Spatial and Spatio-temporal Epidemiology 20 (2017) 27–34

Contents lists available at ScienceDirect

Spatial and Spatio-temporal Epidemiology

journal homepage: www.elsevier.com/locate/sste

Original Research

Relative risk for HIV in India – An estimate using conditional auto-regressive models with Bayesian approach

Chandrasekaran Kandhasamy[1,*], Kaushik Ghosh[1]

Department of Mathematical Sciences, University of Nevada Las Vegas, Las Vegas, NV, USA

ARTICLE INFO

Article history:
Received 14 December 2014
Revised 14 September 2016
Accepted 1 January 2017
Available online 5 January 2017

Keywords:
HIV
Relative risk
Disease mapping
Spatial model
CAR models
Bayesian approach

ABSTRACT

Indian states are currently classified into HIV-risk categories based on the observed prevalence counts, percentage of infected attendees in antenatal clinics, and percentage of infected high-risk individuals. This method, however, does not account for the spatial dependence among the states nor does it provide any measure of statistical uncertainty.

We provide an alternative model-based approach to address these issues. Our method uses Poisson log-normal models having various conditional autoregressive structures with neighborhood-based and distance-based weight matrices and incorporates all available covariate information. We use R and WinBugs software to fit these models to the 2011 HIV data. Based on the Deviance Information Criterion, the convolution model using distance-based weight matrix and covariate information on female sex workers, literacy rate and intravenous drug users is found to have the best fit.

The relative risk of HIV for the various states is estimated using the best model and the states are then classified into the risk categories based on these estimated values. An HIV risk map of India is constructed based on these results. The choice of the final model suggests that an HIV control strategy which focuses on the female sex workers, intravenous drug users and literacy rate would be most effective.

1. Introduction

India's progress in meeting the commitment of realizing zero new HIV infections is encouraging from the fact that from the year 2000 to 2011, the decreased from 0.27 million to 0.13 million. However, based on the National AIDS Control Organization Annual Report 2011–12 (NACO, 2012), India still has the third highest number of people living with HIV/AIDS. According to this report, in the year 2011, there was an estimated 2.09 million people with HIV/AIDS living in India, resulting in an adult prevalence rate of 0.27%. Additionally, the number of AIDS-related deaths reported was 147,729.

HIV is not curable and these patients are more likely to get opportunistic infections such as Tuberculosis. The HIV/AIDS patients are more prone to social stigma, discrimination, and at times, collective denial are as central to the global AIDS challenge as the disease itself (Mann, 1987). It is thus imperative for health planners and administrators to get an accurate profile of the risk of HIV infection.

The current practice of classification of Indian states into HIV prevalence risk categories is based on the observed prevalence counts, the percentage of attendees in antenatal clinics with the disease, and the percentage of high-risk individuals with the disease. A model-based risk profile will be more precise and useful for health planners to focus additional effort on places where there is high risk

* Corresponding author.
E-mail address: chandrasekarank@nirt.res.in (C. Kandhasamy).
[1] Visiting from Statistics Department, National Institute for Medical Research (Indian Council of Medical Research), Madurai, Tamilnadu, India.

http://dx.doi.org/10.1016/j.sste.2017.01.001
1877-5845/

on a priority basis for a major disease like HIV in India. Also, a lack of model-based approach results in an inability to establish a degree of confidence in the accuracy of risk estimates, and for all intents and purposes, may represent a very inaccurate picture of the disease. In order to achieve the goal of zero new HIV infections, a model-based approach with relevant covariates to identify the areas where there is excess risk will be quite useful for health administrators for efficient planning.

The variations found in the observed HIV across the states and union territories might have been caused by unobserved covariates or may be due to clustered observations. In such cases, the observed values are said to have over-dispersion or extra heterogeneity. Such variations in the HIV incidence may be explained by spatial proximity between regions or inter-differences in covariates. Thus, over-dispersion is always present in the spatially referenced data. The conventional Poisson model is not able to account for this heterogeneity (Lawson, 2008).

In this paper, we estimate the HIV risk profile for all the Indian states and union territories using various types of Conditional Auto Regressive (CAR) models incorporating potential covariates such as proportion of female sex workers, proportion of intravenous drug users, and literacy rate, whereby we account for the extra variation present in the Poisson model. The paper proceeds as follows. In Section 2, we first describe the data on HIV and relevant covariates. We then discuss in detail the various models that we have applied to our data and methods used to fit these models. In Section 3, we present the results of fitting these models and assess the model fits. In Section 4, we briefly discuss the results.

2. Methods

The National AIDS Control Organization of India (NACO) is responsible for implementing the National AIDS Control Program in India. Under this program, it has been supporting many Integrated Counseling and Testing Centers across the country in the form of staff and logistical support. Facility for HIV testing is available free of cost for all, with extra attention given to at-risk groups including intravenous drug users, homosexual males, female sex workers, patients attending sexually transmitted infection clinics, single male migrants, long distance truckers and pregnant women attending antenatal clinics.

We have used the data on HIV for the year 2011 from the 2011 NACO Annual Report (NACO, 2012). NACO conducts annual HIV Sentinel Surveillance in designated sites all over India to monitor HIV trends in various risk groups of population in conjunction with the National Institute of Health and Family Welfare, New Delhi, and National Institute of Medical Statistics-Indian Council of Medical Research (ICMR), New Delhi. For this purpose, they have used the updated Spectrum 4.53 Beta 19 tool, recommended by the Global Reference Group on Estimations, Projections and Modeling for their estimation, which was described in the technical report India HIV Estimates 2010 (Government of India, 2010).

Data on two covariates: (i) proportion of Female Sex Workers (FSW) and (ii) proportion of Intravenous Drug

Fig. 1. India administrative map. Image courtesy Wikipedia.

Users (IDU) were obtained for each administrative region (State or Union Territory) from the technical report India HIV Estimates 2010 (Government of India, 2010) published by the Government of India. Data on a third covariate (iii) literacy rates (LIT) were obtained from the Provisional Population Totals-India, 2011 Census (Government of India, 2011). Note that for the 2011 census, the definition of literacy is a person aged seven or older being able to read or write in any Indian language.

During 2011, there were 28 states and 6 union territories in India, giving rise to a total of $n = 34$ administrative regions. For simplicity, we will not make any distinction between union territories and states and call each of these administrative regions as states. A map of India showing the states is given in Fig. 1.

Data on HIV as well as the three covariates mentioned earlier were obtained for all these states for the year 2011. This data was used for model building and further analyzes.

In this paper we have considered states as the geographical units and used state-level data for modeling purposes. Since some of the states can be quite large, it might be advisable to work with smaller geographical units. However, due to lack of availability of data at smaller unit levels, we have proceeded with state-level modeling.

For region k $(k = 1, \ldots, n)$, let us denote the age- and sex-adjusted expected count by E_k, the vector of covariates by $\mathbf{X}_k$ and the corresponding random effect by ϕ_k. Denote $\mathbf{Y} = (Y_1, \ldots, Y_n)^T$, $\mathbf{E} = (E_1, \ldots, E_n)^T$, $\mathbf{X} = (\mathbf{X}_1^T, \ldots, \mathbf{X}_n^T)^T$ and $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_n)^T$. We will model the incidence $\mathbf{Y}$ as a function of covariates $\mathbf{X}$ and random effects $\boldsymbol{\phi}$. The random effects will be used to account for the spatial