Mapping the Relative Risk of Congenital Hypothyroidism Incidence via Spatial Zero-Inflated Poisson Model in Guilan Province, Iran

Abstract

Background: Congenital hypothyroidism (CH) is one of the most prevalent preventable causes of mental retardation. Studies show that the incidence rate of CH is very high in Iran. Disease mapping is a tool for visually expressing the frequency, incidence, or relative risk of illness. The present study aimed to model CH counts considering the effects of the neighborhood in towns and perform mapping based on the relative risk. **Methods:** In this historical cohort study, data of all neonates diagnosed with CH with TSH level ≥5 mIU/L between March 21, 2017, and March 20, 2018, in health centers in Guilan, Iran were used. The number of neonates with CH was zero in most towns of Guilan Province. The Bayesian spatial zero-inflated Poisson (ZIP) regression model was employed to investigate the effect of the town's neighborhood on the relative risk of CH incidence. Then, the map of the posterior mean of the relative risk for CH incidence was provided. The analysis was performed using OpenBUGS and Arc GIS software programs. **Results:** The relative risk of CH incidence was high in the West of Guilan. Moreover, the goodness-of-fit criterion indicated that it is more appropriate to fit the Bayesian spatial ZIP model to these data than the common model. **Conclusions:** Considering the high relative risk of CH in the Western towns of Guilan Province, it is better to check important risk factors in this region.

Keywords: *Bayes theorem, congenital hypothyroidism, neonates, Poisson distribution, spatial analysis*

Bahare Gholami
Chaboki,
Manijeh Tabrizi[1],
Maryam Heydarpour
Meymeh[2],
Hojjat Alaei[3],
Alireza Akbarzadeh
Baghban[4]

Department of Biostatistics, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran, [1]Pediatric Diseases Research Center, Guilan University of Medical Sciences, Rasht, Iran, [2]English Language Department, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran, [3]Department of Biostatistics, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran, [4]Department of Biostatistics, Proteomics Research Center, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Introduction

One of the most prevalent preventable causes of mental retardation is congenital hypothyroidism (CH), defined as the lack of thyroid hormones from birth. There are two main types of CH: transient and permanent. Permanent CH is the chronic type of CH, requiring lifelong treatment, while transient CH can be treated in the first few months or years of life. Most infants with CH show no apparent symptoms or signs of the disease in the early months after birth. Therefore, screening programs of CH with simple laboratory tests have been established for the better management of the disorder and prevention of abnormal brain development and function. The first newborn screening program for CH was developed in Quebec, Canada, in the early 1970s, and it is currently a routine program worldwide.[1]

Results of previous studies indicate that the incidence of CH is one per 4,500 live births in the US, one per 3,000 live births in Europe, one per 5,700 live births in Japan, and one per 800 live births in Greece. The incidence of CH is very high in Iran, reported in the literature to be one per 1,000 live births.[1-4]

For the first time in Iran, Ordookhani *et al*. (1987) performed newborn CH screening. They collected samples of 20,107 newborns in Tehran and Damavand. Cord blood samples of newborns collected immediately after birth and newborns with cord TSH ≥20 µU/mL were recalled after 7–14 days of life, and then serum TSH and T4 and also urinary iodine were measured. Eventually, the incidence rate of CH was estimated at 1:914 births.[5] The universal screening practice commenced in 2005 in the healthcare programs of Iran.[6]

Several studies have been conducted on CH worldwide. Some of these were case-control retrospective studies and their data were based on those available in CH screening programs in public health centers. Samples included neonates whose medical records were collected.[7] Some other studies were

Address for correspondence:
Prof. Alireza Akbarzadeh Baghban,
Ghods Square, Darband Avenue, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
E-mail: akbarzad@gmail.com

cohort studies collecting the neonatal medical data of CH screening programs. The major purpose of these studies was to determine the risk factors for neonates' CH from among demographic, environmental, and medical factors. Findings of some studies demonstrate sex, very low and very high neonatal weight, maternal age, the kindred relationship of parents, taking special medications during pregnancy, birth season, jaundice at birth, and deficiency or high levels of iodine as risk factors of CH.[1,8,9]

The result of studies conducted in Iran indicates that the prevalence of CH is relatively higher in this country than others. Several studies have been conducted in most provinces of Iran, including Isfahan, Fars, Yazd, Tehran, Kurdistan, and Guilan, reporting the incidence ratio of CH to be 1:370, 1:1433, 1:1608, 1:1000, 1:400, and 1:542 per live birth, respectively. These studies investigated some risk factors of CH such as the type of delivery; weight, length, and head circumference at birth; iodine deficiency; mother's; birth season; sex; and postdate delivery.[1,9-14]

It is noteworthy that, in these studies, the Poisson regression model has been employed to analyze the count data. Variance and mean are equal in this model. However, count data often depart from the Poisson distribution due to the large frequency of extreme observation and excess zero values. Thus, the variance is no longer equal to the mean. In other words, common distributions such as Poisson are not suitable and, therefore, the zero-inflated Poisson (ZIP) regression must be used.[15]

The purpose of the present study was to identify the pattern of CH incidence relative risk among the 51 towns of Guilan Province, Iran. The response variable was the number of neonates with CH in towns between March 21, 2017, and March 20, 2018. The number of neonates with CH was zero in most towns and, due to the adjacency of towns and their specific geographical location in Guilan, it was expected that the CH data would be correlated. As a result, the spatial ZIP regression model using the Bayesian approach was adopted to investigate the effect of the town's neighborhood on the relative risk of CH incidence. After computing the relative risk, CH was mapped based on the relative risk.

Methods

Data in this study

This historical cohort study was approved with an ethics certificate "IR.SBMU.RETECH.REC.1397.1338" at Shahid Beheshti University of Medical Sciences, Iran. The recorded medical data of all neonates diagnosed with CH with the TSH level ≥5 mIU/L between March 21, 2017 and March 20, 2018 in Guilan health centers were used in this study. Inclusion criteria consisted of both categories of transient and permanent. The number of neonates with CH was specified in 51 towns of Guilan. The spatial segmentation of 51 towns was determined based on the 2016 census performed by the Iranian Center for Statistics, based on which there

were 16 townships and 51 towns in this province. The map of Guilan's towns was prepared by Arc GIS version 10.4.1 using longitude and latitude. Moreover, the neighborhood effect was considered in the Bayesian spatial ZIP model, and then mapping was performed for estimated relative risk.

Statistical analysis

A major objective of spatial studies is estimating the disease incidence relative risk in each region, estimated via the maximum likelihood (ML) method. This estimator is the standard incidence ratio $(SIR) \hat{R}_i = \dfrac{Y_i}{E_i}$, observed count of disease divided by its expected count.[16,17] In this study, the observed and expected counts of neonates with CH in each town were represented by Y_i and E_i (i=1,2,...,51), respectively. The expected count of CH was computed by multiplying the count of newborn and provincial incidence ratio.[18] An SIR value >1 indicates an area with a higher risk of disease than the expected risk; inversely, a value of SIR <1 shows a lower risk than expected. The SIR is a useful tool for comparing the relative risk of disease incidence in each region, but it also has a major disadvantage, that is, considering regions independently and not taking into account spatial dependence and the region's neighborhood effects. Furthermore, in cases where regions are small and/or the disease is rare, the SIR will estimate the relative risk misleadingly. Therefore, it is common to estimate disease incidence relative risk using generalized linear models which were introduced by Besag *et al*.[16,17,19]

Since the response variable is based on count observations, a more common model is a spatial Poisson log-linear model as follows:

$$Y_i \sim Poisson(E_i R_i) \qquad i = 1, 2, \ldots, 51$$

$$log(R_i) = x^{'}\beta + \theta_i \tag{1}$$

Where θ_i indicates spatial random effect, $\theta = (\theta_1, \theta_2, \ldots, \theta_{51})$ demonstrate the prior distribution of conditional autoregressive (CAR), $\beta = (\beta_0, \beta_1, \ldots, \beta_k)$ is the vector of coefficients for the vector of covariates, $x^{'} = (1, x_1, \ldots, x_k)$, and k represents the number of covariates. In this study, there are no covariates and $\beta = \beta_0$ is the intercept for the entire towns.[16]

The spatial Poisson model is no longer suitable for count data containing excess zero. Consequently, the spatial ZIP model was adopted for the analysis of these data. Equation (1) was extended as follow:

$$\begin{cases} log(R_i) = x^{'}\beta + \theta_i \\ logit(p_i) = z^{'}\gamma \end{cases} \tag{2}$$

Where p_i is the zero-inflation probability in the ith town, $\gamma = (\gamma_0, \gamma_1, \ldots, \gamma_k)$ is the vector of coefficients for the vector of covariates, $z^{'} = (1, z_1, z_2, \ldots, z_k)$. In this study, there are no covariates and $\gamma = \gamma_0$ is intercept for entire towns.[16,20]

The spatial Poisson model and the spatial ZIP model were employed via the Bayesian approach using the Markov Chain Monte Carlo (MCMC) method in the OpenBUGS 3.2.3 software. The CAR prior distribution was employed for a random effect θ, the prior for the precision parameter in CAR distribution was considered gamma distribution with (0.005, 0.005) in the Poisson model and gamma distribution with (0.0005, 0.0005) in the ZIP model. The normal distribution with (0, 0.00005) was considered for the coefficient of the model. In the Bayesian spatial ZIP model, 100,000 updates were taken with two parallel chains. A burn-in sample of 25,000 was used, and then every 25th sample was kept. In the Bayesian spatial Poisson model, two chains of 40,000 iterations were run; the first 10,000 samples were discarded, and every 10th sample was kept. Convergence was visually evaluated through monitoring the dynamic traces of Gibbs iterations and computing the Gelman-Rubin convergence statistic.

The deviance information criterion (DIC), proposed by Spiegelhalter *et al*. (2002), was adopted for comparing the two models. The general DIC formula is:

$$DIC = pD + \overline{D(\theta)} \quad (3)$$

In which $pD = \overline{D(\theta)} - D(\bar{\theta})$ and $D(\theta) = -2log\left(p\left(y|\theta\right)\right) + c$ (constant).[21] Eventually, the relative risk was mapped using the posterior mean of relative risk and the map of Guilan's towns in this software.

Results

According to the National Organization for Civil Registration of Guilan Province, 29,117 neonates were born between March 21, 2017 and March 20, 2018. Of these, 79 neonates were diagnosed with CH. The CH incidence rate was computed as 27:10,000. The number of neonates with CH was determined in each town of Guilan. Figure 1 illustrates the histogram of the observed counts of CH in towns. Evidently, the frequency of zero is higher than the other counts.

Subsequently, the spatial Poisson model and spatial ZIP model with the Bayesian approach, respectively, were run. The count of CH was predicted in both models for each town. Figure 2 presents the histogram of the predicted count of CH via spatial Poisson and spatial ZIP models using the Bayesian approach. It is clear that about 65% of the CH incidence was predicted as zero in both models. According to Figures 1 and 2, the observed counts were almost close to the predicted counts.

The map of the posterior relative risk of CH incidence based on the spatial Poisson model is depicted in Figure 3a. The relative risk of CH incidence was predicted to be high in the West of the province. Moreover, Figure 3b presents the map for the predicted relative risk of CH incidence according to the spatial ZIP model. In these figures, one can observe that the predicted relative risk based on the spatial Poisson model is similar to the predicted relative risk according to the spatial ZIP model. The estimated RR between 0.8 and 1 had the highest frequency in both models.

The distribution of spatial random effects in Figure 4 shows a clustering structure by towns, where the Western towns of the province have the highest effects. Dark colors in Figures 3 and 4 demonstrate towns with a high propensity for the relative risk of CH incidence.

Table 1 gives the goodness-of-fit criterion for comparing spatial Poisson and ZIP models. The Bayesian spatial Poisson model displayed an inferior fit than the Bayesian spatial ZIP model with a $DIC_{Poisson} = 129.2$ versus $DIC_{ZIP} = 124.6$.

Discussion

In this article, we implemented the spatial Poisson and spatial ZIP model using the Bayesian approach in OpenBUGS software. In these models, the spatial effect was considered by the CAR model. The relative risk of CH incidence was mapped via the predicted relative risk in both spatial Poisson and ZIP models. Based on the findings, the map of the posterior relative risk of CH incidence via spatial Poisson model was similar to the map based on the spatial ZIP model, and in both, the relative risk of CH incidence was high in the West of Guilan. Despite the similar results in the two models, the goodness-of-fit criterion for comparing models indicated that the Bayesian spatial ZIP model fitted these data better than the Bayesian spatial Poisson model.
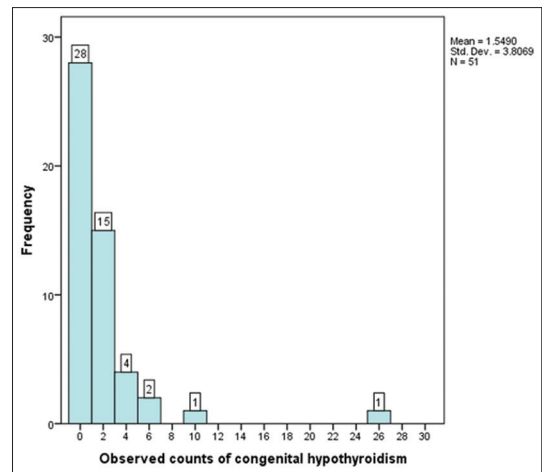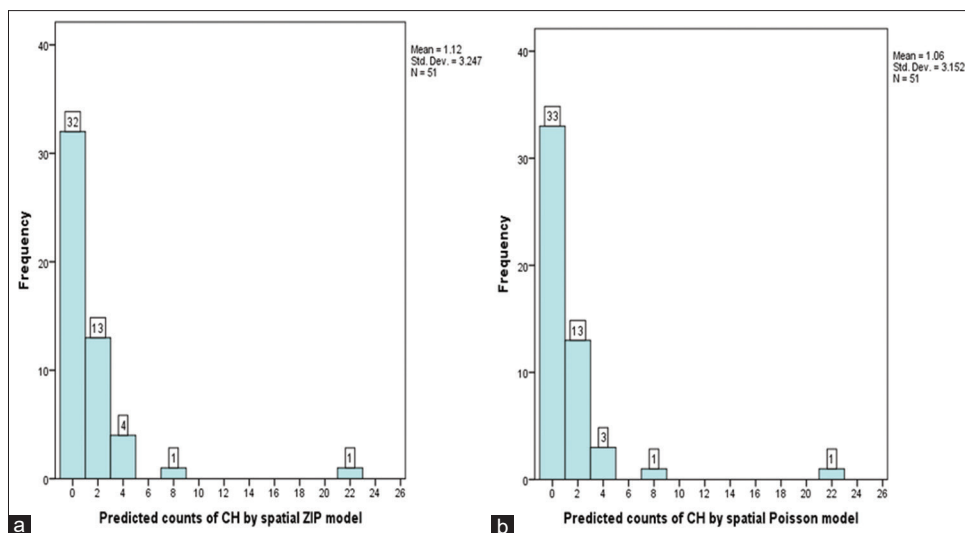

Figure 1: Histogram of observed counts of congenital hypothyroidism (CH) in Guilan province in 1396

Table 1: Goodness-of-fit criterion for comparing spatial Poisson and ZIP Model

Model	Dbar	Dhat	DIC	pD
Poisson	121.8	114.4	129.2	7.407
ZIP	124	123.3	124.6	0.6308

ZIP: Zero-inflated Poisson

Figure 2: Histogram of the predicted count of CH via zero-inflated Poisson (ZIP) and spatial Poisson models. (a) Predicted counts of CH by spatial ZIP. (b) Predicted counts of CH by spatial Poisson
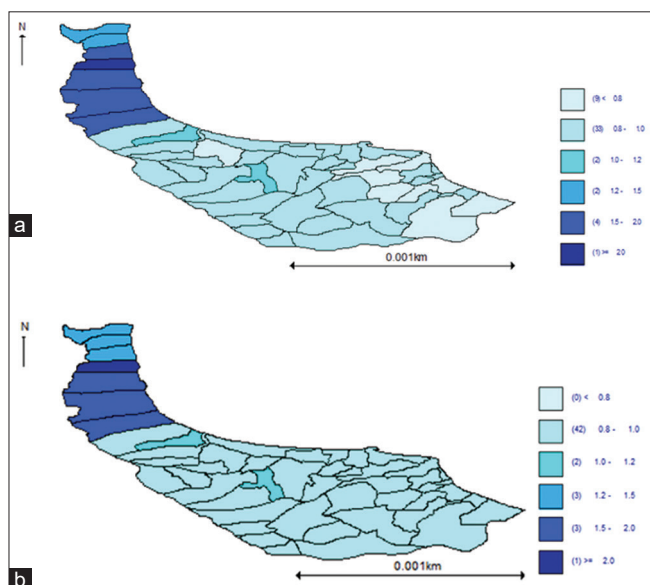
Figure 3: (a) Map of the posterior relative risk of CH incidence based on spatial Poisson model; (b) Map of the posterior relative risk of CH incidence based on spatial ZIP model

Numerous studies were conducted in Iran and worldwide on CH incidence rate and its risk factors from 2001 to 2018. CH incidence rate in Iran was reported to be high relative to many countries. The most important risk factors associated with CH are birth season, twin birth, jaundice at birth, birth weight, age at pregnancy, maternal anemia and goiter, gestational age, etc. in the study conducted by Rezaeian *et al.* (2013);[7] low birth weight, postdate delivery, and macrosomia in the study by Dalili *et al.* (2012);[12] length at birth, mother's age, and birth season in the study by Khammarian *et al.* (2018);[13] and mother's age in the study by Dayal *et al.* (2015).[22]

Some studies have investigated mental signs and growth signs. For instance, anatomic abnormalities, obesity problems, low-fertility in women, partial deafness, vision problems, etc. were studied by Bakopoulos *et al.* (2015).[23] Dalili *et al.* (2014) also studied the growth milestones outcome for neonates with CH. In this study, neonates with CH who were diagnosed and treated had normal growth.[11]

Osuli *et al.* (2008) examined the geographic distribution of CH in Iran. They computed and mapped CH incidence in provinces and townships of Iran using screening program data. In this article, they could not geographically justify the existence of a high CH incidence in Iran.[24]

Moreover, Mehrnejat *et al.* (2015) performed spatial analysis to determine the incidence rate of CH for each town in Isfahan Province during 2010–2013 and applied linear regression for data analysis. According to the results, there was no significant relationship between nitrate concentration in water and CH incidence.[25]

To the best of our knowledge, previous studies did not use the spatial ZIP model for the CH, but this model was implemented in other fields. Mapping of the disease and environmental species by a spatial ZIP model was conducted by Loquiha *et al.* (2018) and Lyashevska *et al.* (2016), respectively. In the study conducted by Lyashevska *et al.* (2016), a general procedure for mapping the abundance of species was provided for zero-inflated data.[26] Furthermore, Loquiha *et al.* (2018) mapped the maternal mortality rate by spatial zero-inflated models for Mozambique. The map of the posterior means of spatial random effect based on the spatial zero-inflated model was also plotted.[20]

In the present study, the relative risk for CH incidence was mapped via the posterior mean of relative risk based on the Bayesian spatial ZIP and Poisson model. Results revealed that the relative risk of CH is higher in Western
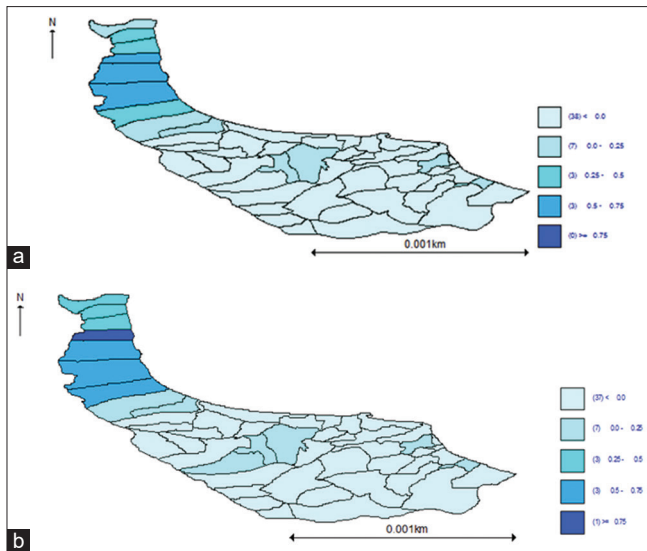
Figure 4: (a) Map of the posterior mean for θ in ZIP model; (b) Map of posterior mean for θ in spatial Poisson model

towns of Guilan than other regions of the province. In fact, few towns in the province had a high relative risk, which may be attributed to the elimination of iodine deficiency as one of the main causes of CH incidence by consuming seafood and iodized salt in this province. It is recommended that more investigations be conducted on risk factors such as iodine deficiency disorder, low birth weight, postdate delivery, and macrosomia in these towns. It should be also noted that we simultaneously examined both types of CH, namely transient and permanent. It is suggested that future studies examine these two types separately. We also recommend a more extended study in different parts of Iran to evaluate the prevalence and causes of CH.

Conclusions

Given the high relative risk of CH incidence in the Western towns of Guilan, it is recommended that the aforesaid risk factors related to CH be examined, especially in the West of Guilan.

Conflicts of interest

There are no conflicts of interest.

References

1. Keshavarzian E, Valipoor AA, Maracy MR. The incidence of congenital hypothyroidism and its determinants from 2012 to 2014 in Shadegan, Iran. Epidemiol Health 2016;38:1-6.
2. Hashemipour M, Hovsepian S, Ansari A, Keikha M, Khalighinejad P, Niknam N. Screening of congenital hypothyroidism in preterm, low birth weight and very low birth weight neonates: A systematic review. Pediatr Neonatol 2018;59:3-14.
3. Ordookhani A, Padyab M, Goldasteh A, Mirmiran P, Richter J, Azizi F. Seasonal variation of neonatal transient hyperthyrotropinemia in Tehran province, 1998–2005. Chronobiol Int 2010;27:1854-69.
4. Skordis N, Toumba M, Savva SC, Erakleous E, Topouzi M, Vogazianos M, *et al*. High prevalence of congenital hypothyroidism in the Greek Cypriot population: Results of the neonatal screening program 1990-2000. J Pediatr Endocrinol Metab 2005;18:453-62.
5. Ordookhani A, Minniran P, Najafi R, Hedayati M, Azizi F. Congenital hypothyroidism in Iran. Indian J Pediatr 2003;70:625-8.
6. Yarahmadi SH, Ali Mohammadzadeh KH, Tabibi SJ, Maleki MR. Presenting mathematics model of cost-benefit calculation of screening for congenital hypothyroidism in Iran. Int Math Forum 2011;6:681-97.
7. Rezaeian S, Poorolajal J, Moghimbegi A, Esmailnasab N. Risk factors of congenital hypothyroidism using propensity score: A matched case-control study. J Res Health Sci 2013;13:151-6.
8. Ahmad N, Irfan A, Al Saedi SA. Congenital hypothyroidism: Screening, diagnosis, management, and outcome. J Clin Neonatol 2017;6:64.
9. Etemad K, Khazaei Z, Pordanjani SR, Shahsavand M, Ajam F, Riahi S-M, *et al*. Evaluation of the therapeutic interventions effects on body growth pattern of infants with congenital hypothyroidism. Biomed Res Therap 2018;5:2194-207.
10. Dalili S, Mohtasham-Amiri Z, Rezvani SM, Dadashi A, Medghalchi A, Hoseini S, *et al*. The prevalence of iodine deficiency disorder in two different populations in Northern Province of Iran: A comparison using different indicators recommended by WHO. Acta Med Iran 2012;50:822-6.
11. Dalili S, Rezvani SM, Dalili H, Amiri ZM, Mohammadi H, Kesh SA, *et al*. Congenital hypothyroidism: Etiology and growth-development outcome. Acta Med Iran 2014;52:752-6.
12. Dalili S, Rezvany SM, Dadashi A, Medghalchi A, Mohammadi H, Dalili H, *et al*. Congenital hypothyroidism: A review of the risk factors. Acta Med Iran 2012;50:735-9.
13. Khammarnia M, Siakhulak FR, Ansari H, Peyvand M. Risk factors associated with congenital hypothyroidism: A case-control study in southeast Iran. Electron Physician 2018;10:6286.
14. Mohtasham AZ, Mousavi SM, Hosein–Zadeh M. Newborn screening for congenital hypothyroidism in Rasht, North of Iran, 2007. Early Human Develop 2008;(84):S122.
15. Hu M-C, Pavlicova M, Nunes EV. Zero-inflated and hurdle models of count data with extra zeros: Examples from an HIV-risk reduction intervention trial. Am J Drug Alcohol Abuse 2011;37:367-75.
16. Anderson C, Ryan L. A comparison of spatio-temporal disease mapping approaches including an application to ischaemic heart disease in New South Wales, Australia. Int J Environ Res Public Health 2017;14:146.
17. Manda SO, Feltbower RG, Gilthorpe MS. Investigating spatio-temporal similarities in the epidemiology of childhood leukaemia and diabetes. Eur J Epidemiol 2009;24:743-52.

18. Mahaki B, Mehrabi Y, Kavousi A, Akbari ME, Waldhoer T, Schmid VJ, *et al*. Multivariate disease mapping of seven prevalent cancers in Iran using a shared component model. Asian Pac J Cancer Prev 2011;12:2353-8.

19. Besag J, York J, Mollié A. Bayesian image restoration, with two applications in spatial statistics. Ann Inst Statist Math 1991;43:1-20.

20. Loquiha O, Hens N, Chavane L, Temmerman M, Osman N, Faes C, *et al*. Mapping maternal mortality rate via spatial zero-inflated models for count data: A case study of facility-based maternal deaths from Mozambique. PLoS One 2018;13:e0202186.

21. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. J Royal Statist Soc 2002;64:583-639.

22. Dayal D, Sindhuja L, Bhattacharya A, Bharti B. Advanced maternal age in Indian children with thyroid dysgenesis. Clin Pediatr Endocrinol 2015;24:59-62.

23. Bakopoulos N, Despotidis O, Saridi M. Congenital hypothyroidism: A variety of clinical and mental signs. Int J Caring Sci 2015;8:819.

24. Osuli M, Hagh dust A, YarAhmadi S, Foruzanfar M, Deini M, Holakuei K. Geographic distribution of congenital hypothyroidism in Iran via geographic information system. 2008. [In Persian].

25. Mehrnejat N, Yazdanpanah H, Fadaei Nobari R, Hashemipour M, Maracy M, Moafi M, *et al*. Spatial analysis of Neo- natal Congenital Hypothyroidism and Nitrate as an environmental pollutant in Isfahan Province during 2010-2013. Int J Prev Med 2015;6:76.

26. Lyashevska O, Brus DJ, van der Meer J. Mapping species abundance by a spatial zero-inflated Poisson model: A case study in the Wadden Sea, the Netherlands. Ecol Evol 2016;6:532-43.