Application of Bayesian Multilevel Space-Time Models to Study Relative Risk of Esophageal Cancer in Iran 2005-2007 at a County Level

Sedigheh Rastaghi1, Tohid Jafari-Koshki2,3, Behzad Mahaki1*

Abstract

Background: Reported age standardized incidence rates for esophageal cancer in Iran are 0.88 and 6.15 for females and males, at fifth and the eighth ranks, respectively, of cancers overall. The present study aimed to map relative risk using more realistic and less problematic methods than common estimators. Materials and Methods: In this ecological investigation, the studied population consisted of all esophageal cancer patients in Iran from 2005 to 2007. The Bayesian multilevel space-time model with three levels of county, province, and time was used to measure the relative risk of esophageal cancer. Analyses were conducted using R package INLA. Results: The total number of registered patients was 7,160. According to the results, the three-level model with adjustment for risk factors of physical activity and smoking had the best fit among all models. The overall temporal trend was significantly increasing. At county level, Ahar, Marand, Salmas, Bojnoord, Saghez, Sarakhs, Shahroud and Torbatejam had the highest relative risks. Physical activity was found to have significant direct association with risk of developing esophageal cancer. Conclusions: Given to great variation across geographical areas, many different factors affect the incidence of esophageal cancer. Conducting further studies at the individual level in areas with high incidence could provide more detailed information on risk factors of esophageal cancer.

Keywords: Disease mapping - esophageal cancer - Bayesian multilevel space-time model - Iran

Introduction

The causes of human death have been shifted from infectious diseases to non-communicable ones; In developed countries, the incidence of cancers is outpacing cardiovascular diseases (Mohammadpour Tahamatan et al., 2013). Cancer is one of the main causes of death and disabling factors all around the world. According to the World Health Organization, up to 2020, the incidence of cancer will be increased 50 percent. In fact, cancer is expected to be the first cause of death in 2030 (Goya, 2007; Zarnegar Nia et al., 2011). Nowadays, cancer is the second cause of death in the world (Scholefield, 2000) and Over 70 percent of deaths occur due to cancer in developing and developed countries (Mathers and Boschi-Pinto, 2000). It is the third cause of death in Iran after cardiovascular diseases and accidents. Hence, it makes necessary to search the treatments and preventive methods as the priorities for health officials and medical community (Naghavi, 2006).

Gastrointestinal tract cancers are prevalent, especially in developing countries (Parkin, 1998). According to the report of International Agency For Research on Cancer, esophageal cancer is the eighth prevalent cancer and sixth cause of death induced by cancer around the world (Mathers and Boschi-Pinto, 2000; Parkin et al., 2005; Semnani et al., 2005; Kamangar et al., 2006; Ma and Yu, 2006; Ferlay et al., 2010; Scarpa et al., 2011). Esophageal cancer is rare in western countries as around 80 percent of all total esophageal cancer cases are from developing countries (Parkin et al., 2005). The prevalence of esophageal cancer is about 5-10 per 100,000 population in the north of America and western Europe (Kamangar et al., 2006; Kamangar et al., 2007). Esophageal cancer is a relatively common kind of cancer in the eastern countries (Delpisheh et al., 2014). Half million deaths is attributable to esophageal cancer and its five year survival rate is less than 10 percent (Ma and Yu, 2006; Fauci et al., 2008; Rasouli et al., 2011). Previous epidemiological studies in Asia have shown considerable changes in the incidence of esophageal cancer. Iran is located in areas of Asia with high rates of incidence (Long et al., 2010; Moore et al., 2010; Moore et al., 2010a; Moore et al., 2010b; Salim et al., 2010). This cancer is more common in some geographical areas, known as the belt of esophageal cancer, that covers areas from north of China, west and center of Asia to the north of Iran (Kamangar et al., 2007). This suggests the existence of some risk factors with a heterogeneous
geographical distribution that are related to the incidence pattern (Corley and Buffler, 2001; Stein et al., 2005).

Incidence of esophageal cancer in Iran is higher than the global rate in both sexes (Moore et al., 2010b; Radmard, 2010), with age-standardized incidence rate of 0.88 and 6.15 for females and males having the fifth and eighth ranks among all cancer sites respectively (E’temed and Gooya, 2011).

Obtaining a geographical pattern for disease has been used since the early decades of the 20th century both for communicable and non-communicable diseases. Its use, nowadays, has been developed so that health designers apply it as a way to do more qualified interventions and help disease prevention. On the other hand, disease registration based on geographical regions, makes it possible to test variability patterns. It is also useful to recognize low and high risk regions (Lawson et al., 2001).

A number of studies have been conducted about the mapping of GI tract cancers in Iran (Mahaki et al., 2011; Asmarian et al., 2012; Asmarian et al., 2013a; Asmarian et al., 2013b; Mahaki et al., 2013), but few investigations have been done about the mapping of esophageal cancer in recent years (Poor Ahmad and Yavar, 2002; Asmarian et al., 2013c).

Incidence studies are descriptive and have been conducted for specific individual years or province. None of them were at the province and county levels and/or for a wider range of time.

Due to high incidence of esophageal cancer and necessity of providing knowledge about its geographical distribution at province and county levels, present study aims to use Bayesian multilevel space-time models to obtain more accurate estimates and maps for relative risk of esophageal cancer.

The merits of this model include the power of tolerance for missing data without need of estimation and loss of data and the mean studied population, the possibility of considering the role of neighboring areas in different levels, the possibility of determining the relative effectiveness of every level hierarchically over the response variable and, finally, the feasibility of determining effective factor over the response variable for each levels. These models provide more precise estimates and more clear responses in comparison with single level analyses (Hosseini and Gohari, 2014).

Materials and Methods

This is an ecological study conducted in Iran in county level. The data were obtained from reports by Non-communicable Diseases Center of Health and Medical Education Ministry’s Diseases Management Center (http://www.ircancer.ir). On esophageal cancer in 320 counties (ICD-10) from 2005 to 2007. The census of the year 2006 and also estimated population of Iranian Statistics Center were used for the population which is potentially at risk (Iran, 2012).

The two geographically nested levels of counties and provinces were nested in time episodes. The expected number of esophageal cancer cases was calculated on the basis of population size and the observed number of cancer cases during this period. In order to prevent assessing geographical distribution from being distorted by other factors, we included risk factors, including physical activity, smoking, fruit and vegetable consumption and weight at province level and obtained adjusted estimates. Smoking was considered as multiplication of the population percent of smokers in the mean number of used cigarettes per day in every province. Also, obesity was regarded as the percent of the population having BMI>25 in every province. Fruit and vegetable consumption were measured as the total of mean fruit and vegetable number used daily in every province. Physical activity was calculated using a combined index called Metabolic Equivalent (MET).

For rare diseases such as cancers, it is usually assumed that the number of patients follows a Poisson distribution in every region. For modeling the data, multilevel model was used where parameters, in every level, can be random and the function of the characteristics of that level or other levels. One variation of the multilevel model is multiple-membership model as follows.

$$
(y_i | x_i) \sim \text{Poisson}(\lambda_i)
$$

$$
\log(\lambda_i) = \log(\epsilon) + x_i \beta + \omega_h(i) + \sum_{j \in \text{neighbor}(i)} \omega_n(j)
$$

$$
u_{i}^{(h)} \sim N(0, \sigma_{h}^{2}) \quad u_{i}^{(n)} \sim N(0, \sigma_{n}^{2})
$$

Where [i] is the county indicator and random effects are defined by u and superscript numbers show corresponding level. Since number 1 is supposed for observation level, the numbering process is started from 2. There are regional and neighborhood effects indicated by number 2 and 3, respectively. Here β and x represent fixed effects and design matrix.

Regions [i] refer to counties and neighbor [j] denote the set of neighboring areas of each county. Therefore, in this model, the specification of the observed count is affected by different predictor variables. Typically the weights in this model are such that \(\sum_{j \in \text{neighbor}(i)} \omega_n(j) = 1 \) \forall i and, generally, all neighbors are given equal weights so that, in fact, \( w_n(i) = 1/|N| \) where N is the number of neighbors to Region [i]. Bayesian estimation is used for fitting the model (Lawson et al., 2003).

The Full Bayesian method is one of the common ways in diseases mapping. Structural and non-structural heterogeneities are regarded by this method. We calculate relative risk of esophageal cancer by use of full Bayesian model suggested by Besag et al. known as BYM model from 2005 to 2007. This spatial model has no parameter to capture the effect of time (Lawson et al., 2003).

We also used Deviance Information Criterion (DIC) and Logarithmic Score (LS) to compare and find the best model. The model with the least LS and DIC is supposed to have the highest quality of prediction among all models. DIC is applied to assess the goodness of fit of Bayesian models and is calculated as \( DIC = D + pD = D + 2pD \) where, \( D, D \) and \( pD \) represent the mean deviance posterior distribution, estimation of pointed variation and the number of effective parameters of the model, respectively. \( pD \) is found by \( pD = D - D \). As a result, a model with a best trade-off between DIC, LS, and \( pD \) would be considered as the best fit for the data (Spiegelhalter et al., 2002; Celeux et al., 2006).
Bayesian Multilevel Space-Time Models to Study Relative Risk of Esophageal Cancer in Iran 2005-2007 at County Level

In this study, analysis was performed using INLA Bayesian approach in R software. INLA approach is used as an effective calculating method. This method is not only faster than McMC methods, but also does not have the challenges related to convergence and other assumptions of McMC (Schrödle and Held, 2011; Blangiardo et al., 2013).

Results

There were the total numbers of 7160 esophageal cancer cases in Iran in 2005-2007. The highest number of cases was in the Mashhad with 684 cases in a three-year period. Table 1 shows the DIC, pD, and LS for different models. The description for each model is stated in the table.

Model 6 has the least amount of DIC, LS and fits the data more appropriately. The results of these models are shown in Table 2. In the model 6, the effect of year was significant and increasing. Physical activity was also significant. The same results were happened in model 7. But the other factors were not significant.

Table 1. Results Indicators of DIC, PD, , LS Models are Fitted To Compare

| Model | DIC  | PD  | pD  | LS  |
|-------|------|-----|-----|-----|
| 1     | 4345 | 310 | 4035| 2.213|
| 2     | 4154 | 216 | 3938| 2.26 |
| 3     | 4347 | 310 | 4036| 2.316|
| 4     | 4114 | 216 | 3895| 2.316|
| 5     | 4157 | 216 | 3941| 2.263|
| 6     | 4111 | 218 | 3894| 2.236|

Table 2. Estimated Coefficients of Risk Factors for Models Fitted

| Model 4 | By adjusting risk factors including smoking and physical activity | Model 6 | By adjusting all studied risk factor |
|---------|-----------------------------------------------------------------|---------|-----------------------------------|
| Parameter estimation | 95% CrI | Parameter estimation | 95% CrI |
| Intercept | 0.4071 | (0.1352, 0.6800) | 0.5333 | (-0.0705, 1.1386) |
| Smoking | -0.0012 | (-0.0025, 0.0001) | -0.0013 | (-0.0027, 0.0001) |
| Weight | -0.0029 | (-0.0147, 0.0088) | -0.0029 | (-0.0147, 0.0088) |
| Physical activity | -0.0243 | (-0.0315, -0.0171) | -0.0244 | (-0.0316, -0.0172) |
| Fruit and vegetable | 0.0056 | (-0.0797, 0.0907) | 0.0056 | (-0.0797, 0.0907) |
| Year | 0.0836 | (0.0462, 0.1210) | 0.0859 | (0.0396, 0.1323) |

Figure 1. Relative Risk of Incidence Esophageal Cancer in Counties. a) multilevel model by adjusting risk factors including smoking and physical activity in 2005; b) multilevel model by adjusting risk factors including smoking and physical activity in 2006; c) multilevel model by adjusting risk factors including smoking and physical activity in 2007; f) multilevel model by adjusting all studied risk factor in 2005; g) multilevel model by adjusting all studied risk factor in 2006; d) multilevel model by adjusting all studied risk factor in 2007.
Table 3. ICC values

|        | Variance | ICC  |
|--------|----------|------|
| province | 0.366    | 57%  |
| county  | 0.267    | 42%  |

Discussion

According to the results, north-east and northwest regions and some parts of central zones (especially Khorasan, North Khorasan, Golestan, Ardabil, West Azerbaijan, East Azerbaijan, Kurdistan and some regions of Semnan provinces) have higher levels of esophageal cancer incidence (relative risk more than 3.5) in comparison with desert and southern regions (especially Sistan and Baluchestan, Bushehrand Hormozgan provinces). It is necessary to investigate the causes of cancer in high-risk areas and the relationship between these factors and the results of this study.

The present investigation is the first mapping of esophageal cancer incidence rate in Iran at county level. Esophageal cancer incidence and related mortality in Iran is one of the important health problems and climate conditions facilitates its incidence and prevalence of esophageal cancer. Hence in the present study, environmental investigation and its mapping were done and relative risks were estimated in the counties during three years. Esfandi (Isfandiyar et al., 2009a; Isfandiyar et al., 2009b), Aazami et al. (2006) and Kamangar et al. (2007) illustrated that Ardebil and Golestan provinces are high risk regions of esophageal cancer.

In their ecological study, Semnani et al. (2010) found that the highest incidence of esophageal cancer in Golestan province was related to the exceeding rate of selenium in the soil of the region. Rahimzadeh Bozorgi et al. (2013) declared that the presence of selenium in the soil and rice of Golestan was a possibly important factor in the high incidence of esophageal cancer. Sadeghieh Ahari et al. indicated the importance of the Sabalan volcano as an important environmental factor in the incidence of esophageal cancer in Ardebil province (Ahari et al., 2013).

As in the work by Aazami et al. (2006) we found no statistically significant association between smoking and vegetable intake and esophageal cancer incidence. In the investigation of Hajizadeh et al. (2012) there was no correlation between vegetable intake and esophageal cancer as well. In contrast to the results of previous studies (Lagergren et al., 1999; Lagergren et al., 2014; Thrift et al., 2014), weight was not a significant factor in esophageal cancer incidence in our study. Physical activity has significant effect, which is in accordance with the results of previous studies (Chen et al., 2014; Dos Santos and Coutinho, 2014; Singh et al., 2014). This risk factor may be accompanied by decreasing incidence risk of esophageal cancer.

Many articles have been published in Iran about esophageal cancer in which this kind of cancer has been studied in a single or a few provinces especially in Northern provinces. They are, mostly, descriptive and cross-sectional (Ghavamzadeh et al., 2001; Radmard, 2010; Sadjadi et al., 2010; Rasouli et al., 2011; Rahimzada Bozorgi et al., 2013; Somi et al., 2014). We found no study in which environmental factors have been taken into consideration the esophageal cancer incidence nationwide. The specialists and health planners should present their beliefs for designing the effective programs, in order to study precisely influencing factors in the incidence of esophageal cancer and the sort of achieved pattern in the variation of this cancer, especially the approach of facing high risk areas.

It should be noted that just pathological reports are

Figure 2. The Relative Risk of Incidence Esophageal Cancer in the Counties with the Bym Model without Adjustment for Risk Factors
taken into account in an Iranian registration system of cancer. So, at best, only 80 percent of cancer cases can be registered. In addition, this registration system was disqualified in some counties and years so that it apparently leads to underestimate the number of cancer cases (Asmarian et al., 2013a). Mapping of a disease at county level will be smoother than province level due to more areas in county level mapping.

Because access to up-to-date data about esophageal cancer is not available, other types of risk factors such as tobacco consumption, alcohol use, warm drinks like tea, socioeconomic condition, unhealthy nutrition, Selenium, shortage, viral and genetic agents were not registered, and also the data related to effective risk factors in county level were not registered and there are missed data in county level. Using more precise and complete data will help further clarifying the process of esophageal cancer in Iranian counties. It is also recommended to increase the period of study or to investigate the impact of each factor in every region by use of varying slope models. To find more clearly risk factors for esophageal cancer incidence, it is suggested conducting further studies in the areas with high incidence.

References
Asmarian NS, Kavousi A, Salehi M, et al (2012). Mapping of stomach cancer rate in Iran using area-to-area poisson kriging. J Health Sys Res, 9, 681-7.
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 Med Sci, 20, 10-7.
Asmarian NS, Kavousi A, Salehi M, et al (2013b). Comparison of point poisson kriging and empirical bayesian methods in mapping of gastrointestinal cancer incidence rate in Iran. J Health Sys Res, 9, 277-85.
Asmarian NS, Ruzitalab A, Amir K, et al (2013c). 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.
Azami A, Yazdanbod A, Abbasgolizadeh N, et al (2006). Esophageal cancer risk factors in relation with different anatomical sites of esophagus in aras Gi clinical, Ardabil. J Ardebil Uni Med Sci, 6, 126-31.
Blangiardo M, Cameletti M, Baio G, et al (2013). Spatial and spatio-temporal models with R-INLA. Spat Spatiotemporal Epidemiol, 7, 39-55.
Celeux G, Forbes F, Robert CP, et al (2006). Deviance information criteria for missing data models. Bayesian Anal, 1, 651-73.
Chen Y, Yu C, Li Y (2014). Physical activity and risks of esophageal and gastric cancers: a meta-analysis. PLoS One, 9, 1-6.
Corley DA, Bulpifer PA (2001). Oesophageal and gastric cardia adenocarcinomas: analysis of regional variation using the cancer incidence in five continents database. Int J Epidemiol, 30, 1415-25.
Delpishah A, Veisani Y, Sayehmiri K, et al (2014). Esophageal carcinoma: long-term survival in consecutive series of patients through a retrospective cohort study. Gastroenterol Hepatol Bed Bench, 7, 101-7.
Dos Santos WM, Coutinho RX (2014). Dietary and physical activity habits associated with risk for the development of cancer: prevalence in adolescents. Sci Med, 24, 254-8.
E’temad K, Gonya MM (2011). Iran cancer registry report 2009, Iran ministry of health and medical education.
Fauci AS, Kasper DL, Braunwald E, et al (2008). Harrison’s Principles of Internal Medicine, McGraw-Hill Medical New York.
Ferlay J, Shin HR, Bray F, et al (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer, 127, 2893-917.
Ghavamzadeh A, Moussavi A, Jahani M, et al (2001). Esophageal cancer in Iran. Semin Oncol, 28, 153-7.
Goya M (2007). Iranian annual cancer registration report 2005/2006. ministry of health and medical education, health deputy, center for disease control and prevention [In Persian].
Hajizadeh B, Sedaqhat F, Rashidi Khani B, et al (2012). Association between types of food consumed and the risk of esophageal squamous cell carcinoma in Kurdistan Province. J School Public Health Inst Public Health Res, 10, 1-10.
Hosseini AF, Gohari MR (2014). Application of multilevel model in determining the effective factors in the length of stay among appendectomy patients. Razi J Med Sci, 20, 70-7.
Islami F, Kamangar F, Nasrollahzadeh D, et al (2009a). Oesophageal cancer in Golestan Province, a high-incidence area in northern Iran—a review. Eur J Cancer, 45, 3156-65.
Islami F, Pourshams A, Nasrollahzadeh D, et al (2009b). Tea drinking habits and oesophageal cancer in a high risk area in northern Iran: population based case-control study. BMJ, 338, 1-8.
Kamangar F, Dores GM, Anderson WF (2006). Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. J Clin Oncol, 24, 2137-50.
Kamangar F, Malekzadeh R, Dawsey SM, et al (2007). Esophageal cancer in Northeastern Iran: a review. Arch Iran Med, 10, 70-82.
Lagergren J, Bergstrom R, Nyren O (1999). Association between body mass and adenocarcinoma of the esophagus and gastric cardia. Ann Intern Med, 130, 883-90.
Lagergren J, Mattsson F, Nyren O (2014). Gastroesophageal reflux does not alter effects of body mass index on risk of esophageal adenocarcinoma. Clin Gastroenterol Hepatol, 12, 45-51.
Lawson AB, Browne WJ, Vidal Rodeiro CL (2003). Disease mapping with WinBUGS and MLwiN, John Wiley & Sons.
Long N, Moore M, Chen W, et al (2010). Cancer epidemiology and control in north-East Asia-past, present and future. Asian Pac J Cancer Prev, 11, 107-48.
Ma X, Yu H (2006). Cancer issue: global burden of cancer. Yale J Biol Med, 79, 85-94.
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.
Mahaki B, Mehrabi Y, Kavousi A, et al (2013). Applying gamma-poisson, lognormal, and bym models in comparing relative risk of suicide among provinces of ilam, Iran. J Health Sys Res, 9, 86-95.
Mathers C, Boschi-Pinto C (2000). Global burden of cancer in the year 2000: Version 1 estimates. Global Program on Evidence for Health Policy (GPE), Global burden of disease2000. 186-95.
Mohammadpour Tahamtan RA, Yazdani J, Zare S, et al (2013). Geographical mapping of gastrointestinal cancer mortality rates in Mazandaran and Golestan provinces, Iran. J Mazand Univ Med Sci, 23, 125-35.
Moore MA, Ariyaratne Y, Badar F, et al (2010a). Cancer...
epidemiology in South Asia-past, present and future. *Asian Pac J Cancer Prev*, 11, 49-66.

Moore MA, Eser S, Igisinov N, et al (2010b). Cancer epidemiology and control in North-Western and Central Asia-past, present and future. *Asian Pac J Cancer Prev*, 11, 17-32.

Naghavi M (2006). Transition in health status in the islamic republic of Iran. *Iranian J Epidemiol*, 11, 49-66.

Parkin D (1998). Epidemiology of cancer: global patterns and trends. *Toxicol lett*, 102, 227-34.

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

Poor Ahmad A, Yavar B (2002). Medical geography of esophageal cancer in Iran. *J Entrepreneurship Res*, 41, 13-24.

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

Rahimzada Bozorgi H, Beyrami S, Roshandel G, et al (2013). High concentrations of selenium in rice cultivation in high-risk areas for esophageal cancer. Sixteenth National Conference on Environmental Health-Persian date Mehr 1392. *Tabriz Uni Med Sci* [Epub ahead of print].

Rasouli M, Ghadimi MR, Mahmoodi M, et al (2011). Survival analysis of patients with esophageal cancer using parametric cure model. *Asian Pac J Cancer Prev*, 12, 2359-63.

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

Salim EI, Moore MA, Bener A, et al (2010). Cancer epidemiology in South-West Asia-past, present and future. *Asian Pac J Cancer Prev*, 11, 33-48.

Scarpa M, Valente S, Alfieri R, et al (2011). Systematic review of health-related quality of life after esophagectomy for esophageal cancer. *W J Gastroenterol*, 17, 4660.

Schleefeld JH (2000). ABC of colorectal cancer: screening. *Br Med J*, 321, 1004-6.

Schrdle B, Held L (2011). A primer on disease mapping and ecological regression using INLA. *Comput stat*, 26, 241-58.

Semnani S, Besharat S, Abdolahi N, et al (2005). Factors associated with esophageal cancer in the southeast part of the Caspian Sea. *J Med Facul Guilan Univ Med Sci*, 52, 28-4.

Semnani S, Roshandel G, Keshkar A, et al (2010). Relationship between soil selenium level and esophageal cancer: an ecological study in Golestan province of Iran. *J Gorgan Uni Med Sci*, 12, 51-6.

Singh S, Devanna S, Varayil JE, et al (2014). Physical activity is associated with reduced risk of esophageal cancer, particularly esophageal adenocarcinoma: a systematic review and meta-analysis. *BMC Gastroenterol*, 14, 1-11.

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 DJ, Best NG, Carlin BP, et al (2002). Bayesian measures of model complexity and fit. *J R Stat Soc Series B Stat Methodol*, 64, 583-639.

Stein HJ, von Rahden BH, Siewert JR (2005). Survival after oesophagectomy for cancer of the oesophagus. *Langenbeck’s Arch Surg*, 390, 280-5.

Thrift AP, Shaheen NJ, Gammon MD, et al (2014). Obesity and risk of esophageal adenocarcinoma and Barrett’s esophagus: a Mendelian randomization study. *J Natl Cancer Inst*, 106, 252.

Zarnegar Nia Y, Alavi Majd H, Azodi M, et al (2011). Using partitioning and non-partitioning clustering algorithms for included proteins sequences in esophagus, stomach and colon cancer. *J Paramed Sci*, 2, 9-16.