International Journal of Innovative Research and Advanced Studies (IJIRAS)
Volume 6 Issue 10, October 2019

ISSN: 2394-4404

Bayesian Spatial Cox Proportional Hazard Model For HIV Infected Tuberculosis Cases In Chennai

Srinivasan R

Ponnuraja C

Pradeep A M

Dept. of Statistics, National Institute for Research in Tuberculosis, ICMR, Chennai, India

Moeng S.R.T

Sivasamy R

Botswana University, Botswana

Venkatesan, P

Sri Ramachandra University, Chennai, India

Abstract:

Background: Spatial survival model is used to obtain information on spatial inequalities for studying time to event data. Spatial data are complex in nature, and require high level modeling like, Bayesian spatial random effect model in which spatial model incorporate random effects at neighboring locations are allowed to exhibit spatial dependence using conditionally autoregressive (CAR) prior for HIV infected tuberculosis patients in Chennai ward.

Methods: Data was collected from National Institute for Research in Tuberculosis (NIRT). Monte Carlo Markov Chain (MCMC) techniques were used to estimate the parameters. Stata and WinBUGS software were used for Bayesian Survival models. The Deviance information Criterion (DIC) criterion was used to evaluate the various competing models.

Results: The Cox proportional models with frailty revealed that weight at baseline was statistically significant for death in HIV associated Tuberculosis (TB) patients. Also, spread of tuberculosis in Chennai ward has been diverse, with many wards having a low autocorrelation of infection and the epidemic being most extreme in north east of Chennai wards.

Conclusion: The minimum value of Deviance Information Criterion revealed that spatial model was important for spatial correlated time to event data and also, there were unmeasured risk factors associated with death in all the region.

I. BACKGROUND

Survival analysis has been widely used in biostatistics, econometrics, and many other areas where time-to-event data occur [1-3]. This method is used when outcome variable of interest is time until an event occurs for individuals. The time indicates any unit of time for their time to end point. The event of interest is death in an area or relapse of TB, etc. In the spatial survival model, the importance of the spatial inequalities was considered including, the impact of difference in area level characteristics and individual level risk factors [4-5]. Bayesian model are very popular since the last three decades, after revolution of computer and software packages. Bayesian model minimize the estimation bias and increase the accuracy of the parameter estimates. Bayesian spatial model were developed to explicitly incorporate spatial correlation between areas while describing spatial survival pattern across area [6].

Regions or wards are assumed to be independent in conventional survival analysis, but in spatial survival analysis regions are spatially arranged and closer proximity to each other might also be similar in magnitude [7]. It also concerns identically and independently distributed (i.i.d) concepts and robustness of the estimation of the parameter estimation. To overcome these problem like spatial dependence and heterogeneity within individual, Bayesian survival model was used in this study in such a way that they allowed other factors, like random effects and spatial dependence effects to be included in the analysis[8]. Bayesian computation is providing stable estimates for each region in the spatially

ISSN: 2394-4404

arranged regions[9]. It also allowed for unexplained heterogeneity to be investigated in the disease maps.

Parametric proportional hazards model with a Weibull formulation for the baseline hazards, placing a univariate CAR structure on the frailty intercept terms was studied for infant mortality data[10]. Semi-parametric setup under the usual Cox proportional hazards model was further extended to the spatiotemporal case, using a univariate CAR[11]. Bayesian spatial survival models for political event process were extensively studied in spatial, non-spatial model with in parametric and semiparametric distribution and proved that spatial dependence in the random effects also produces changes in the effects of covariates [12]. The Cox proportional models with frailty for HIV were explored for other diseases [13-14]. Spatial and spatio temporal models for tuberculosis disease were extensively studied for Chennai using Bayesain models. [15-17].

II. METHODS

Bayesian model is based on prior knowledge about distribution with full likelihood which gives the posterior distribution of the parameters using Markov Chain Monte Carlo (MCMC) estimation method[18]. In this model random effects at neighboring locations are allowed to exhibit spatial dependence that is incorporated by specifying a conditionally autoregressive (CAR) prior developed by Besag, et al.,[19] for application in effects in time-to-event data across neighboring units, with the neighbors defined via an adjacency matrix where each neighbour of a unit is given a weight of 1, while each non-neighbour of a unit is given a weight of 0. CAR prior distribution is to capture correlations across both geographic regions and considers the random effects for a given region.

The Bayesian semi-parametric model of Cox proportional hazard model was used to explain the event of death occurring at a given time is affected by covariates viz., age, sex, treatment regimen, and weight at baseline for HIV associated tuberculosis cases for right censored data in Chennai ward. The advantage of this model is referred to be semi-parametric because no parametric distribution specified is for the baseline hazard ($h_0(t)$) and covariates are assumed constant over time in this model.

III. SURVIVAL DATA

Data was collected from National Institute for Research in Tuberculosis (NIRT), Chennai where NIRT is conducting clinical trials in TB and HIV since 1953. The data taken for this study is a pilot study of HIV infected Tuberculosis patients admitted in clinical trial who were treated with three types of treatments of 6 months to 8 months duration during 1998-99. The Revised National Control Programme (RNTCP) regimen of CategoryI, Category II, and Category .III regimen [20] were randomly assigned into the above said regimens and followed up for another 12 months. The death occurred in an area is event of interest for this analysis. Street addresses were geo-coded for the entire ward in Chennai ward that contains 256 wards in 9 zones, Chennai district.

STATISTICAL ANALYSIS

Cox proportional hazard model was performed using Stata software to model the hazard rate of death and significance factors associated with death cases in Chennai ward. The same model was performed in WinBUGS software with frailty (random effect) model to account for the heterogeneity at area level. The Bayesian spatial random effect model was also used to find the spatial inequalities using Markov Chain Monte Carlo (MCMC) estimation method. The DIC was used to assess the model selection and evaluation for both spatial and non spatial approach.

IV. RESULTS

Conditional autoregressive model which indicates the existence of spatial dependence on the composition of covariance where is the CAR parameter distribution stating precision or variance inverse of its random effect distribution. The WinBUGS software used for Cox PH model for which observed time, censoring time with other covariate mentioned above and adjancy matrix for Chennai wards were included for analysis. The prior for this model is hyper prior i.e., Gamma prior which is distributed with a small precision, thus taking a larger neighborhood structure into account. The software used for no frailty estimates are Stata and for disease mapping QGIS (10.2) which is open source software. The non-informative priors were considered where $\beta \sim N(0,0.000001)$. The initial 3000 sample were discarded to avoid autocorrelation effect of initial values. Using a burn in of 10000 samples and additional of 10000, from 10000 to 30000 Gibbs samples were drawn, posterior estimates of β's given in the table. The comparisons of the posterior estimates indicate that the convergence has achieved in 30000 iterations.

The semi-parametric Cox proportional hazard model shows the non-frailty, frailty and Spatial frailty for Cox model to HIV infected TB data with considering the covariates age, treatment (cat),sex and weight were given in the table. Need to specify the models used, with theta, γ , σ , τ see attached document. Shoul probably be additional equations following eqtn 1

	Cox Model(Non-Frailty)			Cox Model(Frailty)			Spatial Cox frailty Model		
Cov.	HR	SE	95 % CI	HR	SE	95 % CI	Mean	MC error	Credible Interval
Cat	0.87	0.17	0.59, 1.27	0.87	0.17	0.59, 1.27	0.08	0.00	-0.35 0.50
Age	1.00	0.02	0.97, 1.03	-	-	- -	0.01	0.00	-0.03 0.04
Sex	1.42	0.47	0.74, 2.73	1.42	0.47	0.74, 2.73	0.46	0.00	-0.26 1.18
wt0m	0.96*	0.02	0.93, 0.99	0.96*	0.02	0.93, 0.99	-0.07	0.00	-0.10 -0.04
Theta	-	-	- -	0.00	0.00	- -	3.50	0.17	0.62, 10.10
γ	-	-	- -	-	-	- -	0.04	0.00	0.00, 0.12
σ	-	-	- -	-	-	- -	2.56	0.02	1.79, 3.57
τ	-	-	- -	-	-	- -	0.17	0.00	0.07, 0.31

$p<0.05$

Table 1: Summaries Statistics for Cox Model

From the above table, the wards are assumed as an important covariate to account heterogeneity. Hence, the age is clustered into 5 groups (0-25, 26-30, 31-35, 36-40 and above 40). The Cox proportional models with frailty showed

ISSN: 2394-4404

that weight at baseline was statistically significant for death in HIV associated Tuberculosis (TB) patients in all the models. The hazard Rate for sex indicates the male were 1.42 times risk of death in non spatial model. The amount of heterogeneity accounted through age is very minimal and other supporting evidence like SE and hazard ratio (HR) are similar in with and without frailty model. But in the overall, Bayesian spatial Cox frailty model's MC error is less compared to non-spatial frailty and non-frailty model for all the covariates. The mean value for *tau* (τ) is 0.17(0.07-0.31) very less. Spatial Cox with frailty model fits better for this time to event survival disease data.

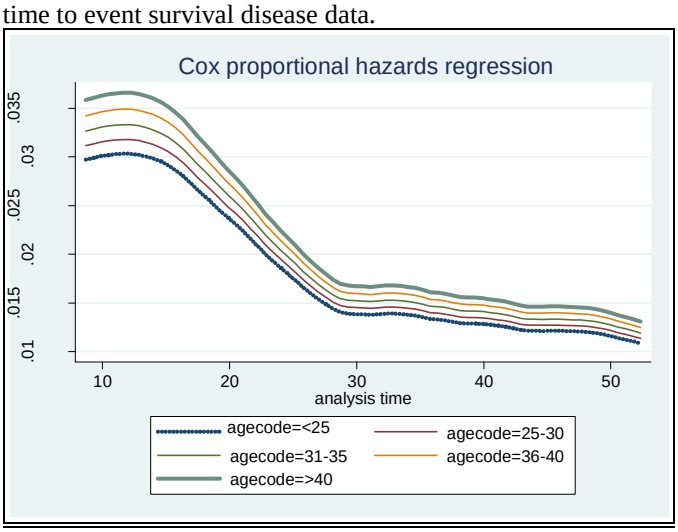

Figure 1: Hazard curve for Cox model for the different age group

Cox proportional hazard rate curve was used to check the PH assumptions for age group and found that the curves were differing. Also, younger age group which is mentioned in blue colour is having higher survival time than the other age groups.

The spatial autocorrelation of W under Cox model are presented in figure

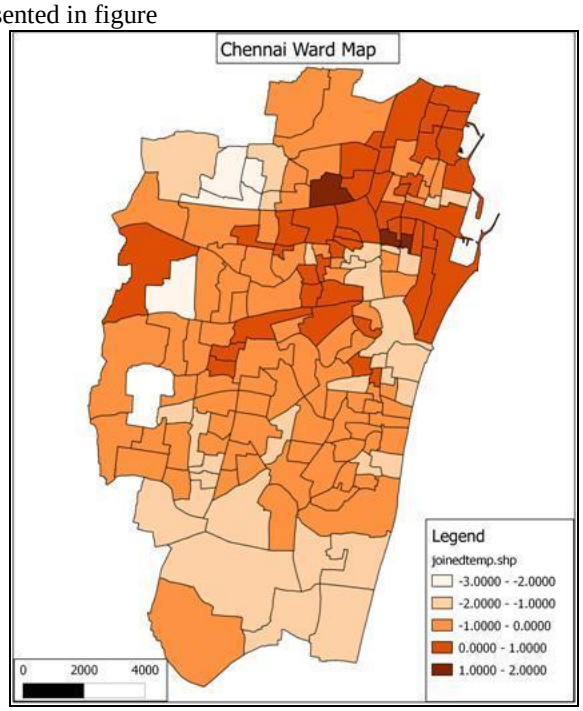

Figure 2: Posterior Mean for spatial dependence on of TB

Figure 2 shows the mean value for spatial autocorrelation classified into five groups with light brown colour to dark brown colour. The dark colour indicates the severity of disease for the Chennai wards using Cox PH model. The spatial autocorrelation is more on north-east of Chennai ward. Probably may include the Table before Table 2, which gives parameter estimates, of betas

Table2 gives the parameter estimates of beta values for the all the variables which includes intercept, treatment category, age, sex and weight.

Beta	Model1	Model2	Model3	Model4	Model5	Model6
Intercept	0.05593	0.1261	0.06266	0.09692	-0.2018	-0.3764
Treatment category	0.006426	-0.00839	0.008862	-0.4144	-	-0.5672
Age	0.4217	0.311	-2.38	-	0.4051	3.45
Sex	-0.07466	-1.685	-	-0.06917	-0.04786	3.345
Weight	-6.978	-	-0.06672	5.091	7.488	-3.338

Table 2: Parameter estimates of the Cox Model

Model1 is the full model that includes spatial dependence (w) and random effect(v) with all the variable of interest; the negative coefficient value of beta for sex and weight indicates the negative trend in the model1. Model2 consists of the all the variable except weight; sex and Treatment category shows the negative coefficient value. In the model 3, age and weight are having negative coefficient value after removing sex in the model. Model4 is the full model except age; sex, Treatment category shows negative coefficient value in our model. Model 5 is the full model except Treatment category in the model; sex only shows the negative value in our model. The model6 does not consist of spatial dependence (w) and random effect(v), here also Treatment category and weight are negatively associated with death in our model.

Model No.	Model	Bayesian Cox frailty			
		Spatial		Non-Spatial	
		pD	DIC	pD	DIC
1	$\beta_0 + \beta_{Tr.Cat} * Tr.Cat + \beta_{Age} * Age + \beta_{Sex} * Sex + \beta_{Wt} * Wt + W_i + V_i$	53.392	623.028	98.721	698.165
2	$\beta_0 + \beta_{Tr.Cat} * Tr.Cat + \beta_{Age} * Age + \beta_{Sex} * Sex + W_i + V_i$	35.696	740.086	88.636	836.176
3	$\beta_0 + \beta_{Tr.Cat} * Tr.Cat + \beta_{Age} * Age + \beta_{Wt} * Wt + W_i + V_I$	42.867	719.335	87.876	830.435
4	$\beta_0 + \beta_{Tr.Cat} * Tr.Cat + \beta_{Sex} * Sex + \beta_{Wt} * Wt + W_i + V_i$	41.333	716.73	56.769	819.451
5	$\beta_0 + \beta_{Age} * Age + \beta_{Sex} * Sex + \beta_{Wt} * Wt + W_i + V_i$	120.86	657.809	67.686	717.567
6	$\beta_0 + \beta_{Tr.Cat} * Tr.Cat + \beta_{Age} * Age + \beta_{Sex} * Sex + \beta_{Wt} * Wt$	85.243	950.181	75.864	1155.45

Table 3: Goodness of fit for Weibull and Cox models

Assessing the model fit was used by DIC for this data, in which first model contains full model with correlated effect were captured through (*W*) and uncorrelated effects were captured through (*V*). Here, the spatial model and non spatial model were compared with in which another six models were compared again. The model with lowest DIC value was considered to be the better model. Among all models, spatial frailty models fits better than the other non-spatial frailty models. Among the spatial frailty model, the lowest amount of DIC is 623.028 for model (1) of Cox model which also identified the better model than the other spatial model. The

model (5) consists of age, sex and weight are having the next lowest DIC (657.81) value which indicates that the effect of Treatment regimen has very less impact in this model as compared with age and sex. The model (4) and model (3) are almost similar in DIC value 716.73 and 719.33 respectively. The model (2) consists of treatment regimen, age and sex is having high value in DIC (740.08) indicates that weight at baseline is important factor in this analysis. The last model(6) is fixed effect model, with highest DIC value of 950.18, indicates that random effects with in region and spatial autocorrelation between ward were important factors for estimating any parameter in spatial analysis.

V. CONCLUSION

All applications of disease mapping methods in the literature are in the context of diseases such as cancers and vector borne diseases. But spatial survival models have rarely been applied in the context of HIV associated Tuberculosis disease modeling. The models help us to spatial variation of a disease in Chennai ward, which may lead to detection of unknown risk factors. From our results it was found that weight at baseline, is one of the factor associated with death in our study and hazard rate of male having two fold comparing to female. The spatial dependence was diverse in Chennai ward, some ward having less dependence and north-eastern wards in Chennai having spatial high spatial dependence. Spatial correlated Random effect model accounts higher heterogeneity (12.3) in our model which indicates that the regional variation and other environmental factors influencing survival pattern of disease in this model.

Even though weight and sex significantly were associated with death, the spatial model of model1 consists of spatial dependence and random effect with the entire variable shows the better model; treatment category and age are positively associated while sex and weight are negatively associated with death. It may be due to unmeasured variable were adjusted when spatial dependence and spatial randomness were considered in our model that also gives the unbiased estimates value.

The overall spatial frailty model fits better other than the non-spatial frailty model. Among the spatial frailty model, the minimum value of DIC in full Cox model. This indicates that adding of spatial frailty dependence on both models, spatial frailty accounts higher heterogeneity compared with non-spatial models and it takes into account the spatial dependence between the state-level frailties. The fixed effect model has the high DIC value for both spatial and non-spatial. The result reveals that the importance of modeling the spatial autocorrelation that is common to spatial data. For accounting regional variation, the frailty model provides unbiased estimates of standard error and gives less biased estimates in our model. It is clear that a non-spatial and non-frailty model understates the unexplained heterogeneity in the data. Hence, spatial frailty model captures the unexplained spatial heterogeneity and it draws accurate inferences about other spatial covariates of interest.

REFERENCES

[1] Banerjee S, Carlin B and Gelfand A E (2004): Hierarchical Modeling and Analysis for Spatial Data. Boca Raton: Chapman & Hall.
[2] Gelfand A Diggle P and Guttorp P (2010): Handbook of spatial statistics, Chapman & Hall / CRC.
[3] Dasgupta P Cramb S M Aitken J F Turrell G and Baade P D (2014): Comparing multilevel and Bayesian spatial random effects survival models to assess geographical inequalities in colorectal cancer survival: a case study, International Journal of Health Geographics,13:36
[4] Allred SR Brainard DH (2013). A Bayesian model of lightness perception that incorporates spatial variation in the illumination. Journal of Vision, 13(7):18.
[5] Besag J (1974): Spatial interaction and the statistical analysis of lattice systems. Journal of Royal Statistical Society, Series-B, 36: 192-236.
[6] Besag J and Kooperberg C (1995): On conditional and intrinsic autoregressions. Biometrika, 82: 733-746.
[7] Tobbler Law
[8] DA Piggott, AD Muzaale, SH Mehta, et al.(2013). Frailty, HIV infection, and mortality in an aging cohort of injection drug users PLoS One, 8, p. e54910
[9] Lawson A, Browne W and Vidal Rodeiro C (2003): Disease mapping with WinBUGS and MLwiN. John Wiley & Sons Ltd.
[10] Banerjee S, Wall M M and Carlin B P (2003): Frailty modeling for spatially-correlated survival data with application to infant mortality in Minnesota. Biostatistics, 4: 123-142.
[11] Li Y and Ryan L (2002): Modeling spatial survival data using semiparametric frailty models. Biometrics, 58, 287-297.
[12] Darmofal, D. 2007. "Bayesian Spatial Survival Models for Political Event Processes," Unpublished: http://people.cas.sc.edu/darmofal/DarmofalBayesianSpatialSurvival.pdf.
[13] Therneau, T. M. and Grambsch, P. M. (2000). Modeling Survival Data: Extending the Cox Model. New York: Springer
[14] Cox D and Oakes D (1984): Analysis of Survival Data. London: Chapman and Hall.
[15] Venkatesan P and Srinivasan R (2008): Bayesian model of HIV/AIDS in India: A spatial Analysis. Applied Bayesian Statistical Analysis: 51-56.
[16] Srinivasan R Venkatesan P: Bayesian Spatio temporal model for tuberculosis in India: A Spatial analysis, Indian Journal of Medical Research, April, 2015.
[17] Venkatesan P, Srinivasan R and Dharuman C (2012): Bayesian conditional autoregressive model for mapping tuberculosis prevalence in India, International Journal of Pharmaceutical Studies and Research, Vol.III, online journal.
[18] MCMC estimation
[19] Besag J, York J C and Mollie A (1991): Bayesian image restoration, with two applications in spatial statistics. Annals of the Institute of Statistical Mathematics, 43: 1-59.

[20] Revised National Tuberculosis Control Programme DOTS-Plus Guidelines (2010). Central TB Division, Directorate General of Health Services, Ministry of Health & Family Welfare, Nirman Bhavan, New Delhi, India.