RESEARCH ARTICLE

# Trends of Breast Cancer Incidence in Iran During 2004-2008: A Bayesian Space-time Model

**Tohid Jafari-Koshki[1], Volker Johann Schmid[2], Behzad Mahaki[1]***

## Abstract

**Background: Breast cancer is the most frequently diagnosed cancer in women and estimating its relative risks and trends of incidence at the area-level is helpful for health policy makers. However, traditional methods of estimation which do not take spatial heterogeneity into account suffer from drawbacks and their results may be misleading, as the estimated maps of incidence vary dramatically in neighboring areas. Spatial methods have been proposed to overcome drawbacks of traditional methods by including spatial sources of variation in the model to produce smoother maps. Materials and Methods: In this study we analyzed the breast cancer data in Iran during 2004-2008. We used a method proposed to cover spatial and temporal effects simultaneously and their interactions to study trends of breast cancer incidence in Iran. Results: The results agree with previous studies but provide new information about two main issues regarding the trend of breast cancer in provinces of Iran. First, this model discovered provinces with high relative risks of breast cancer during the 5 years of the study. Second, new information was provided with respect to overall trend trends o. East-Azerbaijan, Golestan, North-Khorasan, and Khorasan-Razavi had the highest increases in rates of breast cancer incidence whilst Tehran, Isfahan, and Yazd had the highest incidence rates during 2004-2008. Conclusions: Using spatial methods can provide more accurate and detailed information about the incidence or prevalence of a disease. These models can specify provinces with different health priorities in terms of needs for therapy and drugs or demands for efficient education, screening, and preventive policy into action.**

**Keywords:** Breast cancer - incidence - spatio-temporal information - bayesian disease mapping - Iran

*Asian Pac J Cancer Prev,* **15 (4)**, 1557-1561

## Introduction

Cancer is considered as the leading cause of death in developed and the second leading cause of death in developing countries, with estimated 7.6 million deaths worldwide in 2008 (Jemal et al., 2011). Breast cancer was the most frequently diagnosed cancer with 23% out of the total new cancer cases and accounts for 14% of the total cancer deaths in females in 2008 (Ferlay et al., 2010; Jemal et al., 2011). In Asia, 27.2% of all new cancer cases and 19% of all cancer deaths in females is due to breast cancer. Breast cancer was the first site and cause of death among all types of cancer in 2008 (American Cancer Society, 2011).

Global incidence trend of breast cancer is increasing especially in countries with a low rate of incidence (Montazeri et al., 2008). Breast cancer had an increasing trend from 1965-2000 among Iranian women changing the rank of prevalence from the second to the first during these years (Mousavi et al., 2007). According to the last cancer registry reports by Iran Ministry of Health, incidence rate of female breast cancer in Iran has been reported to be 21.4% to 24.4% in 2000-2006 that has the first rank among all diagnosed cancers and it is the fifth common cause of death due to cancer in women (Lamyian et al., 2007; Mousavi et al., 2009; Babu et al., 2011; Noroozi et al., 2011; Taghavi et al., 2012). Iranian women are affected by breast cancer at least 10 years earlier than their counterparts in developed countries (Harirchi et al., 2004). Hence, breast cancer is one of the health priorities in Iran and awareness about the temporal trends and regional patterns of its incidence and prevalence rates, as well as its risk factors, will lead to better health policies.

Health officials use different indices to assess the gravity of a health-threatening issue in a population or area to allocate health budgets and determine reasons as well as manipulations needed to control the problem. Disease mapping, as a useful tool in the analysis of geographical variation in rates of disease incidence or mortality, aims to summarize spatial variation in occurrence of a phenomenon to project spatial heterogeneity and identify areas with low or high rates and formulate and validate etiological hypotheses (Gilks et al., 1996). That is, using these maps one can describe geographic variation of diseases risk, suggest possible risk factors responsible for this variation, assess health inequalities to improve

*[1]Department of Biostatistics and Epidemiology, School of Health, Isfahan University of Medical Sciences, Isfahan, Iran, [2]Bioimaging Group, Department of Statistics, Ludwig Maximimilians, Munich, Germany *For correspondence: behzad.mahaki@gmail.com*

health policy and resource allocation, and finally construct disease atlases (Lawson et al., 2000; 2003).

The main challenge in disease mapping is the choice of an appropriate measure of relative risk (RR) (Clayton and Kaldor, 1987; Maiti, 1998). Standardized morbidity/mortality ratio (SMR) or standardized incidence rate (SIR) is the typical estimate of relative risk (RR) to assess the risk of a disease (MacNab and Dean, 2002; Mahaki et al., 2011). For count data, such as the number of patients, Poisson models are used to produce standardized rates. In this model, observed number of events in each population depends on area-specific RR. In other words, number of patients in population $i$ follows a Poisson distribution with mean $E_i RR_i$, where $E_i$ is the expected number of patients calculated as follows: $E_i=n_i(\Sigma_i y_i/\Sigma_i n_i)$, $i=1,2,...,A$

where ni is the population at risk in area $i$, $yi$ is the observed number of patients in area $i$ and A is the number of areas. In traditional setting, $RR_i$ is estimated by dividing the observed, $y_i$, to the expected number of patients, $E_i$.

This estimator has some drawbacks. First, the observed number of patients exceeds that expected from the model, especially in rare diseases and small areas. This is known as overdispersion or extra-Poisson variation, induced by the fact that individual risks are heterogeneous within each population (Gilks et al., 1996).

The second problem is that this model overlooks spatial patterns of a disease. We expect that geographically closer areas have similar RR and the risk of a disease not vary dramatically in neighboring areas. This model fails to take geographical dependence into account and produces rough maps.

A hierarchical Bayesian approach is used to filter out the Poisson variation from the map and obtain maps which better reflect the true heterogeneity of the RR (Bernardinelli et al., 1995a). Bayesian methodology uses predefined information about the parameter of interest via so called prior distribution and combines it with information available from observed data and produces posterior distribution of the parameters. The mean of this distribution is called the Bayesian estimates and tends to be less dispersed (Bernardinelli et al., 1995a). Clayton and Kaldor (1987) applied empirical Bayesian method to estimate incidence of lip cancer in Scotland. Impact of neighboring areas is introduced to the model by using a conditional autoregressive (CAR) distribution.

In general, studies report the incidence and prevalence of breast cancer in a single period of time or without taking the spatial effect into account. Bernardinelli et al. (1995b), Waller et al. (1997), Knorr-Held et al. (1998), and (Knorr-Held and Besag, 1998) used Bayesian methodology to model the disease risk in space and over time.

Different methods have been used to study various cancers in Iran. Asmarian et al. (2013) used kriging analysis to study incidence of esophageal cancer in Iran in county-level. Kavousi et al. (2009) used the model proposed by Besag et al. (1991) to study the spatial pattern of lip cancer in Iran. Fouladi et al. (2012) and Taghavi et al. (2012) studied the trends of breast cancer incidence and mortality in Ardabil (a north-western province) and Iran in 2003-2010 and 1995-2004. They used traditional estimators to estimate the RR of breast cancer. However, no study has been done in Iran about the temporal trends of breast cancer incidence and prevalence along with the spatial patterns.

In this study, we evaluate the incidence of breast cancer during 2004-2008. We use the method proposed by Bernardinelli et al. (1995b) to allow for spatial and time effects and their interaction to obtain better estimates and produce smoother maps.

## Materials and Methods

In this study data for breast cancer (ICD10 code C50), in 30 provinces of Iran during 2004-2008 were considered. The data has been collected and made available by the Iranian Ministry of Health and Medical Education (Iran Cancer Registry Report, 2004-2008).

To circumvent problems related to SMR and SIR, several methods have been proposed and applied in a variety of disciplines (Besag et al., 1991; Bernardinelli et al., 1995b; Bernardinelli et al., 1997; Knorr-Held, 1998; Maiti, 1998; Bohning et al., 2000; Knorr-Held, 2000; Lawson et al., 2000; MacNab and Dean; 2002). Bayesian models have received great attention amongst the others due to their ability in modeling complex spatial data. Besag et al. (Besag, 1974; Besag et al., 1991) proposed CAR models to take into account spatial correlation of disease rates among neighboring areas. Other models have been suggested to analyze disease rates over apace and time simultaneously (Bernardinelli et al., 1995b; Waller et al., 1997; Knorr-Held, 1998; Knorr-Held and Besag, 1998). These models try to estimate RR in areas over time. We used generalized linear mixed model (GLMM) proposed by Bernardinelli et al. (1995b) to study breast cancer prevalence in Iran during 2004-2008. This model considers two sources of heterogeneity among provinces: independent heterogeneity, inherent in each area, and correlated heterogeneity due to correlation among neighboring areas as follows.

Let $i$ denote $i$th province and $k$ be the time indicator. Then RR is modeled as:

$$O_{ik} \sim Poisson(E_{ik} RR_{ik})$$

$$log(RR_{ik}) = \alpha + logE_{ik} + u_i + v_i + \beta t_k + \delta_i t_k$$

where $O_{ik}$ and $E_{ik}$ are the observed and the expected number of cases for province $i$ in year $k$, as before. $\alpha$ is an overall level of the relative risk, $v_i$ is the correlated heterogeneity and $u_i$ is the uncorrelated heterogeneity. $\beta$ introduces the effect of time at time $t_k$. Here $\delta_i$ is a parameter for comparing area-specific trend with overall mean trend; negative and positive value of which indicates less steep and steeper trend than country-wide trend, respectively. Besag et al. (1991) supposed a normal prior distribution for ui to account for uncorrelated heterogeneity. CAR-normal prior distribution is assumed for $v_i$ and $\delta_i$ to allow for spatial correlated heterogeneity by borrowing potential effects from adjacent areas. In the literature it has been accepted to assign equal weight to all adjacent areas. The last term is the interaction between time and area. Bernardinelli et al. (1995b) used this model to investigate IDDM prevalence in Sardinia.

We used this model in Bayesian setting and for normal and CAR-normal distributions we set the mean equal to

0 and allowed the variance to follow an inverse-Gamma distribution with $\alpha$=0.5 and $\gamma$=0.0005 (Besag et al., 1991; Lawson et al., 2003). MCMC methods are used extensively to fit the complex models like this and different packages readily implement them. First, we used the BYM model (Besag et al., 1991) to estimate overall RR of breast cancer during in 2004-2008 taking the population of 2006 as reference. BYM is a spatial model and has no parameter to capture the effect of time. We used open-access software OpenBUGS 3.2.2 to fit the models (Lunn et al., 2009).

## Results

Figure 1 shows the post-estimate of the overall incidence of breast cancer in each province over 2004-2008 estimated using BYM model. This model takes entire 5-year period as a single period of time and the population size of 2006 as default. Darker areas show the provinces with higher incidence rates. This figure suggests that the incidence of breast cancer is the highest for provinces Tehran (province 8 in Figure 1), Isfahan (P9), and Yazd (P10) with estimated value of 1.45, 1.38, and 1.27 respectivelys.

To investigate which provinces have less steep or steeper trend than the overall trend in the country, during 2004-2008, we use the estimated values of $\boldsymbol{\delta}_i$. This parameter is an indicator of the pace of breast cancer RR in each province relative to countrywide overall pattern; positive value of which indicates higher speed (Bernardinelli et al., 1995b). Figure 2 shows the results. That is, this figure reflects temporal patterns of incidence and serves as a tool to compare the relative speed of breast cancer occurrence in provinces to each other. Darker areas show the provinces with the most increasing pattern of incidence rate. In general, this figure suggests that the incidence of breast cancer has the highest increasing trend in north-eastern provinces and in East-Azerbaijan (P2).

To have a closer look at the trend, we plotted estimated RR in each province for each year. Figure 3 shows the estimated RR during 2004-2008. It can be inferred from this figure that, in general, incidence of breast cancer is very smoothly decreasing over time as the maps are getting

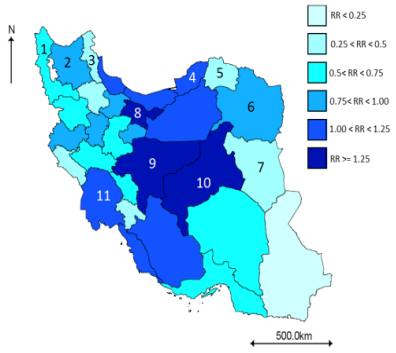


**Figure 1. Overall Post-Estimates of RR of Breast Cancer in Iran During 2004-2008**

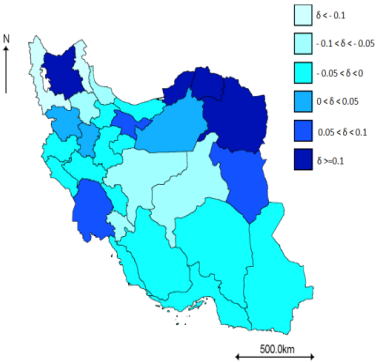


**Figure 2. Post-Estimate values of δi**

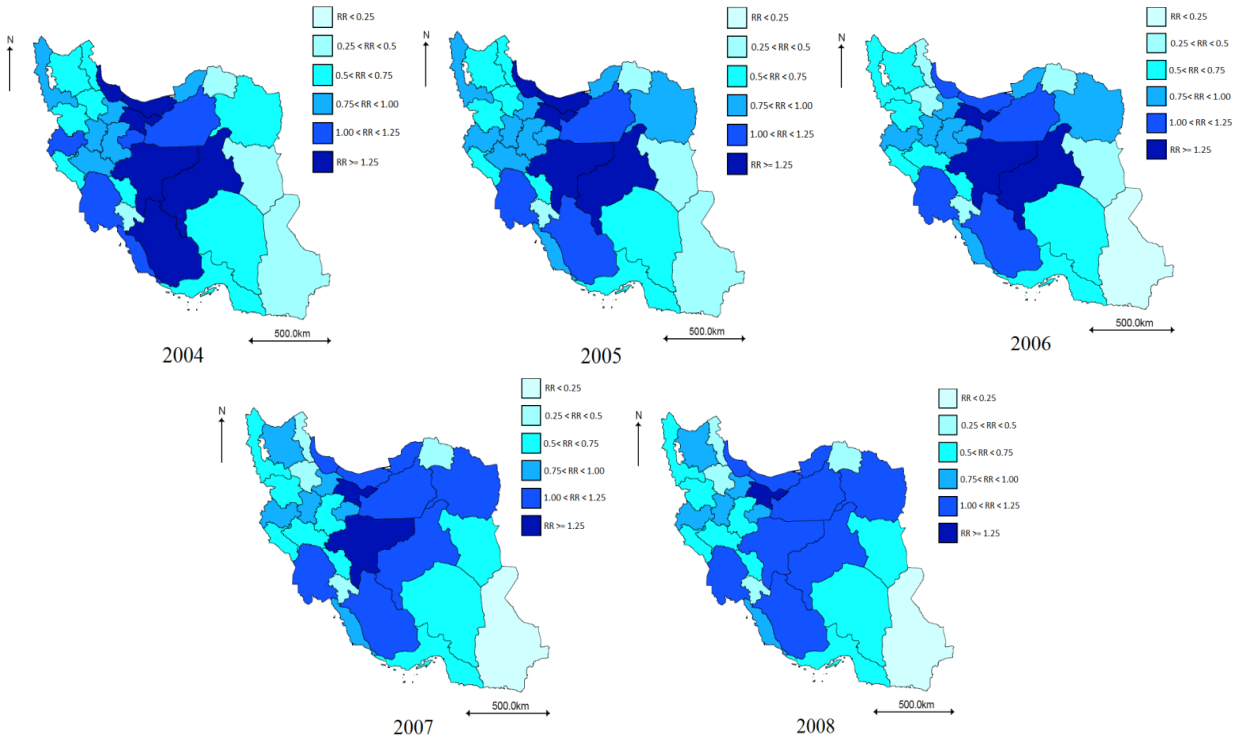


**Figure 3. Estimated RR for Breast Cancer in Iran; 2004-2008**

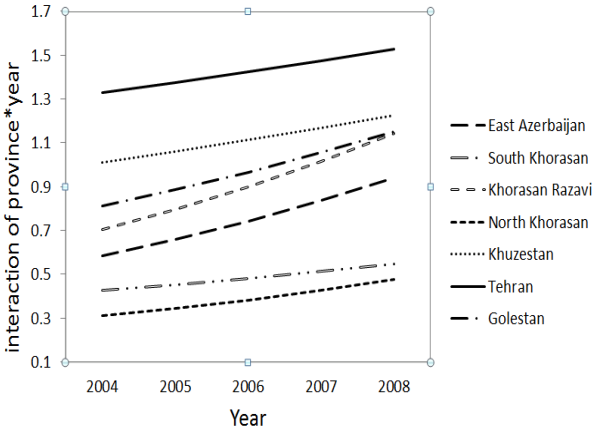


**Figure 4. Estimated Trends of Breast Cancer RR for 7 Provinces from 2004-2008**

lighter by passing the time. This fact is supported by the estimated value of-0.0195 for $\beta$ indicating that estimated overall incidence of breast cancer has a very smooth decreasing pattern in Iran in 2004-2008.

The results presented in Figure 2 are apparent from this figure. The increasing pattern is clear for East-Azerbaijan (P2), Golestan (P4), and Khorasan-Razavi (P6). The pattern is not so clear for North-Khorasan (P5). This is due to low rate of incidence in this province for all years. However, this does not contradict the high increasing pattern of incidence in North-Khorasan.

We used posterior probabilities (PP), Bayesian equivalent to p-value, to conduct a significance test about estimated province×year RR (Bernardinelli et al., 1995b). According to PP, Khorasan-Razavi (P6), East-Azerbaijan (P2), North-Khorasan (P5), Golestan (P4), South-Khorasan (P7), Khuzestan (P11), and Tehran (P8) (provinces with two most intense colors in Figure 2) have breast cancer RR trend significantly steeper than overall trend in the country during 2004-2008 with estimated temporal trend values of 1.128, 1.126, 1.114, 1.091, 1.063, 1.05, and 1.036 respectivelys.

## Discussion

Many studies have looked at the incidence and prevalence of breast cancer in Iran. Most of these studies are constrained to a single time period and their results are in accordance with our results (Sadjadi et al., 2003; 2009a; 2009b; Mousavi et al., 2008; Semnani et al., 2008; Somi et al., 2008; Taheri et al., 2012). Mahaki et al. (2011) studied the incidence of 7 common cancers in Iran in 2007 by allowing for spatial effects. There has been no study in this area to notice shortcomings of traditional methods in RR estimation and consider spatial along with time effects in the study. We used the model proposed by Bernardinelli et al. (1995b) to allow for spatial correlation among closer provinces as well as area and time interaction. Mousavi et al. (2009) reported that the Age-adjusted rate (ASR) of cancer in the north of Iran is generally more than in its southern provinces. This claim is partially supported in Figure 1 with a darker pattern in north with respect to south part.

From Figure 1 and 2 it is evident that RR of breast cancer in Ardabil (P3) and West-Azerbaijan (P1) is low in 2004-2008. Ardabil has been reported to be an area with lowest breast cancer incidence in Iran. Sadjadi et al. (2003 and 2009) reported that breast cancer in Ardabil women during 1996-1999 had low ASR of 7.6 compared to global ASR of 35.7 in 2000. However East-Azerbaijan had a high estimated value of 1.12 for RR temporal trend. This is in accordance with the results of Somi et al. (2008). They reported that ASR of breast cancer in East-Azerbaijan was 23.47 in 2006-2007, a higher value compared to that of Ardabil. Since these provinces are very similar in various characteristics, causes of this discrepancy remain as a question and a comparative study seems necessary to obtain some answer.

In Golestan (P4) breast cancer is the third most prevalent cancer in total population and the first female cancer with crude and age-standardized rate of 22.3 and 27.1 in 2006 (Semnani et al., 2008). Semani et al (2008) reported that the cancer incidence was increasing in Golestan in 2004-2006 but there was no significant change from 2006 to 2008. Taheri et al. (2012) reported high rate of breast cancer for this province (28 per 100,000 person year) during 2004-2009. A similar pattern is present in Figure 3. It can be seen from this figure that this province has maintained its class in 2004-2006, but has experienced a change from 2006 to 2007 and no changes thereafter. However, Golestan notably experiences an increasing incidence rate of breast cancer, as shown in Figure 2.

Figure 4 plots the profiles of estimated provincial trends of breast cancer rates for 2004-2008 for 7 provinces with highest increasing patterns. These provinces have trends significantly steeper than country-wide trend. The slope of the trend is steepest for Khorasan-Razavi and Golestan which agree with previous studies. Figure 2 and Figure 4 suggest that more attention is required for East-Azerbaijan, Golestan, North-Khorasan, and Khorasan-Razavi in terms of research for plausible risk factors and adoption of efficient preventive policy, screening, and public education.

Several studies have been performed in Iran about potential risk factors of breast cancer in various provinces. Almost all of these studies are restricted to small hospital-size populations. It would be very helpful to design and conduct a larger study over provinces identified with increasing patterns or stable high rate of incidence. Such studies could shed a light on the common causes responsible for each of these issues.

We studied the breast cancer profile in 2004-2008 and tried to identify the provinces with highest rates. The advantages of spatial models would be better perceived in small area studies. Therefore, designing a research to include county level data is suggested. Enhancing the time span and inclusion of data from recent years would result in a better view of breast cancer in Iran. Data for Alborz, newly established province, was not available and this province was not included in our study.

Cleary, it is worth conducting more researches considering potential risk factors and their trends to provide a better insight into what happens in each area to help policy makers in controlling cancer incidence.

## Acknowledgements

The authors wish to thank the office of non-communicable disease, Ministry of Health and Medical Education, Iran.

## References

Asmarian NS, Ruzitalab A, Kavousi A, et al (2013). Area-to-area poisson kriging analysis of mapping of county-level esophageal cancer incidence rates in Iran. *Asian Pac J Cancer Prev*, **14**, 11-3.

Babu GR, Samari G, Cohen SP, et al (2011). Breast cancer screening among females in Iran and recommendations for improved practice: a review. *Asian Pac J Cancer Prev*, **12**, 1647-55.

Bernardinelli L, Clayton D and Montomoli C (1995a). Bayesian estimates of disease maps: how important are priors? *Stat Med*, **14**, 2411-31.

Bernardinelli L, Clayton D, Pascutto C, et al (1995b). Bayeian analysis of space-time variation in disease risk. *Stat Med*, **14**, 2433-43.

Bernardinelli L, Pascutto C, Best NG, et al (1997). Disease mapping with errors in covariates. *Stat Med*, **16**, 741-52.

Besag J (1974). Spatial interaction and the statistical analysis of lattice systems. *J Royal Stat Soci. Series B (Methodological)*, **36**, 192-236.

Besag J, York J, Mollie A (1991). Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Statist Math*, **43**, 1-59.

Bohning D, Dietz E, Schlattmann P (2000). Space-time mixture modelling of public health data. *Stat Med*, **19**, 2333-44.

Clayton D and Kaldor J (1987). Empirical Bayes estimates of age-standardized relative risk for use in disease mapping. *Biomet*, **43**, 671-81.

Ferlay J, Shin H-R, Bray F, et al (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*, **127**, 2893-917.

Fouladi N, Pourfarzi F, Amani F, et al (2012). Breast Cancer in Ardabil province in the North-West of Iran: an epidemiological study. *Asian Pac J Cancer Prev*, **13**, 1543-5.

Gilks WR, Richardson S, Spiegelhalter DJ (1996). Markov Chain Monte Carlo in Practice. Chapman & Hall.

Harirchi I, Karbakhsh M, Kashefi A, et al. (2004). Breast cancer in iran: results of a multi-center study. *Asian Pac J Cancer Prev*, **5**, 24-7.

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

Kavousi A, Meshkani MR, Mohammadzadeh M (2009). Spatial analysis of relative risk of lip cancer in Iran: a Bayesian approach. *Environmet*, **20**, 347-59.

Knorr-Held L (2000). Bayesian modelling of inseparable Space-time variation in disease risk. *Stat Med*, **19**, 2555-67.

Knorr-Held L, Besag J (1998). Modelling risk from a disease in time and space. *Stat Med*, **17**, 2045-60.

Lamyian M, Hydarnia A, Ahmadi F, et al (2007). Barriers to and factors faciliating breast cancer screening among Iranian women: a qualitative study. *East Medit Health J*, **13**, 1160-9.

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

Lawson AB, Browne WJ, Vidal Rodeiro CL (2003). Disease Mapping with WinBUGS and MLwiN. John Wiley & Sons Ltd.

Lunn D, Spiegelhalter D, Thomas A, et al (2009). The BUGS project: evolution, critique, and future directions. *Stat Med*, **28**, 3049-67.

MacNab YC, Dean CB (2002). Spatio-temporal modelling of rates for the construction of disease maps. *Stat Med*, **21**, 347-58.

Mahaki B, Mehrabi Y, Kavousi A, et al (2011). Multivariate disease mapping of seven prevalent cancers in Iran using a shared component model. *Asian Pac J Cancer Prev*, **12**, 2353-8.

Maiti T (1998). Hierarchical Bayes estimation of mortality rates for disease mapping. *J Stat Plan Inf*, **69**, 339-48.

Montazeri A, Vahdaninia M, Harirchi I, et al (2008). Breast cancer in Iran: need for greater women awareness of warning signs and effective screening methods. *Asia Pac Fam Med*, **7**, 6.

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

Mousavi SM, Mohagheghi MA, Mousavi Jerrahi A, et al. (2008). Outcome of breast cancer in iran: a study of Tehran cancer registry data. *Asian Pac J Cancer Prev*, **9**, 275-8.

Mousavi SM, Montazeri A, Mohagheghi MA, et al (2007). Breast cancer in Iran: an epidemiological review. *Breast J*, **13**, 383-91.

Noroozi A, Jomand T, Tahmasebi R (2011). Determinants of breast self-examination performance among Iranian women: an application of the health belief model. *J Canc Educ*, **26**, 365-74.

Sadjadi A, Hislop TG, Bajdik C, et al (2009). Comparison of breast cancer survival in two populations: Ardabil, Iran and British Columbia, Canada. *BMC Cancer*, **9**, 381.

Sadjadi A, Malekzadeh R, Derakhshan MH, et al (2003). Cancer occurence in ardabil: results of a population-based cancer regisry from Iran. *Int J Cancer*, **107**, 113-8.

Sadjadi A, Nouraie M, Ghorbani A, et al (2009). Epidemiology of breast cancer in the Islamic Republic of Iran: first results from a population-based cancer registry. *East Medit Health J*, **15**, 1426-31.

Semnani S, Roshandel G, Keshtkar A, et al (2008). Annual Report of Golestan Population-based Cancer Registry; Number 3, 2006.

Somi MH, Farhang S, Mirinezhad SK, et al (2008). Cancer in east azerbaijan, Iran: results of a population-based cancer registry. *Asian Pac J Cancer Prev*, **9**, 327-30.

Taghavi A, Fazeli Z, Vahedi M, et al (2012). Increased trend of breast cancer mortality in Iran. *Asian Pac J Cancer Prev*, **13**, 367-70.

Taheri NS, Bakhshandeh Nosrat S, Aarabi M, et al (2012). Epidemiological pattern of breast cancer in Iranian women: is there an ethnic disparity? *Asian Pac J Cancer Prev*, **13**, 4517-20.

Waller LA, Carlin BP, Xia H, et al (1997). Hierarchical spatio-temporal mapping of disease rates. *J Amer Stat Assoc*, **92**, 607-17.