

Modeling spatial and temporal variability by Bayesian multilevel model

Xiaoxiao Song^{*1,2}, Yan Li^{1,2}, Wei Liu^{1,2}, Le Cai^{1,2}, Wenlong CUI^{1,2} and mingyue wang^{1,2}

¹School of Public Health, Kunming Medical University, Kunming, China; ²Yunnan Provincial Collaborative Innovation Center for Public Health and Disease Prevention and Control, Kunming, China

Objective

The purpose of this article was to quantitative analyses the spatial variability and temporal variability of influenza like illness (ILI) by a three-level Poisson model, which means to explain the spatial and temporal level effects by introducing the random effects.

Introduction

The early detection of outbreaks of diseases is one of the most challenging objectives of epidemiological surveillance systems. In order to achieve this goal, the primary foundation is using those big surveillance data for understanding and controlling the spatiotemporal variability of disease through populations. Typically, public health's surveillance system would generate data with the big data characteristics of high volume, velocity, and variety. One common question of big data analysis is most of the data have the multilevel or hierarchy structure, in other word the big data are non-independent. Traditional multilevel or hierarchical model can only deal with 2 or 3 hierarchical data structure, which bound health big data further research for modeling, forecast and early-warning in the public health surveillance, in particular involving complex spatial and temporal variability of Infectious Diseases in the reality.

Methods

All the data based the ISSC project from April 1 2012 through March 31 2014 in the China. We adopted Markov Chain Monte Carlo algorithm (MCMC) in Bayesian hierarchical (multilevel) model, which means to explain the spatial and temporal level effects by introducing the random effects. In order to calculate the geographical variations and temporal variation of ILI cases during two years surveillance, we constructed spatial and temporal model of three levels, which was day-in-months → months-in-two-year → Monitoring Units (Fig-1). Level one was repeated measures within every month, which was referred as day-in-months and the maximum value was 31 days. Level two was the variation tendency of months which was 24 months. Level three was the effect of spatial distribution of monitoring units, which took the spatial heterogeneity into account rather than dependence. This model was then adopted to evaluate and improve the early warning capacity of syndromic surveillance.

Results

We adopted multilevel spatio-temporal model (day-in-months → months-in-two-year → Monitoring Units) to analyze the points data collected from 2 counties in China, including two hospitals at county level, 15 central hospital at township level and 152 health care units in the villages. The analysis of totally 108163 pieces of point data on ILI case indicated there are significant spatial and temporal variation among these cases. Among two thirds of the variation attributes to the difference of geographical locations of these monitoring sites. The remaining one third of the variation attributes to the time dimensions, such as seasonal effect.

Conclusions

The variation of monitoring data collected from health care units mainly attributes to the difference of geographic locations for

monitoring sites, yet only one third of the change attribute to the time change, such as seasons, holidays and festivals. Therefore, it is critical to select the location of monitoring site, which is more rational to select the hotspots with representative characters rather than try to cover the whole monitoring area.

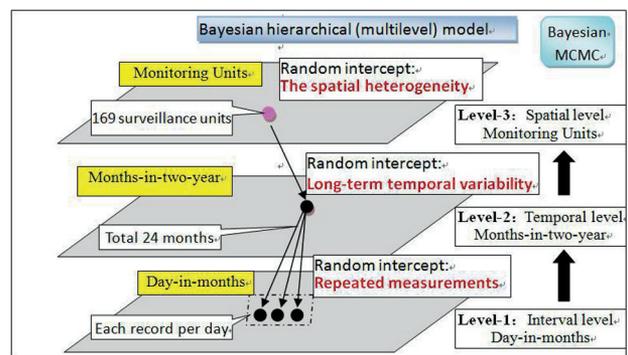


Fig-1 data analysis framework for Bayesian hierarchical (multilevel) model.

Keywords

Bayesian multilevel model; spatial variability; temporal variability; Influenza like illness

References

1. Goldstein H. Multilevel statistical model[M]. 922. John Wiley & Sons, 2011
2. Leckie, G. and Charlton, C. (2013). runmlwin - A Program to Run the MLwiN Multilevel Modelling Software from within Stata. Journal of Statistical Software, 52 (11),1-40.

*Xiaoxiao Song

E-mail: chinaxx@gmail.com

