# Spatio-temporal epidemiologic mapping, modeling and prediction of tuberculosis incidence rate in northeast of Iran

Abbas Abbasi-Ghahramanloo[1], Saeid Safiri[2], Ali Gholami[3], Yaser Yousefpoor[4], Saleh Babazadeh[5], Javad Torkamannejad-Sabzevari*[6]

[1] *Department of Epidemiology, School of Public Health, Iran University of Medical Sciences, Tehran, Iran*

[2] *Department of Public Health, School of Nursing and Midwifery, Maragheh University of Medical Sciences, Maragheh, Iran*

[3] *Department of Public Health, School of Public Health, Neyshabur University of Medical Sciences, Neyshabur, Iran*

[4] *Department of Nanotechnology, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran*

[5] *Department of Global Health Management and Policy Tulane, School of Public Health and Tropical Medicine, Louisiana, USA*

[6] *Department of Public Health, School of Public Health, Mashhad University of Medical Sciences, Mashhad, Iran*

**Abstract**

*Introduction:* Tuberculosis is a major public health problem in the world. The aim of the present study was to determine the incidence rate of tuberculosis through modeling and predict the disease incidence rate using spatio-temporal Kriging method in three endemic regions (Kashmar, Khalilabad and Bardeskan) in the northeast of Iran.

*Methods:* This cross-sectional study was conducted during 2007-2012. The diagnosis of tuberculosis in patients who had signs and symptoms of infection was confirmed using sputum smear test. According to fitted variogram function and Kriging method, we predicted tuberculosis incidence for all spatial and temporal points of regions in the study area.

*Results:* Between 2007 and 2012, 155 cases of tuberculosis were observed. Among all patients, 70 (45.0%) were men and 94.0% were rural residents. Mean age of patients with tuberculosis was 64 years and 151 (97.5%) of patients were 55 years old and above. Based on the geographical coordinates, we identified the place of residence for each patient.

*Conclusion:* Our study showed that the downward trend of the incidence rate of tuberculosis indicates good but inadequate progress with tuberculosis control. The findings of this study can be used for planning and evaluating interventions by considering the risk factors of tuberculosis infection in the northeast of Iran.

## Introduction

Tuberculosis has probably affected human's life through most of their history. Tuberculosis is the leading cause of death from a curable infectious disease worldwide despite the existence of effective and affordable chemotherapy from more than 50 years ago.[1-3]

Tuberculosis remains one of the most destructing and widespread disease in the world and have high priority in international statistics mostly due to killing of young adults. According to statistics, over 80% of the burden of tuberculosis, as measured in terms of disability-adjusted life years lost (DALYs), is due to premature death rather than illness. Despite the availability of effective and useful treatment, this infection killed 1.3 million people in 2012.[4-6]

According to the World Bank data, the greatest rate of dispersion of tuberculosis in Iran is in the eastern provinces such as Sistan

and Balouchestan and Khorasan Razavi, Iran (Figure 1, red regions) which are bordered by Afghanistan and Pakistan.[7] Khorasan Razavi province in Iran is recognized as a risky region regarding this disease.[8] Therefore, preventive measures to identify and reduce the incidence rate of tuberculosis in this area are very important.

Statistical methods nowadays have a significant role in many studies in the fields of healthcare, biomedical, clinical medicine, epidemiology, and modeling of disease in order to provide analytical strategies to identify risk factors, detect mechanisms and identify the regions with high risk of disease spread.[9]

It is undeniable that the description of the disease relies on some variables like time, place and factors associated with the disease. In the past two decades, we have witnessed an increasing interest in the use of space-time models for a wide range of environmental and epidemiological problems. The spatio-temporal epidemiology is one of the most important tools for epidemiologists to detect, monitor and predict public health disease patterns. Identifying key nodes in the spread of an infection, defined by time spent at unique locations, can help us understand past infectious episodes, and predict future developments.

In this paper, we present a spatio-temporal model for tuberculosis incidence rate based on best linear unbiased spatio-temporal prediction method or Kriging method.[10] Kriging is a technique of making best linear unbiased estimates of regionalized variables at unsampled locations and time points using the hypothesis of stationary and structural properties of the covariance and the initial set of data values.[10] Examples include efforts to investigate spatial analysis of tuberculosis[11,12] and spatio-temporal modeling of salmonella surveillance,[13] and spatio-temporal patterns of campylobacter colonization.[14] It is worth noting that most of the disease-based modeling techniques employed in many researches had used classical statistical methods or clustering and spatial regression models. Whereas, in this study, we employed spatio-temporal Kriging that can be applied to model and predict disease incidence in each coordinate of areas at each arbitrary time. In this paper, we estimated the incidence rate of tuberculosis in three endemic regions (Kashmar, Khalilabad and Bardeskan) in the northeast of Iran for all coordinates of these three regions over all seasons in 2013.

## Methods

This was an observational longitudinal ecological study conducted during 2007-2012 in three regions (Kashmar, Khalilabad and Bardeskan) of Khorasan Razavi province of Iran. This study was approved by the committee of ethics under the number of 910675 in Mashhad University of Medical Sciences, Mashhad, Iran. The name and contact information of all patients presented to the medical centers affiliated to Mashhad University of Medical Sciences were extracted from the tuberculosis registry.
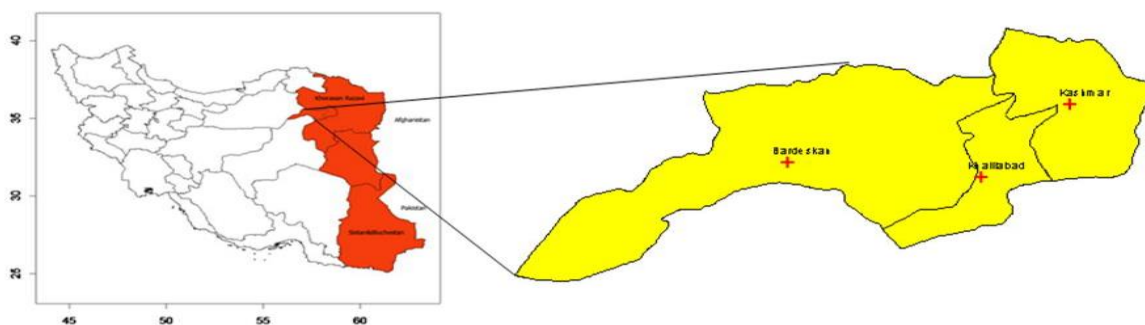


**Figure 1.** Iran map (red regions have greatest rate of dispersion of tuberculosis) and selected regions show the under study area

Patients suspected of having tuberculosis (who had signs and symptoms of tuberculosis) were diagnosed based on the result of the sputum smear test and pathological examinations by the physicians of the health centers or private clinics and were referred to tuberculosis registry centers for medical treatment. All the data regarding 155 patients with tuberculosis who were suffering from various forms of tuberculosis were collected from 41 areas in the study regions during 6 years.

*Spatio-temporal variogram and Kriging:* In view of spatio-temporal epidemiology, disease-specific variogram parameters could be used to infer variation of disease across the time and space. In simple terms, a variogram can be defined as a key function in geostatistics that describes spatial and/or temporal patterns of the observed phenomenon. It has a long and exhaustive history in scientific geostatistical studies.[10] Assuming a value Z $(s_i, t_j)$ of variable Z, observed in a certain sub-district i $(s_i$, represented through a geometric central point) for time $(t_j)$, this value can be correlated with the incidences observed in previous time periods for the same area, and with incidents observed at neighboring sub-districts during the same or previous time periods. The spatio-temporal variogram for some spatial and temporal points can be characterized using a mean spatio-temporal empirical variogram, computed by averaging the empirical values in each time and space lag by:

$$\hat{\gamma}(h, u) = \frac{1}{2N(h,u)} \sum_{N(h,u)} \left[ z(s_i, t_i) - z(s_j, t_j) \right]^2$$

The set N (h,u) consists of the points that are within spatial distance h and time lag u of each other. The spatio-temporal empirical variogram is commonly represented as a graph that shows the variance behavior $\hat{\gamma}(h, u)$ against the distance or time. Usually and in presence of spatial-temporal dependencies, the variogram initially rises from some point on the y-axis (nugget effect) and reaches a threshold (sill) at a certain

location (defining the range). The spatio-temporal variogram is fitted on empirical variogram and optimum values of nugget, sill and range are specified. To model and predict values in space and time, the Kriging or optimal prediction method is applied based on the variogram function. In particular, the prediction values of Z $(S_0, t_0)$ at unsampled space–time location consider a linear combination based on space–time observations, which are given by:

$$\hat{Z}(s_0, t_0) = [\gamma(s - s_0, t - t_0) + 1m]^T \Gamma^{-1} Z(s, t)$$

According to fitted variogram function and Kriging method, we were able to predict the values of tuberculosis incidence for all spatial and temporal points of regions in study area.

**Results**

Between 2007 and 2012, 155 cases of tuberculosis were observed from patients referred to the microbiology labs of three studied regions (Kashmar, Bardeskan and Khalilabad). Among 155 patients, 70 (45.0%) were men and 94% were rural residents. Our result showed that the mean age of patients with tuberculosis was 64 years and 151 (97.5%) of patients were 55 years old and over. Also 4 patients were under 10 years old, 11 patients were 10-25 years old, 13 patients were 25-40 years old, 18 patients 40-55 were years old, 29 patients were 50-70 years old, 85 patients were 70-85 and 37 patients were over 80 years of age (Figure 2).
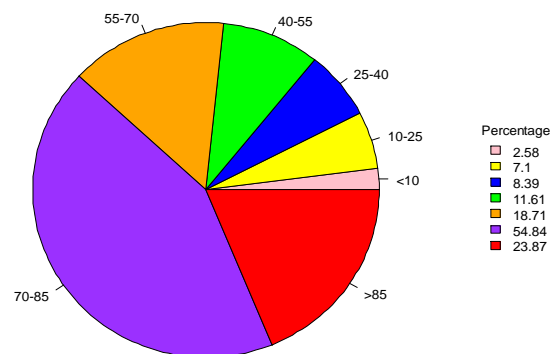


**Figure 2.** Age distribution of the patients with tuberculosis in Kashmar, Bardeskan and Khalilabad from 2007 to 2012

The highest rate of infection belonged to Esmatiyeh, Hafeziye and Abazar with 11, 10 and 9 cases, respectively. While in many rural regions, only one case of tuberculosis infection was reported. In a five year period from 2007 to 2012, the sum values of tuberculosis in each year indicate a downtrend in 2009 to 2012 (Figure 3).

After determining the geographical coordinates and the date of disease occurrence and identifying the patients, spatio-temporal analysis of the incidence of the disease was performed using the R software (version 3.1.0). Based on geographical coordinates, we identified the place of residence for each patient which has

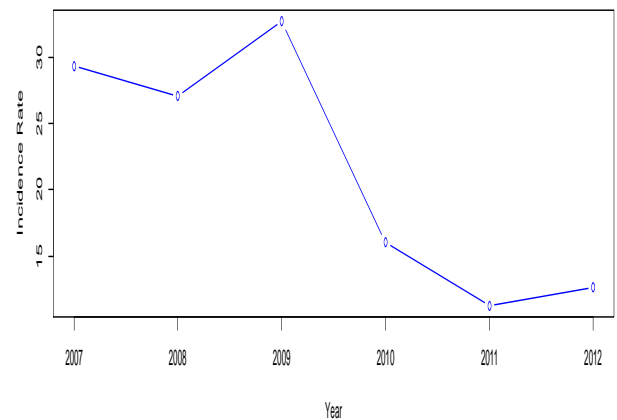been marked in red in figure 4.



**Figure 3.** Incidence rate of tuberculosis in Kashmar, Bardeskan and Khalilabad from 2007 to 2012
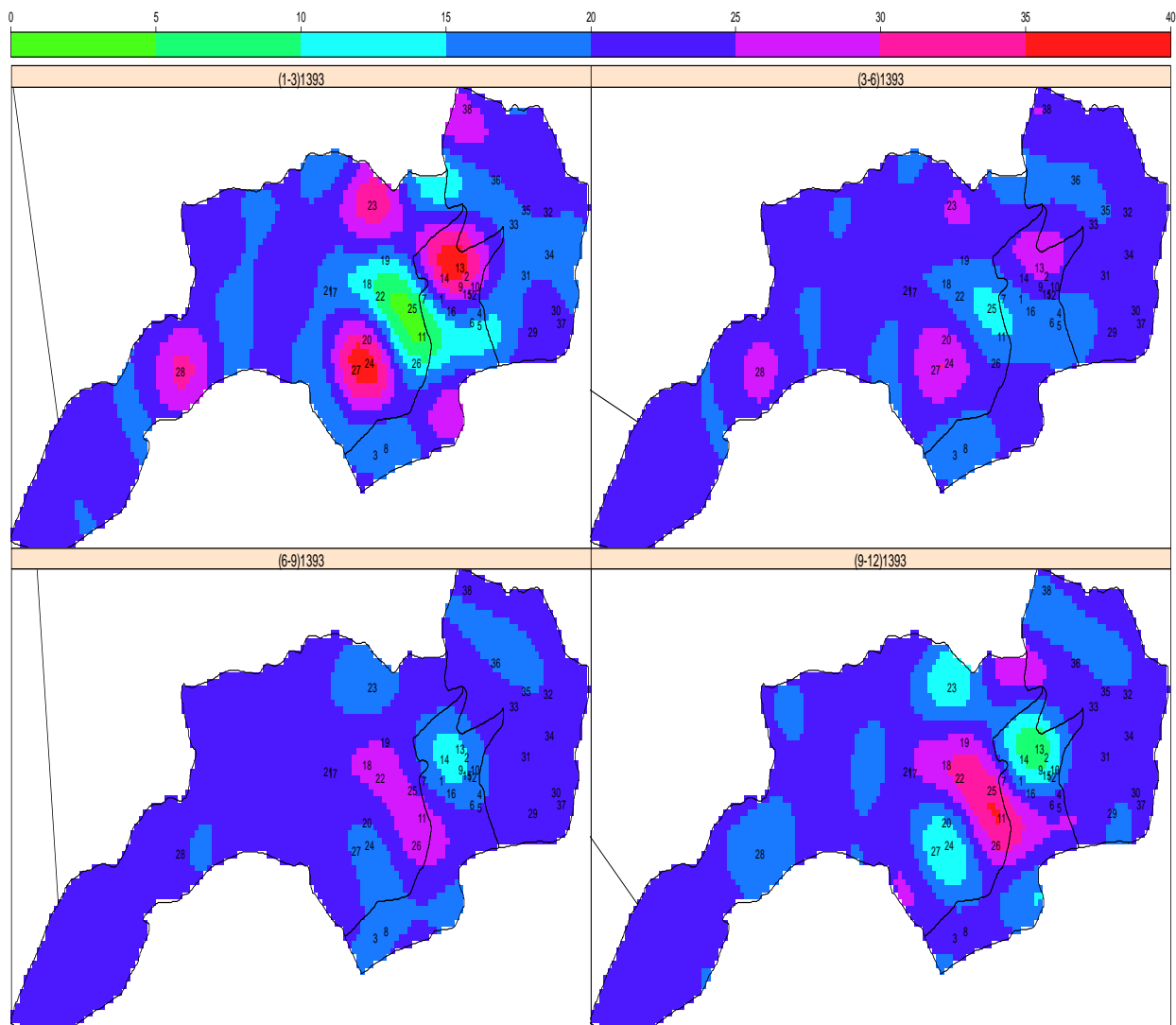


**Figure 4.** Predicted spatio-temporal incidence rate of tuberculosis in Kashmar, Khalil Abad and Bardeskan (purple, pink and red areas are endemic and high-risk)

## Discussion

The average incidence of tuberculosis was 22 cases/100000 inhabitants in three studied regions (Kashmar, Khalilabad and Bardeskan) from 2007 to 2012. Moosazadeh et al.[15] indicated that during 2005-2011, 63568 cases of tuberculosis were recorded in Iran. Based on the results of this study which used a real data, incidence rate of tuberculosis varied from 12.15 in 2005 to 13.78 in 2011.

A population-based study in Golestan, northeast of Iran showed that the incidence of tuberculosis in 2005 was 20.88 per 100000.[16] Rahimi et al. showed that the incidence rate of tuberculosis in West Azerbaijan, Iran was 8.7 per 100000 people in 2006 and 9.91 per 100000 in 2010.[17] The results of a review of the literature indicated that Sistan and Balouchestan and Golestan had the highest incidence and prevalence of tuberculosis among all provinces in Iran over the last 45 years.[18] According to the latest data from the Ministry of Health of Iran, the incidence of tuberculosis in Zabol and Zahedan were reported 109.7 and 36.6 per 100000 populations, respectively.[8] These comparisons indicated that the incidence of tuberculosis in Kashmar, Khalilabad and Bardeskan was higher than the national rate, but tuberculosis is endemic in the east of Iran. The highest incidence rate of tuberculosis infection in Khorasan Razavi as in 2009 with 31 cases and decreased with a steep slope and reached to 7 cases in 2012 (Figure 2) which can be due to precautionary and therapeutic actions taken along these years. This shows that the tuberculosis control programs for early detection of the cases were more effective here.

The empirical variogram method is not applicable to prediction, because it has no specific function. Therefore, we need a functional form that fits the empirical variogram. Thus, it can be used in the forecast Kriging. Using the software, we detected that followed model had a minimum square error and was diagnosed as the best fit to the empirical variogram:

$$\gamma(h, u) = \tau_{st}^2 + \sigma_s^2\big(1 - \varphi_s \sin(\varphi_s h)\big) + \sigma_t^2\big(1 - \varphi_u \sin(\varphi_u h)\big) + \sigma_{st}^2\big(1 - \varphi_s \sin(\varphi_s h)\big)\big(1 - \varphi_u \sin(\varphi_u h)\big)$$

The model parameter values were calculated using the software. This function is calculated as follows:

$$\gamma(h, u) = 0.14 + (0.85\big(1 - 4999.05 * sin(-4999.05h)\big) + 0.14\big(1 - 125.48 sin(125.48u)\big) + 2014.58\big(1 - 4999.05 * sin(-4999.05h)\big)\big(1 - 125.48 sin(125.48u)\big)$$

By achieving the function of spatio-temporal changes in the incidence of tuberculosis in three regions, we will be able to predict the incidence of the disease in desired timescales. Using fitted variogram function and Kriging method, the incidence rate of tuberculosis was predicted for each season of 2013. Predicted values in the three studied regions were drawn with geographic information system (GIS) layers that were produced by the spectrum. Figure 5 shows the predicted rate of incidence of tuberculosis in the three regions separately for each season of 2013.
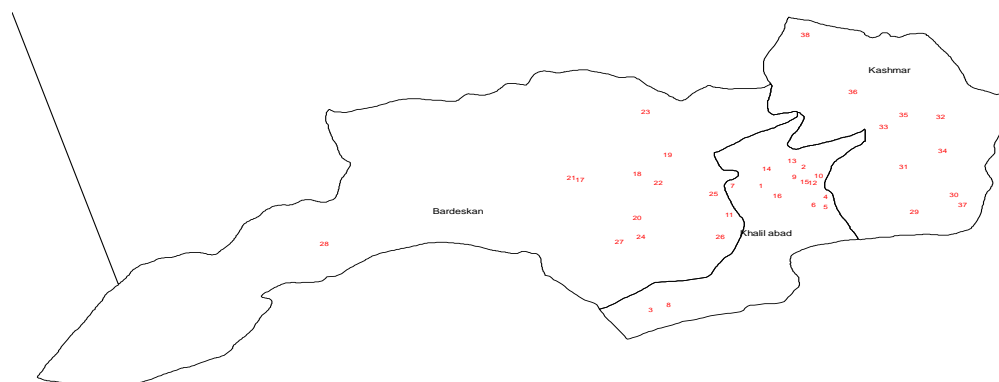


**Figure 5.** Geographical coordinates of identified cases in Kashmar, Bardeskan and Khalilabad from 2007 to 2012

The results showed that the average incidence of tuberculosis in three regions would be 21.14 [95% confidence interval (CI) 21.09-21.18] per 100000. In addition, results indicated the highest incidence of tuberculosis would be in the spring with the average rate of 28.37 per 100000. Finding of spatio-temporal prediction values showed a higher rate of incidence of tuberculosis in 2013 than in the years 2005 to 2012. In some villages of the three mentioned regions, the number of detected cases is higher than at risk population. This suggests the need for special attention in order to identify cases and implementation of interventions in these regions.

*Limitations:* This study showed some epidemiological features of tuberculosis in the Kashmar, Khalilabad and Bardeskan. Similar to other cross-sectional studies incidence rate could not be estimated accurately yet. Longitudinal studies are suggested in order to determine and monitor the incidence rate of tuberculosis and its correlations with other variables.

## Conclusion

Our study showed that the incidence rate of tuberculosis is generally high in Khorasan Razavi. The main factors increasing tuberculosis in this province in the last decade are believed to be proximity with Afghanistan and plenty of Afghan immigrants and their traffic in the province, given that Afghanistan has a high burden of tuberculosis in the world.[19] Population movements from these countries to Khorasan Razavi appear to be the main challenge to tuberculosis control. Nevertheless, the downward trend of the incidence rate of tuberculosis indicates good but inadequate progress with tuberculosis control. The findings of this study can be used for planning and evaluating interventions by considering the risk factors of tuberculosis infection in Khorasan Razavi.

## Conflict of Interest
Authors have no conflict of interest.

## Ethical Approval
This study was approved by the ethics committee under the number of 910675 in the Mashhad University of Medical Sciences.

## References

1. Holloway KL, Henneberg RJ, de Barros LM, Henneberg M. Evolution of human tuberculosis: a systematic review and meta-analysis of paleopathological evidence. Homo 2011; 62(6): 402-58. DOI: 10.1016/j.jchb.2011.10.001
2. Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, et al. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nat Genet 2013; 45(10): 1176-82. DOI: 10.1038/ng.2744
3. World Health Organization. The world health report 2004 - changing history. Geneva, Switzerland: WHO; 2004.
4. Kochi A. The global tuberculosis situation and the new control strategy of the World Health Organization. 1991. Bull World Health Organ 2001; 79(1): 71-5. DOI: 10.1016/0041-3879(91)90017-M
5. Corbett EL, Watt CJ, Walker N, Maher D, Williams BG, Raviglione MC, et al. The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. Arch Intern Med 2003; 163(9): 1009-21. DOI: 10.1001/archinte.163.9.1009
6. Dye C, Watt CJ, Bleed DM, Hosseini SM, Raviglione MC. Evolution of tuberculosis control and prospects for reducing tuberculosis incidence, prevalence, and deaths globally. JAMA 2005; 293(22): 2767-75. DOI: 10.1001/jama.293.22.2767
7. World Health Organization Global tuberculosis

report 2013. Geneva, Switzerland: WHO; 2013.

8. Mafi E, Yousefi Bohlouli Ahmadi GM. Comparative pathology of the space dispersion rate of oriental sore and tuberculosis in Mashhad. Journal of Geography and Regional Development 2009; 13: 301-21. [In Persian].

9. Best N, Wakefield J, Best N, Briggs D. Spatial epidemiology: Methods and applications. 1st ed. Oxford, UK: Oxford University Press; 2001.

10. Cressie N. Statistics for spatial data. Hoboken, NJ: Wiley; 1993.

11. Kelly GE, McGrath GE, More SJ. Estimating the extent of spatial association of Mycobacterium bovis infection in badgers in Ireland. Epidemiol Infect 2010; 138(2): 270-9. DOI: 10.1017/S0950268809990471

12. Kelly GE, More SJ. Spatial clustering of TB-infected cattle herds prior to and following proactive badger removal. Epidemiol Infect 2011; 139(8): 1220-9. DOI: 10.1017/S0950268810002323

13. Domingues AR, Vieira AR, Hendriksen RS, Pulsrikarn C, Aarestrup FM. Spatio-temporal analysis of Salmonella surveillance data in Thailand. Epidemiol Infect 2014; 142(8): 1614-24. DOI: 10.1017/S095026881300215X

14. Chowdhury S, Themudo GE, Sandberg M, Ersboll AK. Spatio-temporal patterns of Campylobacter colonization in Danish broilers. Epidemiol Infect 2013; 141(5): 997-1008. DOI: 10.1017/S0950268812001446

15. Moosazadeh M, Nasehi M, Bahrampour A, Khanjani N, Sharafi S, Ahmadi S. Forecasting tuberculosis incidence in iran using box-jenkins models. Iran Red Crescent Med J 2014; 16(5): e11779. DOI: 10.5812/ircmj.11779

16. Rafiee S, Besharat S, Jabbari A, Golalipour F, Nasermoaadeli A. Epidemiology of tuberculosis in northeast of Iran: A population-based study. Iran J Med Sci 2009; 34(3): 193-7.

17. Rahimi FA, Farzianpour F, Tavana A, Rasouli J, Hosseini S. The 10-year trend of TB rate in West Azerbaijan province, Iran from 2001 to 2010. Iran J Public Health 2014; 43(6): 778-86

18. Metanat M, Sharifi-Mood B, Alavi-Naini R, Aminianfar M. The epidemiology of tuberculosis in recent years: Reviewing the status in south-eastern Iran. Zahedan J Res Med Sci 2012; 13(9): 1-7.

19. Jafari-Koshki T, Arsang-Jang S, Raei M. Applying spatiotemporal models to study risk of smear-positive tuberculosis in Iran, 2001-2012. Int J Tuberc Lung Dis 2015; 19(4): 469-74. DOI: 10.5588/ijtld.14.0459