Mapping of Stomach, Colorectal, and Bladder Cancers in Iran, 2004–2009: Applying Bayesian Polytomous Logit Model

Abstract

Background: According to the last report of Iran cancer registry, stomach, colorectal, and bladder cancers are the most prevalent cancers. The present study focused on separating the latent risk surface into shared and disease-specific components. Methods: In this study, data consisting of stomach, colorectal, and bladder cancers in 30 provinces of Iran during 2004–2009 are considered. These data are analyzed by polytomous logit model. The incidence of stomach cancer acts as the reference category (the surrogate for smoking). Then, the log odds are decomposed into shared and specific structured spatial and unstructured spatial components. These latent components help to detect spatial patterns of shared and disease-specific risk factors. Results: Central, Southern, Eastern, and Southwestern provinces are supposed as high-risk regions for shared risk factor for colorectal and bladder cancers. This shared risk factor is slightly associated more with bladder than with colorectal cancer. Northern, northwestern, and central regions and also three borderline provinces in southwestern are high-risk regions for colorectal cancer. Central, eastern, southern, and western strip of the country except Ilam are found as the high-risk regions of bladder cancer. Conclusions: After considering known shared risk factor of the three cancers, it turns out that colorectal and bladder cancers have unknown shared risk factor. The significant difference in their lifestyle and eating habits could be an assumption of the risk factor.

Keywords: Colorectal neoplasms, disease mapping, polytomous logit model, shared component model, stomach neoplasms, urinary bladder neoplasms

Introduction

Cancer is the second leading cause of deaths worldwide so that there were about 15 million incident cancer cases and over 8 million cancer deaths worldwide in 2013.[1,2]

In Iran, cancer is the third cause of death. According to the last report of Iran cancer registry, stomach, colorectal, and bladder cancers are the most prevalent cancers.[3]

Estimating geographical distribution of cancer can lead to recognize environmental risk factors, determination of high-risk regions, and also obtain more knowledge about the process of developing cancer in the world or regional levels. For achieving precise estimation of disease distribution in recent years, different mapping methods have been used by the researchers.[3,4]

Disease mapping is a series of statistical methods in which their aim is to gain precise estimation of incidence, prevalence, survival, or mortality rate of disease and regulate them in the form of geographical maps.[3] Most of the studies in the field of disease mapping have focused on the modeling of one disease-relative risk.[4]

However, joint modeling of disease is preferred more than other univariate models due to its simple interpretation, considerable improvement in the deviance information criteria, and increase in estimations’ preciseness. By considering the present correlation between the regions and disease, more desirable patterns of disease changes in different regions and the relation of the diseases with each other are found as well.[5] In recent years, a number of studies have been conducted about mapping cancer in Iran among which few ones have been focused on mapping with emphasis on the role of control risk factors.[6-13]

The main aims of the study are to model jointly the incidence rates of stomach, colorectal, and bladder cancers in provinces of Iran during 2004–2009 and separate the latent risk surface into shared and disease-specific components. By applying polytomous logit (PL) model,[14] we separate

common and uncommon spatial patterns of colorectal and bladder cancers; where stomach cancer is considered as the reference category and also the surrogate for shared risk factor. We also estimate the relative weight of the shared component for each disease.

Knowledge of the spatial variation of the shared and specific risk factors can help politicians to evaluate healthcare system performance; better allocation of resources and implementation of prevention policies in sensitive regions.

Methods

This is an ecological study, based on data including stomach (ICD10 code C16), colorectal (C18–C20, C26), and bladder (C67) cancers, in 30 provinces of Iran during 2004–2009. The data have been collected and made available by the Iranian Ministry of Health and Medical Education.[15-20]

To joint modeling of spatial variation related to stomach, colorectal, and bladder cancers, PL model is used with Bayesian approach.[14] It is supposed that is nominal responsive variable including the observed count of stomach, colorectal, and bladder cancers cases in 30 provinces and follows a multinomial distribution with parameters and probability vector \( \pi_{ik} = (\pi_{i1}, \ldots, \pi_{ik}) \), where

\[
\sum_{k=1}^{K} v_{ik} = \pi_{ik} = 1.
\]

Thus, the types of colorectal and bladder cancers are compared with stomach cancer as a reference category; each log odd is modeled as follows:

\[
\log(\phi_{ik}) = \alpha_k + u_{ik} + v_{ik}
\]

\[
\phi_{ik} = \pi_{ik}/\pi_{ik}
\]

Where \( \alpha_k \) is disease-specific intercept of every disease and \( u_{ik} \) and \( v_{ik} \) are structured and unstructured spatial terms, respectively. The amount of parameters was zero for reference category (stomach cancer) \( (\alpha_3 = v_{ij} = u_{ij} = 0) \).

Considering stomach cancer as the reference, the effect of the shared risk factors on spatial changes is adjusted. Flat noninformative distribution is specified to disease-specific intercept \( \alpha_k \). Clustering component \( u_{ik} \) is modeled by use of Gaussian Markov random fields (GMRFs) to cover spatial structure. The heterogeneity terms \( v_{ik} \) are independent, each \( v_{ik} \) being \( N(0, \lambda^{-1}) \) \( (\lambda_{ik} \) represents the precision parameter).[21]

By use of shared component model,[22] spatially structured effect is analyzed into shared and specific effects so that every spatially component includes a coefficient to assess the relative contribution of this component to each related disease.

\[
u_{ik} = u_{ij} \cdot \delta_1 + u_{ij}
\]

\[
u_{ik} = u_{ij} \cdot \delta_2 + u_{ij}
\]

The \( u_{ij} \) term presents the shared clustering component and \( u_{ij} \) and \( u_{ij} \) present colorectal and bladder cancers’ specific clustering component, respectively. Both \( u_{ij} \) and \( u_{ij} \) are modeled by use of GMRFs.

Terms \( \log \delta_1 \) and \( \log \delta_2 \), constrained to \( \sum_{k=1}^{2} \log \delta_k = 0 \) are assumed to be multivariate normal distributed with zero mean and variance-covariance matrix \( \Sigma_\delta = \sigma_\delta^2 \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \).

Model fitting to the data was done by use of Bayesian estimation in OpenBUGS version 3.2.1 (rev 781) software. The estimation of posterior distribution was made by Monte Carlo Markov Chain for the model parameters.

Results

In Figure 1, the map of estimating \( \exp(\delta_s) \) related to the shared structured component is shown from 2004 up to 2009 for colorectal and bladder cancers. Since this surrogate of risk factor is distributed in all the provinces, the country can be divided into high-risk and low-risk regions. Northern regions including northeastern and western provinces and also those ones at the Caspian seas’ margin are known as low-risk regions. Central, southern, southwestern, and eastern provinces are supposed as high risky regions. Since the shared risk factor is adjusted, it can be correlated with other kinds of risk factors which affect the prevalence of these two diseases rather than shared risk factors in the western regions of the country [Figure 1].

The posterior medians for scale parameters associated with shared component are \( \delta_1 = 0.954 (0.79–1.14) \) and \( \delta_2 = 1.047 (0.87–1.25) \) for colorectal and bladder cancers, respectively; \( \delta_1/\delta_2 \) which indicates that shared risk factor is slightly associated more with bladder than with colorectal cancer.

Figure 1: Map of estimating \( \exp(\delta_s) \) representing the surrogate component of shared risk factor.
In Figure 2, the distribution of spatially specific structured component for colorectal and bladder cancers (exp \( \exp[i_1] \) and \( \exp[i_2] \)), in all provinces, are presented. Two observed patterns are totally different due to adjusting the shared risk factor effect of two cancers. The regions with the amount more than 1 are called as the regions exposed to risk of these factors.

North, North-Western, and central regions of the country and also three borderline provinces were found to have the highest risk for colorectal cancer in the whole period of time. These regions are so affected by hypoactivity risk factors.

Central, eastern, southern, and western strip of the country except Ilam are found as high-risk regions of bladder cancer. The severity of this factor worsens by getting far from the northern regions of the country [Figure 2].

**Discussion**

PL model was used to joint modeling of stomach, colorectal, and bladder cancers in provinces of Iran during 2004–2009 and was separated the latent risk surface into a shared and a disease-specific components.

Several studies suggest that smoking is a major known risk factor for three mentioned diseases.\(^{24-28}\)

The findings show the possibility of presence of an unknown risk factor, mostly, in central and southern regions of the country. The significant difference in their lifestyle and eating habits could be an assumption. However, recognizing this risk factor needs to conduct epidemiological studies in this field. The relative effect of this risk factor is more for bladder cancers rather than colorectal cancer.

Based on the previous studies, two risk factors, including, hypoactivity\(^{11,24,25,28}\) and occupational exposure\(^{28-30}\) are considered as the specific risk of colorectal and bladder cancers, respectively.

The provinces at the margin of Caspian Sea (Golestan, Gilan, and Mazandaran), Tehran, Qazvin, and Qom are called high-risk specific risk factor for colorectal cancer. These regions are so affected by hypoactivity risk factors. The same results were found in the studies conducted in Iran previously.\(^{11,31}\)

Sports development and exercise interventions are one of the most useful ways to fight colorectal cancer. The studies show that the people being active physically are caught by this disease, 20%–30% less than the others.\(^{12,33}\) Unfortunately, 69% of Iranian mature people do not have a daily regular physical activity.\(^{34}\)

Central, eastern, southern regions, and western strip of the country except Ilam province are known as high-risk regions for bladder cancer and are affected by occupational exposure risk factor during the study period. Boushehr, Fars, Kerman, and South Khorasan are known as the highest risk regions.

Numerous studies show that 20% of bladder cancer are caused by occupational exposure to bladder carcinogens that occur in industrial regions processing paint, dye, metal, pharmaceutical, and petroleum products.\(^{29,30}\) In some developing countries, by controlling occupational exposures to bladder carcinogens, cases of bladder cancer have decreased.\(^{35,36}\)

The PL model provides the possibility of analyzing the rates of some diseases jointly without having information of population at risk.\(^{14}\) Since the interpretation, in this model, is done according to log odd, the possibility of comparing and determining the priorities is provided to implement the interventions and specify the resource.

By changing the baseline category, it is probable to recognize other shared and specific risk factors for these series of cancers.

It is suggested to perform this model on smaller regions such as counties to access more precise estimations of the shared and specific effects.

It is also suggested to consider the time dimension for a larger dataset in order to recognize the temporal trend, in addition to the spatial pattern of the diseases or to account the possible difference on latency time between the first exposure to risk factors and incidence of these diseases.\(^{11,14}\)
By applying PL model, without having information of risk factor and population at risk, the surrogate of shared and specific risk factors are estimated and are regulated them in the form of geographical maps. Based on the previous studies, two risk factors, including, hypoactivity and occupational exposure are considered as the specific risk of colorectal and bladder cancers, respectively. The possibility of presence of an unknown shared risk factor of colorectal and bladder cancers is reported. The significant difference in their lifestyle and eating habits could be an assumption. However, recognizing this risk factor needs to conduct epidemiological studies in this field.

Knowledge of the spatial variation of the shared and specific risk factors can help politicians to evaluate healthcare system performance; better allocation of resources and implementation of prevention policies in sensitive regions have been detected.

Since this is an ecological study and accompanies with some biases, the result is not used in personal level, but they can be followed as an appropriate goal to find the causes.

Conclusions

After considering known shared risk factor of the three cancers, it turns out that colorectal and bladder cancers have unknown shared risk factor. The significant difference in their lifestyle and eating habits could be an assumption of the risk factor.

Acknowledgments

The authors are grateful to express their sincere thanks to the office of Noncommunicable disease, Ministry of Health and Medical Education, Iran. This work was partially supported by the student Research Center, School of health, Isfahan University of Medical Sciences, Isfahan, Iran (code: 3941060). This article has been extracted from MSc thesis in Biostatistics at School of health at Isfahan University of Medical Sciences.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

Received: 16 Jan 17 Accepted: 09 Jan 18 Published: 05 Dec 18

References