Spatial Bayesian semi-parametric Cox-Leroux modelling of stroke patient hospitalization: aspects on survival

Aswi Aswi,[1] Bobby Poerwanto,[1] Nurussyariah Hammado,[2] Nurwan,[1] Oktaviana,[3] Siti Djawijah,[4] Susanna Cramb[5]

[1]Statistics Department, Universitas Negeri Makassar, Indonesia; [2]Medical Faculty, Universitas Negeri Makassar, Indonesia; [3]Mathematics Department, Universitas Negeri Makassar Indonesia; [4]Stroke Center hospital, Makassar, Indonesia; [5]Australian Centre for Health Services Innovation & Centre for Healthcare Transformation, Queensland University of Technology, Australia

Abstract

Survival analysis consists of a set of statistical methods used to analyse data where the outcome variable is the time until an event occurs. When such data are collected across distinct spatial regions, incorporating spatial information into survival models can be beneficial. A common approach is to apply an intrinsic Conditional Autoregressive (CAR) prior to an area-level frailty term to account for spatial correlation between regions. We extend the Bayesian Cox semi-parametric model by incorporating a spatial frailty term using the Leroux CAR prior. The aim was to improve the model's ability to describe stroke hospitalisations at the Stroke Centre Hospital in Makassar, Indonesia with a focus on understanding the geographic distribution of hospitalisations, Length of Stay (LOS) and factors influencing patient outcomes. The dataset was obtained from medical records of stroke patients admitted to this hospital (April 2021-June 2024). Variables included LOS, discharge outcomes, sex, age, stroke type, uric acid levels, hypertension, hypercholesterolemia, and diabetes mellitus. Our findings indicate that diabetes, stroke type and the presence of hypercholesterolemia significantly influence recovery rates in stroke patients. Specifically, patients with diabetes had lower recovery, while those with hypercholesterolemia and ischemic stroke patients had faster recovery compared to those with haemorrhagic strokes.

Correspondence: Aswi Aswi, Statistics Department, Universitas Negeri Makassar, Indonesia.
E-mail: aswi@unm.ac.id

Key words: Bayesian, spatial, survival, semiparametric Cox-Leroux, Indonesia.

Conflict of interest: the authors declare no potential conflict of interest, and all authors confirm accuracy.

Availability of data and materials: The data and R code will be made available upon request.

Acknowledgements: the authors gratefully acknowledge the funding from the Ministry of Education, Culture, Research, and Technology (KEMENDIKBUDRISTEK) of Indonesia through DRTPM's Regular Fundamental Research program in 2024, under contract numbers 065/E5/PG.02.00.PL/2024 and 2840/UN36.11/LP2M/2024. We also extend our gratitude to Alya Raodhatul Jannah, Nur Fhatika Sari, Ajriansyah, Nurul Mutmainnah CH., Muh. Yusran, Muh. Alwi Sanur, Dwi Yuniarty Nur Zainal, Muh. Mahadtir, and Meyrna Vidya Shanty, who are students from the Statistics Department at Universitas Negeri Makassar, for their valuable assistance in collecting stroke patient data from various hospitals in Makassar, Indonesia.

Received: 11 February 2025.
Accepted: 23 March 2025.

Geospatial Health 2025; 20:1380
doi:10.4081/gh.2025.1380

Introduction

Stroke is a cerebrovascular disease and the second leading cause of death worldwide, following heart disease, with 5.8 million fatal cases annually (Cai *et al.*, 2019). It is also a major contributor to disability (Willeit *et al.*, 2020), with the Global Burden of Disease Study 2017 estimating 80.1 million people suffer from stroke each year and 6.2 million of these cases resulting in death (GBD 2016 Causes of Death Collaborators, 2017; GBD 2017 Causes of Death Collaborators, 2018; GBD 2019 Stroke Collaborators, 2021). Indonesia likewise has high mortality from stroke (Hussain *et al.*, 2016), with an increasing prevalence from 7 per 1,000 individuals in 2013 to 10.9 per 1,000 in 2018 (Kemenkes, 2023). Stroke has become a major financial burden, ranking third in healthcare expenditure after heart disease and cancer and costs reaching 3.23 trillion rupiah (around 195.65 million USD) in 2022 (Kesehatan, 2023). This represents a substantial increase from the 1.91 trillion rupiah (around 115.59 million USD) spent in 2021 (Kesehatan, 2023). Stroke risk factors are typically categorized into two types: modifiable risk factors, which can be altered or controlled through lifestyle changes or medication, and non-modifiable risk factors, which are related to genetic predisposition (Johansson *et al.*, 2021; da Silva Paiva *et al.*, 2022). To mitigate the severe impact of stroke, analysing modifiable risk factors that influence the clinical recovery of stroke patients is essential.

Survival models have often been used to investigate aspects of stroke outcomes. Cox models have been applied to evaluate the relationship between stroke risk factors and hospitalisation frequency (Yu *et al.*, 2021), recurrence and recurrence-free survival after stroke (Flach *et al.*, 2020) and mortality (Li & Li, 2022). Stroke recurrence, risk factors and long-term survival have also been examined using parametric or flexible parametric survival models (Rodan *et al.*, 2012; Romain *et al.*, 2020; Peng *et al.*, 2022; Elhefnawy *et al.*, 2023). Stroke recovery among hospitalised patients is significantly influenced by geographic factors, with timely access to treatment playing a crucial role in patient outcomes. Disparities in healthcare infrastructure such as the availability of stroke centres, neurologists, and emergency medical services affect the likelihood of receiving prompt interventions. Research indicates that patients in rural or underserved areas experience treatment delays due to longer travel distances and limited emergency response systems, resulting in poorer recovery outcomes (Hammond *et al.*, 2020). A study analysing rural-urban disparities in stroke mortality in the United States aimed to identify factors contributing to differences in case-fatality rates (Georgakakos *et al.*, 2022). The findings revealed that rural stroke patients were less likely to receive intravenous thrombolysis or endovascular therapy and exhibited higher in-hospital mortality compared to their urban counterparts.

Socioeconomic status, which varies by region, also influences healthcare access, affordability of post-stroke rehabilitation, medication adherence and lifestyle factors critical to recovery. Patients from lower-income areas often face challenges in accessing follow-up care, rehabilitation services, and preventive measures, leading to worse functional recovery (Fernandez-Lazaro *et al.*, 2019; Yadav *et al.*, 2022). Multiple small-area studies examining stroke mortality rates have found that outcomes were closely linked to socioeconomic factors, including educational attainment and income (Karp *et al.*, 2016; Yadav *et al.*, 2022).

Despite the extensive research on stroke outcomes, and the potential importance of time to treatment on outcomes, spatial effects are rarely considered. One study examined in-hospital stroke mortality with a Gaussian random field for spatial effects in a Bayesian log-logistic proportional odds model (Nazar *et al.*, 2023). Bayesian models are widely used in disease mapping since Bayesian analysis improves the precision and interpretability of the results when prior distributions are integrated (Lawson, 2018b;2018a). The conditional autoregressive prior is particularly popular for spatial effects (Banerjee *et al.*, 2014) and has been applied in survival analysis for cancer (Dasgupta *et al.*, 2014) and dengue fever (Aswi *et al.*, 2020; Thamrin *et al.*, 2021), yet there is limited research involving this approach for stroke.

Survival analysis examines the time until an event occurs, often dealing with censored data when not all individuals experience the event during the study. This issue is well-documented in the literature (Kleinbaum & Klein, 2012; Banerjee *et al.*, 2014). Another challenge arises when events vary by geographic region, making the incorporation of spatial data crucial. A notable example is the Bayesian hierarchical spatial survival model developed by Banerjee, Wall, and Carlin (Banerjee *et al.*, 2003), which integrated Conditionally Autoregressive (CAR) frailty effects into a Weibull baseline hazard. We adapted and extended the Bayesian approach introduced by Osnes and Aalen (Osnes & Aalen, 1999) for semi-parametric Cox proportional hazards models to examine stroke hospitalisations in Makassar, Indonesia. This study aimed to model the recovery rate of hospitalised stroke patients using a Bayesian spatial survival Cox-Leroux model, incorporating spatial data to account for geographic influences on patient outcomes. A key innovation is the integration of spatial factors to assess how geography affects stroke progression and modifiable risk factors.

Materials and Methods

Data collection

Data on stroke patients from the Stroke Centre Hospital in Makassar, Indonesia, located in the Mamajang District, were retrieved from the medical records covering the period from April 2021 to June 2024. Participants in this study were individuals whose data were routinely collected during their hospitalisation for stroke in Makassar. Key variables included the length of hospital stay (in days), discharge status (recovered/improved, not recovered, deceased or transferred to another hospital), age, gender, uric acid levels, presence of hypertension, hypercholesterolemia, diabetes mellitus, stroke type and residential address. The selection of these variables was based on previous studies (Sweetnam *et al.*, 2012; Weng *et al.*, 2013; Bergh *et al.*, 2016; Shang *et al.*, 2020; Setyopranoto *et al.*, 2022; Shah *et al.*, 2023) but constrained by the availability of hospital data.

The residential addresses of the patients at the time of diagnosis were geocoded and assigned to one of the 15 districts in Makassar City in 2024. The Stroke Centre Hospital is a major referral facility for stroke patients in Eastern Indonesia and therefore also receives patients from outside Makassar, such as from Kalimantan Island. However, as this study focuses solely on Makassar City, only patients residing within Makassar were included in the analysis. Additionally, patients with a hospital stay of less than one day (*i.e.* those not admitted as inpatients) were excluded.

Prior to analysis, the dataset underwent a thorough cleaning process, with its accuracy verified by cross-checking discrepancies by the hospital's medical records officer. The initial dataset comprised 1,327 stroke patients but as several key variables for some patients (28 (2.1%) lacked clinical data, 201 (15.1%) had no recorded residential address and 2 (0.2%) had a length of stay (LOS) recorded as zero), they were excluded along with 642 patients (48.4%) residing outside Makassar City, which resulted in a final analytical dataset of 454 patients.

The response variable in this study was the length of hospital stay for stroke inpatients admitted during the study period, with the following conditions: i) if a patient was discharged after recovery or improvement within the study period, their hospital stay was treated as uncensored survival data, as they experienced the event of interest (successful discharge); ii) if a patient died, was transferred to another hospital, exceeded the study period end date (June 30, 2024) or self-discharged, their LOS was considered censored.

The data were right-censored, as in some cases the exact time of recovery was not recorded—either because patients left the hospital before recovery or remained hospitalised at the end of the study. Written permission was obtained from the hospital for data usage.

Model formulation

In this study, we used a semi-parametric Bayesian spatial survival model by incorporating a spatial random effect using Leroux CAR priors within the Cox proportional hazards framework. The Cox models were initially estimated using R version 4.3.2 (Team, 2019), and we assessed the Proportional Hazards (PH) assumption for each covariate through statistical tests based on scaled Schoenfeld residuals, using the survival package and the *cox.zph* function in R. The covariates sex, age, stroke type, uric acid levels, hypertension, hypercholesterolemia and diabetes mellitus were found to satisfy the PH assumption. Details on variable definitions are given in Table 1.

Bayesian Cox-Leroux model

The semi-parametric Cox model (Cox, 1972) is a widely applied method for analysing time-to-event data, as it does not require the specification of a parametric baseline hazard function. In this study, we build upon the approach introduced Osnes and Aalen (Osnes & Aalen, 1999) to extend the semiparametric Cox proportional hazards model. We model the LOS using a proportional hazards function, also known as the intensity function, for patients $k = 1, 2, \ldots, K$ incorporating covariate X_k (Osnes & Aalen, 1999):

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

where $h_0(t_k)$ represents the unknown baseline hazard rate treated nonparametrically; β a vector of regression parameters; with both β and X_k are assumed to remain constant over time with $j = t_k$. The analysis of counting process data, such as survival data, was typically conducted through intensity modelling: for a patient $k = 1, 2, \ldots, K$, I_{kj} represents the intensity function (or hazard function); Y_{kj} whether the patient is at risk during the j-th time interval; the time of observation for patient k within this interval t_k; N (t_k ; X_k) the number of events (failures) occurring in the interval [0, j]; and dN_{kj} the increment of N_k over the small-time interval $[j, j + d_j]$.

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

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

$$dN_{kj} \sim \text{Poisson}(I_{kj}) \quad \text{Eq. 2}$$

Consider the multiplicative intensity model presented in Eq.1:

$$I_{kj} = Y_{kj}\exp\{\beta^T X_k\}h_0(t_k) \quad \text{Eq. 3}$$

We extend the model in Eq. 3 by incorporating a spatial frailty term, μ_i, which follows a Leroux CAR prior. Consequently, Eq. 3 becomes:

$$I_{kj} = Y_{kj}\exp\{\beta^T X_k + u_i\}h_0(t_k) \quad \text{Eq. 4}$$

where the baseline hazard function, $h_0(t_k)$, is assigned a Gamma prior distribution, Gamma($ch_0^*(t_k)$, c), with c = 0.001; $h_0^*(t_k)$ represents a prior estimate of the hazard function; c reflects the degree of certainty in that estimate, and the regression coefficients β follow a normal prior distribution, $\beta \sim N(0,100)$ (Osnes & Aalen, 1999; Aswi *et al.*, 2020). We applied the Leroux CAR prior (Brian *et al.*, 2000), which includes a single frailty term that controls the degree of spatial autocorrelation among neighbouring areas through a constant parameter ρ, which can be estimated, with ρ taking values between 0 and 1, and has three different hyperpriors. The conditional distribution of u_i, given u_j for $i \neq j$, is:

$$(u_i | u_j, i \neq j, \sigma_u^2) \sim \text{N}\left(\frac{\rho\sum_j u_j \omega_{ij}}{\rho\sum_j \omega_{ij} + 1 - \rho}, \frac{\sigma_u^2}{\rho\sum_j \omega_{ij} + 1 - \rho}\right) \quad \text{Eq. 5}$$

where $\omega_{ij} = 1$ if areas i and j are adjacent; and ω_{ij} otherwise = 0. The prior distribution for ρ is uniform (Unif) between 0 and 1:

$$\rho \sim \text{Unif}(0,1)$$

(Lee, 2013; Aswi *et al.*, 2020)

Table 1. Research variables.

Variable	Name	Description
Y	Survival time	The time a stroke patient receives hospital treatment until recovery or improvement, measured in days
X_1	Age	The age of the stroke patient at the time of hospital admission, measured in years.
X_2	Sex	The gender of the stroke patient 0 = Female 1 = Male
X_3	Hypertension	0 = No hypertension 1= The presence of hypertension
X_4	Diabetes	0 = No diabetes 1 = The presence of diabetes
X_5	Hypercholesterolemia	0 = No hypercholesterolemia 1= The presence of hypercholesterolemia
X_6	Uric Acid	0 = No uric acid 1= The presence of uric acid
X_7	Stroke Types	0= haemorrhagic 1= ischemic
S	Status	The outcome status of the stroke patient. 0 = No event (not recovered/deceased) 1 = Event occurrence (recovery/improvement)

and the variance σ_u^2 follows an inverse gamma (IG) distribution:

$$\sigma_u^2 \sim \mathrm{IG}(1{,}0.1)$$

(Lee, 2013)

when $\rho = 1$, this prior simplifies to the intrinsic CAR model and when $\rho = 0$, it reduces to the independence model.

A sensitivity analysis was conducted to examine the impact of different priors on the posterior estimates. We evaluated three different prior options for the variance of the spatial frailty, σ_u^2, for Cox–Leroux models: IG(1, 0.1), IG(0.5, 0.05) and IG(2,0.2) (Aswi *et al.*, 2020). These hyperparameter choices yielded priors ranging from highly concentrated to highly diffuse, relative to the data and posterior distribution. IG(1, 0.1) is weakly informative, allowing for greater spatial variability and heterogeneity, particularly useful when the strength of spatial correlation is uncertain. IG(0.5, 0.05) is more diffuse, accommodating higher variance. IG(2, 0.2) is more informative, assuming moderate spatial variation, which may be appropriate if stroke care practices are relatively standardized across regions.

The models were assessed using two methods. First, the predictive fit was compared between models using the Watanabe-Akaike Information Criterion (WAIC) (Watanabe, 2010), where lower WAIC values indicate better fit. WAIC is a fully Bayesian criterion that relies on the entire posterior distribution rather than a single point estimate. It offers a robust assessment of a model's predictive accuracy while mitigating the risk of overfitting making it particularly suitable for complex hierarchical models (Gelman *et al.*, 2014). WAIC has demonstrated some advantages over traditional model selection criteria (Ariyo *et al.*, 2020). Second, model goodness-of-fit, which is less frequently evaluated, was assessed through graphical comparisons between the observed data and the model predictions. While WAIC provides an overall measure of model fit, it does not offer insights into localised discrepancies or the model's adequacy for specific subgroups within the data. Graphical comparisons serve as an intuitive approach to evaluating how well the model replicates patterns in the observed data, helping to identify potential discrepancies that numerical metrics may overlook. In the chosen model, variables were deemed significant if the 95% posterior Credible Interval (CInt) for the un-exponentiated covariate coefficient did not include zero (*i.e.*, the corresponding CInt for the exponentiated covariate coefficient did not include one). Model parameters were estimated using the R2WinBUGS package (Sturtz *et al.*, 2005) in R version 3.6.1 (Team, 2019). Posterior estimates for the parameters were derived from 5,000 MCMC samples after a burn-in period of 5,000 samples, without thinning. Convergence was assessed using trace plots of the parameters and σ^2. The data and R code is available upon request.

Comparing models

Three distinct prior distributions were applied to the spatial survival Cox-Leroux model to assess their influence on posterior estimates. The selection of priors was guided by relevant literature (Aswi *et al.*, 2020). Model performance was evaluated using goodness-of-fit criteria, specifically the WAIC.

Results

Descriptive analysis

Table 2 presents the distribution of stroke admissions across 15 subdistricts showing that the highest proportion of patients (17.0%) originates from Rappocini District, followed by Tamalate district (14.8%). The average LOS for stroke patients was 6 days, as shown in Figure 1, with the most frequent LOS being 6 days (19.8%), followed by 5 days (19.4%) and 4 days (16.5%). The LOS ranged from a minimum of 1 day to a maximum of 21 days. The majority of stroke patients admitted to the Stroke Centre were older adults, with a mean age of 60.4 years and a median age of

Table 2. The distribution of stroke patients admitted to the stroke centre during the study period based on district of residence.

No	District	Number of stroke patients	Proportion (%)
1	Mamajang	31	6.83
2	Manggala	42	9.25
3	Mariso	19	4.19
4	Sangkarrang	3	0.66
5	Rappocini	77	16.96
6	Tamalate	67	14.76
7	Makasar	57	12.56
8	Ujung Pandang	8	1.76
9	Panakukkang	37	8.15
10	Bontoala	21	4.63
11	Wajo	8	1.76
12	Ujung Tanah	12	2.64
13	Tallo	50	11.01
14	Tamalanrea	9	1.98
15	Biringkanaya	13	2.86
Total		454	100

60.0 years, ranging from 26 to 92 years. Most patients were male (54.0%) and had comorbidities, including diabetes (21.4%), hypercholesterolemia (17.0%), hyperuricemia (3.0%), and hypertension (20.0%). Of the total stroke cases, 13.7% were diagnosed as haemorrhagic stroke, with the remaining 86.3% as ischemic stroke (due to thrombosis).

Posterior summaries and spatial frailty estimation in Cox-Leroux models

A sensitivity analysis was conducted to evaluate the impact of different prior distributions on the posterior estimates. According to the results, the Cox-Leroux model with the IG(0.5, 0.05) hyperprior produced the lowest WAIC value (1926.1), potentially slightly outperforming the models with the IG(2, 0.2) hyperprior (WAIC = 1929.6) and the IG(1, 0.1) hyperprior (WAIC = 1927.6). Posterior hazard ratios for key parameters of Cox-Leroux models are given in Table 3.

The covariates significantly associated with LOS were found to include diabetes, hypercholesterolemia and type of stroke (Table 3). Specifically, patients with diabetes had lower recovery (HR: 0.89), while those with hypercholesterolemia and ischemic stroke patients had faster recovery times compared to those with haemorrhagic strokes, with a Hazard Ratio (HR) of 1.15 and 1.27, respectively. Stroke patients with diabetes had an 11% lower hazard of recovery compared to those without diabetes. In contrast, patients with hypercholesterolemia had a 15% higher hazard of recovery, and ischemic stroke patients had a 27% higher hazard of recovery compared to haemorrhagic stroke patients. These effects were statistically significant. The spatial dependency parameter, ρ (HR = 1.76; 95% CInt: 1.04-2.67), represents the degree of spatial correlation within the data. A positive estimate indicates a significant spatial dependence in patient outcomes relative to location, as evidenced by the CI excluding 1. This finding suggests that geographical factors play a statistically significant role in influencing patient outcomes. The 95% CInts for all spatial random effects included one, indicating that on average, stroke patients from all districts admitted to the stroke centre exhibit similar recovery rates and discharge times. This regional uniformity, after adjusting for covariates, is visualized in Figure 2. Based on Figure 2, districts with a slightly higher risk (HR > 1) include Tamalate (HR = 1.061), Manggala (HR = 1.031), and Sangkarrang (HR = 1.017). Estimation of spatial random effect (spatial frailty) hazard ratios for Cox-Leroux models are given in Table 4. Plot of fitted values versus observed values of the Cox-Leroux model is given in Figure 3. The model fit the data quite well (Figure 3).

Discussion

The primary objective of this study was to analyse the geographic distribution of hospitalisations, LOSs and factors influencing patient outcomes. To do so, we enhance the capability in modelling stroke hospitalisations by extending the Bayesian survival Cox semi-parametric model with a spatial frailty term based on the Leroux CAR prior. A key contribution of this work was the integra-

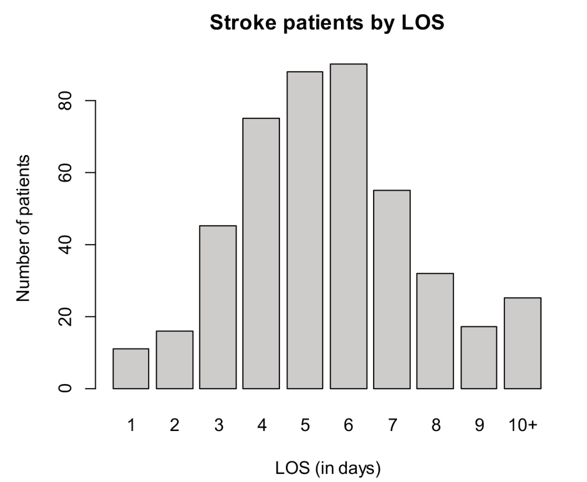

Figure 1. Division according to the number of days in hospital for the stroke patients.

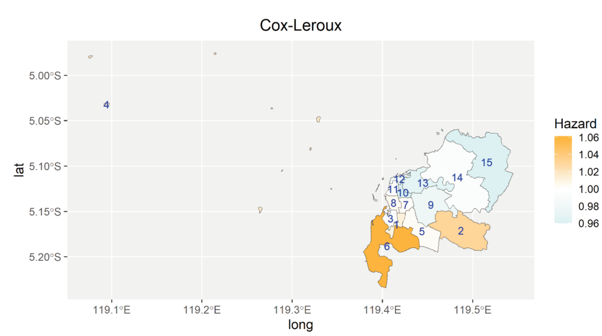

Figure 2. Distribution by district of spatial hazard ratios for each for Cox-Leroux model in Makassar, Indonesia.

Table 3. Posterior hazard ratios for key parameters of Cox-Leroux models.

No	Parameter	Mean	Exp (mean)	95% CInt
1	Age	-0.04	0.96	0.87, 1.06
2	Sex	-0.02	0.98	0.89, 1.08
3	Diabetes mellitus	-0.12	0.89	0.80, 0.98
4	Hypertension	-0.02	0.98	0.89, 1.09
5	Hypercholesterolemia	0.14	1.15	1.04, 1.27
6	Uric Acid Levels	-0.06	0.94	0.85, 1.03
7	Ischemic vs Haemorrhagic stroke	0.24	1.27	1.13, 1.44
	ρ	0.57	1.76	1.04, 2.67
	σ^2	0.03	1.03	1.001, 1.15

Cint, Credible Interval; Exp, Exponential.

tion of spatial components to assess the influence of geographic factors on stroke progression and modifiable risk factors. A sensitivity analysis was performed to evaluate the effect of different prior specifications on posterior estimates. Specifically, we examined three inverse-gamma priors for the variance of the spatial frailty term (s_u^2) in the Cox–Leroux model. Several covariates were found to be significantly associated with LOS, particularly the presence of comorbidities (diabetes, hypercholesterolemia), and stroke type. In contrast, age, sex, the presence of hypertension, and uric acid levels did not show a significant association with LOS. The findings that age and sex are not significant predictors of LOS align with previous research. For example, a study conducted in Nebraska found no significant difference in overall survival in a multivariate analysis that accounted for age, comorbidities, and rehabilitation factors (DeVries *et al.*, 2013). In that study, the Cox proportional hazards model was employed to estimate overall survival. After adjusting for age and other relevant variables, the study concluded that mortality rates were statistically equivalent between men and women. Our study found that patients with diabetes exhibited slower recovery, whereas those with hypercholesterolemia and ischemic stroke experienced faster recovery compared to patients with haemorrhagic stroke. This finding is consistent with previous studies, which have demonstrated that individuals with diabetes have approximately twice the risk of developing ischemic stroke (Sweetnam *et al.*, 2012; Shang *et al.*, 2020). In contrast, patients with hypercholesterolemia had a higher likelihood of recovery, while those with ischemic stroke had an even higher likelihood of recovery compared to individuals with haemorrhagic stroke. A study examining the impact of hypercholesterolemia on recovery outcomes following acute ischemic stroke (Weng *et al.*, 2013) indicated that patients with hypercholesterolemia presented with lower stroke severity at admission but experienced less favourable recovery during hospitalisation, irrespective of age. A limitation of this study is that the results are based solely on data from a single major hospital, which may result in limited case representation for certain districts. We acknowledge that incorporating data from multiple hospitals could enhance the generalisability of the findings and potentially influence the results. Future research should consider expanding the dataset by incorporating data from multiple hospitals across different districts or regions to improve the generalisability of the findings. Additionally, exploring advanced modelling techniques, such as hierarchical Bayesian models or machine learning approaches, may also enhance predictive accuracy and robustness in identifying key determinants of patients. Despite this limitation, our study offers key strengths, particularly the incorporation of spatial factors to assess the influence of geographic variations on disease progression and clinical outcomes in stroke patients, with a particular emphasis on modifiable risk factors. Additionally, the application of the Bayesian method in estimating the spatial survival model enhances the reliability of predictions by providing posterior distributions of the parameters. The use of prior distributions in Bayesian analysis further improves the accuracy and interpretability of the results.

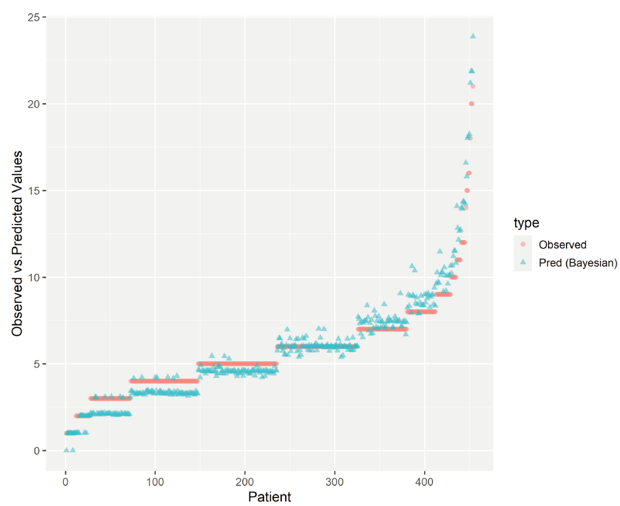

Figure 3. Plot of fitted values versus observed values of the Cox-Leroux model.

Table 4. Estimation of spatial random effect (spatial frailty) hazard ratios for Cox-Leroux models.

No	District	Mean	Exp(mean)	95% CInt
1	Mamajang	0.012	1.012	0.818, 1.259
2	Manggala	0.031	1.031	0.834, 1.348
3	Mariso	-0.008	0.992	0.769, 1.241
4	Sangkarrang	0.017	1.017	0.809, 1.333
5	Rappocini	0.004	1.004	0.817, 1.238
6	Tamalate	0.059	1.061	0.877, 1.395
7	Makasar	0.002	1.002	0.809, 1.236
8	Ujung Pandang	0.005	1.005	0.805, 1.261
9	Panakukkang	-0.022	0.978	0.774, 1.184
10	Bontoala	-0.041	0.959	0.731, 1.160
11	Wajo	-0.006	0.994	0.764, 1.275
12	Ujung Tanah	-0.021	0.979	0.750, 1.219
13	Tallo	-0.032	0.968	0.760, 1.167
14	Tamalanrea	-0.006	0.994	0.775, 1.265
15	Biringkanaya	-0.040	0.961	0.675, 1.249

Conclusions

The key finding of this study is that the Bayesian spatial survival Cox-Leroux model with the IG (0.5, 0.05) hyperprior outperformed the other models. The analysis revealed that covariates significantly associated with stroke hospitalisation included diabetes, hypercholesterolemia and stroke type. Patients with diabetes had lower recovery, while those with hypercholesterolemia and ischemic stroke patients had faster recovery times compared to those with haemorrhagic strokes. Additionally, stroke patients from all districts admitted to the stroke centre exhibited, on average, similar recovery rates and discharge times. Future research could expand the analysis to include additional hospitals and compare the length of stay across facilities.

References

Ariyo O, Quintero A, Muñoz J, Verbeke G, Lesaffre E, 2020. Bayesian model selection in linear mixed models for longitudinal data. J Appl Stat 47:890-913.

Aswi A, Cramb S, Duncan E, Hu W, White G, Mengersen K, 2020. Bayesian spatial survival models for hospitalisation of Dengue: A case study of Wahidin hospital in Makassar, Indonesia. Int J Environ Res Public Health 17:878.

Banerjee S, Carlin BP, Gelfand AE, 2014. Hierarchical modeling and analysis for spatial data, Chapman and Hall/CRC, 584 pp.

Banerjee S, Wall MM, Carlin BP, 2003. Frailty modeling for spatially correlated survival data, with application to infant mortality in Minnesota. Biostatistics 4:123-42.

Bergh C, Udumyan R, Appelros P, Fall K, Montgomery S, 2016. Determinants in adolescence of stroke-related hospital stay duration in men: a national cohort study. Stroke (1970) 47:2416-8.

Brian GL, Lei X, Breslow N, Halloran M, Elizabeth BD, 2000. Estimation of disease rates in small areas: a new mixed model for spatial dependence. In: Statistical models in epidemiology, the environment, and clinical trials. pp. 179-191.

Cai Y, Towne JSD, Bickel CS, 2019. Multi-level factors associated with social participation among stroke survivors: China's health and retirement longitudinal study (2011-2015). Int J Environ Res Public Health 16:5121.

Cox DR, 1972. Regression models and life-tables. J R Stat Soc Series B Stat Methodol 34:187-202.

da Silva Paiva L, de Alcantara Sousa LV, Oliveira FR, de Carvalho LEW, Raimundo RD, Correa JA, de Abreu LC, Adami F, 2022. Temporal trend of the prevalence of modifiable risk factors of stroke: an ecological study of Brazilians between 2006 and 2012. Int J Environ Res Public Health 19:5651.

Dasgupta P, Cramb SM, Aitken JF, Turrell G, Baade PD, 2014. Comparing multilevel and Bayesian spatial random effects survival models to assess geographical inequalities in colorectal cancer survival: a case study. Int J Health Geogr 13:36.

DeVries DMS, Zhang YP, Qu MP, Ma JMS, Lin GP, 2013. Gender difference in stroke case fatality: an integrated study of hospitalization and mortality. J Stroke Cerebrovasc Dis 22:931-7.

Elhefnawy M, Noor Harun S, Leykhim T, Tangiisuran B, Zainal H, et al., 2023. A parametric time-to-event modelling of recurrent ischemic stroke after index stroke among patients with and without diabetes mellitus: implementation of temporal validation of the model. Cureus 15:e50794.

Fernandez-Lazaro CI, Adams DP, Fernandez-Lazaro D, Garcia-González JM, Caballero-Garcia A, Miron-Canelo JA, 2019. Medication adherence and barriers among low-income, uninsured patients with multiple chronic conditions. Res Social Adm Pharm 15:744-53.

Flach C, Muruet W, Wolfe CDA, Bhalla A, Douiri A, 2020. Risk and secondary prevention of stroke recurrence: a population-base cohort study. Stroke 51: 2435–2444.

GBD 2016 Causes of Death Collaborators, 2017. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 390:1151-210.

GBD 2017 Causes of Death Collaborators, 2018. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 392:1736-88.

GBD 2019 Stroke Collaborators, 2021. Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 20:795-820.

Gelman A, Hwang J, Vehtari A, 2014. Understanding predictive information criteria for Bayesian models. Stat Comput 24:997-1016.

Georgakakos PK, Swanson MB, Ahmed A, Mohr NM, 2022. Rural stroke patients have higher mortality: an improvement opportunity for rural emergency medical services systems. J Rural Health 38:217-27.

Hammond G, Luke AA, Elson L, Towfighi A, Maddox KEJ, 2020. Urban-Rural Inequities in Acute Stroke Care and In-Hospital Mortality. Stroke 51: 2131-2138.

Hussain MA, Al Mamun A, Peters SAE, Woodward M, Huxley RR, 2016. The burden of cardiovascular disease attributable to major modifiable risk factors in Indonesia. J Epidemiol 26: 515-521.

Johansson A, Drake I, Engström G, Acosta S, 2021. Modifiable and non-modifiable risk factors for atherothrombotic ischemic stroke among subjects in the Malmö Diet and Cancer Study. Nutrients 13: 1952.

Karp DN, Wolff CS, Wiebe DJ, Branas CC, Carr BG, Mullen MT, 2016. Reassessing the stroke belt: using small area spatial statistics to identify clusters of high stroke mortality in the United States. Stroke 47:1939-42.

Kemenkes, 2023. Kenali Stroke dan Penyebabnya. Retrieved from https://ayosehat.kemkes.go.id/kenali-stroke-dan-penyebabnya

Kesehatan KDJP, 2023. World Stroke Day 2023, Greater Than Stroke, Kenali dan Kendalikan Stroke. Retrieved from https://yankes.kemkes.go.id/read/1443/world-stroke-day-2023-greater-than-stroke-kenali-dan-kendalikan-stroke

Kleinbaum DG, Klein M, 2012. Survival analysis: a self-learning text, Springer Science & Business Media, 712 pp.

Lawson A, 2018a. Bayesian disease mapping : hierarchical modeling in spatial epidemiology, CRC Press.

Lawson A, 2018b. Bayesian disease mapping: hierarchical modeling in spatial epidemiology, Chapman and Hall/CRC, 486 pp.

Lee D, 2013. CARBayes: An R Package for Bayesian Spatial Modeling with Conditional Autoregressive Priors. J Stat Softw 55:1-24.

Li X-D, Li M-M, 2022. A novel nomogram to predict mortality in patients with stroke: a survival analysis based on the MIMIC-

III clinical database. BMC Med Inform Decis Mak 22:92.
Nazar E, Esmaily H, Yousefi R, Jamali J, Ghandehari K, Hashtarkhani S, Jafari Z, Shakeri MT, 2023. A spatial variation analysis of in-hospital stroke mortality based on integrated prehospital and hospital data in Mashhad, Iran. Arch Iran Med 26:300-9.
Osnes K, Aalen OO, 1999. Spatial smoothing of cancer survival: a Bayesian approach. Statistics Med 18:2087-99.
Peng Y, Ngo L, Hay K, Alghamry A, Colebourne K, Ranasinghe I, 2022. Long-term survival, stroke recurrence, and life expectancy after an acute stroke in Australia and New Zealand from 2008–2017: a population-wide cohort study. Stroke 53:2538-48.
Rodan L, McCrindle BW, Manlhiot C, MacGregor DL, Askalan R, Moharir M, deVeber G, 2012. Stroke recurrence in children with congenital heart disease. Ann Neurol 72: 103-111.
Romain G, Mariet AS, Jooste V, Duloquin G, Thomas Q, Durier J, Giroud M, Quantin C, Béjot Y, 2020. Long-term relative survival after stroke: the Dijon Stroke Registry. Neuroepidemiology 54:498-505.
Setyopranoto I, Upoyo AS, Isworo A, Sari Y, Vidyanti AN, Olcay A, 2022. Awareness of being at risk of stroke and its determinant factors among hypertensive patients in Banyumas, Indonesia. Stroke Res Treatment 2022:1-7.
Shah M, Buscot M-J, Tian J, Phan HT, Marwick TH, Dwyer T, Venn A, Gall S, 2023. Sex differences in the association between stroke risk factors and pre-clinical predictors of stroke in the childhood determinants of Adult Health study. Atherosclerosis 384:117171.
Shang Y, Fratiglioni L, Marseglia A, Plym A, Welmer A-K, Wang H-X, Wang R, Xu W, 2020. Association of diabetes with stroke and post-stroke dementia: A population-based cohort study. Alzheimers Dement 16:1003-12.
Sturtz S, Ligges U, Gelman A, 2005. R2WinBUGS: A Package for Running WinBUGS from R. J Stat Softw 12:1-16.
Sweetnam D, Holmes A, Tennant KA, Zamani A, Walle M, Jones P, Wong C, Brown CE, 2012. Diabetes impairs cortical plasticity and functional recovery following ischemic stroke. J Neurosci 32:5132-43.
Team RC, 2019. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org
Thamrin SA, Aswi, Ansariadi, Jaya AK, Mengersen K, 2021. Bayesian spatial survival modelling for dengue fever in Makassar, Indonesia. Gac Sanit 35:S59-S63.
Watanabe S, 2010. Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. J Mach Learn Res 11:3571-94.
Weng W-C, Huang W-Y, Su F-C, Chien Y-Y, Wu C-L, Lee T-H, Peng T-I, 2013. Less favorable neurological recovery after acute stroke in patients with hypercholesterolemia. Clin Neurol Neurosurg 115:1446-50.
Willeit P, Toell T, Boehme C, Krebs S, Mayer L, Lang C, Seekircher L, Tschiderer L, Willeit K, Rumpold G, Schoenherr G, Griesmacher A, Ferrari J, Knoflach M, Lang W, Kiechl S, Willeit J; STROKE-CARD study group, 2020. STROKE-CARD care to prevent cardiovascular events and improve quality of life after acute ischaemic stroke or TIA: A randomised clinical trial. EClinicalMedicine 25:100476.
Yadav RS, Chaudhary D, Avula V, Shahjouei S, Azarpazhooh MR, Abedi V, Li J, Zand R, 2022. Social Determinants of Stroke Hospitalization and Mortality in United States' Counties. J Clin Med 11:4101.
Yu S, Alper HE, Nguyen AM, Maqsood J, Brackbill RM, 2021. Stroke hospitalizations, posttraumatic stress disorder, and 9/11-related dust exposure: Results from the World Trade Center Health Registry. Am J Ind Med 64:827-36.