$

$$f(u_{0j}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{u_{0j}^2}{2\sigma^2}\right) \quad (11)$$

Then, the likelihood function can be derived from equation (8) by substituting the values from equation (9) – (11).

In this paper, standard numerical integration method of adaptive Gaussian quadrature, has been implemented in the SAS procedure Proc NLMIXED (Pinherio and Bates, 1995) due to the complexity of the integral part of the equation (8).

3. Analysis, Results and Discussion

The data is on dengue from 2006 – 2008. For this period monthly climate data has been used. District is used as a clustering factor. Only data from 10 high incidence districts have been considered. The districts studied consists of Colombo, Galle, Gampaha, Kalutara, Kandy, Kegalle, Kurunegala, Matara, Puttlam and Ratnapura. In this study, discrete time approach is appropriate as the duration is recorded as whole days. Therefore, it can be assumed that duration reflects a discrete time measurement. Thus, initially, the time event; length of stay is categorized into three categories; 0-4 days, 4-6 days and 6-10 days which are known as febrile phase, critical phase and recovery phase respectively according to the clinical manifestation of dengue (Yacoub, et.al., 2014). Hence, each individual studied within 10 days and the recorded destination such as discharge, transfer or dead for each individual within the time period was used. If any individuals do not show any event up to 10 days, they are considered as censored observations.

As the next step, data need to be transformed to carry out discrete time competing risk modelling from the standard one-person, one-record data set (the person-data set) in to a one person, multiple period data set (the person-period data set) (Singer and Willett, 1993). That is indicator variables need to be introduced to the time variable and other categorical variables. Table 1 gives a list of variables used for this analysis. Also, reference category for the categorical variables is highlighted in the Table 1.

Table 1: Description of Data

Variable	Category	Notation	Code
LOS	0-4	T1	1
	4-6	T2	2
	6-10	T3	3
	>10	T4	4
Event	Censored		0
	Discharged		1
	Transferred		2
	Dead		3
Year	2006	Yr2006	1
	2007	Yr2007	2
	2008	Yr2008	3
Month	Jan	Month1	1
	Feb	Month2	2
	Mar	Month3	3
	Apr	Month4	4
	May	Month5	5
	June	Month6	6
	July	Month7	7
	Aug	Month8	8
	Sep	Month9	9
	Oct	Month10	10
	Nov	Month11	11
	Dec	Month12	12
Age			
Sex	Male	Sex1	1
	Female	Sex2	2
Ethnicity	Sinhala	Ethnicity1	1
	Tamil	Ethnicity2	2
	Moor	Ethnicity3	3
	other	Ethnicity4	4
Place Treated	Government	PlaceTreated1	1
	Private	PlaceTreated2	2
	other	PlaceTreated3	3
Classification	DF	Classification1	1

	DHF1	Classification2	2
	DHF2	Classification3	3
Packed cell volume		PCV	
Logarithm of Platelet count		Log_Platelet	
Rainfall		Rainfall	
Rainfalllag1		Rainfalllag1	
Rainfalllag2		Rainfalllag2	
Temperature		Temp	
Temperature_lag1		Templag1	
Temperature_lag2		Templag2	
Humidity		Humid	
Humidity lag1		Humidlag1	
Humidity lag2		Humidlag2	

Initial values for formulating the joint association between LOS of a dengue patients and log of platelet count were obtained by fitting the univariate models without considering the random effects in the model. The different destination such as discharge, transfer and dead, is a competing risk event and a nominal category and multinomial logistic regression model was fitted by using “PROC LOGISTIC” procedure under the “glogit” function in SAS. Three different models were formulated from the multinomial logit model as discharge vs censored, transfer vs censored and dead vs censored. Out of the 19 explanatory variables, only 8 variables were significant under the chi squared test at 5% level of significance. These were time indicators for the LOS, year, age, classification, rainfall, humid, humidlag1 and temp. The other response is log platelet and it is a continuous variable. Univariate model for the log platelet was fitted through “PROC Genmod” procedure in SAS. Eight variables which are year, month, age, sex, place treated, classification,, humid and temp were significant for log platelet response.

Then, the joint model was initiated through “PROC NLMIXED” procedure, with the significant variables from the univariate models. When formulating the joint model, the shared random intercept between responses with an inflation factor (λ) is used to account for the different scale at which these are measured (Ivanova, Molenberghs, and Verbeke, 2015)

Table 2 presents parameter estimates of the joint model. This table shows that standard deviation of the random effect (district clustering effect (RI.sd)), residual (Res.Sd) and scale parameter (λ).

Table 2: Parameter estimates for the joint model

Response		Variable	Coefficient	P value	
Competing risk	Discharge	Intercept	0.4405	<0.0001	
		T1	-0.7842	0.002	
		T2	-0.8727	0.012	
		Age	-0.00368	0.0002	
		Classification3	-0.03641	<0.0001	
		Rainfall	-0.005	<0.0001	
		Humid	-0.02247	<0.0001	
		Temp	0.008773	<0.0001	
	Transfer	Intercept	2.2512	<0.0001	
		T1	-1.0513	0.0003	
		T2	0.803	<0.0001	
		Age	0.0228	0.0005	
		Temp	-0.3844	<0.0001	
	Dead	Intercept	-16.9346	<0.0001	
		T1	-0.7969	0.024	
		T2	0.9909	0.035	
		Yr2007	2.1742	<0.0001	
		Rainfall	2.0058	0.0023	
		Temp	0.3565	<0.0001	
		RI.sd		10.1404	<0.0001
	Log_platelet	Intercept		10.3194	0.00817
Yr2007			0.1806	<0.0001	
Yr2008			0.1876	<0.0001	
Month10			-0.08258	0.0057	
Month12			-0.2108	<0.0001	
Age			-0.1007	<0.0001	
Sex1			-0.1007	<0.0001	
Placetreated2			0.1831	<0.0001	
Placetreated3			0.27	0.0237	
Classification2			-0.1163	0.0025	
Classification3			-0.2982	<0.0001	

	Humid	0.01374	<0.0001
	Temp	-0.05368	<0.0001
	scale (λ)	0.4599	<0.0001
	Res. sd(σ_e)	0.8488	<0.0001

As per the parameter estimates obtained in Table 2, the marginal models for competing risk event (multilevel multinomial logit with discrete time hazard) and the log platelet count (multilevel normally distributed model) can be written as,

$$\log(p_{discharge}/P_{censored}) = 0.44 - 0.7842 * T1 - 0.8727 * T2 - 0.00368 * Age - 0.03641 * classification3 - 0.0050 * Rainfall - 0.002247 * Humid + 0.008773 * Temp \quad (12)$$

$$\log(p_{transfer}/P_{censored}) = 2.2512 - 1.0513 * T1 + 0.803 * T2 + 0.0228 * Age - 0.3844 * Temp \quad (13)$$

$$\log(p_{dead}/P_{censored}) = -16.9346 - 0.7969 * T1 + 0.9909 * T2 + 2.1742 * yr2007 + 0.0058 * Rainfall + 0.3565 * Temp \quad (14)$$

$$\log_{platelet} = 10.3194 + 0.1806 * yr2007 + 0.1876 * yr2008 - 0.08258 * Month10 - 0.2108 * Month12 - 0.1007 * Age - 0.1007 * Sex1 + 0.1831 * Placetreated2 + 0.27 * Placetreated3 - 0.1163 * Classification2 - 0.2982 * Classification3 + 0.01374 * Humid - 0.05368 * Temp \quad (15)$$

The equation (12)-(14) with equation (15), emphasizes that month, sex and place treated is only associated with the log of platelet count. Age is negatively associated with the discharge patients while positively associated with transfer patients. Temperature is related with all three destination of discharge, transfer and dead. The risk of death among the patients who were in febrile phase is 0.451 times lower than the patients who were in recovery phase, while 2.69 times higher in critical phase with respect to the recovery phase. For Discharge a patient at the febrile phase is 0.47 times lower than the recovery phase and discharging a patients at the critical phase is 0.418 lower than those at the recovery phase. The transferring patient at the febrile phase higher by 0.35 times lower compare to the recovery phase, while transferring patients who are in critical phase is higher by amount of 2.23 times with respect to the recovery phase. Temperature, rainfall and humidity were associated as geographical factors for the discharge patients. However rainfall and temperature were related to the death of dengue patients, temperature is only geographical factor among the transferred patients. Rainfall is highly associated to the dead patients when compared with discharged patients. In year 2007, the risk of death of a dengue patients are 8.8 times higher than year 2008.

Platelet count is related with October and December only among the months. Also, it indicates that male patients are having higher platelet count than the female patients. DHF1 and DHF2 patients show lesser platelet counts than the DF patients, while lowest values is having for the DHF2. So, it implies that the risk is higher for the patients who are classified as DHF2 for the dengue patients.

Another important result is that the random effect related to the district and the parameter λ are both highly significant, indicating that taking the district as a cluster variable is justified.

3.1 Checking for Supremacy of the Joint Model

It is important to compare the performance of the joint model with the two separate univariate models to express the necessity and performance of the joint model. Summation of the model fit statistics of the two univariate models was used to compare with the model fit statistics of the joint model to check the advantage of using the joint model over the univariate models. This is given in table3.

Table 3: Goodness of Fit test for Proposed Model

Model fit statistics	Joint model	Univariate models		Total
		Competing Risk-Multilevel Multinomial logistic model	Multilevel normal regression	
-2 Log Likelihood	34957	15315.23	27082.56	42397.79
AIC	35013	15381.23	27128.56	42509.79
AICC	35013	15381.44	27128.66	42510.1
BIC	35021	15621.70	27296.15	42917.85

According to table 3, it can be seen that the joint model has much lower AIC, AICC and BIC than the sum of model fit statistics from univariate models. Thus the supremacy of the joint model over the univariate models can be concluded.

4. Conclusion

The main objective of this study was to analyze the dengue patients through the length of stay of dengue patients under the clustering nature. Since LOS of dengue patients is highly associated with the platelet count, joint model was formulated to analyze the dengue patients through the different destination of LOS and platelet count. Here platelet count was considered by converting to the logarithm to achieve normality. The proposed shared parameter model was fitted through two sub models that is multilevel discrete time competing risk model and multilevel normal model by using PROC NLMIXED procedure in SAS. Here multilevel discrete time competing risk model was fitted by the multinomial logistic model after converting the person-data set into the person-period data set. As the multinomial distribution was not supported by PROC NLMIXED, the likelihood of two univariate models were programmed to fit the joint model.

Factors that have an effect on different destination of LOS are, time indicators, year, age, classification, rainfall, temperature and humidity, while age, sex, classification, year place treated, rainfall, temp and humidity are associated factors for the log platelet count of a dengue patients. This study implies that the risk of death is higher for the patients who are at the critical phase rather than recovery phase. Also, there was a higher relation with the rainfall for the dead patients. Moreover, the patients who are from DHF2 are having higher risk.

The results of the univariate and joint model for checking the best model, proved that the joint model is better than two univariate models.

Acknowledgements

The authors would like to thank the University of Colombo for funding this research study, under its 'Research Grant 2014'. Further, we are grateful to Epidemiology unit, Medical Statistics Bureau, Colombo 10 and Dengue Management Unit, IDH, Colombo for helping us to get the necessary data to carry out this study.

References

1. Agresti, A. (1984). Analysis of Ordinal Categorical Data. *Wiley*.
2. Deslandes, E., & Chevret, S. (2010). Joint modeling of multivariate longitudinal data and the dropout process in a competing risk setting: application to ICU data. *BMC Med. Res. Methodol.*

3. *Disease Surveillance*. (2018). Retrieved 07 2018, from Epidemiology Unit, Ministry of Health: www.epid.gov.lk
4. Elashoff, R. M., Li, G., & Li, N. (2007). A joint model for longitudinal measurements and survival . *Biometrics*, 762-771.
5. Elashoff, R. M., Li, G., & Li, N. (2007). An approach to joint analysis of longitudinal measurements and competing risks failure time data. *Statist. Med*, 1999-2011.
6. Fernando, S.M. and Sooriyarachchi, M.R. (2018). Bivariate Negative Binomial Modelling of Epidemiological Data. *Open Science Journal of Statistics and Application* 5(3) : 47 - 57.
7. Gerhard, T. (1995). Competing risks models in discrete time with nominal or ordinal categories of response. *Quality and Quantity*, 405-420.
8. Gould, A., Boye, M. E., Crowther, M. J., Lbrahim, J. G., Quartey, G., Micallef, S., & Bois, F. Y. (2015). Joint modeling of survival and longitudinal non-survival data: current methods and issues: Report of the DIA Bayesian joint modeling working group. *Statist.Med*, 2181-2195.
9. Gueorguieva, R., Rosenheck, R., & Lin, H. (2012). Joint modelling of longitudinal outcome and interval-censored competing risk dropout in a schizophrenia clinical trial. *Statist. Soc*, 417-433.
10. Gueorguieva, R. V., & Agresti, A. (2001, Sep.). A Correlated Probit Model for Joint Modeling of Clustered Binary and Continuous Responses. *Journal of the American Statistical Association*, 96(455), 1102-1112. Retrieved from <http://www.jstor.org/stable/2670256>
11. Haller, B. (2014). *The Analysis of Competing Risks Data with a Focus on Estimation of Cause-Specific and Subdistribution Hazard Ratios from a Mixture Model*. Fakultät für Mathematik, Informatik und Statistik der Ludwig-Maximilians-Universität München.
12. Ivanova, A., Molenberghs, G., & Verbeke, G. (2015). Mixed models approaches for joint modelling of different types of responses. *Journal of Biopharmaceutical Statistics*.
13. Jayanetti, W and Sooriyarachchi, M.R.(2015). A multilevel study of dengue epidemiology in Sri Lanka: Modeling survival of dengue patients. *International Journal of Mosquito Research* 2(3) : 114-121.

14. Karunaratna, G.H.S., & Sooriyarachchi, M. R. (2017). Competing Risk Model for Dengue Epidemiology in Sri Lanka: Modeling Length of Stay in Hospital. *Tropical Agricultural Research*, 28(2), 200-210.
15. McCullagh, P. (1980). Regression model for ordinal data (with discussion). *Journal of the Royal Statistical Society*, 109-127.
16. Pinheiro, J. C., & Bates, D. M. (1995). Approximations to the log likelihood function in the non linear mixed effects models. *Journal of computation and graphical statistics*, 4, 12-35.
17. Pooransingh, S., Teeluksingh, S., & Dialsingh, I. (2016). Dengue Deaths: Associated Factors and Length of Hospital Stay. *Advance in Preventive medicine*. Retrieved from <http://dx.doi.org/10.1155/2016/6807674>
18. Prentice, R. L., & Gloeckler, L. A. (1978). Regression analysis of grouped survival data with application to breast cancer data. *Biometrics*, 57-67.
19. Prentice, R. L., Kalbfleish, J. D., Peterson, A. V., Flurnoy, N., Farewell, V. T., & Breslow, N. E. (1978). The Analysis of Failure Times in the presence of Cometing Risks. *Biometrics*, 541-544.
20. Sa, C., Dismuke, C. E., & Guimaares, P. (2007, November). Survival Analysis and Competing Risk Models of Hospital Length of Stay and Discharge Destination. *ResearchGate*. doi:10.1007/s10742-007-0020-9
21. Singer, J. D., & Willett, J. B. (1993). It's about time: Using discrete-time survival analysis to study duration and the timing of even. *Journal of Educational Research*, 18(2), 155-195.
22. Steele, F., & Goldstein, H. (n.d.). A General multilevel multistate competing risks Model for event History Data, with an Application to a study of contraceptive use dynamics. *Journal of Statistical Modeling*, 4(2), 145-159.
23. Steele, F., Diamond, I., & Wang, D. (1996). The determinants of the duration of contraceptive use in China:A multilevel multinomial discrete hazards modelling approach. *Demography*, 33, 12-33.
24. WHO. (2016). *Dengue and severe dengue*. WHO. Retrieved from <http://www.who.int/en/news-room/fact-sheets/detail/dengue-and-severe-dengue>
25. Wickremasuriya, S. and Sooriyarachchi, M.R. (2014). A Multilevel Analysis to Determine the impact of Demographic, Clinical and

Climatological Factors on Type of Dengue. *International Journal of Biological Science and Engineering* 5(2) : 43-57.

26. Williamson, P. R., Kolamunnage, D. R., Philipson, P., & Marson, A. G. (2008). Joint modeling of logitudinal and competing risk data. *stat med*, 27(30), 6426-38. doi:10.1002/sim.3451
27. Yacoub, S., Wertheim, H., Simmons, P. C., Screatton, G., & Wills, B. (2014). Cardiovascular manifestations of the emerging dengue pandemic. *Nature Reviews Cardiology*, 11, 335-345. doi:10.1038/nrcardio.2014.40