Bayesian Cox Semiparametric Spatial Survival for Hospitalisation of Dengue

Aswi Aswi [1,*], Susanna Cramb [1,2], Earl Duncan [1], Wenbiao Hu [2], Gentry White [1] and Kerrie Mengersen [1]

[1] ARC Centre of Excellence for Mathematical and Statistical Frontiers, Queensland University of Technology, 2 George St, Brisbane, Queensland 4001, Australia

[2] School of Public Health and Social Work, Institute of Health and Biomedical Innovation, Queensland University of Technology, Victoria Park Road, Kelvin Grove, Queensland 4059, Australia;

Environmental Health under Infectious Disease (EHID) Group

3 February 2020

Introduction

- Time-to-event or survival analysis is a set of statistical procedures for analysing data for which the outcome variable is time until an event occurs.
- In biostatistics, the event is often death, recovery or disease incidence, and survival time is usually defined in days, weeks, or years.
- Data used in survival analyses are often collected over distinct spatial regions.
- If there is the potential for survival time to vary between locations, it may be useful to include spatial information in the survival model.

Introduction

- Bayesian spatial survival models have recently emerged in the literature.
- An early example is the study of breast cancer and malignant melanoma patients using a fully Bayesian Cox model by incorporating spatial autocorrelation between neighbouring areas [1].
- We follow and extend the Bayesian approaches of Osnes and Aalen [1] for a semiparametric Cox proportional hazards model to describe hospitalisation for dengue fever in a hospital in Makassar, Indonesia.

Introduction

- Although spatial patterns in dengue fever have been the subject of considerable research effort, there appears to have been little research so far into modelling time-to-discharge for dengue hospitalisations.
- Hospitalisation of dengue fever patients is expensive [2] and understanding the geographic pattern of hospitalisation, duration of hospitalisation, influential factors affecting hospitalisation, is critical for hospital management.

Introduction

- A number of studies have used Bayesian survival models to investigate the hospitalisation for dengue in Indonesia.
- Two studies used a Weibull spatial survival model with a CAR frailty introduced by Banerjee et al. However, these papers implemented only one form of spatial prior (a CAR prior) and focused only on important factors associated with dengue.

Introduction

- Other authors used a Bayesian Weibull survival model for dengue patients in Makassar, Indonesia, but this study did not consider including a spatial term in their model.
- None of these papers appeared to examine how well the predictive values from the model fit the observed data which is critical for estimation and prediction.

Objective

- The overall objective of this study is to expand on the research described above in order to better describe spatial time to event for hospitalisations of dengue, with a focus on understanding dengue fever in Makassar, Indonesia.

Aims:

- Evaluate the suitability of Cox Model for estimation and prediction of LOS,
- Compare different spatial priors, and
- determine factors that significantly affect the duration of hospital stay for dengue fever patients.

Data

- Data on admitted dengue cases were obtained from patient medical records from a major public hospital in Makassar, namely Dr Wahidin Sudirohusodo,
- The data include inpatient dengue fever and dengue haemorrhagic fever (DHF) cases admitted during 1 January 2013 to 31 July 2018.

Data

The information collected:

- the LOS (*t*) for dengue fever patients in days,
- the patient's condition when discharged from the hospital (recovered/improved, not recovered, died, transferred to another hospital),
- sex,
- age,
- white blood cell (WBC) count,
- red blood cell (RBC) count,
- levels of haemoglobin (HGB),
- haematocrit (HCT),
- platelet (PLT) count, and
- address of patient residence.

Data

- Residential addresses at admission were geocoded and assigned to a district, which was one of 14 districts in Makassar city in 2017.
- As our spatial analysis was focused on Makassar, only residents of Makassar city were included in the study.
- Hospitalisation was defined as stay in hospital for at least one day, so outpatients (those who visited the hospital but did not stay overnight) were not included.

Data

- The initial dengue hospitalisation from Wahidin hospital comprised 2500 patients.
- Of these, 1603 (64%) patients came from outside Makassar city and
- 20 (0.8 %) patients had a LOS of zero days.
- There were 171 (6.8 %) dengue cases without clinical variables.
- There was one outlier, a patient who stayed 48 nights, that was also excluded.
- As a result, the final dengue fever patient dataset used in the analysis included 705 patients.

Data

- The response variable is the LOS for inpatients with dengue fever who enter during the study period, with the following conditions:
- If an inpatient is released from the hospital because they have recovered and this discharge is within the study period, their survival time (i.e., LOS) is categorised as uncensored survival data,
- If an inpatient instead experiences one of the following: Death, transfer to another hospital, exceeds the study end date (31 July 2018), or they self-discharge, then the hospital LOS is censored.

Model Formulation

The model of the LOS is assumed to follow a proportional hazard function (also called intensity function), for patient *k* = 1, 2, …, *K* with covariate X_k

$$h(t_k; x_k) = h_{kj} = h_0(t_k)\exp\{\beta^T X_k\} \qquad (1)$$

where $h_0(t_k)$ is the unknown baseline hazard rate (nonparametric),

β is the vector of regression parameters, and

both β and X are assumed to be constant over time $j = t_k$.

Model Formulation

.

For patient k = 1, 2, …, K,

I_{kj} is the intensity function (hazard function),

Y_{kj} is the at risk indicator, the time of observation of patient k within the j-th interval of time,

$N(t_k; X_k) = N_{kj}$ is the number of failures in interval [0, j].

dN_{kj} is the increment of N_k over the small-time interval $[j, j + d_j]$.

$$dN_{kj} = \begin{cases} 1 & ; \text{if patient } k \text{ is discharged well in time interval } j \\ 0 & ; \text{otherwise} \end{cases}$$

Model Formulation

Where dN_{kj} is assumed to have Poisson distribution with a mean of $I_{kj} = Y_{kj}h_{kj}$

$$dN_{kj} \sim \text{Poisson}(I_{kj}).$$

Consider the multiplicative intensity model in (1):

$$I_{kj} = Y_{kj}\exp\{\beta^T X_k\}h_0(t_k) \qquad (2)$$

We extend this model (2) by adding the spatial frailty term u_i using three different priors, namely ICAR, Leroux, and independent priors.

Model Formulation

$$I_{kj} = Y_{kj}\exp\{\beta^T X_k + u_j\}h_0(t_k)$$
$$h_0(t_k) \sim \text{Gamma}(ch_0^*(t_k), c)$$

c = 0.001

$\beta \sim N(0,100)$

$h_0^*(t_k)$ is as a prior guess of the hazard function, and c is the degree of confidence in that guess.

Spatial Prior Formulation

$$\left(u_i \middle| u_j, i \neq j, \tau_u^2\right) \sim N\left(\frac{\rho \sum_j u_j \omega_{ij}}{\rho \sum_j \omega_{ij} + 1 - \rho}, \frac{\sigma_u^2}{\rho \sum_j \omega_{ij} + 1 - \rho}\right)$$

ω_{ij} = 1 if *i*, *j* are adjacent, ω_{ij} = 0 otherwise.

$$\rho \sim \text{Unif}(0,1)$$
$$\sigma_u^2 \sim \text{IG}(1,0.1)$$

When ρ = 1 this prior reduces to the intrinsic CAR and when ρ = 0 it reduces to the independent prior.

For the Leroux model, the value of ρ is estimated,

For the intrinsic CAR and independent models, it is held fixed.

Statistical Analysis

- Deviance Information Criterion (DIC) and Watanabe–Akaike information criterion (WAIC) was used to compare predictive fit between models
- graphical examination of the observed data and modelled results is used to compare the model goodness of fit, which is less commonly examined
- for the selected model, variables were considered to be important if the 95% posterior credible interval (CI) for the un-exponentiated covariate coefficient does not contain zero (i.e., the corresponding CI for the exponentiated covariate coefficient does not contain one).

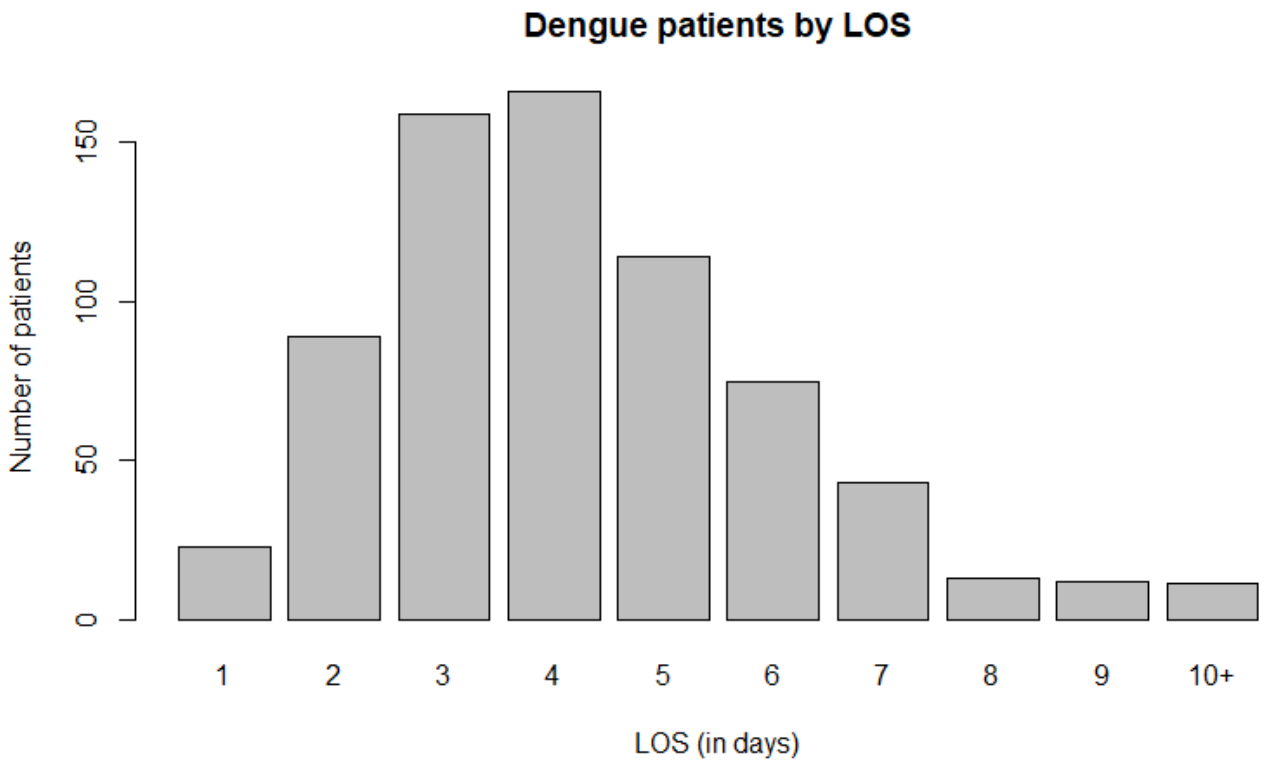
Descriptive Analysis of Wahidin hospital

	Districts	Number of dengue patients	Proportion
1	Biringkanaya	117	16.6
2	Bontoala	9	1.3
3	Makassar	19	2.7
4	Mamajang	15	2.1
5	Manggala	73	10.3
6	Mariso	13	1.8
7	Panakkukang	59	8.3
8	Rappocini	117	16.6
9	Tallo	17	2.4
10	Tamalanrea	219	31.1
11	Tamalate	28	3.9
12	UjungPandang	12	1.7
13	UjungTanah	<5	<0.7
14	Wajo	<5	<0.7
	Total	705	100

Descriptive analysis of continuous demographic and clinical data details for dengue patients in Wahidin hospital

Variables	Min	Q1	Median	Mean	Q3	Max
LOS (days)	1.00	3.00	4.00	4.28	5.00	16
Age (years)	0.23	9.34	18.28	20.50	26.06	79
WBC (10^3 μL)	0.60	2.80	4.10	4.96	6.10	42.80
RBC (10^6 μL)	1.96	4.35	4.77	4.78	5.18	8.06
HGB (gr/dl)	5.70	12.00	13.20	13.24	14.60	22.20
HCT (%)	18	36	40	39.61	43	61
PLT (10^3 μL)	4	47	90	102	138	361

Dengue patients in Wahidin hospital by the length of stay (LOS)



Spatial hazard ratios in each district for Cox Bayesian models with the Leroux prior.

Posterior hazard ratios for key parameters of the Cox models

Parameters	ICAR		Leroux		Independent	
	Mean	95% CI	Mean	95% CI	Mean	95% CI
Age	0.82	0.75; 0.89	0.82	0.75; 0.89	0.82	0.75; 0.89
Sex	1.01	0.94; 1.09	1.00	0.93; 1.09	1.01	0.93; 1.09
WBC	1.03	0.96; 1.1	1.03	0.96; 1.10	1.03	0.95; 1.10
RBC	1.22	1.04; 1.44	1.21	1.02; 1.42	1.21	1.03; 1.42
HGB	1.14	0.86; 1.5	1.12	0.82 ;1.53	1.12	0.83; 1.53
HCT	0.79	0.56; 1.09	0.80	0.55; 1.16	0.80	0.56; 1.15
PLT	0.89	0.82; 0.97	0.89	0.82; 0.96	0.89	0.82; 0.96
ρ	-	-	1.82	1.05; 2.67	-	-
σ^2	48.39	3.99; 2683.83	1.04	1.0003; 1.21	2.70	1.03; 37.38

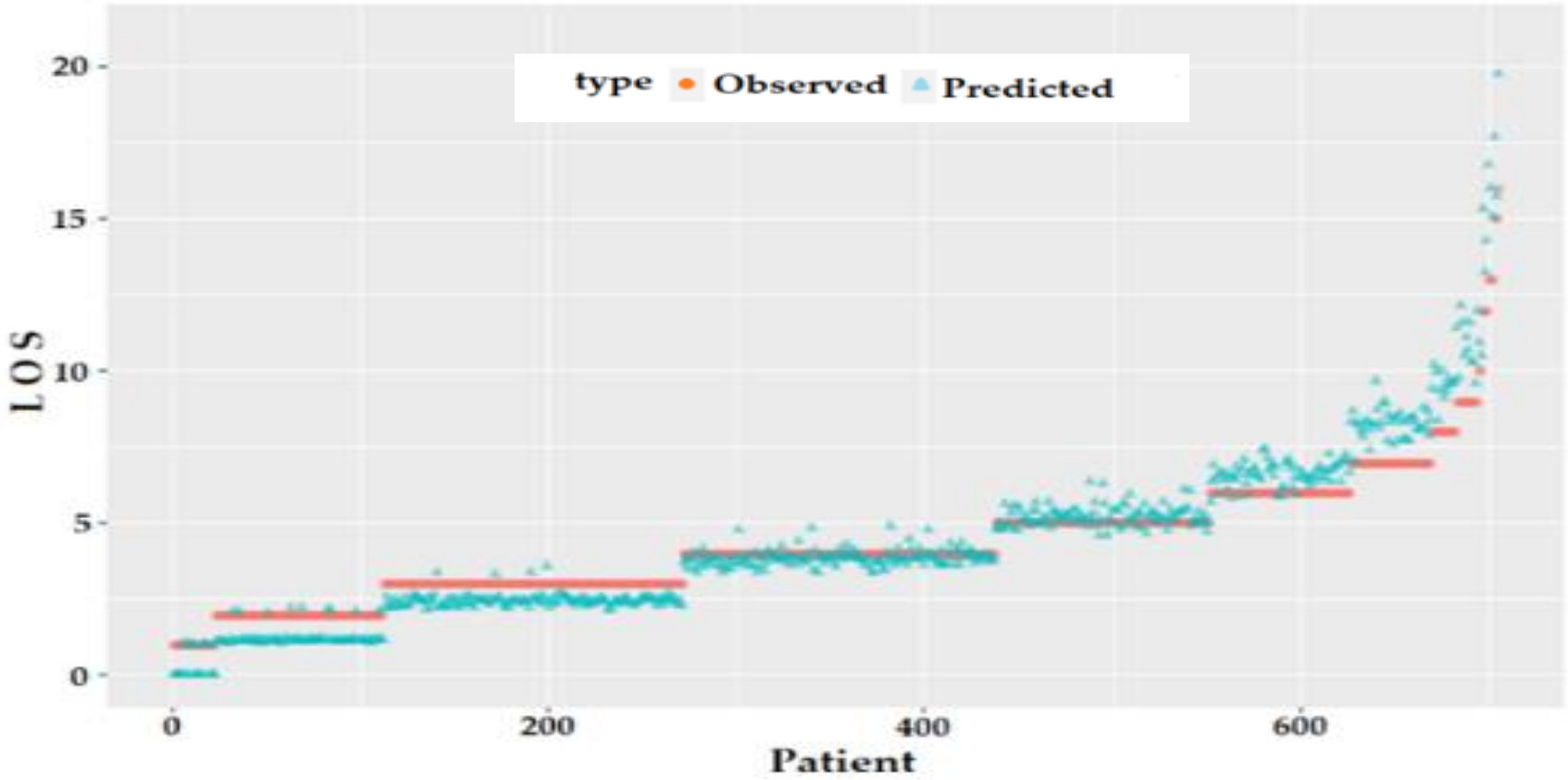
Estimation of spatial random effect (spatial frailty) hazard ratios for Cox models.

District	ICAR		Leroux		Independent	
	Mean	95% CI	Mean	95% CI	Mean	95% CI
1	1.08	0.89; 1.31	1.04	0.85; 1.34	1.05	0.90; 1.26
2	1.10	0.76; 1.63	1.03	0.79; 1.39	1.03	0.82; 1.38
3	0.82	0.56; 1.16	0.94	0.66; 1.20	0.94	0.71; 1.15
4	0.93	0.63; 1.32	0.98	0.72; 1.30	0.97	0.75; 1.21
5	0.94	0.75; 1.16	0.98	0.77; 1.23	0.97	0.79; 1.14
6	1.01	0.71; 1.42	1.00	0.77; 1.33	1.00	0.80; 1.26
7	1.07	0.84; 1.37	1.04	0.81; 1.36	1.04	0.88; 1.28
8	0.98	0.80; 1.18	1.00	0.80; 1.25	0.99	0.84; 1.16
9	1.20	0.84; 1.73	1.05	0.82;1.48	1.06	0.86; 1.44
10	0.93	0.79; 1.10	0.97	0.77; 1.18	0.95	0.80; 1.09
11	0.98	0.70; 1.35	1.01	0.77; 1.34	0.99	0.79; 1.24
12	0.90	0.58; 1.33	0.98	0.72; 1.29	0.97	0.75; 1.20
13	1.07	0.70; 1.65	1.02	0.78; 1.39	1.02	0.80; 1.37
14	1.07	0.50; 2.20	1.02	0.70; 1.58	1.01	0.78; 1.35

Goodness of fit of Cox models

PRIORS	Cox	
	DIC	WAIC
ICAR	2976.4	2967.77
Leroux	2963.8	2957.88
Independent	2963.5	2957.43

Plots of fitted values versus observed values of the Cox models using a Leroux CAR prior



- Cox models for all three priors indicate that age and platelet count were negatively associated with the LOS, while RBC were positively associated with the LOS for all three types of priors.
- The negative association between age, PLT, and the LOS means that dengue fever duration and the LOS were prolonged in children.
- One possible reason is that the immune system in children tends to be more vulnerable than for mature people. Furthermore, the lower the platelet count of dengue patients, the longer the LOS

- The positive association between RBC and the LOS means that dengue patients stay longer in the hospital when their RBC count is high.
- This is in line with previous studies that have found that patients with a low platelet count with high haematocrit levels are at very high risk of developing severe dengue.
- Although our models had the association for RBC count, rather than haematocrit, there was high correlation between these two measures, so our results are consistent

- The DIC and WAIC results were similar for the Leroux and independent priors, despite the ρ values indicating that some spatially structured variation was included. This may be because the sample size is small.
- To our knowledge, no previous dengue survival models have used the Leroux prior to the model spatial frailty in survival.
- This study demonstrates that including spatial structure may outperform an independent prior even when there is no substantive spatial variation present.
- The importance of a spatial frailty model is in agreement with some other studies. For example, Darmofal [26] reported a result similar to that found in our study, that the Cox model with spatial shared frailties (in the form of an ICAR prior) outperforms the Cox model with no spatial shared frailties.