RESEARCH ARTICLE

Risk Estimation for Lung Cancer in Libya: Analysis Based on Standardized Morbidity Ratio, Poisson-Gamma Model, BYM Model and Mixture Model

Maryam Ahmed Alhdiri[1]*, Nor Azah Samat[2], Zulkifley Mohamed[2]

Abstract

Cancer is the most rapidly spreading disease in the world, especially in developing countries, including Libya. Cancer represents a significant burden on patients, families, and their societies. This disease can be controlled if detected early. Therefore, disease mapping has recently become an important method in the fields of public health research and disease epidemiology. The correct choice of statistical model is a very important step to producing a good map of a disease. Libya was selected to perform this work and to examine its geographical variation in the incidence of lung cancer. The objective of this paper is to estimate the relative risk for lung cancer. Four statistical models to estimate the relative risk for lung cancer and population censuses of the study area for the time period 2006 to 2011 were used in this work. They are initially known as Standardized Morbidity Ratio, which is the most popular statistic, which used in the field of disease mapping, Poisson-gamma model, which is one of the earliest applications of Bayesian methodology, Besag, York and Mollie (BYM) model and Mixture model. As an initial step, this study begins by providing a review of all proposed models, which we then apply to lung cancer data in Libya. Maps, tables and graph, goodness-of-fit (GOF) were used to compare and present the preliminary results. This GOF is common in statistical modelling to compare fitted models. The main general results presented in this study show that the Poisson-gamma model, BYM model, and Mixture model can overcome the problem of the first model (SMR) when there is no observed lung cancer case in certain districts. Results show that the Mixture model is most robust and provides better relative risk estimates across a range of models.

Keywords: Spatial models- disease mapping- lung cancer- poisson- gamma model- relative risk

Asian Pac J Cancer Prev, **18 (3)**, 673-679

Introduction

Globally, cancer, also called malignancy, is an abnormal growth of cells. There are more than 200 different types of cancer, including lung cancer, prostate cancer, breast cancer, liver cancer, and skin cancer. It is one of the most important health phenomena of today. Its high mortality rate, the disabilities it leaves behind and the high cost of medication are the causes of heavy loss in national economy and labour. Symptoms vary depending on the type of cancer and cancer treatment may include radiation, chemotherapy, and/or surgery (Cancer Research UK, 2016). Fortunately, cancer can be controlled if detected early. Therefore, in order to prevent and control this disease before it occurs, both of the society and government must be cooperative to eradicate this deathly disease (WHO, 2016).

In this research, lung cancer is the focus of our attention. Generally, the number of lung cancer cases seems to be increasing from year to year worldwide, and this includes developing countries, Libya, specifically. Lung cancer is the most common cancer in the world after skin cancer, breast cancer in women and prostate cancer in men, which causes most of the death of cancer (Siegel et al., 2011; Ferlay et al., 2010). In addition, studies confirmed that there is a great variation in the geographical distribution of lung cancer in the world. Approximately 70% of all new cases of lung cancer occur in developed countries, especially in Arabic countries. It is still the most common cancer in men worldwide (1.61 million new cases, 12.7% of all new cancers). Also, lung cancer is the most common cause of death from cancer, with 1.38 million deaths (18.2% of the total). The majority of the cases now occur in the developing countries (55%) (John and Ross, 2010; Salim, Jaziel, and Moore, 2011; Washington, 2007).

In Libya, cancer is becoming a major problem; however, epidemiological data with its distributions are sparse. In the case of lung cancer, it is the most prevalent kind of cancer among the Libyan society (Abusaa, 2008;

*[1]Department of Statistics, Faculty of Science, University of Tripoli, Alfernag Tripoli, Libya, [2]Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, Tanjung Malim, Perak, Malaysia. *For Correspondence: m.alhdiri@yahoo.com*

Bodalal et al., 2014; Bodalal and Bendardaf, 2014). Abusaa (2008) mentioned that the most common cancer in Libya is cancer of breast (23%), closely followed by cancer of lung (15%). Together, lung cancer and breast cancer comprise about 40% of all cancer patients. Tobacco-related cancer is the second most common cancer in males and fifth most common cancer in females. Annual reports of the African Oncology Institute (AOI) clearly show that cigarette smoking is a prevalent male habit in Libyan society. In Libya, the ratio of females who are smoking is almost negligible (Abusaa, 2008).

Regarding the importance of cancer prevention and considerable variation of lung cancer incidence in different parts of the country, it is necessary to recognize regions with high incidence of lung cancer and evaluate the role of potential risk factors by utilise of advanced statistical models with cancer maps. These models and maps can be used to map where people diagnosed with a particular disease live. In this work, disease mapping is the focus of our attention. Disease mapping is a method to display the geographical distribution of disease occurrence, given 'noisy' observed data on disease rates. In other words, disease mapping is a process of displaying the geographical variability of disease on maps using different colours or shading. The idea is not new, but the advent of computers and computer graphics has made it simpler to apply and it is now widely used in descriptive epidemiology. For example, to display morbidity or mortality information for an area. Good maps of disease risk have been recognized as an important tool for disease control. Disease maps may be useful, especially for government agencies to allocate resources (Meza, 2003). The maps can also be used for evaluating the performance of public health interventions. In any case, the maps must be designed to communicate effectively among the public, health researchers and decision makers (Bell et al., 2006). The biggest challenge is to ensure the maps not being misinterpreted.

Recently, several types of research focused on cancer disease mapping have been familiar with the health service geographical information system (GIS) (Elebead et al., 2012; George et al., 2013). It works by linking data to maps, via a geographical or spatial link. There are many of the advantages of this software, which are producing and providing the clear depiction of relative risk estimations which analyse the spatial data. Also, GIS can be a useful presentation tool for disease maps. With the GIS software, the selected models in this study were fitted to the data using full Bayesian estimation using WinBUGS version 1.4.3. This software is specialized in BUGS (Bayesian inference Using Gibbs Sampling) and can be run under Windows. It assumes a Bayesian probability model, in which all unknown parameters are treated as random variables.

In this review, since the cancer is the ever increasing health problem and most common cause of medical deaths in Libya, in order to manage this problem in Libya, we discuss the lung cancer situation in general and in different districts in Libya. Also this study discusses the most common models used in the field of disease mapping. These involve the analysis for relative risk estimation based on the SMR method, Poisson Gamma Model, Besag, York and Mollie (BYM) Model and Mixture model, with their application to lung cancer in Libya, in order to demonstrate and identify a better method of estimating lung cancer risk. We conducted this study to learn more about the geographical distributions of lung cancer for twenty-two administrative districts in Libya.

This paper is organized as follows. First, the classical method in estimation the SMR method, is described; also its drawbacks are considered. This is followed by an overview of the earliest application of the Bayesian methodology called the Poisson-gamma model, which is based on the model suggested by (Lawson et al., 2003). Likewise, the definitions and overview for a set of models used to improve the estimate of relative risk are considered. These are Poisson gamma model, BYM model and Mixture model. All these models are used to reduce in noise from the SMRs method of smoothing tool and modelling tools of relative risk to get estimates. The relative risk which are estimated, will be displayed in maps to depict the high and low risk areas. These maps can be produced using the ArcGIS software. Results of the analysis will be presented in tables and maps. Finally, these four methods or models are applied to observe lung cancer data in Libya to identify a better method of estimate lung cancer risk by using a common goodness-of-fit called Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002).

Materials and Methods

Disease maps are important tools in public health research. In the case of cancer, mapping can be viewed as a descriptive presentation of the cancer burden in some geographical areas and can help point out the areas where health policy should be improved. In this section, the existing models used in the study of disease mapping will be explained and applied. In fact, there are many models that commonly been used which includes the non-spatial models, spatial models, and space-time models. However, the classical method (SMR) and three models (spatial and non-spatial) were used in this work. The analysis of relative risk estimation begins with the common method used in the study of disease mapping which is based on SMR in order to assess the status of an area with respect to disease incidence. This is followed by using Poisson-gamma, BYM model, and the Mixture models to estimate relative risk for lung cancer risk in Libya.

Standardized Morbidity Ratio (SMR) Method

SMR is the most common method used by researchers in the choice of appropriate measures of relative risk in disease mapping. In this research, SMR compares the observed incidence with the expected incidence, which has been traditionally used for the analysis of counts within tracts as described by Lawson (2006). In disease mapping, suppose that the study area to be mapped is divided into P exclusive regions (i=1, 2…, P), each region has its own observed number of cases O_i and expected number of cases Ei. Using O_i and E_i as obtained from the available data, the relative risk Θ_i, which is the SMR, for area i, can be

calculated using the following equation.

$$\theta_i = (O_i / E_i) \quad i = 1, 2..., P \qquad (1)$$

A study by Samat and Percy (2008) used Equation (1) in their discussion on standardized morbidity ratio and its application to dengue disease mapping in Malaysia. Although the SMR is commonly used as an index to measure relative risk, it has several disadvantages. According to (Lawson et al., 2003), since SMR is based on a ratio estimator the mean and variance of SMR are highly dependent upon Ei. The SMR is very large in areas where the expected number of cases are small, and small for areas where the expected number of cases are large. Furthermore, in areas where there are no observed count data or cases, the SMR is necessarily zero.

Poisson-gamma Model

Problems of SMR in disease mapping such as if there are areas with no observed count data or cases, then the SMR is necessarily zero in these areas, encouraged many researchers to produce a variety of methods to estimate the relative risk of a disease, including the use of Bayesian methods. Clayton and Kaldor (1978) was the first person who proposed the Poisson-Gamma model which assumed that the relative risk have the Poisson distribution. The Poisson - Gamma model is one of the earliest example of Bayesian mapping (Lawson et al., 2003). In this model, for i=1,2,..,*p* study areas, Let O_i, E_i, Θ_i be the observed count, expected count and relative risk parameter in the ith area. The numbers of new cases O_i are assumed to follow a Poisson distribution with period time, as equation 2.

$$O_i / E_i, \theta_i \sim Poisson(E_i \theta_i), \quad \forall i \qquad (2)$$

If the prior distribution is a gamma distribution with parameters a and b (*Gamma (a, b*) ,then the relative risks or the posterior distribution of the relative risk Θ_i has a gamma distribution with parameters a and b (*Gamma* $(a + O_i, b + E_i)$). However, this model also has some drawbacks. One of them is its inability to cope with the spatial correlation. This drawback of this model can be overcame by using BYM models which consider the neighboring area.

Besag, York and Mollie Model

To overcome the problem of SMRs, this research uses the BYM model to demonstrate the data. BYM model, introduced by Clayton and Kaldor, (1978) and developed by Besag et al., (1991). The main idea for this model is to produce more reliable estimates for relative risks 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 are decomposed which are divided into two components. The first component is ui that takes into account the effects that vary in a structured manner in space (clustering or correlated heterogeneity). The second component is vi that takes into account the effects that vary in an unstructured way between areas (uncorrelated heterogeneity). The model is formulated as follows:

$$O_i / E_i, \theta_i \sim Poisson(E_i \theta_i), \quad \forall i \qquad (3)$$
$$\log \theta_i = \alpha + u_i + v_i,$$

where α is an overall level of the relative risk, O_i, 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:

$$[u_i / u_j, i \neq j, \tau_u^2] \sim N(\bar{u}_i, \tau_i^2)$$

where:

$$\bar{u}_i = \frac{1}{\sum_j \omega_{ij}} \sum_j u_i \omega_{ij}, \; \tau_i^2 = \frac{\tau_u^2}{\sum_j \omega_{ij}}, \omega_{ij} = 1$$

if *i, j* are adjacent (or 0 if they are not).

In the field of full Bayesian inference, prior distributions 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.

Mixture Model

This research considers the mixture model, which proposed by Lawson and Clark (2002). This model assumed that the log-relative risk could be written as follows,

$$Log\theta_i = \alpha + v_i + p_i u_i + (1 - p_i)\varphi_i \qquad (4)$$

The log function of θ_i is influenced by three additive components. v_i (follows a Normal distribution ($v_i \sim N(0, 1/\tau_v^2)$) is the fixed component which represents the unstructured heterogeneity and measure the overdispersion in a individual region. In the case of other two mixing components are (*u, φ)*, represent different aspects of spatial correlation.

Special cases of this formulation arise depending on the value of pi. If all $p_i = 1 \; \forall_i$, the Equation (4) become which is same with the BYM model in Equation (3) (we obtain the BYM model). The other case is a pure jump model, which arises if all $p_i = 0 \; \forall_i$ (Lawson and Clark, 2002; Lawson et al., 2000).

Most researchers noticed that the maps produced by the mixture model were very clear and visually closer, compare with maps which were produced by the SMRs and maps produced by the BYM model (Lawson et al., 2003; Lawson and Clark, 2002).

In this study based on three common models, which are the Poisson-gamma model, BYM model and Mixture model, the results of the analysis contain the posterior expected relative risks for all districts and for the time periods of the study.

Application to Lung Cancer Data in Libya

In this section, the outcomes of the applications of common relative risk estimation methods conformable to the traditional method based on the standardized morbidity ratio (SMR), Poisson-gamma, BYM and Mixture model using observed lung cancer data of Libya are displayed and compared in graphs and maps.

The Data Set

In this study, all observed cancer incidence and population data in Libya were provided by AOI Sabratha for the years 2006 to 2011. This data set gives the number of incidences per year. The area under study covered twenty-two major locations (local authority districts in Libya), namely, Alnikat, Zawia, Aljafara, Tripoli, Almergaib, Mustrata, Sirt, Benghazi, Almarg, Aljabal Alakhader, Darna, Albatnan, Nalut, Aljabal Algarbi, Wadi Shatee, Aljufra, Ejdabiya, Ghat, Wadi Alhiya, Sabha, Morzuk and Alkufra. The districts consisted of urban, suburban and rural populations (as shown in Figure 1).

Results

The outcomes of total number of 517 cases of lung cancer by years and districts are displayed in Figures 2 and 3, respectively. Figure 2 presents the incidence of lung cancer cases in Libya during 6 years from 2006 to 2011. There were minor fluctuations of lung cancer cases between 2009 and 2010, during which period the worst outbreak occurred in 2009 with 104 cases, followed by 2007 with 100 cases. The number of cases decreased in 2011 to 51 cases. It can be seen from the bar graph in Figure 3 that the capital of city, Tripoli, which is located at the north-west of Libya, recorded the highest number of cases in 145, followed by the district of Zawia, which recorded the second highest number of lung cancer with 136 cases. The other two districts that recorded the highest number of cases where the district of Alnikat with 40 cases, and the northeast district of Aljabal Algarbi with 38 cases. The total numbers of cases reported in other districts are in the range between 0 and 33 cases.

The outcomes of relative risk estimation using the existing models, which are the SMR method, Poisson gamma model, BYM model and Mixture model for all 22 districts of Libya for six years from 2006 to 2011, are displayed in Table 1. Table 1 shows numeric values for the relative risk based on four models. For all years of study, it can be seen that there are a number of districts have a relative risk (RR) less than one. Therefore, that means that the O_i (observed number) of lung cancer cases

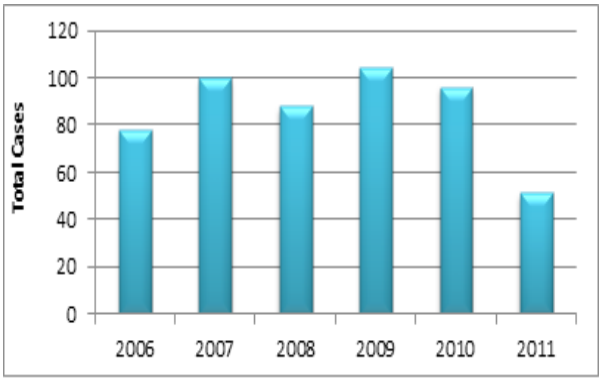

Figure 2. Incidence of Lung Cancer Cases (Clinical) from 2006 to 2011

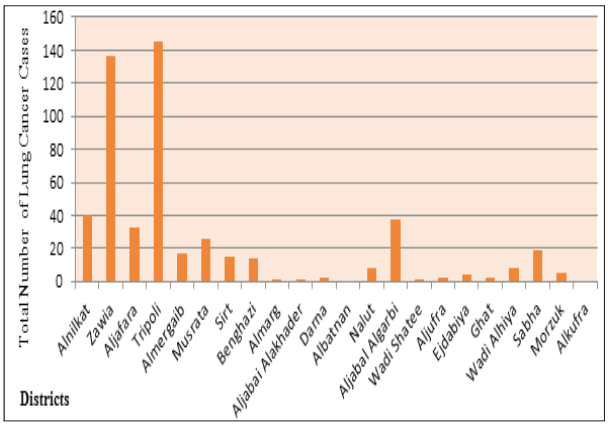

Figure 3. Total Number of Lung Cancer Cases Reported for Each District in Libya from 2006 to 2011

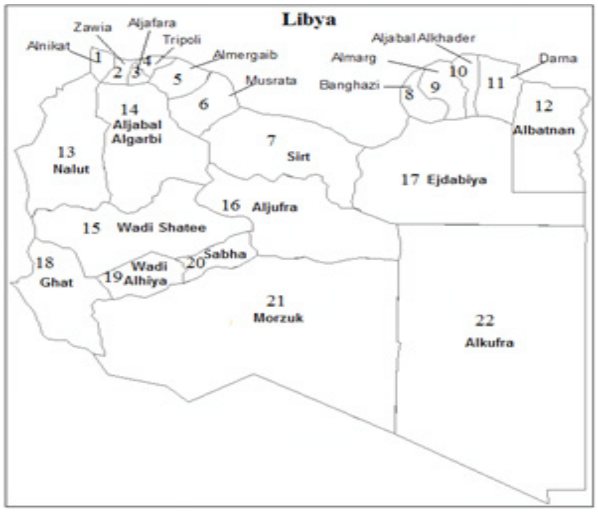

Figure 1. 22 Authority Districts in Libya (Source: Alhdiri et al., 2016)

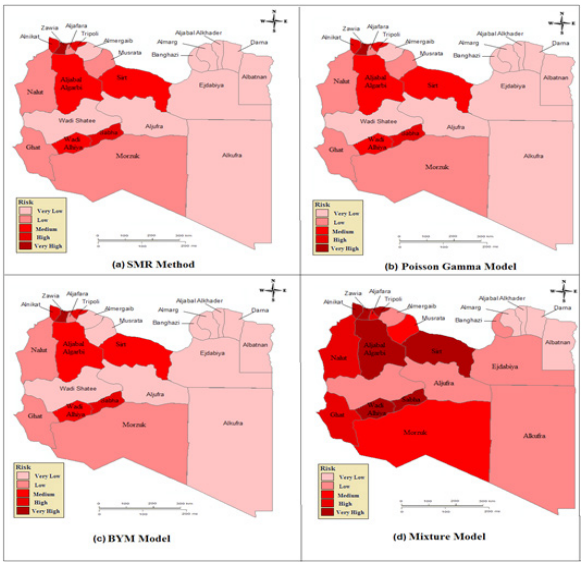

Figure 4. SMR, Poisson gamma, BYM Maps for Lung Cancer Cases during the years 2006 to 2011

is smaller than the E_i (expected number) of cancer cases in those districts by using four models. Among those districts, ten districts have relative risk less than one for all four models, with most located in the eastern part of the country. These are Almergaib, Benghazi, Almarg, Aljabal Alakhader, Darna, Albatnan, Wadi Shatee, Aljufra, Ejdabiyaand Alkufra. Seven districts have a relative risk greater than one. This means that the observed number of cases is larger than the expected number of cases, they are Alnikat, Zawia, Tripoli, Sirt, Aljabal Algarbi, Wadi Alhiya and Sabha. However, the five districts have less than one for SMR, Poisson gamma and BYM model, which have greater than one by using the Mixture model. addition, this table for lung cancer shows several differences in terms of high and low cancer occurrence areas for the districts in Libya by using four different methods. It can be seen that when using the SMR model, the estimated relative risk becomes exactly zero when there are no observed lung cancer cases in certain district, such as in the both districts of Albatnan and Alkufra have a relative risk equal to zero because there is no report for lung cancer cases during 2006 to 2011. This is the disadvantage of the SMR approach. Conversely, the relative risk estimation based on three other models (Poisson-Gamma model, BYM model and Mixture model) can overcome the problem of SMR when there are no observed lung cases in certain regions. It can be seen clearly from Table 1 that the most

Table 1. Relative Risk Based on the SMR Model and Posterior Expected Relative Risks Based on the Existing Models (Poisson Gamma model, BYM Model and Mixture Model) for Lung Cancer

District	SMR	Model		
		Poisson-gamma	BYM	Mixture
Alnikat	1.53	1.5	1.49	3.33
Zawia	5.16	4.99	5.09	11.37
Aljafara	0.83	0.84	0.83	1.81
Tripoli	1.51	1.5	1.51	3.34
Almergaib	0.43	0.44	0.43	0.96
Musrata	0.53	0.53	0.53	1.17
Sirt	1.15	1.13	1.11	2.45
Benghazi	0.24	0.25	0.25	0.54
Almarg	0.06	0.11	0.13	0.26
Aljabal Alakhader	0.05	0.09	0.11	0.23
Darna	0.13	0.17	0.18	0.39
Albatnan	0.00	0.06	0.09	0.17
Nalut	0.91	0.89	0.86	1.88
Aljabal Algarbi	1.35	1.33	1.33	2.94
Wadi Shatee	0.14	0.24	0.24	0.53
Aljufra	0.32	0.40	0.38	0.84
Ejdabiya	0.23	0.27	0.27	0.59
Ghat	0.72	0.75	0.68	1.48
Wadi Alhiya	1.16	1.11	1.07	2.39
Sabha	1.64	1.57	1.56	3.47
Morzuk	0.71	0.72	0.67	1.48
Alkufra	0.00	0.14	0.17	0.33

Table 2. Deviance Information Criterion (DIC) for Relative Risk Based on the Poisson- Gamma Model, BYM Model and Mixture Model to Estimate Relative Risk of Lung Cancer

	Model		
	Poisson-gamma	BYM	Mixture
DIC	122.09	126.26	120.44

significant difference is in the district of Zawia, which has the highest relative risk of contracting lung cancer by using all models considered.

This indicates that susceptible people within this district is more likely to contract lung cancer compared with people in the overall population. The corresponding values of relative risk are approximately 5.16, 4.99, 5.09 and 11.37, respectively. The highest risks in cancer incidence in these districts is probably related to the tendency of oil installations in this area such as Mellitah Oil and Gas B.v, the Azawia Oil Refining Company and Bouri Oil Field, as well as the electrical power stations (Alsaker, 2013). Susceptible people within districts located in the Eastern part of the country have the lowest risk of all models when compared to the people in the overall population. The district with the lowest risk was Albatnan. The corresponding values of relative risk are approximately 0.00, 0.06, 0.09 and 0.17, respectively.

In this study, Deviance Information Criterion (DIC) is applied as the goodness-of-fit (GOF) measure for the models used. This GOF is used because it can be evaluated easily in the WinBUGS software that we used and study by Lawson (2009) showed some disadvantages of the other measures such as AIC and BIC when is involving a model with several random effects. The DIC was proposed by Spiegelhalter et al., (2002). According to Lawson et al., (2003) and Spiegelhalter et al. (2002) the model that has the smallest DIC is considered as the best model for data. Table 2 shows the DIC values for the models for all districts in Libya with lung cancer during 2006 to 2011.

From the DIC in Table 2, it can be seen that there are differences in the terms of the DIC value between different types of cancer disease. As compered among the models, the Mixture model shows a better model for lung cancer because it has the smallest DIC. As a conclusion of this outcome, this means that the Mixture model is the best model to be used in the analysis specifically for estimation of relative risk for lung cancer data. This conclusion gives some insight into the possible advantages of collecting other cancer data in the future and which model is the best to use to give better relative risk estimation results.

Maps of the SMR, Poisson Gamma, BYM and Mixture Model for Lung Cancer in Libya

From one of the goals of this study is to demonstrate the graphical distributions of the districts, therefore, the maps of disease have been used as a means of graphical statistical results to estimate the relative risk based on SMR method, Poisson gamma model, BYM model and Mixture model. The comparison between them by using the DIC method for the number of lung cancer cases is reported for each district in Libya. In this application, for the purpose of clarity in the interpretation, we used

thematic maps with multiple colours in this analysis in order to show and differentiate between different levels of the disease in each area, especially, the high and low risk area. In disease mapping, there is no definitive way to choosing the interval levels of risk, so each district is assigned one of five different levels of relative risk which are very low (the lighter regions), low, medium, high and very high risks (the darker regions), with respective intervals of [0,10), [10, 20), [20, 30), [30,40) and [40, ∞) respectively for summary map, and [0.0,0.5), [0.5,1.0), [1.0,1.5), [1.5,2.0) and [2, ∞) respectively for relative risks map.

Figure 4 shows thematic lung cancer risk maps to estimate the relative risk using four models in 22 districts in Libya during the years from 2006 to 2011. These maps give a clear presentation about the high and low risk area of lung cancer occurrences. The maps illustrated in Figure 4 depicts huge differences in terms of high and low lung cancer occurrence areas for the districts in Libya, particularly in the north-west and south-west of the country. The most important difference is no difference between the maps of lung cancer for SMR model, Poisson gamma model and BYM model. These models showed that only a few districts show the very high and high risk of lung cancer, while Mixture map clearly show the high and low risk area of lung cancer occurrence. For instance, the SMR map, Poisson gamma map and BYM map show that the district of Zawia has a very high risk and the districts of Sabha, Tripoli and Alnikat have high risk, which means that people in these districts are more likely to have lung cancer compared to people in the overall population. The Mixture map shows that there are seven districts with very high risk which are Zawia Alnikat, Sirt, Tripoli, Aljabal Algarbi, Sabha and Wadi Alhiya. This is followed by the districts of Nalut, Ghat and Aljafara with high risk and the eleven other districts have medium and low risk of lung cancer, most of them located in the eastern part of the country.

Comparisons among the SMR map, Poisson-gamma map, BYM map and Mixture map specifically from 2006 to 2011, demonstrate no obvious differences in terms of the estimated risks generated by three methods considered. DIC shows that Mixture model is the best model to estimate relative risk for lung cancer, compare to other models. These maps are primarily intended to be a good presentation tool for identifying districts with very high and high risk, so that further attention could be given to these districts, in the future.

This study has been carried out to identifying the hot-points areas (the high risk or low risk) for lung cancer incidences in Libya by displaying the risks in the map using the classical method and three statistical models. The relative risk estimations obtained by the selected models (Poisson-gamma model, BYM model and Mixture models) that are suggested in this study provide an important approach to assess future risk. The estimated relative risk values can be displayed in a lung cancer risk map to show the high and low risk area of lung cancer risk. The higher the risk of a particular area, the more serious attention of government policy and financial support are needed. It can be concluded that those models can be implemented using lung cancer data in Libya during six years from 2009 to 2011.

The results of the analysis show that the use of the spatial model which is Mixture model in estimating the relative risk in maps provides better high-low risk appearances in maps compared to SMR method, Poisson gamma model, BYM model. Mixture model can be chosen as the better model that fit the data, since it has the smallest DIC. Additionally, the other two spatial and non- spatial models (BYM model, Poisson-gamma model), can also be applied to model the relative risks estimation, but are not as good as the Mixture model, since these models have greater DIC.

These findings suggest that the Mixture model can be used as a basic procedure for estimating relative risk instead of using the number of lung cancer cases alone. The methodology of this research might be useful in the analysis of geographical variances in cancer morbidity and cancer mortality, which has been discovered in cancer record data. These maps should be regarded as tools for forming hypotheses leading to attempts to start investigating and developing further analysis to improve upon current models in order to inform and direct government strategy for monitoring and controlling lung cancer disease.

Acknowledgments

All the authors of this research would like to thank the staff members in the African Oncology Institute (AOI) Sabratha at the Libyan Cancer Centre, and the staff in the Universiti Pendidikan Sultan Idris, Malaysia, for their continued support for the completion of this study. This article was a part of Ph.D. thesis in applied statistics at Department of Mathematics, Faculty of Science and Mathematics at Universiti Pendidikan Sultan Idris supported it.

References

Abusaa A (2008). First annual report population-based cancer registry 2006. Sabratha, African Oncology Institute, pp 1-60.

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. University of Tripoli, pp 1-232.

Bell BS, Hoskins RE, Pickle LW, Wartenberg D (2006). Current practices in spatial analysis of cancer data: Mapping health statistics to inform policymakers and the public. I*nt J Health Geogr*, **5**, 49.

Besag J, York J, Mollie A (1991). Bayesian image restoration with two applications in spatial statistics. *Ann Inst Stat Math*, **43**, 1–59.

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 Gastrointest Oncol*, **6**, 98–103.

Cancer Research UK. What is cancer?. Retrieved 10 January 2017, Available at: http://www.cancerresearchuk.org/

about-cancer/what-is-cancer.
Clayton DG, Kaldor J (1987). Empirical bayes estimates of age-standardised relative risks for use in disease mapping. *Biometrics*, **43**, 671–91.
Elebead FM, Hamid A, Hilmi HSM, Galal H (2012). Mapping cancer disease using geographical information system (GIS) in Gezira state-Sudan. *J Community Health*, **37**, 830-9.
Ferlay J, Hai RS, Freddie B, et al (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer,* **127**, 2893–917.
George JM, Po-Huang C, Tyler S, et al (2013). Use of GIS mapping as a public health tool from cholera to cancer. *Health Serv Insights*, **6**, 111–16.
John R, Ross H (2010). The global economic cost of cancer. Atlanta, GA: American Cancer Society and Livestrong Organization, pp 1-10.
Lawson A B (2006). Statistical methods in spatial epidemiology. England: Wiley, pp 1-424.
Lawson AB (2009). Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology. *Int Stat Rev*, **77**, 325–6.
Lawson AB, Biggeri AB, Boehning D, et al (2000). Disease mapping models: An empirical evaluation. *Stat Med*, **19**, 2217–41.
Lawson AB, Browne WJ, Vidal Rodeiro CL (2003). Disease mapping with winBUGS and MLwiN. England: John Wiley and Sons, pp1-296.
Lawson AB, Clark A (2002). Spatial mixture relative risk models applied to disease mapping. *Stat Med*, **21**, 359–70.
Meza JL (2003). Empirical bayes estimation smoothing of relative risks in disease mapping. *J Stat Plan Inference*, **112**, 43–62.
Salim E, Jaziel AR, Moore AM (2011). Lung cancer incidence in the Arab league countries: Risk factors and control. *Asian Pac JCancer Prev*, **12**, 17–34.
Samat NA, Percy DF (2008). Standardized mortality and morbidity ratios and their application to Dengue disease mapping in Malaysia. In Proceedings of the Salford postgraduate annual research conference, 200–10. Manchester: Salford University. doi:9781905732715.
Siegel R, Elizabeth W, Otis B, Ahmedin J (2011). Cancer Statistics, 2011 the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA Cancer J Clin*, **61**, 212–36.
Spiegelhalter DJ, Best NG, Carlin BP, Der Av (2002). Bayesian measures of model complexity and fit (with discussion). *J R Stat Soc Series B*, **64**, 583–640.
Washington DC (2007). Food, nutrition, physical activity, and the prevention of cancer: A global perspective / world cancer research fund; American institute for cancer research. doi:978-0-9722522-2-5.
World Health Origination (WHO), Cancer: Early detection of cancer. Retrieved 13 March 2017, Available at: http://www.who.int/cancer/detection/en/.