Indian J Med Res 141, April 2015, pp 478-480

Correspondence

Bayesian spatio-temporal model for tuberculosis in India

Sir,

India has about 2.2 million new cases of tuberculosis (TB) annually[1]. TB is the most common HIV-related opportunistic infection, and caring for patients with both diseases is a major public health challenge. Spatial data are correlated with space and tendency of close areas to have similar disease rates, and ignoring this spatial correlation leads to bias and inefficient estimation of disease risk rate. Actual rates of disease of all the States are to be achieved from different spatial regions and various time points by using Bayesian methods. The resulting estimates involve a weighted average of disease rates from the specific region with neighbourhood areas.

Bayesian spatial method is useful in minimizing bias and variance compared to conventional statistical methods[2]. The spatial and temporal model provides spatial distribution and temporal changes of relative risk of a disease across a study area. Most Bayesian methods propose extensions of the purely spatial models postulated by others[2-5]. Bernardinelli and others[6] suggested a model in which both area specific intercept and temporal trend were modeled as random effects. Daiane and *et al*[7] used this method for finding spatio-temporal patterns of TB. The authors and their group have earlier[8-10] proposed several spatial models for multivariate data based on general univariate conditional autoregressive (CAR) model for TB and HIV disease mappings.

The aim of the study was to correlate the disease rate among the States in India and assess the rate of neighbourhood TB infection and its causes by the Baysian model including CAR approach of spatial and temporal pattern changes for the two survey periods. The data of TB cases were taken from National Family Health Surveys (NFHS) conducted in a representative sample of households aged between 15-54 yr throughout India during the years 1992-1993, 1998-1999 and 2005-2006[11]. The survey collected information on population, nutrition, quality of health, family planning services, *etc*., for 29 States and seven Union Territories of India. The number of cases of TB disease Y_i occurring in area S_i was taken, where the set of areas $\{S_i\}$, i = 1,…,n represented a partition of the region under study. For each area S_i, the expected number of cases E_i was also computed using regional reference rates for the disease prevalence. Each region is defined as a polygon in a map file and region is associated with a unique index.

We considered Y_{it} and $E_{it,}$ the observed and expected disease counts in region i during time period 't'[12]. The Poisson model is then $Y_{it} \mid \mu_{it} \sim Po\ (E_{it} e^{\mu i l t})$ where μ_{it} is the log relative risk of disease for region i, and year t. The model is; $(\mu_{it}) = \alpha + \mu_i + v_i + \beta_i t\ [j] + \delta_{i\,t}\ [j]$ where α is an intercept, μ is a spatial correlated heterogeneity effect, v_i is a spatial uncorrelated heterogeneity, $\beta_i t[j]$ is a linear trend term in time t[j], and δ_i is an interaction random effect between space and time. The CAR approach to an aerial unit of State level analysis was used for spatial correlated model that takes larger neighbourhood structure into account in our model.

The WinBUGS software[13] was used for space time model analysis. It implements models for data that are collected within discrete regions and smoothing is done based on Markov random field for the neighbourhood structure of the regions relative to each other. To fit the model in WinBUGS, observed count, and expected count with adjancy matrix for India were included for analysis. The Gamma prior distributed with mean 0.5 and precision (τ) taking a larger neighbourhood structure into account was used for spatial correlated model. Model fitting was carried out using two separate chains starting from different initial values. Convergence was checked using time series plots of

Table. Posterior estimates for the parameter in the space time model

Parameters	Mean	Standard deviation	Median	95% Credible interval	
α	-0.67	0.28	-0.70	-1.11	-0.16
β	0.93	0.07	0.93	0.81	1.10
τ.δ	0.04	0.01	0.04	0.02	0.06
τ.μ	0.02	0.01	0.02	0.01	0.04
τ.ν	0.47	0.28	0.40	0.12	1.22

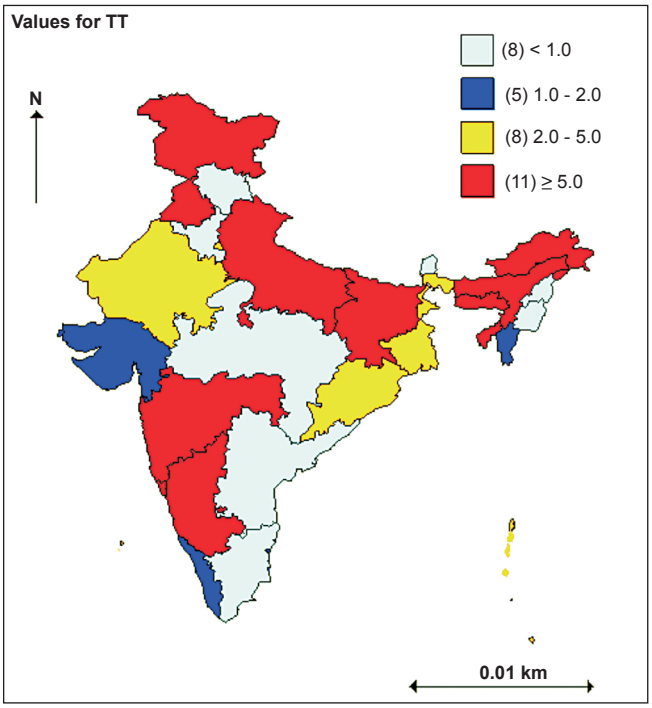

Fig. Posterior expected temporal trend (TT) of relative risk for India. The different colours show the areas with TT of relative risk. Values in parentheses indicate number of States.

samples for each chain and by computing the Gelman and Rubin diagnostic[13]. Thinning was used to avoid the autocorrelation of the data. The first 5000 samples were discarded as a burn-in and the convergence was reached after 40000 Markov chain Monte Carlo (MCMC) iterations[13]. The posterior expected temporal trend (TT) for India is given in the Figure.

The Bayesian TT of relative risk was classified into four groups in which areas with TT > 1 were shaded in red and other colours. The neighbourhood relative risk among 24 States was very high with TT≥1. Of these, five States had TT 1.0-2.0, eight States 2.0-5.0 and 11 States had TT≥5.

The posterior estimates for the parameter of the space time model are given in the Table. The trend coefficient β(0.93) was significant. The risk exp(β) was 1.93, indicating about two-fold increase during the period. The mean of spatial correlated heterogeneity (τ.μ) was 0.02, while mean of spatial uncorrelated heterogeneity (τ.ν) was 0.47 where τ was the precision.

In conclusion, a two- fold increase of TB disease risk in space time variation between the two survey periods was observed. The spatio-time changes of TB relative risk was increased in 25 States. However, the precision of the variance was small indicating the risk in any given area was similar to that in neighbouring areas and time. Also, it was observed that Bayesian spatial CAR method gave smoothening of relative risk of disease in all the regions based on neighbourhood structure and provided variance reduction among the States, which would be more stable estimate of the pattern of underlying risk of disease rather than the raw estimates. This approach effectively borrows information from neighbouring areas than from areas far away and smoothening local rates toward local, neighbouring values[16]. It reduces the variance in the associated estimates and allows for the spatial effect of regional differences in regional populations. The limitation of the study was that NHFS data consisted of the age group between15-49 yr for female and 15-54 yr for female and prevalence of TB for this survey was based on questionnaire only.

R. Srinivasan[1] & P. Venkatesan[1,*]
[1]Department of Statistics, National Institute for Research in Tuberculosis (ICMR), Mayor Ramanathan Road, Chetpet, Chennai 600 031, Tamil Nadu, India
[*]*For correspondence*: venkaticmr@gmail.com

References

1. Central TB Division, Ministry of Health and Family Welfare. Annual Status Report (2014). Revised National Tuberculosis Control Programme. Available from: *http://www.tbcindia.nic.in/pdfs/TB%20INDIA%202014.pdf*, accessed on May 22, 2015.
2. Besag J, York J, Mollie A. Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Stat Math* 1991; *43* : 1-59.
3. Clayton D, Kaldor J. Empirical Bayes estimates of age-standardised relative risks for use in disease mapping. *Biometrics* 1987; *43* : 671-81.
4. Knorr-Held L, Besag J. Modelling risk from a disease in time and space. *Stat Med* 1998; *17* : 2045-60.
5. Waller L A, Carlin B P, Xia H, Gelfand AE. Hierarchical spatiotemporal mapping of disease rates. *J Am Stat Assoc* 1997; *92* : 607-17.
6. Bernardinelli L, Clayton D G, Pascutto C, Montomoli C, Ghislandi M, Songini M. Bayesian analysis of space-time variation in disease risk. *Stat Med* 1995; *14* : 2433-43.
7. Daiane LDR, Maria do Gullaci GCB, Edson ZM. Spatio-temporal patterns of tuberculosis incidence in Ribeirao Preto, State of Sao Paulo, southest Brazil, and their relationship with social vulnerability : a Bayesian analysis. *Rev Soc Bra Med Trop* 2012; *45* : 607-15.
8. Venkatesan P, Srinivasan R. Bayesian model of HIV/AIDS in India: A spatial analysis. *Appl Bayesian Stat Anal* 2008; *8* : 51-6.
9. Venkatesan P, Srinivasan R. Bayesian intrinsic conditional autoregressive random effects model of tuberculosis: an application. *Recent Trends Stat Comput Appl* 2011; *11* : 195-200.
10. Venkatesan P, Srinivasan R, Dharuman C. Bayesian conditional autoregressive model for mapping tuberculosis prevalence in India. *Int J Pharm Stud Res* 2012; 3 : 1-3.
11. International Institute for Population Sciences. National Family Health Survey Reports. Available from: *http://www.rchiips.org/NFHS/index.shtml*, accessed on May 22, 2015.
12. Banerjee S, Carlin B, Gelfand AE. *Hierarchical modeling and analysis for spatial data.* Boca Raton: Chapman & Hall; 2004.
13. Lawson AB, Browne WJ, Vidal Rodeiro CL. *Disease mapping with WinBUGS and MLwiN*, 2nd ed. West Sussex, England: John Wiley & Sons Ltd; 2003.