RESEARCH ARTICLE

# Identifying High-Risk Clusters of Gastric Cancer Incidence in Iran, 2004 - 2009

**Amir Kavousi[1], Yousef Bashiri[2]*, Yadollah Mehrabi[3], Korosh Etemad[3], Amir Teymourpour[2]**

## Abstract

**Background: Gastric cancer is considered as the second most prevalent cancer in Iran. The present research sought to identify high risk clusters of gastric cancer with mapping using space-time scan statistics. Materials and Methods: The present research is of descriptive type. The required data were gathered from the registered cancer reports of Cancer Control Office in the Center for Non Communicable Disease of the Ministry of Health (MOH). The data were extracted at province level in the time span of 2004-9. saTScan software was used to analyse the data and to identify high risk clusters. ArcGIS10 was utilized to map the distribution of gastric cancer and to demonstrate high risk clusters. Results: The most likely clusters were found in Ardabil, Gilan, Zanjan, East-Azerbaijan, Qazvin, West-Azerbaijan, Kurdistan, Hamadan, Tehran and Mazandaran between 2007 and 2009. It was statistically significant at the p-value below 0.05. Conclusions: High risk regions included Northern, West-North and central provinces, particularly Ardabil, Kurdistan, Mazandaran and Gilan. More screening tests are suggested to be conducted in high risk regions along with more frequent epidemiological studies to enact gastric cancer prevention programs.**

**Keywords:** Gastric cancer - space-time scan statistics - cluster identification - mapping

*Asian Pac J Cancer Prev,* **15 (23)**, 10335-10338

## Introduction

According to the estimations of the International Agency for Research on Cancer (IARC) in 2012, gastric cancer is the fifth most common cancer in the world. Half of the cases occur yearly in East Asia and mostly in China. Gastric cancer is third cancer mortality of both sexes in the worldwide. Moreover, in 2012 in Iran gastric cancer accounted for 11.4% of all cases of cancer occurrence and was considered as the second most prevalent cancer. In the same year, gastric cancer which accounted for 15.5 percent of all mortalities caused by cancers was in fact the most deadly cancer in Iran (GLOBOCAN., 2012). Cancers of the metabolic system are among the most deadly cancer types. According to physicians, environmental/ biological factors and nutritional culture are involved in its occurrence (Gelaman et al., 2003). In developed countries the occurrence of this cancer is being reduced due to proper interventions such as healthcare education, nutrition recognition and control of stimulating behaviors. In developing countries, however, due to aging, nutrition habits, improper behaviors such as alcohol consumption or cigarette smoking, the incidence of gastric cancer is increasing (Asmarian et al., 2012).

Recognizing the important role of spatial location in human health has been the center of thinkers' attention from long ago. Over 2400 years ago, Hippocrates regarded life environment and life style among the key determinants of public health in society. Avicenna is one of the first Iranian scientists who discussed the effect of life environment on human health. Since diseases and injuries do not have the same distribution in society, in recent years there has been a closer investigation of the living place in human health by using geographical distribution map of health indices. This kind of maps that is related to the health has become more important than before for societies' healthcare decision makers in order to prevent and treat patients more than before.

In 1989, Knox defined a disease clustering as a series of events in one geographical region which is higher in frequency and density than being merely induced by chance. If the occurrence cases of a disease are high in some regions than others and such cases also vary within time, the consequent cluster would be called a spatial-temporal cluster.

Scientific investigations of disease clustering are essential for the etiology, provision, maintenance and promotion of public health. Any sort of deficiency in carrying out such investigations could have irreparable consequences for public health. In investigating different types of disease clustering through simple and classical methods, the majority of disease clusters could be unreal.

*[1]Department of Basic Sciences, School of Health, Safety and Environment, [2]Department of Biostatistics, School of Paramedical Sciences, [3]Department of Epidemiology, School of Public Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran*
**For correspondence: yousef_bashiri@yahoo.com*

Therefore, a group of scientists believe that research on disease clustering ought not to impose high costs on medical and healthcare systems (Rezarian, 2012). In many applied fields, both false positive reports (incorrect report of clustering) and false negative reports (not reporting an actual clustering) would have high and irreparable costs (Neill., 2006). Therefore, in the present research using space-time scan statistics, only the statistically significant and high-risk clusters were identified.

The first disease mapping was done by John Snow in 1854 who used cholera epidemic map in London and proved the hypothesis that the factor for spreading cholera had been contaminated water (Snow., 1855). Among other studies on the correlation of cancer and geographic location, mention can be made of the correlation of breast cancer and exposure to pesticides (Brody et al., 2004), and the risk of cancer and residence near a petrochemical factory (Belli et al., 2004). In 1965, Naus demonstrated, for the first time, the function of scan statistics in one-dimensional and two-dimensional clustering of random spots (Naus.,1965a; Naus.,1965b). Blood cancer was the first type of cancer investigated through spatial scan statistics in 1990 by Turnbull et al. (Turnbull et al., 1990). In 1998, Kulldorffdorff et al. for the first time used scan statistics for spatial-temporal data known as a spatial shuttle (Kulldorff et al., 1998). In 2005, prospective space-time scan statistics were normally used to investigate disease distribution by Kuldorff et al. in Hypothyroidism cancer (Kulldorff et al., 2005). In 2007, DeChello et al. used space-time scan statistics to investigate geographical variations of colon cancer in Massachusetts (DeChello et al., 2007). In 2012, Luginaah et al. delved into the environmental effects of breast cancer using space-time scan statistics (Isaac et al., 2012). In 2014, Katamaya et al. (2014) used spatial scan statistics to detect a cluster of high breast cancer incidence in Kanagawa and also in this year, Azhar Shah et al. (2014) demonstrated spatial patterns of colorectal cancer in Kuala Lumpur.

The present research was conducted to determine the geographic pattern of gastric cancer occurrence and to identify high-risk areas for this cancer to occur within time. Findings of this study can be used as a basis for purposeful interventions by healthcare programmers and experts in society.

## Materials and Methods

This research is of a descriptive type. The data were collected according to the International Classification of Diseases (ICD-10) in Iran. In this study, the cases of gastric cancer occurrence in the whole country were 29,474 people between 2004 and 2009 in 30 provinces. They were obtained from the recorded gastric cancer reports of the Bureau of Cancer, center for Non communicable disease of the Ministry of health and medical education: (http://www.ircancer.ir).

In this study, all provinces were considered as regions. At-risk population was obtained from the census of 2006 and also the estimated populations were obtained from the website of Statistical Center of Iran (http:// www.sci.org.ir).

For rare diseases such as the different types of cancer, it is assumed that the number of patients in each region follows a Poisson distribution considering the fact that the data are numerical and discrete. Spatial scan statistics are used to identify regions where the number of patients is significantly higher than the expected. Therefore, in spatial scan statistics, the assumption of the sameness of disease risk inside and outside the cluster is tested against the assumption of a higher risk of disease inside the cluster than outside. Then, based on the maximum value of logarithm likelihood ratio would be identified the high risk cluster.

Space-time scan statistics is a simple generalization of spatial scan statistics. Space-time scan statistics is defined as one having a cylinder-like window and a circular base and a height corresponding to time. This window moves in time and space for each geographical region in the whole country. It includes all possible time intervals for the occurrence of a disease from the time the study begins till the present time. The size of the scan window is decided upon by the importance of the target disease (Neill and Moore , 2004). In the present study, its size is set to be 50% of the whole population at risk.

The required data were firstly entered into Excel software. Then, they were prepared for statistical analysis by means of SaTScan software (Kulldorff., 2014). Once the high-risk clusters were spotted, the test of cluster significance was run using Monte Carlo method. The cluster mapping of the occurrence cases of gastric cancer were illustrated in ArcGIS10.

## Results

Using the space-time analysis technique, a high-risk cluster was detected which includes 10 provinces: Ardabil, Gilan, Zanjan, East-Azerbaijan, Qazvin, West-Azerbaijan, Kurdistan, Hamadan, Tehran and Mazandaran between 2007 and 2009. It was seen that 10,856 out of 19,714 gastric cancer incident cases occurred in this cluster, while


**Figure 1. Clustering of the Occurrence Cases of Gastric Cancer Using Space-Time Scan Statistics**

**Table 1. High-risk Region Cluster for the Occurrence of Gastric Cancer in Iran using Space-Time Scan Statistics**

| Provinces | Time span | Observed cases | Expected cases | Relative risk | P-value |
|---|---|---|---|---|---|
| Ardabil, Gilan, Zanjan, Easte-Azerbaijan, Qazvin, West-Azerbaijan, Kurdistan, Hamadan, Tehran, Mazandaran | 2007-2009 | 10856 | 7824 | 1.57 | $10^{-17}$ |

7,824 cases were expected. The incidence rateof gastric cancer was computed as 11.1 per 100,000. Moreover, in the high-risk cluster the proportion of observed to expected cases was 1.39 and relative risk was estimated as 1.57 with P-value of $10^{-17}$. Relative risk was defined as the proportion of inside cluster incidence rate to that of outside the cluster. That is to say that, the chance of occurrence of gastric cancer inside the cluster was 1.57 times of the other regions (Table 1). In Figure 1, the high-risk cluster is illustrated in a gray background while other provinces are shown in yellow.

## Discussion

Occurrence of gastric cancer and the mortality caused by that is among primary healthcare issues in Iran. Climatic and environmental conditions of each region pave the way for the occurrence and prevalence of cancer. Therefore, in the present research we investigated this matter from environmental aspects and mapping using space-time scan statistics.

Among men, gastric cancer is two times as prevalent as among women. Above the age of 50, it rises drastically (Asmarian et al., 2012). A myriad of research has been conducted so far on different types of gastric cancer in Iran which have examined cancers individually (stomach, esophagus, colon). Or they have been restricted to one particular province or a group of provinces (especially Northern provinces). One such research was conducted by Mousavi Jarrahi who indicated that gastric cancer is the first most prevalent cancer among men and the fourth most prevalent cancer among women in Iran (Mosavi-Jarahi et al., 2001). Somi et al. (2014) reported a high prevalence of gastric cancer in Eastern Azerbaijan (Somi et al., 2014). In the same year, Bahrampour et al. regarded diet, environmental/biological factors and pesticides as among the underlying factors of high occurrence of gastric cancer in the north of Iran (Bahrampour et al., 2014). In 2008, Mohebbi et al. used the recorded data of cancer in the north of Iran and investigated the spatial patterns of the occurrence of gastric cancer. They found that gastric cancer in the southern coastline of the Caspian Sea has high-risk clusters and a regular geographic pattern (Mohebbi et al., 2008). In 2012, Asmarian et al. used Poisson Kriging method and estimated the rate of gastric cancer regionally in all prevences of Iran. They suggested using Poisson Kriging method in mapping heterogeneous regions in order to take into account the spatial base in the estimation of the disease risk (Asmarian et al., 2012).

The results of the present research indicate a high occurrence rate of gastric cancer in northern, north-western and some central regions of Iran (particularly Ardabil, Kurdistan, Mazandaran and Gilan). The reason of such high incidence must be traced in factors related to these regions. Besides the type of nutrition, nitrate-filled soil, nitrogen and agricultural fertilizers used in the north among the reasons for the high prevalence of gastric cancer (Salamatnews., 2012).

To precisely investigate the intervening factors of gastric cancer occurrence and the type of the obtained pattern of distribution in this study, healthcare experts and decision makers should ponder and comment. In accordance with the existing facilities, the required planning should be done. More screening tests, similar to those in Japan, of high-risk regions are recommended. Besides that, more epidemiologic research is suggested to be done in order to enact disease prevention plans.

## Acknowledgements

This article is taken from an MS thesis of Biostatistics from Shahid Beheshti University of Medical Science.

## References

Asmarian NS, Kavousi A, Salehi M, et al (2012). Mapping of stomach cancer rate in Iran using area-to-area poisson kriging. *Health System Res*, **8**, 681-7 (in Persian).

Behnampour N, Hajizadeh E, Zayeri F, et al (2014). Modeling of influential predictors of gastric cancer incidence rates in Golestan Province, North Iran. *Asian Pac J Cancer Prev*, **15**, 1111-7.

Belli S, Benedetti M, Comba P, et al (2004). Case-control study on cancer risk associated to residence in the neighbourhood of a petrochemical plant. *Eur J Epidemiol*, **19**, 49-54.

Brody JG, Aschengrau A, McKelvey W, et al (2004). Breast cancer risk and historical exposure to pesticides from wide-area applications assessed with GIS. *Environ Health Perspect*, **112**, 889.

Cancer IAfRo (2014). GLOBOCAN 2012: estimated cancer incidence, mortality and prevalence worldwide in 2012. World Health Organization. http://globocan. iarc. fr/Pages/fact_sheets_cancer. aspx. Accessed on, 9.

Neill DB (2006). Detection of spatial and spatio-temporal clusters: Doctoral dissertation, University of South Carolina.

Neill DB, Moore AW (2004). Detecting space-time clusters: prior work and new directions: report.

DeChello LM, Sheehan TJ (2007). Spatial analysis of colorectal cancer incidence and proportion of late-stage in Massachusetts residents: 1995-1998. *Int J Health Geogr*, **6,** 20.

Gelman A, Carlin JB, Stern HS, et al (2003). Bayesian Data Analysis, (Chapman & Hall/CRC Texts in Statistical Science).

Kulldorff M. 2014. SaTScan: software for the spatial, temporal, and space-time scan statistics, version 9.3 [computer program] [Online].

Kulldorff M, Athas WF, Feurer EJ, et al (1998). Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. *Am J Public Health*, **88**, 1377-80.

Kulldorff M, Heffernan R, Hartman J, et al (2005). A space-time permutation scan statistic for disease outbreak detection. *PLoS medicine*, **2**, 59.

Luginaah IN, Gorey KM, Oiamo TH, et al (2012). A geographical analysis of breast cancer clustering in southern Ontario: generating hypotheses on environmental influences. *Int J Environ Health Res*, **22**, 232-48.

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.

Mosavi-Jarrahi A, Mohagheghi M, Zeraatti H, et al (2001). Cancer registration in Iran. *Asian Pac J Cancer Prev*, **2**, 25-9.

Naus JI (1965a). The distribution of the size of the maximum cluster of points on a line. *J Am Stat Assoc*, **60,** 532-8.

Naus JI (1965b). Clustering of random points in two dimensions. *Biometrika*, 263-7.

Rezaeian M (2012). Geographical Epidemiology, Arak, Nevisandeh (in Persian).

Salamatnews. 2012.

Snow J (1855). On the mode of communication of cholera. John Churchill.

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.

Turnbull BW, Iwano EJ, Burnett WS, et al (1990). Monitoring for clusters of disease: application to leukemia incidence in upstate New York. *Am J Epidemiol*, **132**, 136-43.