RESEARCH ARTICLE

Disease Mapping for Stomach Cancer in Libya Based on Besag-York-Mollié (BYM) Model

Maryam Ahmed Salem Alhdiri^{1*}, Nor Azah Samat², Zulkifley Mohamed²

Abstract

Globally, Cancer is the ever-increasing health problem and most common cause of medical deaths. In Libya, it is an important health concern, especially in the setting of an aging population and limited healthcare facilities. Therefore, the goal of this research is to map of the county' cancer incidence rate using the Bayesian method and identify the high-risk regions (for the first time in a decade). In the field of disease mapping, very little has been done to address the issue of analyzing sparse cancer diseases in Libya. Standardized Morbidity Ratio or SMR is known as a traditional approach to measure the relative risk of the disease, which is the ratio of observed and expected number of accounts in a region that has the greatest uncertainty if the disease is rare or small geographical region. Therefore, to solve some of SMR's problems, we used statistical smoothing or Bayesian models to estimate the relative risk for stomach cancer incidence in Libya in 2007 based on the BYM model. This research begins with a short offer of the SMR and Bayesian model with BYM model, which we applied to stomach cancer incidence in Libya. We compared all of the results using maps and tables. We found that BYM model is potentially beneficial, because it gives better relative risk estimates compared to SMR method. As well as, it has can overcome the classical method problem when there is no observed stomach cancer in a region.

Keywords: BYM model- standardized morbidity ratio- disease mapping- relative risk- stomach cancer- Libya

Asian Pac J Cancer Prev, 18 (6), 1479-1484

Introduction

Cancer is an important public health problem worldwide. Cancer disease impacts everyone regardless of age, gender and class. There are several risk factors that cause cancer which are alcohol use, physical inactivity, tobacco use and unhealthy diet are the main cancer risk factors, as well as some chronic infections, which have main relevance in low- and middle-income countries, including, Libya. In Libya, cancer is the third leading cause of death after car accident and cardiovascular disease, therefore cancer is an important problem in public health in Libya, particularly in the setting of an aging population and limited healthcare facilities. For instance, stomach cancer has the highest prevalence of disease digestive cancer in Libya especially West and North districts of Libya (Alhdiri et al., 2016; Bodalal and Bendardaf, 2014; Bodalal et al., 2014).

Medically, stomach cancer known as gastric cancer and it is a cancer that starts in the stomach. As with other cancers, the cause of stomach cancer is not known with certainty, but always associated with peptic ulcers, inflammation of the stomach is accompanied by shrinking the stomach.

A "stomach cancer risk factor" is any type of factor that increases the risk of developing stomach cancer. Many of the most important risk factors for stomach cancer are beyond the control, such as age, family history, and medical history. Although it is uncertain what the cause of gastric cancer, however, factors that may increase the risk of stomach cancer are:

- Consumption of salted and smoked foods
- Rarely consume fruits and vegetables
- A family medical history of gastric cancer in which
- Infections caused by Helicobacter pylori, a bacterium living in the mucous lining of the stomach
- · Chronic gastric inflammation, which refers to a long-term stomach inflammation.
- Pernicious anaemia, which is a decrease in the number of red blood cells that occurs when the digestive tract cannot absorb vitamin B12 properly.
 - Smoking

In recent years, a number of studies have been conducted to map the geographical spread of stomach cancer incidence, such as a study done by Mohebbi and his colleagues (2008; 2011) using adjusted age-specific standardized incidence ratio (SIR) in the southwest of the

Department of Statistics, Faculty of Science, University of Tripoli, Alfernag, Tripoli, Libya, Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, Malaysia. *For Correspondence: M.alhdiri@uot.edu.ly, m.alhdiri@yahoo.com

Caspian Sea from 2001 to 2005. In addition, the previous studies have considered counties of Northern provinces of Iran as clustered and has high-risk incidence rate cluster of stomach cancer (Mohebbi et al., 2008; 2011). In Libya, studies on geographical analysis for cancer in the country at all is poor. This is because only the Eastern part of the country were considered. According to Abdel-Naser et al., (2012) this is due to stomach cancer with Helicobacter pylori infection in prominent in Eastern Libya during the period from 2000 to 2002.

El-Mistiri et al., (2006) and Elzouki and Alkhomsi (2005) showed that stomach cancer is the second most prevalent disease in Benghazi (the largest district in eastern Libya) after colorectal cancer. The age-standardized rates of stomach cancer in the Benghazi cancer registry were 11.6 new cases per 100,000 men-years and 8.8 per 100,000 women-years, respectively (Elzouki and Alkhomsi, 2005). No such published works or studies were constructed to study or determine the pattern geographical distributions of stomach cancer in Libya at all.

Therefore, the goal of this research is to characterize some geographical features that associated with stomach cancer in Libya. To the best of our insight, such results on Libyan database have not been distributed or published already (for the first time in a decade). In other words, the main aim of this study is to discuss and compare the relative risk estimation for stomach cancer disease mapping based on two different approaches. These involve the analysis for relative risk estimation based on the SMR method BYM and, the application of these methods to observed stomach cancer data from Libya to describe geographic of disease risk and identifying unusual high and low risk areas.

The areas or regions of Libya vary in size, shape and population size (Figure 1). In this study, we discuss and demonstrate the most common methods used in the study of disease mapping, which is the classical method that known as Standardized Morbidity Ratio and BYM model. We will focus on its application to stomach cancer data in Libya. This study is organized as follows. First, in Section 2, we review and describe the classical method in estimating relative risk using SMR method. This includes the definition of SMR and its drawbacks. Then we move to describe and overview of the common application of Bayesian methodology, that called as "BYM model" which is applied and used by Lawson et al., (2003), will be in next section. Section 3 describes the cancer data used in our application and presented findings. As well as several results presented in this section based on these two methods, which are applied to observed stomach cancer data in Libya in order to demonstrate and identify a better method of estimating stomach risk. Finally, we close with some discussion in Section 4.

Materials and Methods

Standardized Mortality Ratio (SMR) Method

In disease mapping, SMR is the common statistics used in spatial studies. The main aim of the SMR is to estimate the RR of a certain disease in a certain map, which may be interpreted as the probability that a person

within a specified region contracts the disease divided by the probability that a person in the population contracts the disease. In the field of epidemiology studies, SMR represents either standardized mortality ratio or the standardized morbidity ratio, when mortality refers to death while morbidity refers to the incidence.

However, in general notation, suppose that O_i , where i=1, 2, ..., H (H indexes the areas or regions), indicates the observed cases of a certain disease of the study, and let E_i represents the expected cases or expected number of cases. Using these values as obtained from the available data, we can calculate the relative risk θ_i for area i, which is the SMR defined as

$$S\widehat{M}R = \widehat{\theta}_i = \frac{O_i}{E_i} \tag{1}$$

For the expected value E_i , it could be calculated by using a particular formula as below:

$$E_{i} = N_{i} \frac{\sum\limits_{i=1}^{H} O_{i}}{\sum\limits_{i=1}^{H} N_{i}}, \quad then \quad \hat{\theta}_{i} = \frac{O_{i}}{\sum\limits_{i=1}^{H} O_{i}} = \frac{\left(\frac{O_{i}}{N_{i}}\right)}{\left(\sum\limits_{i=1}^{H} N_{i}}\right)}$$

Where N_i the is the population of district/area i. Here standardization is completed by the total population at risk, assuming everybody is equal at risk. Consequently, we estimate the relative risk using formula:

$$S\widehat{M}R = \widehat{\theta}_i = \frac{\left(O_i/N_i\right)}{\left(\sum_{i=1}^H O_i / \sum_{i=1}^H N_i\right)}$$

which is defined to be the probability that a person within the district/area contracts the disease divided by the probability that a person in the population contracts the disease.

Samat and Percy (2008) used the Equation (1) for the SMR in their study and applied it to dengue disease mapping in Malaysia. According to Lawson et al., (2003), although SMR is used commonly as measure to estimate the true relative risk, but at the same time, it has some problems associated with the use of it. SMR is based on a ratio estimator, the mean and variance (SMR_i, SMR_i/E_i) of SMR are very highly dependent on expected count Ei. Furthermore, if there are areas with no observed count data, mathematically the SMR is necessarily zero. Meza (2003), showed that this problem of SMR makes the interpretation of SMR difficult and it should be done with caution and also points out other problems of using SMR, which is that the SMR is a reliable measure of relative risk for large geographical regions such as countries or states, but is unreliable for small areas such as counties.

However, to be able to overcome these problems of using of SMR method, many researchers have produce other alternative methods to estimate the relative risk of the disease. One of these methods was the use of Bayesian methods. In this research, we suggested very common method to estimate the relative risk of a disease, as will be discussed in next section.

Besag-York-Mollié (BYM) Model

To address the problem of SMRs, in this research the BYM model will be used to analysis the data, as well as to consider the information of the adjacent neighbours of each district (area). The main idea for this model is to produce a more reliable estimation for relative risks and for small areas or rare disease. This is by borrowing required information from the neighbouring areas. In this model, the relative risk is modelled with additional consideration. Area-specific random effects which are divided into two components. The first component is u_1 that takes into account the effects that vary in a structured manner in space (clustering or correlated heterogeneity). The second component is v_i that takes into account the effects that vary in an unstructured way between areas (uncorrelated heterogeneity). Therefore, the model introduced by Clayton and Kaldor (1987) and developed by Besag et al., (1991), is formulated as follows:

$$O_i \sim Poisson(E_i \theta_i), \quad Log(\theta_i) = \alpha + u_i + v_i$$
 (2)

where $\theta_i = \exp(\alpha + u_i + v_i)$, α is an intercept (an overall level of the relative risk) O_i and E_i and θ_i be the observed count, expected count and relative risk parameter in the *ith* area respectively, u_i is the correlated heterogeneity and v_i is the uncorrelated heterogeneity. The uncorrelated heterogeneities are assumed to follow a normal distribution, as follows:

$$v_i \sim N(0, \tau_v^2)$$

For the first component, which is the clustering component, a spatial correlation structure is used, where estimation of the risk in any area depends on neighbouring areas. The conditional autoregressive (CAR) model proposed by Besag et al. (1991) will be used to model the distribution of the correlated heterogeneity as:

$$[u_i / u_k, i \neq k, \tau_u^2] \sim N(\overline{u}_i, \tau_i^2)$$
 where

$$\overline{u}_{j} = \frac{\sum_{k} u_{j} \, \omega_{k}}{\sum_{k} \omega_{k}}, \quad \tau_{j}^{2} = \frac{\tau_{u}^{2}}{\sum_{k} \omega_{k}}; \quad \omega_{k} = \begin{cases} 1 & i, k, are \ adjectent \\ 0 & f \end{cases}, k, are \ not \ adjectent$$

For the parameters, in this case are τ_u and τ_v , must be specified. These parameters control the variability of u and v. From this source, u and v are considered to have Gamma distribution.

Results

Relative Risk Of Disease Using SMR Method And BYM Model To Stomach Cancer Mapping

This section explains and displays the outcome of the applications of existing relative risk estimation methods, corresponding to the classical model based on the standardized morbidity ratio and one of earliest examples of Bayesian mapping methods based on the BYM model using observed stomach cancer in Libya. Models were fitted to data using WinBUGS software. Then, all of these outcomes are compared and displayed in the table and maps, to reveal the best-fitted model for relative risk

Cancer Data

In this applied, ecological research, information of the Libyan districts for one year 2007 was analyzed, which is obtained from Africa Oncology Institute (AOI) (Sabratha Cancer Registry, 2008; Ministry of Health, 2012). These administrative districts are Alnikat, Zawia, Aljafara, Tripoli, Almergaib, Musrata, Sirt, Benghazi, Almarg, Aljabal Alakhader, Darna, Albatnan, Nalut, Aljabal Algarbi, Wadi Shatee, Aljufra, Ejdabiya, Ghat, Wadi Alhiya, Sabha, Morzuk and Alkufra. A geographical system of Libya's districts is explained in Figure 1, where ID, name and population for each district in the map are shown. Stomach cancer disease data is used to illustrate the SMR model and BYM model to estimate the relative risk of disease.

The outcomes for the relative risk estimation using the BYM mode and the SMR for 2007 are displayed in Table 1. Table 1 presents two obvious differences in terms of the value of relative risk. Eleven districts have value of relative risk equal to zero, in the absence of cases of observed stomach cancer based on analysis using the SMR method. These districts are Musrata, Benghazi, Almarg, Aljabal Alakhader, Darna, Albatnan, Ejdabiya, Ghat, Wadi Alhiya, Sabha and Alkufra. Based on SMR method, susceptible people within the district of Nalut have the highest risk of contracting stomach cancer, while susceptible people within the above districts have the lowest risk of stomach cancer when compared with people in the overall population. The corresponding values of relative risk are approximately 4.935 and 0, respectively. Conversely, by using BYM model, the finding shows opposite results where the value of relative risk is not zero although there are districts with no observed counts of stomach cancer cases, which can be a disadvantage of the SMR approach. However, the BYM model does not suffer from this drawback and generates positive estimates of relative risk in districts that have no observed case.

In addition, from Table 1, Alnikat and Wadi Shatee have a high value of relative risk during 2007. Estimation

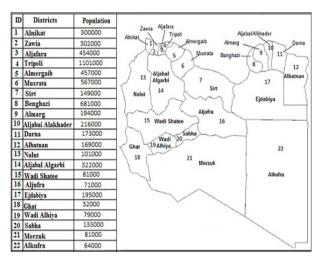


Figure 1. Names of 22 Geographic Boundaries, Code on the Map and Population of All Districts in Libya (Source: Alhdiri et al., 2016)

Table 1. Comparison between the Classical Method and Smoothed Relative Risk Estimates Based on BYM Model and Their Associated Standard Deviations of Stomach Cancer Disease for the Year 2007

No.	District	O_i	E_{i}	Relative Risk based on SMR Method		Relative Risk based on BYM model	
				RR	SD	RR	SD
1	Alnikat	5	1.013	4.935	2.207	1.504	1.089
2	Zawia	3	1.019	2.941	1.698	1.107	0.398
3	Aljafara	2	1.533	1.304	0.922	1.103	0.409
4	Tripoli	3	3.7183	0.806	0.466	0.996	0.302
5	Almergaib	1	1.543	0.648	0.648	0.966	0.319
6	Musrata	0	1.915	0	0	0.902	0.329
7	Sirt	1	0.503	1.987	1.987	1.061	0.503
8	Benghazi	0	2.299	0	0	0.903	0.298
9	Almarg	0	0.655	0	0	0.921	0.338
10	Aljabal Alakhader	0	0.729	0	0	0.9102	0.339
11	Darna	0	0.584	0	0	0.914	0.337
12	Albatnan	0	0.571	0	0	0.918	0.357
13	Nalut	1	0.341	2.932	2.932	1.015	0.429
14	Aljabal Algarbi	1	1.087	0.919	0.919	0.978	0.321
15	Wadi Shatee	1	0.274	3.656	3.656	1.114	0.771
16	Aljufra	1	0.239	4.1704	4.1704	1.07	0.545
17	Ejdabiya	0	0.659	0	0	0.936	0.341
18	Ghat	0	0.108	0	0	0.975	0.394
19	Wadi Alhiya	0	0.267	0	0	0.971	0.413
20	Sabha	0	0.449	0	0	0.968	0.379
21	Morzuk	1	0.274	3.656	3.656	1.057	0.487
22	Alkufra	0	0.216	0	0	0.942	1.089

RR: Relative Risk; SD: standard deviation; values highlighted in bold: highest or lowest relative risk or expected cases.

using the BYM model shows that susceptible people within the district of Alnikat have the highest risk of about 1.504, while susceptible people within the districts of Musrata have the lowest risk of about 0.902.

In addition, the results in Table 1 displayed that a small population (small number of people in the ith district) has low expected counts, but SMR and standard error are high (see 18th district, Ghat). Conversely, district with high population (high number of people in the ith district) has high expected counts, however SMR and standard error are low (see 4th district, Tripoli). Generally, SMRs have the greatest uncertainty because they have small population but standard error are high. So, we can say that the relative risks based on BYM provide a more stable risk estimate such as yielding low standard error than using the classical method.

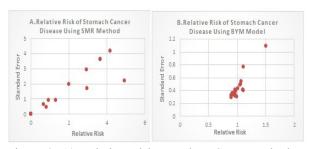


Figure 2. A) Relative Risk Based on SMR Method vs Standard Error, B) Relative Risk Based on BYM Model vs Standard Error

Figure 2 shows the relationship between the relative risk and standard error based on SMR method and BYM model. From this Figure 2A, it can be seen clear that the relative risk increases as standard error increases. While in Figure 2B shows that the relative risk of disease using BYM model produced higher precision of estimate than classical method, because they have smaller standard error.

Disease Maps of Relative Risk Estimates for Stomach Cancer in Libya

Disease maps are used to represent the different levels of risk for stomach cancer, which covered all 22 districts in Libya. Diseases maps graphically display statistical outcomes for relative risk estimation and are an inferential and fundamental decision-making tool. In this

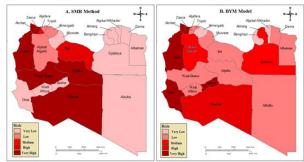


Figure 3. Diseas Maps of Estimated Relative Risk Baed on A) SMR Method and B) BYM Model

Disease Mapping for Stomach Cancer in Libya: Bayesian Study

study, multiple colors are applied in the maps to identify and display among the hot areas which have high or low risk of stomach cancer incidents for all districts in Libva. The results represent the hot points of stomach cancer incidents. Figure 3 shows the risk maps for stomach cancer based on the SMR method and the BYM model.

Mapping issues related to aggregated data are discussed by Lawson (2006), as well as interpretation and representation issues for disease risk in maps in other works (Pickle et al., 1999; Lewandowsky et al., 1993; Mungiole et al., 1999). To explain the results more effectively, multiple colors are used in the maps to highlight areas of high-low risks districts of stomach cancer occurrences. In other words, in our application, for the purposes of results explanation, multiple colors are used in the maps in order to display and identify between the areas with high and low risk of stomach cancer occurrences for each district in Libya.

Many studies point out that there is no definitive way of choosing the levels of risk, therefore it is fairly arbitrary (McGrory and Titterington, 2008). Therefore, each district will be categories into five different levels of hazard. These levels are very high, high, medium, low and very low with their intervals of (<0.5), (0.5,1), (1,1.5), (1.5,2) and $[(2,\infty)]$ respectively. In our application, depending on the concept, and definition of relative risk that is given in several studies, we choose these intervals to cover the range of observed values. In addition, the dark shade in the map indicate the highest risk (very high) and by the lightest shade indicate the very lowest risk (very low).

Figure 3A for the SMR map shows that there are six districts with very high risk of stomach cancer, which are Alnikat, Zawia, Nalut, Wadi Shatee, Aljufra, Ejdabiya and Morzuk. This is followed by only one district which have high risk, which was Sirt. Similarly, only one district has medium risk, which was in Aljafara. The districts of Tripoli, Almergaib and Aljabal Algarbi with low risk and the districts of Musrata, Benghazi, Almarg, Aljabai Alakhader, Darna, Albatnan, Ejdabiya, Ghat, Wadi Alhiya, Sabha and Alkufra have very low risk. The BYM model map in Figure 3B shows that the district of Alnikat has high risk of stomach cancer occurrences, while no districts have very high risk. The districts of medium risk are Zawia, Aljafara, Sirt, Nalut, Wadi Shatee, Aljufra and Morzuk. while the other fourteen districts with low risk, include the districts of Tripoli, Almergaib, Musrata, Benghazi, Almarg, Aljabal Alakhader, Darna, Albatnan, Aljabal Algarbi, Ejdabiya, Ghat, Wadi Alhiya, Sabha and Alkufra. Clearly, there are no districts have very low risk.

Comparisons between the SMR method and BYM model for only one year 2007 show some evident differences in terms of the estimated risks based on both maps considered. Therefore, disease maps are mostly meant to be a better presentation tool for identifying areas, which have very high, or high risk of stomach cancer disease, so that further interest could be provided to these priority districts.

Discussion

Bayesian disease mapping techniques with BYM

model gives smoother relative risk, especially when rare diseases are investigated in an area which has a small population.

The findings based on BYM model offered better estimates of relative risk compared to the SMR method. These finding showed that the BYM model can overcome the drawback of SMR especially when there is no observed stomach cancer case in certain districts. The maps show that the high risks are concentrated in the north-west part of the study area (the country) and least in the South and East. It is identical in terms of population concentration congestion in the north and the least concentration of population in the South and East.

The districts with the highest risks are located in the west probably due to oil installations in this area such as Mellitah Oil and Gas B.V, the Zawia Oil Refining Company and Bouri Oil Field, as well as the electrical power stations (Alsaker, 2013). The high risk of a specific region, the more focus is needed by the government and financial support are required. Susceptible people within districts located in the eastern part of the country have the lowest risk based on both methods when compared to the people in the overall population.

Acknowledgements

The authors would like to thank the staff members of the African Oncology Institute (AOI) Sabratha at the Libyan Cancer Centre and the staff in the Universiti Pendidikan Sultan Idris, Malaysia, for their continued help and support to complete this research. This study has been extracted from PhD thesis in Applied Statistics at Mathematics Department at Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, Malaysia.

References

Alhdiri MA, Samat NA, Mohamed Z (2016). Mapping Libya's prostate cancer based on the SMR method: A geographical analysis. Geografia Malays J Soc Space, 12, 118-25.

Alsaker MH (2013). Natural human factors plays in the incidence of malignant tumors in the north - west of Libya. Ph.D. thesis, University of Tripoli.

Besag J, York J, Mollie A (1991). Bayesian image restoration with two applications in spatial statistics. Ann I Stat Math,

Bodalal Z, Azzuz R, Bendardaf R (2014). Cancers in eastern Libya: First results from Benghazi medical center. World J Gastroenterol, 20, 6293-6301.

Bodalal Z, Bendardaf R (2014). Colorectal carcinoma in a Southern Mediterranean country: The Libyan scenario. World J Gastroenterol, 6, 98-103

Clayton DG, Kaldor J (1987). Empirical Bayes estimates of age-standardised relative risks for use in disease mapping. Biometrics, 43, 671-91.

El-Mistiri M, Verdecchia A, Rashid I (2003). Cancer incidence in eastern Libya: the first report from Benghazi cancer registry, 2003. Int J Cancer, 120, 392-7.

Elzouki A-N, Alkhomsi S (2005). Pattern of gastrointestinal tract cancer in the Eastern part of Libya. Garyounis Med J, 22, 27–31.

Lawson AB (2006). Statistical methods in spatial epidemiology.

- (2nd ed.). England: John Wiley and Sons, pp 1-424.
- Lawson AB, Browne WJ, Rodeiro CLV (2003). Disease mapping with WinBUGS and MlwiN. England: Wiley, pp 1-296.
- Lewandowsky S, Hermann D, Behrens J, et al (1993). Perception of clusters in statistical maps. Appl Cognitive Psych, 7,
- McGrory CA, Titterington DM, Reeves R, Pettitt AN (2009). Variational Bayes for estimating the parameters of a hidden Potts model. Stat Comput, 19, 329.
- Meza JL (2003). Empirical Bayes estimation smoothing of relative risks in disease mapping. J Stat Plan Infer, 112, 43-62.
- Ministry of health- information and documentation center (2012). Annual statistical report 2012. Tripoli. file:///C:/ Users/malhd/Downloads/_report%202012.pdf, WWW. SEHA.LY.
- Mohebbi M, Mahmoodi M, Wolfe R (2008). Geographical spread of gastrointestinal tract cancer incidence in the Caspian Sea region of Iran: Spatial analysis of cancer registry data. BMC Cancer, 14, 8-137.
- Mungiole M, Pickle L, Simonson K (1999). Application of a weighted head-banging algorithm to mortality maps. Stat Med, 18, 3201-09.
- Pickle L M, Mungiole GJ, White A (1999). Exploring spatial patterns of mortality: the new atlas of United States mortality. Stat Med, 18, 3211-20.
- Sabratha cancer registry (2008). First annual report, 2006. Edited by African Oncology Institute, Sabratha, Libya.
- Samat NA, Percy DF (2008). Standardized mortality and morbidity ratios and their application to dengue disease mapping in Malaysia. Proc. of the Salford postgraduate annual research conference (Salford University, UK), ISBN: 9781905732715.
- Mohebbi M, Wolfe R, Jolley D (2011). The spatial distribution of esophageal and gastric cancer in Caspian region of Iran: An ecological analysis of diet and socio-economic influences. Int J Health Geogr, 15, 10-13.
- Abdel-Naser YE, Soad IB, Akram A, Salah H, Annie JS (2012). Gastric cancer and Helicobacter Pylori infection in the Eastern Libya: A descriptive epidemiological study. Arab J Gastroenterol, 13, 85-8.