Spatio-temporal pattern of two common cancers among Iranian women: An adaptive smoothing model

Mehdi Raei, Volker Schmid, Majid Moayyed, Behzad Mahaki

1Student Research Committee, School of Health, Isfahan University of Medical Sciences, Isfahan, Iran; 2Bioimaging Group, Department of Statistics, Ludwig-Maximilians-Universität München, Munich, Germany; 3Department of Economic Studies, School of Business, University of Dundee, Dundee, UK; 4Department of Epidemiology and Biostatistics, School of Health, Isfahan University of Medical Sciences, Isfahan, Iran; 5Department of Epidemiology and Biostatistics, School of Health, Kermanshah University of Medical Sciences, Kermanshah, Iran.

Summary

Purpose: Considering the increase in incidence of breast and cervix uteri cancers in Iran, this study investigates spatio-temporal patterns of the incidence of these two cancers by estimating the step changes between pairs of adjacent regions and between the Iranian women from 2004 to 2009.

Methods: Using an adaptive smoothing model, spatio-temporal mapping of the breast and cervix uteri cancers and their changes were studied. Identification of step changes between the neighboring spatial units was carried out by modeling adjacency matrix elements as random variables.

Results: There was a high relative risk of breast cancer around the central northern half of Iran, and a high relative risk of cervix uteri cancer was seen in the northeastern part of Iran. Northwest and southeast of Iran had a relatively low risk of breast and cervix uteri cancer. In general, step changes were largely similar between the two diseases with an agreement coefficient of 56%. This was observed in the Chaharmahal & Bakhtiari, and Kohgiluye & Boyerabadi provinces on the central band of Iran, as well as some eastern and northern regions on the map that were distinct from their adjacent provinces from the aspect of relative risk of both cancers.

Conclusion: Identifying areas with high/low incidence risk can help health authorities to make better decisions to prevent and control breast and cervix uteri cancers and allocate resources more efficiently. In addition, determining and identifying the step changes in unexplained components of the disease risk can lead to a deeper understanding of the spatial structure of unmeasured confounding factors.

Key words: adaptive smoothing model, breast cancer, cervix cancer, disease mapping, Iran, spatio-temporal analysis

Introduction

Women are about half of the world’s population, and their proper health is of particular importance as it ensures the health of the community [1]. According to the World Health Organization, 25% of women’s deaths are due to malignant tumors [2]. Breast cancer is the most common malignancy of women throughout the world and accounts for 50% of all cancers in women [3]. In 2000, more than one million breast cancer cases were diagnosed among women, with the lowest incidence of breast cancer in Asia and the highest in the United States and in northern Europe [4]. Although the incidence of breast cancer in Iranian women (22.2 per 100000) is about one-fifth of the incidence rate in western countries, more than 7000 patients are added annually and the incidence of
breast cancer among Iranian women is increasing [3,4].

Cervical cancer has been one of the most common cancers in recent years, and at the same time, breast cancer has grown dramatically over the years [5]. Cancer of the cervix is the second most common malignant neoplasm in women after breast cancer and is the third most common cancer of the female genital system. Although this cancer is highly preventable, it is still considered as the most common cause of cancer deaths among women in most of the developing countries, as well as in some parts of the world including Africa and South Asia [6]. Every year, approximately half a million new cases are diagnosed. In Iran, it is one of the most common cancers among women as well and is the third most common cause of death after heart disease and accidents [7]. Although the incidence of cervix uteri cancer in Iran is not higher than the other countries in the world, its mortality rate is about 40% of patients [8].

Although breast and cervical cancer are more easily treated in case of early diagnosis and treatment and it is possible to control and treat them completely, the incidence of these two cancers in developing countries is increasing significantly [9]. It is estimated that more than 270,000 and 508,000 women die every year from cervix and breast cancer worldwide, respectively [10]. Several factors have an effect on the incidence of breast and cervical cancers, including genetic factors and race, obesity, environmental and geographical factors, cultural factors, life habits, diet, ethnicity, alcohol intake, smoking, and physical activity [11].

Cancer information is being reported in many countries, including Iran, where data are frequently and annually recorded in all provinces. Data collection from a similar population in consecutive periods and neighborhood effects are important factors in the development of spatial and temporal autocorrelations within these data [12]. Without considering the spatio-temporal correlations of the data reduces the accuracy of the results of the statistical analysis [13].

In recent years, improvements have been made in the statistical models and spatial data analysis techniques. Spatial models have now been expanded to various spatio-temporal models [14]. In this case, and in models for disease mapping, spatial patterns are estimated in the incidence of disease. The purpose of the disease mapping studies is to express the spatial variation of the disease risk in order to evaluate and determine the amount of spatial heterogeneity and related patterns. These patterns represent the risk surface for each time period considering the known covariates and a set of spatially smooth random effects. Spatially smooth random effects represent the unmeasured spatial confounders that are modeled in a hierarchical framework by a prior Gaussian Markov random field (GMRF) [15,16]. In GMRF models, the spatial structure of random effects is often characterized by a spatially smooth evolution between the pairs of adjacent regional units, while in practice, some of the other pairs of regions indicate large step changes (a significant difference in the relative risks of the two regions/areas that are located beside each other or in proximity to each other). This spatial heterogeneity is not compatible with conventional smoothing models [16].

Recently, boundary analysis is used to determine the areas that are significantly different from other areas among a wide range of researchers in different fields. In fact, the boundary analysis focuses on identifying and locating abrupt changes [17]. For this purpose, the generalization of GMRF models has been investigated in various studies, in which there is a possibility of a change in the spatio-temporal smoothing in different regions [18-20]. Some other researchers have considered the extension and generalization of these models by considering non-zero elements of the adjacency matrix as a random variable [21,22]. In this regard, adaptive smoothing model presented by Rushworth et al. offers advantage in two parts including fitting the model and estimating the incidence of disease risk in the simulated study [16]. This adaptive smoothing model is used to map the relative risk (RR) of diseases and in the spatio-temporal data, in which the regions differing from each other are distinguished.

Distribution of the incidence in some cancers and the existence of changes in their temporal trend are confirmed in various studies in different provinces of Iran [23-29]. However, spatio-temporal patterns are not properly considered in these studies. On the other hand, no attention is paid to determine the step changes in the unexplained component of disease risk. Using the appropriate statistical model and accurate estimation of relative risk of the disease can be effective in determining the correct pattern of it. Moreover, due to the increasing incidence of breast and cervical cancer in Iran, and because many women are referred during the advanced stages of the disease [30], in this study an adaptive smoothing model was used for the first time to estimate the spatio-temporal pattern of breast and cervix uteri cancer in different regions of Iran. In addition, step changes in disease risk between the pairs of adjacent regions during the period of 2004 to 2009 were also investigated.
Methods

Data
We carried out an applied ecological study that concentrated on spatiotemporal distribution through spatiotemporal adaptive smoothing model. We used information on two common cancers of breast (C50) and cervix uteri (C53) in women from 30 provinces of Iran during 2004-2009 that were collected from the annual reports of the Iran National Cancer Registry published by the Center for Disease Control of the Ministry of Health and Medical Education. These data are available until 2009.

Modeling
The observed and expected counts for the incidence of cancer in the ith province and jth year are denoted by \( Y_{ij} \), \( E_{ij} \) respectively and the Poisson log-linear model for these data are considered as follows:

\[
Y_{ij} \mid E_{ij}, RR_{ij} \sim \text{poisson } (E_{ij} \cdot RR_{ij}) \quad i=1,...,30, j=1,...,6
\]

\[
\ln (RR_{ij}) = \Phi_{ij} + \Psi_{ij}
\]

\( RR_{ij} \) represents the relative risk and \( \Psi_{ij} \) indicates the latent component for the ith areal unit and jth time period, which includes spatio-temporally autocorrelated random effects. In fact, spatio-temporal autocorrelation is considered in response variables via latent random effects components (Ψ) and using CAR type prior distributions. In the above model offsets are shown with \( O_{ij} \).

Likelihood and random-effects model for \( (Y_{ij}, \Psi_{ij}) \)

The adaptive smoothing spatio-temporal model presented by Rushworth et al. (2017) was used in this study [16]. Using the Bayesian method for modeling these data, the spatio-temporal structure was modeled through a set of autocorrelated random effects. In these cases, the extension of the conditional auto-regressive priors (CAR), which are special cases of GMRF, were assigned to random effects components (Ψ) and using CAR type prior distributions. In the above model offsets are shown with \( O_{ij} \).

At the first level of the proposed model, following were considered:

\[
\Psi_{ij} = \Phi_{ij} + \Phi_{ij-1} + \Phi_{i-1,j} - N (\rho_{ij}\Phi_{ij-1}, \tau^{2}\{Q(W, \rho_{ij})\}^{-1}) \quad j=2,...,6
\]

\[
\Phi_{ij} = N (0, \tau^{2} \{Q(W, \rho_{ij})\})
\]

\[
\tau^{2} \sim \text{Inverse - Gamma } (1, 0.01)
\]

\[
\rho_{ij}, \rho_{ij} \sim \text{Uniform } (0,1)
\]

The conditional expectation of \( \Phi \) (space-time effect) is a weighted average (which is determined by \( W \)) of the adjacent \( \Phi \), that induces smoothness across the surface. Temporal autocorrelation and spatial autocorrelation were considered through \( \rho_{ij}\Phi_{ij-1} \) and \( \tau^{2}\{Q(W, \rho_{ij})\}^{-1} \) in the model, respectively. Precision matrix was also defined as \( Q(W, \rho_{ij}) = \rho_{ij}\text{diag}(W1)W + (1-\rho_{ij})I \) [32]. The temporal and spatial correlations were controlled by \( \rho_{ij} \in [0,1] \) and \( \rho_{ij} \in [0,1] \), respectively.

The spatial smoothing between random effects was achieved through the GMRF priors using the W-adjacency matrix. W matrix elements represented spatial closeness between the two regional units. Typically, the W-matrix elements were defined as binary [33], however, in this study, step changes between neighboring spatial units was achieved through modeling W elements, i.e. \( W = \{W_{ij} \mid i \sim k\} \), that were considered as random variables. Areas (i, k) are neighbors if areas i and k share a common border and is denoted by i \sim k. In this case, the estimating \( W_{ij} \in W \) near 1 led to partial auto-correlation and consequently, smoothing between \( (\Phi_{ij}, \Phi_{ik}) \) for all time periods of j. Conversely, if \( W_{ij} \) was estimated close to zero, then \( (\Phi_{ij}, \Phi_{ik}) \) were conditionally independent for all time periods of j, and no spatial smoothing would occur. In the second case, there was a step change at the level of random effects between the regional units (i,k) at all time periods of j.

In order to determine the step changes, as reported in the study of Lu and Carlin (2005) [34], initially the posterior probability of \( W_{ij} \) was considered less than 0.5, equal to \( \rho_{ij} \), and as follows:

\[
\rho_{ij} = P (W_{ij} < 0.5 \mid Y)
\]

Then, based on \( \rho_{ij} \), values more than the threshold of 0.75 were considered as a step change.

The model was fitted in a Bayesian framework using R 3.4.4 software. All the inferences were based on the Markov Chain Monte Carlo sampling. For the model, the Markov chain was executed with 100,000 iterations. Burn-in period was performed with 20,000 repetitions and the results were reported on the basis of 80,000 remaining posterior samples with thin 10 to reduce the temporal autocorrelation (after which convergence was assessed to have been reached). Convergence was checked by the calculation of the Geweke diagnostic [12]. All the maps were produced with ARCGIS 10.4.1 software.

Results

The incidence of breast cancer was 3803, 4621, 5578, 6181, 8145 and 7840 in 2004-2009, respectively. Also, the incidence of cervix uteri cancer during these 6 years was 394, 358, 436, 533, 630, and 572, respectively. Based on these results, despite slight fluctuations in some of the years, the incidence of both cancers in the whole areas of Iran showed an increasing trend.

The temporal trend in the relative risk of breast and cervix uteri cancer is shown in Figure 1. According to Figure 1, changes in the level of spatial variations from year to year, as well as larger amounts of spatial variation before 2006 can be seen.

The posterior median relative risk based on the adaptive smoothing model was computed and then plotted for the 6 years (Figure 2 and Figure 3). The maps show clear changing spatial pattern in relative risks over time.

According to the results, in the case of breast cancer, the provinces of Tehran, Esfahan, Mazandaran, Guilan, Khuzestan, Fars, Yazd and Razavi Khorasan were considered as high-risk areas.
The provinces of North Khorasan, Kohgiluyeh & Buyerahmad and Sistan & Baluchestan had the lowest relative risk level during each of the six periods.

In the case of cervix uteri cancer, the provinces of Tehran, Golestan, Razavi Khorasan, North Khorasan, Khuzestan, Esfahan, and Kerman showed the higher relative risk level. After these provinces, Mazandaran, Fars, Yazd and Hormozgan provinces were considered as more risky areas than the other provinces. The provinces of Tehran, Golestan and Razavi Khorasan were at the highest relative risk level during all the six periods. Chaharmahal & Bakhtiari and Sistan & Baluchestan provinces had the lowest relative risk during all the six years.

Estimates of parameters, model fitting indices, and number of step changes in the unexplained component of the risk surface are presented in Table 1. Also, Convergence diagnostics revealed an acceptable degree of convergence.

![Figure 1](image1.png)

**Figure 1.** Boxplots showing the temporal trend in the relative risk of breast (A) and cervix uteri (B) cancers between 2004 and 2009.

![Figure 2](image2.png)

**Figure 2.** Maps showing the changing spatial evolution of the posterior median estimated relative risk due to breast cancer for 6 years.
Spatio-temporal pattern of two common cancers in Iranian women

Figure 3. Maps showing the changing spatial evolution of the posterior median estimated relative risk due to cervix uteri cancer for 6 years.

Figure 4. Maps showing the average relative risk of breast cancer (A) and cervix uteri cancer (B) between 2004 to 2009: step changes that have been identified by using a cut-off of $P_{ik}>0.75$.

Table 1. Diagnostics for adaptive smoothing model for the breast and cervix uteri cancers

<table>
<thead>
<tr>
<th>Diagnostics</th>
<th>Breast cancer</th>
<th>Cervix uteri cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>median</td>
<td>Geweke.diag</td>
</tr>
<tr>
<td>DIC$^1$</td>
<td>2463.22</td>
<td>-</td>
</tr>
<tr>
<td>$pD^1$</td>
<td>163.48</td>
<td>-</td>
</tr>
<tr>
<td>% of borders with $P_{ik}&gt;0.75$</td>
<td>32.39</td>
<td>-</td>
</tr>
<tr>
<td>$\tau^2$</td>
<td>0.11</td>
<td>-1.74</td>
</tr>
<tr>
<td>$p_r$</td>
<td>0.86</td>
<td>-0.5</td>
</tr>
<tr>
<td>$p_s$</td>
<td>0.27</td>
<td>-1.95</td>
</tr>
</tbody>
</table>

$^1pD$: effective number of parameters, DIC: deviance information criterion
The mean relative risk over 6 years is displayed in Figure 4. As seen, relatively similar spatial patterns with Pearson correlation coefficient of 0.51 were found between the two cancers. The results showed that the average relative risk varied over provinces with values between 0.24 and 1.51, and 0.42 and 1.55 respectively for breast and cervix uteri cancer.

According to Table 1, the step changes at the borders for breast and cervix uteri cancer were 52% and 25%, respectively, as shown in Figure 4 (a) and 4 (b), with black lines on the maps. Therefore, in the case of relative risk for both cancers, some provinces including south Khorasan, and Sistan & Baluchestan in the southeast, and Chaharmahal & Bakhtiar and Kohgiluye & Boyerahmad were distinguished on the map from adjacent provinces. In addition, the provinces of Mazandaran, Tehran, and Guilan in the central band of northern Iran were different in the relative risk of breast cancer from their adjacent areas. Generally, step changes were largely similar between the two cancers with 56% agreement between their locations.

**Discussion**

This study aimed to determine the spatio-temporal patterns of breast and cervix uteri cancers among Iranian women by adaptive smoothing model. The advantage of this model was that we could estimate the spatio-temporal patterns of diseases as well as estimating step changes between neighbouring provinces. Oversmoothing of the step changes in conventional models [35-37] leads to poor estimates of the relative risk of disease, failure to identify step changes, and subsequently the unidentification of potentially unmeasured risk factors.

Based on the results of this study, there was a high relative risk of breast cancer around the central band of Iran. More access to cancer recognition facilities, environmental pollution issues, and changes in the life habits of general population in these areas can be considered as the probable causes. The southeast and the western regions had a relatively low risk of breast cancer. The high/low relative risk of breast cancer in various Iranian provinces is reported in other studies; however, differences in the results of cancer mapping in Iran can be justified by the manner in which different models are employed in these studies [26,28,35].

According to the results of this study, there was a high relative risk of cervical cancer in the northeastern part of Iran. Northwest and southeastern parts of Iran had a relatively low risk of cervical cancer. The epidemiology of cancer in each region depends on gender, age, race, nutritional status, cultural and social customs and individual living patterns [38]. Therefore, the incidences of breast and cervical cancers vary between different parts of the world [39]. Differences between countries across the globe and fluctuations in numbers and values within a country are affected by the change in the prevalence of risk factors and the secular trends in the diagnosis of the disease.

In general, although the relative risk values of breast and cervical cancer were somewhat different, a relatively similar spatial pattern existed for the relative risk of these two cancers. The relatively high values of temporal and spatial correlation parameters in this study also showed that there are some spatio-temporal autocorrelations in the data. With the increase in life expectancy and increased index of aging in the Iranian population, it is expected that the incidence of various cancers during the next years will also increase [3]. This will lead to increased burden of human and financial investment. It is believed that the tendency to the Western lifestyle is one of the main causes of this disease in women [1].

As the adaptive smoothing models are capable of identifying step changes, in this study the location of abrupt changes and the distinction of adjacent regions with significant differences in relative risks were clearly seen on the map. The similarity of many features in Sistan & Baluchestan and South Khorasan, Chaharmahal & Bakhtiari and Kohgiluye & BoyerahGand provinces and also in Tehran, Mazandaran and Guilan can be a major factor in distinguishing these provinces from their adjacent areas. The results are nearly consistent with the results of the Ministry of Interior’s latest zoning [40]. This zoning has been done according to neighborhoods and common factors between the provinces. In health sciences, it is important to identify areas that vary in terms of disease or mortality from other areas, because it can help health authorities to make better decisions to prevent and control the disease and even allocate resources more efficiently [17]. In most of the cases, estimating and determining step changes in unexplained components of the disease can lead to a deeper understanding of the spatial structure of unmeasured confounding factors. It can also help to identify the geographical extent of regions with higher unexplained level of risks [16]. In addition, the association between the estimated relative risks and identified step changes between breast and cervical cancers can be indicative of the effect of the same unobserved risk factors on these two diseases.
Conclusions

Disparities in the relative risks between different provinces for breast and cervix uteri cancers suggest the presence of substantial health inequalities. Considering the obvious differences and distinctions between high/low risk provinces, and identifying the pattern of breast and cervical cancers, the results of this study can help health policymakers to prevent these diseases based on the geographic, cultural and social characteristics of the country. According to the results, and considering the impact of socioeconomic factors on the development of these cancers [41], it is necessary to inform women about their health, raise awareness about the mortality of this disease, encourage participation in the cancer screening and comprehensive training programs especially for the deprived economic class. In addition, publish health professionals can identify potential risk factors for the two cancers by searching for the risk factors that show step changes in similar locations.

Acknowledgments

This study as a part of PhD thesis was supported by the Research Deputy of Isfahan University of Medical Sciences (Study project number: 396547). The authors want to thank the Vice Chancellor for Research and Technology, Isfahan University of Medical Sciences.

Conflict of interests

The authors declare no conflict of interests.

References


