RESEARCH ARTICLE

Spatial Analysis of Common Gastrointestinal Tract Cancers in Counties of Iran

Ali Soleimani[1], Jafar Hassanzadeh[1]*, Ali Ghanbari Motlagh[2], Hamidreza Tabatabaee[1], Elham Partovipour[3], Sareh Keshavarzi[1], Mohammad Hossein

Abstract

Background: Gastrointestinal tract cancers are among the most common cancers in Iran and comprise approximately 38% of all the reported cases of cancer. This study aimed to describe the epidemiology and to investigate spatial clustering of common cancers of the gastrointestinal tract across the counties of Iran using full Bayesian smoothing and Moran I Index statistics. Materials and Methods: The data of the national registry cancer were used in this study. Besides, indirect standardized rates were calculated for 371 counties of Iranand smoothed using Winbug 1.4 software with a full Bayesian method. Global Moran I and local Moran I were also used to investigate clustering. Results: According to the results, 75,644 new cases of cancer were nationally registered in Iran among which 18,019 cases (23.8%) were esophagus, gastric, colorectal, and liver cancers. The results of Global Moran's I test were 0.60 (P=0.001), 0.47 (P=0.001), 0.29 (P=0.001), and 0.40 (P=0.001) for esophagus, gastric, colorectal, and liver cancers, respectively. This shows clustering of the four studied cancers in Iran at the national level. Conclusions: High level clustering of the cases was seen in northern, northwestern, western, and northeastern areas for esophagus, gastric, and colorectal cancers. Considering liver cancer, high clustering was observed in some counties in central, northeastern, and southern areas.

Keywords: Spatial - epidemiology - clustering - cancer - Iran counties

Asian Pac J Cancer Prev, **16 (9)**, 4025-4029

Introduction

Cancer is one of the most important causes of death in the world. Globally there are 14 million new cases of cancer recorded each year, of which approximately 8.2 million cause death (Bernard and Christopher, 2014). In Iran, cancer is the third leading cause of death after cardiac diseases and injuries caused by accidents (Mousavi et al., 2009). Furthermore, gastrointestinal tract cancers are among the most common cancers in Iran and comprise approximately 38% of all reported cases of cancer. (Mohebbi et al., 2008) 14% of all deaths are attributed to cancer throughout the country, of which gastrointestinal tract cancers contribute approximately 44.4% (Yazdanbod et al., 2004).

Studying geographical differences among chronic diseases and research about risk factors that contribute to these differences have a long history in epidemiology. (Doll, 1980) The purpose of mapping studies related to disease distribution is to determine spatial differences in terms of risks for occurrence of diseases. This work was done to assess and quantify spatial heterogeneity for cancer diseases and to determine existing patterns in order to identify areas at low or high risk from clues about possible causes of cancer diseases in areas ranked with higher prevalence (Best et al., 2005). Standardized rates of morbidity and mortality were ranked and marked by color on maps to show risk of diseases. There is high heterogeneity of populations that may lead to different estimates of relative risk, particularly in sparsely populated areas and conventional modeling methods do not always consider spatial correlation between observations. Accordingly, various methods have been suggested to overcome this problem (Clayton and Kaldor, 1987; Besag et al., 1991; Rytkonen, 2004).

Smoothing or modeling methods are usually used for producing reliable estimates across different geographical regions. Spatial correlation of such observations requires application of reliable estimates to models in order to preserve data in a hierarchical structure. Therefore, generalized or Bayesian models that take into account correlations among data are used for maximum efficiency in such studies (Lawson et al., 2000; Lawson et al., 2003). These models consider previous information on variability

[1]*Departeman of Epidemiology, Research Center for Health Sciences, School of Health, Shiraz University of Medical Science, Shiraz,* [2]*Non-Communicable Diseases, Deputy of Health, Ministry of Health and Medical Education-Cancer Research Center, Shahid Beheshti University of Medical Sciences,* [3]*Non-Communicable Diseases, Deputy of Health, Ministry of Health and Medical Education, Tehran, Iran *For correspondence: jhasanzad@sums.ac.ir.*

rate of disease occurrence across areas using the number of diseases observed in each area. These models also consider spatial correlation for rate of occurrence of diseases among neighboring areas (Lawson et al., 2003). Achieving robust and direct estimates of neighboring areas, due to estimating stable and reliable nodes and better quantification for variance of estimators, are among the most prominent advantages of Bayesian models compared to other methods (Lawson et al., 2003; Wakefield, 2007; Banerjee et al., 2014). Bayesian type models, including full Bayes and empirical Bayes models that full Bayes are capable of obtaining estimates of higher accuracy (Bernardinelli and Montomoli, 1992).

Incidence of cancer including gastrointestinal tract type cancers is not uniformly distributed across different parts of Iran (Ansari et al., 2006; Malekzadeh et al., 2009; Sadjadi et al., 2010). Knowledge of spatial patterns of diseases can be useful for planning disease prevention and control. Techniques of exploratory spatial data analysis (ESDA) introduce a set of tools and techniques that are helpful for investigating spatial patterns of diseases (Anselin, 1994). ESDA map geographical differences, identify locations with outlier and disease clustering (Anselin, 1999; Anselin and Getis, 2010). Measuring spatial autocorrelation constitutes an important part of ESDA and the Moran I Index is the most commonly used statistics device applied to measure spatial autocorrelation (Anselin and Getis, 2010). Bayesian and ESDA smoothing methods are often used to describe spatial epidemiology of diseases in different studies (Mohebbi et al., 2008; Abdullayev et al., 2012; Zhang et al., 2014).

This study was conducted to describe the epidemiology and to investigate spatial clustering of common cancers of the gastrointestinal tract across the counties of Iran using the Full Bayesian smoothing method and the Moran I Index statistic.

Materials and Methods

Data collection

The registration system for reporting incidence of cancer in Iran began in 1984 when a law was passed that made it compulsory to report cases of cancer. Data on patients with cancer was recorded in the system in two ways; one through reports from pathological centers and the other through a system based on population records, whereby information was collected from mortality registers, hospital records, therapeutic health centers, and specialist physicians' clinics. In total, 589 centers collected data on first and last names, age, gender, residence, anatomical location of the cancer and morphology and grade of the tumor, through special forms designed for this purpose. Collected data were sent to the county's health center where information was checked and then corrected and modified if it was found to be incomplete. Data was then entered onto the software used for registration of cases of cancer. Data was transferred from the county to the provincial health center. In provincial health centers, data collected for the whole county was checked once again and any duplicated data was removed in the provincial center and transferred to the National Office of Cancer. The process of removing duplicated data on all data collected from across the whole country was implemented in the National Office of Cancer and data collection was prepared for analysis in the National Office for Cancer. It was necessary to remove duplicated data because a particular case may have been reported in more than one center, such as in different county health centers or different provincial health centers. Data on age, gender, and county of residence for records of esophagus cancer (C15), gastric cancer (C16), colorectal cancer (C18-21) and liver cancer (C22)in year of 2010 used in this study were collected nationwide in Iran by the above-mentioned registration system.

Data analysis

Firstly, Age Standardized Rate (ASR) was calculated using direct standardization method and the standard population in WHO 2000 for both genders in order to compare statistics on cancer in Iran with those for the rest of world (Ahmad et al., 2001). Next, in order to compare incidences of each type of cancer among the different counties of Iran, age and gender Standardized Incidence Ratio (SIR) for each type of cancer were computed using state incidences of the same cancer for different counties and by using in standardization method and STATA 11 software. However, according to SIR, when there is low population in an area ,the number of cases of a disease may be zero in that region, and then problems occur in the calculations because of the underlying structure of data. Therefore, SIR was smoothed in the first stage using the method proposed by Besag, York, and Mollie (BYM), known as the Full Bayesian estimator (Clayton and Kaldor, 1987; Besag et al., 1991; Lawson et al., 2000). The spatial Poisson model was used with 2 random effect components, in this method one of its components represented the spatial structured component in a location or spatial homogeneity and another was representative of changes without a spatial structure or heterogeneity and the model is defined as follows:

$$Oi \sim \text{Poisson}(\mu i\ Ei)$$

$$\log \mu i = \alpha + bi + hi$$

Where μi is SIR in a county i, Oi is the observed value of any cancer in the county i, and Ei is the expected value of a cancer according to indirect standardization. hi and bi describe spatial heterogeneity and homogeneity respectively, and non-informative normal and autoregressive conditional (CAR) prior distributions were used for bi and hi respectively (Clayton and Kaldor, 1987; Spiegelhalter et al., 2004).

Bayesian statistics software (Winbug 1.4) was used to calculate the smoothed SIR by the BYM method (Spiegelhalter et al., 2004). For this purpose, the model was fitted using the Monte Carlo simulation method of Markov chains that operates by choosing previous appropriate values for a chain (Best et al., 2005). Initial values for stochastic nodes were selected to provide disperse initial values without additional over dispersion, so that zero value was chosen for the beginning chain for the common intercept factor and hi and bi random effects (Mohebbi et al., 2008).

After burning 50000 iterations, the next 150,000

Table 1. Crude and Age Adjusted Incidence of Esophagus, Stomach, Colorectal and Liver Cancer in Iran in 2010

Cancer	Crude incidence		Adjusted incidence (95% confidence interval)		Age (Mean ± standard deviation)	
	Male	Female	Male	Female	Male	Female
Esophagus	4.5	3.7	5.2 (5.0 5.5)	4.3 (4.1 4.5)	66.4±14.4	64.7±14.0
Stomach	13.9	6.1	16.3 (15.8 16.8)	6.9 (6.6 7.2)	66.8±14.2	63.9±15.8
Colorectal	9.1	7.1	10.7 (10.3 11.1)	7.9 (7.6 8.2)	60.5±16	58.9±15.7
Liver	1.7	1.1	2 (1.8 2.2)	1.2 (1.1 1.3)	61.5±18.5	61.8±19.9

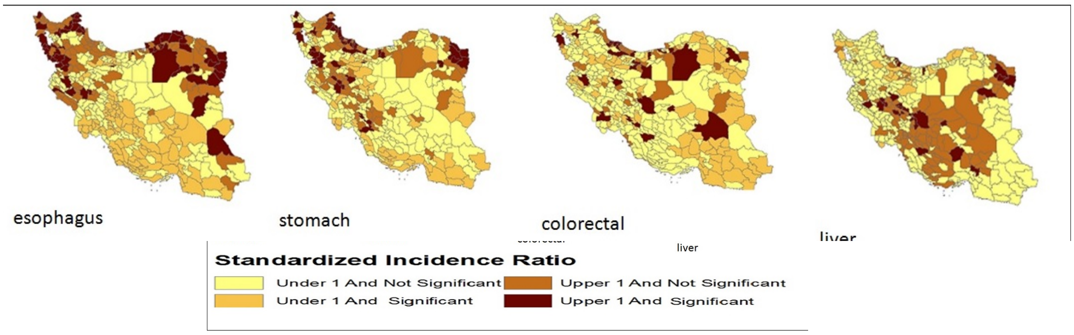

Figure 1. Mapping of Smoothed Standard Incidence of Esophagus, Stomach, Colorectal and Liver Cancer in Iran in 2010

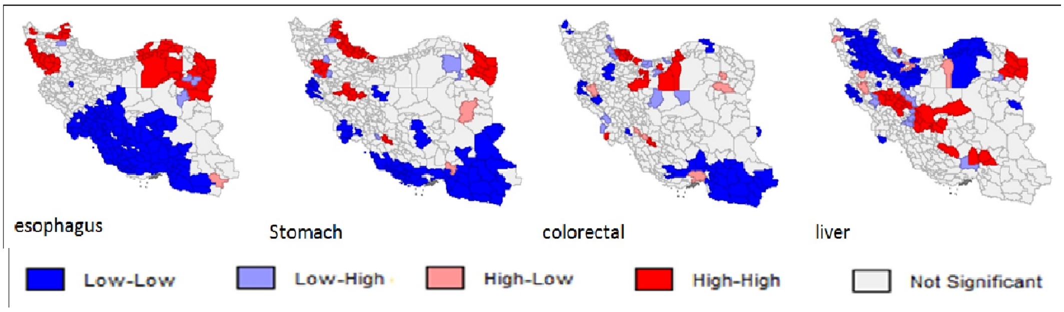

Figure 2. Clustering Map of Smoothed Standard Incidence of Esophagus, Stomach, Colorectal and Liver Cancer in Iran in 2010. low- low: clustering or positive autocorrelation and the rate of the disease in the colored regions and their neighbors are Low high- high: clustering or positive autocorrelation and the rate of the disease in the colored regions and their neighbors are high, high -low: no clustering in the colored region, while the rate of the disease is high in the colored regions but low in their neighbors, low- hogh : no clustering in the colored region, while the rate of the disease is low in the colored regions but high in their neighbors

iterations were used to select any 100 iterations as estimates for the smoothed SIR. Convergence of the chain was investigated using the coda software package and the Geweke test in R software that showed converging and static values of iteration (Plummer et al., 2006). Results obtained by the Winbug software were transferred into the geographical information system software to map the smoothed SIR.

Calculations were made using Global Moran's I and Local Moran's I to detect clustering using Gedoa 1.6.017 software. Global Moran's I was used to investigate existence of clustering, and Local Moran's I was used to identify counties that had clustering of gastrointestinal tract cancers. Neighborhood criterion for adjoining regions was considered together.

Results

75,644 new cases of cancer had been registered nationally in Iran and among these 18,019 (23.8%) related to four types of cancer; esophagus, gastric colorectal and liver cancers. 1583 of patients (8.8%) among the mentioned cases of cancer had were registered as having unknown county of residence and were excluded from analyses related to location.

Table 1 shows that the highest and lowest ASRs related to men with gastric cancer and women with liver cancer respectively and that averages for minimum and maximum age of incidence related to men with gastric cancer and women with colorectal cancer, respectively.

Figure 1 shows smoothed SIR mapping of the four studied cancers at national level in Iran. Results of the Global Moran's I test were 0.60 (P=0.001), 0.47 (P=0.001), 0.29 (p=0.001) and 0.40 (p=0.001) respectively for esophagus, gastric, colorectal and liver cancers, which show clustering at the national level of Iran for each of the four studied cancers and Figure 2 shows local results of the Moran I test for cancers and indicates counties with high and low level clustering as identified by the map.

Discussion

This study was done to investigate epidemiologic characteristics of patients with gastrointestinal tract cancers and to investigate geographical distribution of such cases at the national level in Iran according to data relating to the different counties, collected in 2010.

Esophagus cancer: Results of the present study demonstrated higher ASR values for men in comparison with women and high level clustering of cases was seen for cancer of the esophagus in the north, northwest and west and northeastern areas.

The results of a study conducted by Mousavi et al in 2005-2006 reported age standardized incidences of esophagus cancer for men and women at 6.2 and 5.8 per 100,000 persons respectively, statistics that are close to estimates obtained in this study (Mousavi et al., 2009). In global terms, esophagus cancer has widespread geographical distribution so that its age standardized rate was above 100 in some areas of China, 29 among black Americans, 24 among French and 3 among American Caucasians per 100,000 persons in a population. High incidence of esophagus cancer has also been reported in some areas of Iran such as Gonbad (Sadjadi et al., 2010). However, recent studies conducted in the north show a significant decrease relative to the report published in 1970, but this area of Iran has high incidences of cases of esophagus cancer relative to other parts of the country (Semnani et al., 2006). Another study conducted by the Poisson Kerijing method at the national level in Iran has reported many cases of esophagus cancer in northern, north western, and north eastern Iran (Asmarian et al., 2013b). Studies conducted in northern Iran show that risk factors were low-level consumption of fresh fruit and vegetables, low economic-social status, opium and hot tea consumption in this region of Iran (Radmard, 2010).

Gastric cancer: Results of the present study show that ASR was more prevalent in men than in women and high level clustering was seen for incidence of esophagus cancer in northern, northwestern, western, and northeastern regions. Incidences of gastric cancer among regions of the world show geographical difference.

High-risk areas, according to rates of age-adjusted incidence of over 20 in 100,000 persons per year include Japan, China, Brazil, Costa Rica and Korea. Areas with a moderate risk and standardized incidence of between 20 to 10 per 100, 000 persons include England, Germany, New Zealand and Turkey. Areas with low risk and adjusted incidence of less than 10 include America, Sweden, Denmark and India (Cancer, 2009). In spite of a decline that has occurred in incidences of gastric cancer in Iran during recent years, according to results of the present study, Iran is placed in the group of countries with a moderate risk of gastric cancer(Malekzadeh et al., 2009). Furthermore, different regions of the country also have obvious geographical differences. Clustering of cancer incidence has been reported in a study conducted in Iran in 2008 in counties at the margin of the Caspian Sea (Mohebbi et al., 2008). Another study in Ardebil province in Iran also showed that the province was a high risk area in terms of global statistics (Sadjadi et al., 2003) and another study investigated the status of the gastric cancer among provinces of Iran from 2004 to 2009 as clustering of incidences and showed that gastric cancer had clustering status in northern, north-western and western provinces (Kavousi et al., 2014).

It should be noted that the study conducted in Ardebil, a high-risk area, has shown that higher rate of cardia gastric cancer incidence than non-cardia type of cancer in the area and that that this was also evident in other high-risk areas of the world such as Japan. Furthermore, another study has shown that in Iran, in low-risk areas, 85% of cases of gastric cancer were non-cardia in these areas (Malekzadeh et al., 2009). Several factors such as infection of Helicobacter pylori, smoking, high salt consumption and inadequate intake of antioxidants were among the known risk factors for gastric cancer (Radmard, 2010).

Colorectal cancer: Results of this study show that ASR evaluation was higher for men than for women and clustering of high cases of the disease were observed in the northern counties of Iran. Colorectal cancer is the third and second most common type of cancer among men and women of the world, respectively, which in terms of ranking means that incidence of colorectal cancer in Iran is similar to that for global statistics (Jemal et al., 2011). Epidemiologic characteristics and age distribution of colorectal cancer have abundant geographical differences at the global level (Center et al., 2009). A study conducted in Iran by Ansari et al has estimated that age standardized incidences for both men and women were 8.0 and 7.2 per 100,000 persons respectively (Ansari et al., 2006); in this study, estimates showed slightly higher rates of incidence for men, this difference may be because of a time difference in conducting the two studies or because of difference in data applied to analyses because data used in this study covered all provinces but data from only 5 provinces was used in the other mentioned study.

Other studies have shown that incidence of colorectal cancer is rising in Iran (Moghimi-Dehkordi et al., 2008; Atrkar-Roushan et al., 2013). In addition, another study has investigated geographical distribution of colorectal cancer in Iran, according to county by the Poisson Kerijing method and demonstrated high incidence of colorectal cancer in northern and central areas of Iran, and low incidence in the southeast, which corresponds with the results of this study(Asmarian et al., 2013a). High total energy consumption, consumption of animal fat, low fiber intake and lack of physical activity were identified as the most important risk factors for colorectal cancer (Somi et al., 2014).

Liver cancer: Comparison showed that ASR of this cancer was higher in men than in women. In addition, this type of cancer showed high clustering for disease cases in some counties of central, northeastern and southern areas of Iran. This cancer is the sixth most common cancer at the global level. Its incidence achieves less than 10 per 100,000 persons in some North American and Western European areas and up to 50 to 150 per 100,000 persons in some of African and Asian areas (Parkin et al., 2005). Iran with incidence of less than 5 per 100,000 persons, it is ranked as a low risk region along with risk of liver cancer(Gomaa et al., 2008). The study of Mohebbi et al showed no high rate of incidence of liver cancer in northern areas of Iran(Mohebbi et al., 2008). Another study showed that incidence of liver cancer in southeastern Iran was above the national average and our study has also reported high level presence of cases by clustering results in a few counties (Moghaddam et al., 2010).

Acknowledgements

This paper has been extracted from a Ali Soleimani's thesis (title of thesis: Bayesian factor analysis for spatial correlated data: application to prevalent cancers in Iran in 2010) from an epidemiologic PhD course, No. 93-7180, recorded in the research department of Shiraz Medical Sciences University. In addition, data used in this study was received from the National Cancer Registration Center. Hereby, the effort of all participating colleagues is appreciated and acknowledged.

References

Abdullayev R, Kracalik I, Ismayilova R, et al (2012). Analyzing the spatial and temporal distribution of human brucellosis in Azerbaijan (1995-2009) using spatial and spatio-temporal statistics. BMC infectious diseases, **12**, 10.

Ahmad OB, Boschi-Pinto C, Lopez AD, et al (2001). Age standardization of rates: a new WHO standard, World Health Organization Geneva.

Ansari R, Mahdavinia M, Sadjadi A, et al (2006). Incidence and age distribution of colorectal cancer in Iran: results of a population-based cancer registry. *Cancer Letters*, **240**, 143-7.

Anselin L (1994). Exploratory spatial data analysis and geographic information systems. *New tools for spatial analysis*, 54.

Anselin L (1999). Interactive techniques and exploratory spatial data analysis. Geographical Information Systems: Principles, Techniques, Management and Applications, eds.

Anselin L, Getis A (2010). Spatial statistical analysis and geographic information systems. In 'Perspectives on Spatial Data Analysis', *Eds Springer*, 35-47

Asmarian NS, Kavousi A, Salehi M (2013a). Mapping of large intestine cancer incidence rate using area-to-area poisson kriging in Iran from 2003 to 2007. *Razi J MedSci*, **20**, 7.

Asmarian NS, Ruzitalab A, Amir K, et al (2013b). Area-to-area poisson kriging analysis of mapping of county-level esophageal cancer incidence rates in Iran. *Asian Pac J Cancer Prev*, **14**, 3.

Atrkar-Roushan Z, Kazemnejad A, Mansour-Ghanaei F, et al (2013). Trend analysis of gastrointestinal cancer incidences in Guilan province: comparing rates over 15 years. *Asian Pac J Cancer Prev*, **14**, 7587-93.

Banerjee S, Carlin BP, Gelfand AE (2014). Hierarchical modeling and analysis for spatial data, Crc Press.

Bernard W, Christopher P (2014). World cancer report 2014. international agency for research on cancer. World Health Organization, Lyon, France.

Bernardinelli L, Montomoli C (1992). Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. *Stat Med*, **11**, 983-1007.

Besag J, York J, Mollie A (1991). Bayesian image restoration, with two applications in spatial statistics. *Ann Institute Statist Math*, **43**, 19.

Best N, Richardson S, Thomson A (2005). A comparison of Bayesian spatial models for disease mapping. *Stat Methods Med Res*, **14**, 35-59.

Cancer IAfRo (2009). Cancer incidence in five continents, international agency for research on cancer. **[DETAIL]**

Center MM, Jemal A, Ward E (2009). International trends in colorectal cancer incidence rates. *Cancer Epidemiol Biomarkers Prev*, **18**, 1688-94

Clayton D, Kaldor J (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*, **[ISSUE]** 671-81

Doll SR (1980). The epidemiology of cancer. *Cancer*, **45**, 2475-85.

Gomaa AI, Khan SA, Toledano MB, et al (2008). Hepatocellular carcinoma: epidemiology, risk factors and pathogenesis. *World J Gastroenterol*, **14**, 4300.

Jemal A, Bray F, Center MM, et al (2011). Global cancer statistics. CA: Cancer J Clin, **61**, 69-90.

Kavousi A, Bashiri Y, Mehrabi Y, et al (2014). Identifying high-risk clusters of gastric cancer incidence in iran, 2004-2009. *Asian Pac J Cancer Prev*, **15**, 10335-7.

Lawson A, Biggeri A, Boehning D, et al (2000). Disease mapping models: an empirical evaluation. Disease Mapping Collaborative Group. *Statist Med*, **19**, 2217-41.

Lawson AB, Browne WJ, Rodeiro CLV (2003). Disease mapping with WinBUGS and MLwiN, John Wiley & Sons.

Malekzadeh R, Derakhshan MH, Malekzadeh Z (2009). Gastric cancer in Iran: epidemiology and risk factors. *Arch Iran Med*, **12**, 576-83.

Moghaddam SD, Haghdoost AA, Hoseini SH, et al (2010). Incidence of hepatocellular carcinoma in southeast iran. *Hepat Mon*, **10**, 270.

Moghimi-Dehkordi B, Safaee A, Zali MR (2008). Prognostic factors in 1,138 Iranian colorectal cancer patients. *Intern J Colorectal Dis*, **23**, 683-8.

Mohebbi M, Mahmoodi M, Wolfe R, et al (2008). Geographical spread of gastrointestinal tract cancer incidence in the Caspian Sea region of Iran: spatial analysis of cancer registry data. *BMC Cancer*, **8**, 137.

Mousavi SM, Gouya MM, Ramazani R, et al (2009). Cancer incidence and mortality in Iran. *Ann Oncol*, **20**, 556-63.

Parkin DM, Bray F, Ferlay J, et al (2005). Global cancer statistics, 2002. *CA: Cancer J Clin*, **55**, 74-108.

Plummer M, Best N, Cowles K, et al (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, **6**, 7-11.

Radmard AR (2010). Five common cancers in Iran. *Arch Iran Med*, **13**, 143-6.

Rytkonen MJP (2004). Not all maps are equal: GIS and spatial analysis in epidemiology. *Intern J Circumpolar Health*, **63**.

Sadjadi A, Malekzadeh R, Derakhshan MH, et al (2003). Cancer occurrence in Ardabil: Results of a population-based Cancer Registry from Iran. *Int J Cancer*, **107**, 113-8.

Sadjadi A, Marjani H, Semnani S, et al (2010). Esophageal cancer in Iran: A review. *Middle East J Cancer*, **1**, 5-14.

Semnani S, Sadjadi A, Fahimi S, et al (2006). Declining incidence of esophageal cancer in the Turkmen Plain, eastern part of the Caspian Littoral of Iran: a retrospective cancer surveillance. *Cancer Detect Prev*, **30**, 14-9

Somi MH, Golzari M, Farhang S, et al (2014). Gastrointestinal cancer incidence in East Azerbaijan, Iran: update on 5 year incidence and trends. *Asian Pac J Cancer Prev*, **15**, 3945-9.

Spiegelhalter D, Thomas A, Best N, et al (2004). WinBUGS User Manual version 1.4 January 2003 MRC Biostatistics Unit. Inst. of Public Health, Cambridge, England.

Wakefield J (2007). Disease mapping and spatial regression with count data. *Biostatistics*, **8**, 158-83.

Yazdanbod A, Nasseri-Moghaddam S, Malekzadeh R (2004). Upper gastrointestinal cancer in Ardabil, North West of Iran: A review. *Arch Iranian Med*, **7**, 173-7.

Zhang J, Yin F, Zhang T, et al (2014). Spatial analysis on human brucellosis incidence in mainland China: 2004-2010. *BMJ open*, **4**.